

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
CÂMPUS CORNÉLIO PROCÓPIO  
DIRETORIA DE GRADUAÇÃO E EDUCAÇÃO PROFISSIONAL  
DEPARTAMENTO DE COMPUTAÇÃO  
CURSO DE ENGENHARIA DE COMPUTAÇÃO

WILLIAM RENAN DEBRINO

**ESTUDO DE MÉTODOS DE RE-RANKING EM IMAGENS**

TRABALHO DE CONCLUSÃO DE CURSO

CORNÉLIO PROCÓPIO

2016

**WILLIAM RENAN DEBRINO**

**ESTUDO DE MÉTODOS DE RE-RANKING EM IMAGENS**

Trabalho de Conclusão de Curso apresentada à Disciplina de Trabalho de Conclusão de Curso II, Departamento de Computação da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de Bacharel em Engenharia de Computação

Orientador: Prof. Dr. Pedro Henrique Bugatti

**CORNÉLIO PROCÓPIO**

**2016**



---

## **TERMO DE APROVAÇÃO**

### ESTUDO DE MÉTODOS DE RE-RANKING EM IMAGENS

por

William Renan Debrino

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Engenheiro em Engenharia da Computação” e aprovado em sua forma final pelo Programa de Graduação em Engenharia de Computação da Universidade Tecnológica Federal do Paraná.

Cornélio Procópio, 15/06/2016

---

Prof. Dr. Pedro Henrique Bugatti  
Universidade Tecnológica Federal do Paraná

---

Prof. Dra. Priscila Tiemi Maeda Saito  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Silvio Ricardo Rodrigues Sanches  
Universidade Tecnológica Federal do Paraná

## **AGRADECIMENTOS**

A Deus por ter me dado força para superar as dificuldades.

Ao meu orientador pelo suporte, correções e incentivos.

A minha família pelo incentivo e apoio incondicional nas horas mais difíceis.

A universidade, seu corpo docente e direção que oportunizaram este trabalho, além de proporcionar o apoio à execução deste trabalho por meio do Programa de Bolsas de Fomento às Ações de Graduação.

E a meus amigos, que me ajudaram direta ou indiretamente ao longo de minha formação.

## RESUMO

DEBRINO, William Renan. ESTUDO DE MÉTODOS DE RE-RANKING EM IMAGENS. 54 f. Trabalho de Conclusão de Curso – Curso de Engenharia de Computação, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

Em coleções de imagens, a medida que são adicionadas novas imagens, torna-se cada vez mais desgastante para o usuário encontrar uma imagem específica em meio a tantas outras. Assim, faz-se necessário que sejam utilizados métodos para recuperação de uma imagem baseada em critérios definidos pelo usuário. Para isso, é aplicada a abordagem de recuperação de imagens baseada em conteúdo. Diferentemente das abordagens textuais, um processo de Recuperação de Imagens Baseadas no Conteúdo possibilita a utilização de métodos para recuperar as imagens de acordo com critérios estabelecidos pelo usuário, baseando-se no conteúdo da imagem, otimizando assim o processo de busca. Estes sistemas buscam imagens similares em coleções de imagens, sendo que a similaridade pode ser medida de acordo com o conteúdo das imagens. Neste contexto, algoritmos de *Re-Ranking* têm sido utilizados com o objetivo de aumentar a acurácia destes sistemas. Este tipo de algoritmo pode ser utilizado para explorar informações contextuais presentes nos relacionamentos entre algumas imagens. O objetivo deste trabalho foi analisar e estudar um algoritmo de *Re-Ranking*, para utilizá-lo em diferentes estudos de caso para aumentar sua precisão. As principais contribuições foram a análise e a implementação de um algoritmo aplicado a diferentes casos melhorando assim a precisão dos mesmos. Nos experimentos foram utilizados diferentes descritores, funções de distância e medidas de comparação em duas bases de imagens. Os resultados demonstraram que o algoritmo mostrou-se eficiente para a maioria dos casos, mas para cada caso foi necessária a adequação dos parâmetros devido às características de cada caso.

**Palavras-chave:** Busca por Similaridade, Recuperação de Imagens Baseada em Conteúdo, *Re-Ranking*, Imagem

## ABSTRACT

DEBRINO, William Renan. STUDY OF RE-RANKING METHODS IN IMAGE. 54 f. Trabalho de Conclusão de Curso – Curso de Engenharia de Computação, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

The image search processing leads to some issues when the image dataset exponentially grows. Thus, it is necessary to develop techniques to automatically retrieve such images according to the user intention, based on their similarity. To do so, the Content-based Image Retrieval approach is applied. Considering this context, re-ranking algorithms has been used in order to increase de precision of the similarity queries. In this context, algorithms it Re-Ranking has been used in order to increase the accuracy of these systems. This type of algorithm can be used to explore contextual information present in the relationships between some images. The goal of this research was to study and to analyze a Re-Ranking algorithm in order to it in different descriptors to increase its accuracy. The main contributions were the analysis and implementation of an algorithm applied to different cases, seeking to improve the precision of CBIR systems. In the experiments we utilized different descriptors, distance functions and comparison of measures in two image databases. The results showed that the algorithm is effective for most cases, however for each case, it was necessary to adequate the algorithm parameters due to the image database and methods employed.

**Keywords:** Similarity Queries, Content-Based Image Retrieval, Re-Ranking, Image

## LISTA DE FIGURAS

FIGURA 1	– Estrutura de um CBIR .....	15
FIGURA 2	– Histograma de Cores .....	17
FIGURA 3	– Operador LBP .....	18
FIGURA 4	– <i>Chain Codes</i> para 4 e 8 direções .....	20
FIGURA 5	– Forma para <i>Chain Codes</i> .....	20
FIGURA 6	– Consulta por Abrangência e k-Vizinhos Mais Próximos .....	22
FIGURA 7	– Representação das formas geométricas da família <i>Minkowski</i> .....	23
FIGURA 8	– Cálculo da medida $\psi$ : intersecção entre <i>ranked-lists</i> .....	33
FIGURA 9	– Imagens de uma classe na base <i>Illumination Direction</i> .....	38
FIGURA 10	– Imagens de uma classe na base <i>Illumination Color</i> .....	38
FIGURA 11	– Classes utilizadas nos experimentos .....	39
FIGURA 12	– Pipeline de execução .....	41

## LISTA DE TABELAS

TABELA 1	– Resultados iniciais para cada estudo de caso .....	43
TABELA 2	– Precisão e NDCG média por Classe para o melhor estudo de caso. ....	44
TABELA 3	– Ganho/Perda na precisão e NDCG ao aplicar o algoritmo para <i>Illumination Direction</i> .....	46
TABELA 4	– Ganho/Perda na precisão e NDCG ao aplicar o algoritmo para <i>Illumination Color</i> .....	47
TABELA 5	– NDCG utilizando as 10 primeiras posições para <i>Illumination Direction</i>	49



## LISTA DE SIGLAS

CBIR	Content-Based Image Retrieval
CAD	Computer-Aided Diagnosis
BIC	Border/Interior Classification
RGB	Sistema de cores formado por <i>Red, Green e Blue</i>
LBP	Local Binary Pattern
SGBDs	Sistemas de Gerenciamento de Banco de Dados
RQ	Range Query
MAP	Mean Average Precision
NDCG	Normalized Discount Cumulative Gain
kNN	k-Vizinhos mais Próximos
MkNN	Mutual k-Vizinhos mais Próximos
ALOI	Amsterdam Library of Object Images
fd	Função de distância
mdc	Medida de comparação entre as <i>top k lists</i>

## LISTA DE SÍMBOLOS

$P(i, j)$	Valor do ponto (i,j) na matriz de co-ocorrência
$O$	Conjunto de Objetos ( $o_1, o_2, o_3, \dots, o_n$ )
$Do$	Domínio
$d()$	Função de distância
$o_q$	Objeto de consulta
$r_q$	Distância máxima
$k$	Número de vizinhos em relação ao Objeto de Consulta
$L_p$	Funções de distância da família <i>Minkowski</i>
$\{x_1, \dots, x_n\}$	Coordenadas de valores reais
$d_{L1}$	Distância <i>Manhattan</i>
$d_{L2}$	Distância Euclidiana
$L_\infty$	Distância <i>Chabychev</i>
$Pn$	Precisão na posição
$AP$	Precisão média para uma consulta
$N$	Número de elementos recuperados
$rel(n)$	Função que define a relevância de um elemento
$C$	Conjunto de imagens
$D$	Descritor
$\rho$	Função de distância
$img_i$	Imagem com índice $i$
$A$	Matriz de distâncias
$\sigma_q$	Permutação de um conjunto de imagens
$R_q$	<i>Ranking</i> de uma consulta
$\mathcal{R}$	Conjunto de <i>Rankings</i>
$\hat{A}$	Matriz de distâncias mais eficiente
$f_r$	Função que relaciona toda coleção de imagens
$\hat{\mathcal{R}}$	Conjunto de <i>Rankings</i> melhorados
$\mathcal{X}$	Domínio de um objeto
$\mathcal{Y}$	Rótulo de um objeto
$Pr(x, y)$	Probabilidade conjunta de objetos x e y
$l$	Instâncias rotuladas
$u$	Instâncias não rotuladas
$\mathcal{N}(i)$	Vizinhança de uma imagem que contém imagens semelhantes a ela
$top\ k$	Primeiras posições de um <i>ranked-list</i>
$\tau_i$	<i>Top k list</i> de índice $i$
$\rho_c$	Medida de distância contextual
$(t)$	Expoente que indica a iteração atual para uma <i>top k list</i>
$T$	Número máximo de iterações
$\hat{\rho}$	Nova medida de distância baseada na medida de distância contextual
$\mathcal{N}_{kNN}(i)$	Conjunto de vizinhos obtido usando o método kNN
$\tau_{i_{kNN}}(j)$	<i>Top k list</i> definida utilizando kNN

$\mathcal{N}_{MkNN}(i, k)$	Conjunto de vizinhos obtido usando o método MkNN
$k_c$	Medida de superposição
$\psi$	Intersecção da medida de similaridade
$d_\tau(\tau_i, \tau_j, k)$	Medida de distância <i>Tau Kendall</i>
$\bar{K}_{x,y}(\tau_i, \tau_j)$	Função que determina se as imagens estão em ordem comparando seus <i>ranked-lists</i>
$\lambda$	Número das primeiras posições de cada <i>ranked-list</i>
$k_s$	Número inicial de k-vizinhos utilizado no algoritmo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
1.1	MOTIVAÇÃO	12
1.2	OBJETIVOS	13
1.3	ORGANIZAÇÃO DO TEXTO	13
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>14</b>
2.1	RECUPERAÇÃO DE IMAGENS BASEADAS EM CONTEÚDO - CBIR	15
2.1.1	Extração de Características	16
2.1.1.1	Descritores de cor	16
2.1.1.2	Descritores de textura	17
2.1.1.3	Descritores de forma	19
2.1.2	Consultas por Similaridade	21
2.1.2.1	Consulta por Abrangência	21
2.1.2.2	Consulta aos Vizinhos Mais Próximos	21
2.1.3	Avaliação de Desempenho nas Consultas por Similaridade	24
2.1.3.1	Mean Average Precision	24
2.1.3.2	Normalized Discount Cumulative Gain	25
2.2	RE-RANKING	25
2.2.1	Definição de Re-Ranking em Imagens	26
2.2.2	Principais Abordagens Utilizadas	27
2.2.2.1	Abordagens Supervisionadas	27
2.2.2.2	Abordagens Semi-Supervisionadas	27
2.2.2.3	Abordagens Não Supervisionadas	28
2.3	ALGORITMO DE RE-RANKING	28
2.3.1	Medidas de Distância Contextuais baseadas em Ranked-Lists	29
2.3.2	Comparação nos Ranked-Lists	31
2.3.2.1	Conjunto de Vizinhos	31
2.3.2.2	Medidas de Distância Entre Top k Lists	32
2.3.3	O Algoritmo RL-Sim Re-Ranking	34
2.4	CONSIDERAÇÕES FINAIS	36
<b>3</b>	<b>METODOLOGIA DE DESENVOLVIMENTO</b>	<b>37</b>
3.1	BASE DE IMAGENS	37
3.2	EXTRAÇÃO DE CARACTERÍSTICAS, FUNÇÕES DE DISTÂNCIA E MEDIDAS DE COMPARAÇÃO	39
3.3	CENÁRIOS DE APLICAÇÃO	39
3.4	AVALIAÇÃO DOS RESULTADOS	39
3.5	PIPELINE DE EXECUÇÃO	40
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>42</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>50</b>
5.1	SUGESTÕES PARA TRABALHOS FUTUROS	51
	REFERÊNCIAS	52

## 1 INTRODUÇÃO

Com a constante evolução das tecnologias digitais, os recursos multimídia estão cada vez mais acessíveis. Isto vem proporcionando melhorias significativas na aquisição de dados multimídia. Surgem assim cada vez mais grandes coleções com estes tipos de dados, principalmente imagens.

Em coleções de imagens, a medida que são adicionadas novas imagens torna-se cada vez mais desgastante para o usuário encontrar uma imagem específica em meio a tantas outras. Assim, faz-se necessário que sejam utilizados métodos que recuperem imagens com base em critérios especificados pelo usuário, otimizando assim o processo de busca.

Os sistemas mais utilizados para este tipo de busca podem ser baseados na recuperação das imagens baseadas na sua descrição textual e também podem ser recuperadas baseadas em seu conteúdo, os sistemas *Content-Based Image Retrieval* (CBIR). Estes sistemas permitem uma busca por imagens similares em um banco de dados. A similaridade entre imagens pode ser medida de várias formas, podendo ser baseadas em características como forma, cor e textura.

### 1.1 MOTIVAÇÃO

Quando são utilizadas abordagens de recuperação de imagem com base em palavras-chave e metadados textuais enfrentam-se vários problemas. A descrição do conteúdo de uma imagem baseando-se na sua descrição textual é muito difícil, e essa tarefa torna-se ainda mais desafiadora devido ao crescimento e diversidade das coleções de imagens. Aplicações que lidam com grandes coleções de imagens encontram dificuldades na obtenção de descritores textuais, uma vez que a descrição manual é muito custosa, escassa, nem sempre confiável e além disso é impossível antecipar as necessidades do usuário (FARIA et al., 2010).

Diferentemente da abordagem textual, abordagens do tipo CBIR analisam pares de imagens para a geração de listas de resultados, todavia, podem ignorar importantes informações nos relacionamentos entre as imagens. Com o objetivo de aumentar a acurácia de sistemas

CBIR, algoritmos de reclassificação (*Re-Ranking*) (WANG et al., 2011) e agregação de listas (*rank aggregation*) (ZHANG; YE, 2010) têm sido propostos. Algoritmos de *Re-Ranking* são utilizados para explorar informações contextuais intrínsecas em relacionamentos entre imagens enquanto métodos de *rank aggregation* têm sido propostos para combinar resultados produzidos por diferentes descritores de imagens (PEDRONETTE; TORRES, 2013).

## 1.2 OBJETIVOS

O presente trabalho tem como objetivo principal o estudo e análise de métodos de *Re-Ranking* aplicados a uma combinação de diferentes descritores utilizados em sistemas CBIR. Espera-se assim aumentar a acurácia nestes sistemas, bem como elaborar uma análise comparativa dos diferentes métodos utilizados.

A partir do objetivo geral temos os seguintes objetivos específicos:

- Implementação e análise comparativa dos diferentes métodos de um algoritmo de *Re-Ranking* aplicado a diferentes descritores.
- Possíveis propostas de adaptações e melhorias ao algoritmo utilizado.

## 1.3 ORGANIZAÇÃO DO TEXTO

O presente trabalho apresenta a seguinte organização:

- No capítulo atual, foram apresentadas a introdução, motivação para o desenvolvimento do trabalho, bem como os objetivos referentes ao seu desenvolvimento e, por fim, a organização do texto.
- No Capítulo 2 são introduzidos os conceitos básicos de sistemas CBIR, bem como sua estrutura, métodos, além de diferentes abordagens utilizadas em algoritmos de *Re-Ranking*.
- No Capítulo 3, é apresentada a metodologia de desenvolvimento para este trabalho, onde são elucidados as características da base, cenários utilizados e também o pipeline de execução.
- No Capítulo 4 são apresentados os resultados obtidos para os diferentes estudos de casos.
- No Capítulo 5 são apresentadas a conclusão e as sugestões para trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

No presente capítulo serão apresentados os principais conceitos relacionados a Sistemas CBIR, bem como as principais etapas que vão desde o processo de aquisição da imagem, até as listas de imagens retornadas após a busca.

Sistemas CBIR estão sendo cada vez mais utilizados, principalmente na área médica. A agricultura também vem sendo beneficiada com diferentes métodos de processamento de imagem, sobretudo na parte de classificação de sementes e identificação de pragas em grandes áreas cultivadas.

Na área médica, diversos métodos vem sendo propostos voltados para partes isoladas de sistemas CBIR, como por exemplo extratores de características, ou também propostas para utilização deste sistema como um todo, atuando em aplicações computacionais para o auxílio de diagnóstico, as chamadas aplicações *Computer-Aided Diagnosis* (CAD). Tais aplicações têm como principal objetivo auxiliar o profissional da área médica na detecção precoce de patologias diversas. Em Paris (2008) são implementados algoritmos de extração de características utilizando as Transformadas de *Wavelet* e *Wavelet Packet*. Paris (2008) afirma que a Transformada *Wavelet Packet* revelou vantagens expressivas em relação a Transformada *Wavelet*, sendo comparados através avaliadores de desempenho.

Em Felipe (2005) são apresentados novos métodos de extração de características, bem como novas famílias de funções de distância. Estes novos métodos propostos foram aplicados aos descritores de baixo nível, sendo eles baseados em cor, forma e textura. Os experimentos realizados mostraram que os ganhos em precisão são maiores para os métodos propostos, quando comparados com algoritmos tradicionais.

Em Bueno et al. (2002), um novo método de extração de características é proposto, chamado de histograma métrico, sendo desenvolvida também uma nova função de distância métrica. Bueno et al. (2002) afirmaram que ao usar o método de histograma métrico como uma primeira técnica de extração, reduz-se o número de imagens a serem comparadas.

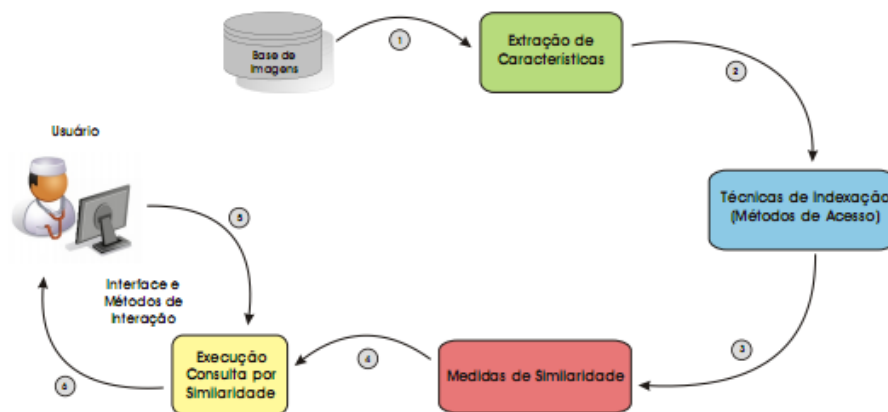
## 2.1 RECUPERAÇÃO DE IMAGENS BASEADAS EM CONTEÚDO - CBIR

Os Sistemas de Recuperação de Imagens baseados no Conteúdo, tem como principal objetivo satisfazer as necessidades dos usuários a partir de especificações de consulta (ANANTHARATNASAMY et al., 2013). Estes sistemas visam recuperar imagens mais similares em uma coleção de imagens, dada uma imagem como parâmetro de consulta (PEDRONETTE; TORRES, 2013).

Os Sistemas CBIR embasam-se em dois principais aspectos: características mais relevantes dos dados e quantificação da similaridade entre imagens. O grande desafio é como alcançar a melhor integração entre estes dois aspectos chaves com intuito de obter maior precisão nas consultas por similaridade (BUGATTI, 2008).

A grande maioria dos sistemas CBIR, são baseados em uma arquitetura básica. Esta estrutura contempla a etapa de armazenamento, indexação e extração de características, e também a obtenção das medidas de similaridade entre as imagens. Estas etapas são integradas por um mecanismo de recuperação de imagens que possibilita a realização de consultas baseadas em conteúdo (MÜLLER et al., 2004).

A Figura 1, apresenta a arquitetura básica de um sistema CBIR.



**Figura 1: Estrutura básica utilizada em sistemas CBIR.**

**Fonte: Bugatti (2008)**



Nesta estrutura, a partir de uma base de imagens são extraídas as características das imagens sendo estas armazenadas em um vetor. De posse deste vetor, aplicam-se as técnicas e métodos a fim de comparar os vetores aos pares ou compará-los em relação a base, para que possam ser feitos os cálculos para suas respectivas medidas de similaridade. Por fim são realizadas as consultas de forma que o usuário define uma especificação de consulta e o sistema retorna as imagens recuperadas.

Neste capítulo, serão introduzidos os conceitos relacionados as etapas principais dos sistemas CBIR, bem como os conceitos relacionados aos métodos de *re-ranking* aplicados a estes sistemas de consulta.

### 2.1.1 EXTRAÇÃO DE CARACTERÍSTICAS

Após sua aquisição, a imagem passa pela etapa de extração de características. Esta etapa tem como objetivo determinar uma medida numérica de uma determinada característica visual da imagem, podendo ela ser focada em pequenas regiões ou analisando a imagem como um todo.

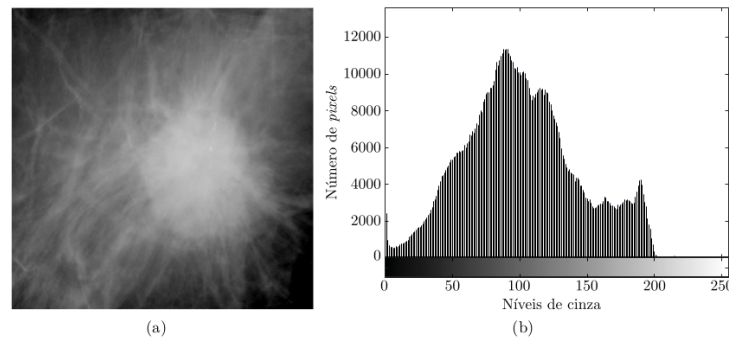
O processo de extração pode ser dividido em três grandes grupos dentro do conjunto de características chamadas primitivas: cor, forma e textura. Muitos sistemas de recuperação utilizam estas primitivas para representar uma imagem, assim sua recuperação está baseada na similaridade das características derivadas delas (EAKINS; GRAHAM, 1999).

#### 2.1.1.1 DESCRITORES DE COR

A cor é uma característica que está presente em todas as imagens. Dada uma imagem, seu histograma determina a distribuição de cores em uma certa região, definindo também seus níveis de intensidade. Imagens com distribuições de cores similares podem conter algum grau de similaridade (SWAIN; BALLARD, 1991).

Os extratores de cor tem um baixo custo computacional no seu processo de extração e baseiam-se geralmente em histogramas. Um histograma descrito em Swain e Ballard (1991) é obtido pela quantização do espaço de cores e pela contagem de pixels que cada cor quantizada possui em uma dada imagem. A Figura 2 apresenta um exemplo de histograma de cores de uma imagem de mamografia que foi quantizada com 256 níveis de cinza. A utilização de histogramas é vantajosa em termos computacionais, devido a seu baixo custo, nas suas propriedades de invariância a transformações de escala, rotação e translação nas imagens (SILVA, 2011). Segundo Silva (2011) existe uma desvantagem do histograma de cores: o fato dele não apresentar

informação sobre a distribuição espacial das cores.



**Figura 2: Histograma de Cores: (a) Imagem de mamografia em 256 níveis de cinza; (b) Histograma de níveis de cinza da imagem (a).**

**Fonte: Silva (2011)**

Um extrator muito utilizado também é o *Border/Interior Classification* (BIC), um método que considera a representação RGB na análise de cores, quantizadas em 64 níveis de intensidade e rotula os pixels da imagem em dois histogramas distintos, sendo um de borda e outro de interior. Em Ponti e Escobar (2013) é apresentada uma análise de desempenho e classificação utilizando o extrator BIC em várias resoluções.

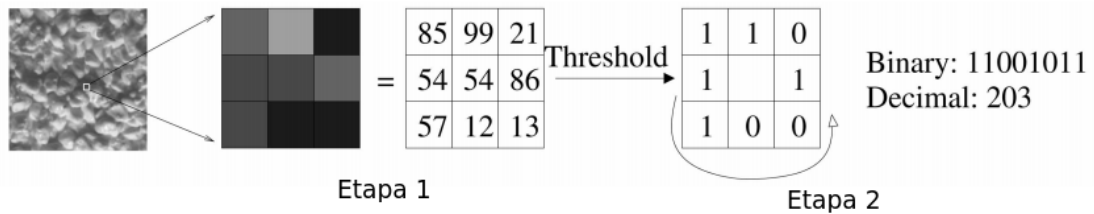
Na literatura existem outras propostas para o processo de extração de características baseadas em cor como o *Colour Space Selection*, *Quantization of Colour Space*, *Colour Descriptor Selection*, todos encontrados em Anantharatnasamy et al. (2013).

### 2.1.1.2 DESCRITORES DE TEXTURA

A textura também é um elemento importante no processo de extração de características. Através da análise deste atributo pode-se identificar padrões como uniformidade, densidade, aspereza, regularidade e intensidade, sendo estes muitos utilizados para distinguir as classes das imagens (PARIS, 2008).

O *Local Binary Pattern* (LBP) é um extrator de textura muito utilizado em reconhecimentos faciais. Geralmente este extrator utiliza em sua execução uma vizinhança-8, percorrendo a imagem e atribuindo valores 0 e 1 de forma que valores da vizinhança que possuem intensidade maior que o pixel central são definidos como 1 e os valores com intensidade menor que o pixel central são definidos como 0, gerando ao final um número binário que descreve o padrão local de distribuição de pixels da imagem (AHONEN et al., 2006).

A Figura 3 apresenta um operador básico LBP. Na primeira etapa deste operador, os níveis de cinza de uma região são quantizados e posteriormente, na etapa 2, são atribuídos os valores '1' para os pixels com um valor de intensidade maior que o valor de intensidade do pixel central e é atribuído o valor '0' para os pixels com valores de intensidade menores que o pixel central.



**Figura 3: Descrição de um Operador LBP**

**Fonte: Adaptado de Ahonen et al. (2006)**

Ainda se tratando de descritores de textura, Haralick et al. (1973) propôs 14 fórmulas numéricas para se descrever a textura de uma imagem. Estas fórmulas dão origem a uma matriz denominada de matriz de co-ocorrência (CENTENO; PACHECO, 2011).

A matriz de co-ocorrência considera a relação entre dois pixels sendo um chamado de pixel referencial e outro de pixel vizinho. O pixel vizinho pode estar em qualquer direção em relação ao pixel de referencial. Todos os pixels da imagem tornam-se pixel referenciais para a elaboração da matriz de co-ocorrência.

De posse da matriz de co-ocorrência, podem ser aplicados os descritores de Haralick et al. (1973). A seguir são apresentadas as principais equações utilizadas nos descritores de Haralick et al. (1973). Nestas equações,  $(i,j)$  são os índices da matriz de co-ocorrência e  $P(i, j)$  o valor do respectivo ponto na posição  $(i, j)$  desta matriz.

A Equação 1 apresenta o descritor relacionado a variância. Quando os níveis de cinza de uma imagem estão equalizados, o descritor apresentará uma variação local de pequena intensidade. Porém se os níveis de cinza estiverem com variação elevada, a imagem possuirá uma distribuição maior de níveis de cinza.

$$Variância = \sum_i \sum_j (i - j)^2 P(i, j) \quad (1)$$

Já a Equação 2 tem o intuito de calcular a entropia, que nada mais é que o grau de desordem ou dispersão dos níveis de cinza da imagem. Neste descritor, quando a entropia é baixa, os valores da matriz de co-ocorrência tendem a igualdade, diferentemente de quando a entropia é alta quando a textura não é uniforme.

$$Entropia = \sum_i \sum_j P(i-j) \log(P(i,j)) \quad (2)$$

Já a energia, expressa pela a Equação 3, mede a quantidade de repetições de pares de pixels para caracterizar a uniformidade da imagem. Assim, se os valores da matriz de co-ocorrência estiverem concentrados, a energia terá valores altos, caso contrário a energia terá um valor mais baixo.

$$Energia = \sum_i \sum_j P^2(i,j) \quad (3)$$

A Equação 4 representa a homogeneidade, que é sensível a valores próximos a diagonal principal da matriz de co-ocorrência. A homogeneidade é a medida inversamente relacionada à energia e o contraste da imagem.

$$Homogeneidade = \sum_i \sum_j \frac{1}{1+(i+j)^2} \quad (4)$$

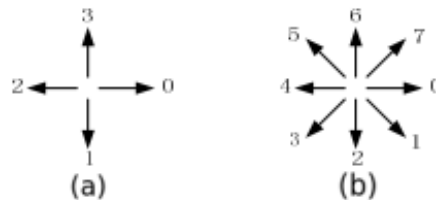
Por fim, a Equação 5 representa o contraste. O descritor de contraste mostra que uma pequena diferença entre as intensidades de pixels localizados em regiões próximas indicam a ocorrência de baixo contraste.

$$Contraste = \sum_i \sum_j (i-j)^2 P(i,j) \quad (5)$$

### 2.1.1.3 DESCRITORES DE FORMA

Quando comparado aos descritores citados anteriormente, um descritor baseado em forma é muito mais eficaz na caracterização do conteúdo de uma imagem, porém tem uma alta complexidade para extrair com exatidão a forma de uma imagem, segundo Anantharatnasamy et al. (2013). Mesmo com sua alta eficiência em relação aos outros descritores, em sistemas CBIR, este tipo de descritor apresenta um alto custo computacional, pois há a necessidade de segmentar e reconhecer o tamanho dos objetos contidos na imagem (PARIS, 2008).

*Chain Codes Method*, é um exemplo de um descritor baseado em forma, usado para representar o contorno de um objeto através de uma sequência de segmentos de retas com comprimentos e direções específicas. Tais representações baseiam-se em vizinhança-4, também podendo ser representados por vizinhança-8, sendo codificados usando a direção destes segmentos. A Figura 4 apresenta esta codificação dos segmentos usando vizinhança-4 e vizinhança-8.



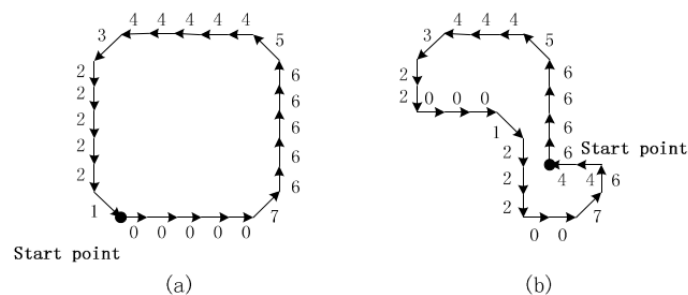
**Figura 4:** Números de direção para (a) código de cadeia de 4 direções; (b) código de cadeia para 8 direções.

**Fonte:** Junding e Heli (2009)

O código pode ser gerado seguindo a borda de um objeto, para isso, adota-se um pixel para ser o ponto inicial, um sentido e atribui-se uma direção para segmentos que unem os pares de pixels. A Figura 5 são apresentadas as formas para dois códigos, adotando-se sentido horário, e para cada código o ponto inicial é referente ao primeiro elemento do código. Para a representação das formas foram utilizados os dois códigos a seguir:

(a) 000007666665444443222221;

(b) 666654443220001222007644.



**Figura 5:** Contornos para os dois códigos (a) e (b)

**Fonte:** Junding e Heli (2009)

*Area of an Object, Horizontal and Vertical Distances*, são outros exemplos de métodos baseados em forma encontrados em Anantharatnasamy et al. (2013).

## 2.1.2 CONSULTAS POR SIMILARIDADE

Cálculos de similaridade são exclusivamente dependentes das características extraídas, e são fundamentais para o processo de consulta. Nestes casos, uma função de distância é definida para calcular o grau de similaridade, retornando um valor numérico que quantifica o quão similar estas imagens são (BUGATTI, 2008).

A grande maioria dos Sistemas de Gerenciamento de Banco de Dados (SGBDs) fazem suas consultas manipulando dados numéricos, alfanuméricos e textos curtos, baseando-se em operadores, sejam eles de igualdade ou de ordem. No entanto, este tipo de consulta não se torna viável para dados multimídia, pois estes são de natureza complexa. Devido a esta característica, a busca por similaridade torna-se a maneira mais adequada para a realização de consultas a esses tipos de dados. Neste tipo de consulta, é realizada a comparação de todos os objetos com o objeto de consulta, onde são retornados apenas os elementos que atendam um certo critério de similaridade definido anteriormente.

A seguir serão apresentadas os dois tipos de consultas mais utilizados: a consulta por abrangência e a consulta aos vizinhos mais próximos.

### 2.1.2.1 CONSULTA POR ABRANGÊNCIA

As Consultas por Abrangência (*range query* (RQ)), buscam objetos que tenham determinado nível de similaridade com o objeto de referência da busca. Para um dado conjunto de objetos  $O = \{o_1, o_2, o_3, \dots, o_n\}$ , pertencentes a um domínio  $Do$ , com uma função de distância  $d()$ , um objeto de consulta  $o_q$  também pertencente a este domínio e uma tem uma distância máxima  $r_q$ , sendo estes descritos pela Equação 6.

$$RQ(o_q, r_q) = B = \{b | b \in O, d(b, o_q) \leq r_q\} \quad (6)$$

A Figura 6 (a) apresenta a resposta de uma consulta por abrangência em um domínio bidimensional usando uma função de distância Euclidiana.

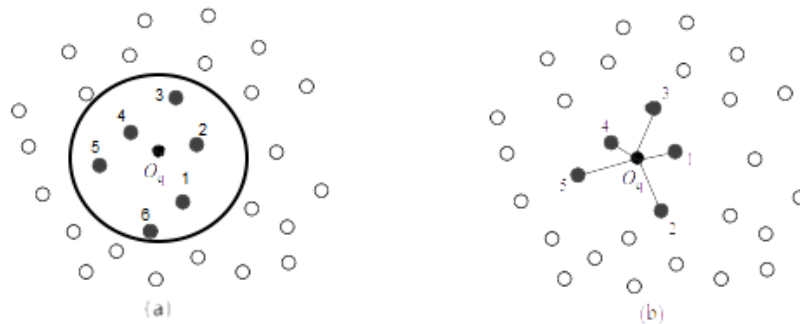
### 2.1.2.2 CONSULTA AOS VIZINHOS MAIS PRÓXIMOS

Nem sempre é possível identificar com precisão a distância máxima sem um prévio conhecimento da distribuição dos dados e das funções de distância. Ainda se tratando de consultas por abrangência, uma escolha inadequada pode retornar inúmeros elementos ou até mesmo ne-

nhum. A Consulta aos  $k$  Vizinhos Mais Próximos ( $k$  - *Nearest-Neighbor query*) consiste em buscar um número de  $k$  vizinhos em relação ao objeto de consulta. Para isso, dado um conjunto de objetos  $O = \{o_1, o_2, o_3, \dots, o_n\}$ , pertencentes a um domínio  $D0$ , com uma função de distância  $d()$ , um objeto de consulta  $o_q$  também pertencente a este domínio e um número inteiro  $k$ , terá seu resultado a partir da Equação 7.

$$kNN(o_q, k) = B = \{b | b \in O, \forall o \in O - B, d(o_q, b) \leq d(o_q, o), |B| = k\} \quad (7)$$

A Figura 6 (b) apresenta a resposta de uma consulta pelos  $k$ -vizinhos mais próximos (sendo  $k = 5$  vizinhos) em um domínio bidimensional usando uma função de distância Euclidiana, sendo que o elemento central  $O_q$  é o objeto de consulta, e os objetos destacados com a cor cinza são os objetos retornados após a consulta segundo um raio de abrangência  $r_q$  ou os  $k$  vizinhos mais próximos.



**Figura 6: Consulta por similaridade utilizando a função Euclidiana, onde (a) é uma consulta por abrangência e (b) representa uma consulta pelos 5-vizinhos mais próximos.**

**Fonte: Adaptado de Bugatti (2008)**

### Funções de Distância

A análise da similaridade de dados multidimensionais é conseguida através de funções de distâncias que são a chave para recuperação de imagens por conteúdo e algoritmos de agrupamento (ZHANG; LU, 2003). Existem diversas funções de distância relacionadas na literatura (WILSON; MARTINEZ, 1997). As mais utilizadas são as funções da família *Minkowski* ( $L_p$ ), sendo aplicadas a domínios vetoriais.

Os objetos que pertencem a um espaço vetorial de dimensão  $n$  ( $n$ -dimensional), são representados por  $n$  coordenadas de valores reais  $\{x_1, \dots, x_n\}$ . Sendo assim o conjunto de funções  $L_p$  definido como:

$$L_p((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p} \quad (8)$$

Quando o parâmetro  $p=1$ , obtemos a função  $L_1$ , também conhecida como *City Block* ou *Manhattan* ( $d_{L1}$ ) que pode ser expressa pela Equação 9.

$$d_{L1}(X, Y) = \sum_{i=1}^n |x_i - y_i| \quad (9)$$

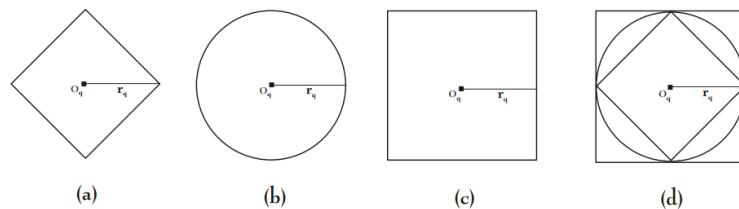
Já quando temos  $p=2$ , temos a função  $L_2$ , denominada de *Euclidiana* ( $d_{L2}$ ), sendo expressa pela Equação 10.

$$d_{L2}(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (10)$$

Aplicando o cálculo do Limite na Equação 8, quando  $p$  tende ao infinito, obtêm-se  $L_\infty$ , que é denominada de como *Infinity* ou *Chebychev*, definida formalmente pela Equação 11.

$$L_\infty((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_{i=1}^n |x_i - y_i| \quad (11)$$

Na Figura 7, são representadas as formas geométricas em um espaço bidimensional com as mesmas distâncias geradas pelas funções *Manhattan* Figura 7(a), *Euclidiana* Figura 7(b) e *Chebychev* Figura 7(c). A sobreposição das representações geométricas das funções da Figura 7(d) distingue os subespaços distintos as quais as diferentes funções de distância englobam.



**Figura 7: Representação das formas geométricas geradas conforme  $L_p$  utilizada e em relação aos pontos equidistantes à distância  $r_q$  a partir do objeto central  $o_q$ .**

**Fonte: Bugatti (2008)**

Em Vasconcelos e Lippman (2000) são apresentadas outras funções de avaliação como à distância de *Mahalanobis*, a divergência de *Kullback-Leibler*, a distância Quadrática e outras,



formulando o problema de recuperação como um problema de classificação.

### 2.1.3 AVALIAÇÃO DE DESEMPENHO NAS CONSULTAS POR SIMILARIDADE

Quando se fala em desempenho na recuperação de imagens baseadas em conteúdo, existem diversas maneiras que podem ser analisadas. Para o presente trabalho seram utilizadas duas técnicas, sendo elas a *Mean Average Precision*(MAP) e a *Normalized Discount Cumulative Gain*(NDCG) que serão apresentadas a seguir.

#### 2.1.3.1 MEAN AVERAGE PRECISION

A precisão na posição  $n$  ( $P_n$ ) mede a relevância dos  $n$  primeiros documentos retornados após determinada consulta. Seu cálculo é definido pela Equação 12.

$$P_n = \frac{\text{número de elementos relevantes nos } n \text{ primeiros resultados}}{n} \quad (12)$$

Por exemplo, se são retornados 5 elementos para uma determinada consulta (relevante, irrelevante, irrelevante, relevante, irrelevante), então a precisão da  $P_1$  até a  $P_5$  é dada por  $(1, \frac{1}{2}, \frac{1}{3}, \frac{2}{4}, \frac{2}{5})$  respectivamente. Assim para um conjunto de consultas, a média das  $P_n$  para cada consulta são calculadas. Já a precisão média para uma única consulta ( $AP$ ) é definida como a média dos valores de  $P_n$  para todos os documentos relevantes, representa a partir da Equação 13.

$$AP = \frac{\sum_{n=1}^N (P_n * rel(n))}{\text{total de elementos relevantes para a consulta}} \quad (13)$$

Nesta Equação,  $N$  representa o número de elementos recuperados e  $rel(n)$  é uma função que define se um elemento é ou não relevante para a consulta. Esta função é dada pela Equação 14.

$$rel(n) = \begin{cases} 1, & \text{se o documento na posição } n \text{ é relevante} \\ 0, & \text{outros} \end{cases} \quad (14)$$

Desta forma, a *Mean Average Precision* é dada pela média dos valores obtidos em  $AP$  para todas as consultas expressa pela Equação 15.

$$MAP = \frac{\sum AP(img_i)}{\text{número de consultas}} \quad (15)$$

### 2.1.3.2 NORMALIZED DISCOUNT CUMULATIVE GAIN

Ao se utilizar a  $Pn$  e a  $MAP$  é possível ter apenas duas opções de classificação: relevante e irrelevante. A Normalized Discount Cumulative Gain, proposta por Järvelin e Kekäläinen (2002), busca lidar com vários níveis de relevância na classificação. Para avaliação de um *ranking*, a NDCG segue dois princípios:

1. Documentos altamente relevantes tem peso maior que elementos menos relevantes;
2. Documentos nas últimas posições do *ranking* tem um peso menor, pois são menos suscetíveis de serem examinadas pelo usuário.

Seguindo estas regras, o valor da NDCG para um *ranking* até a posição  $n$  é calculado pela Equação 16.

$$NDCG = Z_n \sum_{j=1}^n \begin{cases} 2^{r(j)} - 1, & j = 1 \\ \frac{2^{r(j)-1}}{\log(j)}, & j > 1 \end{cases} \quad (16)$$

Na Equação 16,  $r(j)$  é a classificação do documento na posição  $j$  do *ranking* e  $Z_n$  é a constante de normalização de modo que para a NDCG seja igual a 1. Cabe ressaltar que para o cálculo da NDCG é necessário definir a classificação de cada documento, sendo que para o presente trabalho foi utilizado 0 para elementos irrelevantes e 1 para elementos relevantes.

## 2.2 RE-RANKING

Um *ranking* pode ser definido como a ordem das imagens retornadas após ser realizada uma determinada consulta. Quando uma consulta é executada, as imagens mais similares em relação a imagem de consulta são recuperadas e são retornadas em ordem, sendo a ordem de retorno das imagens denominada de *ranking*. Este processo agrega informações para a obtenção de melhorias na acurácia dentro do processo de recuperação.

O principal objetivo destes algoritmos é a classificação baseada na suposição que a informação contextual codificada nas listas de similaridade podem oferecer informações úteis para melhorar a efetividade nos descritores utilizados em sistemas CBIR (PEDRONETTE, 2012). Pelo fato do algoritmo de *re-ranking* não depender de distâncias e nem de *scores* de similaridades, eles podem ser usados para diferentes tarefas na recuperação da informação.

### 2.2.1 DEFINIÇÃO DE RE-RANKING EM IMAGENS

Dado um conjunto de imagens  $C = \{img_1, img_2, img_3, \dots, img_n\}$  e seja  $D$  um descritor, pode-se definir uma tupla  $(e, \rho)$ , onde:

- $e : \hat{I} \rightarrow \mathbb{R}^n$ , é uma função que extrai um vetor de características  $v_i \in \mathbb{R}^n$  em uma imagem  $\hat{I}$ .
- $\rho : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , a função de distância entre duas imagens calculada como a distância entre seus respectivos vetores de características.

Para obtenção do cálculo da distância entre duas as imagens  $img_i$  e  $img_j$ , é necessária a realização do cálculo de  $\rho(e(img_i, img_j))$ . Para simplificar será usada a notação  $\rho(img_i, img_j)$  para este cálculo.

A distância  $\rho(img_i, img_j)$  entre todas as imagens  $img_i, img_j \in C$ , são calculadas para obter-se uma matriz  $A$  de distância com dimensões  $N \times N$  onde  $A[i, j] = \rho(img_i, img_j)$ .

Para uma determinada imagem de consulta  $img_q$ , pode-se calcular um *Ranking*  $R_q$  em resposta a consulta baseado na matriz de distância  $A$ . O *Ranking* calculado pode ser definido como a permutação do conjunto de imagens  $C$ . A permutação  $\sigma_q$  é uma bijeção da coleção  $C$  em  $[N] = \{1, 2, \dots, N\}$  onde  $N$  é a cardinalidade de  $|C|$  na coleção de imagens  $C$ . Para a permutação  $\sigma_q$ , interpreta-se  $\sigma_q(i)$  como a posição da imagem  $img_i$  no *Ranking*  $R_q$ . Assim se  $img_x$  é rankeada antes de  $img_y$ , isto é,  $\sigma_q(x) < \sigma_q(y)$  então  $\rho(img_q, img_x) < \rho(img_q, img_y)$ .

Utilizando cada  $img_i \in C$  como sendo uma imagem de consulta  $img_q$ , é possível obter um conjunto conjunto de *rankings*  $\mathcal{R} = \{R_1, R_2, \dots, R_n\}$  para cada imagem da coleção  $C$ .

O algoritmo de *Re-Ranking* de uma imagem tem como entrada a matriz  $A$  e o conjunto de *rankings*  $\mathcal{R}$  para calcular uma nova e mais efetiva matriz de distância  $\hat{A}$ . O algoritmo de *Re-Ranking* considera as relações entre toda coleção de imagens para representar uma função  $f_r$ , apresentada na Equação 17.

$$\hat{A} = f_r(A, \mathcal{R}) \quad (17)$$

Assim, para a nova matriz de distância  $\hat{A}$ , um novo conjunto  $\hat{\mathcal{R}}$  é obtido, contendo as novas posições rankeadas para toda a coleção de imagens, ou seja, as imagens são re-ranqueadas. Cabe ressaltar que a função  $f_r$  explora todas as relações entre  $A$  e  $\mathcal{R}$ .

## 2.2.2 PRINCIPAIS ABORDAGENS UTILIZADAS

As abordagens utilizadas para o aprendizado das medidas de distância podem ser divididas em três categorias dependendo do tipo de treinamento. São elas a abordagem supervisionada, semi-supervisionada e não-supervisionadas.

### 2.2.2.1 ABORDAGENS SUPERVISIONADAS

No aprendizado supervisionado, as amostras de treinamento são usadas para inferir uma função. Cada amostra treinada é constituída de um par contendo um objeto de entrada e um valor de saída esperado. Em Zhu e Goldberg (2009) define-se  $\mathcal{X}$  como o domínio dos objetos e  $\mathcal{Y}$  como o domínio dos rótulos. Dado  $Pr(x, y)$ , a distribuição de probabilidade conjunta de objetos e rótulos é dada por  $\mathcal{X} \times \mathcal{Y}$ . Dada uma amostra de treinamento  $\{(x_i, y_i)\}_{i=1}^n$ , sendo aprendizado supervisionado, este treina a função  $f: \mathcal{X} \rightarrow \mathcal{Y}$ , como um objeto de  $f(x)$ , a qual preve um rótulo verdadeiro em futuros dados de  $x$ .

No aprendizado supervisionado as medidas de distância abordam tentativas de aprender métricas que mantêm todos os dados de uma mesma classe agrupados e agrupa todos os dados diferentes em classes distintas. No aprendizado de métricas de distância a informação rotulada é especificada geralmente em pares restritos de dados: (1) restrições equivalentes, isto para dado par semanticamente similares; (2) restrições não equivalentes, isto para dado par semanticamente não similar.

Uma abordagem baseada em *learning-to-rank* é apresentada em Faria et al. (2010). Nela são usados uma combinação de diferentes descritores a fim de melhorar o desempenho do *ranking*, e um conjunto de imagens é fornecido como entrada para o aprendizado dos algoritmos. Além disso, para cada imagem consultada, um conjunto de imagens de amostras é fornecido. A relevância de uma imagem com a imagem de consulta também é informada como entrada. Estas informações são usadas para o treinamento, para isso os algoritmos de aprendizagem produzem funções de *ranking* mapeando o nível de similaridade da coleção de imagens para uma imagem de consulta definida. Em uma nova consulta, dada uma nova imagem, a relevância das imagens retornas é estimada de acordo com a função de aprendizado, usando algoritmos supervisionados.

### 2.2.2.2 ABORDAGENS SEMI-SUPERVISIONADAS

A abordagem semi-supervisionada está entre as abordagens supervisionada e não-supervisionada. O aprendizado semi-supervisionado se refere ao uso de dados rotulados e não

rotulados para o treinamento. O treinamento dos dados consiste em  $l$  instâncias rotuladas  $\{(x_i, y_i)\}_{i=1}^l$  e  $u$  instâncias não rotuladas  $\{x_j\}_{j=l+1}^{l+u}$ . Enquanto o rótulo do dado  $(x, y)$  é difícil de se identificar, os dados não rotulados de  $x$  estão disponíveis em grande escala.

Abordagens semi-supervisionadas consideram a hipótese antes de sua ocorrência, isto é, pontos próximos provavelmente tem o mesmo rótulo, e também pontos em uma mesma estrutura provavelmente tem os mesmos rótulos.

Em Hoi et al. (2008), uma métrica de distância semi supervisionada apresenta uma abordagem com foco em aplicações em sistemas CBIR. Nele o conjunto de dados utilizados para o treinamento são explorados pelo histórico de realimentação de relevância dos dados. Esta abordagem utilizando o histórico de dados utilizados para a recuperação é definido como *Collaborative Image Retrieval*.

### 2.2.2.3 ABORDAGENS NÃO SUPERVISIONADAS

Devido ao grande custo de rotular os objetos manualmente, considerando também que os dados não rotulados são mais acessíveis, o aprendizado não supervisionado torna-se atraente em diferentes cenários.

*Clustering* é um exemplo clássico de aprendizado não supervisionado. De maneira geral, *clustering* organizam os objetos em grupos (*clusters*), deste modo, objetos de um mesmo grupo são mais similares entre si do que comparado a grupos diferentes. Em Gupta et al. (2008) um método de abordagem não supervisionada baseada em *clustering* é proposto, realizando a filtragem e agrupamento de dados.

Abordagens *clustering* são utilizadas em métodos de *re-ranking* com intuito de explorar informações contextuais em sistemas CBIR. Em Park et al. (2005) um algoritmo de *re-ranking* usando *clustering* para sistemas CBIR é proposto. Nele, primeiramente as imagens são recuperadas utilizando recursos visuais. Depois, são agrupadas utilizando métodos de *clustering* e a classificação dos resultados é ajustada de acordo com a distância de um grupo para a imagem de consulta.

## 2.3 ALGORITMO DE RE-RANKING

Nesta seção serão apresentados alguns conceitos para a implementação do algoritmo de *Re-Ranking*, sendo elas:

- Abordagem de aprendizado não supervisionada baseada na similaridade de *ranked-lists*;

- Métodos para comparação de *ranked-lists*;
- E por fim, o algoritmo *RL-Sim Re-Ranking* proposto por Pedronette (2012).

### 2.3.1 MEDIDAS DE DISTÂNCIA CONTEXTUAIS BASEADAS EM RANKED-LISTS

As medidas de distância contextuais são a base para a implementação do algoritmo de *Re-Ranking*, pois para uma dada imagem, e um descritor  $D$  pode-se calcular a distância  $\rho(img_i, img_j)$  entre duas imagens  $img_i, img_j \in C$ , assim esta medida de distância considera apenas duas imagens (PEDRONETTE, 2012).

Para calcular os *ranked-lists*  $R_i, R_j$  para as imagens  $img_i, img_j$  as distâncias entre a imagem e toda a base de imagens devem ser calculadas. Desta forma, os *ranked-lists* representam uma descrição contextual das imagens em relação a toda a base. As imagens nas primeiras posições dos *ranked-lists* apresentam as imagens mais relevantes em relação a busca. Sendo assim, dadas duas imagens e seus *ranked-lists* respectivos, pode-se calcular uma nova e eficiente medida de distância considerando apenas as imagens nas primeiras posições dos *ranked-lists*.

De maneira geral, as primeiras posições dos *ranked-lists* apresentam as imagens mais similares em relação a imagem de consulta, mas algumas imagens não similares também podem ser retornadas. Estas imagens muitas vezes são semelhantes entre si, havendo assim muitas imagens em comum no *ranked-list* respectivo de cada imagem, fazendo com que este conjunto de imagens apareça nos *ranked-lists* de todas as imagens que compõe este conjunto. Este comportamento não se repete quando se comparam as primeiras posições das listas de imagens não similares. Por este motivo, a medida de distância contextual recebe um valor menor quando ela é calculada, pois existem poucas imagens em comum nas primeiras posições das listas de imagens não similares. Neste contexto o objetivo do algoritmo de *re-ranking* é mover as imagens menos similares para o final dos *ranked-lists*, resultando *ranked-lists* melhorados (PEDRONETTE, 2012).

Uma nova proposta de medida de distância contextual, usando aprendizado não supervisionado é proposta em Pedronette (2012). Esta medida de distância é incorporada às informações contextuais, a fim de melhorar a recuperação dos resultados. Considerando uma vizinhança  $\mathcal{N}(i)$  em uma imagem  $img_i$  que contém imagens semelhantes a ela de acordo com uma distância  $\rho$  definida por um descritor. O conjunto  $\mathcal{N}(i)$  pode ser obtido usando a abordagem dos k-vizinhos mais próximos, onde a cardinalidade do conjunto é denotada por  $|\mathcal{N}(i)| = k$ .

Seguindo esta abordagem, as primeiras posições dos *ranked-lists* são chamadas de

*top k lists*. Define-se um *ranked-list*  $R_i$  como a permutação da coleção  $C$ , dada por uma bijeção  $\sigma_i$  para uma coleção  $C$  no conjunto  $[N] = \{1, 2, \dots, N\}$ . Da mesma forma, uma *top k list*  $\tau_i$  é uma bijeção para o domínio  $\mathcal{N}(i)$  para  $[k] = \{1, 2, \dots, k\}$ . Neste contexto, uma imagem  $img_j$  aparece em uma *top k list*  $\tau_i$  se a  $img_j \in \mathcal{N}(i)$ . Assim  $\tau_i(j)$  é a posição (classificação) da imagem  $img_j$  em  $\tau_i$ . Para a abordagem dos  $k$ -vizinhos mais próximos, uma imagem  $img_1$  é ranqueada antes de uma imagem  $img_2$  ( $\tau_i(1) < \tau_i(2)$ ), então  $\rho(img_i, img_1) \leq \rho(img_i, img_2)$ .

Utilizando  $\tau_i$  e  $\tau_j$  sendo as *top k lists* calculadas para as imagens  $img_i$ ,  $img_j$  respectivamente, as medidas de similaridade podem ser definidas. Sendo  $d(\tau_i, \tau_j, k)$  que determina a medida de distância que compara as *top k lists*, a medida de distância contextual não iterativa  $\rho_c(img_i, img_j)$  baseada na comparação das *top k lists* será dada pela 18.

$$\rho_c(img_i, img_j) = d(\tau_i, \tau_j, k) \quad (18)$$

Supondo que a medida de distância contextual  $\rho_c$  representa uma distância mais efetiva entre as imagens, é possível calcular a distância entre todas as imagens da base. Desta forma um novo conjunto de *ranked-lists* pode ser obtido, de modo que a distância contextual pode ser recalculada. Este processo pode ser repetido de maneira iterativa, onde  $(t)$  é o expoente que indica a iteração,  $\tau_i^{(t)}$  é uma *top k list* de uma imagem  $img_i$  na iteração  $t$ , e  $\rho_c^{(t)}$  é a medida de distância contextual para a iteração  $t$ . Desta forma a Equação 19 define esta medida contextual.

$$\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k) \quad (19)$$

Desta maneira, a medida de distância contextual melhora ao longo das iterações, melhorando assim também a eficácia dos *ranked-lists* (PEDRONETTE, 2012). As imagens não relevantes são removidas das primeiras posições dos *ranked-lists* e,  $k$  é incrementado para considerar mais imagens. Desta forma, um alto valor de  $k$  é considerado para calcular a *top k list* ao longo das iterações, sendo estes expressos pela Equação 20.

$$\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k+t) \quad (20)$$

Depois de um número de  $T$  iterações, uma nova distância  $\hat{\rho}$  é calculada baseada na distância contextual  $\rho_c$ , dada pela Equação 21.

$$\hat{\rho}(img_i, img_j) = \rho_c^{(T)}(img_i, img_j) \quad (21)$$

Ao final destas etapas, uma nova matriz de distância  $\hat{A}$  é calculada baseada em  $\hat{\rho}$  tal que para todas as imagens  $img_i, img_j \in C$  obtendo  $\hat{A}_{ij} = \hat{\rho}(img_i, img_j)$ . Baseado em  $\hat{A}$ , um novo conjunto de *ranked-lists*  $\hat{\mathcal{R}}$  é calculado completando o processo de *Re-Ranking*.

### 2.3.2 COMPARAÇÃO NOS RANKED-LISTS

A comparação entre *ranked-lists* é baseada na proposta de medidas de distância contextuais. Nesta etapa, os processos de comparação serão divididos em duas partes, sendo eles: (a) recuperação do conjunto de vizinhos para uma imagem  $img_i$ , que será utilizado para a composição da *top k list*  $\tau_i$ ; (b) cálculo da distância  $d(\tau_i, \tau_j, k)$  entre as *top k lists*. Estes processos (a) e (b) serão descritos a seguir na Seção 2.3.2.1 e na Seção 2.3.2.2 respectivamente.

#### 2.3.2.1 CONJUNTO DE VIZINHOS

Nesta Seção serão apresentados duas abordagens para o cálculo das *top k lists* para uma dada imagem: o Método dos k-Vizinhos mais Próximos (kNN), e o método *Mutual k-Vizinhos mais Próximos* (MkNN).

##### **k-Vizinhos Mais Próximos**

Dada uma imagem  $img_i \in C$ , o conjunto de  $k$  imagens mais similares serão selecionados. Dado  $\mathcal{N}_{kNN}(i)$  o conjunto de vizinhos é obtido usando o método dos k-vizinhos mais próximos, definidos pela Equação 22.

$$\mathcal{N}_{kNN}(i, k) = \{\mathcal{R} \subseteq C, |\mathcal{R}| = k \wedge \forall x \in \mathcal{R}, y \in C - \mathcal{R} : \rho(i, x) \leq \rho(i, y)\} \quad (22)$$

Assim baseado no conjunto de vizinhos  $\mathcal{N}_{kNN}(i)$  pode-se definir a *top k list*  $\tau_{i_{kNN}}$  usando os k-vizinhos mais próximos. Neste contexto, considerando que o valor de uma medida de distância não se repete, a distância entre uma  $img_i$  e as imagens do conjunto de vizinhos  $\mathcal{N}_{kNN}(i, k)$ , tal que  $\{\forall x, y \in \mathcal{N}_{kNN}(i, k) : \rho(i, x) = \rho(i, y)\} = \emptyset$ . A *top k list*  $\tau_{i_{kNN}}(j)$  é a permutação de  $C$ , considerando a bijeção  $\tau_{i_{kNN}} : C \rightarrow [1, \dots, k]$  dada pela Equação 23.

$$\tau_{i_{kNN}}(j) = |\{j \in \mathcal{N}_{kNN}(i, k), \forall x \in \mathcal{N}_{kNN}(i, k) : \rho(i, x) < \rho(i, j)\}| + 1 \quad (23)$$

##### **Mutual k-Vizinhos Mais Próximos**

Dado  $\tau_{i_{kNN}}(j)$  sendo a posição da imagem  $img_j$  em uma *top k list*  $\tau_{i_{kNN}}$  e dados  $\tau_{j_{kNN}}(i)$



para a posição da imagem  $img_i$  em uma *top k list*, é comum que em sistemas CBIR que  $\tau_{i_{kNN}}(j) \neq \tau_{j_{kNN}}(i)$  (PEDRONETTE, 2012). Contudo, quando a diferença entre essas posições é grande, isto pode indicar uma posição incorreta em uma imagem desta *top k list*. Define-se como *Mutual k-Vizinhos mais Próximos* o método que considera a recíproca das posições das imagens em um *ranked-list* (PEDRONETTE, 2012). De forma geral, seleciona-se os k-vizinhos mais próximos considerando  $c \times k$  vizinhos, sendo  $c$  uma constante. Dado um conjunto de vizinhança  $\mathcal{N}_{kNN}(i, c \times k)$ , seleciona-se os  $k$  itens mais similares no conjunto tomando: (a) a posição das imagens no *ranked-list* da a imagem  $img_i$  e ; (b) a posição da imagem  $img_i$  nos *ranked-lists* destas imagens. A Equação 24 e a Equação 25 expressam respectivamente as definições de (a) e (b) para o método *Mutual k-vizinhos mais próximos*(  $\mathcal{N}_{MkNN}(i, k)$ ).

$$\mathcal{N}_{MkNN}(i, k) = \{\mathcal{R} \subseteq \mathcal{N}_{kNN}(i, c \times k), |\mathcal{R}| = k \wedge \forall x \in \mathcal{R}, y \in C - \mathcal{R} : \tau_{i_{kNN}}(x) + \tau_{x_{kNN}}(i) \leq \tau_{i_{kNN}}(y) + \tau_{y_{kNN}}(i)\} \quad (24)$$

A Equação 25 é definida para uma *top k list*  $\tau_{i_{MkNN}}$  usando o método *Mutual k-vizinhos mais próximos*.

$$\tau_{i_{MkNN}}(j) = |\{j \in \mathcal{N}_{MkNN}(i, k), \forall x \in \mathcal{N}_{MkNN}(i, k) : \tau_{i_{kNN}}(x) + \tau_{x_{kNN}}(i) \leq \tau_{i_{kNN}}(y) + \tau_{y_{kNN}}(i)\}| + 1 \quad (25)$$

### 2.3.2.2 MEDIDAS DE DISTÂNCIA ENTRE TOP K LISTS

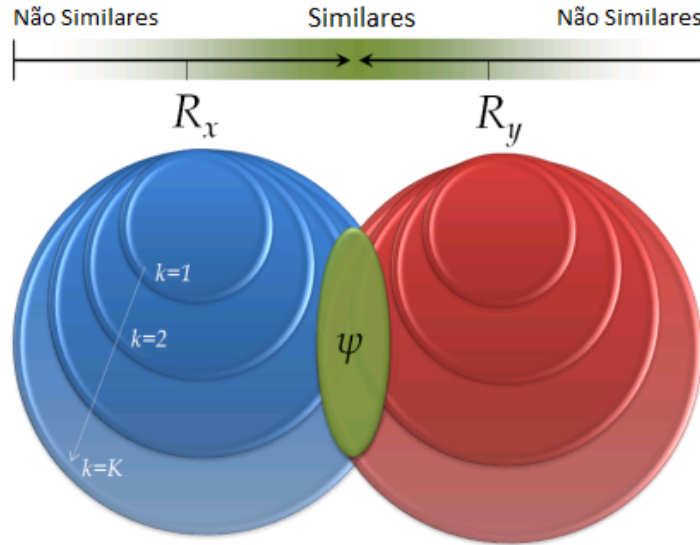
A presente Seção apresenta duas abordagens utilizadas para o cálculo das medidas de distância entre as *top k lists* sendo elas: (a) Medidas de Intersecção e; (b) Medidas de *Kendall*, apresentadas a seguir.

#### **Medidas de Intersecção**

Fagin et al. (2003) define a medida de distância entre duas *top k lists*  $\tau_i$  e  $\tau_j$  como sendo a superposição cumulativa incrementando-se a profundidade. Assim, para cada  $k_c \in \{1, \dots, k\}$  é calculada a superposição de  $k_c$ , e então essas superposições são usadas para calcular a medida de similaridade. A medida atribui valores mais elevados para as primeiras posições da *top k list*. A Equação 26 define a intersecção da medida de similaridade  $\psi$ .

$$\psi(\tau_i, \tau_j, k) = \frac{\sum_{k_c=1}^k |\mathcal{N}(i, k_c) \cap \mathcal{N}(j, k_c)|}{k} \quad (26)$$

Quando dois *ranked-lists* apresentam as mesmas imagens nas primeiras posições, o tamanho da intersecção do conjunto é maior, e o valor de  $\psi$  também é maior. A Figura 8 apresenta o cálculo de  $\psi$  considerando múltiplos valores de  $k$ .



**Figura 8:** Cálculo da medida  $\psi$ : intersecção entre *ranked-lists* com diferentes tamanhos.

Fonte: Adaptado de Pedronette e Torres (2013)

Assim, a medida de distância entre as *top k lists* pode ser definida pela Equação 27.

$$d_\psi(\tau_i, \tau_j, k) = \frac{1}{1 + \psi(\tau_i, \tau_j, k)} \quad (27)$$

### Medida Tau de Kendall

A medida de *Kendall* é uma medida que calcula o grau de semelhança entre dois *rankings* dados um mesmo conjunto de objetos. Seu valor é igual ao número de trocas necessárias no *bubble sort* para converter uma permutação em outra (FAGIN et al., 2003). A medida de *Kendall* ( $d_\tau(\tau_i, \tau_j, k)$ ) normalizada é definida a seguir na Equação 28.

$$d_\tau(\tau_i, \tau_j, k) = \frac{\sum_{x,y \in \mathcal{N}(i,k) \cup \mathcal{N}(j,k)} \bar{K}_{x,y}(\tau_i, \tau_j)}{k \times (k-1)} \quad (28)$$

Neste contexto,  $\bar{K}_{x,y}(\tau_i, \tau_j)$  é a função que determina se as imagens  $img_x$  e  $img_y$  estão

em ordem comparando os *ranked-lists*  $R_i$  e  $R_j$ . Esta função pode ser definida pela Equação 29, onde o valor máximo definido para a medida *Kendall* é dado por  $k \times (k - 1)$ , ocorrendo quando  $\mathcal{N}(i, k) \cap \mathcal{N}(j, k) = \emptyset$  e  $\sigma_i$  é o inverso de  $\sigma_j$ .

$$\bar{K}_{x,y}(\tau_i, \tau_j) = \begin{cases} 0 & \text{se } (\sigma_i(x) \leq \sigma_i(y) \wedge \sigma_j(x) \leq \sigma_j(y)) \vee (\sigma_i(x) \geq \sigma_i(y) \wedge \sigma_j(x) \geq \sigma_j(y)) \\ 1 & \text{caso contrário} \end{cases} \quad (29)$$

### 2.3.3 O ALGORITMO RL-SIM RE-RANKING

O objetivo do algoritmo proposto por Pedronette (2012) é explorar um conjunto de *ranked-lists*  $\mathcal{R} = \{R_1, R_2, \dots, R_N\}$  para calcular uma matriz de distância  $\hat{A}$  mais efetiva e assim, gerar um conjunto mais efetivo de *ranked-lists*  $\hat{\mathcal{R}}$ . O algoritmo *RL-Sim Re-Ranking* é baseado na medida contextual apresentada em  $\rho_c$ , que leva em conta a similaridade entre os *ranked-lists* de forma iterativa (PEDRONETTE, 2012).

Dado um expoente  $(t)$  sendo a iteração atual, um novo e mais eficiente conjunto de *ranked-lists*  $\mathcal{R}^{(t+1)}$  é calculado considerando as distâncias entre as *top k lists*. Este novo conjunto de *ranked-lists* é usado na próxima execução e se repete ao longo de várias iterações aumentando a eficiência a cada iteração. Depois de um número  $T$  de iterações, um *re-ranking* é feito baseado na matriz final de distância  $\hat{A}$  e baseado nesta matriz, o conjunto de *ranked-lists*  $\hat{\mathcal{R}}$  é calculado.

No Algoritmo 1, as distâncias são redefinidas considerando a função  $d(\tau_i, \tau_j, k)$  para as primeiras  $\lambda$  posições de cada *ranked-list*, de forma que  $\lambda \in \mathbb{N}$  e  $0 \leq \lambda \leq N$ . Um dos parâmetros necessários para execução do algoritmo é  $k_s$ , que representa o número de *k-vizinhos* utilizado para a primeira execução do algoritmo. Para as imagens nas posições restantes dos *ranked-lists* uma nova medida é redefinida (Linha 12) basadas na distância atual. Nestes casos, a função  $d(\tau_i, \tau_j, k)$  não necessita ser calculada, considerando que as imagens relevantes devem estar no início dos *ranked-lists*. Desta forma, o cálculo computacional diminui, sendo que nesta etapa o algoritmo não depende do tamanho  $N$  da base de imagens.

Já na Linha 18, a cada iteração  $t$ , é incrementado o número dos  $k$  vizinhos considerados. O motivo inicial desta incrementação é baseada no fato de que a eficácia dos *ranked-lists* aumenta ao longo das iterações. Desta forma, as imagens não relevantes são movidas para as últimas posições dos *ranked-lists* e  $k$  pode ser incrementado para considerar mais imagens.

---

**Algorithm 1:** Algoritmo RL-Sim Re-Ranking
 

---

**Input:** Conjunto original de *Ranked-Lists*  $\mathcal{R}$  e os parâmetros  $k_s, T, \lambda$

**Output:** Conjunto de *Ranked-Lists* melhorado  $\hat{\mathcal{R}}$

**begin**

$t \leftarrow 0$  ;

$\mathcal{R}^{(t)} \leftarrow \mathcal{R}$ ;

$A^{(t)} \leftarrow A$ ;

$k \leftarrow k_s$ ;

**while**  $t < T$  **do**

**forall the**  $R_i \in \mathcal{R}^{(t)}$  **do**

$c \leftarrow 0$ ;

**forall the**  $img_j \in R_i$  **do**

**if**  $c \leq \lambda$  **then**

$A^{(t+1)}[i, j] \leftarrow d(\tau_i, \tau_j, k)$ ;

**else**

$A^{(t+1)}[i, j] \leftarrow 1 + A^{(t)}[i, j]$ ;

**end**

$c \leftarrow c + 1$  ;

**end**

**end**

$\mathcal{R}^{(t+1)} \leftarrow performReRanking(A^{(t+1)})$ ;

$k \leftarrow k + 1$ ;

$t \leftarrow t + 1$ ;

**end**

$\hat{\mathcal{R}} \leftarrow \mathcal{R}^{(T)}$ ;

**end**

---

O algoritmo de *re-ranking* não depende especificamente das medidas entre as *top k lists* (PEDRONETTE; TORRES, 2013). Desta forma uma grande vantagem do algoritmo é a possibilidade de serem utilizadas diferentes abordagens para a recuperação do conjunto dos vizinhos (tais métodos foram apresentados na Seção 2.3.2.1) e diferentes medidas para comparação das *top k lists* (apresentadas na Seção 2.3.2.2). Este algoritmo pode ser facilmente estendido para considerar diferentes e complexas abordagens para o cálculo de similaridade entre as *top k lists*.

## 2.4 CONSIDERAÇÕES FINAIS

Neste Capítulo foram abordados os principais conceitos utilizados no processo de recuperação de imagens baseadas por seu conteúdo. Foram apresentados extratores de característica, funções de distância, bem como abordagens para avaliar o desempenho das consultas, tais conceitos são relacionados aos sistemas CBIR. Além destes, foram apresentadas também as principais abordagens utilizadas em algoritmos de *re-ranking* além de métodos para comparação de *ranked-lists*, e, por fim, A seguir será apresentado o plano de trabalho para o desenvolvimento desta proposta, bem como a metodologia a ser utilizada, os resultados esperados e o cronograma de atividades.

### 3 METODOLOGIA DE DESENVOLVIMENTO

Ao longo do presente trabalho, foram apresentadas diferentes abordagens utilizadas na recuperação de imagens baseadas em conteúdo, desde os conceitos iniciais relacionados aos sistemas CBIR, passando pelo processo de extração de característica realizado pelos descritores, pelas consultas por similaridade e funções de distância. Além disso foram apresentadas as principais abordagens de aprendizado de máquina mais utilizadas e por fim o algoritmo de *re-ranking*.

Vários trabalhos vem sendo propostos utilizando sistemas CBIR, mas nem todos abordam a utilização de um algoritmo de *re-ranking* nos resultados. A proposta deste trabalho visa o estudo e análise de diferentes descritores de imagens para posterior implementação. Após este processo de extração, estes descritores serão combinados e passarão por um algoritmo de *re-ranking* buscando melhorar a precisão no processo de recuperação das imagens.

#### 3.1 BASE DE IMAGENS

Para a realização do presente trabalho, foi utilizada uma base de imagens pública, a *Amsterdam Library of Object Images* (ALOI), uma base de imagens registrada para fins científicos com 1000 objetos, com possibilidades de variação do ângulo de iluminação, variação da cor de iluminação, rotação do ângulo de visualização para cada objeto, entre outras possibilidades. Nesta base, para cada objeto foram registradas mais de 100 imagens por objeto, totalizando 110.250 imagens para a base (GEUSEBROEK et al., 2005).

Uma das principais vantagens de se utilizar esta base foi o fato de não ser necessário realizar o download de toda a base, pois devido ao grande número de imagens e das diferentes possibilidades de variação de cada objeto a base em sua totalidade contém 55GB de dados. Com a utilização desta base, foi possível analisar diferentes cenários em diferentes resoluções que posteriormente serão detalhados.

Inicialmente, foi utilizada a base *Quarter Resolution*, sendo que esta contém imagens

com resolução reduzida em um quarto em relação a imagem original. Ainda nesta base, foram utilizadas as imagens disponíveis em *Illumination Direction* e *Illumination Color*.

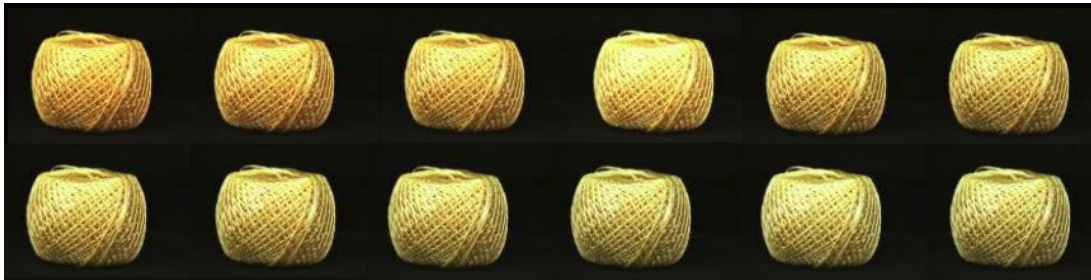
A Figura 9 apresenta as imagens pertencentes a uma das classes da base *Illumination Direction*. A característica principal desta base é a variação no ângulo de iluminação do objeto, sendo que para cada classe existem 24 imagens.



**Figura 9:** Imagens de uma classe na base *Illumination Direction*

Fonte: Geusebroek et al. (2005)

A Figura 10 apresenta uma classe da base *Illumination Color*, sendo que as principais características desta base são: 12 imagens por classe, que variam a iluminação do objeto indo de “tonalidades quentes” a “tonalidades frias” buscando um efeito de temperatura.



**Figura 10:** Imagens de uma classe na base *Illumination Color*

Fonte: Geusebroek et al. (2005)

Para a fase dos experimentos foram utilizadas 30 classes escolhidas aleatoriamente, sendo estas apresentadas na Figura 11. As 30 classes escolhidas foram utilizadas tanto na base *Illumination Color* quanto para a base *Illumination Direction*.



**Figura 11: Classes utilizadas nos experimentos**

**Fonte: Adaptado de Geusebroek et al. (2005)**

### 3.2 EXTRAÇÃO DE CARACTERÍSTICAS, FUNÇÕES DE DISTÂNCIA E MEDIDAS DE COMPARAÇÃO

Nesta etapa foram utilizados os descritores BIC, LBP e Haralick, todos implementados como foram descritos anteriormente. Ao final do processo de extração de características foram utilizadas duas diferentes funções de distância, a dLog e a distância Euclidiana, todas descritas anteriormente. Calculada a distância entre as imagens, o *ranking* para cada imagem foi gerado. De posse deste *ranking*, foram utilizadas duas diferentes medidas de comparação entre as *top k lists*, sendo elas a Medida de Intersecção e tal de *Kendall*. Cabe ressaltar que para a geração destas *top k lists* foi utilizado o método knn com sua devida função de distância para a recuperação dos k vizinhos mais próximos.

### 3.3 CENÁRIOS DE APLICAÇÃO

Devido a utilização de diferentes descritores, diferentes funções de distância e medidas de comparação, foi possível realizar 24 estudos de caso diferentes, sendo que estes serão apresentados na Seção 4. Para cada estudo de caso foram utilizados até 35 diferentes parâmetros para a execução do algoritmo. No Capítulo 4 serão apresentados os resultados iniciais referentes a cada estudo de caso e posteriormente os melhores casos serão mais detalhados.

### 3.4 AVALIAÇÃO DOS RESULTADOS

Para a avaliação dos resultados foram utilizados os seguintes cálculos:

- Precisão e NDCG média para cada estudo de caso;
- Precisão, desvio padrão da precisão, e NDCG média por classe;



- NDCG para as 10 primeiras posições.

Como foi descrito anteriormente, a NDCG mede a qualidade de um *ranking*, e para avaliação dos resultados ela foi utilizada da seguinte forma:

1. Dado um *ranking* de consulta  $Rq$ , a NDCG será calculada com base no número de imagens por classe ou com base em um valor fixo. Inicialmente, é realizado o cálculo do DCG relativo ao *ranking*, atribuindo-se peso 1 as imagens pertencentes a classe da imagem de consulta, e peso 0 para as demais.
2. Em seguida o *ranking* é ordenado da melhor forma possível, isto é, as imagens pertencentes a classe da imagem de consulta serão as primeiras imagens do *ranking* e terão peso 1, e as outras irão para o final do *ranking* e terão peso 0. Após esta ordenação é calculado o IDCG que é o DCG ideal para este *ranking*.
3. Finalizados estes cálculos, é realizado o cálculo da NDCG, que nada mais é que a razão entre o DCG e o IDCG.

### 3.5 PIPELINE DE EXECUÇÃO

Visando o aproveitamento de dados, alta complexidade dos algoritmos e das diferentes possibilidades de execução, o trabalho foi dividido em algumas partes, apresentadas na Figura 12.

1. Inicialmente, foi definida a base, gerando uma lista com as imagens a serem utilizadas;
2. De posse desta lista, foi realizada a extração de características, utilizando BIC, LBP e Haralick. Para cada extrator, foi gerado um conjunto de vetores com as características das imagens;
3. Para o *ranking* das imagens, foi necessário selecionar um dos conjunto de vetores e uma função de distância (dLog ou Euclidiana);
4. O processo de *Re-Ranking* teve como base a lista de *rankings* para a sua execução. Nesta etapa também foi necessária a escolha de uma medida de distância (*Kendall*, Intersecção) para a comparação entre as *top k lists* e os parâmetros de entrada do algoritmo. Outro conceito utilizado no algoritmo foi a busca dos vizinhos utilizando o método kNN, também descrito anteriormente. Após a execução deste, foi realizada a análise dos resultados.

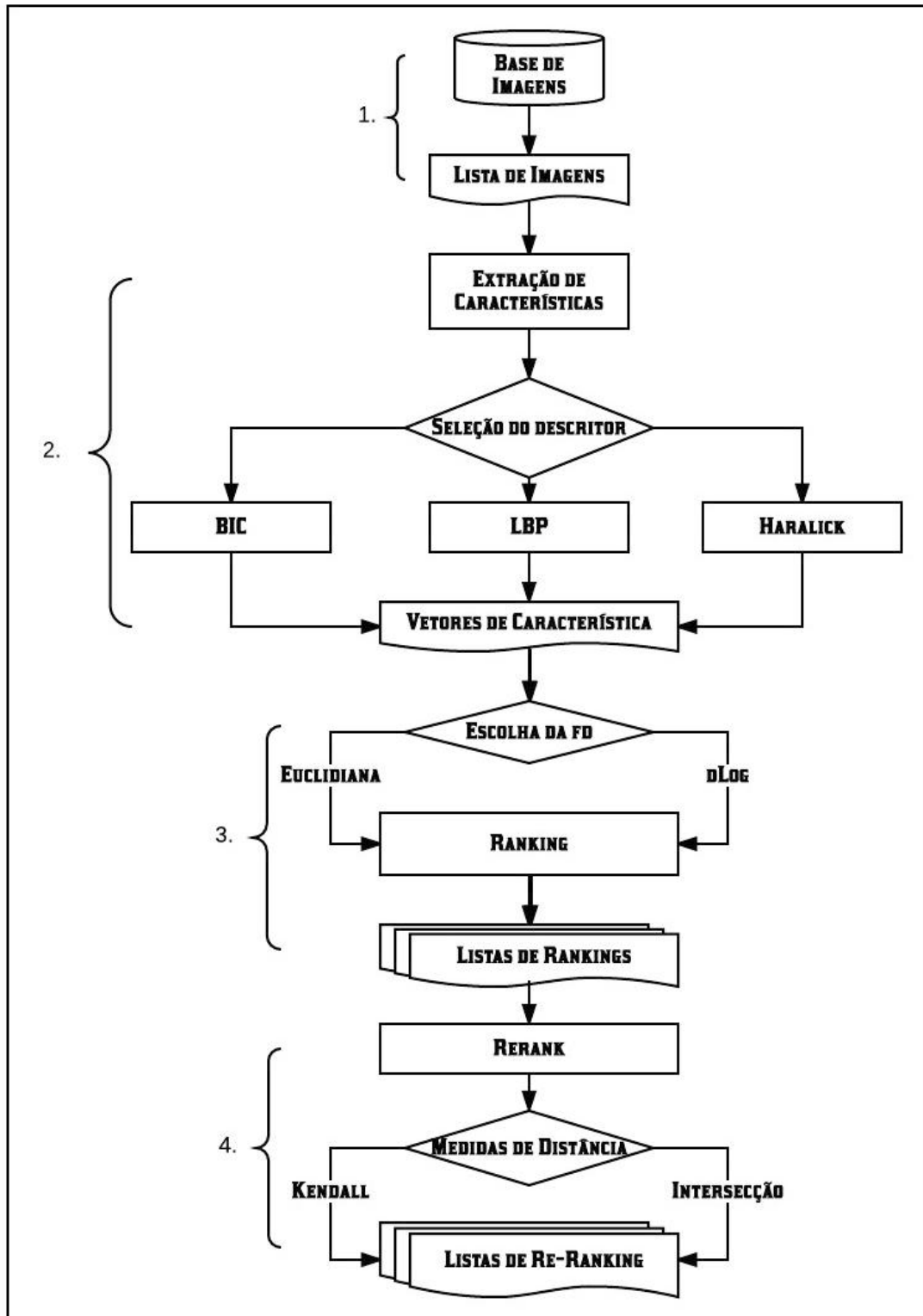


Figura 12: Pipeline de execução

Fonte: Autoria Própria

## 4 RESULTADOS E DISCUSSÕES

Nesta seção serão apresentados os resultados obtidos para todos os possíveis estudos de casos que foram implementados devido à utilização de diferentes bases e métodos.

A Tabela 1 apresenta os resultados obtidos para cada estudo de caso, onde cada linha da tabela representa um estudo de caso. A tabela apresenta a base utilizada, o descritor, a função de distância  $fd$ , a medida de comparação entre as *top k lists*  $mdc$ , os parâmetros de entrada do algoritmo de *Re-Ranking*, sendo eles o número de  $k$ -vizinhos e o número das “ $t$ ” iterações do algoritmo. Nas duas últimas colunas são apresentadas a Precisão Média e a NDCG para cada caso, sendo que estas contém o valor obtido através do *ranking* e o ganho ou perda após a utilização do algoritmo, por exemplo, para o terceiro estudo de caso, obteve-se uma precisão média de 69,76% e uma NDCG de 95,89% executando apenas o *ranking* e ao se utilizar o algoritmo, obteve-se um ganho de 0,34% na precisão média e uma queda de 0,62% na NDCG.

A partir dos resultados da Tabela 1, nota-se que para a maioria dos casos o algoritmo melhorou a precisão média, a NDCG ou ambas. Nota-se também que para o descritor BIC a medida *Kendall* não foi apropriada, pois ao ser utilizada, não obteve-se ganho algum na precisão média e na melhor configuração do algoritmo, a NDCG se manteve igual ou pior que antes da execução do algoritmo. Ainda se tratando do BIC, ao se utilizar a  $fd$  Euclidiana e  $mdc$  Inter(Intersecção), obteve-se os melhores resultados para este descritor com aumento na precisão média e na NDCG. Vale ressaltar que ao se utilizar bases com características diferentes, a precisão mudou, sendo que para a base *Illumination Color* a precisão média e a NDCG foram maiores que o da base *Illumination Direction*. Isto influencia os resultados do algoritmo, pois quando mais alta a precisão, maior é a dificuldade de aumentá-la.

O descritor LBP também obteve alguns resultados positivos, pois na base *Illumination Color*, a NDCG aumentou em todos os cenários e a precisão média se manteve e alguns casos aumentou. Na outra base, ele se manteve ou teve um aumento na precisão média, mais em relação a NDCG ele obteve duas quedas significativas ao se utilizar a medida *Kendall*. Neste descritor, ao se utilizar a medida *Kendall*, a precisão se manteve ou obteve um pequeno ganho,

Tabela 1: Resultados iniciais para cada estudo de caso

Base	Desc	fd	mdc	k e t	Precisão±Ganho	NDCG±Ganho
Illumination Direction	BIC	dLog	Inter	(30;7)	64,39% + 2,65%	94,19% + 0
			Kendall	(7;6)	64,39% + 0	94,19% - 1,16%
		Eucli	Inter	(13;13)	69,76% + 0,34%	95,89% - 0,623%
			Kendall	(7;6)	69,76% + 0	95,89% - 0,81%
	LBP	dLog	Inter	(5;1)	18,73% + 0	74,32% + 0,74%
			Kendall	(34;12)	18,73% + 2,06%	74,32% - 2,59%
		Eucli	Inter	(4;2)	35,96% + 0	87,25% + 0,895
			Kendall	(28;4)	35,96% + 0,49%	87,25% - 8,7%
	Haralick	dLog	Inter	(6;1)	21,07% + 0	76,17% + 0,55%
			Kendall	(70;15)	21,07% + 3,26%	76,17% - 2%
		Eucli	Inter	(37;5)	16,39% + 0,05%	71,70% + 1,21%
			Kendall	(70;15)	16,39% + 7,26%	71,70% + 3,5%
Illumination Color	BIC	dLog	Inter	(14;1)	73,05% + 0,76%	97,09% + 0,18%
			Kendall	(13;1)	73,05% + 0	97,09% - 4,12%
		Eucli	Inter	(13;3)	82,31% + 0,92%	97,80% + 0,4%
			Kendall	(10;3)	82,31% + 0	97,80% - 1,8%
	LBP	dLog	Inter	(10;12)	43,51% + 1,64%	88,89% + 2,85%
			Kendall	(4;2)	43,51% + 0	88,89% + 0,14%
		Eucli	Inter	(11;5)	82,01% + 0,69%	97,08% + 0,17%
			Kendall	(4;2)	82,01% + 0	97,08% + 0,44%
	Haralick	dLog	Inter	(32;1)	40,02% + 0	85,03% + 0,49%
			Kendall	(13;10)	40,02% + 1,85%	85,03% - 3,46%
		Eucli	Inter	(21;2)	24,12% + 0,25%	78,03% + 2,7%
			Kendall	(22;15)	24,12% + 9,12%	78,03% + 3,58%

mas em relação a NDCG os resultados negativos foram bem mais expressivos quando comparados aos resultados positivos.

Ao se utilizar o descritor de Haralick obteve os maiores ganhos em relação a precisão média e a NDCG, quando comparado aos outros descritores. Nota-se também que ao se utilizar este descritor juntamente com a medida de comparação *Kendall*, são obtidos os melhores resultados tanto na precisão quanto na NDCG. Outro ponto importante ao se analisar este descritor, foi que para um aumento na precisão e na NDCG foi necessário a utilização de valores elevados para a geração das *top k lists*, pois devido a baixa precisão do descritor as imagens mais similares a imagem de consulta encontravam-se distantes, e assim ao se utilizar um número alto na busca dos *k* vizinhos, estas imagens foram para as primeiras posições do *ranking*. Cabe ressaltar também que para este descritor, ao se utilizar a distância Euclidiana e *Kendall* a base que obteve o maior ganho foi a *Illumination Color*, mesmo esta sendo a base com o menor número de imagens, diferentemente dos outros casos onde a base *Illumination Direction* apresentou ganhos mais elevados.

Ao final desta análise, foi selecionado o melhor caso entre todos os 24 casos para uma análise mais detalhada, calculando a precisão e a NDCG média para o *ranking* e posteriormente foi calculada a diferença entre a precisão e a NDCG após o *Re-Ranking*. Este melhor caso foi obtido utilizando a base *Illumination Color*, descritor de Haralick, distância Euclidiana e *Kendall* e será apresentado na Tabela 2.

**Tabela 2: Precisão e NDCG média por Classe para o melhor estudo de caso.**

Classe	Ranking		Re-Ranking (22,15)			
	Precisão	NDCG	Precisão	NDCG	Dif-Pre	Dif-NDCG
1	18,7500%	74,3608%	37,5000%	83,7808%	18,7500%	9,4200%
4	17,3611%	76,2127%	20,1389%	72,9461%	2,7778%	-3,2666%
7	14,5833%	77,9146%	34,0278%	76,2245%	19,4445%	-1,6901%
8	29,1667%	72,8392%	38,8889%	83,4804%	9,7222%	10,6412%
11	13,1944%	79,9478%	18,7500%	85,0468%	5,5556%	5,0990%
<b>13</b>	20,8333%	72,0860%	44,4444%	89,9860%	<b>23,6111%</b>	17,9000%
16	15,2778%	79,1646%	20,1389%	73,0165%	4,8611%	-6,1481%
19	14,5833%	78,5952%	27,0833%	79,3513%	12,5000%	0,7561%
23	43,0556%	80,1358%	39,5833%	84,8648%	-3,4723%	4,7290%
<b>26</b>	25,6945%	68,9598%	36,8056%	86,9387%	11,1111%	<b>17,9789%</b>
40	16,6667%	73,6268%	34,7222%	82,2068%	18,0555%	8,5800%
43	26,3889%	70,7648%	29,1667%	79,5604%	2,7778%	8,7956%
46	19,4444%	71,7432%	25,6944%	81,0739%	6,2500%	9,3307%
53	15,9722%	72,2624%	23,6111%	78,2043%	7,6389%	5,9419%
58	30,5556%	72,7734%	52,0833%	81,0852%	21,5277%	8,3118%
60	11,8055%	80,3354%	28,4722%	73,8069%	16,6667%	-6,5285%
62	72,9167%	91,4001%	83,3333%	96,0939%	10,4166%	4,6938%
72	38,1944%	71,9548%	29,1667%	76,0776%	-9,0277%	4,1228%
73	15,2778%	86,1453%	17,3611%	80,8080%	2,0833%	-5,3373%
79	17,3611%	77,3943%	22,9167%	80,3770%	5,5556%	2,9827%
81	20,8333%	80,7068%	38,1944%	84,6274%	17,3611%	3,9206%
<b>83</b>	27,7778%	73,1401%	51,3889%	85,9057%	<b>23,6111%</b>	12,7656%
100	15,2778%	84,8731%	32,6389%	84,5999%	17,3611%	-0,2732%
104	14,5833%	83,6642%	16,6667%	83,8962%	2,0834%	0,2320%
109	15,9722%	82,6184%	35,4167%	74,7403%	19,4445%	-7,8781%
130	62,5000%	81,3995%	37,5000%	77,6294%	-25,0000%	-3,7701%
133	21,5278%	69,9085%	29,1667%	79,2195%	7,6389%	9,3110%
137	39,5833%	76,6408%	59,7222%	88,0198%	20,1389%	11,3790%
141	15,2778%	91,9345%	17,3611%	81,6242%	2,0833%	-10,3103%
152	13,1944%	87,5370%	15,2778%	83,3061%	2,0834%	-4,2309%

Na Tabela 2, estão destacadas as classes que obtiveram um maior ganho, sendo a classe 13 e a classe 83 as classes que obtiveram o maior ganho na precisão chegando a 23,6% e a classe 26 foi a classe que obteve o maior ganho na NDCG, chegando a 17,9%. Ao analisá-las, observou-se que a classe 13 teve um maior ganho na precisão devido ao objeto em estudo ser

uma embalagem com uma ampla variação de textura, o que diferenciou-a das demais classes. Já a classe 83, devido ao objeto em estudo aparentar ser feito de um material diferente dos outros, o diferenciou dos demais objetos. Já para classe 26, ocorreu o mesmo comportamento da classe 13, mas para esta classe, obteve-se um bom resultado na precisão e também o melhor resultado na NDCG. Para este estudo de caso a maioria das classes obtiveram resultados satisfatórios ao se utilizar o algoritmo, sendo que a maioria delas obteve um ganho na precisão e/ou na NDCG, havendo apenas uma classe que obteve queda na precisão e NDCG ao mesmo tempo, a classe 130. Esta queda deve-se ao fato de que o objeto da classe 130 ser fabricado por um material semelhante ao material de outras classes, assim ao ser analisada a textura, imagens de outras classes podem ter sido retornadas devido ao alto número de  $k$  vizinhos utilizados pelo algoritmo.

Ao término desta análise, foram escolhidos os melhores casos de cada descritor para cada base, e realizou-se uma análise semelhante a Tabela 2. Esta análise é apresentada na Tabela 3 e Tabela 4, sendo a primeira referente a base *Illumination Direction* e a segunda a *Illumination Color*. Nestas tabelas, foram apresentadas as classes, juntamente com a diferença da precisão e NDCG comparando-se o melhor caso do algoritmo com o resultado obtido pelo *ranking*. Desta forma, na linha 1 da Tabela 3, a coluna 2 representa a diferença entre o melhor caso do BIC após o algoritmo comparado com o valor obtido no *ranking* inicial utilizando o BIC. O Valor presente nesta célula, de 6,77%, demonstra que após o algoritmo, a precisão média para a classe 1 aumentou 6,77% em relação ao *ranking* inicial. Da mesma forma, ainda na linha 1, na coluna 6, o valor -3,86% demonstra que ao se aplicar o algoritmo no melhor caso do LBP, a NDCG caiu 3,86% em relação ao *ranking* inicial. Os casos escolhidos são descritos a seguir e apresentados nesta mesma ordem na Tabela 3 e Tabela 4.

- *Illumination Direction*:

BIC, dLog, Intersecção,  $k=30$  e  $T=7$ ;

LBP, dlog, *Kendall*,  $k=34$  e  $T=12$ ;

Haralick, Euclidiana, *Kendall*,  $k=70$  e  $T=15$ .

- *Illumination Color*:

BIC, Euclidiana, Intersecção,  $k=13$  e  $T=3$ ;

LBP, dLog, Intersecção,  $k=10$  e  $T=12$ ;

Haralick, Euclidiana, *Kendall*,  $k=22$  e  $T=15$ .

A Tabela 3 mostra que para a maioria das classes obteve-se um ganho na precisão, destacando-se o descritor de Haralick que obteve os resultados mais expressivos referentes a

**Tabela 3: Ganho/Perda na precisão e NDCG ao aplicar o algoritmo para *Illumination Direction***

Classe	Precisão			NDCG		
	BIC	LBP	Haralick	BIC	LBP	Haralick
1	6,7708%	6,0764%	9,2014%	1,3495%	-3,8691%	0,5044%
4	0,0000%	-2,2569%	3,2986%	-0,0357%	-2,2192%	2,4067%
7	1,7361%	0,8681%	11,4583%	0,7927%	-4,2665%	-0,8660%
8	4,8611%	3,1250%	8,5070%	0,6811%	-3,2835%	-2,8651%
11	2,2569%	3,1249%	3,2986%	-0,2168%	-10,8341%	-0,9378%
13	1,9097%	2,6042%	18,9237%	-2,1629%	-0,9778%	8,5523%
16	11,9792%	1,0416%	11,8055%	3,2644%	-1,8171%	2,5269%
19	5,7292%	9,0277%	16,3194%	1,9314%	7,0832%	5,9096%
23	3,1250%	2,4306%	4,5139%	0,9179%	-0,9700%	0,3027%
26	1,7361%	1,0416%	6,4236%	-4,1668%	-0,6503%	-0,1368%
<b>40</b>	3,2986%	5,5556%	12,1528%	0,7243%	<b>14,5843%</b>	5,1884%
43	-2,6041%	0,5208%	3,9930%	1,3220%	1,1396%	-0,9006%
46	-4,1667%	6,7708%	4,5139%	-0,9146%	-12,5882%	1,4286%
53	4,8611%	1,5625%	-1,5625%	1,6441%	2,5156%	7,5088%
58	6,0764%	2,9514%	2,7778%	2,4471%	-5,0381%	4,7693%
60	-0,6945%	4,3403%	15,7986%	-0,9080%	5,0812%	1,9416%
62	3,8194%	-0,8681%	-5,0348%	0,7355%	-8,7223%	3,9168%
72	4,6875%	1,5625%	0,5209%	2,1197%	-12,0492%	7,2600%
73	2,6041%	-2,2569%	3,1250%	0,2375%	-7,0385%	4,8492%
79	4,6875%	1,5625%	5,9028%	1,1212%	-2,2274%	-3,4287%
<b>81</b>	-1,7361%	9,3750%	<b>25,0000%</b>	-1,2755%	7,5715%	10,1181%
83	0,5208%	0,0000%	6,5973%	3,6526%	-6,7960%	5,6927%
100	0,6945%	6,5972%	12,1528%	-4,0590%	14,6163%	11,4252%
104	-2,7778%	-1,5625%	2,4306%	-6,1186%	-4,1708%	5,2477%
109	-2,7778%	0,6944%	-0,8680%	-3,0871%	-10,5639%	-4,0956%
130	3,4722%	6,7709%	19,6181%	0,9179%	7,6738%	11,1792%
133	6,5973%	-5,2084%	1,2152%	1,4742%	-18,9267%	-5,5402%
137	2,7777%	0,6945%	20,1389%	-3,3335%	-6,5987%	13,3610%
141	6,9445%	-1,2153%	-5,3820%	1,5199%	-4,3871%	1,5541%
152	3,1250%	-3,1250%	1,0417%	-0,6825%	-10,0118%	8,0850%

ganhos. Já ao se analisar a NDCG, nota-se que várias classes obtiveram quedas nesta medida, o que se deve ao fato da utilização de altos valores na busca de k vizinhos. Ainda se tratando desta medida, o descritor LBP teve os piores resultados, sendo que apenas 7 das 30 classes obtiveram ganho na NDCG. Em contrapartida o descritor de Haralick obteve queda da NDCG apenas em 8 das 30 classes utilizadas, demonstrando assim que para esta base, este descritor obteve os melhores resultados tanto na precisão média quanto na NDCG. Outro ponto importante foi que as classes 19 e 40 foram as únicas classes que obtiveram ganho em todos os descritores tanto na precisão quanto na NDCG.

A Tabela 4 apresenta os resultados para a base *Illumination Color*.

**Tabela 4: Ganho/Perda na precisão e NDCG ao aplicar o algoritmo para *Illumination Color***

Classe	Precisão			NDCG		
	BIC	LBP	Haralick	BIC	LBP	Haralick
1	2,7778%	-1,3889%	18,7500%	1,6637%	2,9033%	9,4200%
4	-0,6945%	0,3472%	2,7778%	-2,7423%	0,6569%	-3,2666%
7	0,0000%	-2,0833%	19,4445%	-0,3486%	1,6309%	-1,6901%
8	0,0000%	4,1666%	9,7222%	0,0000%	2,2490%	10,6412%
11	2,7778%	-0,6945%	5,5556%	1,1751%	-1,5783%	5,0990%
13	2,0833%	1,3888%	23,6111%	0,9650%	5,4290%	17,9000%
16	-2,7778%	0,3473%	4,8611%	1,8097%	3,2095%	-6,1481%
19	0,0000%	-1,3889%	12,5000%	0,0000%	2,3022%	0,7561%
23	0,6944%	1,7361%	-3,4723%	3,2502%	-0,4850%	4,7290%
26	1,3889%	-0,3472%	11,1111%	0,3102%	1,0116%	17,9789%
<b>40</b>	2,0833%	1,3889%	18,0555%	-0,2943%	<b>2,8135%</b>	8,5800%
43	0,6945%	-2,4305%	2,7778%	-3,6192%	1,2903%	8,7956%
46	2,7777%	4,5139%	6,2500%	1,2330%	1,2724%	9,3307%
53	-2,0833%	0,6945%	7,6389%	-0,5365%	2,4951%	5,9419%
58	0,0000%	1,0417%	21,5277%	-2,0849%	3,2555%	8,3118%
60	0,0000%	-1,3889%	16,6667%	0,0000%	8,7905%	-6,5285%
62	4,1666%	3,4722%	10,4166%	0,8779%	1,7681%	4,6938%
72	2,0833%	2,4305%	-9,0277%	0,4524%	0,2743%	4,1228%
73	-9,0277%	1,7361%	2,0833%	-3,3583%	1,9579%	-5,3373%
79	1,3889%	-2,7777%	5,5556%	0,4497%	-3,6647%	2,9827%
<b>81</b>	-0,6945%	-5,2083%	<b>17,3611%</b>	0,0297%	-2,6842%	3,9206%
83	10,4166%	-1,3889%	23,6111%	6,0513%	0,3065%	12,7656%
100	1,3889%	-2,0834%	17,3611%	0,3240%	-0,2815%	-0,2732%
104	-6,2500%	4,8611%	2,0834%	-4,0828%	1,6683%	0,2320%
109	0,0000%	3,8195%	19,4445%	1,3347%	0,0811%	-7,8781%
130	3,4722%	2,0834%	-25,0000%	0,7985%	1,7130%	-3,7701%
133	4,8611%	5,2083%	7,6389%	0,7228%	0,9459%	9,3110%
137	0,0000%	3,8194%	20,1389%	0,0000%	2,6002%	11,3790%
141	4,8611%	1,0416%	2,0833%	0,5403%	-2,9245%	-10,3103%
152	1,3889%	1,7361%	2,0834%	0,4144%	2,4560%	-4,2309%

Analisando a Tabela 4, em relação a precisão a maioria das classes também obtiveram um ganho na precisão, destacando-se novamente o descritor de Haralick. Ao se analisar a NDCG, o descritor de Haralick também foi o que conseguiu maiores ganhos, mas também para esta base, destaca-se o LBP, que diferentemente da Tabela 3, nesta análise ele obteve resultados bem melhores que o descritor BIC.

E por fim, foi realizada uma análise da medida NDCG utilizando apenas as 10 primeiras posições de cada *ranking*, sendo que esta análise foi feita apenas para a base *Illumination Direction*, pois esta possui 24 imagens por classe. Para esta análise foram selecionados os mesmos casos utilizados na Tabela 3. Esta mesma análise foi feita para a base *Illumination Color*,



mas como esta base possui apenas 12 imagens por classe, os resultados tiveram alterações mínimas.

A Tabela 5 apresenta os valores da NDCG calculada apenas para as 10 primeiras posições, sendo que são apresentados os valores da NDCG igual a 10 para o respectivo *ranking* e a NDCG após a aplicação do algoritmo em sua melhor configuração. Nesta tabela é possível ver que diferentemente das análises anteriores onde o descritor BIC não obteve valores expressivos com relação a ganhos, nesta análise, este descritor quando comparado ao *ranking* teve melhoras expressivas na NDCG, maximizando várias classes, mostrando-se muito eficiente para o cálculo das primeiras posições. Um detalhe importante é que este mesmo descritor no caso geral, apresentado na Tabela 1, apenas aumentava a precisão média e a NDCG se mantinha constante, e ao utilizá-lo buscando apenas as 10 primeiras posições, ele foi o melhor dentre os demais descritores. Já o descritor LBP, mesmo em seu melhor caso, não obteve resultados positivos. Com o descritor de Haralick foi possível ver que para a maioria dos classes ele também conseguiu resultados positivos, mas não tão expressivos quanto os obtidos através do BIC.

Tabela 5: NDCG utilizando as 10 primeiras posições para *Illumination Direction*

Classe	BIC		LBP		Haralick	
	Ranking	Re(30,7)	Ranking	Re(34,12)	Ranking	Re(70,15)
1	94,6285%	100,0000%	83,4894%	71,0095%	90,2170%	82,6793%
4	94,2081%	97,1458%	81,9996%	71,2375%	85,9087%	83,8148%
7	92,8089%	95,4794%	79,9633%	71,3015%	79,4208%	79,4910%
8	94,7856%	100,0000%	78,7896%	69,6844%	80,0573%	82,2411%
11	91,5257%	90,2014%	80,0594%	62,5509%	84,2465%	82,6342%
13	95,8359%	97,5155%	84,5360%	70,7428%	85,2415%	85,3342%
16	92,8036%	97,9288%	76,8909%	67,9800%	82,8029%	76,8631%
19	94,5650%	100,0000%	79,2842%	79,0579%	79,9317%	81,1103%
23	94,8864%	100,0000%	83,3938%	72,8651%	91,8650%	85,7698%
26	96,2609%	98,0848%	83,2780%	71,1229%	80,5610%	84,6992%
<b>40</b>	93,3707%	96,5033%	74,0950%	84,3870%	<b>81,9348%</b>	85,4515%
43	88,0570%	92,2563%	83,7829%	67,0202%	87,5551%	88,3714%
46	95,4969%	97,0421%	90,0541%	74,4932%	87,8185%	90,3633%
53	90,1879%	94,3908%	83,3770%	76,1055%	76,0703%	89,4900%
58	92,0358%	94,9921%	86,9510%	66,8715%	79,4377%	89,6542%
60	93,2419%	93,4815%	84,4127%	78,4295%	81,6766%	83,7231%
62	94,8947%	100,0000%	83,7725%	67,5361%	84,4004%	90,2741%
72	95,4201%	99,7574%	86,3878%	70,6367%	79,2347%	88,0021%
73	93,9109%	98,9957%	82,4390%	65,1933%	81,0931%	83,2192%
79	94,7204%	100,0000%	85,0392%	68,8754%	86,3455%	81,5738%
<b>81</b>	90,1429%	89,7197%	<b>85,5643%</b>	88,3661%	84,9470%	86,7336%
83	93,2835%	98,6046%	80,3653%	64,3148%	79,6296%	77,4946%
100	94,9159%	93,4759%	79,6395%	85,7634%	88,4826%	80,2920%
104	85,3510%	82,3863%	79,8662%	66,5638%	83,3326%	83,1203%
109	89,6599%	90,4431%	82,1066%	66,9262%	80,6749%	80,1564%
130	94,9558%	100,0000%	82,6148%	79,3888%	73,2123%	84,5728%
133	95,2589%	99,9779%	91,2479%	69,1930%	85,5637%	88,8675%
137	95,7232%	97,8727%	80,7419%	67,5697%	83,9763%	86,2212%
141	94,9445%	99,9879%	85,3643%	67,3471%	85,9159%	85,4296%
152	95,8835%	96,8997%	83,9077%	69,4919%	81,7371%	84,2826%

## 5 CONCLUSÃO

Vários trabalhos vem sendo propostos utilizando sistemas CBIR, mas nem todos abordam a questão de utilizar algoritmos de Re-Ranking. Este trabalho, visou o estudo e análise de diferentes descritores de imagens utilizados em Sistemas CBIR, que foram analisados de forma individual e posteriormente analisados aplicando o algoritmo de *Re-Ranking* buscando melhorar a precisão neste tipo de sistema.

Neste trabalho inicialmente foi realizada uma análise utilizando o descritor BIC que para o melhor caso obteve uma melhora de 2,6% na precisão média. Ao se analisar a precisão média por classe, para as duas bases de imagens, este descritor obteve bons resultados, chegando até a um ganho de 11,97% por classe. Já ao se analisar a NDCG apenas para as 10 primeiras posições do *ranking*, este descritor em sua melhor configuração obteve os melhores resultados em relação aos demais descritores. Um ponto negativo foi que para este descritor, ao se utilizar a medida de *Kendall* os resultados não foram satisfatórios.

Para as bases de imagens utilizadas, o descritor LBP foi o que obteve os resultados menos significativos. Em sua melhor configuração, ele obteve um aumento de 1,64% na precisão média e de 2,85% na NDCG. Já ao se analisar a precisão média por classe, este descritor obteve resultados positivos nas duas bases. Ao analisar a NDCG, para a base *Illumination Color* ele de maneira geral obteve ganhos, mas para a base *Illumination Direction*, os resultados foram muito ruins. Um ponto importante é que para a base *Illumination Color*, ele diferentemente dos outros descritores, conseguiu manter ou aumentar a precisão média e a NDCG para todos os casos estudados.

O descritor de Haralick foi o descritor que obteve os maiores ganhos após a execução do algoritmo de *Re-Ranking*, chegando até a um ganho de 9,12% na precisão média e 3,58% na NDCG. Para o melhor caso este descritor conseguiu ganhos de até 23,61% na precisão e até 17,97% na NDCG por classe. Este também foi o descritor que obteve os melhores resultados ao se analisar a precisão média e a NDCG por classe, mostrando-se bem mais eficiente. Ao se analisar a NDCG para as 10 primeiras posições este descritor obteve bons resultados, mas não

foram tão expressivos se comparados ao BIC. Outro ponto importante é que este descritor foi o que obteve os melhores resultados ao se utilizar a medida de *Kendall*.

Assim ao final destas análises, o descritor de Haralick de maneira geral foi o descritor que obteve os maiores ganhos após a implementação do algoritmo de *Re-Ranking*. Um dos motivos deste descritor obter resultados expressivos deve-se ao fato de sua precisão ser baixa, proporcionando assim ganhos mais expressivos tanto na precisão quanto na NDCG. Cabe ressaltar também que a eficiência do algoritmo está relacionada com o contexto que está sendo utilizado, pois as características da base, o descritor e a função de distância influenciam o resultado sendo necessário a calibração dos parâmetros do algoritmo para os diferentes cenários.

## 5.1 SUGESTÕES PARA TRABALHOS FUTUROS

A seguir são apresentadas algumas propostas para a continuidade no desenvolvimento deste trabalho, sendo elas:

- Implementar um algoritmo genético com intuito de obter melhores parâmetros para serem utilizados no algoritmo devido ao fato das características da base e métodos utilizados influenciarem nos resultados de cada caso;
- Implementar o algoritmo de *Re-Ranking* utilizando alguma técnica de processamento paralelo ou distribuído, devido a alta complexidade do algoritmo que tem como base operações com matrizes;
- Implementar um sistema de busca para o usuário, onde algumas configurações estariam disponíveis para a recuperação de imagens utilizando o algoritmo de *Re-Ranking* visto anteriormente;
- Buscar outras medidas de comparação entre as *top k lists* e medidas de avaliação de desempenho nos *rankings* e uma possível implementação de outro algoritmo de *Re-Ranking* para ser comparado com o que foi estudado anteriormente.

## REFERÊNCIAS

- AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face description with local binary patterns: Application to face recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 12, p. 2037–2041, 2006.
- ANANTHARATNASAMY, P. et al. Fusion of colour, shape and texture features for content based image retrieval. In: IEEE. **8<sup>o</sup> International Conference on Computer Science & Education**. [S.l.], 2013. p. 422–427.
- BUENO, J. M. et al. cbpacs: Pac's com suporte à recuperação de imagens médicas baseada em conteúdo. **Unifesp-SP, Setor de Telemedicina**, Universidade Federal de São Paulo, v. 6, p. 7, 2002.
- BUGATTI, P. H. **Análise da influência de funções de distância para o processamento de consultas por similaridade em recuperação de imagens por conteúdo**. Dissertação (Mestrado) — Universidade de São Paulo, 2008.
- CENTENO, J.; PACHECO, A. Classificação de sombra em imagens orbitais e aéreas utilizando uma matriz de co-ocorrência modificada. **Boletim de Ciências Geodésicas**, SciELO Brasil, v. 17, n. 1, p. 75–92, 2011.
- EAKINS, J. P.; GRAHAM, M. E. **A report to the JISC Technology Applications Programme**. Technical Report, University of Northumbria at Newcastle, 1999.
- FAGIN, R.; KUMAR, R.; SIVAKUMAR, D. Comparing top k lists. **Journal on Discrete Mathematics**, Society for Industrial and Applied Mathematics, v. 17, n. 1, p. 134–160, 2003.
- 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. p. 285–294.
- FELIPE, J. C. **Desenvolvimento de métodos para extração, comparação e análise de características intrínsecas de imagens médicas, visando à recuperação perceptual por conteúdo**. Tese (Doutorado) — Universidade de São Paulo, 2005.
- GEUSEBROEK, J.-M.; BURGHOUTS, G. J.; SMEULDERS, A. W. The amsterdam library of object images. **International Journal of Computer Vision**, Springer, v. 61, n. 1, p. 103–112, 2005.
- GUPTA, A. A.; FOSTER, D. P.; UNGAR, L. H. **Unsupervised distance metric learning using predictability**. Technical Reports, Department of Computer and Information Science, 2008. 885 p.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**, n. 6, p. 610–621, 1973.

- HOI, S. C.; LIU, W.; CHANG, S.-F. Semi-supervised distance metric learning for collaborative image retrieval. In: IEEE. **Conference on Computer Vision and Pattern Recognition**. [S.l.], 2008. p. 1–7.
- JÄRVELIN, K.; KEKÄLÄINEN, J. Cumulated gain-based evaluation of ir techniques. **ACM Transactions on Information Systems**, ACM, New York, NY, USA, v. 20, n. 4, p. 422–446, 2002.
- JUNDING, S.; HELI, X. Contour-shape recognition and retrieval based on chain code. In: IEEE. **International Conference on Computational Intelligence and Security**. [S.l.], 2009. v. 1, p. 349–352.
- MÜLLER, H. et al. A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. **International journal of medical informatics**, Elsevier, v. 73, n. 1, p. 1–23, 2004.
- PARIS, A. C. **Análise da Eficiência de Recuperação por Conteúdo de Imagens Médicas, utilizando Extratores de Textura baseados em Wavelet e Wavelet Packet**. Tese (Doutorado) — Universidade de São Paulo, 2008.
- PARK, G.; BAEK, Y.; LEE, H.-K. Re-ranking algorithm using post-retrieval clustering for content-based image retrieval. **Information processing & management**, Elsevier, v. 41, n. 2, p. 177–194, 2005.
- PEDRONETTE, D. C. G. **Exploiting contextual information for image re-ranking and rank aggregation in image retrieval**. Tese (Doutorado) — Universidade Estadual de Campinas - Instituto de Computação, 2012.
- PEDRONETTE, D. C. G.; TORRES, R. d. S. Image re-ranking and rank aggregation based on similarity of ranked lists. **Pattern Recognition**, Elsevier, v. 46, n. 8, p. 2350–2360, 2013.
- PONTI, M.; ESCOBAR, L. Compact color features with bitwise quantization and reduced resolution for mobile processing. In: IEEE. **Global Conference on Signal and Information Processing**. [S.l.], 2013. p. 751–754.
- 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.
- SWAIN, M. J.; BALLARD, D. H. Color indexing. **International journal of computer vision**, v. 7, n. 1, p. 11–32, 1991.
- VASCONCELOS, N.; LIPPMAN, A. A unifying view of image similarity. In: IEEE. **15<sup>o</sup> International Conference on Pattern Recognition**. [S.l.], 2000. v. 1, p. 38–41.
- WANG, R. et al. Re-ranking search results using semantic similarity. In: IEEE. **8<sup>o</sup> International Conference on Fuzzy Systems and Knowledge Discovery**. [S.l.], 2011. v. 2, p. 1047–1051.
- WILSON, D. R.; MARTINEZ, T. R. Improved heterogeneous distance functions. **Journal of Artificial Intelligence Research**, v. 6, p. 1–34, 1997.

ZHANG, D.; LU, G. Evaluation of similarity measurement for image retrieval. In: IEEE. **Proceedings of the International Conference on Neural Networks and Signal Processing**. [S.l.], 2003. v. 2, p. 928–931.

ZHANG, J.; YE, L. Series feature aggregation for content-based image retrieval. **Computers and electrical engineering**, Elsevier, v. 36, n. 4, p. 691–701, 2010.

ZHU, X.; GOLDBERG, A. B. Introduction to semi-supervised learning. **Synthesis lectures on artificial intelligence and machine learning**, Morgan & Claypool Publishers, v. 3, n. 1, p. 1–130, 2009.