

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
CÂMPUS CORNÉLIO PROCÓPIO  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

REGINALDO DA ROCHA

**EXPLORANDO ABORDAGENS EVOLUTIVAS PARA  
RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO**

DISSERTAÇÃO – MESTRADO

CORNÉLIO PROCÓPIO

2016

**REGINALDO DA ROCHA**

**EXPLORANDO ABORDAGENS EVOLUTIVAS PARA  
RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO**

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – UTFPR como requisito parcial para a obtenção do título de “Mestre Profissional em Informática”.

Orientador: Prof. Dr. Pedro Henrique Bugatti

Co-orientador: Profa. Dra. Priscila Tiemi Maeda Saito

**CORNÉLIO PROCÓPIO**

**2016**

---

Dados Internacionais de Catalogação na Publicação

---

- R672 Rocha, Reginaldo da  
Explorando abordagens evolutivas para recuperação de imagens baseada em conteúdo /  
Reginaldo da Rocha. – 2016.  
95 f. : il. ; 30 cm
- Orientador: Pedro Henrique Bugatti.  
Coorientador: Priscila Tiemi Maeda Saito  
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-  
graduação em Informática. Cornélio Procópio, 2016.  
Referências: p. 88 - 95.
1. Realimentação (Eletrônica). 2. Recuperação de imagem. 3. Algoritmos. 4. Imagens médicas  
5. Informática - Dissertações. I. Bugatti, Pedro Henrique, orient. II. Saito, Priscila Tiemi Maeda,  
coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em  
Informática. IV. Título.

CDD (22. ed.) 004



**Título da Dissertação Nº 25:**

**“EXPLORANDO ABORDAGENS EVOLUTIVAS PARA RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO”.**

por

**Reginaldo da Rocha**

Orientador: **Prof. Dr. Pedro Henrique Bugatti**

Esta dissertação foi apresentada como requisito parcial à obtenção do grau de MESTRE EM INFORMÁTICA – Área de Concentração: Computação Aplicada, pelo Programa de Pós-Graduação em Informática – PPGI – da Universidade Tecnológica Federal do Paraná – UTFPR – Câmpus Cornélio Procópio, às 14h do dia 06 de julho de 2016. O trabalho foi \_\_\_\_\_ pela Banca Examinadora, composta pelos professores:

---

Prof. Dr. Pedro Henrique Bugatti  
(Presidente – UTFPR-CP)

---

Prof. Dr. Claiton de Oliveira  
(UTFPR-CP)

---

Prof. Dr. Sylvio Barbon Junior  
(DC-UEL)

---

Profa. Dra. Priscila Tiemi Maeda Saito  
(Coorientadora – UTFPR - CP)

Visto da coordenação:

---

**André Takeshi Endo**

Coordenador do Programa de Pós-Graduação em Informática  
UTFPR Câmpus Cornélio Procópio

A Folha de Aprovação assinada encontra-se na Coordenação do Programa.

Dedico esse trabalho à minha família, em especial a minha esposa Rita e minhas filhas Julia e Beatriz que são as mulheres da minha vida. Sempre estiveram comigo durante essa caminhada, me animando, incentivando e apoiando nos momentos mais difíceis. Aos meus avôs, Raimundo e José, que partiram deixando saudades e muitas lembranças.

## **AGRADECIMENTOS**

Agradeço a Deus por sempre me dar a porção necessária a minha caminhada, me trazendo até aqui, por ter colocado no meu caminho amigos e professores dedicados.

Ao meu orientador, Professor Dr. Pedro Henrique Bugatti, por sua dedicação, algumas vezes me orientando as 6h da manhã, com paciência e conhecimento me ajudando em todas as etapas desse trabalho. Agradeço a Professora Dra. Priscila T. M. Saito pelas valiosas contribuições.

Aos meus familiares, por todo incentivo e apoio. Ao meu amigo Diego Guarnieri pelo companheirismo durante essa jornada. Por fim a todas as pessoas que colaboraram direta ou indiretamente com este trabalho.

Confia no Senhor de todo o teu coração, e não te estribes no teu próprio entendimento. Reconhece-o em todos os teus caminhos, e ele endireitará as tuas veredas. Provérbios 3:5,6.

## RESUMO

ROCHA, Reginaldo. EXPLORANDO ABORDAGENS EVOLUTIVAS PARA RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO. 95 f. Dissertação – Mestrado – Programa de Pós-graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

O processo de organização e recuperação de imagens apresenta inúmeros problemas a serem abordados, compreender o significado subjetivo de uma consulta visual por meio de parâmetros numéricos que podem ser extraídos e comparados por meio de um computador é um dos maiores desafios. A disparidade entre a comparação realizada pela máquina e o que de fato o ser humano interpreta de uma consulta visual é denominado de lacuna semântica. Diversos trabalhos na literatura abordam técnicas para diminuir essa disparidade. Nesse contexto, a abordagem de realimentação de relevância apresenta-se como um meio eficaz para capturar a intenção do usuário e reduzir a diferença entre os conceitos semânticos oriundos da percepção visual do usuário, bem como das características visuais de baixo nível extraídas automaticamente de uma imagem. Entretanto, a lacuna semântica ainda é um desafio a ser vencido. Por tal motivo, o presente trabalho tem por objetivo o estudo, análise e proposta de um arcabouço para recuperação de imagens o qual combina a realimentação de relevância unida a algoritmos evolutivos visando aproximar a expectativa do usuário em relação aos resultados retornados pela recuperação de imagens, por meio da captura de intenção do mesmo e posterior definição dos parâmetros mais adequados. Para tanto, a ideia principal do trabalho é inserir memória de aprendizado ao processo de recuperação de imagens por conteúdo, armazenando os dados da interação do usuário com o sistema em perfis, os quais posteriormente serão utilizados para prover respostas ao usuário de maneira personalizada, conseqüentemente, contribuindo para diminuição da lacuna semântica.

**Palavras-chave:** realimentação de relevância, algoritmos evolutivos, recuperação de imagens por conteúdo, funções de distância, imagens médicas



## ABSTRACT

ROCHA, Reginaldo. EXPLOITING EVOLUTIONARY APPROACHES FOR CONTENT-BASED IMAGE RETRIEVAL. 95 f. Dissertação – Mestrado – Programa de Pós-graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

The process of organization and retrieval of images presents numerous problems to be addressed, to understand the subjective meaning of a visual query through numerical parameters that can be extracted and compared by a computer is a challenge. The disparity between the comparison performed by the machine and that in fact the human being interprets a visual query is known as semantic gap. Several studies in the literature techniques to address this gap. In this context, relevance feedback (RF) is an effective approach to capture user intent and reduce the difference between semantic concepts and the visual characteristics of an image. However, the semantic gap is still a challenge to be overcome, therefore, the present work aims to study, analysis and proposal for a new framework for image retrieval through relevance feedback combined with evolutionary algorithms aiming to bring the user's expectation through the results returned by the retrieval of images through the his intention and definition of the most appropriate parameters. To this end, the main idea of the work is to insert learning memory in a content-based image retrieval system for this by storing the user interaction data with the system profiles so that later these data are able to provide parameters for the system "learn" and respond to personalized user and need. Thus, contributing to reduction of existing semantic gap between the results of a retrieval operation and indeed expected images for a specific context, improving the efficiency and effectiveness of the retrieval process.

**Keywords:** relevance feedback, evolutionary algorithms, content based image retrieval, distance functions, medical imaging

## LISTA DE FIGURAS

FIGURA 1	– Exemplo de ambiguidade em uma consulta textual usando a palavra-chave “banco”. . . . .	19
FIGURA 2	– Fluxo de dados de uma sistema CBIR. Adaptado de Smeulders et al. (2000) . . . . .	20
FIGURA 3	– Extração de características para representação de uma imagem. . . . .	21
FIGURA 4	– Representação visual da fusão de três histogramas de cores no espaço RGB. . . . .	22
FIGURA 5	– Exemplo de divisão de uma imagem para obtenção do histograma local de cor (LCH). . . . .	23
FIGURA 6	– Ilustração das áreas de abrangência das distâncias $L_1, L_2$ e $L_\infty$ . . . . .	28
FIGURA 7	– Esquema conceitual de um descritor de imagens. . . . .	30
FIGURA 8	– Exemplo de uma consulta kNN. . . . .	32
FIGURA 9	– Exemplo de uma consulta por abrangência. . . . .	33
FIGURA 10	– Recuperação de imagens baseada em conteúdo com realimentação de relevância. . . . .	36
FIGURA 11	– Exemplos de movimentação de consulta: (a) de um único centro; (b) múltiplos centros em forma convexa; (c) múltiplos centros em forma cônica. . . . .	38
FIGURA 12	– Estratégias para recuperação de imagens com múltiplas consultas. . . . .	39
FIGURA 13	– Esquema de um GA típico. . . . .	43
FIGURA 14	– Recombinação usando 2 pontos de corte. . . . .	46
FIGURA 15	– Ilustração do operador de mutação. . . . .	46
FIGURA 16	– Conjuntos de dados de uma consulta. . . . .	55
FIGURA 17	– Exemplo de uma Curva de Precisão e Revocação, extraída de Silva et al. (2011). . . . .	56
FIGURA 18	– Esquema proposto para a atualização de perfil do usuário. . . . .	59
FIGURA 19	– Otimização de parâmetros offline. . . . .	61
FIGURA 20	– Menu principal da aplicação implementada para executar os experimentos. . . . .	63
FIGURA 21	– Resultado de uma pesquisa de imagem por conteúdo e realimentação de relevância. . . . .	63
FIGURA 22	– Parametrização de experimentos. . . . .	64
FIGURA 23	– Amostras de imagens presentes em cada classe da base contendo imagens de exames de Papanicolau. . . . .	67
FIGURA 24	– Amostras de imagens presentes em cada classe da base contendo imagens de raios-X de pulmão. . . . .	68
FIGURA 25	– Amostras de imagens presentes em cada classe da base contendo imagens de mamografia. . . . .	68
FIGURA 26	– Amostras de imagens presentes em cada classe da base Corel. . . . .	69
FIGURA 27	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/Haralick Textures (Wavelet). . . . .	74

FIGURA 28	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/LCH. ....	74
FIGURA 29	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Manhattan/ <i>Chebyshev Coefficients (Fourier)</i> . ....	75
FIGURA 30	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Euclidiana/ <i>Comb Moments (Fourier)</i> . ....	76
FIGURA 31	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Manhattan/ <i>Comb Moments (Wavelet)</i> . ....	76
FIGURA 32	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Euclidiana/ <i>Zernike Coefficients</i> . ....	77
FIGURA 33	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Euclidiana/ <i>Haralick Textures (Wavelet)</i> . ....	77
FIGURA 34	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Euclidiana/LBP. ....	78
FIGURA 35	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Canberra/ <i>Tamura Textures (Chebyshev (Fourier))</i> . ....	79
FIGURA 36	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base <i>Corel</i> e descritor Euclidiana/ <i>Comb Moments (Chebyshev)</i> . ..	80
FIGURA 37	– Resultado obtido por meio da otimização de parâmetros offline, utilizando a base <i>Corel</i> e descritor Euclidiana/ <i>Zernike Coefficients</i> . ....	80
FIGURA 38	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/LCH. ....	82
FIGURA 39	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/ <i>Haralick Textures (Wavelet)</i> . ....	82
FIGURA 40	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Manhattan/ <i>Chebyshev Coefficients (Fourier)</i> . ....	83
FIGURA 41	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Euclidiana/ <i>Comb Moments (Fourier)</i> . ....	83
FIGURA 42	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de mamografia e descritor Euclidiana/ <i>Haralick Textures (Wavelet)</i> . ....	84
FIGURA 43	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de mamografia e descritor Euclidiana/LBP. ....	84
FIGURA 44	– Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base <i>Corel</i> e descritor Euclidiana/ <i>Comb Moments (Chebyshev)</i> . ....	85

FIGURA 45 – Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base *Corel* e descritor Euclidiana/*Zernike Coefficients*. ..... 85

## LISTA DE TABELAS

TABELA 1	– Extratores utilizados nos experimentos e a respectiva quantidade de atributos .....	69
TABELA 2	– Os seis primeiros genes do cromossomo definidos para o GA encontrar os melhores descritores .....	71
TABELA 3	– Os dois últimos genes do cromossomo definidos para o GA encontrar os melhores descritores .....	71
TABELA 4	– Descritores com melhor desempenho em cada base de imagem .....	72
TABELA 5	– Média de precisão e tempo de execução .....	72
TABELA 6	– Tamanho do vetor de características após processo de seleção de atributos .....	81
TABELA 7	– Tempo de execução (em segundos) dos experimentos antes e após a seleção de atributos .....	81

## LISTA DE SIGLAS

CBIR	<i>Content Based Image Retrieval</i>
RF	<i>Relevance Feedback</i>
PSO	<i>Particle Swarm Optimization</i>
GA	<i>Genetic Algorithm</i>
RGB	<i>Red, Green, Blue</i>
HSV	<i>Hue Saturation &amp; Value</i>
GCH	<i>Global Color Histogram</i>
LCH	<i>Local Color Histogram</i>
BIC	<i>Border/Interior Classification</i>
CS	<i>Contour Saliences</i>
QPM	<i>Query Point Movement</i>
QEX	<i>Query Expansion</i>
WED	<i>Weighted Euclidean Distance</i>
PRF	<i>Pseudo Relevance Feedback</i>
SSA	<i>Social Spider Algorithm</i>
MAP	<i>Mean Average Precision</i>
CFS	<i>Correlation based Feature Selection</i>
SQL	<i>Structured Query Language</i>
SGBD	Sistema Gerenciador de Banco de Dados

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>15</b>
1.1 CONSIDERAÇÕES PRELIMINARES .....	15
1.2 OBJETIVOS .....	16
1.3 ORGANIZAÇÃO DO TRABALHO .....	16
<b>2 FUNDAMENTOS E CONCEITOS RELACIONADOS</b> .....	<b>18</b>
2.1 SISTEMAS DE RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO ..	18
2.1.1 Características de imagem .....	20
2.1.1.1 Cor .....	21
2.1.1.2 Textura .....	24
2.1.1.3 Forma .....	26
2.1.2 Funções de distância .....	27
2.1.3 Descritores .....	29
2.2 SELEÇÃO DE ATRIBUTOS .....	31
2.3 CONSULTAS POR SIMILARIDADE .....	31
2.3.1 Vizinhos mais próximos .....	32
2.3.2 Abrangência .....	33
2.4 ABORDAGENS PARA REDUÇÃO DA LACUNA SEMÂNTICA .....	33
2.4.1 Perfil de usuário .....	34
2.4.2 Realimentação de relevância .....	35
2.4.2.1 Movimentação do centro de consulta .....	36
2.4.2.2 Reponderação de Consulta .....	40
2.5 ALGORITMOS EVOLUTIVOS .....	41
2.5.1 Algoritmo genético .....	42
2.5.1.1 Cromossomos .....	42
2.5.1.2 Função de Aptidão .....	43
2.5.1.3 Seleção .....	44
2.5.1.4 Operadores de reprodução .....	45
2.5.2 Enxame de Partículas .....	46
2.5.2.1 Inicialização do enxame .....	47
2.5.2.2 Deslocamento das partículas .....	47
2.5.3 Social Spider Algorithm .....	49
2.5.3.1 Aranha .....	49
2.5.3.2 Vibração .....	50
2.5.3.3 Padrão de busca .....	51
2.6 AVALIAÇÃO DE DESEMPENHO .....	54
2.6.1 Precisão e Revocação .....	54
2.7 CONSIDERAÇÕES FINAIS .....	57
<b>3 ABORDAGEM PROPOSTA</b> .....	<b>58</b>
3.1 RECUPERAÇÃO DE IMAGENS COM ALGORITMOS EVOLUTIVOS .....	58
3.1.1 Otimização do Processo de Realimentação de Relevância .....	60
3.1.2 Seleção de atributos .....	62

3.2	IMPLEMENTAÇÕES E RECURSOS COMPUTACIONAIS .....	63
3.3	CONSIDERAÇÕES FINAIS .....	64
<b>4</b>	<b>EXPERIMENTOS .....</b>	<b>65</b>
4.1	BASES DE IMAGENS .....	66
4.2	EXTRAÇÃO DE CARACTERÍSTICAS E PREPARAÇÃO DA BASE .....	69
4.2.1	Escolha dos melhores descritores .....	70
4.3	RESULTADOS .....	72
4.3.1	Redução de dimensionalidade .....	80
4.4	CONSIDERAÇÕES FINAIS .....	86
<b>5</b>	<b>CONCLUSÕES .....</b>	<b>87</b>
5.1	RESULTADOS ALCANÇADOS .....	87
5.2	PUBLICAÇÕES .....	88
5.3	PROPOSTAS PARA TRABALHOS FUTUROS .....	88
	<b>REFERÊNCIAS .....</b>	<b>89</b>



# 1 INTRODUÇÃO

## 1.1 CONSIDERAÇÕES PRELIMINARES

Atualmente, com a massificação de dispositivos equipados com câmeras digitais, o volume de imagens disponíveis nas mais diversas áreas do conhecimento aumentou drasticamente, criando a necessidade de desenvolvimento de mecanismos de busca, navegação e recuperação eficientes de imagens (ZHUO et al., 2014; LIU et al., 2007).

Nesse cenário torna-se ainda maior o desafio para encontrar uma solução que possa recuperar em meio a esse grande volume de informações imagens relevantes em um determinado domínio de conhecimento.

Uma forma de prover tal recuperação é por meio de sistemas de recuperação de imagens por conteúdo (*Content Based Image Retrieval - CBIR*), os quais tem colaborado na recuperação e classificação de imagens usando informações intrínsecas às mesmas, como por exemplo, características de textura, forma ou determinados padrões de cores. Dessa forma, sistemas CBIR conseguem se aproximar, ainda que pouco, da percepção humana.

No entanto, existe uma grande diferença entre a interpretação de uma imagem realizada por um computador e a realizada por um ser humano, o que na literatura é chamada de lacuna semântica (*Semantic Gap*). Para diminuir a lacuna semântica faz-se necessário capturar a intenção de consulta do usuário, aplicando para isso, por exemplo, técnicas de realimentação de relevância (*Relevance Feedback - RF*). Desenvolvida durante os anos sessenta essa técnica foi adaptada e introduzida para recuperação de imagens durante o início dos anos 90 (ZHOU; HUANG, 2003).

As técnicas de RF permitem aproximar o usuário do sistema de recuperação de imagens, uma vez que permite que o usuário informe o grau de relevância das imagens recuperadas na consulta, tal informação atribuída pelo usuário permite que a consulta seja refeita levando em consideração a intenção do usuário (TRAINA et al., 2006), provendo uma estratégia eficiente para mapear a semântica de alto nível.

Mesmo sendo largamente empregadas as técnicas de RF ainda são objetos de pesquisa, pois a lacuna semântica ainda é um desafio a ser vencido. Dessa forma, atualmente na literatura, uma série de novas estratégias incluindo algoritmos de otimização vem sendo aplicados em conjunto com RF à recuperação de imagens. Algoritmos evolutivos, como enxame de partículas (*Particle Swarm Optimization* - PSO) aplicados nos trabalhos de Arevalillo-Herráez et al. (2014), Chang et al. (2013), algoritmos genéticos (*Genetic Algorithm* - GA) aplicados nos trabalhos de Calumby et al. (2014), Torres et al. (2009), entre outros, tem melhorado o desempenho de sistemas CBIR quando acoplados a processos de recuperação de imagens.

Além disso, é extremamente importante que todas as informações oriundas do processo de recuperação de imagens realizada por um dado usuário seja persistida para posterior análise de intenção do mesmo, ou de um grupo dos mesmos, criando assim perfis de usuários.

É neste contexto, em que surgem novas abordagens que ajudam a diminuição da lacuna semântica (PINJARKAR MANISHA SHARMA, 2012; SU et al., 2011; LIU et al., 2007), que foram concentrados os estudos para a execução do presente trabalho.

## 1.2 OBJETIVOS

O objetivo geral desse trabalho foi propor e validar um modelo de recuperação de imagens utilizando realimentação de relevância combinada a algoritmos evolutivos, contribuindo para diminuição da lacuna semântica e especializando a recuperação de acordo com um determinado contexto.

Propondo-se, portanto, um novo arcabouço de realimentação acoplado a algoritmos evolutivos e levando em consideração informações oriundas do processo de consultas por similaridade, como as imagens relevantes e irrelevantes rotuladas durante uma dada consulta, e assim obter dados sobre expectativa do usuário, provendo a personalização do processo de recuperação, bem como o autoajuste dos melhores parâmetros do mesmo. Tal calibramento foi realizado por meio da definição dos melhores descritores em junção com métodos evolutivos, segundo a intenção do usuário. Além disso, os fatores de ponderação do processo de realimentação também foram acoplados a tais abordagens.

## 1.3 ORGANIZAÇÃO DO TRABALHO

A presente dissertação apresenta a seguinte organização:

- No presente capítulo foram apresentadas a introdução, abordando a motivação para o

desenvolvimento do presente trabalho, bem como os objetivos do mesmo;

- No Capítulo 2 são apresentados os conceitos pertinentes para o entendimento e desenvolvimento do presente trabalho, como o processo de recuperação de imagens baseada em conteúdo (CBIR), elucidando as etapas de extração de características, funções de distância, descritores, abordagens de redução da lacuna semântica, bem como medidas de avaliação de desempenho aplicadas ao processo de recuperação. Além disso, são elencadas também abordagens evolutivas utilizadas na literatura, bem como trabalhos que realizam a ligação de tais técnicas com CBIR.
- No Capítulo 3 são introduzidas as definições necessárias para formalização e entendimento da abordagem proposta, bem como a estruturação do arcabouço proposto e suas políticas agregadas. Além disso, é apresentada uma breve descrição de um protótipo para automatização e execução da presente proposta;
- O Capítulo 4 apresenta os experimentos e resultados obtidos ao longo do presente trabalho, bem como é apresentada uma análise em largura e profundidade em relação à diferentes instâncias da abordagem proposta em comparação à principal abordagem da literatura.
- Por fim, no Capítulo 5, são expostas as conclusões do presente trabalho, bem como uma síntese das principais contribuições do mesmo e perspectivas de trabalhos futuros.

## 2 FUNDAMENTOS E CONCEITOS RELACIONADOS

A recuperação de imagens baseada em conteúdo é empregada em diversas áreas do conhecimento, como por exemplo na medicina, nas artes, no entretenimento, na publicidade, na indústria, segurança, sensoriamento remoto, entre outras. Todas essas áreas possuem em comum a geração de um grande volume de dados complexos (i.e. imagens) os quais precisam ser organizados e recuperados de maneira eficaz e eficiente. Portanto, ferramentas que possibilitem tal organização e recuperação são cada vez mais necessárias e utilizadas por especialistas de vários domínios. Vale ressaltar que a recuperação de tais dados apresenta inúmeros desafios, uma vez que não apenas envolve diferentes campos como o processamento de imagens, a recuperação de informações e a visão computacional, bem como está intimamente atrelada à subjetividade do conceito de percepção visual sujeita a cada especialista de um dado domínio específico.

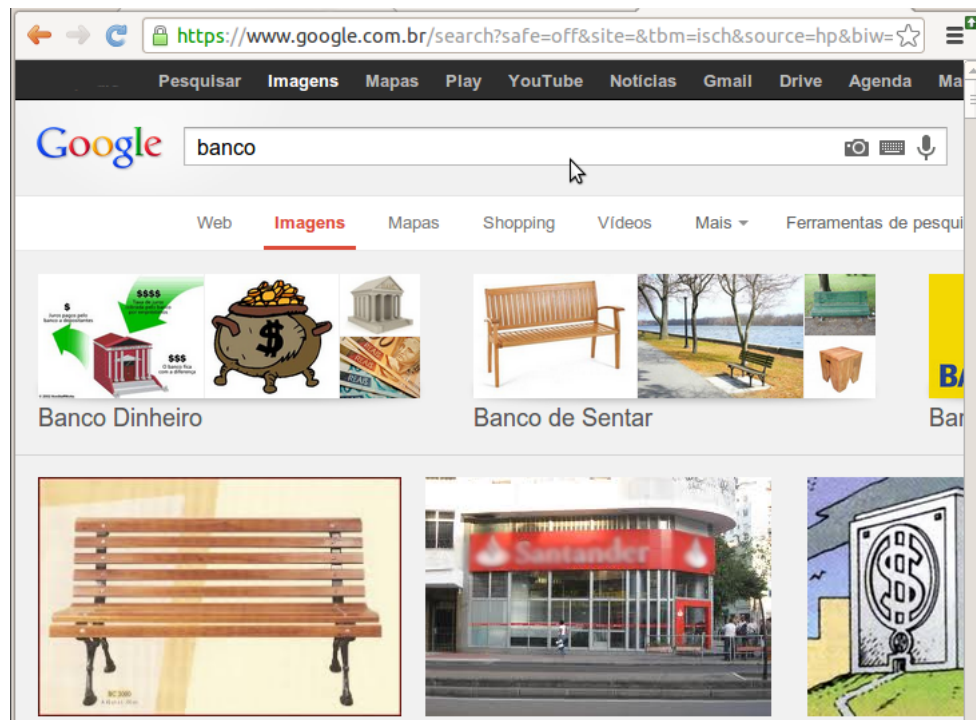
Uma imagem possui um conteúdo complexo carregado de semântica, porém a extração é muito custosa computacionalmente, uma vez que são usadas características visuais para representar as imagens. O desempenho de um sistema de recuperação de imagens por conteúdo depende fortemente das características extraídas. Normalmente, existe uma lacuna semântica entre o conteúdo de uma imagem e de suas características visuais.

Apesar dos diversos esforços despendidos para a diminuição da lacuna semântica, tal problema ainda se mantém como um tema desafiador na área de pesquisa em recuperação de imagens (SMEULDERS et al., 2000; RASHEDI et al., 2013).

### 2.1 SISTEMAS DE RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

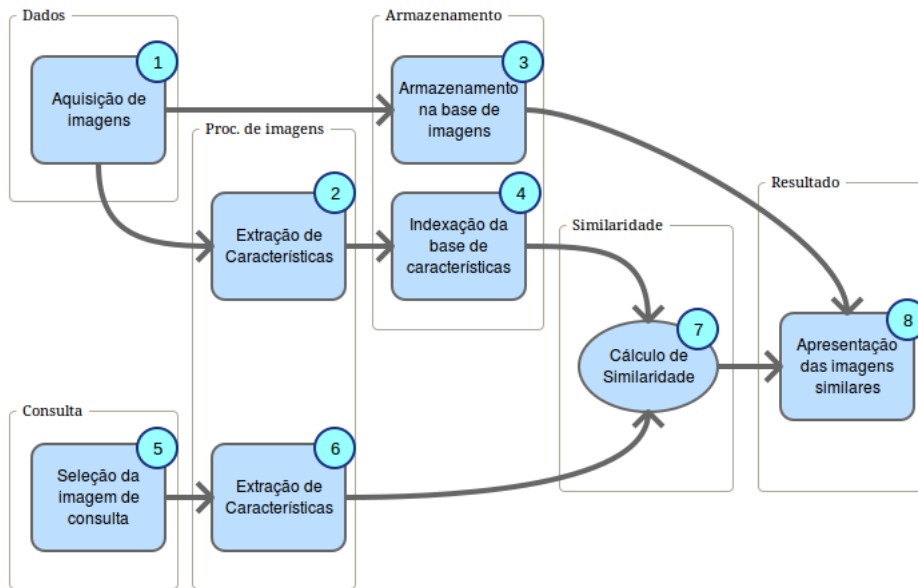
Os primeiros sistemas de recuperação de imagens, iniciados por volta da década de 70, eram baseados em anotações associadas às imagens (i.e. metadados), dessa forma o sistema comparava as palavras providas pelo usuário em uma consulta textual, retornando um conjunto de imagens associadas a tais palavras-chave. Os problemas inerentes a esse tipo de consulta são principalmente o esforço humano necessário para anotar o conjunto de imagens, a subjetividade

de tais anotações e a imprecisão das mesmas devido à ambiguidade das palavras utilizadas. A Figura 1 ilustra tal imprecisão, na qual a partir da especificação da palavra chave “banco” foram recuperadas imagens oriundas de conceitos divergentes, como por exemplo, “banco” como um móvel e “banco” representando capital ou uma instituição financeira. Além de tais problemas, é importante salientar que uma imagem contém mais informações do que se consegue descrever com palavras (DATTA et al., 2008; WANG et al., 2010).



**Figura 1: Exemplo de ambiguidade em uma consulta textual usando a palavra-chave “banco”.**

O termo “recuperação de imagens baseada em conteúdo”, ou *Content-Based Image Retrieval* (CBIR), teve sua origem no início da década de 90, quando foi usado por (KATO, 1992) para descrever experiências de recuperação automática de imagens a partir de um banco de dados, com base nas cores e formas presentes nas imagens (ZHANG; HUANG, 2014). Desde então, para superar desvantagens da recuperação de imagens baseadas em texto, sistemas CBIR realizam consultas baseadas em propriedades visuais da imagem (i.e. características de baixo nível) baseadas em cor, textura e forma. O conteúdo do conjunto de imagens é caracterizado por extratores que codificam a propriedade visual da imagem em vetores de características. Posteriormente as características extraídas são comparadas com um modelo usado para consulta onde métricas mensuram a similaridade entre elas. A partir de tal comparação torna-se possível gerar uma lista ordenada de imagens baseada na similaridade destas em relação a uma dada imagem de referência (i.e. imagem de consulta).



**Figura 2: Fluxo de dados de uma sistema CBIR. Adaptado de Smeulders et al. (2000)**

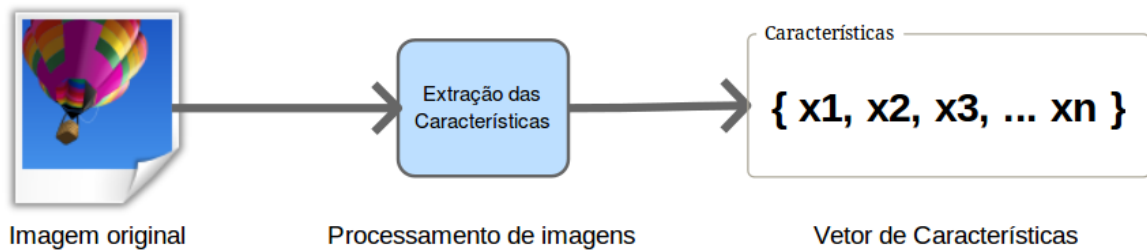
Conforme ilustra a Figura 2, em sistemas CBIR após a etapa de aquisição (passo 1) as imagens são processadas por meio de algoritmos de extração de características que organizam tais características em vetores (passo 2), bem como as imagens são persistidas em uma base de dados (passo 3). Os vetores de características gerados podem ser utilizados para indexação (passo 4), auxiliando na eficiência do processo de recuperação. Quando a base está pronta o sistema CBIR está apto a realizar consultas a partir de um padrão fornecido pelo usuário, sendo nesse caso uma imagem de consulta (passo 5). Tal imagem de consulta também é processada por algoritmos de extração de características (passo 6), e em seguida é aplicada uma métrica (i.e. função de dissimilaridade) para comparar o vetor de características da imagem de consulta, com os vetores de características das imagens provenientes da base (passo 7), recuperando assim as imagens mais similares da base de dados, de acordo com a consulta especificada, retornando as imagens em ordem das mais similares para as menos similares (passo 8) (FARIA et al., 2010).

### 2.1.1 CARACTERÍSTICAS DE IMAGEM

Dados complexos, como imagens, possuem uma grande quantidade de informações tornando difícil a recuperação por parte de sistemas de busca. Uma abordagem utilizada em sistemas CBIR presente na literatura é o uso de algoritmos de processamento de imagens para extração de características, assim é possível representar dados complexos por meio de seus atributos visuais.

Diversos algoritmos são usados para processar uma imagem visando explicitar informações

que possam representar seu conteúdo. Os tipos de características mais comuns utilizadas para representar o conteúdo visual de uma imagem são características de baixo nível baseadas em cor, textura e forma (CHINO et al., 2015). A Figura 3 ilustra o processo de extração de características de uma dada imagem de entrada, gerando ao final um vetor de características com  $n$  posições.



**Figura 3: Extração de características para representação de uma imagem.**

Intuitivamente, pode-se imaginar que quanto maior a quantidade de características extraídas de uma dada imagem, melhor será sua representação. Porém, uma maior dimensionalidade (quantidade de elementos) dos vetores de características diminui o poder discriminatório de tais características, degradando o processo de recuperação. Tal fenômeno é dominado como maldição da alta dimensionalidade (ZHUO et al., 2014).

#### 2.1.1.1 COR

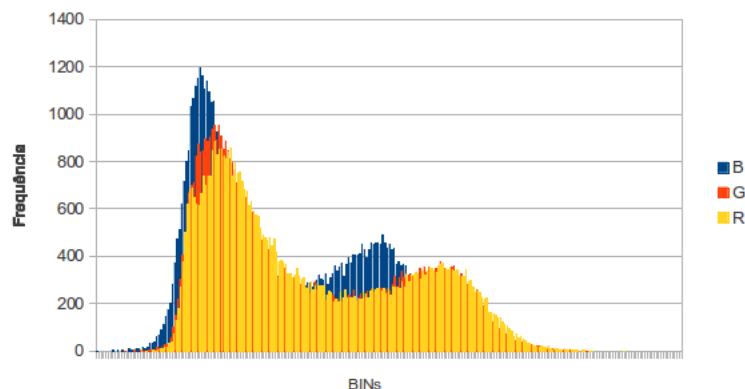
Dentre as características visuais as cores ocupam um papel muito importante na visão humana sendo uma das propriedades mais importantes identificadas. Assim a cor é uma das propriedades mais estudadas na literatura, bem como amplamente aplicada em sistemas de CBIR (SANTOS et al., 2010). A maioria das técnicas de extração de características de imagens utilizam valores de intensidade de cada pixel em suas implementações.

O conceito de cor pode ser representado em diferentes espaços. Por exemplo, uma determinada cor pode ser representada em um espaço composto pelos componentes vermelho, verde e azul (*Red, Green, Blue* - RGB) que representam as três cores primárias na teoria de cores de Goethe (BARROS, 2006; IVANOVA; STANCHEV, 2009). Outras representações que utilizam diferentes componentes como matiz, saturação e valor (*Hue Saturation & Value* - HSV) (AN et al., 2010) podem descrever o sistema de cores de maneira mais intuitiva, segundo Preim e Bartz (2007). No trabalho de Kekre et al. (2010) diversas outras representações do espaço de cores utilizadas em sistemas CBIR são abordadas em um estudo comparativo.

Um primeiro extrator de características que pode ser utilizado para representar uma

imagem é o histograma. O termo estatístico histograma explicita e representa a contagem de ocorrência de certo acontecimento, de maneira mais explícita trata-se da representação gráfica de uma distribuição. A grande vantagem de aplicação de histogramas é sua simplicidade e custo computacional linear, tornando esta técnica uma das mais populares em sistemas CBIR e de processamento de imagens em tempo real (TRAINA et al., 2003). Aplicando o mesmo conceito no contexto de imagens um histograma descreve a frequência de ocorrência de uma dada intensidade de pixel na imagem. Além disso, uma outra característica interessante dos histogramas é que quando normalizados os mesmos apresentam invariância a transformações geométricas.

A Figura 4 ilustra a fusão de três histogramas cada um representando um componente de cor, no caso em questão oriundos do espaço de cores RGB. É importante salientar que o eixo das abscissas refere-se aos níveis de intensidade de pixel em cada um dos componentes de cores, sendo que a cada intervalo discreto do mesmo é dada a denominação de “bin” do histograma. Já o eixo das ordenadas explicitam a frequência de ocorrência de uma determinada intensidade.



**Figura 4: Representação visual da fusão de três histogramas de cores no espaço RGB.**

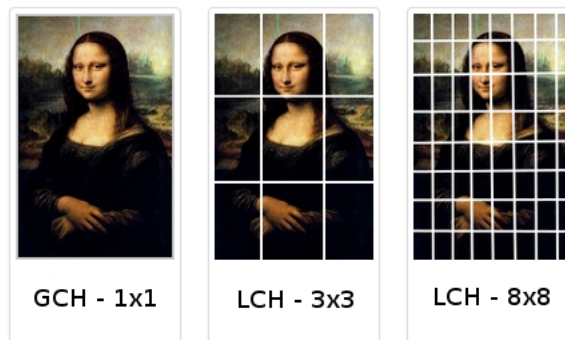
O histograma global de cor (*Global Color Histogram - GCH*) é um descritor muito utilizado na recuperação de imagens, não só por apresentar uma implementação trivial, mas também por possuir um nível de precisão aceitável em muitas aplicações (SWAIN; BALLARD, 1991). O processo de extração de características por meio de tal abordagem consiste num algoritmo que percorre toda a imagem, *pixel a pixel*, gerando os bins da imagem de acordo com o valor de sua intensidade considerando individualmente cada componente de cor. Dessa forma, considerando por exemplo o espaço RGB, são gerados 3 histogramas, um para cada componente, sendo que ao final os mesmos são concatenados para geração do vetor de características.

Já o histograma de cor acumulado (*Cumulative Global Color Histogram - siglaCGCHCumulative*



*Global Color Histogram*) é muito semelhante ao descritor GCH (STRICKER; ORENCO, 1995), a única diferença entre os dois é que o CGCH transforma o mesmo em uma distribuição monotonicamente crescente pois realiza a soma do “*bin*” corrente com seus antecessores, assim o último “*bin*” possui a somatória de todos os anteriores, apresentando portanto valor 1 caso o histograma esteja normalizado.

Uma outra variação da aplicação de histogramas no contexto de imagens é o denominado histograma local de cor (*Local Color Histogram - LCH*), diferente do GCH, o mesmo é extraído por meio de um algoritmo que divide a imagem em janelas (também denominadas como *tiles*), como ilustrado na Figura 5, calculando o histograma de cor de cada janela, ao final é gerado um vetor concatenando todos os histogramas oriundos de cada uma das janelas definidas (LYLE et al., 2012; ALMEIDA et al., 2008).



**Figura 5: Exemplo de divisão de uma imagem para obtenção do histograma local de cor (LCH).**

Além dos extratores baseados em histogramas existem outras propostas mais sofisticadas, porém eficazes e eficientes, como o extrator denominado *Border/Interior Classification* (BIC) (STEHLING et al., 2002), o qual utiliza o espaço de cores RGB uniformemente quantizado em uma determinada cardinalidade, sendo que cada pixel é geralmente representado por 6 bits. Dessa forma apresenta um vetor de características compacto. O BIC tem por finalidade classificar os *pixels* da imagem em: 1) *pixels* de borda e 2) *pixels* de interior, gerando dois histogramas. Um pixel é classificado como de interior se seus quatro vizinhos (vizinhança-4) apresentarem a mesma intensidade, caso contrário é considerado como de borda. Ao final do processo os dois histogramas são combinados em um único vetor de características. Outros tipos de vizinhança (e.g. vizinhança-8), bem como quantizações podem ser aplicadas ao BIC, porém, segundo resultados apresentados pelos autores do mesmo, levando em consideração custo computacional e eficácia, aplica-se vizinhança-4 e *pixels* representados (com profundidade) por 6 bits.

### 2.1.1.2 TEXTURA

A textura é uma propriedade muito importante para a caracterização e reconhecimento de imagens, e por esse motivo diversos trabalhos tem focado seus esforços de pesquisa envolvendo a análise de textura em imagens (MANJUNATH; MA, 1996; XU; ZHANG, 2006; ZEGARRA et al., 2009; VERMA et al., 2015). A textura é uma medida do arranjo estrutural dos *pixels* em uma imagem. Embora não exista nenhuma definição formal para textura, os extratores de textura objetivam mensurar algumas propriedades como regularidade, orientação (direção), suavidade, granularidade, entre outras (GONZALES; WOODS, 2002).

Existem diversas técnicas propostas na literatura para caracterizar a existência de primitivas básicas, cuja distribuição espacial criam padrões visuais definidos em termos de granularidade, direcionalidade e repetição. Essas técnicas podem ser usadas para extrair e representar texturas. Tais técnicas costumam ser classificadas em modelos baseados no espaço, modelos baseados na frequência, e assinaturas de texturas (BIMBO, 1999).

Um exemplo de extrator de características baseado em textura são os descritores de *Haralick* (HARALICK, 1979) os quais utilizam a distribuição dos níveis de cinza por meio da aplicação de matriz de co-ocorrência para avaliar diferenças de textura. As matrizes de co-ocorrência descrevem o relacionamento espacial entre níveis de cinza em uma imagem. Uma célula definida pela posição  $(i, j)$  nessa matriz registra a probabilidade de dois *pixels* de níveis de cinza  $i$  e  $j$  ocorrerem em duas posições relativas. A geração de tal matriz de co-ocorrência leva em consideração uma dada orientação de análise, bem como de distância (i.e. granularidade) entre os *pixels* analisados. Dessa forma, ao definir-se diferentes angulações e distâncias de análise torna-se possível representar relacionamentos entre *pixels* levando em consideração diferentes direcionalidades, granularidades e, por consequência repetições. Após a construção de tais matrizes aplica-se a estas um conjunto de medidas (i.e. energia, entropia e contraste), denominadas como descritores de Haralick, para caracterizar diferentes especificidades das texturas inerentes à uma dada imagem (TORRES; FALCÃO, 2008).

O *Local Binary Pattern* (LBP) é um extrator de características, pode ser usado como um exemplo, que segue a abordagem de métodos baseados na construção de um modelo podendo ser utilizado não somente para descrever a textura, mas também para sintetizá-la. Os parâmetros do modelo capturam as qualidades perceptivas essenciais de uma textura. Por exemplo, podem ser modelados elementos de textura como: um ponto claro ou um ponto escuro, uma transição horizontal ou vertical, cantos ou retas, etc. Extratores que seguem essa abordagem são bons para imagens com texturas definidas (PENATTI, 2009).

Os *Filtros de Gabor* são um conjunto de funções Gaussianas moduladas por funções sinusoidais complexas, bidimensionais. A estrutura local da textura, nos vários canais utilizados, é descrita pela frequência e pela orientação, que são definidas pela função sinusoidal do filtro de Gabor e apresentam características muito úteis para fins de classificação de imagens (GRIGORESCU et al., 2002).

As características representadas por *Multi-scale Histograms* implicitamente inclui a informação espacial da imagem. Esse extrator utiliza um histograma multi-escala, ou seja, um conjunto de frequências de intensidade de uma imagem em múltiplas resoluções (HADJIDEMETRIOU et al., 2001). Assim como os histogramas de cor esse extrator também é invariante para rotações e translações.

*First 4 Moments* (primeiros 4 momentos), média, desvio padrão, assimetria, e curtose são calculadas de imagens, divididas em forma de listras, em quatro diferentes direções (0, 45, 90, 135 graus). Os resultados calculados de cada conjunto de listras são convertidos em um histograma de 3 bins, para cada momento, resultando em 4 momentos x 4 direções x 3 bins obtendo um vetor com 48 características (SHAMIR et al., 2008).

Em Tamura et al. (1978) é proposto o uso de características mais precisas para descrever uma textura, esse conjunto de características são conhecidos na literatura como *Tamura Textures features* e usam índices que medem: a “granularidade”, representando o tamanho das células na textura, as vezes também chamada de “espessura”; o “contraste”, dado por variações de tons de cinza presentes na imagem; a “direcionalidade”, ou seja, a direção predominante entre os elementos constituintes; o “delineamento” que é intensidade da presença de linhas; a “regularidade” com que as primitivas se repetem e a “rugosidade” que mede a sensação de “aspereza” transmitida pela textura.

Outro extrator que também usa estatística para extrair informações da imagem é *Edge Statistics features*, calculado sobre o gradiente de Prewitt combina a média, mediana e variância, tanto da magnitude quanto dos componentes de direção. Outras características de borda são o número total de pixels de orla (normalizado para dimensão da imagem), a homogeneidade da direção e a diferença no sentido de diferentes ângulos (SHAMIR et al., 2008).

*Object Statistics* incluem o número de Euler, mínimo, média, máximo, mediana, variância e um histograma de 10 bins das áreas dos objetos e as distâncias dos mesmos para o centroide da imagem (SHAMIR et al., 2008).

Características de momentos de Zernike (*Zernike features*) são os valores absolutos dos coeficientes da aproximação polinomial de Zernike de uma imagem, essas características são

robustas em relação a compressão e pequenas variações de escala, muitas das características derivadas da transformação morfológica da imagem (i.e, número de objetos, tamanho médio objeto) são também robustas em relação a essas variações (MURPHY et al., 2001).

*Chebyshev-Fourier features*, esse extrator é projetado para caracterizar todas as faixas de domínio espectral de uma imagem - características de baixas a altas frequências. A ideia é a de reter um número finito de termos de expansão, com os coeficientes de expansão a serem utilizados como descritores de imagem. Chebyshev é usada tanto como uma transformada (com ordens correspondentes as dimensões da imagem) e como um conjunto de estatísticas (ORLOV et al., 2007). Características baseadas na *Transformada de Chebyshev* geram um histograma de 32 bins originado de um vetor de 400 elementos produzido pela transformada de uma imagem com ordem de  $N = 20$  (SHAMIR et al., 2008).

### 2.1.1.3 FORMA

A forma é uma propriedade muito relevante, pois possibilita descrever estruturas topológicas de objetos presentes em imagens. Porém, características de forma são mais complexas de serem representadas quando comparadas a cor e textura (TORRES et al., 2008). Apesar disso, uma grande parte da informação semântica de uma imagem está associada aos objetos nela representados, uma pessoa pode até ignorar outras características, como a cor, se sua intenção for buscar por um determinado objeto semelhante (LEE; KIM, 2001).

O reconhecimento de caracteres alfanuméricos, a detecção e reconhecimento de pessoas em sistemas de segurança, o rastreamento de objetos em vídeos, entre outros, são aplicações que inúmeras vezes tiram proveito da identificação, reconhecimento e recuperação de formas em imagens.

Os métodos empregados nos extratores de forma podem se basear em três classes: (1) baseados no contorno do objeto; (2) baseados no conteúdo interno; e (3) baseados em região da imagem. Essas classes se subdividem em estrutural ou global, dependendo se a forma representa o todo ou é dividida em partes menores. A classificação desses métodos podem ainda se ramificar em domínio espacial ou transformado, tendo como base o fato das características de forma serem derivadas de um domínio espacial ou transformado (TORRES; FALCÃO, 2008).

Os extratores globais consideram o contorno ou a região da forma como um todo, sem realizar divisões em subpartes. Os extratores estruturais dividem o contorno ou a região em segmentos e possibilitam a comparação parcial de formas conseguindo, portanto, suportar oclusão de formas. Entretanto, eles são mais complexos computacionalmente, especialmente

para o cálculo de distância (PENATTI, 2009).

Um exemplo de extrator de forma baseado em contorno é o *Curvature Scale Space* (sigla CSS *Curvature Scale Space*). Os contornos simplificados de um determinado objeto na imagem são obtidos por meio de sucessivas suavizações por uma função Gaussiana e cada estágio desta suavização do contorno representa uma escala na curva (*scale-space*). Dessa forma o vetor do CSS representa uma organização multiescala dos pontos de curvatura zero no contorno. Neste sentido, a dimensionalidade do vetor de características varia para formas distintas, requisitando um algoritmo de casamento especial para o cálculo de distância entre dois vetores CSS (OKAL; NUCHTER, 2013).

O extrator *Contour Saliences* (CS) é uma outra abordagem de caracterização de forma, a qual utiliza as saliências da forma do objeto para representação. As saliências de forma são definidas como pontos de maior curvatura ao longo do contorno do objeto. Os vetores de objetos distintos, extraídos por esse algoritmo, analogamente ao extrator CSS, podem possuir tamanhos diferentes, dessa forma o CS necessita utilizar um algoritmo heurístico de casamento entre os contornos, que registra os vetores em relação ao ponto de referência e computa a similaridade (TORRES; FALCÃO, 2008).

A *Trasformada de Radon* é frequentemente utilizada em processamento de imagem para a detecção de retas (ROJBANI et al., 2011) e para obtenção de descritores de forma (TABBONE et al., 2006), calculada para os ângulos de 0, 45, 90, 135 graus, os resultados são usados para se obter um histograma provendo um total de 12-bins.

### 2.1.2 FUNÇÕES DE DISTÂNCIA

Após a extração de características das imagens, as mesmas precisam ser comparadas para determinar, por exemplo o quão próximas duas imagens são. Para tanto são aplicadas as denominadas funções de distância, também conhecidas como funções de dissimilaridade, essas funções tem por objetivo mensurar, como o próprio nome especifica, a distância entre dois elementos, ou em outras palavras o quanto um elemento é distinto de outro em um determinado espaço de dados, no caso em questão imagens representadas por vetores de características  $n$ -dimensionais. Dessa forma, a recuperação de imagens por conteúdo depende intrinsecamente da junção entre as características extraídas das imagens e das funções de distâncias utilizadas para comparar tais características. A distância (i.e. dissimilaridade) entre duas imagens é um número real e positivo, resultado da aplicação de uma função de distância, e quanto menor o valor, menor será a dissimilaridade entre elas (RAZENTE, 2009).

As funções de distância mais utilizadas no processo de recuperação de imagens são as pertencentes à família Minkowski, também conhecida como norma  $L_p$ . A partir de tal família são derivadas as distâncias Manhattan ( $L_1$ ), também conhecida como *City-block*, a distância Euclidiana ( $L_2$ ) e a distância Chebyshev utilizadas em espaços multidimensionais (WILSON; MARTINEZ, 1997), onde, os elementos são representados por  $n$ , um conjunto de números reais identificados pelas coordenadas  $(x_1, \dots, x_n)$ . Considerando dois vetores  $F = (f_1, \dots, f_n)$  e  $G = (g_1, \dots, g_n)$ , então as distâncias  $L_p$  são definidas:

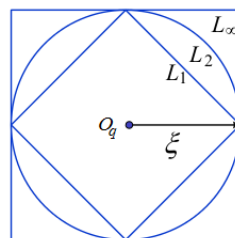
$$L_p(F, G) = \sqrt[p]{\sum_{i=1}^n |f_i - g_i|^p} \quad (1)$$

De acordo com o valor atribuído a  $p$  (Equação 1) obtém-se as variações da família  $L_p$ , por exemplo, quando  $p$  admite o valor 1 obtém-se a distância  $L_1$  (Manhattan), a qual é formalmente definida pela Equação 2. Já quando o valor de  $p$  tender ao infinito obtém-se a distância Chebyshev ( $L_\infty$ ) definida formalmente pela Equação 3. A distância Euclidiana amplamente conhecida e empregada na literatura é obtida quando  $p$  assume valor 2.

$$D(F, G) = \sum_{i=1}^n |f_i - g_i| \quad (2)$$

$$D(f, g) = \max_{i=1}^n |x_i - y_i| \quad (3)$$

As diferentes funções de distância oriundas da família Minkowski apresentam abrangências distintas. A Figura 6 ilustra tais abrangências considerando um espaço bidimensional para um dado centro de consulta ( $O_q$ ) e pontos equidistantes ao mesmo, onde  $\xi$  representa o raio de abrangência.



**Figura 6: Ilustração das áreas de abrangência das distâncias  $L_1$ ,  $L_2$  e  $L_\infty$ .**

Uma outra distância interessante é a distância Canberra, a qual é uma medida muito

sensível a pequenas variações. Tal distância tem sido aplicada em diversas áreas, bem como no contexto CBIR (LAM et al., 2007). Seu comportamento é bastante parecido ao da distância Manhattan, pois é baseada no cálculo das diferenças absolutas de cada dimensão, porém a abrangência da mesma modifica-se à medida que aproxima-se ou afasta-se da origem do espaço. Essa distância é formalmente representada pela equação 4.

$$D(F, G) = \sum_{i=1}^n \frac{|f_i - g_i|}{|f_i + g_i|} \quad (4)$$

Uma outra função de distância aplicada no contexto CBIR é a distância Quadrática, a qual pode ser formalmente definida por:

$$d_A(F, G) = \sqrt{(f - g)^T A (f - g)} \quad (5)$$

onde  $A = [a_{ij}]$  é uma matriz  $N \times N$  de elementos  $a_{ij}$  que denotam a similaridade entre as dimensões  $i$  e  $j$ , e  $f$  e  $g$  dizem respeito a vetores de características. O valor de  $a_{ij}$  é dado por:

$$a_{ij} = 1 - \frac{d_{ij}}{\max[d_{ij}]} \quad \text{onde } d_{ij} = |q_i - c_i| \quad (6)$$

A Distância Quadrática tem sido utilizada em muitos sistemas CBIR para histograma de cores. Foi mostrado que esta distância conduz a resultados perceptualmente mais desejáveis do que a distância Euclidiana, pois considera a similaridade entre diferentes dimensões dos vetores de características.

Uma ampla gama de distâncias pode ser encontrada no trabalho de Cha (2007). Já em Gonçalves et al. (2010) foi realizada uma avaliação sistemática das funções de distâncias utilizadas, dada a influência que as mesmas exercem em sistemas CBIR.

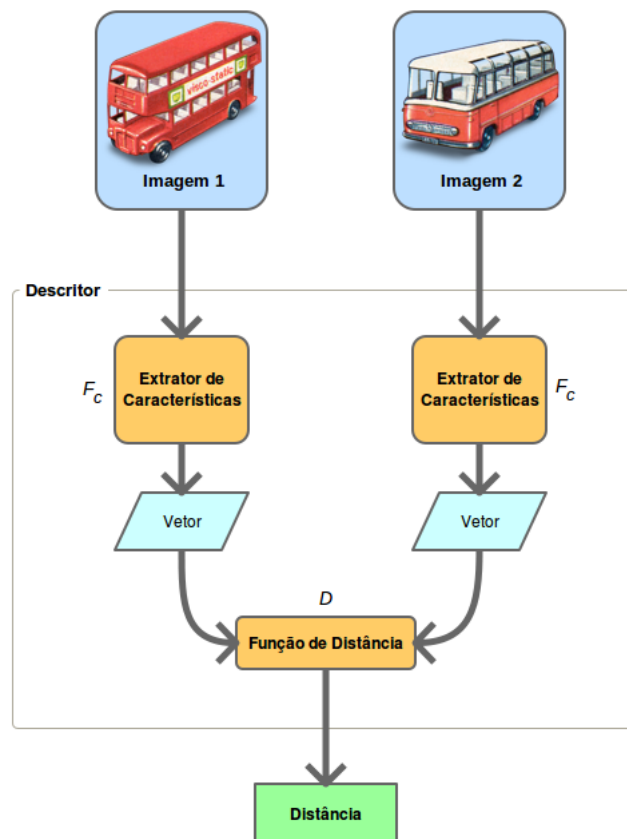
### 2.1.3 DESCRITORES

Para que o processo de recuperação de imagens seja possível, é desejável que as imagens possam ser descritas pelas suas propriedades intrínsecas, tais como forma, textura e cor, normalmente representadas em vetores de características. Neste sentido, os descritores de imagens são utilizados para extrair e comparar esses vetores, tornando possível a busca de imagens (TORRES et al., 2008).

Conceitualmente um descritor pode ser entendido como o responsável por quantificar o quão semelhantes são duas imagens. Formalmente, um descritor de imagens pode ser

definido como uma tupla  $(F_c, D)$ , onde  $F_c$  representa um algoritmo responsável por extrair características visuais da imagem e organizá-las em um vetor de características, e  $D$  uma função de dissimilaridade usada na comparação entre duas imagens, como ilustra a Figura 7.

É importante entender que uma vez que a recuperação de imagens é aplicada em diversas áreas, num determinado domínio de conhecimento (e.g. imagens médicas, sensoriamento remoto, etc) um descritor pode ser considerado melhor que outro de acordo com a quantidade de imagens relevantes que consegue recuperar. Além disso, o mesmo pode ocorrer, obviamente, para diferentes usuários inseridos em um mesmo domínio (i.e. inter-observador), visto que os mesmos podem possuir intenções distintas, bem como percepções visuais diferentes. Essa mesma condição ainda pode ocorrer de maneira que um mesmo indivíduo (i.e. intra-observador) em momentos distintos apresente diferentes intenções para uma dada recuperação, o que pode ocorrer devido a diferentes fatores, tanto positivos (e.g. ampliação do conhecimento) quanto negativos (i.e. cansaço). Dessa forma, a definição de descritores ótimos não é um processo trivial, uma vez que demanda diferentes componentes variáveis.



**Figura 7: Esquema conceitual de um descritor de imagens.**



## 2.2 SELEÇÃO DE ATRIBUTOS

Normalmente as bases de imagens usadas para alimentar sistemas CBIR possuem um grande número de atributos, muitas deles inseridos com o objetivo de uma melhor representação dos grupos presentes na base. Porém, na maioria das situações, muitos desses atributos são redundantes e/ou irrelevantes. Assim, selecionar os atributos de maior relevância oferece benefício tanto a performance quanto a precisão de sistemas CBIR.

Processos de seleção de atributos atuam transformando um determinado espaço de dados em espaço de características, de menor dimensão, mas que ainda possua a maior parte da informação intrínseca dos dados. Dessa forma, o conjunto de dados tem a dimensionalidade reduzida sem perder a precisão em consultas por similaridade. Os métodos de seleção de atributos promovem a escolha, dentre todos os atributos, quais os mais relevantes do ponto de vista da informação (GUYON; ELISSEEFF, 2003), pois vetores de características com uma quantidade alta de dimensões contém um grande número de informações redundantes que possuem certa correlação entre seus valores (ZHUO et al., 2014).

A escolha da métrica de avaliação influencia fortemente o algoritmo, e são essas métricas de avaliação que distinguem entre as três categorias principais de algoritmos de seleção de atributos: *wrappers*, filtros e métodos incorporados. *Wrappers* utilizam algoritmos de aprendizado de máquina como uma caixa preta para selecionar subconjuntos de atributos de acordo com seu poder preditivo. Filtros selecionam subconjuntos de atributos como um passo de pré-processamento, independente do algoritmo de predição escolhido. Métodos incorporados executam a seleção de atributos durante o processo de treinamento e são geralmente específicos a determinados algoritmos de aprendizagem de máquina (CHANDRASHEKAR; SAHIN, 2014).

## 2.3 CONSULTAS POR SIMILARIDADE

Quando um usuário submete uma imagem ou uma amostra para uma consulta à base de imagens ele recebe uma resposta em forma de lista, onde as imagens presentes nessa listagem são ordenadas pela similaridade. Para que um sistema CBIR consiga organizar a resposta ao usuário de tal forma são utilizadas operadores de consulta por similaridade.

As consultas por similaridade visam recuperar objetos, por exemplo dados complexos, dentro de um limite de proximidade definido por uma função de distância. Essas consultas, geralmente, possuem dois tipos de limites definidos a partir de um dado centro de consulta. Tais limites são: (i) definido por raio de abrangência, o qual gera a denominada consulta por

abrangência e (ii) definido pela quantidade de vizinhos mais próximos, o qual gera a chamada consulta aos  $k$ -vizinhos mais próximos.

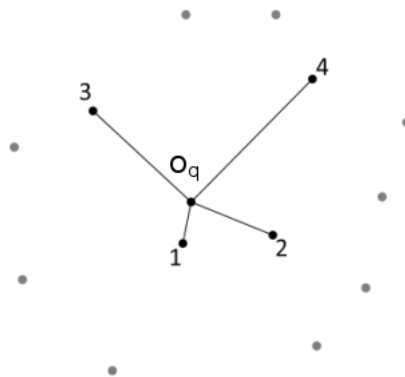
A principal diferença entre os dois tipos de consultas mais amplamente empregados está na forma de se definir o limite de objetos a serem recuperados, no primeiro caso define-se a distância máxima (raio de abrangência) que os objetos podem estar afastados de um determinado centro de consulta. Já no segundo, define-se a quantidade de objetos (vizinhos) mais próximos ao centro de consulta. Assim enquanto num tipo o raio de abrangência é fixo independente do número objetos recuperados, no outro o número de objetos que é fixado independente do raio de abrangência que o centro de consulta possa “alcançar”.

### 2.3.1 VIZINHOS MAIS PRÓXIMOS

A técnica de  $k$ -vizinhos mais próximos ou comumente chamada de  $kNN$  é usada tanto para executar consultas quanto para classificar um dado objeto. A consulta executada é baseada na similaridade ou distância tendo como condição o número de vizinhos mais próximos. Tal consulta é formalmente representada por:

$$kNN(O_q, k) = A = \{a | a \in S, \forall o \in S - A, d(s_q, a) \leq d(s_q, o), |A| = k\} \quad (7)$$

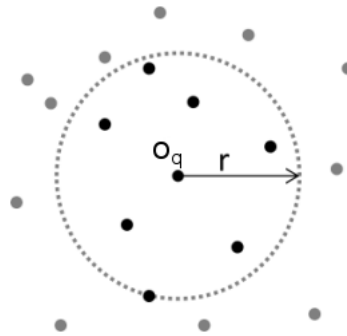
onde dado um conjunto de objetos  $S$ , é escolhido um objeto qualquer deste conjunto,  $s \in S$ . Faz-se então uma consulta aos  $k$ -vizinhos mais próximos  $kNN(O_q, k)$ , onde  $O_q$  representa o objeto de consulta e  $k$  a quantidade de vizinhos, o qual busca-se encontrar o subconjunto  $A \subseteq S$ . (FILHO et al., 2001). Graficamente poderíamos representar uma consulta  $kNN$  com  $k = 4$  conforme ilustrado pela Figura 8.



**Figura 8: Exemplo de uma consulta kNN.**

### 2.3.2 ABRANGÊNCIA

Em uma consulta por abrangência (*range query*) é fornecida uma imagem de exemplo e um raio de alcance, tendo como objetivo recuperar as imagens que possuam distância igual ou menor que o raio. Formalmente a resposta é representada pelo conjunto  $R$  contendo as imagens  $I$  da base que estejam dentro da área de abrangência do raio  $r$  da imagem de consulta  $O_q$ , assim sendo,  $R_{rq} = \{I | distancia(I, C) \leq r\}$ .



**Figura 9: Exemplo de uma consulta por abrangência.**

### 2.4 ABORDAGENS PARA REDUÇÃO DA LACUNA SEMÂNTICA

Um dos maiores desafios que tem sido alvo de muita pesquisa na literatura é a lacuna semântica, definida como a divergência entre a resposta de um sistema CBIR e aquela esperada pelo usuário. As abordagens visando diminuir esse problema são muito importantes e tem incentivado diversos pesquisadores a desenvolverem técnicas com o objetivo de identificar a semântica existente nos contextos das imagens (SILVA, 2014).

A interação com o usuário é usada em sistemas de recuperação como forma de atenuar a lacuna semântica. Apesar de muitos algoritmos sofisticados terem sido projetados para descrever características de baixo nível de uma imagem, eles não são adequados para lidar com a semântica e possuem muitas limitações em uma grande base de dados. Muitos experimentos envolvendo sistemas CBIR tem mostrado que características de baixo nível não são suficientes para descrever conceitos semânticos de alto nível (DARWISH; ALI, 2015).

Existem diferentes abordagens para tratamento da lacuna semântica como por exemplo a utilização de métodos para a personalização de consultas por meio da utilização de perfis de usuário, bem como técnicas de realimentação de relevância provendo a aproximação do usuário com o processo de recuperação de imagens por conteúdo. Além disso, a composição de ambas as abordagens pode ser amplamente explorada.

### 2.4.1 PERFIL DE USUÁRIO

A principal motivação para utilização de perfis de usuário, em sistemas de recuperação de imagens baseada em conteúdo, é justificada pelas diferentes formas de interpretação de uma mesma imagem por usuários diferentes, sendo assim, o resultado obtido está estreitamente ligado com a intenção de um dado usuário em um determinado momento (ABOULMAGD et al., 2009).

A personalização na recuperação de informação consiste em fornecer resultados de pesquisa que se adaptam às necessidades de informação do usuário individual e combinar seus interesses em vez de fornecer os mesmos resultados de uma consulta para todos os usuários (QIU; CHO, 2006). No trabalho de Kacem et al. (2014) são abordadas técnicas de consulta baseadas em tópicos de preferência do usuário, já Wang et al. (2006) propõe uma abordagem que calcula o grau de importância que cada documento multimídia possui, tendo como base informações de perfil do usuário. O trabalho de Miyazawa et al. (2013) traz uma abordagem híbrida, combinando busca por conteúdo em imagens e busca textual baseada em descrições e ações dos usuários, e Song et al. (2014) apresenta uma técnica de recomendação baseada na personalização do usuário com o objetivo de melhorar o desempenho de sistemas CBIR.

Um perfil de usuário pode ser definido como uma coleção de informações pessoais, isto é, um modelo que representa o usuário dentro de um sistema CBIR. Assim, os perfis podem ser utilizados pelo sistema para explorar as preferências do usuário. Os perfis de usuário são comumente usados por sistemas de busca com o objetivo de fornecer recomendações de acordo com as preferências do usuário, utilizando para esse fim técnicas que agrupam usuários com interesses semelhantes. Considerando-se o campo da recuperação de imagens, o conceito de perfil de usuário pode ser usado para capturar e manter a semântica de um determinado especialista em um determinado domínio de conhecimento, por exemplo, imagens médicas (BUGATTI et al., 2014).

Os perfis podem ser divididos em duas categorias principais: perfis estáticos e perfis dinâmicos. No perfil estático os dados armazenados são previamente conhecidos e fornecidos diretamente pelo usuário por meio de um cadastro no sistema ou preenchimento de questionários manuais. São exemplos de dados estáticos a especialidade do usuário, idade, experiência, entre outros, conforme contexto abordado. Já no perfil dinâmico, ao contrário do estático, são capturados dados da interação do usuário com o sistema visando conhecer os interesses do mesmo gerando um perfil comportamental. Posteriormente, a análise desses dados pode fornecer informações importantes que irão ajudar a melhorar a precisão do sistema em futuras consultas realizadas pelo mesmo usuário. Dessa forma, se torna possível, por exemplo,

saber qual descritor irá recuperar imagens mais próximas do interesse/intenção de um determinado usuário.

Portanto a abordagem de perfis é de extrema importância e pode contribuir para melhorar a qualidade das consultas realizadas pelos usuários em sistemas CBIR.

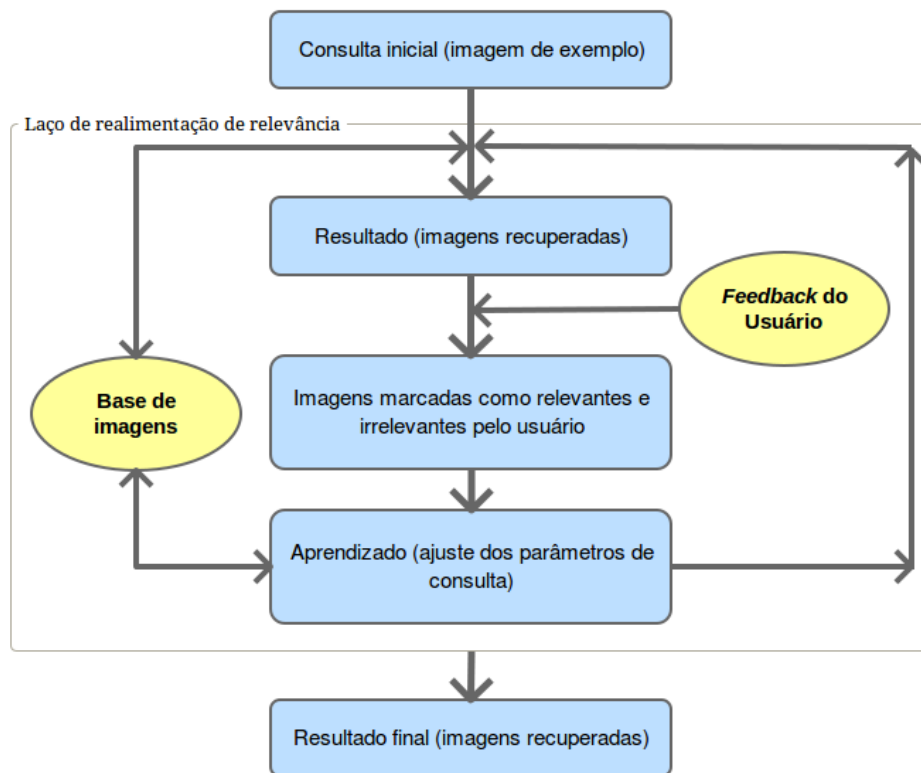
#### 2.4.2 REALIMENTAÇÃO DE RELEVÂNCIA

O métodos de realimentação de relevância caracterizam-se como uma abordagem para diminuição da lacuna semântica. Originalmente foram desenvolvidos para sistemas de recuperação de informações baseados em texto, realimentação de relevância, também chamado de *Relevance Feedback*, é uma técnica implementada para refinar o resultado de uma consulta com base na avaliação do usuário. A ideia principal por trás dessa técnica é mostrar os resultados que são retornados inicialmente a partir de uma determinada consulta e usar a informação sobre a relevância dos elementos retornados para realização de uma nova consulta, assim o sistema itera até que o usuário esteja satisfeito. Essa técnica provê um arcabouço eficiente para mapear semântica de alto-nível para características de baixo-nível usando a interação com o usuário (DOULAMIS; DOULAMIS, 2006), porém somente a realimentação por relevância não é o suficiente para preencher a lacuna semântica entre as características de baixo nível e a semântica na intenção do usuário mas tem se mostrado uma excelente alternativa de ligação entre eles. Basicamente três passos compõe o processo de realimentação por relevância (ZHOU; HUANG, 2003):

1. Por meio de uma imagem fornecida pelo usuário o sistema CBIR apresenta resultados iniciais da consulta;
2. O usuário avalia as imagens retornadas indicando o grau de relevância e/ou irrelevância;
3. Levando em consideração avaliação da consulta, realizada pelo usuário, o sistema processa a consulta novamente e retorna ao passo 2, iterando até que o usuário esteja satisfeito.

Os passos 2 e 3 são repetidos até que o usuário esteja satisfeito com os resultados apresentados pelo sistema. A Figura 10 ilustra um diagrama simples de recuperação de imagens baseada em conteúdo com realimentação de relevância.

Geralmente técnicas de realimentação de relevância são divididas em 2 tipos distintos de ação: Ações que movimentam o centro de consulta, e ações que modificam a medida de similaridade usada para ordenar o resultado, como a reponderação de consulta.



**Figura 10: Recuperação de imagens baseada em conteúdo com realimentação de relevância.**

#### 2.4.2.1 MOVIMENTAÇÃO DO CENTRO DE CONSULTA

Essa abordagem, também conhecida por *Query Point Movement* (QPM), consiste em movimentar o centro de consulta conforme a orientação do usuário por meio de sua interação com o sistema, assim o usuário indica quais são as imagens relevantes e irrelevantes. Com base nessas informações um novo centro de consulta é calculado. As estratégias de Movimentação do Centro de Consulta são divididas em duas principais categorias: Movimentação de um único Centro de Consulta; e Movimentação de Múltiplos Centros de Consulta.

O algoritmo de Rocchio (ROCCHIO, 1971) é usado em estratégias de Movimentação de um único Centro de Consulta, diversos trabalhos na área de recuperação de imagens (ZHANG; WANG, 2012; BUGATTI et al., 2011; DOULAMIS; DOULAMIS, 2006; LIU et al., 2007) aplicam esse algoritmo para realizar movimentações do centro de consulta. Rocchio é um algoritmo clássico de realimentação por relevância baseado em um modelo de espaço vetorial, esse método propõe alterações na consulta original para se obter uma consulta ideal, o vetor de consulta original é atualizado pelos documentos relevantes e irrelevantes. A abordagem de Rocchio é definida formalmente como:

$$\vec{q}_m = \alpha \vec{q}_o + \beta \left( \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j \right) - \gamma \left( \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j \right) \quad (8)$$

onde  $\vec{q}_m$  representa o elemento  $m$  do vetor de características da nova consulta;  $q_o$  é o elemento do vetor de características do centro antigo,  $D_r$  e  $D_{nr}$  são conjuntos contendo os vetores de características das imagens rotuladas como relevantes e irrelevantes, respectivamente, e  $|D_r|$  a quantidade de imagens relevantes e  $|D_{nr}|$  de imagens irrelevantes;  $\alpha$ ,  $\beta$  e  $\gamma$  são respectivamente os fatores de ponderação do centro de consulta antigo, do conjunto de imagens relevantes e do conjunto de imagens irrelevantes.

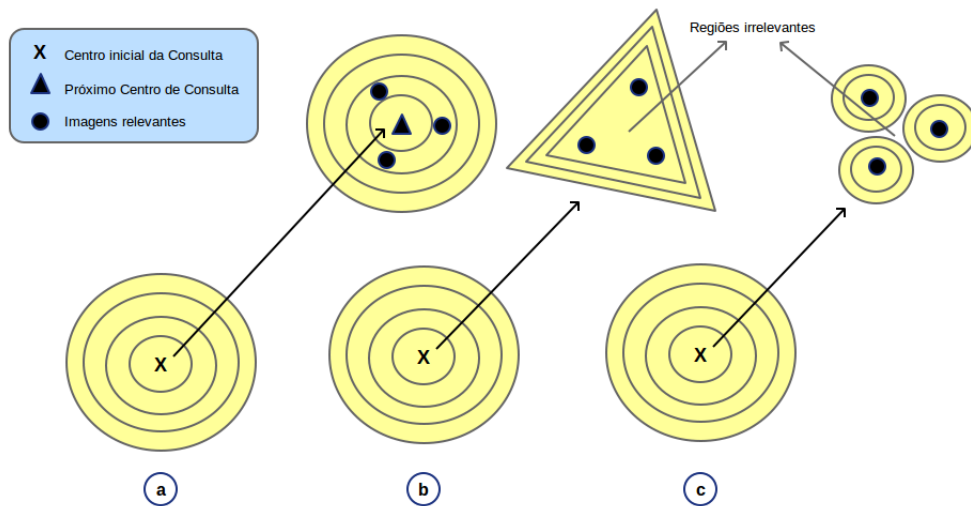
Em outra abordagem, um único centro no espaço de características é calculado usando a média dos vetores de todas as imagens indicadas pelo usuário como relevantes. A Figura 11(a) representa a ideia da abordagem de Movimentação de um Único Centro de Consulta. Essa abordagem é reportada nos trabalhos de Doulamis e Doulamis (2006), Silva et al. (2011), Mosbah e Boucheham (2014). Mais especificamente, em cada iteração de realimentação de relevância, o usuário marca diversas imagens como relevantes atribuindo um grau de relevância a elas. Em seguida, gera-se um centroide, de acordo com o grau de relevância/irrelevância atribuído aos vetores de características das imagens selecionadas como relevantes/irrelevantes e posteriormente o mesmo é utilizado para uma consulta refinada. O cálculo da nova consulta pode ser formalmente representado por:

$$f_{novaConsulta} = \frac{\sum_i n_i y_i}{\sum_i n_i}, \quad (9)$$

onde  $n_i$  é o grau de relevância para a imagem selecionada na posição  $i$ , sendo  $i = 1, \dots, m$  ( $1 \leq m \leq M$ ), e  $y_i$  é o vetor de características dentre  $M$  imagens selecionadas.

Já a abordagem de Movimentação de múltiplos centros de consulta utiliza várias consultas simultâneas, e técnicas de expansão de consulta (*Query Expansion - QEX*) (SIVAKAMASUNDARI; SEENIVASAGAM, 2012; BERNHARD, 2010), *Qcluster* (LIU et al., 2009) e *Top-k* (LIU et al., 2009; RAHMAN et al., 2011) podem ser consideradas como exemplos dessa abordagem.

Na técnica de expansão de consulta (QEX) o usuário rotula diversas imagens como relevantes. Dessa forma um pequeno número de centros representativos são selecionados e usados para definir múltiplas consultas. Para extrair um centro representativo, é usado um algoritmo de agrupamento (PORKAEW; CHAKRABARTI, 1999), e os centroides dos grupos são usados como representativos. Na construção da consulta, imagens consideradas relevantes



**Figura 11: Exemplos de movimentação de consulta: (a) de um único centro; (b) múltiplos centros em forma convexa; (c) múltiplos centros em forma côncava.**

em iterações de RF anteriores são incorporadas nos grupos, tornando o sistema mais confiável em relação às preferências do usuário.

Assim, a medida de similaridade de uma imagem a partir de uma consulta com múltiplos centros é definida como a combinação ponderada das distâncias entre os representantes individuais selecionados na consulta. Uma representação gráfica da expansão de consulta é mostrada na Figura 11(b).

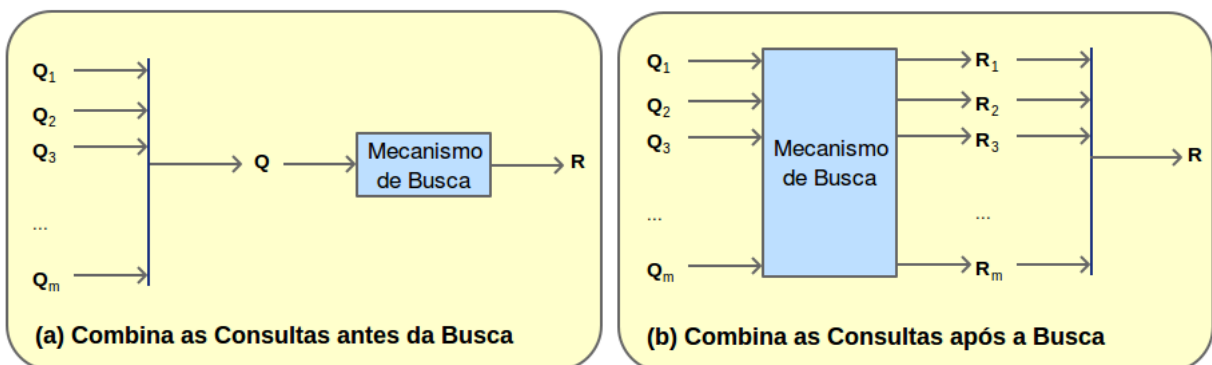
Em contraste com a abordagem de movimentação de um único centro de consulta, QEX implicitamente muda a função de distância pela soma dos centros relevantes e removendo os irrelevantes. Além de que QEX captura grupos locais, enquanto que uma consulta simples trata todas as imagens relevantes como equivalentes (mantendo a proporção do grau de relevância). Porém, QEX possui um custo computacional mais elevado visto que envolve agrupamentos múltiplos (DOULAMIS; DOULAMIS, 2006).

A abordagem QEX assume que imagens relevantes são mapeadas para centros próximos, no que diz respeito a uma medida de similaridade. Por esta razão, um único grande contorno é estabelecido de modo a abranger todos os centros de consulta. No entanto, é muito provável que as imagens relevantes selecionadas acabem por mapear uma grande área no espaço de características por meio dos grupos, dessa forma é estabelecido um único contorno através dos pontos da consulta que provavelmente contenham regiões do espaço de características com imagens irrelevantes. Isso é ilustrado na Figura 11(b), onde o espaço existente entre os pontos de consulta indicam uma região no espaço de características contendo imagens irrelevantes para as preferências do usuário. O contorno único obtido por meio da QEX não consegue excluir essa



região irrelevante. Para superar essa dificuldade, os autores de (KIM; CHUNG, 2003) apresentaram a abordagem *Qcluster*, um novo método de agrupamento adaptativo consistindo em dois processos; o de classificação e o de junção dos grupos. O processo de classificação coloca cada imagem relevante em um dos agrupamentos atuais ou em um novo agrupamento. Então, o processo de junção dos agrupamentos reduz o número de grupos unindo os agrupamentos de conteúdo semelhante reduzindo também o número de centros de consulta em cada iteração. Por fim, os representantes dos agrupamentos gerados a partir das imagens relevantes, no conjunto classificado, compõem o conjunto dos novos centros de consulta. A representação gráfica dessa ideia é ilustrada pela Figura 11(c).

Na abordagem *Top-k* (JIN; FRENCH, 2005) ao invés de se combinar um número de consultas antes da pesquisa, como fazem os métodos anteriores, as pesquisas são executadas individualmente pelo sistema CBIR e os resultados recuperados de cada consulta são devidamente combinados posteriormente para melhorar o desempenho da recuperação. A Figura 12 ilustra uma representação esquemática comparando a abordagem anterior (a) com a abordagem apresentada (b). Neste caso, as  $k$  imagens mais relevantes (incluindo a imagem de consulta) são selecionadas a partir da lista de todas as imagens recuperadas. Após a execução das consultas *Top-k* considerando as imagens rotuladas como relevantes como centros, as imagens são selecionadas entre todas as imagens retornadas pela dissimilaridade entre cada imagem retornada e a imagem usada para a consulta (RAZENTE, 2009). Obviamente, o pior caso é quando não há nada relevante, exceto a consulta original. Já o melhor caso é quando existe um número maior de imagens relevantes que  $k = 1$ , incluindo a consulta original. Essa estratégia reduz significativamente a complexidade computacional. Somente uma parte das imagens são retornadas para o usuário, e ele só visualizará as imagens relevantes nesse conjunto menor.



**Figura 12: Estratégias para recuperação de imagens com múltiplas consultas.**

### 2.4.2.2 REPONDERAÇÃO DE CONSULTA

Uma outra abordagem de realimentação de relevância são as técnicas que atualizam a medida de similaridade. A medida mais comumente usada para recuperação de dados multimídia é a distância Euclidiana. Dentro do espaço de características, no entanto, a distância Euclidiana considera de igual importância todos os elementos do vetor de características. Dessa maneira não é apropriada para esquemas de realimentação de relevância que considere cada dimensão do vetor de característica com diferentes graus de importância. Assim, uma medida de similaridade ponderada é necessária para adaptar a resposta do sistema conforme as intenções e necessidades de informação do usuário. Um exemplo dessa técnica é a distância Euclidiana ponderada (*Weighted Euclidean Distance* - WED), a qual é formalmente definida como:

$$d_w(f_q, f_j) = (f_q - f_j)^T \cdot \Lambda \cdot (f_q - f_j) \quad (10)$$

onde  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$  representa os elementos da diagonal principal da matriz,  $N$  representa o número de características que o vetor possui, sendo  $n = (1, 2, \dots, N)$ , o elemento  $\lambda_n$  expressa a importância da  $n$ -ésima característica do vetor para a função de distância. O vetor  $f_q$  corresponde ao vetor de características da imagem de consulta, enquanto  $f_j$  ao vetor de característica da  $j$ -ésima imagem na base de imagens.

O objetivo da técnica de realimentação de relevância WED é prover uma forma para estimar graus de relevância para cada dimensão do vetor de características. Isto é realizado por meio de um conjunto de imagens relevantes/irrelevantes selecionadas pelo usuário de acordo com suas necessidades e preferências de informação. Uma vez estimados esses parâmetros são armazenados em um vetor  $\lambda$ .

A abordagem heurística mais popular para estimar os parâmetros  $\lambda$  é a arquitetura MARS (RUI et al., 1998). Nesse caso em particular, assume-se que  $m$  imagens das  $M$  melhores (i.e. consulta aos  $k$ -vizinhos) recuperadas são selecionados pelo usuário como relevantes e que  $y_i, i = 1, \dots, m (1 \leq m \leq M)$  contém os vetores de características de  $m$  imagens selecionadas. Diferentes graus de relevância podem ser levados em conta através da multiplicação do vetor  $y_i$  com um fator  $n$  indicando o grau de relevância da amostra selecionada.

O conceito da presente abordagem heurística é que se um elemento de característica em particular, por exemplo, o  $y_{i,n}$ , onde  $n$  representa a  $n$ -ésima característica do vetor, será consistente se o desvio padrão (Equação 11) do respectivo elemento de característica sobre todas as  $m$  amostras relevantes selecionadas, ou seja para todos  $i = 1, \dots, m$ , é pequeno.

$$\sigma_n = \frac{1}{m} \sqrt{\sum_{i=1}^m (y_{i,n} - \mu_n)^2} \quad (11)$$

Na equação 11,  $\mu_n$  é a média do  $n$ -ésimo elemento de característica das  $m$  amostras selecionadas, i.e.,  $\mu_n = (1/m) \sum_{i=1}^m y_{i,n}$ . Caso o valor de  $\sigma_n$  seja alto, isso indica que o respectivo elemento de característica não é interessante para o objetivo do usuário. Como resultado, os parâmetros  $\lambda$  são calculados como o inverso do desvio padrão:

$$\lambda_n = \frac{1}{\sigma_k} \quad (12)$$

Além da atualização de pesos, a abordagem MARS também propõe uma técnica de movimento de consulta, de modo que a nova consulta se mova em direção as imagens relevantes e se afaste das irrelevantes.

Embora a ideia por trás da abordagem MARS seja bastante razoável continua sendo uma técnica *ad hoc*. Por essa razão, existe a necessidade de uma atualização ótima para estimar os parâmetros  $\lambda_n$  (DOULAMIS; DOULAMIS, 2006).

Para realizar uma análise automática de algoritmos de RF, a pseudo realimentação de relevância (*Pseudo Relevance Feedback* - PRF) automatiza a parte manual da realimentação de relevância, simulando a interação com um usuário (LIN et al., 2015). Após a consulta por similaridade, uma fração das imagens recuperadas são definidas como positivas quando pertencerem à classe da imagem de consulta, e como negativas se não pertencerem. Assim, torna-se possível executar experimentos com algoritmos de RF simulando a intenção do usuário.

## 2.5 ALGORITMOS EVOLUTIVOS

Dentro da área de inteligência artificial, os algoritmos evolutivos (AEs) fazem parte de uma categoria chamada de computação evolutiva. A ideia por trás desses algoritmos é aplicar o processo de evolução natural como paradigma de solução de problemas, a partir de sua implementação.

Os primeiros AEs surgiram em trabalhos na década de 1930 (WRIGHT, 1930), onde sistemas evolutivos naturais passaram a ser observados e posteriormente convertidos em algoritmos de exploração de múltiplos picos de uma função objetivo. Apesar disso, foi apenas com o surgimento dos computadores, por volta da década de 1960, que os desenvolvimentos de AEs foram intensificados, possibilitando diversos estudos teóricos e empíricos (GABRIEL;

DELBEM, 2008).

Em geral, esses algoritmos não garantem encontrar a solução ótima, porém, avaliações estatísticas mostram que eles tem alcançado com regularidade a finalidade de retornar uma solução de boa qualidade em tempo computacional aceitável (DORIGO; BLUM, 2005). Para que alcancem tais resultados os evolutivos usam dois processos básicos: *Exploration* e *Exploitation*, respectivamente esses processos são responsáveis por visitar novas regiões do espaço de busca e refinar regiões já visitadas (ČREPINŠEK et al., 2013).

Atualmente percebe-se na literatura um crescente interesse por algoritmos evolutivos na pesquisa de sistemas CBIR, dessa forma diversas alterações tem surgido trazendo ganho de eficiência e eficácia aos sistemas CBIR atuais.

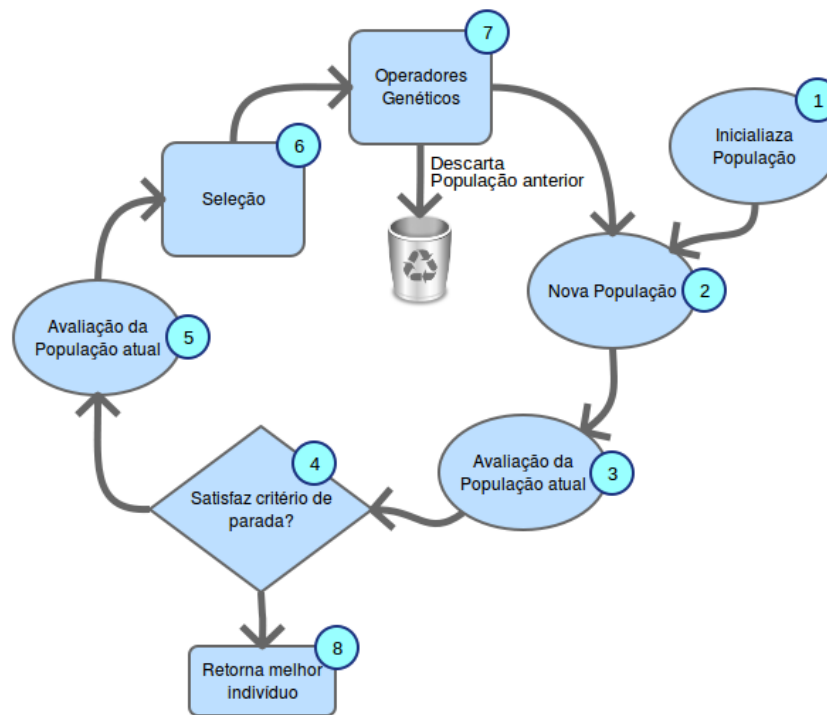
### 2.5.1 ALGORITMO GENÉTICO

O algoritmo genético (*Genetic Algorithm* - GA), é uma heurística de busca que procura reproduzir o processo de evolução natural para resolver problemas. Existe uma grande variedade de modelos computacionais, mas em comum, implementam o conceito de evolução das espécies, operadores de seleção, mutação e reprodução. O resultado está relacionado ao desempenho simulado dos indivíduos dentro de um ambiente computacional. Usualmente, o processo evolutivo dos GAs é encerrado quando os indivíduos da população não mais evoluem, ou quando é atingido o número máximo de gerações preestabelecidas (HAUPT; HAUPT, 2004).

O ciclo de execução de um GA típico é demonstrado pela Figura 13. Inicialmente uma população de possíveis soluções para o problema é gerada de forma aleatória (passos 1 e 2). Os indivíduos dessa população são avaliados (passo 3) para verificar se o critério de parada foi atingido (passo 4), quando isso ocorrer o melhor indivíduo é retornado (passo 8) como uma possível solução do problema. Porém, se o critério de parada não foi alcançado a população é novamente avaliada (passo 5) e o processo de seleção (passo 6) e operadores genéticos (passo 7) são executados. Assim a população antiga é descartada dando origem a uma nova população (passo 2) reiniciando o ciclo até que o critério de parada seja satisfeito.

#### 2.5.1.1 CROMOSSOMOS

Um GA possui aspectos comuns e invariáveis que permitem a reutilização de um mesmo código para problemas distintos. O primeiro passo para usar um algoritmo genético é criar uma representação de uma possível solução para problema que se pretende resolver. Em GA a representação cromossomial, representada por um conjunto de genes, consiste em traduzir



**Figura 13: Esquema de um GA típico.**

o problema para um forma viável de ser processada pelo computador (LINDEN, 2008). Inicialmente os GAs adotavam exclusivamente a representação binária, mas atualmente representações baseadas em números inteiros e reais são bastante utilizadas. A escolha de uma forma de representação depende muito do problema a ser solucionado pelo GA.

### 2.5.1.2 FUNÇÃO DE APTIDÃO

A função de aptidão, também conhecida como *fitness*, é responsável por determinar a qualidade de um indivíduo como solução do problema abordado. Ao longo dos estudos sobre GAs, pesquisas tem mostrado que a especificação de uma função de aptidão apropriada é o ponto crucial para o seu desempenho. É esperado que a mensuração fornecida por essa função seja capaz de diferenciar, na proporção correta, as soluções que são promissoras das que são inadequadas (HAMDANI et al., 2011; SILVA et al., 2009).

A maior demanda computacional na execução de um GA normalmente está associada com a função de aptidão, isso se deve ao fato que novos indivíduos são gerados a cada ciclo evolutivo, sendo que esses indivíduos devem ser avaliados sistematicamente. Com o objetivo de diminuir essa carga computacional, o trabalho de (HAUPT; HAUPT, 2004) propõe alguns cuidados especiais na hora de implementar tal função: 1) não gerar cromossomos idênticos na população inicial; 2) manter a distinção entre os indivíduos, garantindo que a reprodução não irá

gerar cromossomos idênticos e 3) criar uma memória para os GAs, com o objetivo de descartar os indivíduos gerados anteriormente.

### 2.5.1.3 SELEÇÃO

O método de seleção simula o mecanismo de seleção natural onde pais mais capazes geram mais filhos, porém pais menos capazes também geram descendentes. Nesse processo, indivíduos (cromossomos) com notas altas atribuídas pela função de avaliação devem ser privilegiados, mas sem desprezar completamente indivíduos com avaliação muito baixa, pois mesmo esses indivíduos podem possuir características que sejam favoráveis para a geração de um indivíduo ótimo (LINDEN, 2008).

A maneira pela qual os cromossomos são selecionados pode variar, dependendo do operador de seleção utilizado. Os operadores mais populares são: **seleção por roleta** onde para cada cromossomo da população é atribuída uma fatia de uma roleta imaginária, sendo o tamanho dessa fatia proporcional à aptidão do cromossomo, a cada giro dessa roleta um indivíduo é selecionado; **seleção por torneio** onde a ideia é promover uma disputa entre um grupo de  $N(N \geq 1)$  cromossomos tomados da população aleatoriamente (torneio simples) ou usando a roleta (torneio estocástico), o indivíduo com maior valor de aptidão no grupo é selecionado; **seleção elitista (elitismo)** onde  $N(N \geq 1)$  melhores indivíduos da população corrente são selecionados diretamente, este operador é normalmente acoplado a outros operadores de seleção (SILVA, 2011).

As técnicas de seleção podem ser combinadas com outras técnicas auxiliares, essas ajudam o GA a melhorar seu desempenho e/ou contornar situações em que a função de avaliação não discrimina satisfatoriamente os indivíduos. Algumas delas são descritas a seguir:

- **Elitismo:** é utilizado para garantir que o GA melhore seu desempenho no decorrer das gerações, permitindo a preservação do cromossomo dos indivíduos mais aptos. O funcionamento básico dessa técnica é garantir que os  $N$  melhores indivíduos de uma geração não devem ser descartados e possam ter seus genomas preservados para a próxima geração (LINDEN, 2008).
- **Normalização:** é uma técnica que resolve o problema do super indivíduo e o de aglomeração de indivíduos. Para implementar essa técnica deve-se ordenar a população pela avaliação de aptidão, sendo uma ordem decrescente, onde o primeiro é o melhor indivíduo da população e o último o pior. Feito a ordenação o melhor indivíduo recebe um valor

fixo determinado por  $k$ , sucessivamente os indivíduos abaixo recebem os valores iguais ao do indivíduo anterior na lista subtraindo-se uma constante  $t$ .

Usando essa técnica são criados mais dois parâmetros para se otimizar,  $k$  e  $t$  conforme equação 13, sendo esses parâmetros críticos para o desempenho do algoritmo. Se o valor de  $t$  for muito pequeno podemos ter uma aglomeração de indivíduos e um valor muito grande pode criar desigualdades artificiais, tornando distantes indivíduos que originalmente eram próximos (LINDEN, 2008).

$$\begin{aligned}aval_0 &= k \\aval_i &= aval_{i-1} - t\end{aligned}\tag{13}$$

- **Escalonamento Sigma:** é uma técnica que tem por objetivo tornar o GA menos suscetível à convergência prematura. Seu princípio é modificar a função de avaliação de um indivíduo,  $f(i)$ , por uma fórmula dependente tanto do indivíduo quanto da população no instante  $t$ , demonstrada por meio da equação 14.

$$E(i,t) = \begin{cases} 1, & \sigma(t) = 0 \\ 1 + \frac{f(i) - \bar{f}(t)}{2\sigma(t)}, & \sigma(t) \neq 0 \end{cases}\tag{14}$$

onde  $f(i)$  é a avaliação do indivíduo,  $\bar{f}(t)$  é a avaliação média no instante  $t$  e  $\sigma(t)$  é o desvio padrão das avaliações no instante  $t$ .

Se o desvio padrão é igual a zero, então todos os indivíduos têm avaliações iguais, o que implica em que devem receber a mesma chance de ser sorteados para se submeter a um operador genético. Se a função se torna negativa para algum indivíduo (caso de indivíduos cuja avaliação está mais de dois desvios-padrão abaixo da média da população no instante  $t$  podemos atribuir-lhe um valor arbitrário baixo (por exemplo, 0.1), para que eles tenham uma chance, mesmo que pequena, de ser selecionados.

#### 2.5.1.4 OPERADORES DE REPRODUÇÃO

Na biologia a reprodução pode ocorrer de duas formas: a) Assexuada, ou seja, um indivíduo sozinho gera um descendente utilizando o processo de “mutação”; b) Sexuada, isto é, a geração de um ou mais descendentes é realizada utilizando o processo de “recombinação”.

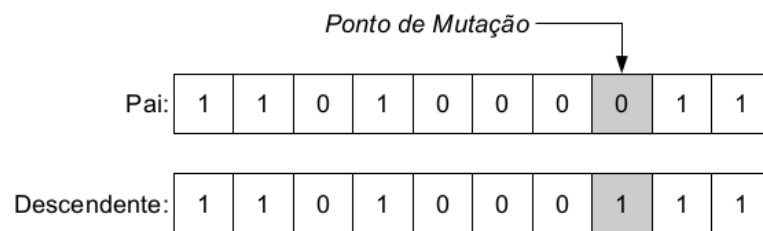
Em um GA típico, uma vez selecionados os pais um ou mais pontos de corte são definidos e então a troca de genes é realizada entre os pais gerando filhos, imitando o processo de recombinação genético encontrado na biologia (SILVA, 2011). A Figura 14 ilustra um exemplo

com 2 pontos de corte, onde o *Filho 1* recebe a partição central do *Pai 2* e as partições à esquerda e à direita dos corte do *Pai 1*, o *Filho 2* recebe a partição central do *Pai 1* e as partições à esquerda e à direita dos corte do *Pai 2*.



**Figura 14: Recombinação usando 2 pontos de corte.**

Outro processo que geralmente é combinado com a troca de genes é a “mutação”, onde um ou mais genes são definidos aleatoriamente para troca do seu valor, inspirado no processo biológico de “mutação”, injetando assim uma mutação genética no indivíduo, assim como ilustrado pela Figura 15. Esse processo é muito importante para se evitar os máximos locais, por intermédio da troca de genes a evolução pode mudar a direção através do espaço de busca.



**Figura 15: Ilustração do operador de mutação.**

## 2.5.2 ENXAME DE PARTÍCULAS

Uma outra abordagem evolutiva diz respeito ao algoritmo enxame de partículas, também conhecido como *Particle Swarm Optimization - PSO*, teve origem na observação da migração de pássaros em busca de alimentos. Cada partícula representa um ponto no espaço de soluções (XU et al., 2009). O conceito básico é o de criar um conjunto (enxame) de partículas que se movem dentro de um determinado espaço de problema em busca do lugar que melhor se adapte



às suas necessidades, nesse caso, dada pela função de aptidão (*fitness*). A especificação desta função depende do problema a ser otimizado, especialmente em suas dimensões.

### 2.5.2.1 INICIALIZAÇÃO DO ENXAME

Uma vez que um espaço de problema é definido um conjunto de partículas é gerado e posteriormente suas posições e suas velocidades são atualizadas iterativamente.

---

#### Algoritmo 2.5.1 Inicializa.

---

**Entrada:** Enxame de Partículas  $S$ , Dimensões  $D$

- 1: **para todo** partícula  $i \in S$  **faça**
  - 2:   **para todo** dimensão  $d \in D$  **faça**
  - 3:      $x_{i,d} \leftarrow Rnd(x_{min}, x_{max})$
  - 4:      $v_{i,d} \leftarrow Rnd(-v_{max}/3, v_{max}/3)$
  - 5:   **fim-para**
  - 6:    $pb_i \leftarrow x_i$
  - 7:   **se**  $f(pb_i) > f(g_b)$  **então**
  - 8:      $g_b \leftarrow pb_i$
  - 9:   **fim-se**
  - 10: **fim-para**
- 

No algoritmo 2.5.1, tem-se um conjunto completamente conectado, o que significa que todas as partículas compartilham informações, qualquer partícula conhece qual é a melhor posição já visitada por qualquer outra do enxame. Na linha 3 e 4 as posições e velocidades das partículas são inicializadas, na linha 4 é utilizada a função de aptidão  $f(x)$ , onde  $x$  é partícula que será avaliada, para inicializar a posição de melhor global considerando a população inicial.

### 2.5.2.2 DESLOCAMENTO DAS PARTÍCULAS

Cada partícula tem uma posição, formalmente definida pela Equação 15, e uma velocidade, definida pela Equação 16, que são calculadas como segue:

$$x_{i,d}(it + 1) = x_{i,d}(it) + v_{i,d}(it + 1) \quad (15)$$

$$\begin{aligned} v_{i,d}(it + 1) = & v_{i,d}(it) \\ & + C_1 * Rnd(0, 1) * [pb_{i,d}(it) - x_{i,d}(it)] \\ & + C_2 * Rnd(0, 1) * [g_{b,d}(it) - x_{i,d}(it)] \end{aligned} \quad (16)$$

onde:  $i$  é o índice usado para identificar uma partícula;  $d$  é a dimensão a ser considerada (cada partícula possui uma posição e uma velocidade para cada dimensão);  $it$  é o número da iteração;

$x_{i,d}$  é a dimensão  $d$  da partícula  $i$ ,  $v_{i,d}$  é a velocidade da partícula  $i$  na dimensão  $d$ ;  $C_1$  é a constante de aceleração para o componente cognitivo (em direção ao melhor global);  $Rnd$  define um valor randômico entre 0 e 1;  $pb_{i,d}$  é a dimensão  $d$  da melhor partícula local;  $C_2$  é a constante de aceleração para o componente social (em direção ao melhor local); e  $gb_d$  é dimensão  $d$  da melhor partícula global.

A atualização da velocidade, que ocorre de acordo com a equação 16, resulta da soma de diferentes componentes, cada um com um significado específico. Na primeira linha da equação temos o componente de impulso, que é a velocidade anterior da partícula. Na segunda linha temos o componente cognitivo, que depende fortemente da distância atual da partícula para a melhor posição já visitada. Finalmente na terceira linha temos o componente social, que depende fortemente da distância da partícula para a melhor posição do enxame na iteração atual.

O Algoritmo 2.5.2 explicita os passos realizados pela abordagem PSO. Na linha 1 é executada a inicialização do enxame, conforme mostra o Algoritmo, as linhas 4 e 5 definem a melhor posição local, enquanto as linhas 7 e 8 definem a melhor posição global e por fim nas linhas 13 e 14 são realizadas as atualizações de velocidade e posição, respectivamente.

---

**Algoritmo 2.5.2** Particle Swarm Optimization.

---

**Entrada:** Enxame de Partículas  $S$ , Dimensões  $D$

```

1: Inicializa
2: repita
3:   para todo partícula  $i \in S$  faça
4:     se  $f(x_i) > f(pb_i)$  então
5:        $pb_i \leftarrow x_i$ 
6:     fim-se
7:     se  $f(pb_i) > f(gb)$  então
8:        $gb_i \leftarrow pb_i$ 
9:     fim-se
10:  fim-para
11:  para todo partícula  $i \in S$  faça
12:    para todo dimensão  $d \in D$  faça
13:       $v_{i,d} \leftarrow v_{i,d} + C_1 * Rnd(0, 1) * [pb_{i,d} - x_{i,d}] + C_2 * Rnd(0, 1) * [gb_d - x_{i,d}]$ 
14:       $x_{i,d} \leftarrow x_{i,d} + v_{i,d}$ 
15:    fim-para
16:  fim-para
17: até que  $it < MAX\_ITERACOES$ 

```

---

### 2.5.3 SOCIAL SPIDER ALGORITHM

Na literatura encontramos algoritmos evolutivos baseados em enxame com base em outros organismos e que promovem otimização com um desempenho satisfatório. Em Yu e Li (2015) propõe o *Social Spider Algorithm* - SSA, um algoritmo evolutivo baseado no comportamento das aranhas. No SSA o espaço de busca do problema de otimização é representado por uma teia de aranha multidimensional. Cada posição na teia representa uma solução viável para o problema de otimização e todas as soluções viáveis para o problema têm posições correspondentes nesta teia. A teia também serve como o meio de transmissão das vibrações geradas pelas aranhas. Cada aranha na teia mantém uma posição e a qualidade (ou a forma física) da solução baseia-se na função de aptidão, representada pelo potencial de encontrar uma fonte de alimento na posição. As aranhas podem se mover livremente na teia. No entanto, elas não podem deixar a teia, as posições fora dela representam as soluções inviáveis para o problema de otimização. Quando uma aranha se move para uma nova posição, ele gera uma vibração que se propaga através da teia. Cada vibração contém as informações de uma aranha, sendo assim, outras aranhas podem obter essas informações ao receber a vibração.

#### 2.5.3.1 ARANHA

As aranhas são os agentes do SSA para executar a otimização. No início do algoritmo, um número pré-definido de aranhas são colocadas na teia. Cada aranha  $s$  detém uma memória onde armazenam as seguintes informações individuais:

- A posição de  $s$  na teia.
- O valor da aptidão da posição atual de  $s$ .
- A vibração do alvo  $s$  na iteração anterior.
- O número de iterações desde que  $s$  teve a vibração de seu alvo modificada pela última vez.
- O movimento realizado na iteração anterior.
- A máscara de dimensão que  $s$  utilizou para guiar o movimento na iteração anterior.

Os dois primeiros tipos de informação descrevem a situação individual de cada aranha, enquanto todos os outros tem o objetivo de direcionar a aranha para novas posições.

Com base em observações, as aranhas possuem sentidos muito precisos de vibração. Além disso, elas podem separar diferentes vibrações propagadas na teia no mesmo sentido e as suas respectivas intensidades. No SSA, uma aranha gera uma vibração quando assume uma nova posição diferente da anterior. A intensidade da vibração está correlacionada com a aptidão da posição. A vibração irá se propagar através da teia e outras aranhas poderão sentir. Dessa forma, as aranhas presentes na teia compartilham suas informações individuais com as outras para formar um conhecimento coletivo e social.

### 2.5.3.2 VIBRAÇÃO

Vibração é um conceito muito importante em SSA e uma das principais características que distinguem esse algoritmo de outras Metaheurísticas. No SSA, são usadas duas propriedades para definir uma vibração, isto é, a posição e a intensidade da fonte de vibração. A posição de origem está definida pelo espaço de busca do problema de otimização, e define-se a intensidade de uma vibração num intervalo  $[0, +\infty]$ . Sempre que uma aranha se move para uma nova posição, ela gera uma vibração em sua posição atual. É definida a posição de uma aranha no tempo  $t$  como  $P_a(t)$ , ou simplesmente como um  $P_a$  se o argumento é  $t$ . Posteriormente é usado  $I(P_a, P_b, t)$  para representar a intensidade da vibração detectada por uma aranha na posição  $P_b$  no tempo  $t$  e a fonte de vibração está na posição  $P_a$ . A partir destas notações podemos usar  $I(P_s, P_s, t)$  para representar a intensidade da vibração gerada pela aranha  $s$  na posição de origem. Esta intensidade na posição de origem está correlacionada com a aptidão da sua posição  $f(P_s)$ , definindo-se o valor da intensidade da seguinte maneira:

$$I(P_s, P_s, t) = \log\left(\frac{1}{f(P_s) - C} + 1\right) \quad (17)$$

onde  $C$  é uma constante pequena com confiança de tal forma que todos os valores possíveis de fitness são maiores do que  $C$ . É importante salientar que as equações estão considerando um valor de minimização. A formulação da Equação 17 toma as seguintes questões em consideração:

- Todas as possíveis intensidades vibracionais do problema de otimização são positivos.
- As posições com melhores valores de aptidão, ou seja, valores menores para problemas de minimização, têm intensidades de vibração maiores do que aquelas com valores de aptidão piores.
- Quando uma solução aborda o ótimo global, a intensidade da vibração não iria aumentar

excessivamente, e causar mau funcionamento do sistema de vibração de atenuação.

Como uma forma de energia, a vibração atenua com a distância. Este fenômeno físico é calculado pelo SSA. A distância entre uma aranha  $a$  e  $b$  é definida como  $D(P_a, P_b)$  usando a distância Manhattan para calcular a distância.

$$D(P_a, P_b) = \|P_a - P_b\|_1. \quad (18)$$

O desvio padrão das posições de todas as aranhas em cada dimensão é representado por  $\delta$ . Com estas definições, a atenuação das vibrações conforme distância pode ser calculada como se segue:

$$I(P_a, P_b, t) = I(P_a, P_a, t) \times \exp\left(-\frac{D(P_a, P_b)}{\delta \times r_a}\right) \quad (19)$$

Na Equação 19, foi introduzido um parâmetro controlado pelo usuário  $r_a \in (0, \infty)$ . Este parâmetro controla a taxa de atenuação da intensidade de vibração conforme a distância. Quanto maior for  $r_a$ , mais fraca a atenuação imposta à vibração.

### 2.5.3.3 PADRÃO DE BUSCA

Há três fases em SSA: inicialização, iteração, e final. Essas três fases são executadas sequencialmente. Cada execução do SSA começa com a fase de inicialização, em seguida, executa-se a busca de uma maneira iterativa e, finalmente, o algoritmo é encerrado retornando as soluções encontradas.

Na fase de inicialização, o algoritmo define a função de aptidão (ou objetivo) e o seu espaço de busca. Os valores para os parâmetros utilizados no SSA também são atribuídos. Depois de definidos os parâmetros, o algoritmo prossegue criando uma população inicial de aranhas para otimização. À medida que o número total de aranhas permanece inalterada durante a simulação do SSA, uma memória de tamanho fixo é alocada para armazenar a sua informação. As posições de aranhas são geradas aleatoriamente no espaço de busca e seus valores de aptidão são calculados e armazenados. A vibração inicial de cada aranha na população é fixada a sua posição atual, e a intensidade da vibração é zero. Todos os outros atributos de cada aranha também são inicializadas com zeros. Assim termina a fase de inicialização e o algoritmo inicia a fase de iteração, que realiza a pesquisa com as aranhas artificiais criadas.

Na fase de iteração, um determinado número de iterações são realizadas pelo algo-

ritmo. Em cada iteração, todas as aranhas na teia mudam para uma nova posição e a cada mudança o valor de aptidão é recalculado. Cada iteração pode ser dividida nas seguintes sub-etapas: avaliação da aptidão, geração de vibração, mudança de máscara, movimento aleatório e tratamento de restrições.

Primeiro o algoritmo calcula os valores de aptidão de todas as aranhas artificiais de acordo com suas posições na teia e atualiza o valor ótimo global. Os valores de aptidão são avaliados uma vez para cada aranha durante cada iteração. Em seguida, estas aranhas geram vibrações em suas posições usando a Equação 17. Depois de todas as vibrações serem produzidas, o algoritmo simula o processo de propagação destas vibrações usando a Equação 18. Neste processo, cada aranha  $s$  receberá  $|pop|$  diferentes vibrações geradas por outras aranhas, onde  $pop$  é a população de aranhas. A informação recebida destas vibrações incluem a posição da fonte de vibração e a sua intensidade atenuada. É usado  $V$  para representar estas  $|pop|$  vibrações. Após a recepção da vibração  $V$ ,  $s$  irá selecionar a vibração mais forte  $v^{best}_s$  de  $V$  e comparar a sua intensidade com a vibração alvo  $v^{tar}_s$  armazenada em sua memória.  $s$  irá armazenar  $v^{best}_s$  como  $v^{tar}_s$  se a intensidade de  $v^{best}_s$  for maior, e  $c_s$ , ou o número de iterações desde que  $s$  modificou a sua vibração alvo pela última vez, é reiniciado com zero; caso contrário, o  $v^{tar}_s$  original é preservado e  $c_s$  é incrementado em um. Usamos  $P_s^i$  e  $P^{tar}_s$  para representar as posições de origem de  $V$  e  $V^{tar}$ , respectivamente, e  $i = 1, 2, \dots, |pop|$ .

O algoritmo então manipula  $s$  para realizar um movimento em direção  $v^{tar}_s$ . Nesse ponto é utilizada uma máscara de dimensão para guiar o movimento. Cada aranha possui uma máscara de dimensão  $m$ , que é um vetor binário de tamanho  $D$ , onde  $D$  é o número de dimensões do problema de otimização. Inicialmente, todos os valores na máscara são zero. Em cada iteração, a aranha  $s$  tem uma probabilidade de  $1 - p_c^{c_s}$  para mudar a sua máscara, onde  $p_c \in (0, 1)$  é um atributo definido pelo usuário que define a probabilidade de mudança da máscara. Se a máscara for alterada, cada bit do vetor tem uma probabilidade de  $p_m$  de ser atribuído com um, e  $1 - p_m$  de ser atribuído com zero.  $p_m$  é também um parâmetro de controle definido pelo usuário como 0 ou 1. Cada bit de uma máscara é alterado de forma independente e não tem qualquer correlação com a máscara anterior. No caso de todos os bits serem zeros, um valor aleatório da máscara é alterado para um. Do mesmo modo, é atribuído zero a um bit aleatório, se todos os valores são uns.

Depois da máscara de dimensão ser definida, uma nova posição  $P^{fo}_s$  é gerada com base na máscara para  $s$ . O valor da  $i$ -ésima dimensão da posição  $P^{fo}_{s,i}$  é gerado como segue.

$$P^{fo}_{s,i} = \begin{cases} P^{tar}_{s,i} & m_{s,i} = 0 \\ P^{fo}_{s,i} & m_{s,i} = 1, \end{cases} \quad (20)$$

onde  $r$  é um número inteiro randômico entre 1 e  $|pop|$ , e  $m_{s,i}$  representa a  $i$ -ésima dimensão da máscara  $m$  da aranha  $s$ . Aqui o número aleatório  $r$  para duas diferentes dimensões com  $m_{s,i} = 1$  é gerado independentemente.

Com a geração de  $P^f o_s$ ,  $s$  executa um movimento aleatório. Esse movimento é calculado por meio da seguinte Equação.

$$P_s(t+1) = P_s + (P_s - P_s(t-1)) \times r + (P_s^{fo} - P_s) \odot R, \quad (21)$$

Onde  $\odot$  denota multiplicação elemento a elemento e  $R$  é um vetor de números aleatórios de ponto de flutuante gerados a de zero até um uniformemente. Antes de seguir  $P^f o_s$ ,  $s$  primeiro se desloca conforme sua direção anterior, que é a direção do movimento na iteração anterior. Então  $s$  se aproxima de  $P^f o_s$  ao longo de cada dimensão com números aleatórios gerados entre 0 e 1. Este fator aleatório para diferentes dimensões é gerado independentemente. Após este movimento,  $s$  armazena seu movimento na iteração atual para a próxima iteração. Isto finaliza o sub-passo do movimento aleatório.

O sub-passo final da fase de iteração é o tratamento de restrições. As aranhas podem se mover para fora da teia durante a etapa de movimento aleatório, dessa forma violando as restrições do problema de otimização. Para tratar as restrições, no SSA, é adotada a abordagem reflexiva para produzir uma fronteira livre de restrição  $P_s(t+1)$  por

$$P_{s,i}(t+1) = \begin{cases} (\bar{x}_i - P_{s,i}) \times rif P_{s,i}(t+1) > \bar{x}_i (P_{s,i} - \underline{x}_i) \times rif P_{s,i}(t+1) < \underline{x}_i, \end{cases} \quad (22)$$

onde  $\bar{x}_i$  é o limite superior do espaço de busca na  $i$ -ésima dimensão, e  $\underline{x}_i$  é o limite inferior da dimensão correspondente.  $r$  é um número randômico de ponto flutuante no intervalo de 0 a 1.

Na fase de iteração o algoritmo executa o laço até que o critério de parada seja satisfeito. O critério de parada pode ser definido como o número máximo de iterações, máximo tempo de execução ou qualquer outro mais apropriado. Após a fase de iteração, o algoritmo devolve como saída a melhor solução com a melhor aptidão encontrada. As três fases que constituem o algoritmo SSA e o seu pseudo-código pode ser encontrado no Algoritmo 2.5.3.

---

**Algoritmo 2.5.3** Social Spider Algorithm.
 

---

- 1: Define valores para parâmetros do SSA.
  - 2: Cria uma população de aranhas  $pop$  e atribui memória para cada uma.
  - 3: Inicializa  $v_s^{tar}$  para cada aranha.
  - 4: **enquanto** critério de parada não satisfeito **faça**
  - 5:   **para todo** aranha  $s \in pop$  **faça**
  - 6:     Avalia o valor de aptidão de  $s$ .
  - 7:     Gera vibração para a posição de  $s$ .
  - 8:   **fim-para**
  - 9:   **para todo** aranha  $s \in pop$  **faça**
  - 10:     Calcula as intensidades das vibrações  $V$  geradas por todas as aranhas.
  - 11:     Seleciona a vibração mais forte  $v_s^{best}$  em  $V$ .
  - 12:     **se** a intensidade de  $v_s^{best}$  é maior que  $v_s^{tar}$  **então**
  - 13:        $v_s^{best} \leftarrow v_s^{tar}$ .
  - 14:     **fim-se**
  - 15:     Atualiza  $c_s$ .
  - 16:     Gera um número randômico  $r$  entre 0 e 1.
  - 17:     **se**  $r > p_c^{cs}$  **então**
  - 18:       Atualiza a máscara de dimensão  $m_s$ .
  - 19:     **fim-se**
  - 20:     Gera  $p_s^{fo}$ .
  - 21:     Execute movimento aleatório.
  - 22:     Resolve quaisquer restrições violadas.
  - 23:   **fim-para**
  - 24: **fim-enquanto**
  - 25: Saída: melhor solução encontrada.
- 

## 2.6 AVALIAÇÃO DE DESEMPENHO

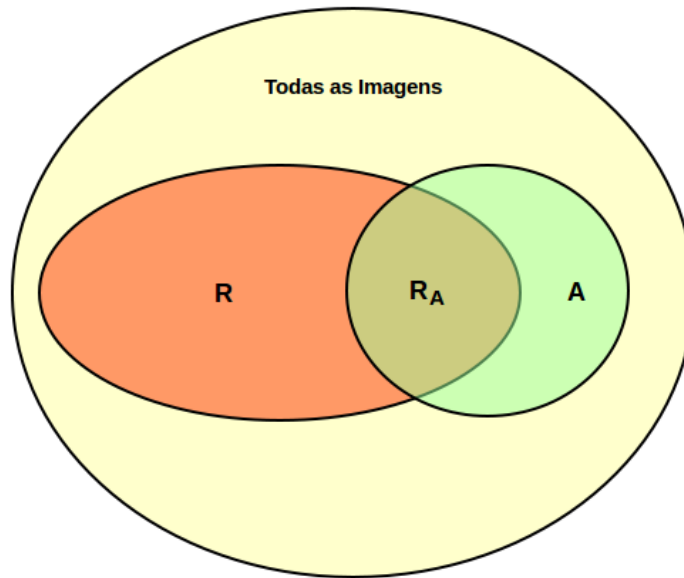
### 2.6.1 PRECISÃO E REVOCAÇÃO

Existem diversas metodologias de avaliação de desempenho recuperação com relação a recuperação de informações. Para avaliar a eficácia de sistemas de recuperação de imagens por conteúdo o método mais utilizado são gráficos de Precisão X Revocação (BAEZA-YATES et al., 2011). A Figura 16 ilustra os principais conjuntos envolvidos em uma consulta utilizados na avaliação. Dado um conjunto  $R$  de imagens relevantes em uma base de dados;  $A$  o conjunto de elementos retornados para uma dada consulta, pode-se observar que intersecção dos conjuntos  $R$  e  $A$  ( $R_A$ ) compreende as imagens relevantes recuperadas. Observa-se que quanto maior for a intersecção entre os conjuntos recuperados e relevantes, maior será a precisão da consulta.

Considerando os conjuntos ilustrados pela Figura 16 define-se a cardinalidade de  $R$ ,  $A$  e  $R \cap A$  denotada como  $|R|$ ,  $|A|$  e  $|R_A|$  respectivamente. A partir desses conjuntos a precisão, também conhecida como *Precision*, é determinada pela fração de imagens recuperadas consideradas relevantes, sendo formalmente definida como:

$$precision = \frac{|R_A|}{|A|} \quad (23)$$





**Figura 16: Conjuntos de dados de uma consulta.**

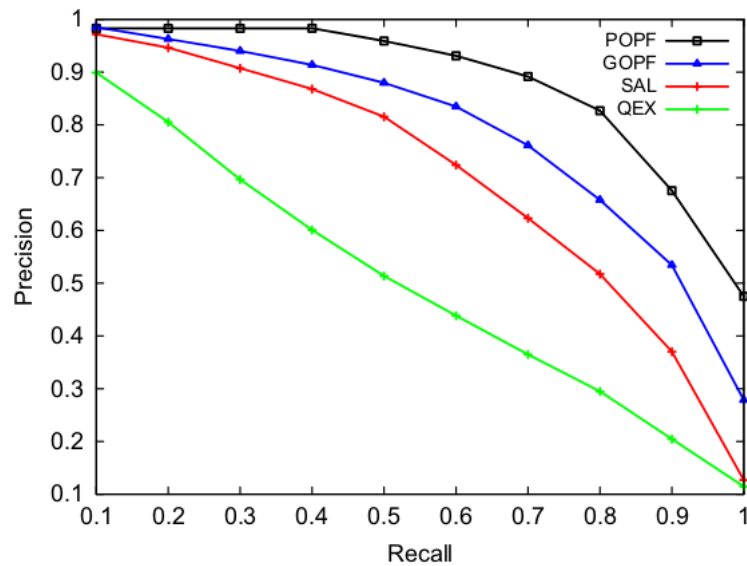
É importante ressaltar que o significado e o uso da precisão no campo da recuperação de imagens é diferente da definição de exatidão e precisão de outros ramos de pesquisa (i.e. matemática estatística). Essa medida leva em consideração todas as imagens recuperadas e em determinados casos, para facilitar a análise, pode ser encontrada na literatura levando em consideração um determinado corte, considerando apenas os resultados mais próximos de um determinado centro de consulta. Nesse caso, essa medida é chamada de *precision* em  $n$ , definida por  $P@n$ , onde  $n$  representa as  $n$  primeiras imagens retornadas para uma dada consulta (ZHANG; WANG, 2012).

A revocação ou *Recall* é a porção do conjunto de imagens relevantes ( $|R|$ ) na base presente no resultado de uma determinada consulta dentre todas as relevantes recuperadas ( $|R_A|$ ). Essa medida pode ser considerada como a probabilidade de uma imagem relevante ser recuperada dada uma consulta. A revocação é formalmente definida por:

$$recall = \frac{|R_A|}{|R|} \quad (24)$$

Para uma avaliação mais abrangente e confiável dos resultados obtidos por um determinado sistema de recuperação de imagens por conteúdo, deve-se estabelecer um conjunto de dados apropriado para avaliação e assim realizar diversas operações de consultas, atribuindo a cada uma delas um elemento diferente da base como centro de consulta. Com as medidas de desempenho dessas diversas consultas é possível construir um gráfico da Curva de Precisão X Revocação (*Precision X Recall Curve*), como ilustrado na Figura 17, onde cada ponto do

gráfico representa a média aritmética (média das precisões) dos desempenhos dessas diversas consultas. Como várias consultas geram gráficos com diferentes coordenadas atribuídas, é necessário realizar-se a interpolação em cada gráfico, onde a precisão interpolada em um dado nível de revocação será  $P(r_j) = \max_{r_j \leq r \leq r_{j+1}} P(r)$  (BAEZA-YATES et al., 2011). A partir do gráfico, a avaliação do desempenho geral do sistema se resume em analisar o formato e altura da curva obtida, sendo que, quanto mais próxima do topo do gráfico a curva estiver, melhor o resultado das operações de busca (SILVA, 2014).



**Figura 17: Exemplo de uma Curva de Precisão e Revocação, extraída de Silva et al. (2011).**

Conforme a taxa de revocação se aproxima de 100% a precisão é deteriorada, dessa forma, conforme mostra a Figura 17, uma avaliação conjunta de precisão e revocação representada por meio do gráfico facilita a interpretação do resultado (BAEZA-YATES et al., 2011).

Num sistema ideal as taxas de precisão e revocação precisam ser altas, porém num sistema real conforme o número de imagens recuperadas aumentam, a precisão diminui, enquanto a revocação aumenta. Por isso ao invés de utilizar simplesmente o valor de precisão ou revocação a curva contendo as duas medidas é frequentemente adotada para avaliar a performance de um sistema CBIR (DOULAMIS; DOULAMIS, 2006).

Além de tais medidas, existem outras pertinentes para a avaliação de sistemas de recuperação de informação. A *F-measure*, definida formalmente pela Equação 25, é uma dessas medidas, a qual trata-se da média harmônica uniformemente ponderada da precisão em níveis de revocação, normalmente dezenas (COSTA et al., 2012).

$$F\text{-Measure} = \frac{2 \times (\textit{Precision} \times \textit{Recall})}{(\textit{Precision} + \textit{Recall})} \quad (25)$$

Essa medida combina as métricas de precisão e revocação. O resultado dessa medida indica que quanto mais próximo de 1 melhor é o resultado, quanto mais próximo de 0 demonstra que os resultados são ruins, ou seja, longe do proposto como ideal.

Outra medida importante na avaliação de sistemas de recuperação de imagens é o valor das precisões médias, ou *Mean Average Precision* (MAP), que é obtida por meio do cálculo do valor médio das médias de precisões de um conjunto consultas (SANTOS et al., 2009), sendo representada formalmente por:

$$MAP = \frac{\sum_{q=1}^Q AveP(q)}{Q} \quad (26)$$

onde  $Q$  representa o número total de consultas realizadas e  $AveP(q)$  a função que calcula a precisão média da consulta  $q$ . De maneira geral essa medida, dentre outras finalidades, serve para calcular a área abaixo da curva de precisão e revocação.

## 2.7 CONSIDERAÇÕES FINAIS

Foram apresentados nesse capítulo os conceitos e metodologias relacionados às etapas de recuperação de imagens baseada em conteúdo (CBIR). Foram apresentadas técnicas de extração de características comumente usadas para representar uma imagem. Descrita a busca por similaridade, que representa uma técnica essencial para a implementação de um sistema CBIR; técnicas de realimentação de relevância e perfil de usuário, que tem por objetivo melhorar as respostas às consultas por similaridade atacando diretamente o problema da lacuna semântica; algoritmos evolutivos, que vem se apresentando como soluções viáveis para problemas de otimização em sistemas CBIR; e por último, medidas de avaliação para recuperação de imagens.

No próximo capítulo será abordada a proposta de trabalho, mostrando em detalhes como pretende-se chegar aos objetivos ao final do trabalho de pesquisa.

### 3 ABORDAGEM PROPOSTA

Conforme exposto ao longo desse trabalho, existem ainda muitos desafios a serem vencidos dentro do campo de pesquisa em recuperação de imagens. Diversos fatores contribuem para a complexidade dessa área, por exemplo a existência de muitos descritores disponíveis na literatura, todos possuindo pontos fortes e fracos. Nesse aspecto, a escolha desses descritores afeta criticamente o desempenho de um sistema CBIR. Diferentes descritores fornecem informações distintas mas complementares sobre a similaridade entre imagens, isso ocorre porque podem representar diferentes tipos de conteúdo (i.e., cor, textura ou forma). Certamente alguns descritores podem ser mais efetivos para alguns tipos de imagem e menos para outros. Não existe um descritor perfeito que possa ser considerado superior sobre todos os outros ou o melhor para todos os casos (FARIA et al., 2010). Outro ponto crítico é a lacuna semântica, em que técnicas de realimentação de relevância atuam para diminuir a distância entre a intenção do usuário e a resposta de um sistema CBIR.

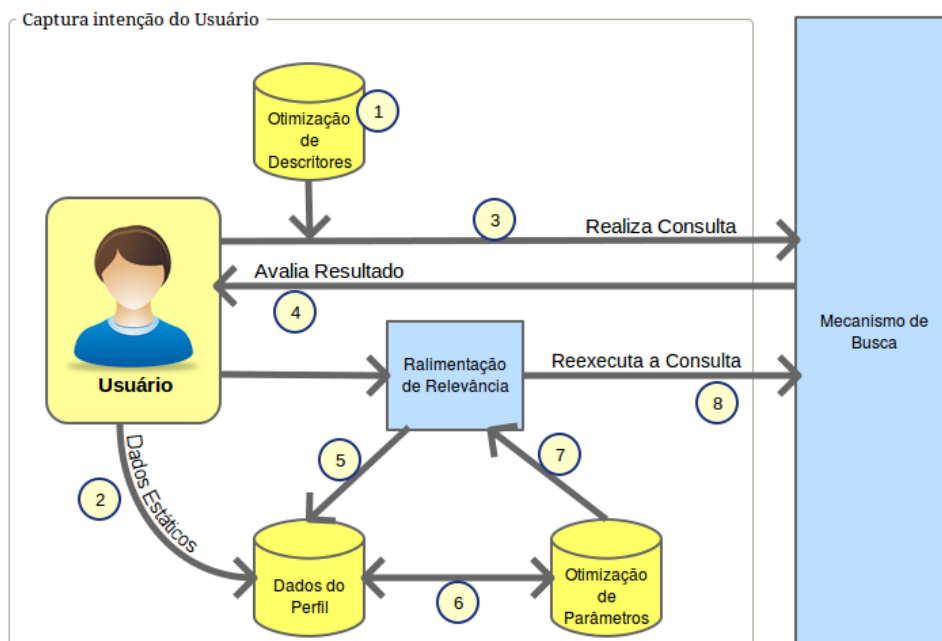
Neste contexto, o trabalho desenvolvido para a presente dissertação proporciona a redução da lacuna semântica em sistemas CBIR, aplicando algoritmos evolutivos juntamente com técnicas de realimentação de relevância e seleção de características. Após um processo minucioso de análise e experimentos chegou-se a proposta de um novo arcabouço que combina as técnicas citadas possibilitando uma maior eficácia e eficiência em todo o processo, abrangendo diversos contextos e áreas do conhecimento da recuperação de imagens.

#### 3.1 RECUPERAÇÃO DE IMAGENS COM ALGORITMOS EVOLUTIVOS

Para se alcançar os resultados/objetivos desse trabalho foi proposto um novo arcabouço para o funcionamento de um sistema CBIR. A ideia principal é a inserção de algoritmos evolutivos nos processos de recuperação de imagens em conjunto com técnicas de realimentação de relevância para a melhoria de desempenho e definição dos descritores mais adequados conforme a intenção do usuário.

Para a captura de tais informações a Figura 18 ilustra o esquema proposto. Inicial-

mente, de modo *off-line*, por meio do algoritmo evolutivo GA, foi possível ao sistema encontrar o melhor descritor (passo 1) para uma determinada base de imagens. Posteriormente, antes de qualquer interação, o usuário foi cadastrado no sistema (passo 2), fornecendo dados para seu perfil estático (veja Seção 2.4.1 do Capítulo 2). Devidamente cadastrado, o usuário realiza consultas (passo 3) no sistema. A cada interação, o usuário avalia os resultados (passo 4) e os dados são salvos (passo 5) em seu perfil dinâmico. Posteriormente, esses dados são processados por um algoritmo evolutivo (passo 6), encarregado de fornecer ao sistema (passo 7) parâmetros otimizados, de acordo com a percepção do usuário, para a realização de uma nova consulta (passo 8), finalizando um ciclo de interação.



**Figura 18: Esquema proposto para a atualização de perfil do usuário.**

Na otimização de descritores, ilustrada na Figura 18, o sistema trabalhou com informações globais, tais como, melhor descritor inicial para uma dada base de imagens. Após a definição foi realizado um processo de seleção de atributos, reduzindo a dimensionalidade dos vetores de características.

Para os dados do perfil estático são armazenadas informações sobre o usuário, tais como, especialidade, nível de conhecimento, entre outros. Com essas informações é possível traçar um perfil prévio do usuário e assim utilizar informações já coletadas para ajustar as primeiras consultas de um novo usuário. Esses dados também são úteis para ajudar na categorização dos especialistas.

Quanto aos dados do perfil dinâmico, são armazenadas informações referentes a imagem de consulta, distância utilizada, extratores de características, valores de  $k$ , quais imagens

foram rotuladas como relevantes e irrelevantes, taxas de precisão e revocação de cada ciclo de realimentação, informação de custo computacional e quais técnicas de realimentação foram utilizadas, dentre outras. Vale lembrar que durante os experimentos essas informações foram geradas por meio da PRF, ou seja, simulando a interação com o usuário de forma automatizada.

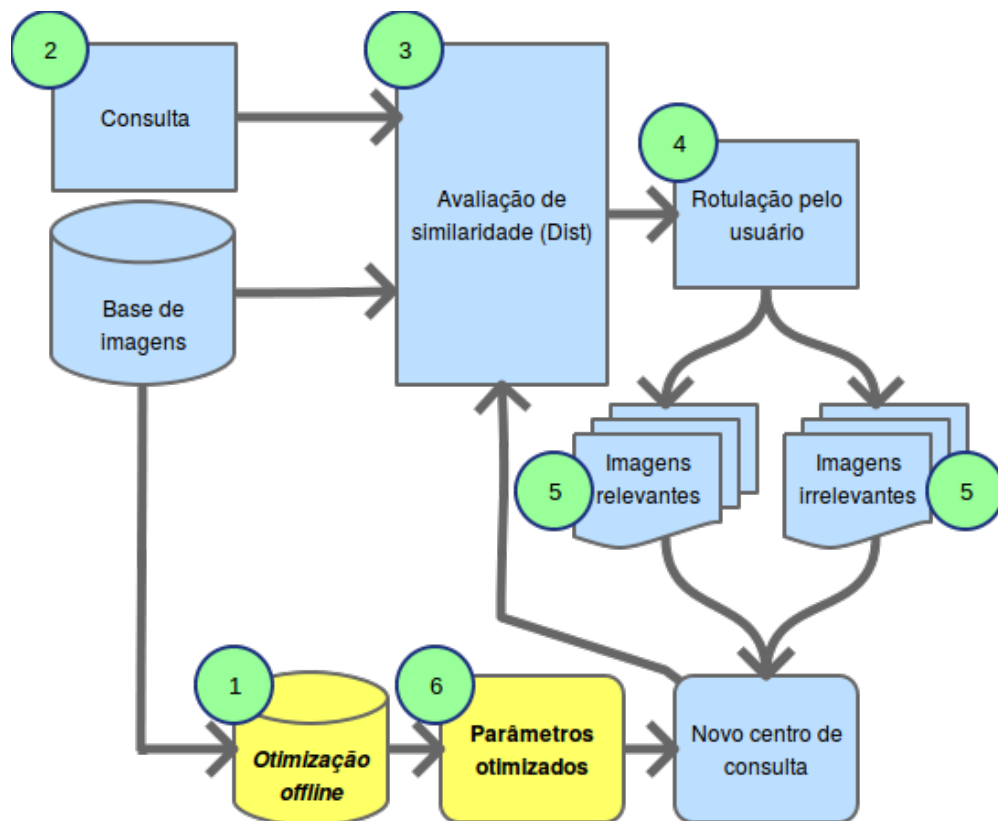
Todas as informações pertinentes às consultas foram armazenadas em perfis de usuário, os quais a partir do acúmulo dos mesmos podem prover formas de personalização das consultas, definindo assim os melhores descritores não apenas para cada usuário em questão, mas também para um grupo de usuários que apresentam comportamentos similares, de acordo com as buscas realizadas. Vale ressaltar que apesar das abordagens evolutivas demandarem um custo de processamento considerável, tais abordagens serão aplicadas de forma *off-line*. Portanto as mesmas podem ser aplicadas no passo anterior à definição da consulta inicial, bem como nos passos intermediários, por meio das informações persistidas nos perfis de usuário, mitigando tal desvantagem.

### 3.1.1 OTIMIZAÇÃO DO PROCESSO DE REALIMENTAÇÃO DE RELEVÂNCIA

Na presente seção é descrita detalhadamente a abordagem proposta, a qual permite otimizar o processo de realimentação de relevância por meio de abordagens evolutivas. O arcabouço proposto apresenta uma grande flexibilidade para incorporar não apenas novos tipos de distâncias, mas também novos métodos de extração de características, abordagens de realimentação de relevância e algoritmos de otimização.

A principal ideia por trás desse arcabouço consiste em usar a informação de intenção do usuário em relação a busca de similaridade para otimizar o processo de realimentação de relevância de acordo com a mesma. A maioria dos trabalhos não leva em consideração os parâmetros iniciais (e.g.  $\alpha$ ,  $\beta$  and  $\gamma$  da fórmula de Rocchio) no processo realimentação de relevância, de acordo com uma dada imagem em um contexto de intenção do usuário, e mantém seus valores iniciais. Esse fato não só leva a uma considerável perda de precisão em consultas de similaridade, mas também degrada o processo de realimentação por relevância em um ritmo muito rápido.

Para mitigar esse problema, o arcabouço implementado permite aplicar estratégias de otimização para escapar de falsos subespaços, usando a intenção do usuário. Dessa forma, ao fazê-lo, os valores das variáveis  $\alpha$ ,  $\beta$  and  $\gamma$  da fórmula de Rocchio, de acordo com o contexto da imagem (e.g. tipos de base de imagens) e o processo de rotulagem realizado pelo usuário durante o processo de realimentação de relevância.



**Figura 19: Otimização de parâmetros offline.**

A Figura 19 ilustra uma instância do arcabouço implementado. Inicialmente, no passo 1, o processo de otimização é executado *offline*, obtendo os melhores parâmetros para o algoritmo de Rocchio de acordo a um dado conjunto de dados. No passo 2, uma imagem de exemplo é submetida a uma consulta de similaridade acoplada ao passo 3. Nos passos 4 e 5, o usuário avalia as imagens recuperadas e indica as relevantes e/ou irrelevantes. Por último, no passo 6, a rotulação realizada pelo usuário é usada como entrada para reposicionar o centro da consulta, usando os parâmetros otimizados. No presente trabalho, foram geradas três instâncias desse arcabouço, aplicando os algoritmos de otimização PSO, GA e SSA para definição dos parâmetros do algoritmo de Rocchio.

Como função de aptidão para os algoritmos evolutivos, foi utilizada a precisão média de um conjunto de consultas por similaridade realizadas para cada iteração de realimentação de relevância. O Algoritmo 3.1.1 descreve como é calculada a aptidão considerando a atual instância do arcabouço proposto. Considerando  $G$  um conjunto de imagens, composto por imagens de diferentes classes, onde cada imagem  $g \in G$  é descrita por um vetor de características (gerado por um dado extrator de característica  $e$ ), o método proposto seleciona as imagens da base (linha 3) considerando a quantidade  $n$  por classe (se  $n = -1$  o algoritmo seleciona todas as imagens da classe), gerando o conjunto *imagens*. Em seguida, os valores das variáveis  $\alpha$ ,  $\beta$  e  $\gamma$  são definidos de acordo com o vetor de parâmetros  $vp$  passado como parâmetro pelo algoritmo

evolutivo.

Dado o conjunto de imagens definido por *imagens*, todas elas são usadas como centro de consulta para uma consulta *k-NN* inicial. Assim, as imagens recuperadas são rotuladas como relevantes ou irrelevantes pelo especialista (i.e., conjuntos de imagens relevantes/irrelevantes). Em seguida, o novo centro de consulta é gerado com base na fórmula de Rocchio, considerando os valores de  $\alpha$ ,  $\beta$  e  $\gamma$ , determinados por *vp* e as imagens rotuladas. Após a geração do novo centro de consulta (*novaConsulta*) é utilizada a função *avaliaConsulta* para calcular a precisão da nova consulta levando em consideração a intenção do usuário. Finalmente, depois de utilizar todas as imagens de consulta, o algoritmo calcula a precisão média de todas as consultas e retorna tal valor como aptidão para o algoritmo evolutivo.

---

#### **Algoritmo 3.1.1** Função de Aptidão

---

**Entrada:** Base de imagens *G*, número de vizinhos *k*, distância *d*, características *e*, imagens por classe *n*, vetor de parâmetros *vp*

**Saída:** Precisão Média *pm*

- 1: *imagens*  $\leftarrow$  obterImagensConsulta(*G*, *n*)
  - 2: *pr*  $\leftarrow$  0 {Inicializa a Precisão}
  - 3:  $\alpha \leftarrow vp_0$
  - 4:  $\beta \leftarrow vp_1$
  - 5:  $\gamma \leftarrow vp_2$
  - 6: **para todo** *img*  $\in$  *imagens* **faça**
  - 7:   *imagensRotuladas*  $\leftarrow$  *k-NN*(*img*, *k*, *d*, *e*)
  - 8:   *novaConsulta*  $\leftarrow$  Rocchio( $\alpha$ ,  $\beta$ ,  $\gamma$ , *imagensRotuladas*)
  - 9:   *pr* + = avaliarConsulta(*novaConsulta*)
  - 10: **fim-para**
  - 11: *pm*  $\leftarrow$  *pr* / count(*imagens*)
- 

### 3.1.2 SELEÇÃO DE ATRIBUTOS

No intuito de melhorar a performance do sistema foi proposto um passo para seleção de atributos. Para executar esse processo foi usada a ferramenta Weka, executando os algoritmos CFS (*Correlation based Feature Selection*)(HALL, 1999) e GA (ver Seção 2.5.1 do Capítulo 2).

O CFS avalia o valor de um subconjunto de atributos ao considerar a capacidade de predição individual de cada atributo, juntamente com o grau de redundância entre eles. Dessa forma, os subconjuntos de características que são altamente correlacionados com a classe e possuem baixa intercorrelação são repassados ao GA para avaliação. Posteriormente, o conjunto de característica selecionado é submetido aos parâmetros otimizados de Rocchio para uma nova avaliação de performance (tempo de execução e precisão). Nesse passo o algoritmo Genético



foi parametrizado da seguinte forma: Taxa de mutação de 0.6, máximo de 20 gerações, taxa de mutação de 0.033 e população de 20 indivíduos.

### 3.2 IMPLEMENTAÇÕES E RECURSOS COMPUTACIONAIS

As implementações para realização dos experimentos foram desenvolvidas usando a linguagem PHP para implementação dos algoritmos evolutivos e da interface com usuário (Figuras 20 a 22), C++ para implementação dos extratores de características e SQL para as funções de distância usadas nas consultas por similaridade em banco de dados (SGBD MySQL).

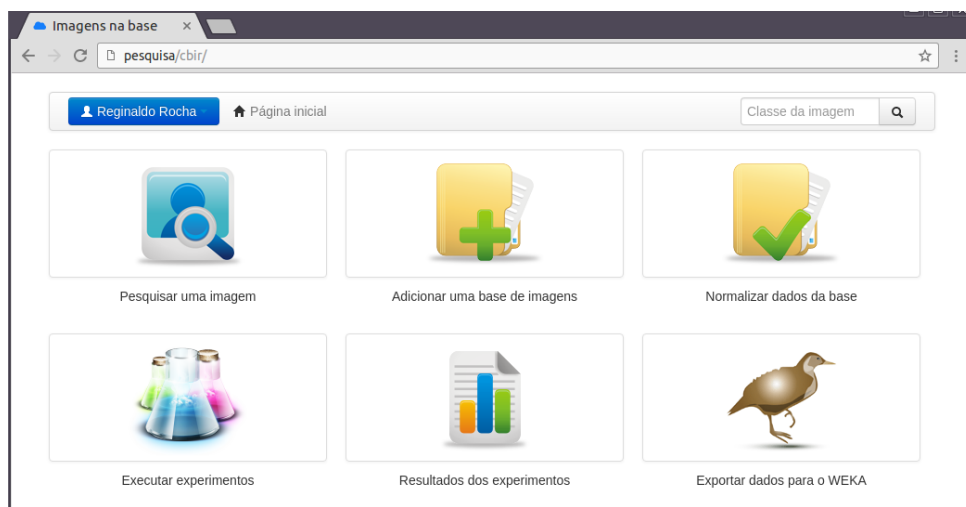


Figura 20: Menu principal da aplicação implementada para executar os experimentos.

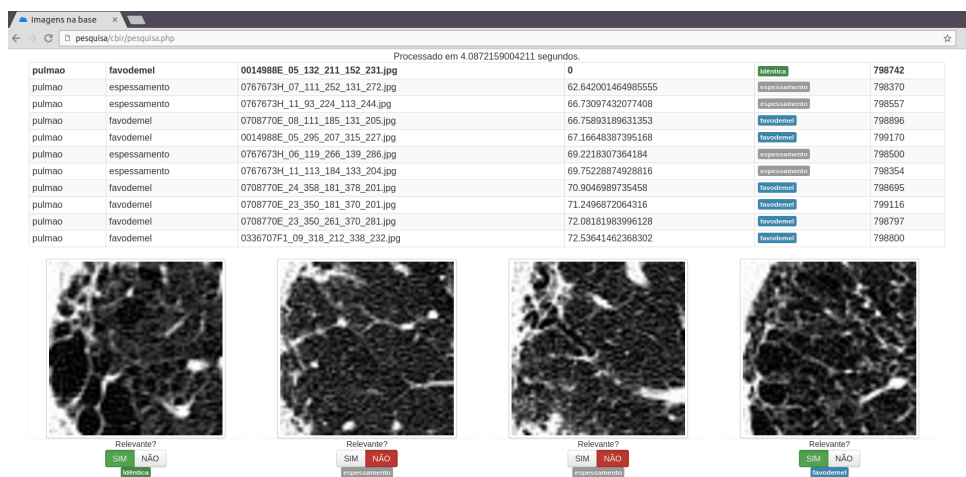
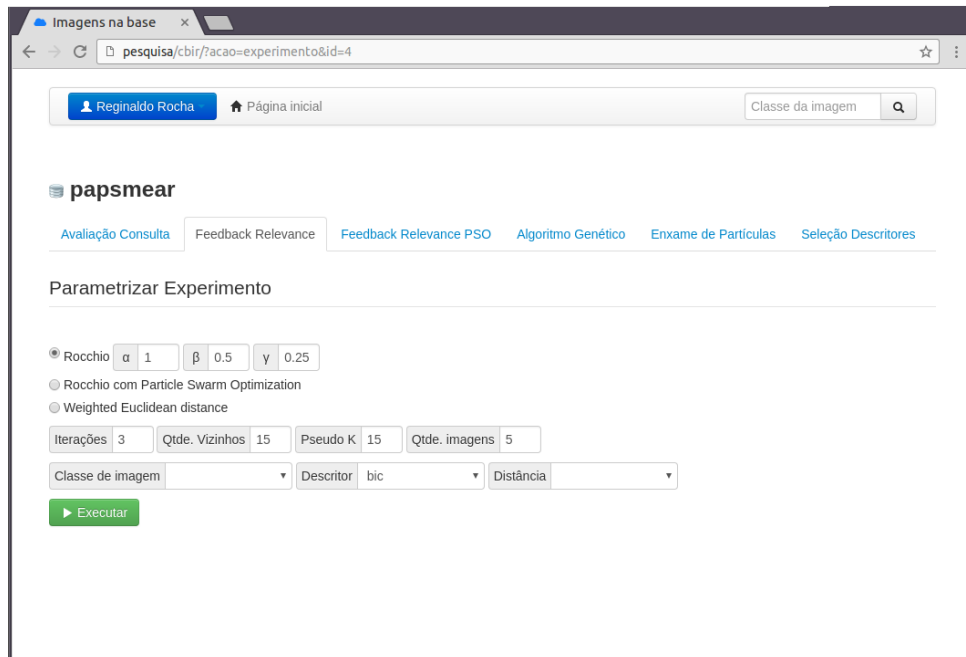


Figura 21: Resultado de uma pesquisa de imagem por conteúdo e realimentação de relevância.

Os experimentos foram executados em uma máquina virtual com 512Mb de memória RAM, Processador de 1 core Intel(R) Xeon(R) de 2.00 GHz, armazenamento SSD de 20Gb. Foi



**Figura 22: Parametrização de experimentos.**

criado um servidor usando Sistema Operacional Linux Ubuntu Server 14.04, Sistema Gerenciador de Banco de Dados MySQL, Servidor web Apache 2 com PHP5 e compiladores GNU C++.

As especificações técnicas citadas também servem como referência mínima de ambiente para hospedar o sistema CBIR implementado para executar os experimentos desse trabalho.

### 3.3 CONSIDERAÇÕES FINAIS

O presente capítulo apresentou a abordagem proposta, a qual aplica o processo de realimentação de relevância em recuperação de imagens acoplado a métodos evolutivos. Tal junção visou a melhor definição de descritores de acordo com a intenção do usuário no processo de recuperação de imagens por conteúdo. Vale ressaltar também que foi agregada à abordagem a possibilidade de seleção de características por meio do auxílio dos métodos evolutivos em junção com a realimentação.

A abordagem proposta proporciona a geração de diferentes instâncias considerando métodos evolutivos, descritores, bem como processos de realimentação de relevância. Além disso, com o intuito de permitir a fusão de diferentes técnicas, foi proposta uma política para definição e codificação de indivíduos baseada na definição de descritores e na otimização dos fatores de ponderação da realimentação de relevância.

## 4 EXPERIMENTOS

Na presente seção, são apresentados os resultados obtidos por meio dos experimentos. Foi avaliada a eficácia dos algoritmos evolutivos, expostos na Seção 2.5 do Capítulo 2, em obter parâmetros ótimos para mover o centro de uma consulta por meio da fórmula de Rocchio.

Os experimentos utilizando a abordagem proposta foram aplicados aos descritores selecionados no passo 1 da Figura 18, combinados com as funções de distância: Canberra, Chebyshev, Euclidiana e Manhattan, descritas na Seção 2.1.2 do Capítulo 2.

Foram executadas consultas *kNN* usando todas as imagens da base como centro de consulta, bem como foi estipulado o valor de  $k = 15$ . Tal valor foi definido por caracterizar como uma boa prática real na análise de imagens realizada por um especialista, visto que um número maior de análises tornaria o processo não fidedigno. Os algoritmos evolutivos foram parametrizados empiricamente da seguinte forma:

- **PSO**: velocidade máxima igual a 0.5; máximo de iterações igual a 20; velocidade na direção do melhor local igual 0.02; velocidade na direção do melhor global igual a 0.02 e quantidade de partículas igual a 10.
- **GA**: quantidade de genes igual 21, ou seja, cromossomo de 21 bits; quantidade de genes envolvidos na mutação igual a 5; população de 40 indivíduos; taxa de conversão igual a 0.9; máximo de gerações estipulado em 100; usa elitismo; não usa normalização; não usa escalonamento; tipo de seleção torneio;
- **SSA**: máximo de iterações igual a 100; taxa de atenuação igual a 1.0; taxa de mudança igual 0.7; máscara igual a 0.1; população de 25 indivíduos;

A realimentação de relevância também foi executada usando valores comumente definidos na literatura (i.e.  $\alpha = 0.75$ ,  $\beta = 0.5$ ,  $\gamma = 0.25$ ) para efeito de comparação com a abordagem otimizada por algoritmos evolutivos.

Além de tais definições foram considerados todos os extratores de características e funções de distância citados nas Seções 2.1.1 e 2.1.2, respectivamente, em conjunto com o processo de realimentação de relevância (Seção 2.4.2 do Capítulo 2). Vale salientar que foram utilizadas diferentes características baseadas em textura, uma vez que as mesmas apresentam uma melhor caracterização quando aplicadas a imagens médicas, diferenciando a orientação e granularidade de diferentes tecidos e estruturas. Já considerando o processo de realimentação foram realizados 10 ciclos, avaliando também dessa forma a saturação apresentada pela abordagem usual e a proposta no presente trabalho.

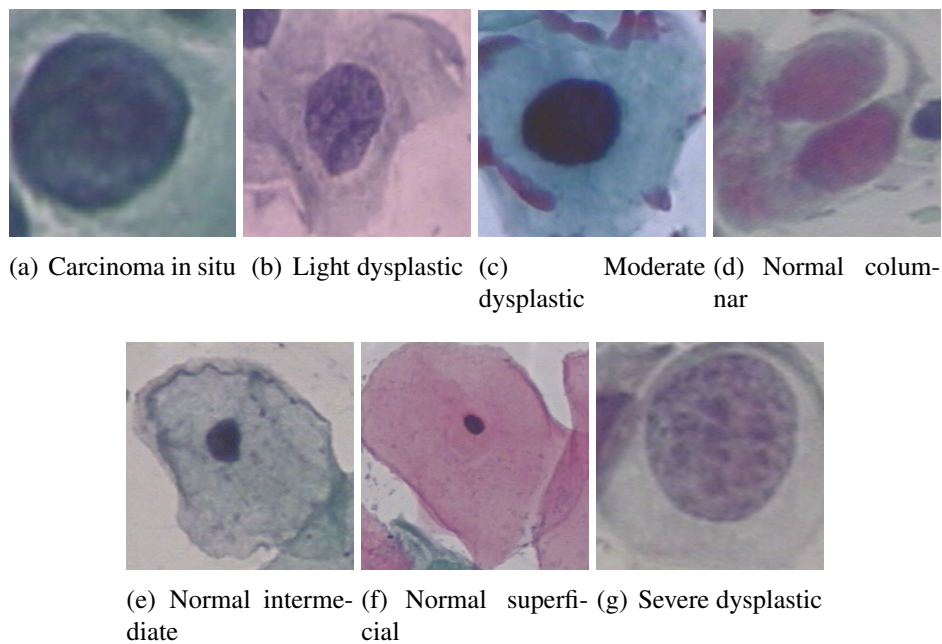
Com o objetivo de realizar uma comparação válida entre o método proposto e a abordagem amplamente utilizada na literatura (definição de pesos de forma empírica), foram realizados testes fixando os melhores valores iniciais do conjunto  $[\alpha, \beta, \gamma]$  como  $[0.5, 0.5, 0]$ ,  $[1, 0.5, 0.25]$ , respectivamente, considerando-se que a realimentação positiva acaba por ser mais valiosa do que a realimentação negativa, e a maioria dos sistemas de recuperação de informações definem  $\beta > \gamma$  (MANNING et al., 2008).

#### 4.1 BASES DE IMAGENS

A escolha de um bom conjunto de imagens é de extrema importância para que possam ser realizados os experimentos e dessa forma avaliar os métodos estudados ao longo do trabalho. Para a definição das bases usadas nesse trabalho foram levados em consideração o foco em imagens médicas, dessa forma foram definidas as bases com imagens esfregaços de lâminas de Papanicolau (Pap-smear), de exames Raios-x de Pulmão e Mamografia. Além do contexto médico, os experimentos também foram executados usando uma base de imagens diversas, nesse caso a Corel 1000, para uma avaliação do arcabouço proposto em outros tipos de imagens abrangendo um maior contexto.

A base Pap-smear (JANTZEN et al., 2005) contém 917 imagens digitalizadas por meio de câmeras digitais acopladas a microscópios do Hospital Universitário de Herlev. As imagens são coloridas e obtidas de esfregaços de lâminas de papanicolau, usadas no exame para diagnóstico do câncer de colo de útero. Conforme exposto na Figura 23, as imagens estão divididas em 7 classes: (a) *Carcinoma in situ* com 150 imagens, (b) *Light dysplastic* com 182 imagens, (c) *Moderate dysplastic* com 146 imagens, (d) *Normal columnar* com 98 imagens, (e) *Normal intermediate* com 70 imagens, (f) *Normal superficial* com 74 imagens e (g) *Severe dysplastic* com 197 imagens.

A Base de imagens de raios-X de pulmão (DIAS et al., 2013) foi obtida por meio

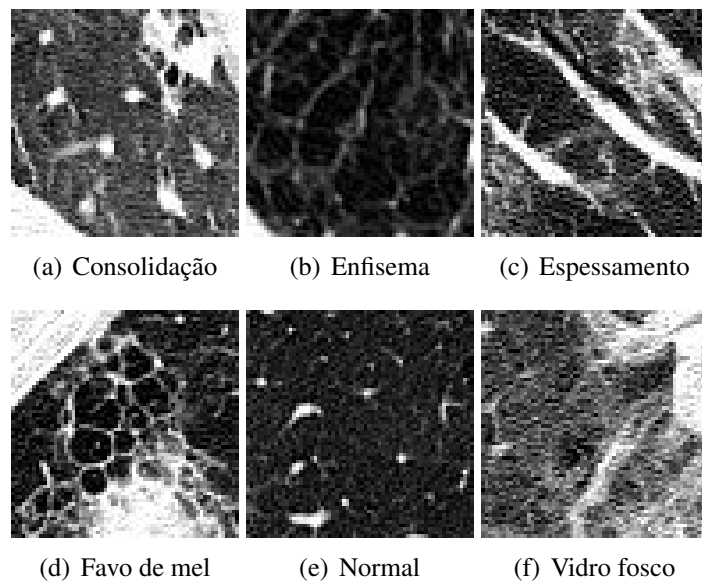


**Figura 23: Amostras de imagens presentes em cada classe da base contendo imagens de exames de Papanicolau.**

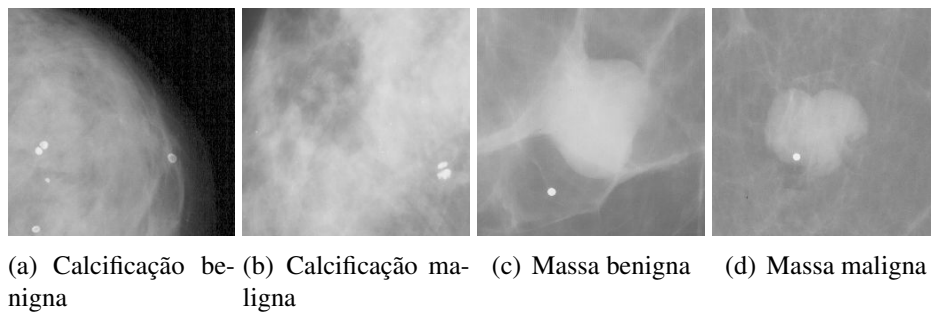
do hospital público de Ribeirão Preto (HCRP) da Universidade de São Paulo. Contém 3.258 imagens em escala de cinza com dimensões de 64 x 64 pixels. As imagens estão divididas em 6 classes distintas, contendo padrões anormais que caracterizam as lesões pulmonares difusas, conforme amostras na Figura 24: (a) Consolidação com 451 imagens, (b) Enfisema com 502 imagens, (c) Espessamento com 590 imagens, (d) Favo de mel com 530 imagens, (e) Normal com 590 imagens e (f) Vidro fosco com 595 imagens. É importante notar que as imagens contendo lesões possuem alterações visuais menos óbvias, aumentando as possibilidades de erros de identificação. Assim, captar a percepção do especialista da melhor forma possível é essencial.

O projeto *Digital Database of Screening Mammography* - DDSM (HEATH et al., 2000) é um esforço colaborativo envolvendo o Hospital Geral de Massachusetts da Universidade do Sul da Flórida, e *Sandia National Laboratories*. Essa base é um recurso disponível muito usado por pesquisadores que investigam técnicas de análise de imagem de mamografia. As imagens foram capturadas em um contexto médico para auxiliar no diagnóstico do câncer da mama. A base contém um número significativo de casos de câncer e exames normais. A versão usada em nossos experimentos possui 2.892 imagens em tons de cinza divididas em 4 classes conforme apresentados na Figura 25, sendo: (a) Calcificação benigna com 615 imagens, (b) Calcificação maligna com 547 imagens, (c) Massa benigna com 906 imagens e (d) Massa maligna com 824 imagens.

Existem bases de imagens públicas disponibilizadas na literatura dentre as quais pode-

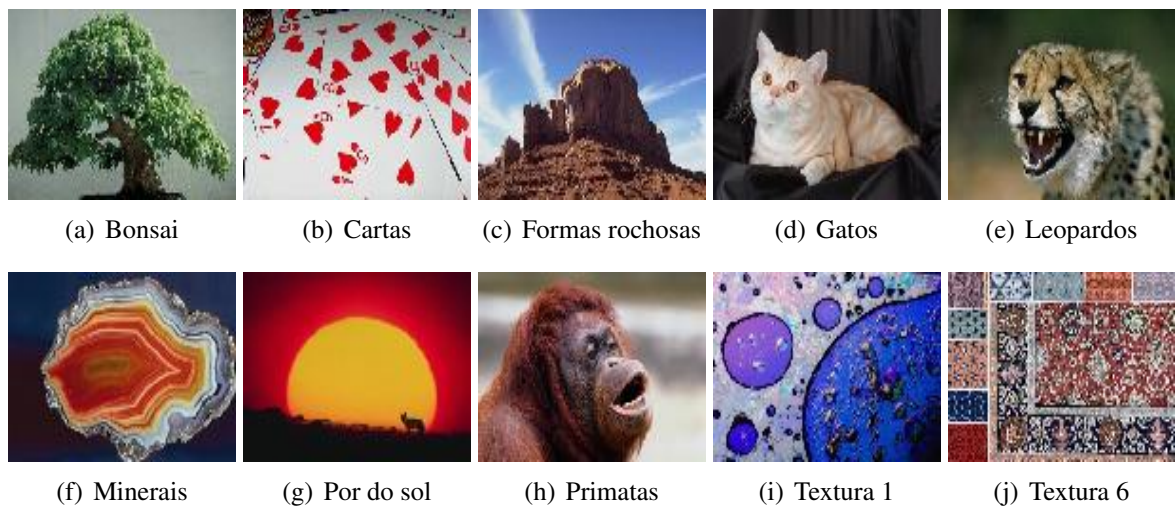


**Figura 24:** Amostras de imagens presentes em cada classe da base contendo imagens de raios-X de pulmão.



**Figura 25:** Amostras de imagens presentes em cada classe da base contendo imagens de mamografia.

se citar a *Corel Dataset* contendo 10.800 imagens coloridas, divididas em grupos conceituais dos quais foram selecionados 10 para uso nos experimentos, conforme amostras da Figura 26: (a) Bonsai, (b) Cartas, (c) Formas rochosas, (d) Gatos, (e) Leopardos, (f) Minerais, (g) Por do sol, (h) Primatas, (i) Textura 1 e (j) Textura 2. Cada grupo contendo 100 imagens coloridas totalizando 1.000 imagens para os experimentos com essa base.



**Figura 26: Amostras de imagens presentes em cada classe da base Corel.**

#### 4.2 EXTRAÇÃO DE CARACTERÍSTICAS E PREPARAÇÃO DA BASE

Para realização dos experimentos, o primeiro passo foi a extração de características intrínsecas às imagens por meio de algoritmos de processamento de imagens, conforme apresentado na Seção 2.1.1 do Capítulo 2. Alguns algoritmos foram desenvolvidos em C++ por Shamir et al. (2008) e outros implementados por meio de referências encontradas na literatura (Seção 2.1.3). Os algoritmos usados na extração e suas respectivas quantidades de atributos são apresentados na Tabela 1.

**Tabela 1: Extratores utilizados nos experimentos e a respectiva quantidade de atributos**

Características	Qtde.	Características	Qtde.
BIC	64	LBP	256
CGCH	255	LCH	240
Chebyshev Coefficients	32	Multiscale Histograms	24
Chebyshev Coefficients (Fourier)	32	Multiscale Histograms (Chebyshev (Fourier))	24
Chebyshev-Fourier Coefficients	32	Multiscale Histograms (Chebyshev)	24
Chebyshev-Fourier Coefficients (Fourier)	32	Multiscale Histograms (Fourier)	24
Comb Moments	48	Multiscale Histograms (Wavelet (Fourier))	24
Comb Moments (Chebyshev (Fourier))	48	Multiscale Histograms (Wavelet)	24
Comb Moments (Chebyshev)	48	Otsu Object Features	34
Comb Moments (Fourier)	48	Radon Coefficients	12
Comb Moments (Wavelet (Fourier))	48	Radon Coefficients (Chebyshev (Fourier))	12
Comb Moments (Wavelet)	48	Radon Coefficients (Chebyshev)	12
Edge Features	28	Radon Coefficients (Fourier)	12
Gabor Textures	7	Tamura Textures	6
GCH	255	Tamura Textures (Chebyshev (Fourier))	6
Haralick Textures	28	Tamura Textures (Chebyshev)	6
Haralick Textures (Chebyshev (Fourier))	28	Tamura Textures (Fourier)	6
Haralick Textures (Chebyshev)	28	Tamura Textures (Wavelet (Fourier))	6
Haralick Textures (Fourier)	28	Tamura Textures (Wavelet)	6
Haralick Textures (Wavelet (Fourier))	28	Zernike Coefficients	72
Haralick Textures (Wavelet)	28	Zernike Coefficients (Fourier)	72
Inverse-Otsu Object Features	34		

Após a extração de características, uma base de dados relacional foi projetada para que pudesse comportar os vetores de características extraídos das imagens. No banco de dados foram implementadas as funções de distância, apresentadas na Seção 2.1.2 do Capítulo 2, permitindo assim a realização de consultas de similaridade por meio de algoritmos desenvolvidos em linguagem tradicional de banco (SQL).

Considerando que atributos de imagens possuem diferentes escalas, que podem interferir no cálculo de distância, para cada conjunto de características foi realizada a normalização dos dados, que trata-se de um pré-processamento das características com objetivo de normalizar os atributos, conforme demonstrado na Equação 27, onde  $V_n$  representa o elemento do vetor,  $v$  é o valor original do elemento,  $max(n)$  e  $min(n)$  os valores máximo e mínimo, respectivamente.

$$V_n = \frac{v - min(n)}{max(n) - min(n)} \quad (27)$$

#### 4.2.1 ESCOLHA DOS MELHORES DESCRITORES

Para selecionar os melhores dentre os descritores, de acordo com o passo 1 ilustrado na Figura 18, foi usado o GA com uma função de aptidão baseada na **F-Measure**, uma medida calculada por meio das medidas de precisão e revocação, conforme descrito na Seção 2.6 do Capítulo 2.

Durante o desenvolvimento da função de aptidão o primeiro passo foi determinar a quantidade e o significado dos genes que cada indivíduo possuiria. Para atender a essa necessidade foi definido que os indivíduos teriam 8 genes, ou seja, 8 bits. Conforme especificado na Tabela 2, os 6 primeiros bits de um indivíduo definem o extrator de característica, e os 2 últimos conforme apresentados na Tabela 3 definem a função de distância do descritor. Para valores fora dos especificados na Tabela 2 a função de aptidão retorna o valor 0, fazendo com que indivíduos com valores inválidos sejam descartados durante o processo de evolução entre as gerações. É importante salientar que nesse passo buscou-se a maximização dos valores de aptidão.



**Tabela 2: Os seis primeiros genes do cromossomo definidos para o GA encontrar os melhores descritores**

Genes	Características	Genes	Características
000001	BIC	011000	LBP
000010	CGCH	011001	LCH
000011	Chebyshev Coefficients	011010	Multiscale Histograms
000100	Chebyshev Coefficients (Fourier)	011011	Multiscale Histograms (Chebyshev (Fourier))
000101	Chebyshev-Fourier Coefficients	011100	Multiscale Histograms (Chebyshev)
000110	Chebyshev-Fourier Coefficients (Fourier)	011101	Multiscale Histograms (Fourier)
000111	Comb Moments	011110	Multiscale Histograms (Wavelet (Fourier))
001000	Comb Moments (Chebyshev (Fourier))	011111	Multiscale Histograms (Wavelet)
001001	Comb Moments (Chebyshev)	100001	Otsu Object Features
001010	Comb Moments (Fourier)	100010	Radon Coefficients
001011	Comb Moments (Wavelet (Fourier))	100011	Radon Coefficients (Chebyshev (Fourier))
001100	Comb Moments (Wavelet)	100100	Radon Coefficients (Chebyshev)
001101	Edge Features	100101	Radon Coefficients (Fourier)
001110	Gabor Textures	100110	Tamura Textures
001111	GCH	100111	Tamura Textures (Chebyshev (Fourier))
010000	Haralick Textures	101000	Tamura Textures (Chebyshev)
010001	Haralick Textures (Chebyshev (Fourier))	101001	Tamura Textures (Fourier)
010010	Haralick Textures (Chebyshev)	101010	Tamura Textures (Wavelet (Fourier))
010011	Haralick Textures (Fourier)	101011	Tamura Textures (Wavelet)
010100	Haralick Textures (Wavelet (Fourier))	101100	Zernike Coefficients
010101	Haralick Textures (Wavelet)	101101	Zernike Coefficients (Fourier)
010111	Inverse-Otsu Object Features		

**Tabela 3: Os dois últimos genes do cromossomo definidos para o GA encontrar os melhores descritores**

Genes	Função de distância
00	Canberra
01	Chebyshev
10	Euclidiana
11	Manhattan

As combinações de características e função de distância que obtiveram melhor desempenho, para cada uma das bases de imagens, estão listadas na Tabela 4.

**Tabela 4: Descritores com melhor desempenho em cada base de imagem**

Base de imagens	Características	Distância	Precisão
<i>pap-smear</i>	LCH	Canberra	56,63%
<i>pap-smear</i>	Haralick Textures (Wavelet)	Canberra	54,08%
Pulmão	Chebyshev Coefficients (Fourier)	Manhattan	66,52%
Pulmão	Comb Moments (Fourier)	Euclidiana	53,64%
Mamografia	Haralick Textures (Wavelet)	Euclidiana	46,54%
Mamografia	LBP	Euclidiana	44,57%
<i>Corel</i>	Zernike Coefficients	Euclidiana	55,78%
<i>Corel</i>	Comb Moments (Chebyshev)	Euclidiana	55,10%

### 4.3 RESULTADOS

Durante a execução dos experimentos, juntamente com os algoritmos evolutivos, para testar o fator randômico, um algoritmo nomeado de *Random* foi implementado para gerar uma população aleatória, medir a aptidão de cada um e retornar o melhor resultado. Assim, o fator “sorte” foi testado e comprovadamente retirado do contexto dos algoritmos evolutivos utilizados. Na Tabela 5 são apresentadas a média de tempo gasto na execução do algoritmo (descontando o tempo da função de aptidão) e a precisão média obtida por meio dos parâmetros otimizados.

**Tabela 5: Média de precisão e tempo de execução**

Algoritmo	Precisão (%)	Tempo (ms)
PSO	0,889	0,0151
Genético	0,848	0,0703
SSA	0,667	1,9664
Random	0,424	0,0005

Em um primeiro momento os experimentos foram realizados combinando apenas algoritmos evolutivos e realimentação de relevância, sem a seleção de descritores ou redução de características.

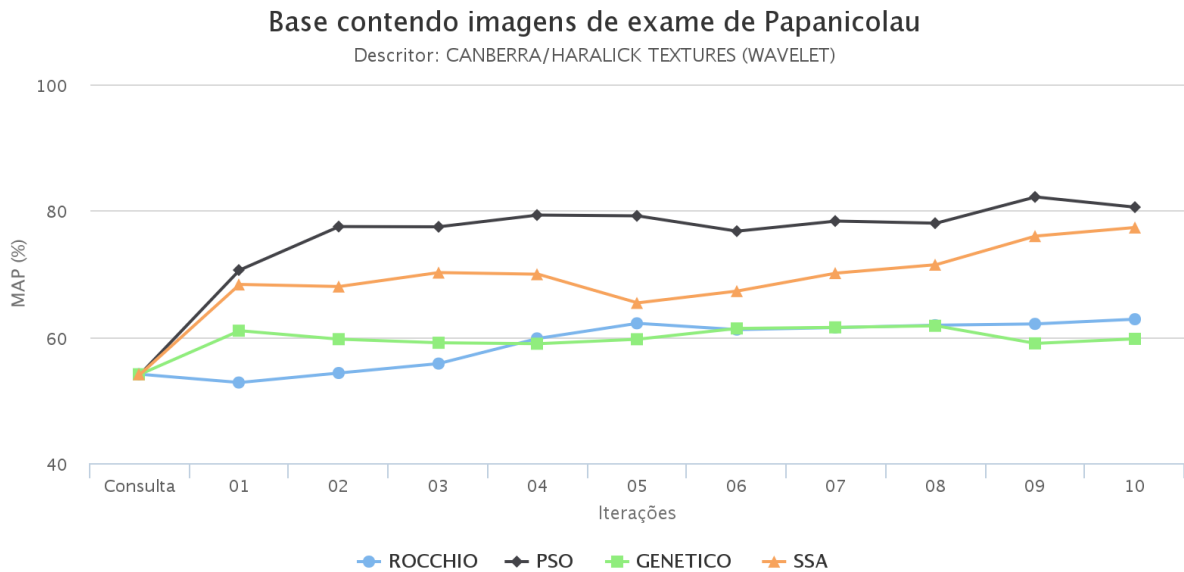
As Figuras 27 a 37 ilustram os resultados obtidos pela abordagem proposta em comparação com a abordagem da literatura (i.e. Rocchio sem otimização de pesos) aplicando as mesmas aos conjuntos de imagens citados na Seção 4.1. O eixo das ordenadas representa os valores de MAP obtidos pelas abordagens. Já o eixo das abscissas representa as iterações do

processo de realimentação (consulta inicial e iterações de 1 a 10). Dessa forma, os gráficos representam os valores de MAP para cada ciclo de realimentação considerando um dado extrator de característica e função de distância.

Nos melhores resultados nota-se que na maioria dos casos o ganho de precisão permanece crescente até a décima iteração, levando em consideração a realimentação baseada em parâmetros otimizados. O mesmo raramente ocorre quando são usados os valores normalmente citados na literatura, como demonstra os gráficos comparando os resultados obtidos levando-se em conta tais valores.

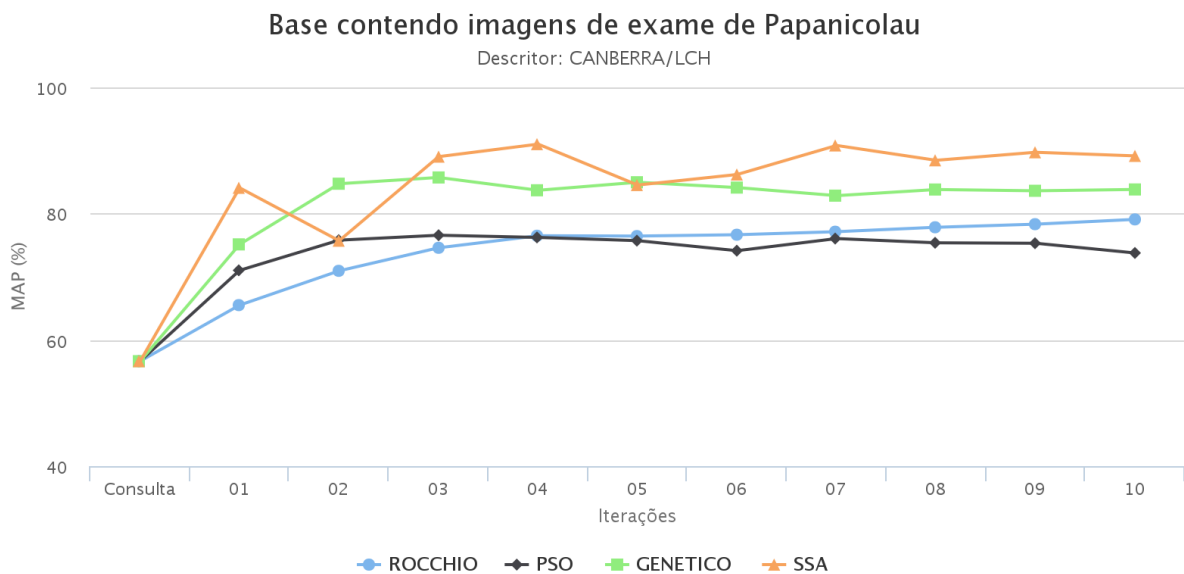
Analisando os resultados da Figura 27, obtidos por meio dos experimentos com a base *pap-smear* percebe-se que a instância da abordagem proposta (i.e. PSO) apresentou ganhos de precisão em todos os ciclos de realimentação quando comparada com a abordagem da literatura (i.e. *Rocchio*). A abordagem proposta, utilizando a distância *Canberra* e o extrator *Haralick Textures (Wavelet)*, atingiu ganhos de precisão de até 28,2% no décimo ciclo. Além disso, nota-se que a instância usando PSO apresentou níveis melhores com relação à saturação das consultas, uma vez que a abordagem da literatura saturou a partir do quinto ciclo de realimentação, ou seja, o ganho de precisão estagnou a partir de tal ciclo e manteve-se em um patamar sem grandes variações nos ciclos posteriores. Dessa forma, apresentando características de aprisionamento da consulta em um mínimo local do espaço. Um comportamento parecido pode ser observado quando analisada a instância usando o SSA, após o quarto ciclo há uma forte queda na precisão, porém há uma recuperação nos ciclos posteriores, chegando ao último com 23,1% de ganho. No caso da instância usando o algoritmo Genético os ganhos se mantiveram apenas nos três primeiros ciclos, mantendo-se estagnado até o último ciclo com uma queda de 4,9% em comparação à abordagem da literatura. Porém, vale ressaltar que em casos de interação com o usuário os primeiros ciclos são os mais importantes, pois a cada novo ciclo a chance do usuário interromper as interações aumentam, seja por satisfazer sua intenção de busca ou por iniciar uma nova busca com novos critérios.

Ainda analisando os experimentos usando imagens da base *pap-smear*, na Figura 28 tem-se os resultados obtidos por meio da função de distância *Canberra* e características LCH, ou seja, um histograma local de cor. Nessa combinação a instância usando SSA se sobressaiu sobre as demais, apresentando um ganho de precisão no décimo ciclo de 12,7% comparado à abordagem com parâmetros da literatura, apesar das quedas de precisão no segundo, quinto e oitavo ciclos. A instância usando o algoritmo Genético apresentou uma performance melhor com o descritor *Canberra/LCH* e também obteve um ganho considerável de precisão com destaque para o terceiro ciclo, onde obteve 13,1% sobre os algoritmos PSO e SSA e 19,7% em



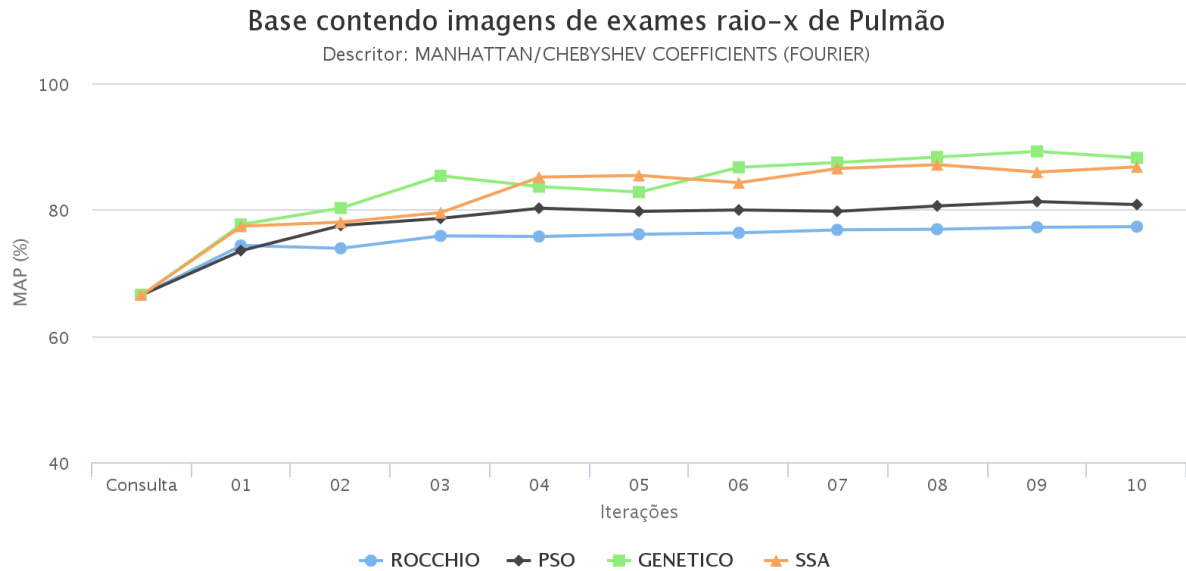
**Figura 27: Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/Haralick Textures (Wavelet).**

comparação ao Rocchio sem otimização. Considerando o último ciclo o Genético obteve 6,3% de ganho. Já a instância usando PSO apresentou uma boa performance apenas nos três primeiros ciclos, estagnando nos demais chegando a perder para a abordagem da literatura.



**Figura 28: Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/LCH.**

O comportamento da abordagem proposta se manteve quando os experimentos foram executados na base Pulmão, conforme exposto na Figura 29, os ganhos de precisão no último ciclo foram de 5,2% para o PSO, 13% para o SSA e 14,3% para o Genético. Para tal análise foram extraídas características por meio das características de *Chebyshev Coefficients (Fourier)* e utilizando a distância Manhattan.

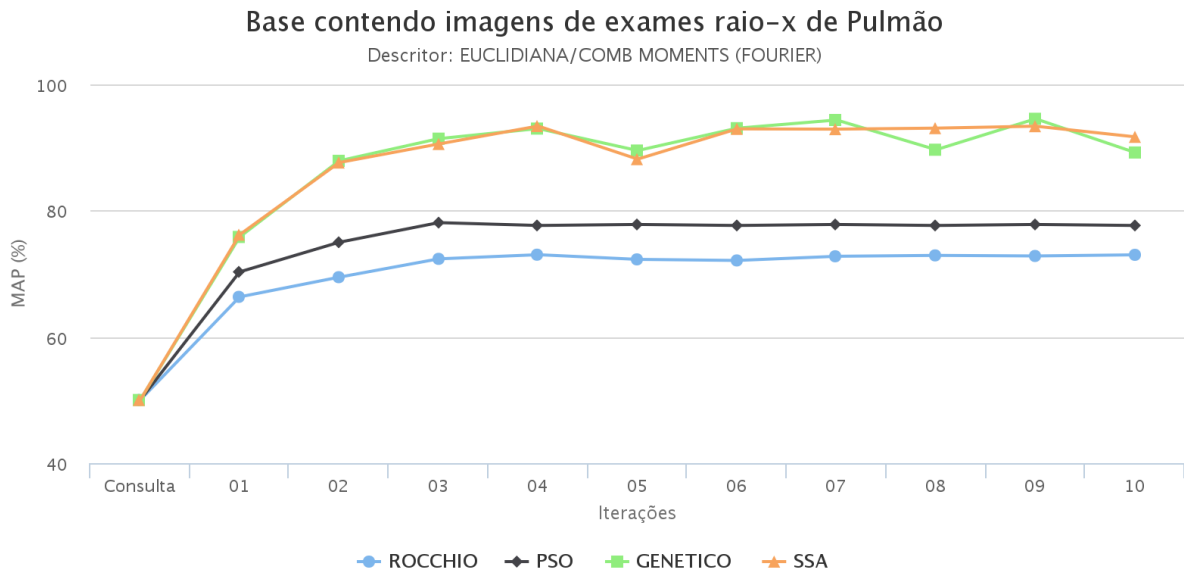


**Figura 29:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Manhattan/*Chebyshev Coefficients (Fourier)*.

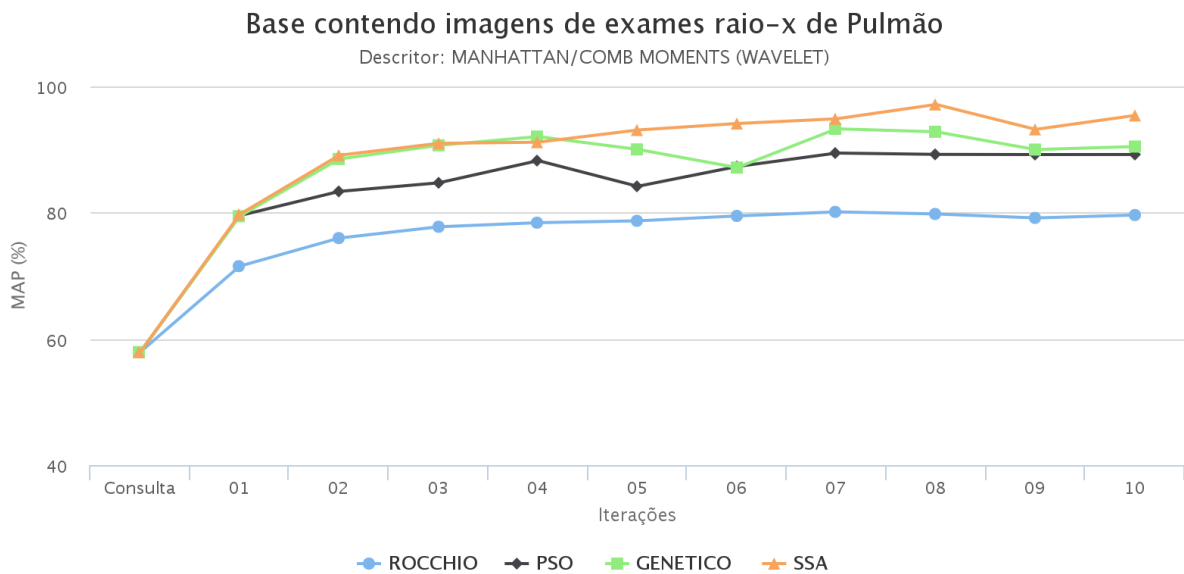
Os experimentos com a característica *Comb Moments (Fourier)* e a função de distância Euclidiana, explicitados na Figura 30, obtiveram o melhor desempenho, destacando o quarto ciclo onde o ganho de precisão do SSA comparado à abordagem da literatura foi de 27,9%, e o ganho do Genético foi de 27,4%. No sétimo ciclo os ganhos chegaram a 27,7% para o SSA e de 29,7% para o Genético. Considerando a instância utilizando o PSO, também obteve-se ganho apesar de menos expressivos. É importante salientar que mesmo com a redução na precisão em alguns ciclos, os ganhos são mantidos altos e permanecem até o último ciclo.

Ainda nos experimentos com imagens da base Pulmão, conforme resultados demonstrados por meio da Figura 31, a instância aplicando o algoritmo PSO apresentou um desempenho melhor para o par *Comb Moments (Wavelet)* e distância Manhattan. Entretanto, o mesmo comportamento onde as instâncias do SSA e Genético se sobressaem foi mantido.

Com relação aos resultados explicitados pela Figura 32, no qual foram utilizadas características oriundas de *Zernike Coefficients* em conjunto com a distância Euclidiana, o comportamento da instância em junção com o PSO acompanha o comportamento dos algoritmos



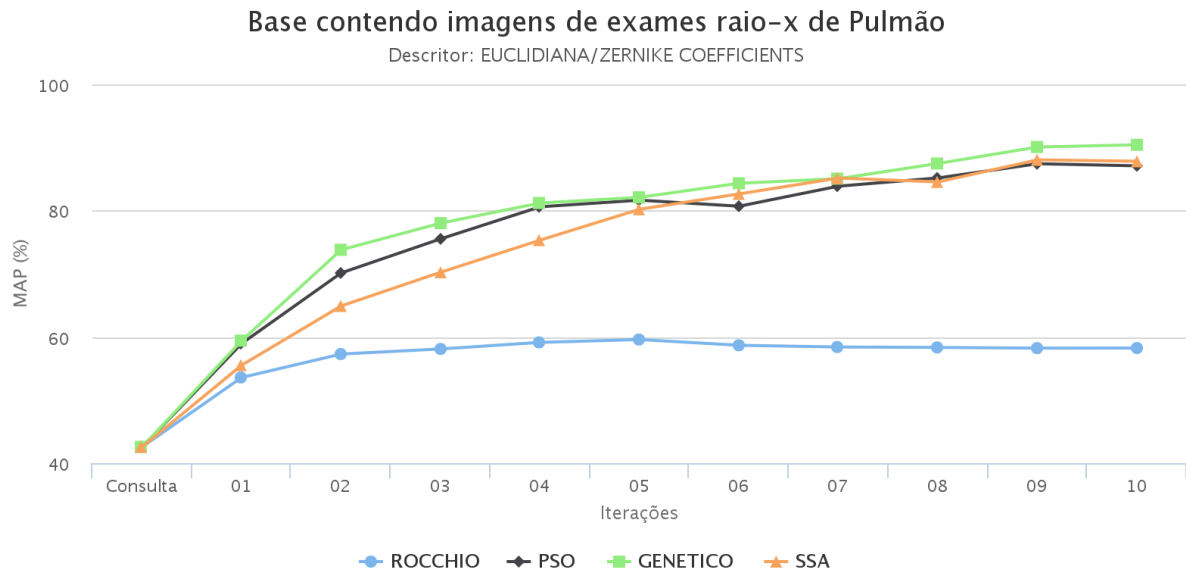
**Figura 30:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor *Euclidiana/Comb Moments (Fourier)*.



**Figura 31:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor *Manhattan/Comb Moments (Wavelet)*.

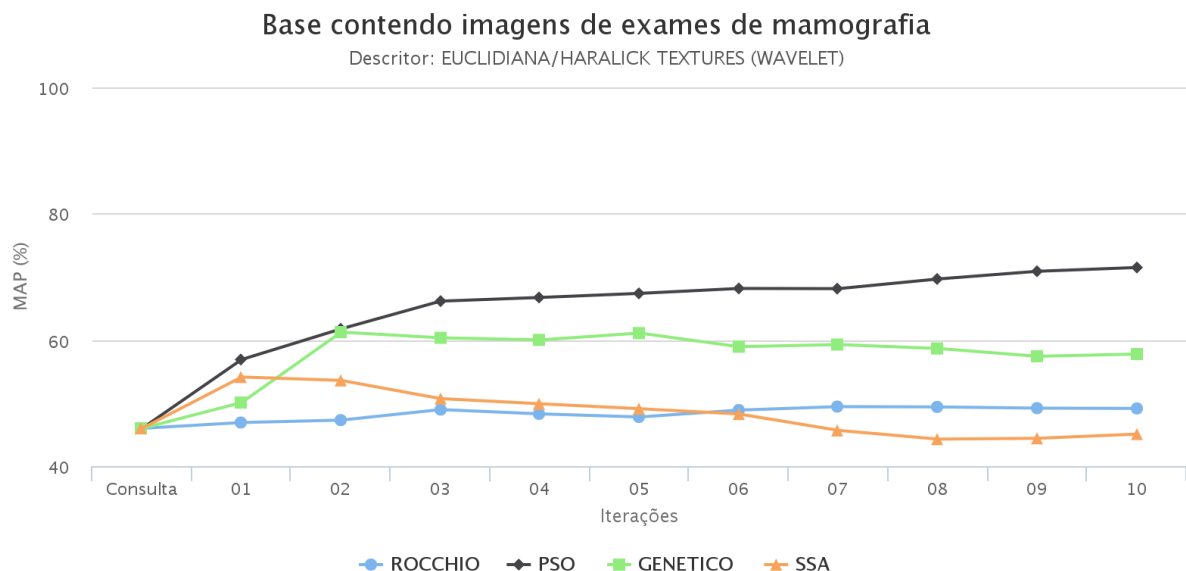
SSA e Genético, alcançando ganhos de precisão de 49,6%, 55,3% e 50,8% respectivamente, em comparação com a abordagem da literatura.

Os experimentos com a base de mamografia utilizando características de textura *Haralick Textures (Wavelet)* e distância Euclidiana mostraram que a instância PSO teve um desempenho superior comparado às outras. Conforme pode ser observado na Figura 33, o PSO já no primeiro ciclo obteve um ganho de 21,3% em comparação com a abordagem da literatura. No



**Figura 32:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Euclidiana/Zernike Coefficients.

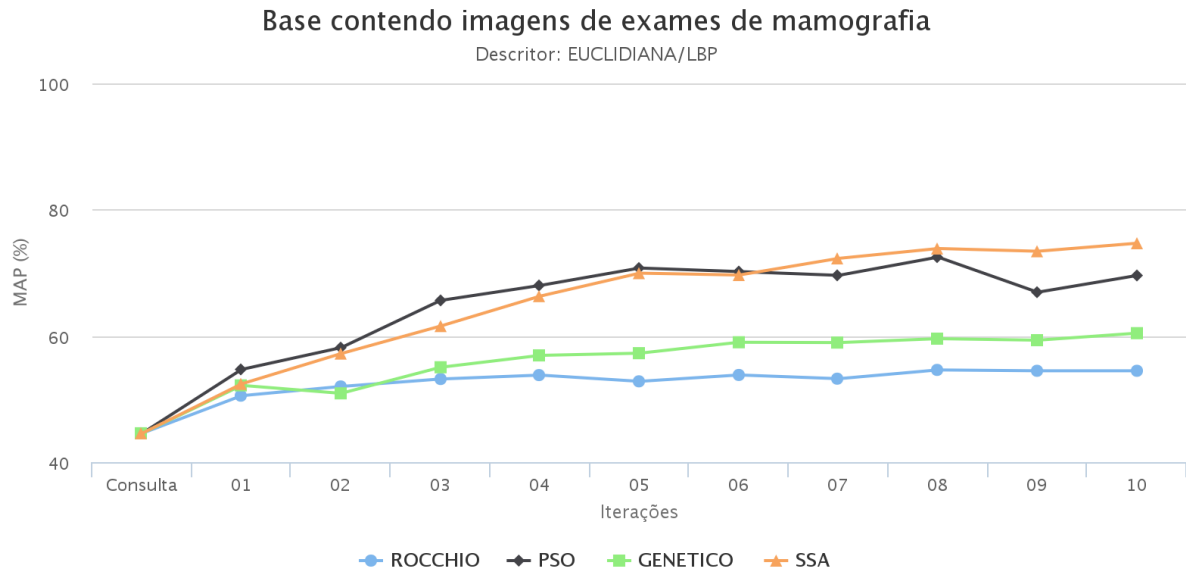
último ciclo, o mesmo obteve ganho de 17,5% quando comparado com a instância do Genético, bem como um ganho de 45,4% contra a abordagem da literatura. A instância com SSA apesar de não manter a vantagem até o ciclo final obteve ganho nos três primeiros ciclos.



**Figura 33:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Euclidiana/Haralick Textures (Wavelet).

Analisando a Figura 34, a qual explicita os resultados obtidos pela instância aplicando o PSO combinado à função de distância Euclidiana e características oriundas do LBP, pode-se verificar que tal instância apresentou um ganho de 8,2% no primeiro ciclo e 23,4% no terceiro,

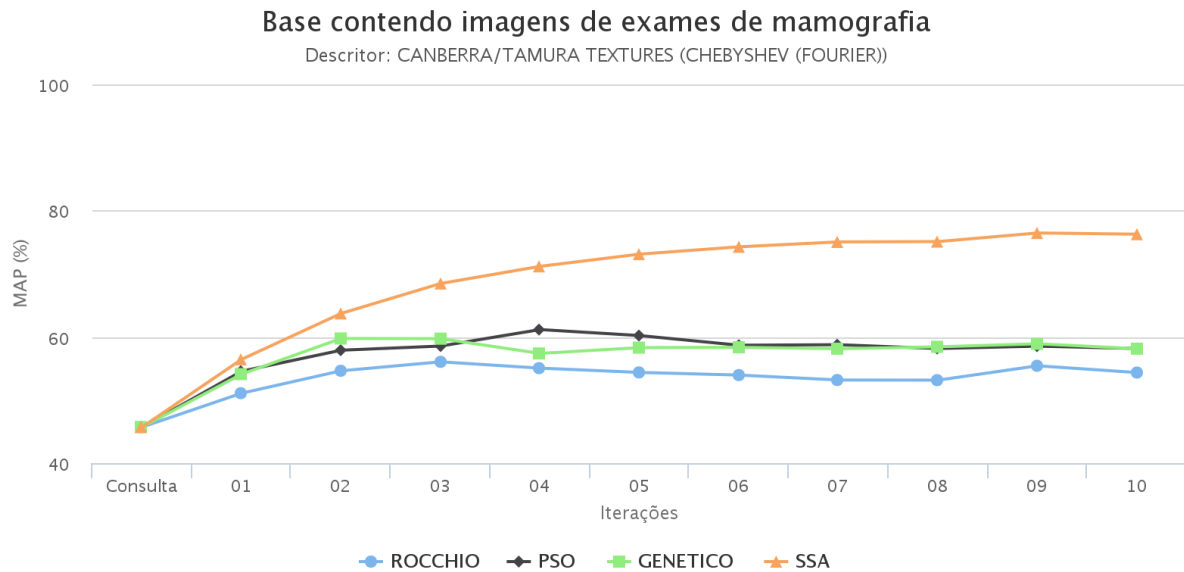
comparando com a abordagem da literatura. Nota-se também que apesar da instância com SSA ter obtido ganhos menores nos primeiros ciclos, no último ciclo a mesma alcançou um ganho de 37,1%. O algoritmo Genético, apesar de apresentar resultados melhores em relação aos da instância da literatura, apresentou uma eficácia mais modesta quando comparada ao PSO e SSA.



**Figura 34: Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Euclidiana/LBP.**

Avaliando os resultados da base de mamografia obtidos por meio das características de textura *Tamura Textures (Chebyshev (Fourier))* e distância Canberra, pode-se perceber que a instância com algoritmo SSA aparece em destaque na Figura 35. O SSA obteve um bom ganho já nos três primeiros ciclos, 10,5% no primeiro, 16,7% no segundo e 22,1% no terceiro. Além de começar com vantagem sobre as outras instâncias, o SSA manteve os ganhos até o último ciclo, que foi de 40,3% comparando com a abordagem sem otimização.

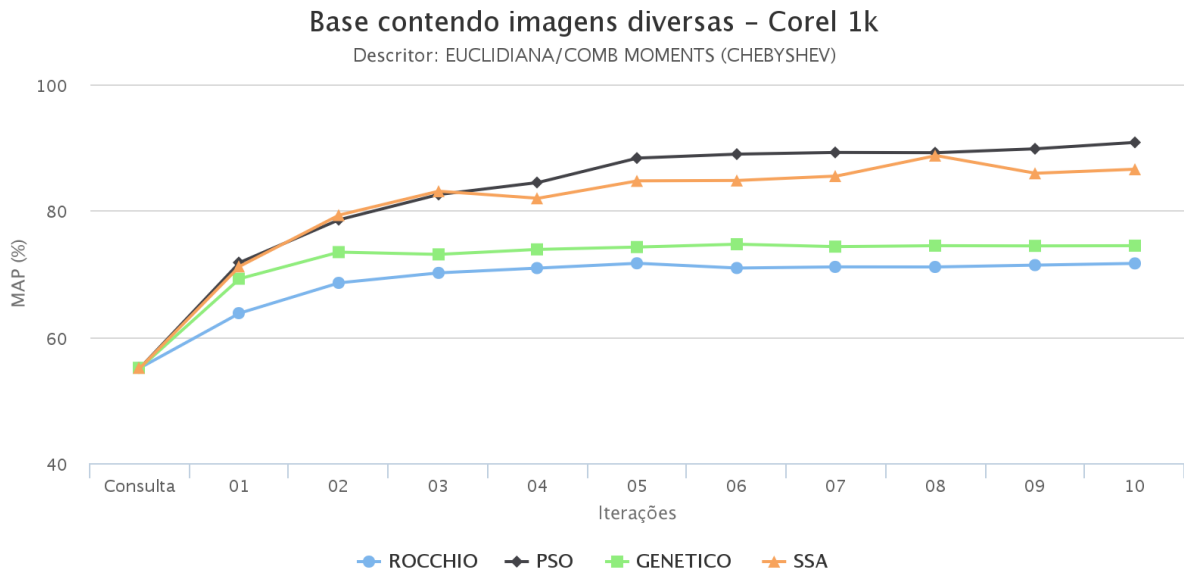




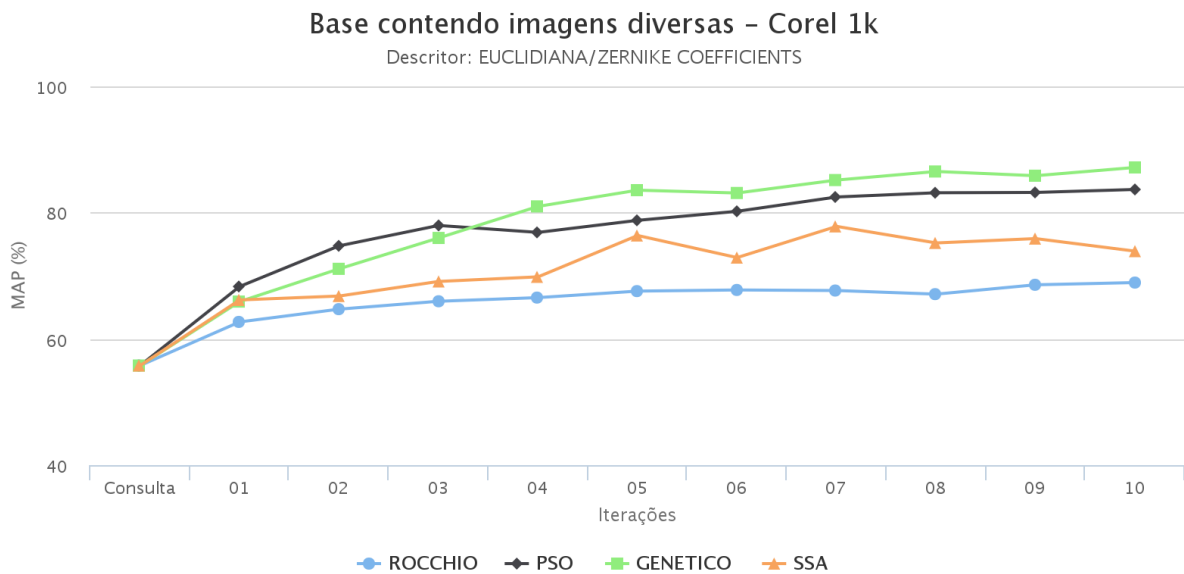
**Figura 35: Resultado obtido por meio da otimização de parâmetros offline, utilizando a base contendo imagens de mamografia e descritor Canberra/Tamura Textures (Chebyshev (Fourier)).**

Para corroborar os resultados e aumentar a robustez e a eficácia da abordagem proposta foram realizados experimentos também utilizando a base de imagens *corel*, assim testando instâncias do arcabouço com imagens de diferentes contextos. A Figura 36 demonstra os resultados dos experimentos aplicando as características obtidas pelo extrator *Comb Moments (Chebyshev)* e distância Euclidiana. Comparando com a instância do Rocchio com valores da literatura, a instância PSO obteve ganho de 12,6% no primeiro ciclo aumentando para 15,7% no segundo, chegando ao último ciclo com um ganho de 26,8%. A instância com SSA também apresentou uma eficácia similar à do PSO, sendo que quando comparada à abordagem da literatura o ganho no primeiro ciclo foi de 11,6% subindo para 15,7% no segundo ciclo e mantendo o ganho até o último ciclo, no qual obteve 20,8%. O Genético apresentou uma eficácia inferior quando comparado ao PSO e SSA, porém, quando comparado com o método *baseline* da literatura apresentou ganho de 8,7% no primeiro ciclo, ganho de 7,1% no segundo, bem como no último ciclo com obteve cerca de 3,9% de ganho na precisão.

Por fim, a Figura 37 ilustra os resultados obtidos utilizando características oriundas do *Zernike Coefficients* em conjunto com a distância Euclidiana. Nesse caso, a instância com PSO obteve os melhores resultados nas primeiras três iterações, com ganhos de 8,9%, 15,5% e 18,2% respectivamente, comparados com a abordagem da literatura. A partir do quarto ciclo a instância do Genético destaca-se, com ganhos superiores ao PSO, chegando ao último ciclo com um ganho de 25,5%. Entre as instâncias com algoritmos evolutivos, o SSA teve o menor desempenho, mesmo assim apresentou maior eficácia que a abordagem tradicional.



**Figura 36:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base *Corel* e descritor *Euclidiana/Comb Moments (Chebyshev)*.



**Figura 37:** Resultado obtido por meio da otimização de parâmetros offline, utilizando a base *Corel* e descritor *Euclidiana/Zernike Coefficients*.

#### 4.3.1 REDUÇÃO DE DIMENSIONALIDADE

Para o segundo experimento foram executados os passos de seleção de atributos, com o objetivo de reduzir a dimensionalidade, conforme algoritmos e parametrização descritos na Seção 3.1.2 do Capítulo 3. A Tabela 6 explicita a redução de atributos para cada característica selecionada para o experimento.

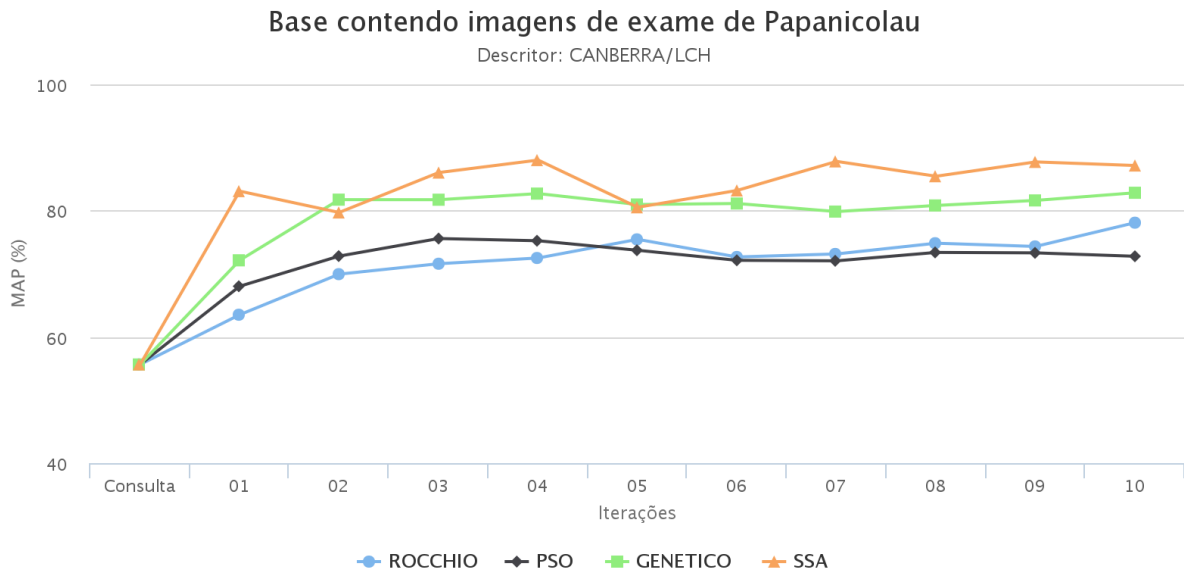
**Tabela 6: Tamanho do vetor de características após processo de seleção de atributos**

<b>Base de imagens</b>	<b>Características</b>	<b>Atributos</b>	<b>Selecionados</b>
<i>pap-smear</i>	Haralick Textures (Wavelet)	28	10
<i>pap-smear</i>	LCH	240	48
Pulmão	Chebyshev Coefficients (Fourier)	32	15
Pulmão	Comb Moments (Fourier)	48	19
Mamografia	Haralick Textures (Wavelet)	28	10
Mamografia	LBP	256	41
<i>Corel</i>	Comb Moments (Chebyshev)	48	17
<i>Corel</i>	Zernike Coefficients	72	33

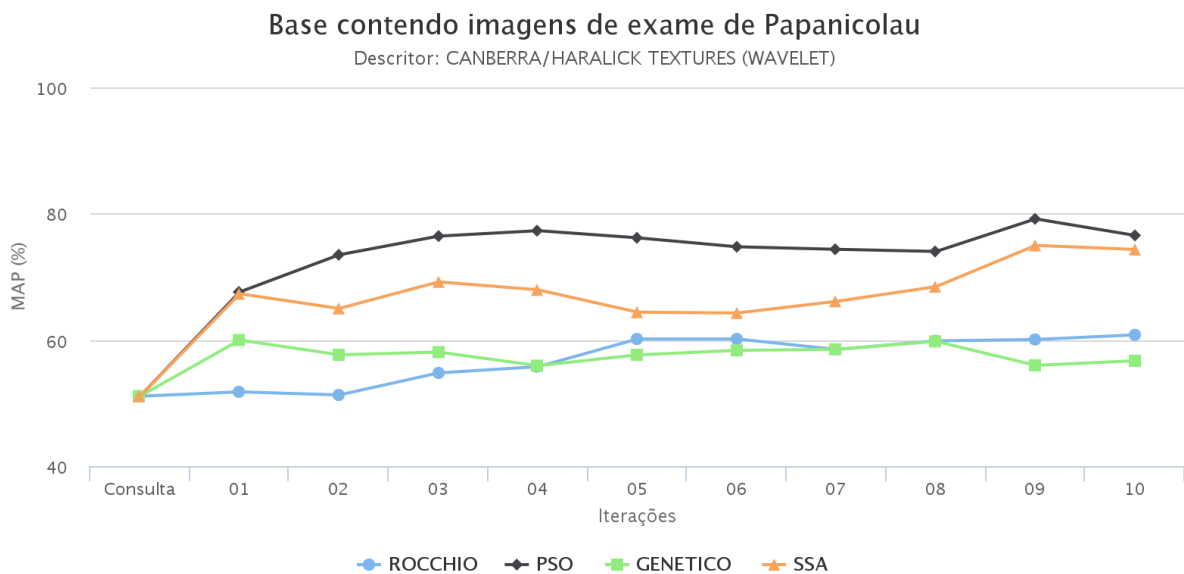
Após executada a redução dos atributos, uma nova iteração de experimentos foi realizada. Os tempos computacionais na execução do experimento antes e após a redução de atributos está descrito na Tabela 7. Com a redução de atributos, os experimentos executaram aproximadamente 5 vezes mais rápido, porém, conforme demonstrado nos gráficos ilustrados nas Figuras 41 a 44, ocorreu uma perda de precisão em torno de 2 a 4%.

**Tabela 7: Tempo de execução (em segundos) dos experimentos antes e após a seleção de atributos**

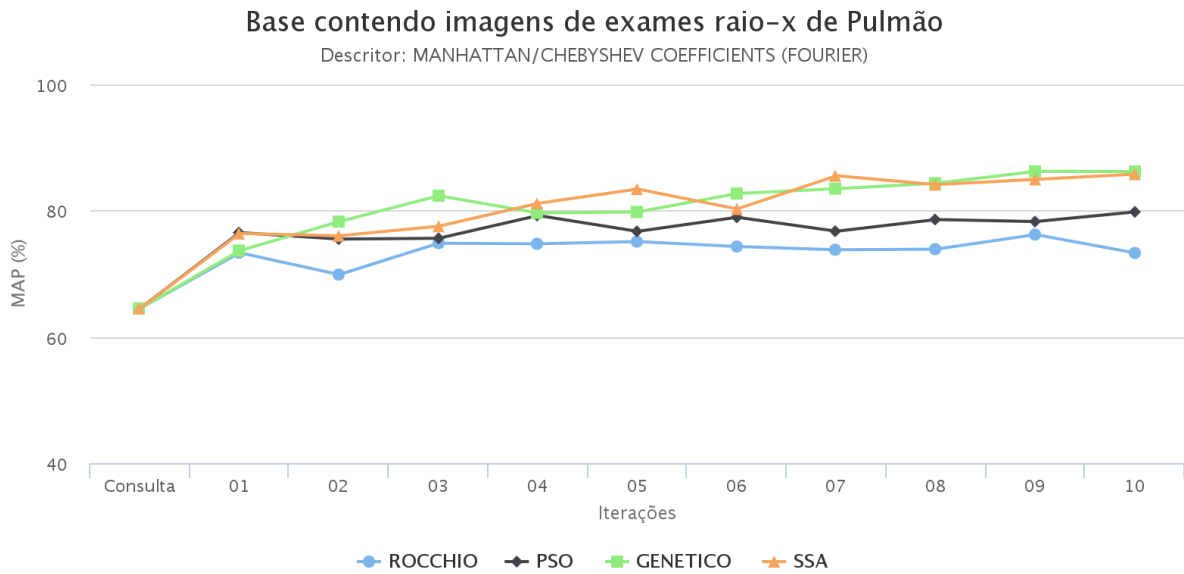
<b>Base de imagens</b>	<b>Características</b>	<b>Distância</b>	<b>Antes</b>	<b>Após</b>
<i>pap-smear</i>	Haralick Textures (Wavelet)	Canberra	3610s	728s
<i>pap-smear</i>	LCH	Canberra	16979s	3424s
Pulmão	Chebyshev Coefficients (Fourier)	Manhattan	16629s	517s
Pulmão	Comb Moments (Fourier)	Euclidiana	5638s	1135s
Mamografia	Haralick Textures (Wavelet)	Euclidiana	3698s	801s
Mamografia	LBP	Euclidiana	63591s	13228s
<i>Corel</i>	Comb Moments (Chebyshev)	Euclidiana	8073s	1608s
<i>Corel</i>	Zernike Coefficients	Euclidiana	1353s	244s



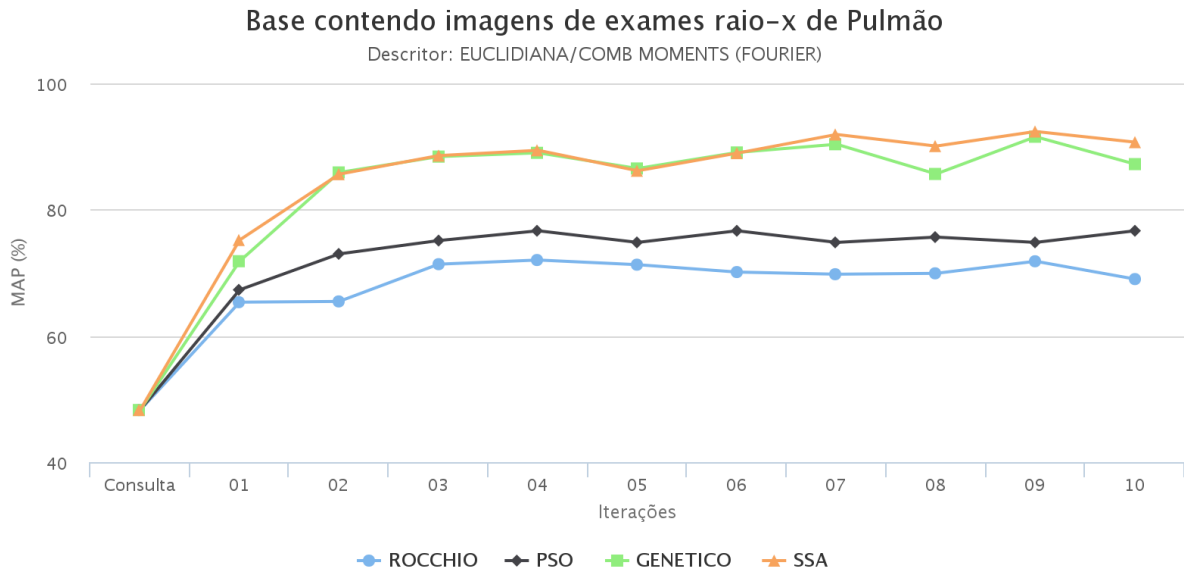
**Figura 38:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/LCH.



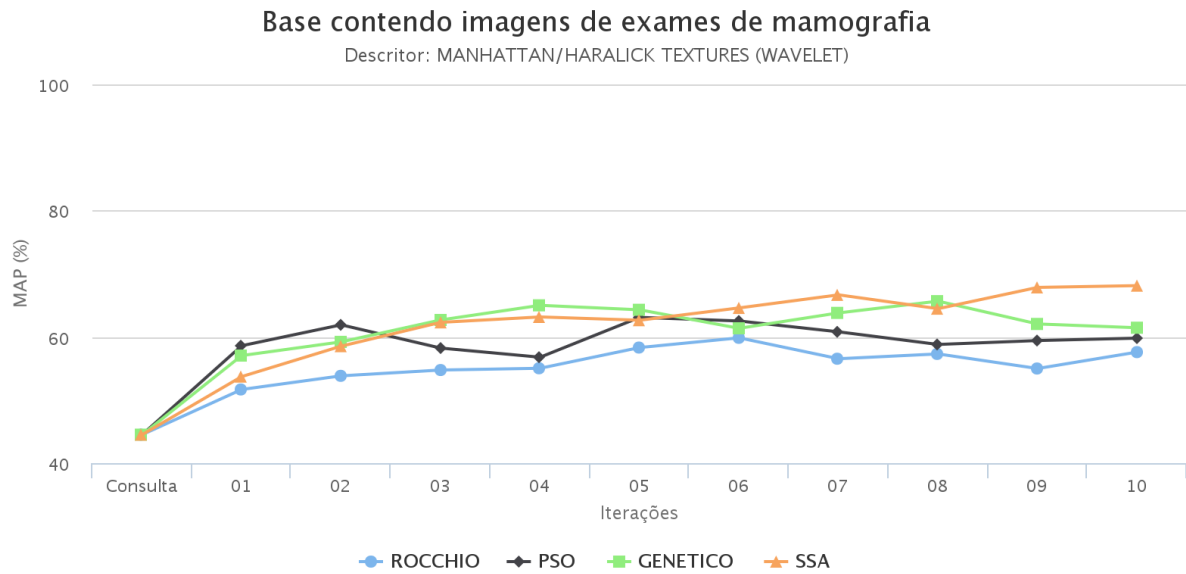
**Figura 39:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de Papanicolau e descritor Canberra/Haralick Textures (Wavelet).



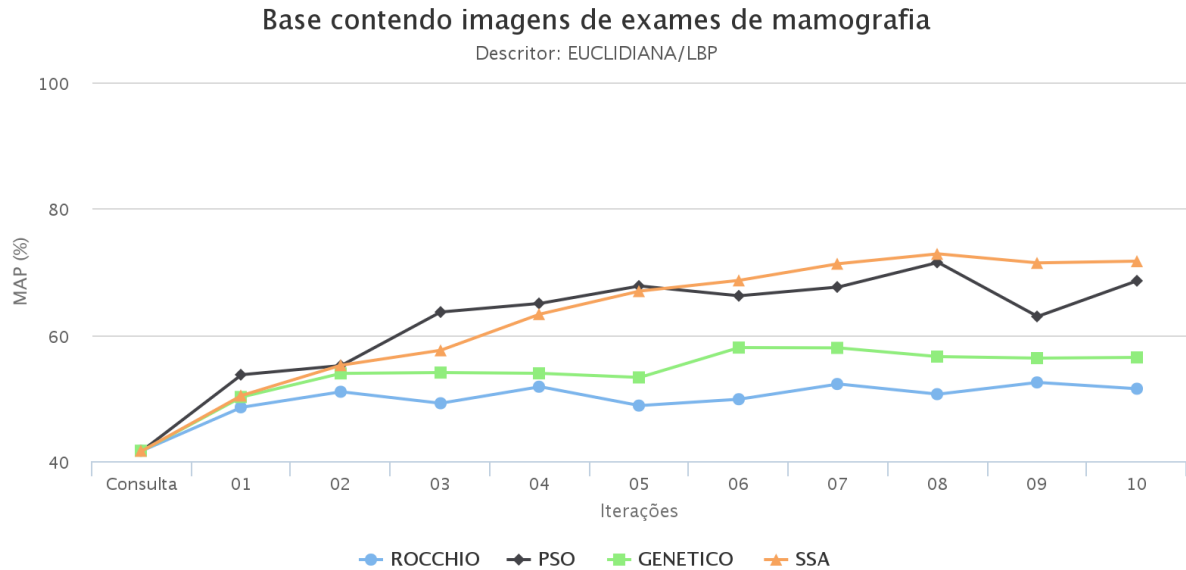
**Figura 40:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Manhattan/Chebyshev Coefficients (Fourier).



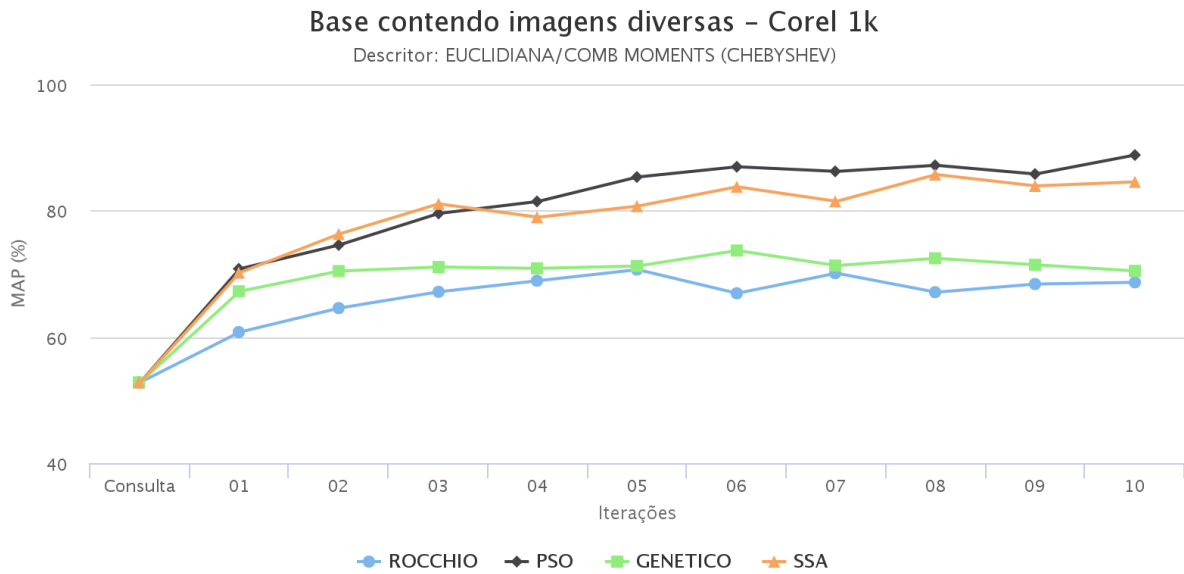
**Figura 41:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de exame de raio-x de pulmão e descritor Euclidian/Comb Moments (Fourier).



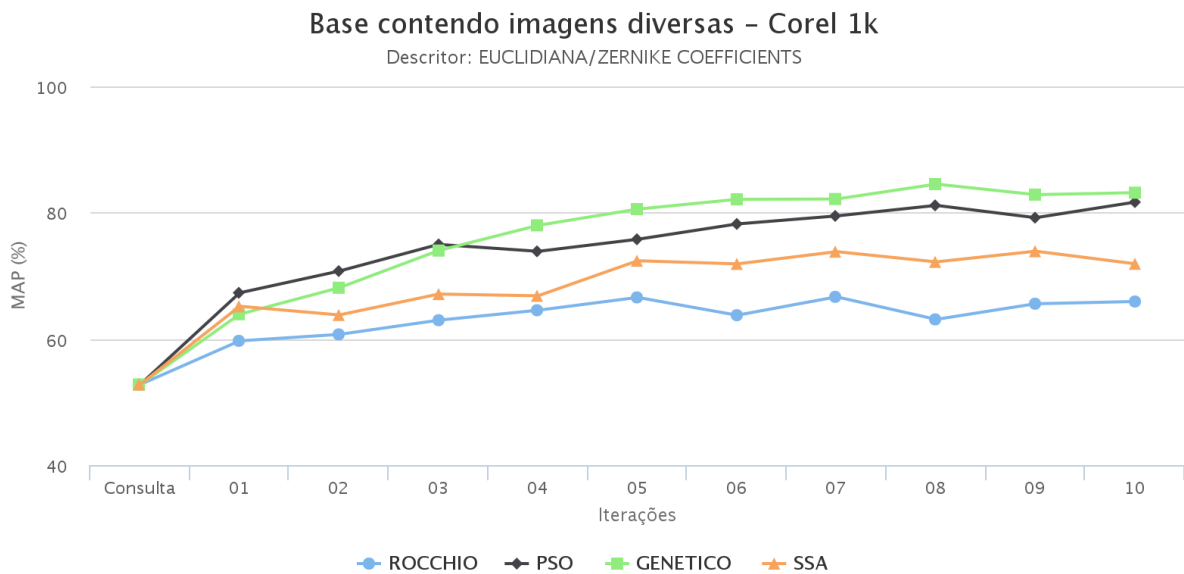
**Figura 42:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de mamografia e descritor Euclidiana/*Haralick Textures (Wavelet)*.



**Figura 43:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base contendo imagens de mamografia e descritor Euclidiana/LBP.



**Figura 44:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base *Corel* e descritor *Euclidiana/Comb Moments (Chebyshev)*.



**Figura 45:** Resultado obtido através da otimização de parâmetros offline com redução de características, utilizando a base *Corel* e descritor *Euclidiana/Zernike Coefficients*.

#### 4.4 CONSIDERAÇÕES FINAIS

No presente capítulo foram explicitados os resultados obtidos pela abordagem proposta, a qual realiza a junção dos processos de realimentação de relevância em sistemas de recuperação de imagens baseada em conteúdo a métodos evolutivos.

Foram aplicadas políticas de calibração e definição dos melhores descritores de acordo com o processo de realimentação, ou seja, segundo a intenção do usuário em relação ao processo de recuperação de imagens. Além disso, foram realizados também experimentos acoplando abordagens de seleção de características ao processo explicitado, com o intuito de reduzir a dimensionalidade dos vetores de características e conseqüentemente o custo computacional das consultas.

Os experimentos realizados, de maneira geral, demonstraram a eficácia e aplicabilidade da abordagem proposta, explicitando a possibilidade de novos caminhos para a aplicação de métodos evolutivos ao processo de realimentação de relevância na recuperação de imagens por conteúdo.



## 5 CONCLUSÕES

O presente trabalho apresentou uma nova abordagem para a realização do processo de realimentação de relevância em recuperação de imagens otimizado por meio de abordagens evolutivas. Diferente das abordagens aplicadas na literatura, a presente proposta teve por objetivo a parametrização do processo de realimentação, de maneira a maximizar a precisão das consultas e atender ao mesmo tempo a intenção do especialista. Fato esse negligenciado pela maioria dos métodos encontrados na literatura, os quais aplicam parametrizações a esmo, sem prévia análise do impacto das mesmas nas precisões das consultas por similaridade.

Os experimentos realizados mostraram que a abordagem proposta alcançou ganhos consideráveis com relação à precisão das consultas, atingindo ganhos de até 55,3% de precisão. Além disso, vale ressaltar que no presente trabalho foram geradas algumas instâncias da abordagem proposta para demonstrar os benefícios da mesma, sendo assim, novas técnicas evolutivas, bem como extratores de características e funções de distância podem ser diretamente acoplados à proposta, gerando novas instâncias da mesma.

### 5.1 RESULTADOS ALCANÇADOS

A partir dos estudos realizados no presente trabalho foram alcançados os seguintes resultados:

- proposta e desenvolvimento de um arcabouço que permite a geração de diferentes instâncias de métodos evolutivos acoplados ao processo de recuperação de imagens;
- proposta de política de definição e combinação do melhor descritor (i.e. par extrator de características/funções de distância) baseada em abordagens evolucionistas;
- aplicação de métodos evolutivos ao processo de ponderação dos fatores de realimentação de relevância por movimentação do centro de consulta;

- fusão dos processos de realimentação e seleção de características às abordagens evolutivas, por meio do arcabouço proposto.

## 5.2 PUBLICAÇÕES

O presente trabalho derivou uma publicação de artigo completo intitulado “*A Novel Framework for Content-Based Image Retrieval Through Relevance Feedback Optimization*” no 20th Iberoamerican Congress on Pattern Recognition (CIARP 2015), bem como um resumo estendido intitulado “*Exploiting Evolutionary Approaches for Content-Based Medical Image Retrieval*” no 28th IEEE International Symposium on Computer-Based Medical Systems (CBMS 2015).

## 5.3 PROPOSTAS PARA TRABALHOS FUTUROS

Para estudos futuros alguns tópicos poderão ser abordados em complementação ao presente trabalho:

- Combinação de descritores por meio de abordagens evolutivas;
- Junção de outras técnicas de realimentação de relevância gerando diferentes instâncias do arcabouço proposto;
- Melhoria na eficiência dos experimentos por meio da implementação de índices voltados ao processo de consultas por similaridade;
- Aplicação do arcabouço proposto a outros contextos de imagem, bem como de dados complexos (e.g. vídeos, sons);
- Acoplagem de uma maior gama de algoritmos de otimização e/ou evolutivos ao arcabouço proposto.

## REFERÊNCIAS

- ABOULMAGD, H. et al. A new approach in content-based image retrieval using fuzzy. **Telecommunication Systems**, Springer, v. 40, n. 1-2, p. 55–66, 2009.
- ALMEIDA, J. et al. Making colors worth more than a thousand words. In: **Proceedings of the 2008 ACM Symposium on Applied Computing**. New York, NY, USA: ACM, 2008. (SAC '08), p. 1180–1186. ISBN 978-1-59593-753-7.
- AN, Y. et al. CBIR based on adaptive segmentation of HSV color space. In: **12th International Conference on Computer Modelling and Simulation (UKSim), 2010**. [S.l.: s.n.], 2010. p. 248–251.
- AREVALILLO-HERRÁEZ, M. et al. Improving distance based image retrieval using non-dominated sorting genetic algorithm. **Pattern Recognition Letters**, n. 0, 2014. ISSN 0167-8655.
- BAEZA-YATES, R. et al. Modern information retrieval: the concepts and technology behind search second edition. **Addision Wesley**, 2011.
- BARROS, L. R. M. **Cor no processo criativo,(A). Um estudo sobre a Bauhaus e a teoria de Goethe**. [S.l.]: Senac, 2006.
- BERNHARD, D. Query expansion based on pseudo relevance feedback from definition clusters. In: **Proceedings of the 23rd International Conference on Computational Linguistics: Posters**. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010. (COLING '10), p. 54–62.
- BIMBO, A. D. **Visual information retrieval**. San Francisco, CA, USA: Morgan and Kaufmann, 1999.
- BUGATTI, P. et al. Improving content-based retrieval of medical images through dynamic distance on relevance feedback. In: **24th International Symposium Computer-Based Medical Systems (CBMS)**. [S.l.: s.n.], 2011. p. 1–6. ISSN 1063-7125.
- BUGATTI, P. H. et al. Prosper: Perceptual similarity queries in medical cbir systems through user profiles. **Computers in biology and medicine**, Elsevier, v. 45, p. 8–19, 2014.
- CALUMBY, R. T. et al. Multimodal retrieval with relevance feedback based on genetic programming. **Multimedia Tools and Applications**, Springer, v. 69, n. 3, p. 991–1019, 2014.
- CHA, S. Comprehensive survey on distance/similarity measures between probability density functions. **International Journal of Mathematical Models and Methods in Applied Sciences**, v. 1, n. 4, p. 300–307, 2007.
- CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. **Computers & Electrical Engineering**, v. 40, n. 1, p. 16–28, 2014. ISSN 0045-7906.

- CHANG, B. et al. Using visual features to design a content-based image retrieval method optimized by particle swarm optimization algorithm. **Engineering Applications of Artificial Intelligence**, v. 26, n. 10, p. 2372 – 2382, 2013. ISSN 0952-1976.
- CHINO, D. Y. et al. Bowfire: Detection of fire in still images by integrating pixel color and texture analysis. In: IEEE. **28th SIBGRAPI Conference on Graphics, Patterns and Images**. [S.l.], 2015. p. 95–102.
- COSTA, A. et al. An efficient algorithm for fractal analysis of textures. In: **25th SIBGRAPI Conference on Graphics, Patterns and Images**. [S.l.: s.n.], 2012. p. 39–46. ISSN 1530-1834.
- ČREPINŠEK, M. et al. Exploration and exploitation in evolutionary algorithms: a survey. **ACM Computing Surveys (CSUR)**, ACM, v. 45, n. 3, p. 35, 2013.
- DARWISH, S. M.; ALI, R. A. Observations on using type-2 fuzzy logic for reducing semantic gap in content-based image retrieval system. In: . [S.l.: s.n.], 2015. v. 7, n. 2, p. 828–843.
- DATTA, R. et al. Image retrieval: Ideas, influences, and trends of the new age. **ACM Comput. Surv.**, ACM, New York, NY, USA, v. 40, n. 2, p. 5:1–5:60, maio 2008. ISSN 0360-0300.
- DIAS, R. L. et al. Reducing the complexity of k-nearest diverse neighbor queries in medical image datasets through fractal analysis. In: IEEE. **IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2013. p. 101–106.
- DORIGO, M.; BLUM, C. Ant colony optimization theory: A survey. **Theoretical Computer Science**, v. 344, n. 23, p. 243 – 278, 2005. ISSN 0304-3975.
- DOULAMIS, N.; DOULAMIS, A. Evaluation of relevance feedback schemes in content-based in retrieval systems. **Signal Processing: Image Communication**, v. 21, n. 4, p. 334–357, 2006. ISSN 0923-5965.
- FARIA, F. F. et al. Learning to rank for content-based image retrieval. In: **Proceedings of the international conference on Multimedia information retrieval**. New York, NY, USA: ACM, 2010. (MIR '10), p. 285–294. ISBN 978-1-60558-815-5.
- FILHO, R. F. S. et al. **Desmistificando o Conceito de Consultas por Similaridade: A Busca de Novas Aplicações na Medicina**. [S.l.]: WIM, 2001.
- GABRIEL, P. H. R.; DELBEM, A. C. B. **Fundamentos de algoritmos evolutivos**. [S.l.]: ICMC-USP, 2008.
- GONÇALVES, V. M. et al. Avaliação de funções de similaridade em sistemas de cbir: Uma revisão sistemática. **VI Workshop de Visão Computacional - UNESP - Presidente Prudente**, 2010.
- GONZALES, R. C.; WOODS, R. E. **Digital Image Processing, 2-nd Edition**. [S.l.]: Prentice Hall, 2002.
- GRIGORESCU, S. E. et al. Comparison of texture features based on gabor filters. **IEEE Transactions on Image Processing**, v. 11, n. 10, p. 1160 – 1167, Oct 2002. ISSN 1057-7149.
- GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. **Journal of Machine Learning Research**, JMLR.org, v. 3, p. 1157–1182, mar 2003. ISSN 1532-4435.

- HADJIDEMETRIOU, E. et al. Spatial information in multiresolution histograms. In: **Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001**. [S.l.: s.n.], 2001. v. 1, p. I-702–I-709 vol.1. ISSN 1063-6919.
- HALL, M. A. **Correlation-based feature selection for machine learning**. Tese (Doutorado) — The University of Waikato, 1999.
- HAMDANI, T. M. et al. Hierarchical genetic algorithm with new evaluation function and bi-coded representation for the selection of features considering their confidence rate. **Applied Soft Computing**, Elsevier, v. 11, n. 2, p. 2501–2509, 2011.
- HARALICK, R. M. Statistical and structural approaches to texture. **Proceedings of the IEEE**, IEEE, v. 67, n. 5, p. 786–804, 1979.
- HAUPT, R. L.; HAUPT, S. E. **Practical genetic algorithms**. [S.l.]: John Wiley & Sons, 2004.
- HEATH, M. et al. The digital database for screening mammography. In: CITESEER. **Proceedings of the 5th international workshop on digital mammography**. [S.l.], 2000. p. 212–218.
- IVANOVA, K.; STANCHEV, P. Color harmonies and contrasts search in art image collections. In: **First International Conference on Advances in Multimedia (MMEDIA)**. [S.l.: s.n.], 2009. p. 180–187.
- JANTZEN, J. et al. Pap-smear benchmark data for pattern classification. **Nature inspired Smart Information Systems (NiSIS 2005)**, p. 1–9, 2005.
- JIN, X.; FRENCH, J. C. Improving image retrieval effectiveness via multiple queries. **Multimedia Tools and Applications**, Springer, v. 26, n. 2, p. 221–245, 2005.
- KACEM, A. et al. Time-sensitive user profile for optimizing search personalization. In: **User Modeling, Adaptation, and Personalization**. [S.l.]: Springer, 2014. p. 111–121.
- KATO, T. Database architecture for content-based image retrieval. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology**. [S.l.], 1992. p. 112–123.
- KEKRE, H. et al. Improving performance of multileveled btc based cbir using sundry color spaces. **International Journal of Image Processing**, v. 4, n. 6, p. 620–630, 2010.
- KIM, D.-H.; CHUNG, C.-W. Qcluster: relevance feedback using adaptive clustering for content-based image retrieval. In: ACM. **Proceedings of the 2003 ACM SIGMOD international conference on Management of data**. [S.l.], 2003. p. 599–610.
- LAM, M. et al. Content-based image retrieval for pulmonary computed tomography nodule images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Medical Imaging**. [S.l.], 2007. p. 65160N–65160N.
- LEE, D.-H.; KIM, H.-J. A fast content-based indexing and retrieval technique by the shape information in large image database. **Journal of systems and software**, Elsevier, v. 56, n. 2, p. 165–182, 2001.
- LIN, W.-C. et al. The effect of low-level image features on pseudo relevance feedback. **Neurocomputing**, v. 166, p. 26–37, 2015. ISSN 0925-2312.

- LINDEN, R. **Algoritmos Genéticos (2a edição)**. [S.l.]: BRASPORT, 2008. ISBN 9788574523736.
- LIU, D. et al. Fast query point movement techniques for large cbir systems. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 21, n. 5, p. 729–743, 2009.
- LIU, Y. et al. A survey of content-based image retrieval with high-level semantics. **Pattern Recognition**, Elsevier, v. 40, n. 1, p. 262–282, 2007. ISSN 0031-3203.
- LYLE, J. R. et al. Soft biometric classification using local appearance periocular region features. **Pattern Recognition**, v. 45, n. 11, p. 3877 – 3885, 2012. ISSN 0031-3203.
- MANJUNATH, B. S.; MA, W.-Y. Texture features for browsing and retrieval of image data. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 18, n. 8, p. 837–842, 1996.
- MANNING, C. D. et al. **Introduction to Information Retrieval**. New York, NY, USA: Cambridge University Press, 2008. ISBN 0521865719, 9780521865715.
- MIYAZAWA, Y. et al. Context-aware recommendation system using content based image retrieval with dynamic context considered. In: **International Conference on Signal-Image Technology Internet-Based Systems (SITIS)**. [S.l.: s.n.], 2013. p. 779–783.
- MOSBAH, M.; BOUCHEHAM, B. Relevance feedback within cbir systems. **International Journal of Computer, Information, Systems and Control Engineering**, v. 8, n. 4, 2014.
- MURPHY, R. F. et al. Searching online journals for fluorescence microscope images depicting protein subcellular location patterns. In: IEEE. **Proceedings of the IEEE 2nd International Symposium on Bioinformatics and Bioengineering Conference**. [S.l.], 2001. p. 119–128.
- OKAL, B.; NUCHTER, A. Sliced curvature scale space for representing and recognizing 3d objects. In: IEEE. **16th International Conference on Advanced Robotics (ICAR)**. [S.l.], 2013. p. 1–7.
- ORLOV, N. et al. **Computer vision for microscopy applications**. [S.l.]: INTECH Open Access Publisher, 2007. 221-244 p. ISBN 987-3-902613-05-9.
- PENATTI, O. A. B. **Estudo comparativo de descritores para recuperação de imagens por conteúdo na web**. Dissertação (Mestrado) — Universidade Estadual de Campinas - Instituto de Computação, 2009.
- PINJARKAR MANISHA SHARMA, K. M. L. Comparison and analysis of content based image retrieval systems based on relevance feedback. **Journal of Emerging Trends in Computing and Information Sciences**, Journal of Emerging Trends in Computing and Information Sciences, v. 3, n. 6, p. 833–837, July 2012.
- PORKAEW, K.; CHAKRABARTI, K. Query refinement for multimedia similarity retrieval in mars. In: ACM. **Proceedings of the seventh ACM international conference on Multimedia (Part 1)**. [S.l.], 1999. p. 235–238.
- PREIM, B.; BARTZ, D. **Visualization in medicine: theory, algorithms, and applications**. [S.l.]: Morgan Kaufmann, 2007.

- QIU, F.; CHO, J. Automatic identification of user interest for personalized search. In: **ACM. Proceedings of the 15th international conference on World Wide Web**. [S.l.], 2006. p. 727–736.
- RAHMAN, M. M. et al. A learning-based similarity fusion and filtering approach for biomedical image retrieval using svm classification and relevance feedback. **IEEE Transactions on Information Technology in Biomedicine**, IEEE, v. 15, n. 4, p. 640–646, 2011.
- RASHEDI, E. et al. A simultaneous feature adaptation and feature selection method for content-based image retrieval systems. **Knowledge-Based Systems**, v. 39, n. 0, p. 85–94, 2013. ISSN 0950-7051.
- RAZENTE, H. L. **Adequando consultas por similaridade para reduzir a descontinuidade semântica na recuperação de imagens por conteúdo**. Tese (Doutorado) — Universidade de São Paulo, 2009.
- ROCCHIO, J. J. Relevance feedback in information retrieval. Prentice-Hall, Englewood Cliffs NJ, 1971.
- ROJBANI, H. et al.  $R\theta$ -signature: A new signature based on radon transform and its application in buildings extraction. In: **2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)**. [S.l.: s.n.], 2011. p. 490–495. ISSN 2162-7843.
- RUI, Y. et al. Relevance feedback: a power tool for interactive content-based image retrieval. **IEEE Transactions on Circuits and Systems for Video Technology**, IEEE, v. 8, n. 5, p. 644–655, 1998.
- SANTOS, J. A. et al. Evaluating the potential of texture and color descriptors for remote sensing image retrieval and classification. In: **International Conference on Computer Vision Theory and Applications (VISAPP)**. [S.l.: s.n.], 2010. p. 203–208.
- SANTOS, K. C. et al. Recuperação de imagens da web utilizando múltiplas evidências textuais e programação genética. In: **24th Brazilian Symposium on Database (SBBDB)**. [S.l.: s.n.], 2009. p. 91–105.
- SHAMIR, L. et al. Wndchrm - an open source utility for biological image analysis. **Source code for biology and medicine**, BioMed Central, v. 3, n. 1, p. 1, 2008.
- SILVA, A. T. et al. Active learning paradigms for CBIR systems based on optimum-path forest classification. **Pattern Recognition**, v. 44, n. 12, p. 2971 – 2978, 2011. ISSN 0031-3203.
- SILVA, M. P. **Sistematização da percepção médica na construção de sistemas para recuperação de imagens por conteúdo**. Tese (Doutorado) — Universidade de São Paulo, 2014.
- SILVA, S. F. **Seleção de características por meio de algoritmos genéticos para aprimoramento de rankings e de modelos de classificação**. Tese (Doutorado) — Universidade de São Paulo, 2011.
- SILVA, S. F. et al. Ranking evaluation functions to improve genetic feature selection in content-based image retrieval of mammograms. In: **IEEE. 22nd IEEE International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2009. p. 1–8.

- SIVAKAMASUNDARI, G.; SEENIVASAGAM, V. Different relevance feedback techniques in cbir: A survey and comparative study. In: **2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)**. [S.l.: s.n.], 2012. p. 1115–1121.
- SMEULDERS, A. et al. Content-based image retrieval at the end of the early years. **IEEE Transactions Pattern Analysis and Machine Intelligence**, v. 22, n. 12, p. 1349–1380, 2000. ISSN 0162-8828.
- SONG, Y. et al. Research on user-personalized image retrieval method. **International Journal of Multimedia & Ubiquitous Engineering**, v. 9, n. 6, 2014.
- STEHLING, R. O. et al. A compact and efficient image retrieval approach based on border/interior pixel classification. In: ACM. **Proceedings of the eleventh international conference on Information and knowledge management**. [S.l.], 2002. p. 102–109.
- STRICKER, M. A.; ORENGO, M. Similarity of color images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology**. [S.l.], 1995. p. 381–392.
- SU, J.-H. et al. Efficient relevance feedback for content-based image retrieval by mining user navigation patterns. **IEEE Transactions Knowledge and Data Engineering**, v. 23, n. 3, p. 360–372, 2011. ISSN 1041-4347.
- SWAIN, M. J.; BALLARD, D. H. Color indexing. **International journal of computer vision**, Springer, v. 7, n. 1, p. 11–32, 1991.
- TABBONE, S. et al. A new shape descriptor defined on the radon transform. **Computer Vision and Image Understanding**, v. 102, n. 1, p. 42–51, 2006. ISSN 1077-3142.
- TAMURA, H. et al. Textural features corresponding to visual perception. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 8, n. 6, p. 460–473, June 1978. ISSN 0018-9472.
- TORRES, R. d. S.; FALCÃO, A. X. Recuperação de imagens baseadas em conteúdo. In: **Revista de Informática Teórica e Aplicada - RITA**. [S.l.: s.n.], 2008. v. 15.
- TORRES, R. d. S. et al. A genetic programming framework for content-based image retrieval. **Pattern Recognition**, v. 42, n. 2, p. 283 – 292, 2009. ISSN 0031-3203.
- TORRES, R. d. S. et al. Recuperação de imagens: Desafios e novos rumos. In: **Anais do XXVIII Congresso da SBC - SEMISH - Seminário Integrado de Software e Hardware**. [S.l.: s.n.], 2008. p. 223–237.
- TRAINA, A. et al. Fighting the semantic gap on cbir systems through new relevance feedback techniques. In: **19th IEEE International Symposium Computer-Based Medical Systems (CBMS)**. [S.l.: s.n.], 2006. p. 881–886. ISSN 1063-7125.
- TRAINA, A. J. M. et al. Efficient content-based image retrieval through metric histograms. **World Wide Web**, Springer, v. 6, n. 2, p. 157–185, 2003.
- VERMA, M. et al. Local extrema co-occurrence pattern for color and texture image retrieval. **Neurocomputing**, v. 165, p. 255 – 269, 2015. ISSN 0925-2312.



- WANG, D. et al. A structure-based approach for multimedia information filtering. **Multimedia Tools and Applications**, Springer, v. 29, n. 1, p. 73–89, 2006.
- WANG, H. H. et al. Semantic gap in cbir: Automatic objects spatial relationships semantic extraction and representation. **International Journal of Image Processing**, v. 4, n. 3, p. 192, 2010.
- WILSON, D. R.; MARTINEZ, T. R. Improved heterogeneous distance functions. **CoRR**, cs.AI/9701101, 1997.
- WRIGHT, S. The genetical theory of natural selection a review. **Journal of Heredity**, Am Genetic Assoc, v. 21, n. 8, p. 349–356, 1930.
- XU, F.; ZHANG, Y.-J. Evaluation and comparison of texture descriptors proposed in mpeg-7. **Journal of Visual Communication and Image Representation**, Elsevier, v. 17, n. 4, p. 701–716, 2006.
- XU, X. et al. The application of particle swarm optimization in relevance feedback. In: **FBIE International Conference on Future - BioMedical Information Engineering**. [S.l.: s.n.], 2009. p. 156–159.
- YU, J. J.; LI, V. O. A social spider algorithm for global optimization. **Applied Soft Computing**, v. 30, p. 614 – 627, 2015. ISSN 1568-4946.
- ZEGARRA, J. A. M. et al. Wavelet-based fingerprint image retrieval. **Journal of Computational and Applied Mathematics**, v. 227, n. 2, p. 294–307, 2009. ISSN 0377-0427.
- ZHANG, C.; HUANG, L. Content-based image retrieval using multiple features. **CIT. Journal of Computing and Information Technology**, SRCE-Sveučilišni računski centar, v. 22, n. LISS 2013, p. 1–10, 2014.
- ZHANG, W.-j.; WANG, J.-y. The study of methods for language model based positive and negative relevance feedback in information retrieval. In: **International Symposium - Information Science and Engineering (ISISE)**. [S.l.: s.n.], 2012. p. 39–43. ISSN 2160-1283.
- ZHOU, X. S.; HUANG, T. S. Relevance feedback in image retrieval: A comprehensive review. **Multimedia Systems**, Springer-Verlag, v. 8, n. 6, p. 536–544, 2003. ISSN 0942-4962.
- ZHUO, L. et al. A comparative study of dimensionality reduction methods for large-scale image retrieval. **Neurocomputing**, v. 141, n. 0, p. 202 – 210, 2014. ISSN 0925-2312.