

DISSERTAÇÃO DE MESTRADO

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM
CIÊNCIA DA COMPUTAÇÃO

**“Inclusão de Diversidade em Consultas aos
Vizinhos Mais Próximos Usando Descritores
Distintos para Similaridade e Diversidade”**

**ALUNA: Ana Claudia Cardoso
ORIENTADOR: Prof. Dr. Renato Bueno**

**São Carlos
Abril/2017**

CAIXA POSTAL 676
FONE/FAX: (16) 3351-8233
13565-905 - SÃO CARLOS - SP
BRASIL

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**INCLUSÃO DE DIVERSIDADE EM
CONSULTAS AOS VIZINHOS MAIS PRÓXIMOS
USANDO DESCRITORES DISTINTOS PARA
SIMILARIDADE E DIVERSIDADE**

ANA CLAUDIA CARDOSO

ORIENTADOR: PROF. DR. RENATO BUENO

São Carlos – SP

Abril/2017

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**INCLUSÃO DE DIVERSIDADE EM
CONSULTAS AOS VIZINHOS MAIS PRÓXIMOS
USANDO DESCRITORES DISTINTOS PARA
SIMILARIDADE E DIVERSIDADE**

ANA CLAUDIA CARDOSO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Engenharia de Software, Banco de Dados e Interação Humano Computador

Orientador: Prof. Dr. Renato Bueno

São Carlos – SP

Abril/2017



UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de Dissertação de Mestrado do candidato Ana Claudia Cardoso, realizada em 18/04/2017.

Prof. Dr. Renato Bueno
(UFSCar)

Profa. Dra. Marcela Xavier Ribeiro
(UFSCar)

Prof. Dr. Humberto Luiz Razente
(UFU)

Certifico que a sessão de defesa foi realizada com a participação à distância do membro Humberto Luiz Razente, depois das arguições e deliberações realizadas, o participante à distância está de acordo com o conteúdo do parecer da comissão examinadora redigido no relatório de defesa do aluno Ana Claudia Cardoso.

Prof. Dr. Renato Bueno
Presidente da Comissão Examinadora
(UFSCar)

AGRADECIMENTOS

Este trabalho é resultado de muito esforço e perseverança dispensados nestes últimos anos. Um sentimento de orgulho e gratidão por ter chegado ao fim desta etapa tão importante para mim. Gostaria de expressar minha sincera gratidão,

- à minha família pelo apoio e compreensão em todos os momentos não me deixando desistir de alcançar o que sempre busquei;
- ao meu orientador Prof. Dr. Renato Bueno, pelo apoio, confiança, sabedoria e coerência com que guiou este trabalho. Além disso, pela enorme paciência e amizade;
- aos colegas Renata, Cleiton e especialmente ao Felipe pelo apoio durante as disciplinas cursadas e no decorrer do desenvolvimento do trabalho;
- aos professores e funcionários do Departamento de Computação da UFSCar;
- ao Prof. Dr. Humberto Razente por ter gentilmente compartilhado o código do sistema DivDB para alteração e testes realizados neste trabalho;
- aos meus colegas e amigos da Universidade do Estado de Minas Gerais-campus Passos, pelo apoio e força para que eu conseguisse chegar até aqui;
- a Deus pela força interior que apesar de não conseguir explicar é um sentimento que me ajuda a procurar melhorar um pouco a cada dia de minha vida.

Obrigada!

*“Não importa aonde você parou...
Em que momento da vida você cansou...
O que importa é que sempre é possível e necessário "Recomeçar".
Recomeçar é dar uma chance a si mesmo...
É renovar as esperanças na vida e o mais importante...
Acreditar em você de novo.”*

Carlos Drummond de Andrade

RESUMO

Uma das formas para se recuperar imagens em banco de dados, é através de consultas por similaridade. Utilizando características extraídas dessas imagens, como cor, forma ou textura, busca-se identificar semelhanças a um elemento central de consulta. No entanto, os resultados nas consultas podem ser muito semelhantes entre si, o que nem sempre é o resultado esperado. Além da redundância nos resultados, deve-se destacar o problema do ‘gap semântico’, que é a divergência na avaliação da similaridade entre imagens realizada pelo computador considerando a sua representação numérica (características de baixo nível) e a percepção humana sobre a imagem (características de alto nível). Com o objetivo de melhorar a qualidade dos resultados nas consultas buscou-se minimizar a questão da redundância e do ‘gap semântico’ através da utilização de mais de um descritor nas consultas por similaridade. Buscou-se explorar a inclusão de diversidade utilizando-se um descritor para tratar a similaridade e outro descritor para tratar a diversidade, mais genericamente, um espaço métrico para similaridade e outro para a diversidade. Para a implementação da consulta por similaridade utilizou-se a consulta aos vizinhos diversos mais próximos. Considerando-se que os descritores utilizados podem ser distintos e que um deles possa ter maior representatividade numérica do que o outro, foi necessário fazer a normalização, sendo considerados os métodos da normalização pela maior distância e normalização pela maior distância aproximada com balanceamento pela dimensão intrínseca. Para a realização dos testes utilizou-se um algoritmo de busca exaustiva. Os experimentos foram realizados em uma base de dados classificada. Para avaliar a qualidade semântica dos resultados foi proposta uma medida que avalia a inclusão de diversidade considerando a diversidade presente na consulta apenas considerando a similaridade e a diversidade máxima que pode ser incluída. Foi feita uma comparação entre o resultado obtido e o considerado ideal, que refere-se ao valor de λ definido pelo próprio usuário. Comparando-se os resultados alcançados com os resultados obtidos nas consultas para um único descritor, a avaliação da diversidade incluída acompanhou a tendência de λ , o que permite dizer que a normalização e balanceamento é necessário. Além disso, pretende-se futuramente estudar novas formas de normalizar.

Palavras-chave: Recuperação de Imagens por Conteúdo, Consulta por Similaridade, Múltiplos Descritores, Diversidade em Consultas por Similaridade

ABSTRACT

One of the ways to recover images in a database is through similarity queries. Using characteristics extracted from these images, such as color, shape or texture, this work seeks to identify similarities to a central query element. However, the results may be very similar to each other, which is not always the expected result. In addition to the redundancy in the results, the problem of the 'semantic gap', which is a divergence in the evaluation of similarity between images performed by the computer considering its numerical representation (low level characteristics) and the human perception about the image (high level characteristics). In order to improve the quality of the results, we sought to minimize the issue of redundancy and the 'semantic gap' through the use of more than one descriptor in queries for similarity. We sought to explore the inclusion of diversity using one descriptor to treat similarity and another descriptor to treat diversity, more generally a metric space for similarity and another for diversity. For the implementation of the query by similarity was used the consultation to several neighbors closer. Considering that the descriptors may be distinct and one of them may have greater numerical representativeness, it was necessary to do the normalization, considering the methods of normalization by the greater distance and normalization by the greater approximate distance with balancing by the intrinsic dimension. An exhaustive search algorithm was used to perform the tests. The experiments were carried out in a classified database. To evaluate the semantic quality of the results, a measure was proposed that evaluates the inclusion of diversity considering the diversity present in the query only considering the similarity and the maximum diversity that can be included. A comparison was made between the result obtained and the considered ideal, which refers to the value of λ defined by the user himself. By comparing the results obtained with the results obtained in the queries for a single descriptor, the evaluation of the included diversity followed the trend of λ , which allows to say that normalization and balancing is necessary. In addition, it is intended in the future to study new ways of normalizing.

Keywords: Content Based Image Retrieval, Similarity Queries, Combination of Multiple Features, Query Result Diversification

LISTA DE FIGURAS

2.1	Consulta por abrangência	21
2.2	Consulta k-NNq	22
3.1	Conjunto $S_{expandido}$	36
3.2	Resultado da consulta k-NDN para $\lambda = 0$	37
3.3	Resultado da consulta k-NDN para $\lambda = 0.3$	37
3.4	Resultado da consulta k-NDN para $\lambda = 0.6$	38
3.5	Resultado da consulta k-NDN para $\lambda = 0.9$	38
4.1	Resultado da média R_{DIDM} para um único descritor de forma	44
4.2	Resultado da média R_{DIDM} para um único descritor de cor	45
4.3	Comparação da média R_{DIDM} para um único descritor	46
4.4	Resultado da média R_{DIDM} para os descritores de forma e cor sem normalização	47
4.5	Exemplo - forma e cor sem normalização	48
4.6	Resultado da média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância	49
4.7	Resultado média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância aproximada	50
4.8	Exemplo - forma e cor com normalização pela maior distância aproximada	51
4.9	Resultado da média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância com balanceamento FPM	52
4.10	Exemplo - forma e cor com normalização pela maior distância com balanceamento FPM	53

4.11	Resultado média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância aproximada com balanceamento FPM	54
4.12	Comparação dos resultados média R_{DIDM} - Normalização pela maior distância, maior distância aproximada e resultado ideal	55
4.13	Comparação dos resultados média R_{DIDM} - Normalização pela maior distância com balanceamento FPM, maior distância aproximada com balanceamento FPM e resultado ideal	56
4.14	Resultado da média R_{DIDM} para os descritores de cor e forma sem normalização	57
4.15	Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância	58
4.16	Resultado média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância aproximada	59
4.17	Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância com balanceamento FPM	60
4.18	Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância aproximada com balanceamento FPM	61
4.19	Comparação dos resultados média R_{DIDM} entre normalização pela maior distância, maior distância aproximada e resultado ideal	62
4.20	Comparação dos resultados média R_{DIDM} entre normalização pela maior distância com balanceamento FPM, maior distância aproximada com balanceamento FPM e resultado ideal	63

GLOSSÁRIO

- BRID** – *Better Results with Influence Diversification*
- BridGE** – *Better Result with Influence Diversification to Group Elements*
- CBIR** – *Content Based Image Retrieval*
- FPM** – *Fractal-scaled Product Metric*
- GMC** – *Greedy Marginal Contribution*
- GNE** – *Greedy Randomized with Neighborhood Expansion*
- K-NDNq** – *k-Nearest Diverse Neighbor Queries*
- K-NNq** – *k-Nearest Neighbor Queries*
- MAM** – *Método de Acesso Métrico*
- MDP** – *Maximum Dispersion Problem*
- MMC** – *Maximum Marginal Contribution*
- MMR** – *Maximal Marginal Relevance*
- MSD** – *Max-Sum Dispersion*
- NNq** – *Nearest neighbor queries*
- ROT** – *Relação de Ordem Total*
- Rdidm** – *Razão entre Diversidade Incluída e Diversidade Máxima*
- Rq** – *Range Query*
- SGBDs** – *Sistemas de Gerenciamento de Banco de Dados*

SUMÁRIO

GLOSSÁRIO

CAPÍTULO 1 – INTRODUÇÃO	12
1.1 Contexto	12
1.2 Motivação	14
1.3 Objetivo	15
1.4 Organização da Dissertação	15
CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS	17
2.1 Recuperação de Imagens por Conteúdo	17
2.2 Descritor de Imagens	19
2.3 Medidas de Similaridade	19
2.4 Consulta por Similaridade	21
2.5 Múltiplos Descritores	22
2.6 Diversidade	24
2.7 Considerações Finais	28
CAPÍTULO 3 – INCLUSÃO DE DIVERSIDADE UTILIZANDO MÚLTIPLOS DESC- CRITORES	30
3.1 Introdução	30
3.2 O Modelo de Diversidade	32
3.3 Validação dos Resultados	35

3.4	Considerações Finais	39
CAPÍTULO 4 – EXPERIMENTOS		40
4.1	Materiais e Métodos	40
4.2	Experimentos	41
4.2.1	Experimento 1: Único Descritor para Similaridade e Diversidade	43
4.2.2	Experimento 2a: Similaridade pela Forma e Diversidade pela Cor	46
4.2.3	Experimentos 2b: Similaridade pela Cor e Diversidade pela Forma	56
4.3	Resumo dos Resultados	63
4.4	Considerações Finais	64
CAPÍTULO 5 – CONSIDERAÇÕES FINAIS		65
5.1	Contribuições	66
5.2	Trabalhos Futuros	66
REFERÊNCIAS BIBLIOGRÁFICAS		68

Capítulo 1

INTRODUÇÃO

Este capítulo apresenta o contexto em que este trabalho está inserido, a motivação e os objetivos que deram origem a este projeto de pesquisa. Em seguida é apresentado a organização desta dissertação de mestrado.

1.1 Contexto

Imagens são dados complexos de interpretação subjetiva. Em geral, bases de dados de imagens tem grande volume, o que torna um desafio fazer consultas por imagens em banco de dados. Muitas destas consultas podem ajudar a melhorar a qualidade dos serviços prestados em diferentes áreas, tais como a área acadêmica, comercial, saúde dentre outras.

Inicialmente, os Sistemas de Gerenciamento de Banco de Dados (SGBDs) foram criados para tratar dados numéricos e alfanuméricos. Para manipular esses tipos de dados, são utilizados operadores de comparação por igualdade ou operadores relacionais. Normalmente os operadores de igualdade podem ser aplicados a qualquer domínio de dados, uma vez que é possível comparar e decidir se dois elementos são iguais ou não. Já os operadores relacionais ($<$, \leq , $>$ e \geq), que atendem à relação de ordem total (ROT), não são aplicáveis ao domínio de dados complexos, uma vez que dados complexos geralmente não atendem à ROT. Em SGBDs tradicionais, esses são os tipos de operadores mais comuns utilizados.

No entanto, apesar dos operadores de igualdade poderem ser utilizados em qualquer domínio de dados, eles tem pouca utilidade para lidar com dados complexos como imagens. A possibilidade de duas imagens serem idênticas é muito pequena, mesmo que tenham sido obtidas pelo mesmo recurso digital, em um mesmo instante.

Imagens apresentam em seu conteúdo características como cor, forma, textura, os quais

são utilizadas em sua descrição ao serem armazenadas em banco de dados. Para auxiliar o seu acesso e manipulação, foram criados os sistemas de recuperação de imagens em banco de dados.

A partir da década de 70, muitas pesquisas começaram a ser desenvolvidas envolvendo sistemas de recuperação de imagens, principalmente na área de gerenciamento de banco de dados e visão computacional (RUI; HUANG; CHANG, 1999). Estas duas áreas estudam recuperação de imagens sobre diferentes ângulos, sendo um baseado em texto e outro baseado no conteúdo da imagem. Os sistemas de recuperação de imagem baseado em texto evoluíram muito, porém, ainda existem algumas dificuldades, especialmente quando o volume de dados é muito grande. A primeira delas é o trabalho exigido na anotação manual sobre as imagens. Outra dificuldade é decorrente do rico conteúdo das imagens e a subjetividade da percepção humana, ou seja, para uma mesma imagem, diferentes pessoas podem ter diferentes percepções. A imprecisão na anotação sobre cada imagem e a subjetividade na percepção humana podem causar problemas no processo de recuperação de imagens (RUI; HUANG; CHANG, 1999).

No início dos anos 80, com o crescente aumento das bases de dados de imagens e na tentativa de resolver os problemas de recuperação de imagens citados anteriormente, foi proposto a recuperação de imagem baseado em conteúdo (CBIR - *Content Based Image Retrieval*). Desta forma, no lugar de anotar manualmente palavras chaves para recuperar imagens, estas seriam indexadas por seu conteúdo, tais como cor, textura e forma (LIU et al., 2007).

Os sistemas para CBIR utilizam algoritmos de processamento de imagens para extrair as características das imagens. Uma vez definidas as características a serem representadas, o algoritmo passa a extraí-las de cada um dos elementos conforme vão sendo armazenados na base de dados. Em seguida, cada elemento passa a ser representado por seu vetor de características. Então, aplicam-se medidas de similaridade para buscar as imagens na base de dados. As operações de busca são realizadas comparando-se as características extraídas dos dados.

Quando mais de uma característica é utilizada para representar as imagens no banco de dados, os vetores são concatenados em um super vetor de características que juntamente com funções de distância formam o que se denomina múltiplos descritores. O uso de múltiplos descritores tem sido muito comum nas consultas por similaridade.

Para que as consultas por similaridade possam ser realizadas, é necessário determinar o grau de similaridade (ou a dissimilaridade) entre os elementos na consulta. Comumente utiliza-se uma função de distância métrica para o cálculo da similaridade, que associadas a um domínio de dados criam o que é conhecido como espaço métrico. A única informação disponível nos espaços métricos são os objetos e a distância entre os mesmos.

Com a facilidade de obter imagens digitais através de várias mídias disponíveis e o aumento crescente no volume de dados, os resultados nas consultas por similaridade podem ser muito semelhantes entre si, e às vezes até redundantes. Uma das maneiras de tratar este problema é incluir diversidade nas consultas.

A diversidade é utilizada para encontrar elementos que são ao mesmo tempo similares ao elemento central de consulta, mas diversos entre si, melhorando assim o conjunto resposta final em uma consultas por similaridade.

1.2 Motivação

Uma das dificuldades na recuperação de dados complexos como imagens em banco de dados é resolver problemas de ambigüidade e redundância de informação. Normalmente o problema de ambigüidade é mais comum de se encontrar em consultas por palavras chave, ou texto. No entanto, em consultas na qual utiliza-se uma imagem como elemento central de consulta, pode ser mais comum encontrar imagens que são muito parecidas entre si.

Outra dificuldade na recuperação de imagens está relacionado ao ‘gap semântico’ que é conhecido como a diferença entre percepção humana sobre uma imagem e como o computador consegue armazenar e recuperar esta imagem. O ser humano reconhece e descreve um imagem através da visão e percepção que ele tem sobre a mesma que são conhecidas como características de alto nível, enquanto o computador reconhece uma imagem através das características de baixo nível que é a representação numérica das características extraídas e armazenadas de uma imagem em um banco de dados.

Tendo em vista que as consultas por similaridade tornaram-se uma das soluções mais adequadas na recuperação de imagem por conteúdo como pode ser visto em Faloutsos (1997), e que a combinação de múltiplos descritores com o objetivo de determinar a similaridade tem se apresentado como um bom método a ser utilizado, este projeto de pesquisa visa explorar a inclusão de diversidade através de múltiplos descritores nessa consultas.

Na tentativa de evitar que problemas com imagens muito semelhantes façam parte do conjunto resposta, neste trabalho ampliou-se a combinação de múltiplos descritores, para aumentar a expressividade que os mesmos podem conferir às consultas, utilizando-os em papéis diferentes para compor as consultas. Entende-se como papéis diferentes, a combinação de tipos diferentes de descritores para tratar a similaridade e a diversidade.

Estudar maneiras de melhorar as consultas por similaridade torna-se um desafio interes-

sante, uma vez que podem contribuir muito para diferentes áreas de atuação, tais como a medicina, a meteorologia, a agricultura, a área acadêmica e muitas outras.

1.3 Objetivo

Este trabalho explora a combinação de múltiplos descritores para incluir diversidade nas consultas por similaridade. O objetivo principal dessa dissertação é utilizar um descritor para tratar a similaridade e outro descritor para tratar a diversidade nas consultas. Espera-se assim que o resultado apresentado nas consultas sejam menos redundantes e também minimizem o ‘gap semântico’.

Ao realizar uma consulta, a ideia é que a mesma retorne elementos que sejam similares considerando um descritor de característica (por exemplo a forma) e diversos considerando outro descritor (descritor de cor por exemplo).

Para atingir o objetivo desta pesquisa, alguns objetivos específicos tornam-se necessários. Propõe-se utilizar um fator de balanceamento que atribui pesos para a similaridade e a diversidade, proposto inicialmente em Carbonell e Goldstein (1998). Esta dissertação propõe-se a alterar a função objetivo encontrada em Vieira et al. (2011b) que utiliza o fator de balanceamento entre similaridade e diversidade nas consultas aos k-vizinhos diversos mais próximos. A proposta é alterar a função objetivo para incluir diversidade através do uso de múltiplos descritores nas consultas por similaridade.

Como os descritores podem ser distintos, propõe-se neste trabalho fazer a normalização para balancear as distâncias obtidas por cada descritor na função objetivo.

Para avaliar a qualidade dos resultados retornados nas consultas propõe-se definir uma medida para analisar o comportamento da diversidade nas consultas.

Pretende-se assim, melhorar a qualidade das respostas nas consultas por similaridade.

1.4 Organização da Dissertação

Além da introdução, esta dissertação está organizada da seguinte forma:

Capítulo 2: O capítulo trata dos principais conceitos e metodologias relacionados a este trabalho e tem por objetivo fornecer uma base teórica para o leitor. São contextualizadas

as definições sobre Recuperação de Imagens baseado em Conteúdo, Descritor de Imagens, Medidas de Similaridade, Consulta por Similaridade e Diversidade;

Capítulo 3: Este capítulo apresenta a proposta desta dissertação, a inclusão de diversidade utilizando múltiplos descritores;

Capítulo 4: Este capítulo apresenta os experimentos realizados para avaliar a proposta dessa dissertação;

Capítulo 5: Este capítulo apresenta as considerações finais do trabalho, suas contribuições e trabalhos futuros.

Capítulo 2

FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS

Este capítulo tem por objetivo discutir os principais conceitos no âmbito de Recuperação de Imagens por Conteúdo, Descritor de Imagens, Medidas de Similaridade e algumas Funções de Distância, Consulta por Similaridade, Múltiplos Descritores, e por fim Diversidade e Novidade. Em síntese, a ideia é prover uma base teórica sobre os aspectos gerais relacionados ao trabalho. Alguns trabalhos correlatos serão apresentados para os tópicos Múltiplos Descritores e Diversidade.

2.1 Recuperação de Imagens por Conteúdo

Em geral, base de dados de imagens são de grande volume, contém centenas de milhares de imagens. Imagens são dados não-estruturados e de difícil interpretação. Para fazer a sua manipulação, apesar dos operadores de igualdade poderem ser utilizados em qualquer domínio de dados, eles tem pouco utilidade para lidar com dados complexos, como imagens. Já os operadores relacionais, que atendem à ROT, estes não podem ser usados para o domínio de dados complexos (BUGATTI; TRAINA; TRAINA JR., 2008).

Para lidar com dados complexos, é necessário desenvolver sistemas de informação adequados para gerenciar eficientemente essa base, para isso foram criados os sistemas de recuperação de imagens em banco de dados.

De acordo com Datta et al. (2008) recuperação de imagem por conteúdo (CBIR - *Content-based Image Retrieval*) é qualquer tecnologia que ajuda a organizar arquivos de imagens digitais através de seu conteúdo visual. Desde uma função para calcular a similaridade entre imagens até a mais poderosa engenharia para fazer anotação sobre imagens, tudo está no âmbito de

CBIR.

Existem dois tipos de abordagens na recuperação de imagens: uma baseada em texto e outra baseada no conteúdo da imagem.

Em sistemas baseados em texto, para cada imagem é associada um metadado em texto e para se fazer uma consulta ao banco de dados é utilizado uma palavra-chave. Entretanto, é necessário uma anotação prévia no banco de dados, o qual é muito trabalhosa e requer tempo para ser feita. Além disso, o processo de anotação não costuma ser eficiente (TORRES; FALCAO, 2006).

Em sistemas CBIR as imagens podem ser recuperadas de acordo com o seu conteúdo, onde o mesmo pode referir-se a distribuição de cores, textura, forma (SHAPIRO; STOCKMAN, 2001) ou qualquer outra informação que possa ser derivada da imagem propriamente dita (GUDEWAR; RAGHA, 2012). Estes sistemas, utilizam técnicas de processamento de imagens para extrair características visuais e então aplica-se medidas de similaridade para obter as imagens semelhantes durante uma consulta (GUDEWAR; RAGHA, 2012).

Os algoritmos de processamento de imagens, normalmente são automáticos e extraem as características visuais (cor, textura, forma) que representam as propriedades da imagem (TORRES; FALCAO, 2006); essas características são conhecidas como características de baixo nível e são armazenadas na forma de vetores multidimensionais. Em seguida, medidas de similaridade ou dissimilaridade entre dois vetores de características são definidos para cada característica (GUDEWAR; RAGHA, 2012).

Comumente banco de dados de imagens são indexados de acordo com seus vetores de características usando estruturas como M-tree (CIACCIA; PATELLA; ZEZULA, 1997) ou Slim-tree (TRAINA et al., 2002a) para acelerar a recuperação e cálculo da similaridade (TORRES; FALCAO, 2006).

Segundo opinião de Smeulders et al. (2000) a maioria das decepções relacionadas a sistemas de recuperação de imagem estão relacionadas a falta de reconhecimento da existência do ‘gap semântico’ e suas consequências para parametrização do sistema.

Em Smeulders et al. (2000) o ‘gap semântico’ é a diferença entre a informação de uma imagem obtida pela visão humana e sua interpretação para a mesma imagem. Um usuário procura por imagens contendo determinados objetos ou que lhe transmite uma determinada mensagem. Por outro lado, a descrição de uma imagem baseia-se nas características extraídas da imagem (características de baixo nível) e que podem não estar relacionadas com a percepção humana (características de alto nível) sobre a mesma imagem.

Uma típica solução CBIR requer a construção de um descritor de imagem, o qual é caracterizado pelas características da imagem (vetor de características) e medidas de similaridade para comparar duas imagens (TORRES; FALCAO, 2006).

2.2 Descritor de Imagens

Imagens podem ser descritas por suas propriedades ou características principais. As propriedades são agrupadas em um vetor e um critério específico é estabelecido considerando a essência da imagem (TRAINA et al., 2003). Um vetor de características pode ser considerado como um vetor que representa, de forma concisa, um critério específico ou propriedade da imagem. As propriedades de cor, forma e textura são as principais características ou propriedades consideradas.

Conforme pode ser visto em Torres e Falcao (2006) um descritor de imagem D pode ser definido como uma tupla $(\varepsilon_D, \delta_D)$ onde:

- $\varepsilon_D : \hat{I} \rightarrow \mathbb{R}^n$ é uma função, que extrai o vetor de característica \vec{v}_I de uma imagem \hat{I} ;
- $\delta_D : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função de similaridade (baseada na distância métrica) que calcula a similaridade entre duas imagens como o inverso da distância entre seus vetores de características.

2.3 Medidas de Similaridade

Em muitas aplicações, para saber o quanto uma imagem satisfaz uma consulta no banco de dados de imagens, depende-se das medidas de distância ou métricas usadas para verificar a similaridade (SHAPIRO; STOCKMAN, 2001).

Medida de similaridade é uma função que determina o grau de similaridade para um par de imagens, representado por seus vetores de características, normalmente definida como o inverso da função de distância, onde quanto maior o valor de distância, menor o grau de similaridade (TORRES; FALCAO, 2006).

Domínios de dados que possuem uma função de distância métrica são chamados domínios ou espaços métricos. Para conjuntos de dados em espaços métricos somente são considerados os elementos e as distâncias entre eles. As estruturas de indexação aplicáveis a esse domínio de dados são chamadas de método de acesso métrico (MAM).

Um espaço métrico é um conjunto $S \subset \mathbb{S}$ de elementos pertencente a um domínio e uma função de distância $d()$ para estes objetos, $M = (S, d)$ é chamado espaço métrico (TRAINA et al., 2002a).

Uma função de distância métrica $d()$ pode ser definida como: Dado 3 objetos, s_i, s_j e s_k pertencentes a um domínio de dados \mathbb{S} , esta função é dita métrica se satisfizer as seguintes propriedades (TRAINA et al., 2002a):

Simetria: $d(s_i, s_j) = d(s_j, s_i)$;

Não Negatividade: $0 < d((s_i, s_j) < \infty, s_i \neq s_j$ e $d(s_i, s_j) = 0$;

Desigualdade Triangular: $d(s_i, s_j) \leq d(s_i, s_k) + d(s_k, s_j)$.

A função de distância depende das características extraídas dos dados que serão usadas. A escolha da função de distância pode melhorar significativamente a qualidade dos dados complexos a serem retornados (BUGATTI; TRAINA; TRAINA JR., 2008).

Dentre as funções de distância, a família Minkowski são amplamente utilizadas e trabalham sobre espaços vetoriais (BUGATTI; TRAINA; TRAINA JR., 2008).

Conforme pode ser visto em Bugatti, Traina e Traina Jr. (2008) em um espaço vetorial, os objetos são identificados com n coordenadas de valores reais x_1, x_2, \dots, x_n . As funções de distância L_p são definidas como:

$$L_p((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = \sqrt[p]{\sum_{i=1}^n [x_i - y_i]^p}$$

De acordo com o valor associado a p tem-se as variações da família L_p . As distâncias L_p são aditivas, no sentido em que cada característica contribui independentemente para a medida de distância (BUGATTI; TRAINA; TRAINA JR., 2008).

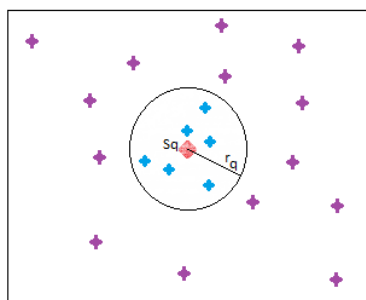
Segundo Bugatti, Traina e Traina Jr. (2008), a família Minkowsky trata cada atributo do vetor de característica independentemente, não leva em consideração o fato que determinados pares de atributos correspondem a características que são mais importantes para noção de percepção de similaridade do que outros.

2.4 Consulta por Similaridade

A similaridade entre elementos tem se tornado a propriedade mais importante ao se fazer consultas por imagens, devido à natureza do tipo de dados das imagens. Comumente, utilizam-se algoritmos de processamento de imagens para extrair características intrínsecas dessas imagens, relacionadas a cor, textura ou forma, representando-as em vetores de características, e mede-se a (dis)similaridade entre elas aplicando-se funções de distância sobre esses vetores. No caso mais genérico, as imagens podem ser consideradas como elementos em espaços métricos, em que a única informação disponível é a distância entre os elementos (TRAINA et al., 2002b).

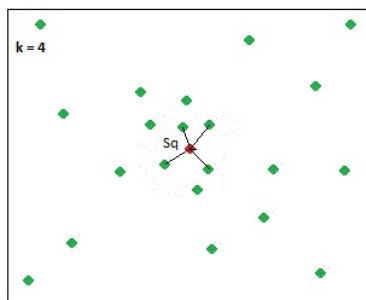
Normalmente, na recuperação de imagens por conteúdo são disponibilizados os dois tipos principais de consulta por similaridade em domínios métricos: consulta por abrangência (*range query - Rq*) e consulta aos vizinhos mais próximos (*nearest neighbor queries - NNq*) (TORRES; FALCAO, 2006).

Figura 2.1: Consulta por abrangência



Fonte: Adaptado de Bueno (2009)

A consulta por abrangência, como mostra a figura 2.1, retorna um conjunto de elementos cuja distância é menor ou igual à distância máxima r_q do elemento central da consulta s_q . O parâmetro de abrangência, referente à distância máxima r_q (ou grau de dissimilaridade máximo) para a resposta, representa a tolerância para recuperação das informações (TRAINA et al., 2002b) (AMATO et al., 2003). A consulta ao vizinho mais próximo retorna o elemento mais próximo do elemento de busca s_q . Uma variação desta consulta é a consulta aos k -vizinhos mais próximos (*k-nearest neighbor queries - k-NNq*), onde k é a quantidade de elementos que serão retornados, como pode ser visto figura 2.2.

Figura 2.2: Consulta k-NNq

Fonte: Adaptado de Bueno (2009)

Tanto a consulta por abrangência quanto a consulta aos vizinhos mais próximos, recebem como parâmetro um elemento do domínio de dados s_q , sendo este o elemento central da consulta, ou também chamado elemento de referência, e a partir deste elemento são calculados quais objetos são mais semelhantes (TORRES; FALCAO, 2006).

2.5 Múltiplos Descritores

Comumente, em banco de dados multimídia para calcular a similaridade utiliza-se uma abordagem baseada no vetor de características. A idéia básica é derivar um vetor de valores numéricos para cada objeto armazenado, extraíndo-se as propriedades que melhor representam a noção de similaridade na aplicação (BUSTOS et al., 2004).

Quando mais de uma propriedade ou característica é usada para compor o vetor de características, as mesmas podem ser concatenadas em um supervetor, em seguida aplica-se uma função de distância para compará-las (BUENO, 2009).

Na maioria das situações reais, empregar mais de um descritor em consultas por similaridade aumenta a precisão nos resultados, pois estes fornecem características complementares que contribuem para representação do conteúdo da imagem (BUENO et al., 2009).

Para combinar múltiplos descritores é preciso uma estratégia que possibilite uma melhor integração entre função de distância e vetor de características para cada descritor, e estabelecer um balanço apropriado dos descritores no cálculo da similaridade (BUENO et al., 2010).

O uso de um grande número de extratores e conseqüentemente, um grande número de características, leva ao aumento da dimensionalidade (*'maldição da dimensionalidade'*), e conseqüentemente diminui a velocidade de recuperação e a precisão na consulta. Portanto, fica evidente a necessidade de melhores abordagens para combinar extratores de características sem

comprometer os resultados (BUENO et al., 2010).

Em Bustos et al. (2004) propõe-se um método que calcula dinamicamente o balanceamento entre os descritores para cada consulta, baseado nas características de consultas pré-classificadas, método conhecido como *purity selection*. O método usa uma combinação de vetores de características através da recuperação da ordem relativa dos vetores de características nas consultas, o que significa que nenhum extrator de características simples sobressai a todos os outros vetores de características em todas as consultas, e que a combinação linear de múltiplos vetores fornecem uma melhoria significativa nas consultas mesmo quando comparado com o melhor vetor de característica simples (BUSTOS et al., 2004).

Em Bueno et al. (2009), propõe-se uma técnica de balanceamento não supervisionada para múltiplos descritores conhecido por *Fractal-scaled Product Metric* (FPM). Este método identifica as contribuições de cada descritor para o cálculo da similaridade dentre todas as imagens - as dimensões intrínsecas dos espaços de características - aproximados por suas respectivas dimensões fractais. Assim, a dimensão fractal de correlação fornece uma estimativa do número mínimo de 'características' necessárias para manter as características essenciais dos dados em consultas por similaridade. Além do mais, a composição da função de distância não é limitada aos extratores de imagem que geram os vetores (dados inerentemente imersos em um espaço multidimensional) e possibilita usar métodos de acesso para aumentar a velocidade na pesquisa de dados. Os resultados deste método comprovaram um aumento na precisão dos resultados. A vantagem em ponderar cada descritor nas consultas por similaridade é garantir uma maior participação de descritores que caracterizem melhor os elementos do conjunto, melhorando os resultados na consulta por similaridade.

Com o objetivo de melhorar a qualidade semântica em consultas por similaridade, em Barroso et al. (2013) foi proposto um método que combina múltiplos descritores usando condições de contorno. O método usa condições de contorno para delimitar um subconjunto de dados dentro de um conjunto maior. Para cada subconjunto estabelece-se a melhor combinação de descritores para ser utilizado em consultas por similaridade, que é feito por um especialista no domínio da aplicação. Entende-se como condições de contorno qualquer informação associada a imagem que possa ser utilizada para estimar limites para o subconjunto, de forma que seja melhor destacado para um ou mais descritores. Para se estabelecer a combinação de descritores é feita um treinamento em imagens pré-classificadas e o resultado é a combinação dos melhores descritores para cada condição de contorno. O uso de condições de contorno para combinar múltiplos descritores em consultas por similaridade mostrou-se promissor para tratar o problema do 'gap semântico'.

Em Barroso et al. (2015) foi proposto um método que busca melhorar a combinação de múltiplos descritores para diferentes condições de contorno fazendo um balanceamento em pares começando pelos melhores descritores. Uma fase inicial define os descritores que poderão ser combinados, é feito um ranqueamento que estabelece os melhores descritores candidatos, o balanceamento é realizado em pares iniciando pelo melhor descritor. A consulta por similaridade utiliza a melhor combinação de descritores de acordo com a condição de contorno e retorna o resultado para o usuário

2.6 Diversidade

Em consultas por similaridade os resultados podem retornar elementos muito semelhantes ao elemento central de consulta. O problema da diversidade pode ser definido em como retornar elementos similares ao elemento central da consulta, mas diversos entre si no conjunto resultado final da consulta, de tal forma que os resultados sejam mais heterogêneos e úteis ao usuário (SANTOS et al., 2013a).

De acordo com Castells, Vargas e Wang (2011), diversidade e novidade são termos diferentes, porém relacionados. Normalmente, novidade refere-se a uma informação que seja diferente das informações que já foram vistas previamente por um usuário. Enquanto a diversidade se aplica a um conjunto de elementos e está relacionado ao quão estes elementos são diferentes entre si (CASTELLS; VARGAS; WANG, 2011).

Em Razente (2009) uma consulta k-NDN tem como objetivo encontrar um conjunto resposta de k elementos que minimize suas distâncias ao elemento central de consulta s_q para tratar a similaridade, e maximize a sua distância entre os elementos do conjunto resposta para tratar a diversidade. Uma consulta k-NDN pode utilizar uma ou duas funções de distância, é definido um domínio de dados \mathbb{S} que é compartilhado por dois espaços métricos $\mathcal{M}_{sim} = \langle \mathbb{S}, \delta_{sim} \rangle$ e $\mathcal{M}_{div} = \langle \mathbb{S}, \delta_{div} \rangle$ (RAZENTE, 2009).

Em Razente (2009) a partir de um elemento de consulta $s_q \in S$ e um conjunto de k elementos $R \subseteq S$, a medida de similaridade $sim : S \times S^k \rightarrow R^+$ de R com relação à s_q é definida como:

$$sim(s_q, R) = \sum_{i=1}^k \delta_{sim(s_q, R_i)}$$

A partir de um conjunto $R \subseteq S$, a medida de diversidade $div : S^k \rightarrow R^+$ de R é dada por:

$$div(R) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \delta_{div(R_i, R_j)}$$

Para calcular a diversidade em uma consulta k-NDN em Razente (2009) utiliza-se uma função onde a partir de um elemento de consulta s_q e um conjunto de dados $S \subseteq \mathbb{S}$, o resultado da consulta k-NDN é o conjunto $R \subseteq S$, $|R| = k$ que minimiza a função: $S \times S_k \rightarrow R^+$ em relação ao elemento s_q .

$$F(s_q, R) = (1 - \lambda) \cdot \text{sim}(s_q, R) - \lambda \cdot \text{div}(R) \quad (2.1)$$

A abordagem adotada em Razente (2009) é a mesma abordagem adotada em Vieira et al. (2011b) descrita logo a seguir, porém em Vieira et al. (2011b) maximiza-se a similaridade ao invés da distância. A medida de similaridade dada por $\text{sim}(s_q, R)$ em Razente (2009) considera a somatória das distâncias, ou seja, a dissimilaridade, porém foi adotada a mesma terminologia.

Em Razente (2009) e alguns trabalhos como Vieira et al. (2011a); Vieira et al. (2011b); Carbonell e Goldstein (1998), utiliza-se um parâmetro λ para fazer o balanceamento entre diversidade e similaridade, que estabelece um peso que varia entre 0 e 1. Quando o peso for igual a zero, significa que será considerado somente a similaridade, o que neste caso, representa uma consulta tradicional k-NN. Quando o peso for igual a um, significa que deseja-se encontrar o máximo de diversidade possível, que é conhecido como o problema da diversidade máxima, que consiste em selecionar um conjunto ótimo de k elementos diversos.

Em termos de complexidade, como pode ser visto em Razente (2009) a função objetivo $F(s_q, R)$ para um conjunto de k elementos apresenta complexidade de tempo de $O(k^2)$. Para encontrar um conjunto de k elementos de um conjunto S que busque o menor valor para a função $F(s_q, R)$, existem ${}^n C_k = n! / (k! \cdot (n - k)!)$ combinações de k possíveis elementos a serem avaliados, o que resulta em uma complexidade de tempo em $O(n^k)$.

Conforme definido em Vieira et al. (2011b), o problema da diversificação de resultados em consultas por similaridade pode ser estabelecido como um problema de otimização, que busca o balanceamento entre encontrar os elementos similares à consulta, mas diversos no conjunto resultado. Considere um conjunto $S = (s_1, s_2, \dots, s_n)$ de n elementos, s_q um elemento de consulta e $k \leq n$. Considere a relevância (similaridade) de cada elemento $s_i \in S$ especificada pela função $\delta_{\text{sim}(s_q, s_i)}, \delta_{\text{sim}} : s_q \times S \rightarrow R^+$, e a diversidade entre dois elementos $s_i, s_j \in S$ dada pela função $\delta_{\text{div}(s_i, s_j)} : S \times S \rightarrow R^+$. Informalmente, o problema da diversidade pode ser descrito como: dado um conjunto S e uma consulta utilizando o elemento central s_q , deseja-se encontrar $R \subset S$ de tamanho $|R| = k$ onde cada elemento em R é relevante para s_q considerando a δ_{sim} e, ao mesmo tempo, diverso em relação a outros elementos em R considerando a δ_{div} .

Em Vieira et al. (2011b), Santos et al. (2015) a busca pelo conjunto R foi modelada definindo-

se um valor para $\lambda \in [0, 1]$ que especifica a proporção desejada entre similaridade e diversidade, o conjunto dos k -elementos similares e diversos R contém k elementos em S tal que:

$$R = \operatorname{argmax} F(s_q, S')$$

$$S' \subseteq S, k = |S'|$$

onde:

$$F(s_q, S') = (k-1)(1-\lambda) \cdot \operatorname{sim}(s_q, S') + 2\lambda \cdot \operatorname{div}(S')$$

A função objetivo $F(s_q, S')$ considera a ‘força atrativa’ entre os k elementos em S' e o centro de consulta s_q , e a ‘força repulsiva’ entre os elementos k em S' (VIEIRA et al., 2011b). A variável λ é definida pelo usuário, dando flexibilidade para enfatizar a relevância (similaridade) ou a diversidade, nos resultados em uma determinada consulta. Caso o usuário deseje consultar apenas os elementos similares ao elemento central de consulta, utiliza-se a variável $\lambda = 0$, neste caso o conjunto resultado depende somente do elemento central de consulta, e o cálculo é direto. Para os casos em que $\lambda > 0$ este é um problema de complexidade computacional NP-difícil. Utilizando-se um algoritmo de força bruta para avaliar a função objetivo que testa todas as possibilidades para encontrar os maiores valores para o subconjunto $R \subseteq S$ de tamanho $|R| = k$ seu tempo computacional é igual a $O(|S|^k k^2)$ onde existem $O(|S|^k)$ possíveis resultados em S de tamanho k para serem avaliados, cada um deles calcula $O(k^2)$ distâncias.

A maximização da similaridade e diversidade é equivalente ao ‘Problema da Diversidade Máxima’ (*Maximum Dispersion Problem - MSD*). De uma maneira geral, em Kuo, Glover e Dhir (1993) o MSD consiste em selecionar um conjunto de elementos em um conjunto maior de tal forma que todos os elementos selecionados apresentam a maior variedade de características encontradas.

Encontrar elementos que são ao mesmo tempo similares e diversos entre si tem despertado bastante interesse na comunidade acadêmica Santos et al. (2015), Santos et al. (2014), Dias, Bueno e Ribeiro (2013), Vieira et al. (2011b).

Em Drosou e Pitoura (2010) foi feito um levantamento sobre as diferentes abordagens para tratar diversificação de resultados, que tem sido propostas na literatura. A partir deste levantamento Drosou e Pitoura (2010) as classificaram em três diferentes categorias: (i) conteúdo (similaridade), isto é, itens que são dissimilares a outros como em Zhang e Hurley (2008). (ii) novidade, isto é, itens que contém nova informação quando comparado com outros previamente vistos como em Clarke et al. (2008), Zhang, Callan e Minka (2002) e (iii) cobertura, isto é, itens que pertencem a diferentes categorias como em Agrawal et al. (2009).

Em Santos et al. (2015) foi proposto um método que procura melhorar a fase inicial de pré-seleção da consulta k-NDN que tradicionalmente limita o espaço de busca aos elementos da consulta k-NN. Sua proposta busca encontrar candidatos que melhorem a diversidade garantindo similaridade mas evitando incluir elementos que pouco contribuem para a fase final que calcula a diversidade, usando técnicas de agrupamento e seleção baseada no conceito de influência.

Em Santos et al. (2014) foi proposto um método '*Better result with influence diversification to Group Elements*' (BridGE) que agrupa hierarquicamente elementos de consulta que são próximos e/ou duplicados, reduzindo a necessidade do usuário em fazer um refinamento na consulta.

Em Santos et al. (2013b) foi proposto um modelo '*Better Results with Influence Diversification*' (BRID) que consiste em inserir diversidade nas consultas por similaridade utilizando apenas propriedades básicas do espaço métrico. Este método é independente de parâmetros, de domínio e de operação de busca. O método usa o conceito de influência, que automaticamente estima a diversidade inerente entre o conjunto de elementos do resultado, sem a necessidade de definição de parâmetros por usuários. Além disso, esta técnica pode ser aplicada sobre qualquer dado representado em um espaço métrico. A técnica BRID é utilizada como base para o algoritmo consulta ao k-diverso vizinho mais próximo e consulta diversa por abrangência, os quais executam consulta ao k-vizinho mais próximo e consulta por abrangência com variedade, mostrando que a técnica pode ser aplicada para diversificar qualquer tipo de consulta por similaridade.

Em Dias, Bueno e Ribeiro (2013) foi proposto um método para calcular os k-vizinhos diversos mais próximos usando dimensão fractal no conjunto de dados. Este método reduz o espaço de busca na consulta baseando-se nas características do conjunto de dados.

Em Vieira et al. (2011b) foram propostos dois métodos, '*Greedy Marginal Contribution*' (GMC) e '*Greedy Randomized with Neighborhood Expansion*' (GNE). Estes métodos utilizam uma função *maximum marginal contribution (MMC)* para calcular a contribuição de cada elemento candidato ao conjunto de resultados. A função mmc combina a relevância de um novo elemento, sua diversidade com relação aos elementos já selecionados no conjunto de resultados e sua máxima contribuição na função considerando os elementos restantes no conjunto. A diferença entre os dois métodos é que no método GMC, o elemento de maior contribuição é sempre escolhido para fazer parte da solução, enquanto que o GNE um elemento é escolhido aleatoriamente dentre os que estão no topo da lista dos elementos que foram elencados para ser incluído na solução.

Para tratar diversidade em consulta por imagem, em Leuken et al. (2009) foi proposto um método para agrupar imagens usando as características visuais sobre um conjunto de imagens previamente recuperadas. Neste método, dada uma consulta feita por um usuário, primeiramente o método determina dinamicamente os pesos das características visuais para capturar melhor os aspectos discriminativos do conjunto de imagens recuperado. Estes pesos são usados em uma função que dinamicamente classifica, a qual consiste em uma técnica de agrupamento de peso para obter uma classificação diversa baseada em agrupamentos representativos. São propostos três algoritmos de agrupamentos, o primeiro agrupa os itens que tem maior representatividade, o segundo que é conhecido por maxmin, busca por uma diversidade visual máxima da representatividade. E por último a escolha recíproca, que permite com que cada imagem elencada vote por outras imagens que sejam melhor representadas, esta é uma estratégia próxima da intuição em um agrupamento. Um exemplo clássico encontrado em Leuken et al. (2009), é a consulta pelo palavra ambígua 'jaguar', no contexto de recuperação de imagens na web.

Agrawal et al. (2009) estudam a diversidade buscando resolver o problema de consultas ambíguas em um conjunto onde existe uma taxonomia de informação em que tanto as consultas como os documentos podem pertencer a mais de uma categoria de acordo com esta taxonomia. Seu método considera a relevância dos documentos através de algoritmos de classificação utilizados na maioria dos algoritmos de busca e a busca pela diversidade de resultados, por meio da categorização para se definir uma taxonomia. Seu objetivo é balancear a relevância e diversidade buscando minimizar o risco de insatisfação dos usuários.

2.7 Considerações Finais

Este capítulo abordou os principais conceitos relacionados a esta dissertação.

Foram pesquisados conceitos sobre recuperação de imagens por conteúdo para um melhor conhecimento sobre o assunto.

Conceitos básicos como descritor de imagens, medidas de similaridade e funções de distância que são de fundamental importância entender como são realizadas as consultas por similaridade.

Foram pesquisados os tipos de consultas por similaridade e o uso de múltiplos descritores para melhorar as consultas. Em seguida foram apresentados conceitos de diversidade.

No próximo capítulo é apresentado o modelo de diversidade que foi desenvolvido neste

trabalho.

Capítulo 3

INCLUSÃO DE DIVERSIDADE UTILIZANDO MÚLTIPLOS DESCRITORES

Este capítulo apresenta o modelo proposto para tratar diversidade em consultas por similaridade, o qual utiliza múltiplos descritores para inclusão de diversidade. Em seguida são apresentados os detalhes de implementação.

3.1 Introdução

Consultas por similaridade tem sido muito utilizadas na recuperação de imagens por conteúdo. Porém, seus resultados podem retornar imagens muito parecidas entre si que podem ser redundantes. Apresentar resultados redundantes pode não atender ao que se espera em uma consulta. Como alternativa, explorar a inclusão de diversidade pode enriquecer o conjunto de respostas, buscando assim, atender melhor às expectativas dos usuários. Através da inclusão da diversidade nas consultas por similaridade, espera-se que as imagens recuperadas sejam similares ao elemento central de consulta, porém diversas entre si, minimizando a redundância nos resultados.

Considere como exemplo uma consulta aos k -vizinhos diversos mais próximos (k -NDN) em que a imagem de um objeto qualquer é utilizada como elemento central. Ao se realizar a consulta tradicional k -NN utilizando um descritor por exemplo, o descritor de forma, prova-

velmente as imagens retornadas tendem a ser semelhantes quanto à forma. Mas, pode ser que os resultados recuperados sejam muitos similares entre si. A inclusão de diversidade pode dar mais expressividade ao conjunto de resposta. Considere a equação 2.1 para tratar a diversidade, considerando-se o mesmo descritor de forma usado para tratar a similaridade, provavelmente as imagens recuperadas sejam diferentes (diversos quanto à forma). Mas o objetivo da consulta pode ser recuperar imagens similares quanto à forma, mas diversas quanto à cor: nesse caso, imagens de um mesmo objeto (com o mesmo formato), mas com cores diferentes. Buscou-se neste trabalho utilizar descritores diferentes em papéis diferentes. Suponha que se queira consultar uma imagem de um carro em uma base de dados. O carro seria o elemento central da consulta. Através da consulta k-NDN, utiliza-se por exemplo um descritor de forma para tratar a similaridade e pode-se optar por um descritor de cor para tratar a diversidade no resultado final da consulta. Com isso, espera-se recuperar imagens de carros que sejam semelhantes quanto a forma e diversas quanto a cor.

Vários algoritmos podem ser usados para inclusão de diversidade em consultas por similaridade (VIEIRA et al., 2011b). Porém, a utilização de descritores semanticamente diferentes para inclusão de diversidade ainda não foi explorada, segundo o conhecimento do autor. Portanto, além de avaliar a qualidade semântica dos resultados obtidos, procurou-se adaptar algoritmos existentes para utilização no cenário específico proposto nesse trabalho.

Explorar diferentes combinações de descritores tanto para consulta por similaridade quanto para a diversidade, demonstra ser uma pesquisa interessante e promissora na busca por melhores resultados em consultas de imagens.

Mais genericamente, pode-se considerar que um espaço métrico é utilizado para representar a similaridade entre os elementos e um outro espaço métrico é utilizado para representar a diversidade. Dessa maneira, pode-se considerar que uma combinação de descritores pode ser utilizada para computar a similaridade e outra combinação de descritores pode ser utilizada para computar a diversidade.

Também podem ser considerados casos em que não necessariamente descritores de imagens sejam utilizados para computar diversidade (ou até mesmo similaridade). Considere uma

consulta sobre imagens médicas similares à imagem central da consulta considerando um descritor qualquer de imagem, mas que a diversidade seja computada baseando-se em informações temporais da imagem: por exemplo, a idade do paciente. Contanto que as informações possam ser representadas em espaços métricos, o modelo proposto pode ser utilizado.

3.2 O Modelo de Diversidade

O modelo dessa dissertação considera dois espaços métricos, sendo que cada espaço métrico pode ser representado por um descritor, α para tratar similaridade e β para a diversidade. Utilizou-se a consulta por similaridade k-NDN e a equação 2.1.

Na consulta k-NDN o conjunto resultado final é calculado em duas etapas. A primeira etapa calcula um subconjunto $S_{expandido}$ que limita o espaço de busca baseado em uma quantidade de elementos aqui representado por $k_{expandido}$ através da consulta tradicional $k_{expandido} - NN$. Em seguida, a segunda etapa calcula o conjunto resposta R sobre o subconjunto $S_{expandido}$ utilizando a equação 2.1 que faz o balanceamento entre similaridade e diversidade.

O modelo de diversidade aqui apresentado foi baseado no trabalho de Razente (2009), onde considerando-se um domínio de dados \mathbb{S} e um conjunto $S \subseteq \mathbb{S}$ busca-se encontrar um conjunto de k elementos em S que minimize as distâncias que tratam a similaridade em relação ao elemento central de consulta $s_q \in \mathbb{S}$ e maximize a diversidade que é dada pela soma das distâncias entre todos os pares de elementos do conjunto resposta R .

Em Razente (2009) a partir de um elemento de consulta $s_q \in S$ e um conjunto de k elementos $R \subseteq S$, a medida de similaridade $sim : S \times S^k \rightarrow R^+$ de R com relação à s_q é definida como:

$$sim(s_q, R) = \sum_{i=1}^k \delta_{sim(s_q, R_i)} \quad (3.1)$$

A partir de um conjunto $R \subseteq S$, a medida de diversidade $div : S^k \rightarrow R^+$ de R é dada por:

$$div(R) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \delta_{div(R_i, R_j)} \quad (3.2)$$

A diferença desta proposta para o trabalho de Razente (2009) reside em encontrar elementos

em um conjunto S que sejam similares ao elemento central de consulta s_q considerando-se um descritor α e diversos considerando-se um descritor β . A função δ_{sim} é calculada considerando-se o descritor α enquanto que a função δ_{div} considera o descritor β . No contexto geral, a equação 2.1 computa a similaridade baseada em um espaço métrico e a diversidade em outro espaço métrico.

Para realizar uma consulta é necessário que o usuário forneça: o elemento central de consulta s_q , o valor de k que indica a quantidade de elementos a serem retornados, um valor $k_{expandido}$ que limita o espaço de busca para consulta $k_{expandido} - NN$ (que pode ser definido em função do parâmetro k) e um valor λ que representa o fator de balanceamento entre a similaridade e diversidade. Uma vez definidos estes parâmetros pode-se iniciar a consulta. A primeira fase inicia-se com a consulta $k_{expandido} - NN$ que calcula o subconjunto $S_{expandido}$ em relação ao elemento s_q baseado na quantidade $k_{expandido}$, esta consulta considera apenas a similaridade δ_{sim} . A partir do subconjunto $S_{expandido}$ aplica-se δ_{sim} sobre o espaço métrico α e δ_{div} sobre o espaço métrico β . O conjunto resposta R tem k elementos e é calculado utilizando-se uma dentre as funções objetivo definidas a seguir (equação 3.3, equação 3.4, equação 3.5, equação 3.6).

Considerando-se que os vetores de características utilizados em cada descritor podem ser muito distintos e que um deles possa ter maior representatividade numérica do que o outro, é necessário fazer a normalização.

Utilizou-se a normalização pela maior distância e normalização pela maior distância aproximada. A normalização pela maior distância é calculada sobre o conjunto total S , para cada par de elementos $s_i, s_j \in S, i \neq j$ calcula-se a maior distância para o descritor α aqui representada por $maxdist_\alpha$ e, em seguida, para o descritor β , representado por $maxdist_\beta$. Os resultados obtidos são considerados para o cálculo da função objetivo na equação 3.3 da seguinte forma:

$$F(s_q, R) = (1 - \lambda) \cdot \frac{sim(s_q, R)}{maxdist_\alpha} - \lambda \cdot \frac{div(R)}{maxdist_\beta} \quad (3.3)$$

A normalização pela maior distância aproximada é calculada sobre o conjunto $S_{expandido}$ que foi calculado pela consulta tradicional $k_{expandido} - NN$. Para cada par de elementos $s_i, s_j \in$

$S_{expandido}$, $i \neq j$ calcula-se a maior distância $aprox_{\alpha}$ para o descritor α , em seguida, para o descritor β , representado por $aprox_{\beta}$. Os resultados obtidos são considerados para o cálculo da função objetivo na equação 3.4.

$$F(s_q, R) = (1 - \lambda) \cdot \frac{sim(s_q, R)}{aprox_{\alpha}} - \lambda \cdot \frac{div(R)}{aprox_{\beta}} \quad (3.4)$$

Considerou-se também o método proposto em Bueno (2009), que utiliza a teoria dos fractais para fazer a normalização considerando o balanceamento pelo fractal entre múltiplos descritores, neste caso os descritores α e β , em consultas por similaridade.

Um fractal é um conjunto de pontos que apresenta a propriedade da ‘autosimilaridade’, independente de escala (FALOUTSOS; KAMEL, 1994). Pode ser interpretado como uma miniatura de um conjunto maior. A dimensão fractal é uma medida de similaridade associada a um fractal.

A dimensão fractal de correlação para balanceamento entre múltiplos descritores, método conhecido como *Fractal-scaled Product Metric* (FPM) encontrado em Bueno (2009) procura identificar a contribuição de cada descritor para o cálculo da similaridade global entre as imagens. Esse valor fornece uma estimativa do número mínimo de ‘características’ necessários para manter as características essenciais dos dados em consultas por similaridade (BUENO, 2009).

Devem ser calculadas as dimensões fractais para o conjunto S considerando-se o descritor α e β , representados por f_{α} e f_{β} . Os resultados obtidos são aplicados tanto na normalização pela maior distância quanto para a maior distância aproximada. Assim, a função objetivo nas duas equações são definidas como:

$$F(s_q, R) = (1 - \lambda) \cdot (f_{\alpha} \cdot \frac{sim(s_q, R)}{normal_{\alpha}}) - \lambda \cdot (f_{\beta} \cdot \frac{div(R)}{normal_{\beta}}) \quad (3.5)$$

$$F(s_q, R) = (1 - \lambda) \cdot (f_{\alpha} \cdot \frac{sim(s_q, R)}{aprox_{\alpha}}) - \lambda \cdot (f_{\beta} \cdot \frac{div(R)}{aprox_{\beta}}) \quad (3.6)$$

Desta forma foram consideradas: normalização pela maior distância - equação 3.3, normalização pela maior distância aproximada - equação 3.4, normalização pela maior distância

com balanceamento FPM - equação 3.5 e normalização pela maior distância aproximada com balanceamento FPM - equação 3.6.

Através da inclusão de múltiplos descritores em papéis diferentes, buscou-se minimizar o problema do ‘gap semântico’ discutido na seção 2.1, procurando apresentar um conjunto resposta de acordo com a preferência do usuário, ou seja, de acordo com o valor definido para λ pode-se escolher a representatividade desejada entre similaridade e diversidade.

3.3 Validação dos Resultados

Para avaliar o modelo de diversidade proposto nesta dissertação é importante avaliar a qualidade dos resultados retornados. Buscou-se avaliar semanticamente a diversidade calculando-se quantas classes de elementos foram incluídas no conjunto resposta R ou seja, a quantidade de classes de elementos r_i que foram incluídas no conjunto resultado R que são diversos em relação aos outros elementos r_j .

Para tentar avaliar semanticamente os resultados apresentados, foi necessário utilizar uma base de dados pré-classificada para que fosse possível identificar a classe de cada elemento no conjunto R e calcular a quantidade de classes de elementos retornadas nas consultas.

Definiu-se uma equação para calcular a Razão entre Diversidade Incluída e Diversidade Máxima identificada por R_{DIDM} , que representa a razão entre a diversidade incluída no conjunto final de respostas R em relação a diversidade máxima possível de ser incluída considerando o subconjunto $S_{expandido}$. Esta medida serve para identificar semanticamente a diversidade que realmente foi incluída no conjunto resposta R , uma vez que todo o cálculo realizado foi baseado na quantidade de classes encontradas no subconjunto $S_{expandido}$ em relação às classes encontradas no conjunto resposta R . A equação é definida como:

$$R_{DIDM} = \frac{C_{kndn} - C_{kmn}}{C_{max} - C_{knn}} \quad (3.7)$$

onde:

C_{knn} representa a quantidade de classes encontradas nos k elementos da consulta tradicional $k - NN$.

$C_{expandido}$ representa a quantidade de classes encontradas no subconjunto $S_{expandido}$.

C_{kndn} representa a quantidade de classes encontradas na consulta k -NDN.

Considerando-se que o total de classes $C_{expandido}$ pode ser maior que a quantidade total de elementos k retornados na consulta, considera-se $C_{max} = \min(k, C_{expandido})$ que representa a quantidade máxima de classes possível de ser encontrada no conjunto resposta R .

Considere o exemplo de uma consulta a um elemento central s_q , para $k = 4$, $k_{expandido} = 40$, $\lambda = 0$. Suponha que nesse exemplo, seja considerado o descritor de forma para tratar a similaridade, e o descritor de cor para tratar diversidade. Após realizar a consulta $k - NDN$, foram obtidos os seguintes resultados: $C_{expandido} = 40$ e $C_{knn} = 4$ e $C_{kndn} = 4$.

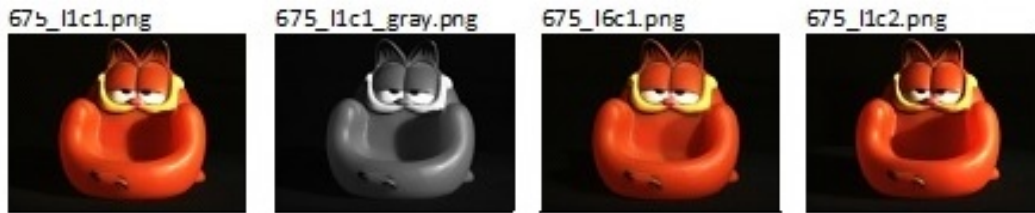
A figura 3.1 representa o subconjunto $S_{expandido}$.

Figura 3.1: Conjunto $S_{expandido}$

Conjunto $S_{expandido}$

1º - 675_l1c1.png	11º - 675_l7c1.png	21º - 675_l7c1_gray.png	31º - 675_l1c1_r.png
2º - 675_l1c1_gray.png	12º - 675_l1c1_gb.png	22º - 675_l2c3_gray.png	32º - 675_l5c1.png
3º - 675_l6c1.png	13º - 675_l2c3.png	23º - 675_l2c2.png	33º - 675_l7c2.png
4º - 675_l1c2.png	14º - 675_l6c2.png	24º - 675_l8c2_gray.png	34º - 675_l7c3.png
5º - 675_l6c1_gray.png	15º - 675_l8c3.png	25º - 675_l4c1.png	35º - 675_l3c1.png
6º - 675_l1c2_gray.png	16º - 675_l8c3_gray.png	26º - 675_l5c3_gray.png	36º - 675_l8c3_gb.png
7º - 675_l2c1.png	17º - 675_l6c3_gray.png	27º - 675_l2c2_gray.png	37º - 675_l7c3_gray.png
8º - 675_l6c3.png	18º - 675_l6c1_gb.png	28º - 675_l5c3.png	38º - 675_l8c1.png
9º - 675_l6c2_gray.png	19º - 675_l1c3.png	29º - 675_l4c3.png	39º - 675_l4c1_gray.png
10º - 675_l2c1_gray.png	20º - 675_l7c2_gray.png	30º - 675_l2c1_gb.png	40º - 675_l7c3_gb.png

A figura 3.2 apresenta o resultado da consulta k -NDN para $\lambda = 0$, ou seja, uma consulta que considera somente a similaridade.

Figura 3.2: Resultado da consulta k-NDN para $\lambda = 0$ 

Utilizando-se a equação 3.7 no exemplo anterior temos:

$$R_{DIDM} = \frac{2-2}{4-2} = 0$$

Quando o resultado de $R_{DIDM} = 0$ pode-se dizer que a quantidade de classes encontrada na consulta k-NDN foi a mesma que a quantidade de classes encontrada na consulta k-NN, ou seja, não houve inclusão de diversidade em relação a consulta k-NN.

Considere o exemplo de uma consulta a um elemento central s_q , para $k = 4$, $k_{expandido} = 40$, $\lambda = 0,3$. Suponha que nesse exemplo, seja considerado o descritor de forma para tratar a similaridade, e o descritor de cor para tratar diversidade. Após realizar a consulta $k - NDN$, foram obtidos os seguintes resultados: $C_{expandido} = 4$ e $C_{knn} = 2$ e $C_{kndn} = 2$.

A figura 3.3 apresenta o resultado da consulta k-NDN para $\lambda = 0.3$.

Figura 3.3: Resultado da consulta k-NDN para $\lambda = 0.3$ 

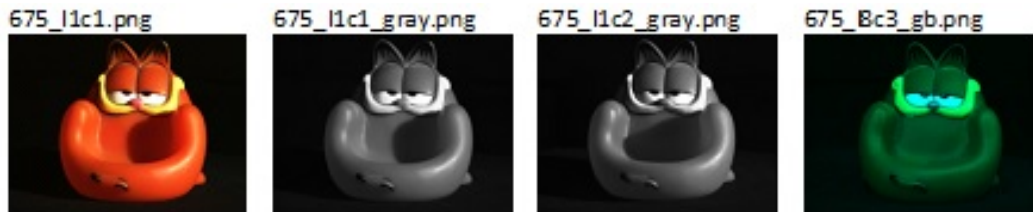
Utilizando-se a equação 3.7 no exemplo anterior temos:

$$R_{DIDM} = \frac{2-2}{4-2} = 0$$

Considere a consulta k-NDN para $k = 4$, $k_{expandido} = 40$, $\lambda = 0,6$. Após a consulta

foram obtidos os seguintes resultados: $C_{expandido} = 4$ e $C_{knn} = 2$ e $C_{kndn} = 3$ conforme pode ser visto na figura 3.4.

Figura 3.4: Resultado da consulta k-NDN para $\lambda = 0.6$



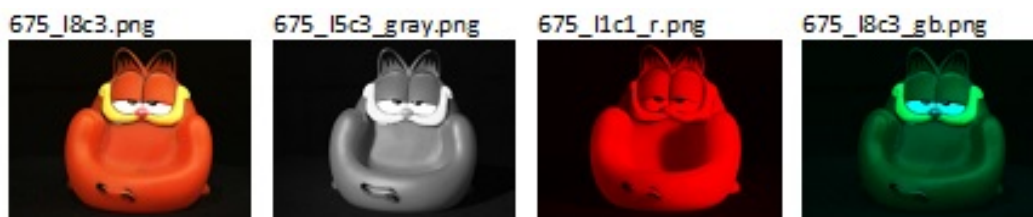
Utilizando-se a equação 3.7 no exemplo anterior temos:

$$R_{DIDM} = \frac{3-2}{4-2} = 0.5$$

Ou seja, considerando-se $\lambda = 0,6$, o resultado obtido para R_{DIDM} foi 0.5. O que pode ser interpretado como inclusão de 50% da diversidade possível de ser incluída, de acordo com $S_{expandido}$ e k .

Considere a consulta k-NDN para $\lambda = 0,9$. Após a consulta foram obtidos os seguintes resultados: $C_{expandido} = 4$ e $C_{knn} = 2$ e $C_{kndn} = 4$ conforme pode ser visto na figura 3.5.

Figura 3.5: Resultado da consulta k-NDN para $\lambda = 0.9$



Utilizando-se a equação 3.7 no exemplo anterior temos:

$$R_{DIDM} = \frac{4-2}{4-2} = 1.0$$

Quando o resultado de $R_{DIDM} = 1$ pode-se interpretar que foi obtido o maior número de classes possível no resultado final da consulta, ou seja, foi incluído 100% da diversidade

possível de ser incluída, de acordo com $S_{expandido}$ e k .

Um resultado negativo de R_{DIDM} significa dizer que o resultado da consulta k-NDN considerando-se diversidade, foi menos diverso do que o resultado da consulta k-NN com somente similaridade.

Para os casos em que $C_{max} - C_{knn} = 0$, ou seja, o denominador foi zero, pode-se dizer que não existe a possibilidade de inclusão de diversidade. Nos experimentos deste trabalho, as consultas que apresentaram esse comportamento foram desconsideradas no cálculo de R_{DIDM} e substituídas com outro elemento central de consulta.

Através da fórmula, é possível calcular o quanto consegue-se incluir de diversidade no resultado final da consulta comparando-se com a consulta tradicional k-NN.

3.4 Considerações Finais

Ao final deste capítulo foi apresentado o método de inclusão de diversidade considerando-se múltiplos descritores proposto nesta dissertação. O método foi implementado utilizando a consulta k-NDN e foi proposta uma função objetivo considerando um tipo de descritor para similaridade e outro tipo de descritor para tratar diversidade. Foi necessário propor uma normalização para evitar uma maior representatividade numérica de um descritor sobre o outro.

Para validar a proposta, definiu-se uma equação que avalia os resultados das consultas calculando a proporção entre a diversidade incluída na consulta k-NDN e a diversidade encontrada na consulta k-NN.

No próximo capítulo serão apresentados os testes realizados nessa dissertação. Espera-se que os resultados alcançados possam contribuir para a melhorar as consultas por similaridade minimizando o problema do ‘gap semântico’, buscando-se assim uma maior satisfação dos usuários.

Capítulo 4

EXPERIMENTOS

Este capítulo apresenta os experimentos realizados para avaliar o modelo proposto em consultas por similaridade utilizando múltiplos descritores. São apresentados as especificações técnicas utilizadas na implementação da proposta, a base de dados utilizada para realizar os testes e a análise do resultados obtidos..

4.1 Materiais e Métodos

Para o desenvolvimento dessa monografia os algoritmos foram implementados em linguagem C++ e o utilizou-se a IDE NetBeans v8.1. A configuração do notebook na qual os experimentos foram executados, foi um notebook Dell Inspiron 14 Série 5000 Special Edition, um único processador Intel Core i7, com 8GB de memória RAM, disco rígido de 1TB com 8GB SSD, placa de video dedicada de 2GB. O sistema operacional utilizado foi o Windows 7.

Os experimentos foram realizados utilizando a base de imagens *Amsterdam Library of Object Images (ALOI)*¹ em sua variação *Illumination Direction* (GEUSEBROEK; BURGHOUTS; SMEULDERS, 2005). A base ALOI é uma coleção de 1000 objetos fotografados cujas imagens são coloridas. O conjunto *Illumination Direction* é composto por 24 configurações de cada objeto com variação de câmera e ângulo, totalizando 24000 imagens, maiores detalhes podem ser encontrados em (GEUSEBROEK; BURGHOUTS; SMEULDERS, 2005).

¹<http://aloi.science.uva.nl/>

Para a validação da proposta dessa dissertação, ampliou-se a base de imagens original para que a mesma imagem fosse representado em diferentes classes de cores. No conjunto original, cada imagem está armazenada considerando o padrão de cores *Red, Green and Blue* (RGB). Para obter classes de cores diferentes, foram removidos os canais G e B resultando imagens apenas com a cor vermelha (R). Para obter imagens em verde e azul (GB), foi removido o canal R. Na base ALOI as imagens são disponibilizadas em RGB e Grayscale. Assim, a base de testes está composta por imagens representadas em quatro classes de cores: RGB, R, GB e Grayscale. Sendo 24000 imagens coloridos em RGB, 24000 imagens em grayscale, 24000 imagens em R e 24000 imagens em GB, totalizando 96000 imagens.

Para extração dos vetores de características dos objetos, foram utilizadas as bibliotecas *Lucene Image Retrieval* (LIRE)² e *JFeatureLib*³, ambas de código livre.

Após a realização dos testes, os resultados obtidos foram plotados em gráficos utilizando-se o Microsoft Excel para melhor visualização, comparação e análise dos resultados.

4.2 Experimentos

Os experimentos foram realizados utilizando um algoritmo de busca exaustiva para consulta k-NDNq, em Razente (2009) identificado por k-NDNq-exaustivo. Foram definidos os descritores de forma e cor para a realização dos testes.

Optou-se por utilizar o *Color Histogram* (SMITH; CHANG, 1997) com dimensão de 512 atributos e métrica L2 para o descritor de cor e *Pyramid of Histograms of Orientation Gradients* (PHOG) definido em Bosch, Zisserman e Munoz (2007) com dimensão de 40 atributos e métrica L2 como descritor de forma.

Para a normalização a distância máxima entre dois objetos para o descritor de forma foi de 1,39 e para o descritor de cor, a distância máxima entre dois objetos foi de 559,68. As dimensões intrínsecas consideradas para fazer o balanceamento FPM para o descritor de forma foi de 6,8 e para o descritor de cor 3,2.

²<http://www.lire-project.net/>

³<https://www.openhub.net/p/JFeatureLib>

As bases de teste foram classificadas. Para o descritor de cor, as classes são identificadas conforme a seguinte descrição:

1. cor - verde e azul (GB)
2. cor - cinza (*Grayscale*)
3. cor - vermelha (R)
4. todas as cores - (RGB)

Considerando o descritor de forma PHOG, a mesma base de imagens foi classificada de acordo com o objeto fotografado (cada um dos 1000 objetos armazenados tem um número único).

O primeiro conjunto de testes descrito a seguir foi realizado com o objetivo de mostrar como a equação 3.7 se comporta para um único descritor. Os demais foram realizados para validação da proposta desta dissertação.

Foram realizados os seguintes conjuntos de testes:

1. consulta k-NDN utilizando um descritor apresentados na seção 4.2.1;
2. consulta k-NDN utilizando múltiplos descritores considerando:
 - (a) similaridade pela forma e diversidade pela cor apresentados na seção 4.2.2
 - (b) similaridade pela cor e diversidade pela forma apresentados na seção 4.2.3
 - sem normalização
 - normalização pela maior distância
 - normalização pela maior distância aproximada
 - normalização pela maior distância com balanceamento FPM
 - normalização pela maior distância aproximada com balanceamento FPM

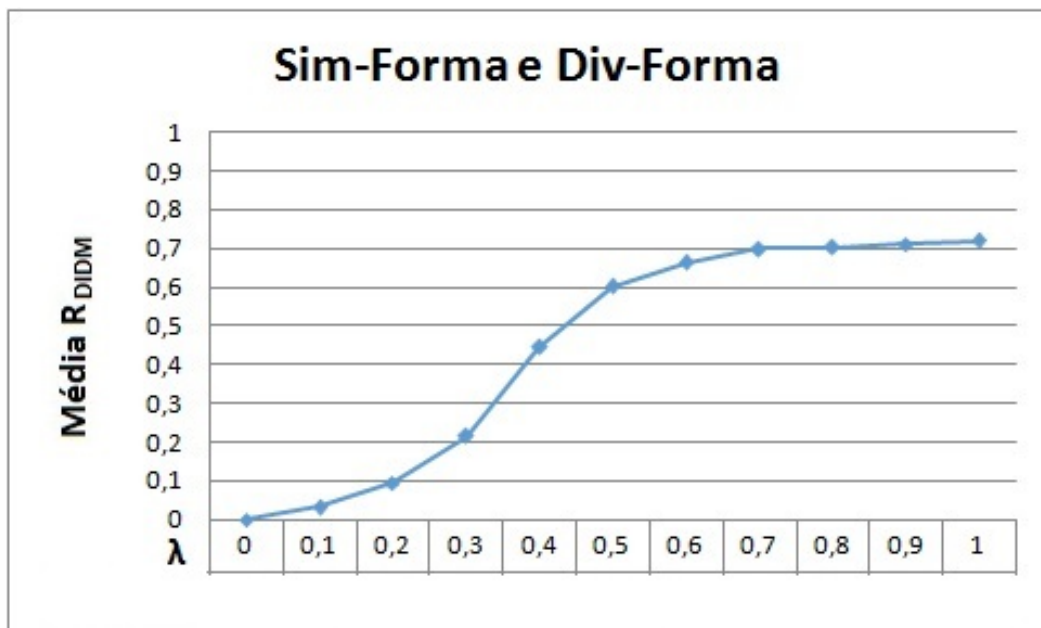
Nos experimentos realizados, cada valor representa a média de 250 consultas (com diferentes centros de consulta). Considerou-se o valor inicial para $\lambda = 0$ (apenas similaridade, desconsiderando a diversidade) variando em 0.1 até o valor $\lambda = 1$ (diversidade máxima). Para as consultas considerando similaridade pela forma e diversidade pela cor adotou-se $k = 4$ e $k_{expandido} = 40$. Para as consultas considerando similaridade pela cor e diversidade pela forma adotou-se $k = 8$ e $k_{expandido} = 40$.

Para avaliar os resultados utilizou-se a equação R_{DIDM} que, conforme descrito na seção 3.3 representa o quanto foi incluído de diversidade na consulta $k - NDN$ em relação à consulta tradicional $k - NN$. Os resultados foram analisados comparando-se o valor λ , que representa o percentual de diversidade que deseja-se incluir no conjunto resposta R , com o resultado da média R_{DIDM} obtido. Assim, quando $\lambda = 0$ deseja-se encontrar 0% de diversidade na consulta, ou seja, não inclui diversidade. Quando $\lambda = 1$ deseja-se encontrar 100% de diversidade na consulta, ou seja, diversidade máxima possível. Quando $\lambda = 0.3$ deseja-se encontrar 30% de diversidade na consulta. Assim sendo, considerou-se como resultado ideal o próprio valor de λ .

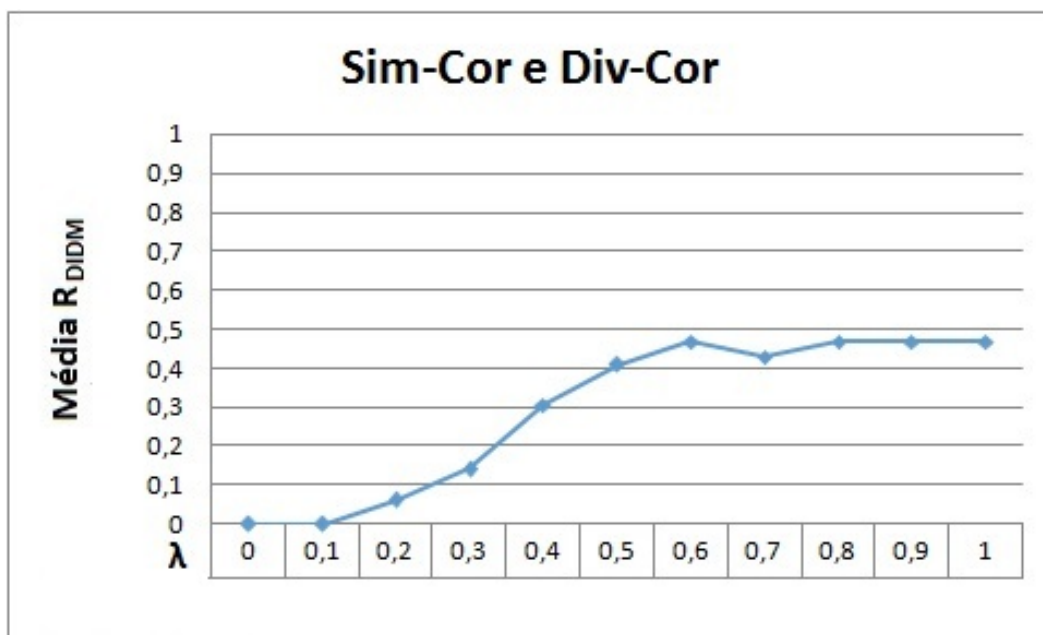
4.2.1 Experimento 1: Único Descritor para Similaridade e Diversidade

O objetivo deste experimento foi verificar o comportamento da equação R_{DIDM} aplicada em consultas k-NDN utilizando um único descritor para servir como referencial para os experimentos desta dissertação. Espera-se que o comportamento do algoritmo k-NDNq-exaustivo considerando-se dois tipos de descritores seja no mínimo semelhante ao obtido por apenas um descritor.

A figura 4.1 apresenta o gráfico com os resultados obtidos pela consulta k-NDNq para um único descritor de forma. O eixo horizontal representa o valor de λ utilizado para fazer as consultas. O eixo vertical representa o valor médio R_{DIDM} calculado para cada conjunto de 250 consultas utilizando somente o descritor de forma (tanto no papel de similaridade como de diversidade).

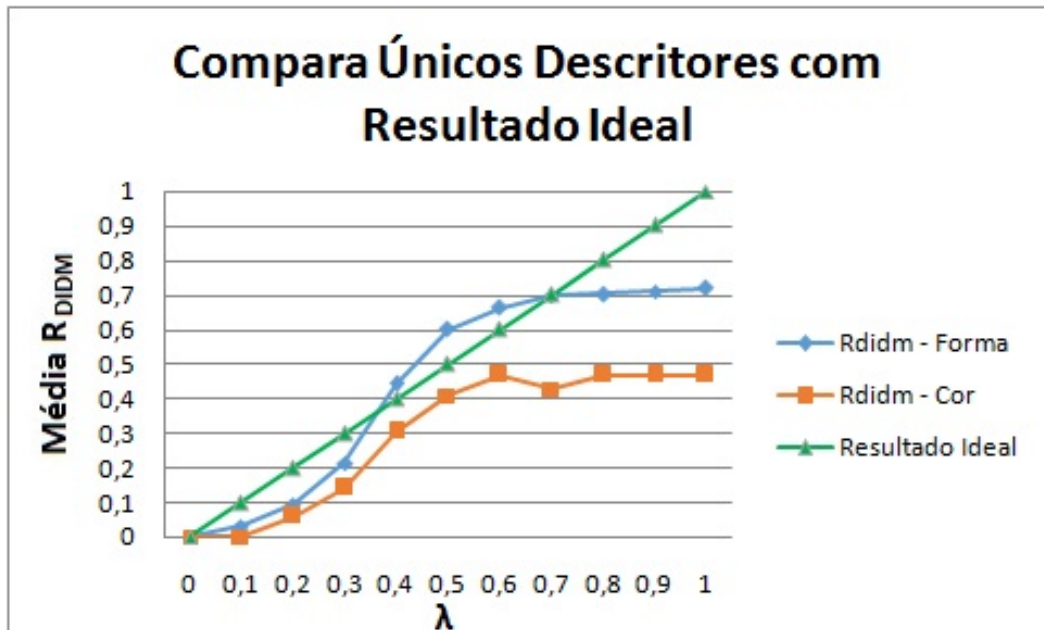
Figura 4.1: Resultado da média R_{DIDM} para um único descritor de forma

De acordo com os resultados apresentados na figura 4.1, pode-se perceber uma curva ascendente da média R_{DIDM} no intervalo entre $0,2 \leq \lambda \leq 0,6$. Pode-se dizer que para este intervalo o percentual de inclusão de diversidade praticamente acompanhou a variação do λ . Para $\lambda \geq 0,6$ os valores praticamente não se alteram, ou seja, o percentual de inclusão de diversidade para estes valores permanecem no mesmo nível.

Figura 4.2: Resultado da média R_{DIDM} para um único descritor de cor

A figura 4.2 apresenta os resultados obtidos na consulta k-NDN considerando-se similaridade e diversidade pelo descritor de cor. Os resultados mostraram que a média R_{DIDM} para todos os valores de λ foi inferior ao próprio valor de λ (considerado ideal).

O gráfico apresentado na figura 4.3 faz uma comparação entre os resultados obtidos para um único descritor de forma, único descritor de cor e o resultado considerado ideal. O valor λ identifica o percentual de diversidade que deseja-se incluir no conjunto resposta R . Assim sendo, considerou-se como resultado ideal o próprio valor de λ . Espera-se que os resultados da média R_{DIDM} obtidos nas consultas estejam próximos do valor de λ . Conforme pode-se perceber para $\lambda \leq 0,6$ os valores de R_{DIDM} acompanham os valores de λ . O descritor de forma esteve mais próximo do valor esperado do que o descritor de cor.

Figura 4.3: Comparação da média R_{DIDM} para um único descritor

4.2.2 Experimento 2a: Similaridade pela Forma e Diversidade pela Cor

A seguir são apresentados os experimentos para a consulta k-NDN considerando o descritor de forma para tratar a similaridade e o descritor de cor para tratar a diversidade.

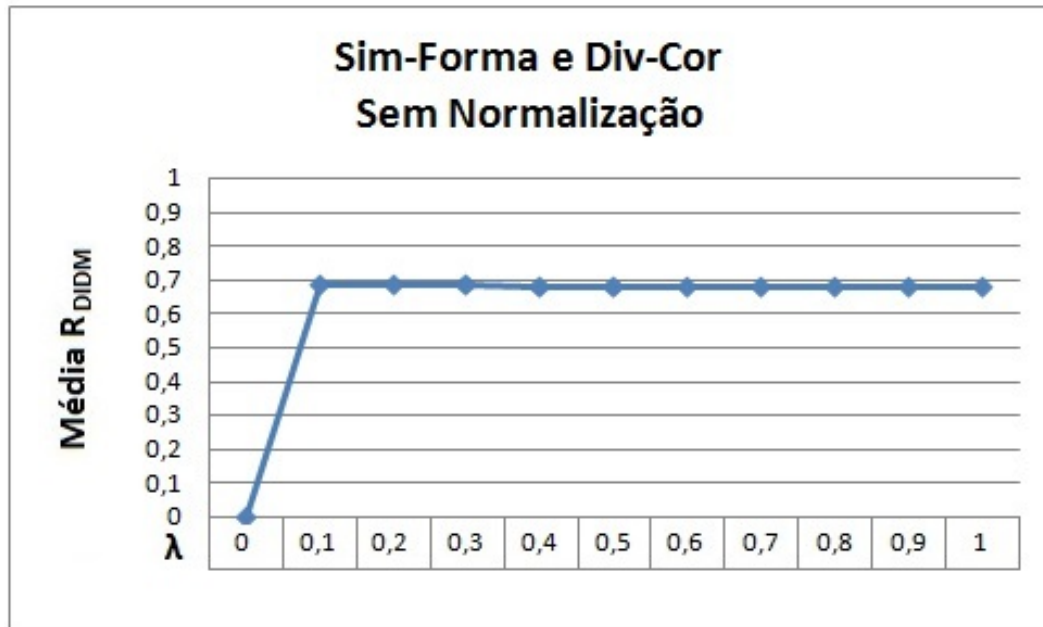
4.2.2.1 Sem Normalização

Para a consulta k-NDN considerando o descritor de forma para similaridade e o descritor de cor para diversidade sem fazer a normalização, a média R_{DIDM} manteve-se constante para todos os valores de λ , conforme pode ser observado na figura 4.4. Considerando-se os resultados obtidos, pode-se dizer que o descritor de cor teve maior representatividade que o descritor de forma. É importante ressaltar que a amplitude das distâncias usando o descritor de cor foram muitos maiores: - usando o descritor de forma, a distância máxima entre dois objetos foi de 1,39 - usando o descritor de cor, a distância máxima entre dois objetos foi de 559,68.

Através da análise dos resultados da consulta k-NDN sem fazer a normalização, pode-se perceber que utilizando-se descritores diferentes, ou mais genericamente, dois espaços métricos diferentes, um para tratar a similaridade e outro para a diversidade, os resultados não foram

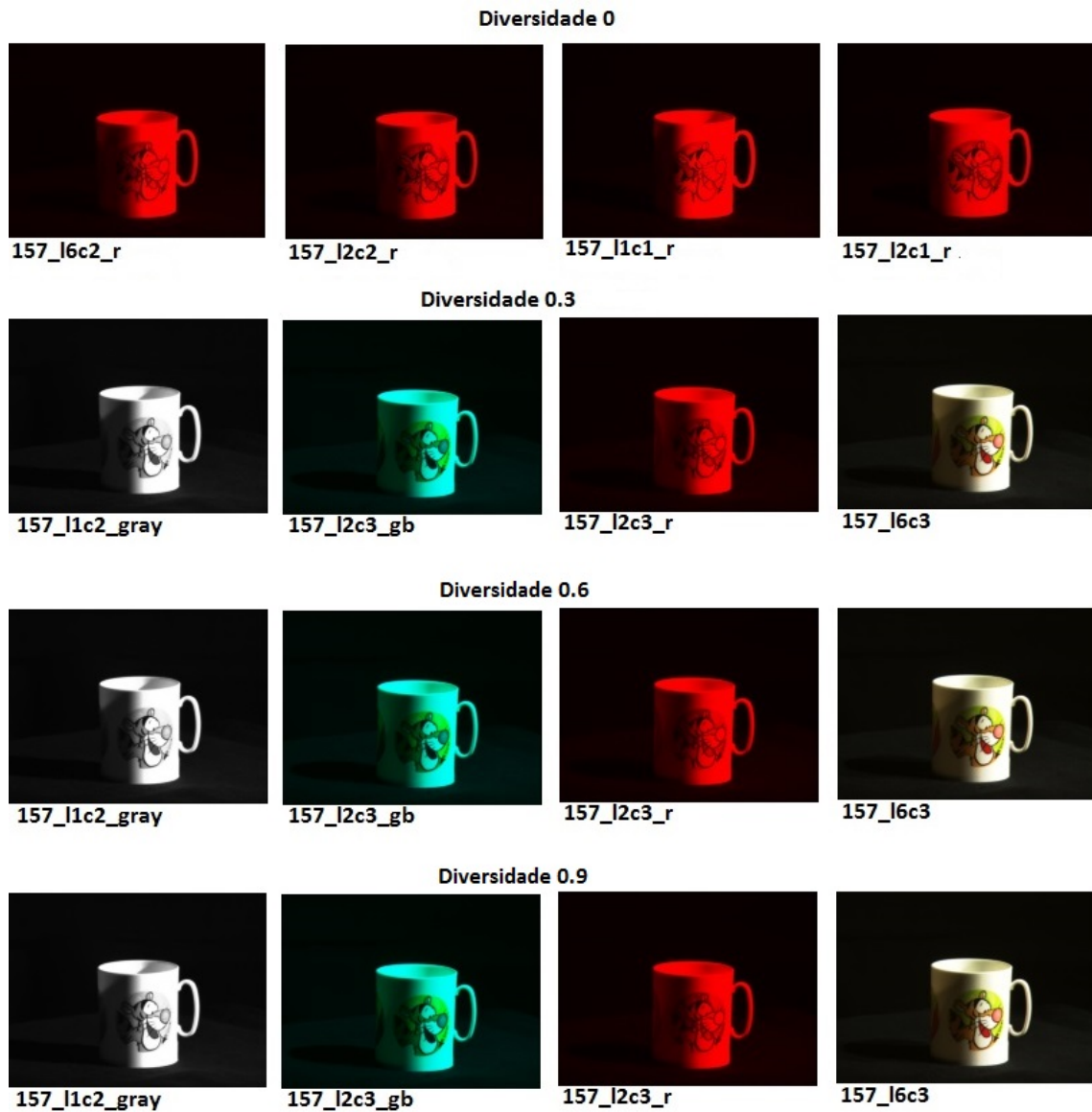
bons. O que mostra a necessidade de fazer a normalização.

Figura 4.4: Resultado da média R_{DIDM} para os descritores de forma e cor sem normalização



Conforme pode ser observado no exemplo da figura 4.5, para $\lambda = 0$, ou seja, 100% de similaridade, os resultados mostram elementos pertencentes à mesma classe de cor do elemento central s_q . Porém, considerando-se $\lambda = 0,3, 0,6$ e $0,9$, o conjunto resposta é o mesmo independente do valor λ escolhido.

Figura 4.5: Exemplo - forma e cor sem normalização

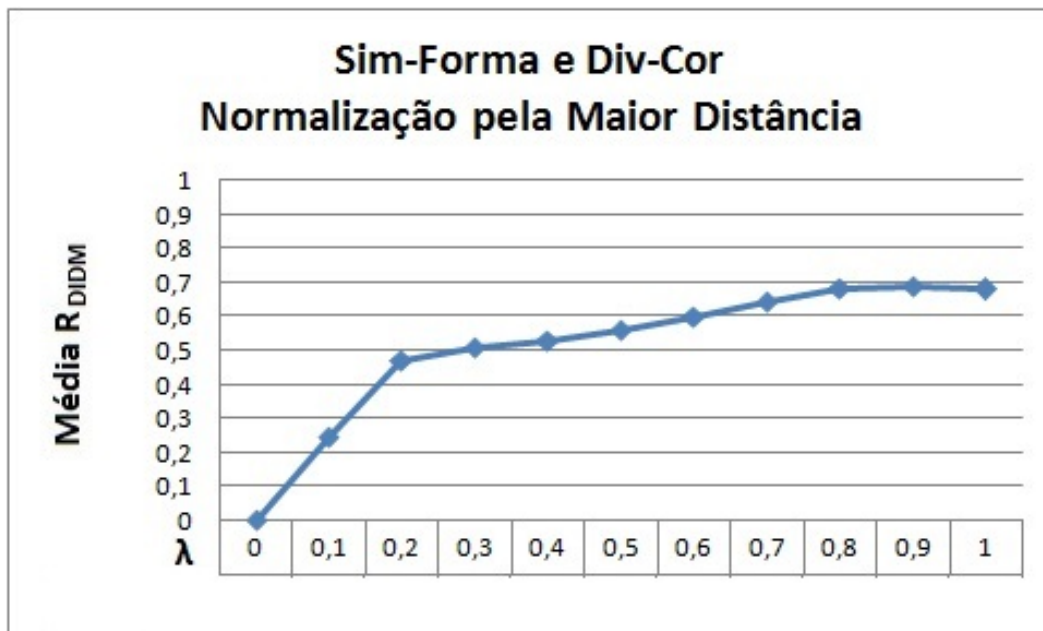


4.2.2.2 Normalização pela maior distância

Quanto aos resultados da consulta k-NDN considerando o descritor de forma para similaridade e cor para diversidade fazendo-se a normalização pela maior distância, conforme pode ser observado na figura 4.6, os resultados apresentados para $\lambda \leq 0,4$ estiveram acima do próprio valor de λ , ou seja, a diversidade obtida foi acima do valor definido na consulta. Para $\lambda \geq 0,7$ os resultados estiveram abaixo do próprio valor de λ ou seja, a diversidade obtida foi abaixo do valor definido na consulta. Para $\lambda = 0,5$ e $\lambda = 0,6$ os resultados acompanharam o valor de λ ,

nesta dissertação considerou-se o próprio valor de λ como valor ideal de resposta.

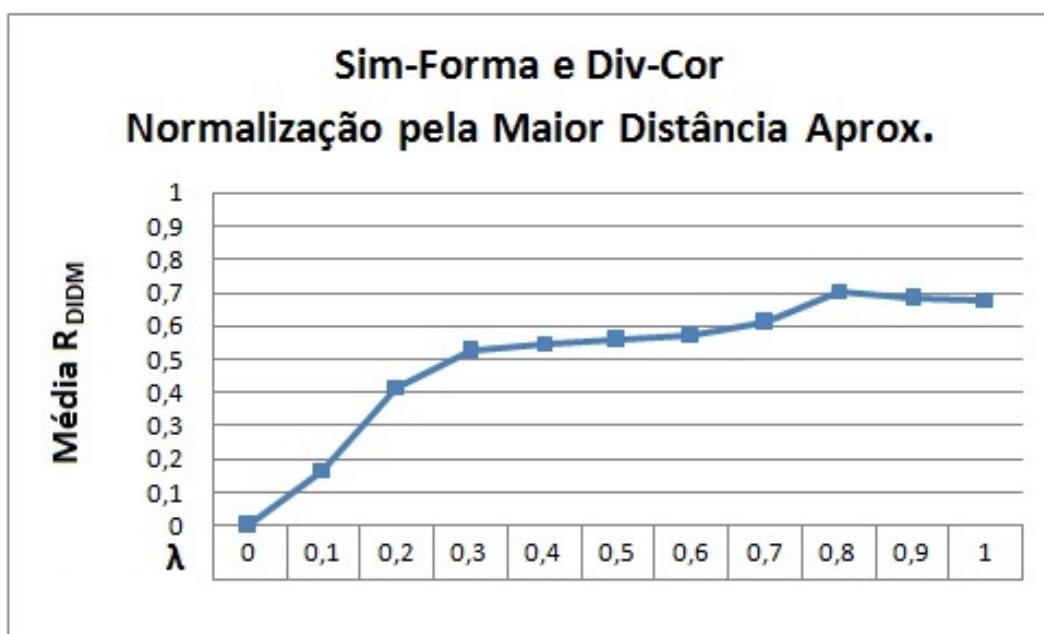
Figura 4.6: Resultado da média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância



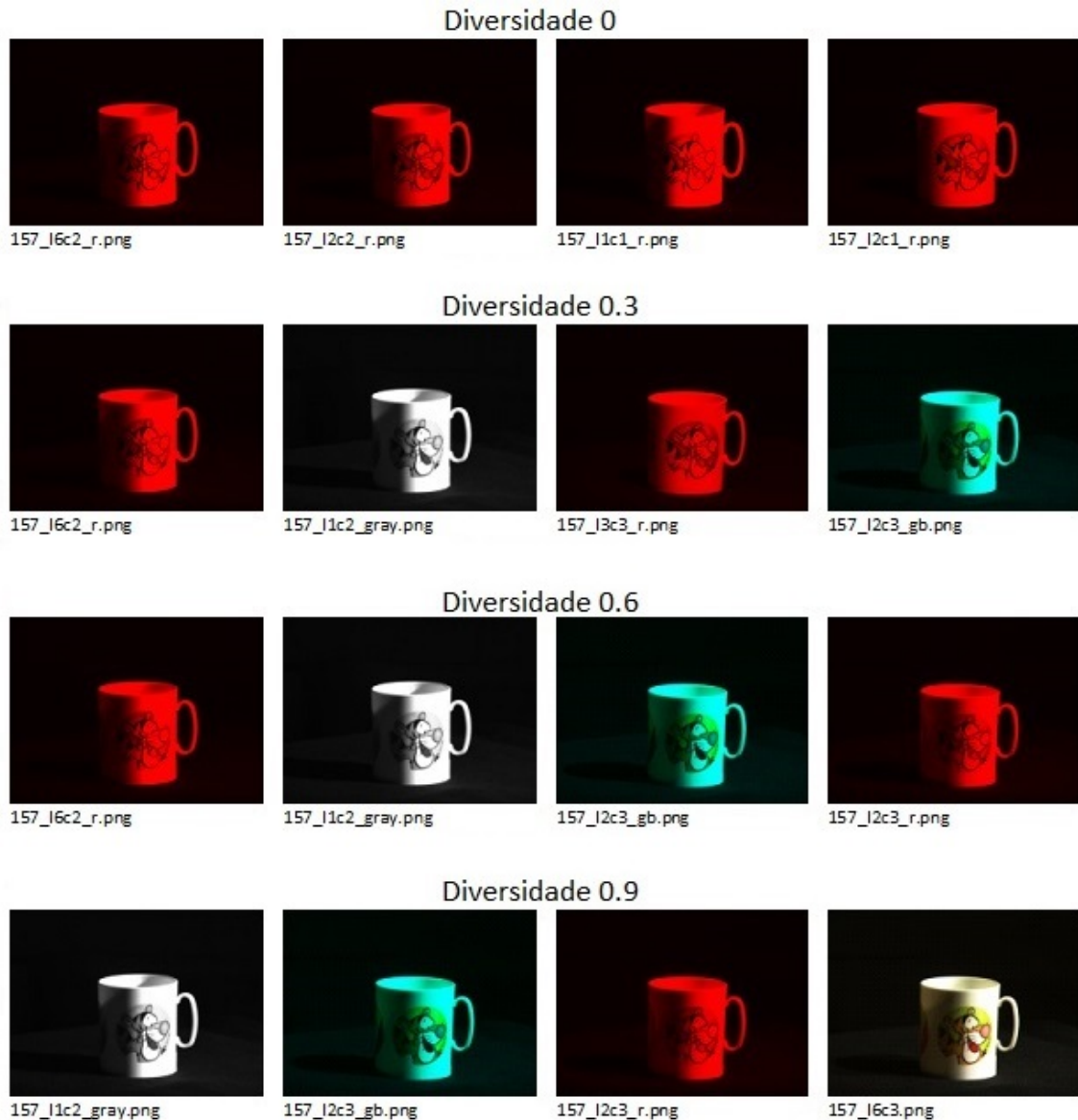
4.2.2.3 Normalização pela maior distância aproximada

A figura 4.7 apresenta o gráfico dos resultados referente a consulta k-NDN considerando similaridade pela forma, diversidade pela cor e normalização pela maior distância aproximada. Pode-se perceber que os resultados apresentados para $\lambda \leq 0,4$ estiveram acima do próprio valor de λ , ou seja, a diversidade obtida foi acima do valor definido na consulta. Para $\lambda \geq 0,7$ os resultados estiveram abaixo do próprio valor de λ ou seja, a diversidade obtida foi abaixo do valor definido na consulta. Para $\lambda = 0,5$ e $\lambda = 0,6$ os resultados acompanharam o valor de λ , nesta dissertação considerou-se como patamar ideal de resposta o próprio valor de λ .

Figura 4.7: Resultado média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância aproximada



As imagens da figura 4.8 exemplificam os resultados apresentados no gráfico da figura 4.7.

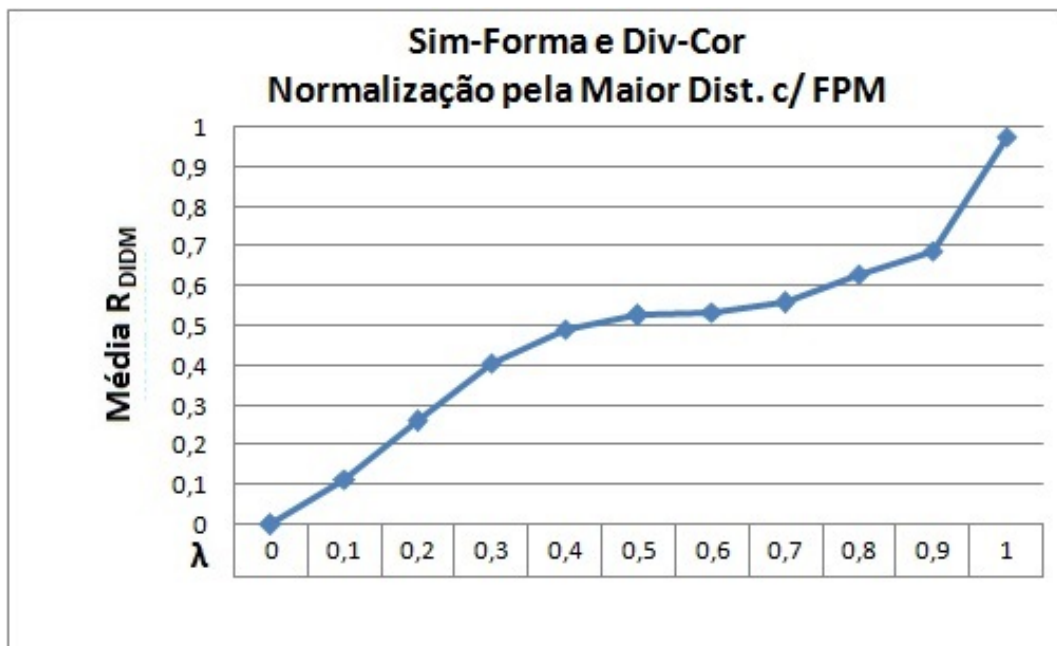
Figura 4.8: Exemplo - forma e cor com normalização pela maior distância aproximada

4.2.2.4 Normalização pela maior distância com balanceamento FPM

As dimensões intrínsecas consideradas para fazer o balanceamento FPM para o descritor de forma foi de 6,8 e para o descritor de cor 3,2.

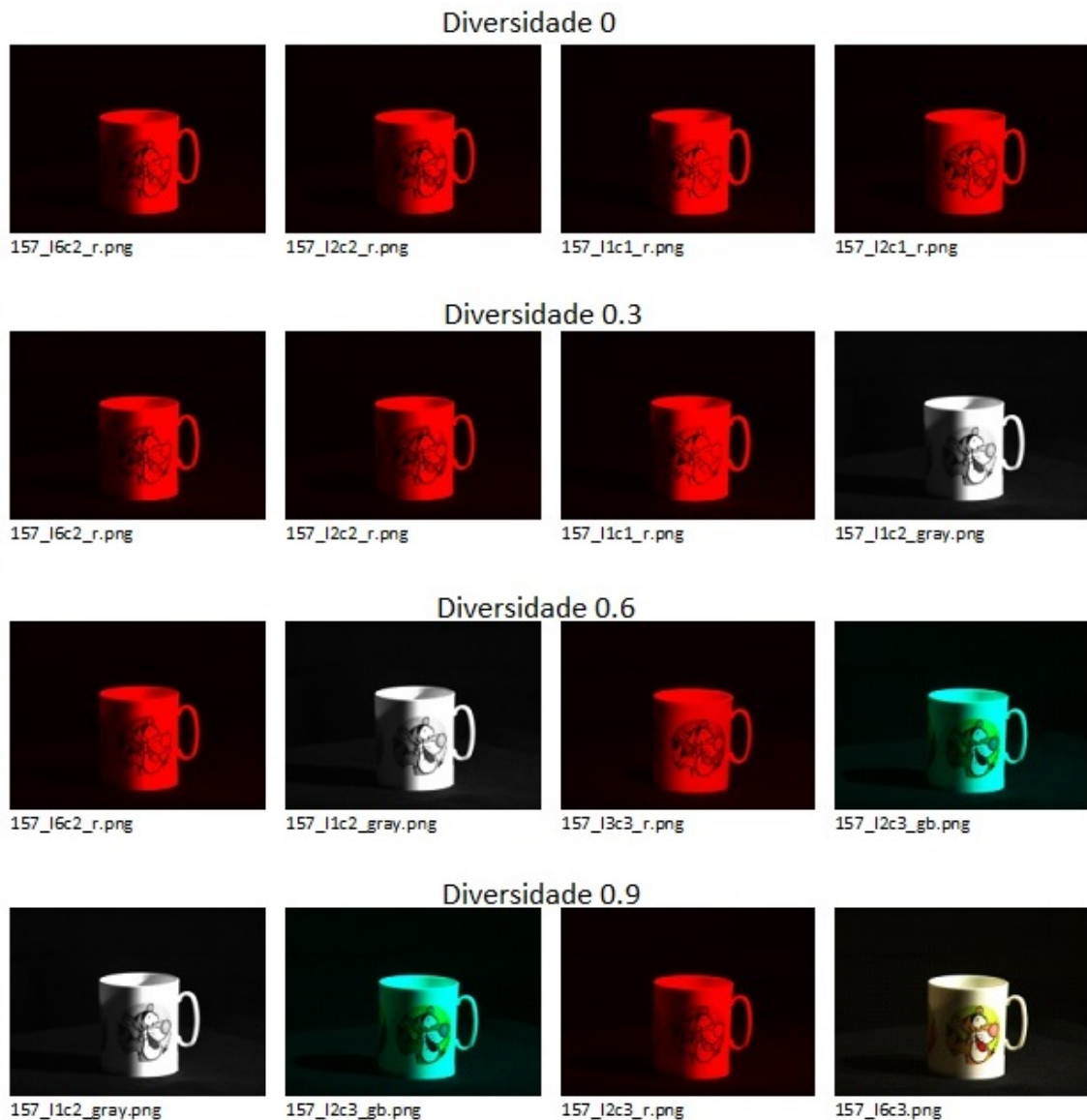
Quanto aos resultados da consulta k-NDN considerando o descritor de forma para similaridade, cor para diversidade e normalização pela maior distância com balanceamento FPM, conforme pode ser observado na figura 4.9 os resultados apresentados mantiveram-se próximos do resultado considerado ideal.

Figura 4.9: Resultado da média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância com balanceamento FPM



Conforme pode ser observado na figura 4.10, as imagens apresentados demonstram o resultado no gráfico 4.9. O número de classes de cores obtidas para todos os casos exemplificados acompanharam os valores de λ considerado ideal.

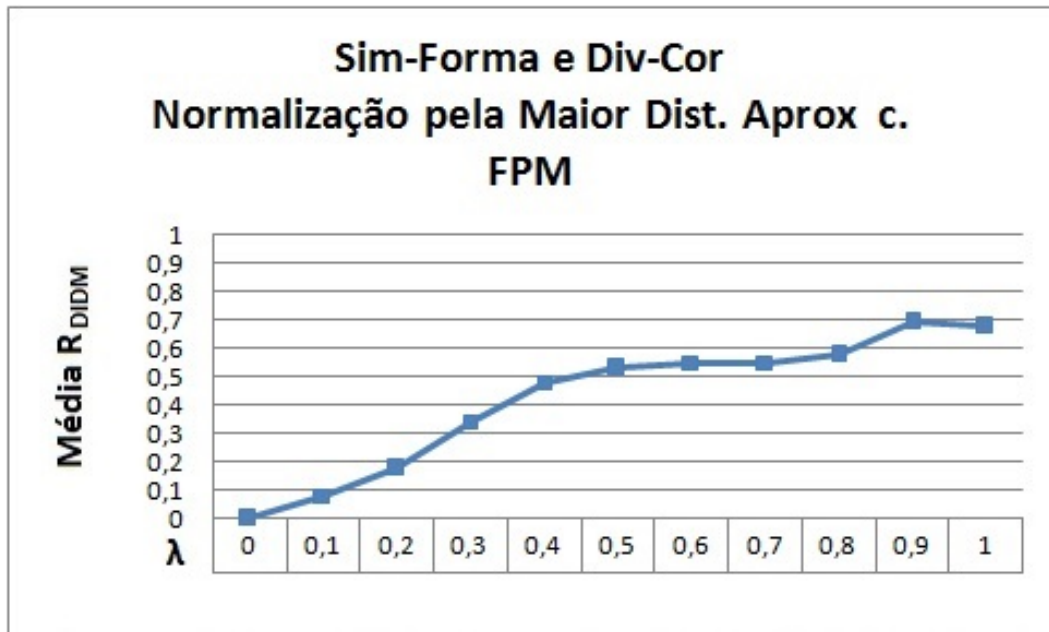
Figura 4.10: Exemplo - forma e cor com normalização pela maior distância com balanceamento FPM



4.2.2.5 Normalização pela maior distância aproximada com balanceamento FPM

A figura 4.11 apresenta os resultados da consulta k-NDN considerando similaridade pela forma, diversidade pela cor e normalização pela maior distância aproximada com balanceamento FPM. Conforme pode ser observado na figura 4.9 os resultados apresentados estiveram sempre próximos do resultado considerado ideal.

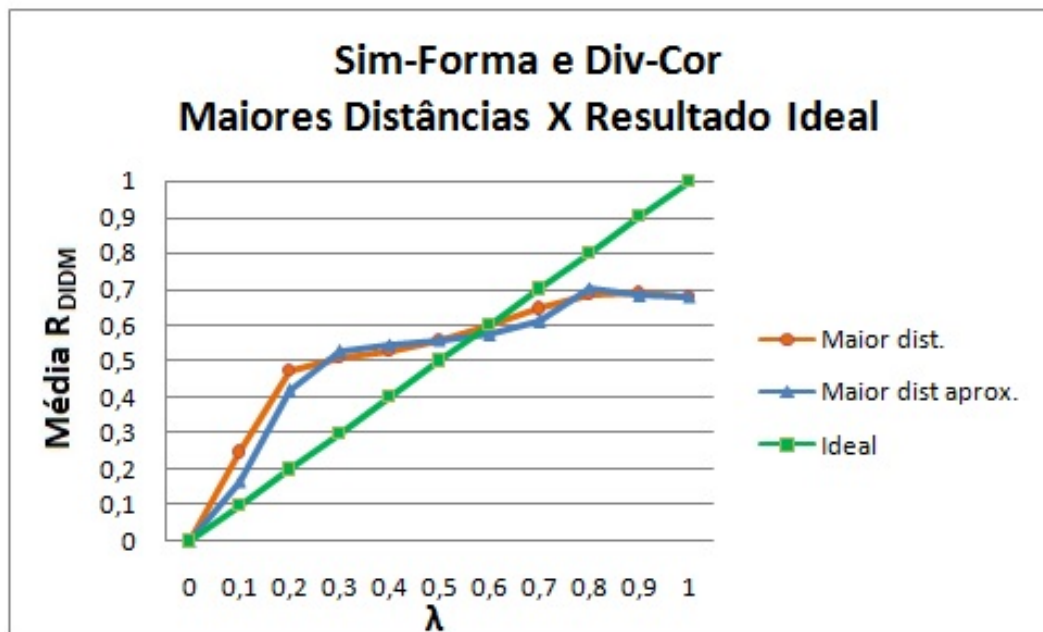
Figura 4.11: Resultado média R_{DIDM} para os descritores de forma e cor com normalização pela maior distância aproximada com balanceamento FPM



4.2.2.6 Comparação dos resultados para k-NDNq considerando similaridade pela forma e diversidade pela cor

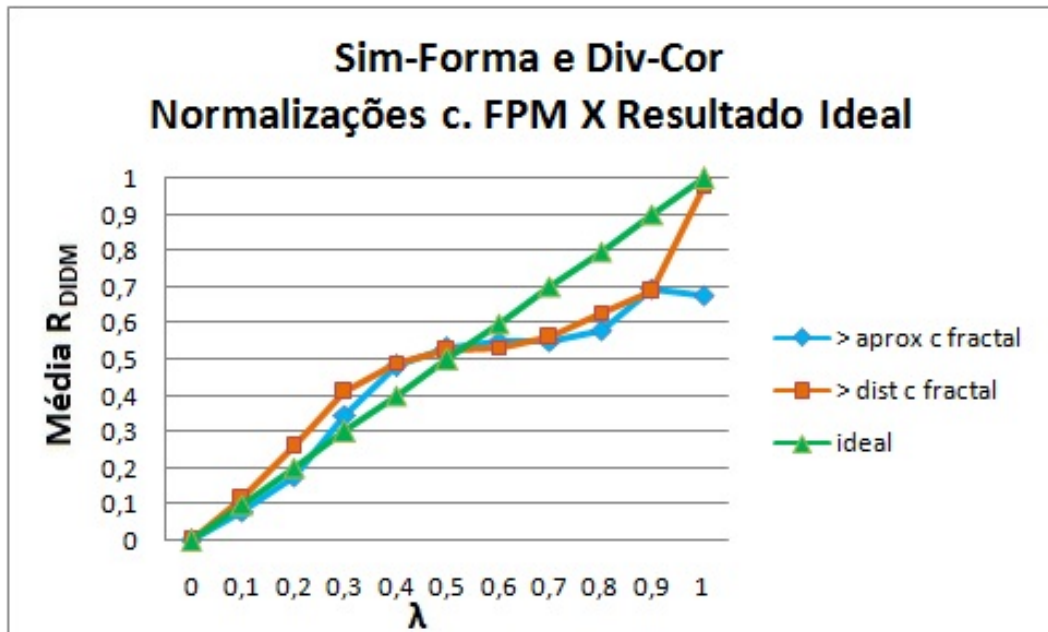
A figura 4.12 apresenta uma comparação da consulta k-NDN considerando similaridade pela forma e diversidade pela cor para normalização pela maior distância, maior distância aproximada com o resultado considerado ideal. Pode-se observar que para todos os valores de λ os resultados foram praticamente os mesmos para os dois tipos de normalização.

Figura 4.12: Comparação dos resultados média R_{DIDM} - Normalização pela maior distância, maior distância aproximada e resultado ideal



A figura 4.13 apresenta uma comparação dos resultados considerados ideal com os resultados da consulta por similaridade pela forma e diversidade pela cor, considerando-se normalização pela maior distância com balanceamento FPM e maior distância aproximada com balanceamento FPM. Pode-se observar que os resultados estiveram sempre próximos do esperado. Para os dois tipos de normalização os resultados se mantiveram nos mesmos patamares.

Figura 4.13: Comparação dos resultados média R_{DIDM} - Normalização pela maior distância com balanceamento FPM, maior distância aproximada com balanceamento FPM e resultado ideal



4.2.3 Experimentos 2b: Similaridade pela Cor e Diversidade pela Forma

A seguir são apresentados os experimentos para a consulta k-NDN considerando o descritor de cor para tratar a similaridade e o descritor de forma para tratar a diversidade.

4.2.3.1 Sem Normalização

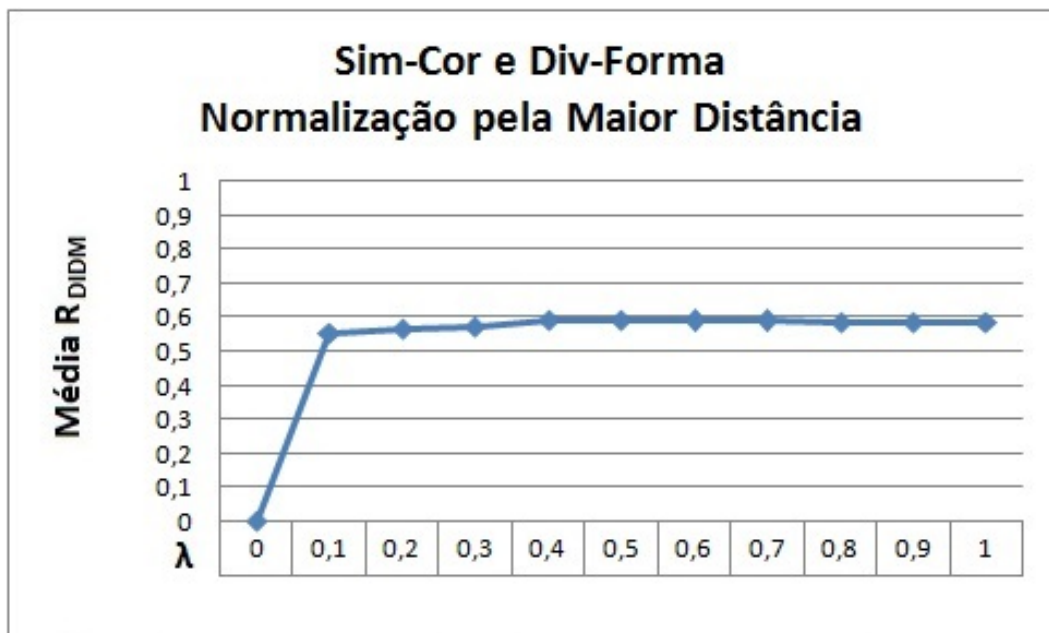
Para a consulta k-NDN considerando similaridade pela cor e diversidade pela forma sem considerar a normalização, as médias R_{DIDM} mantiveram-se sempre abaixo do valor considerado ideal para todos os valores de λ , conforme pode ser observado na figura 4.14.

Figura 4.14: Resultado da média R_{DIDM} para os descritores de cor e forma sem normalização

4.2.3.2 Normalização pela maior distância

Quanto aos resultados da consulta k-NDN considerando o descritor de cor para similaridade e forma para diversidade e normalização pela maior distância, conforme pode ser observado na figura 4.15, os resultados apresentados foram praticamente os mesmos para todos os valores de λ . Pode-se sugerir que um descritor teve uma maior representatividade do que o outro, ou seja, o objetivo da normalização não foi plenamente alcançado.

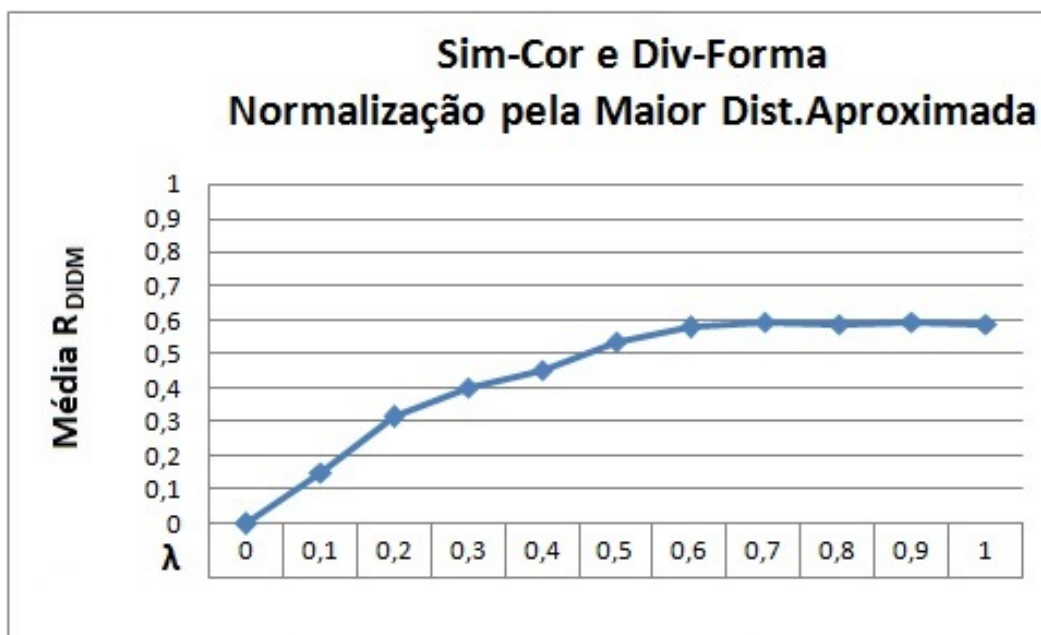
Figura 4.15: Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância



4.2.3.3 Normalização pela maior distância aproximada

Para as consultas k-NDN considerando similaridade pela cor, diversidade pela forma e normalização pela maior distância aproximada, os resultados podem ser observados na figura 4.16. Pode-se observar que os resultados estiveram bem próximos dos resultados esperados.

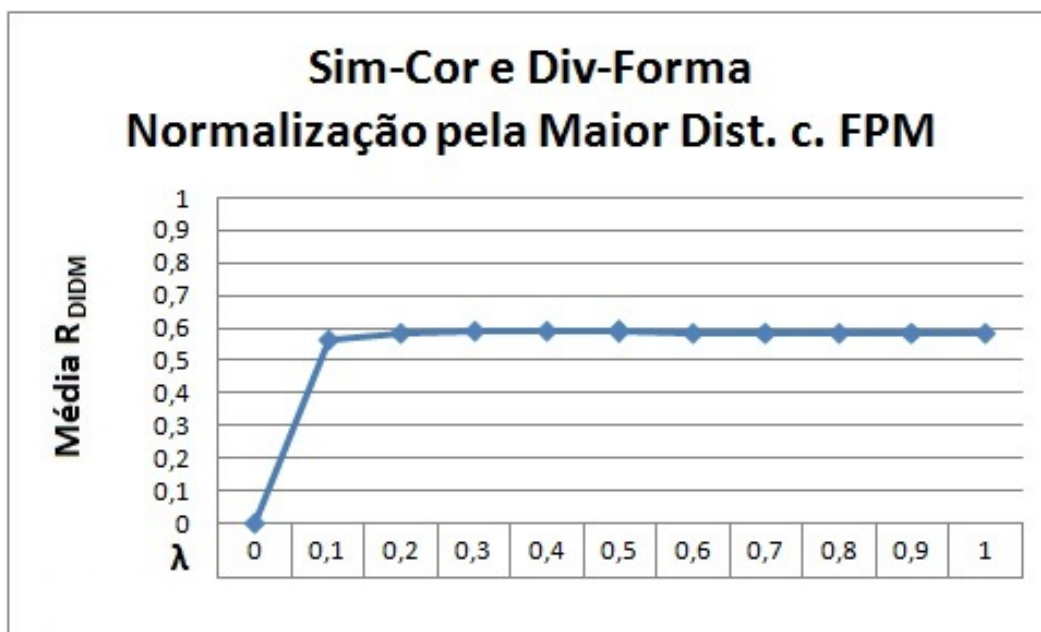
Figura 4.16: Resultado média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância aproximada



4.2.3.4 Normalização pela maior distância com balanceamento FPM

Quanto aos resultados da consulta k-NDNq considerando o descritor de cor para similaridade, forma para diversidade e normalização pela maior distância com balanceamento FPM, conforme pode ser observado na figura 4.17, os resultados apresentados foram praticamente os mesmos para todos os valores de λ , semelhante ao resultado da consulta k-NDNq considerando similaridade pela forma e diversidade pela cor sem normalização conforme apresentado na figura 4.4.

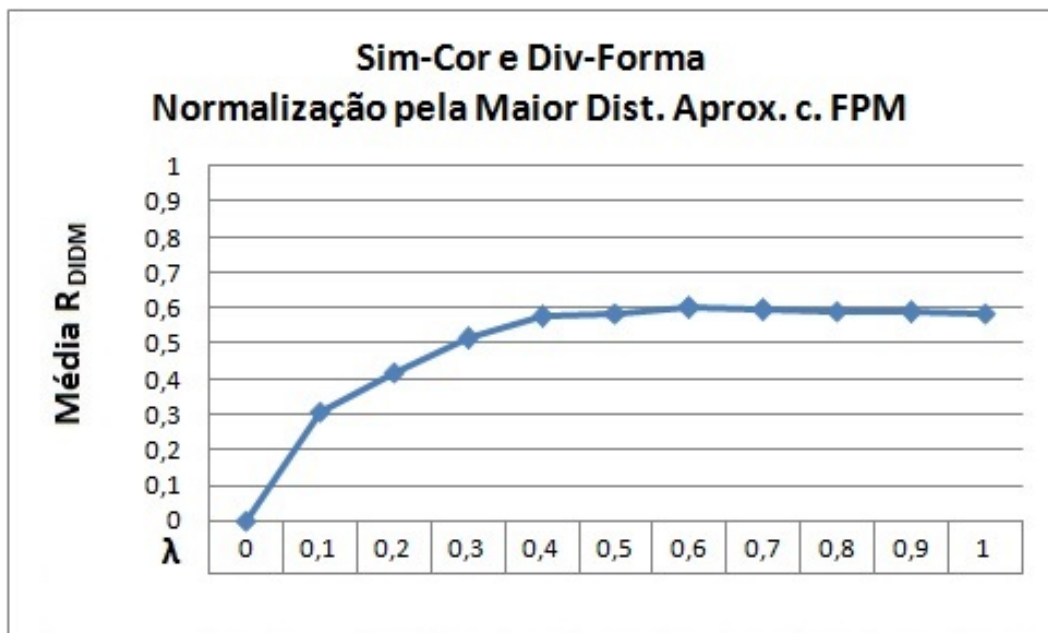
Figura 4.17: Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância com balanceamento FPM



4.2.3.5 Normalização pela maior distância aproximada com balanceamento FPM

Na figura 4.18 pode-se observar os resultados da média R_{DIDM} para todas as consulta k-NDNq considerando o descritor de cor para similaridade, forma para diversidade e normalização pela maior distância aproximada com balanceamento FPM. Neste caso, os resultados estiveram acima do considerado ideal para valores de $\lambda \leq 0,5$, e abaixo do considerado ideal para valores de $\lambda \geq 0,7$. Somente para $\lambda = 0,6$ o resultado foi considerado ideal.

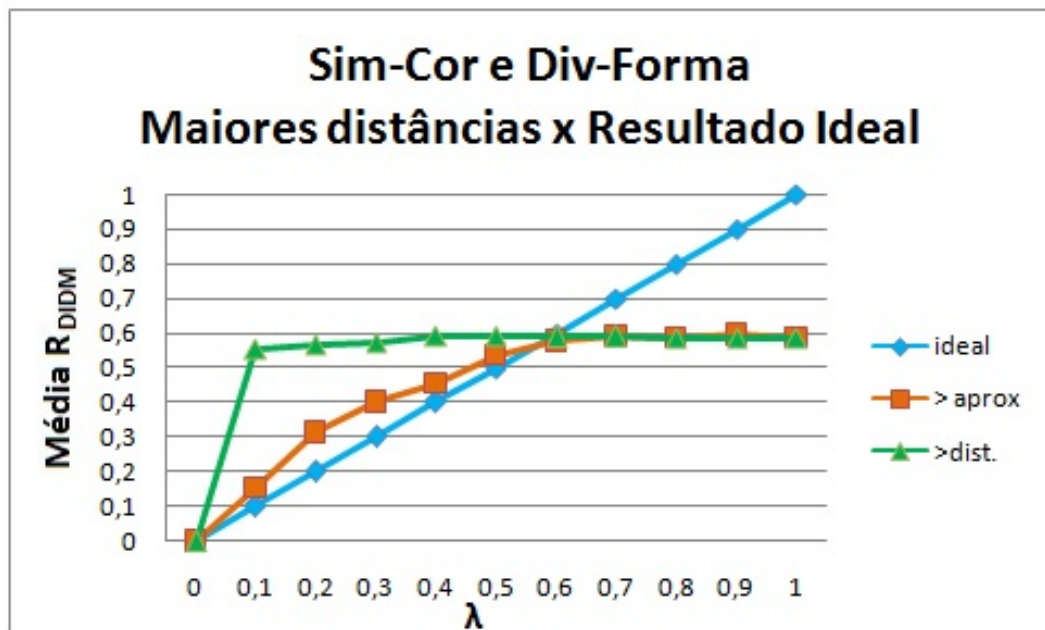
Figura 4.18: Resultado da média R_{DIDM} para os descritores de cor e forma com normalização pela maior distância aproximada com balanceamento FPM



4.2.3.6 Comparação dos resultados para k-NDNq considerando similaridade pela cor e diversidade pela forma

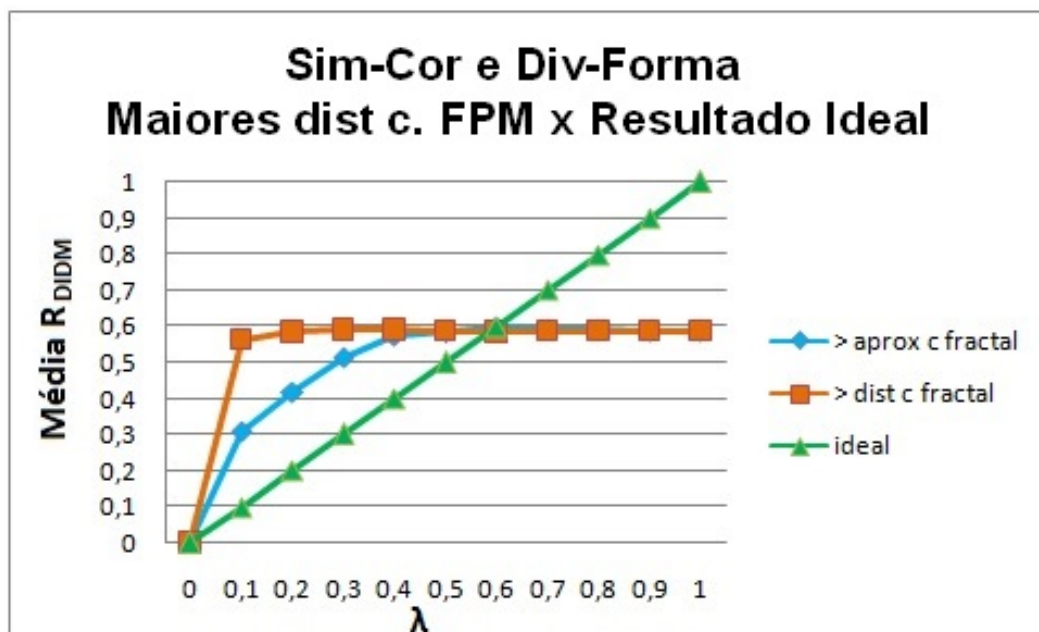
A figura 4.19 apresenta uma comparação dos resultados das consultas k-NDN considerando similaridade pela cor, diversidade pela forma e normalizações pela maior distância, maior distância aproximada e resultado considerado ideal. Neste caso, pode-se observar que os resultados da normalização pela maior distância aproximada acompanharam o resultado esperado para todos os valores de $\lambda \leq 0,6$, enquanto que os resultados da normalização pela maior distância estiveram acima do valor esperado para os mesmos valores de λ . Para valores de $\lambda \geq 0,7$ os resultados mantiveram-se abaixo do resultado ideal para os dois tipos de normalização considerados, o que pode ser observado na maioria dos testes realizados nesta dissertação.

Figura 4.19: Comparação dos resultados média R_{DIDM} entre normalização pela maior distância, maior distância aproximada e resultado ideal



A figura 4.20 apresenta uma comparação dos resultados considerados ideal com os resultados obtidos para as consultas k-NDN considerando similaridade pela cor, diversidade pela forma e normalizações pela maior distância com balanceamento FPM e maior distância aproximada com balanceamento FPM. Neste caso, pode-se observar que os resultados estiveram fora do resultado considerado ideal para a maioria dos valores de λ . Apenas para $\lambda = 0.6$ obteve-se o resultado esperado.

Figura 4.20: Comparação dos resultados média R_{DIDM} entre normalização pela maior distância com balanceamento FPM, maior distância aproximada com balanceamento FPM e resultado ideal



4.3 Resumo dos Resultados

Através dos resultados apresentados pelos experimentos realizados para as consultas k-NDN considerando dois descritores em papéis diferentes, de uma maneira geral observou-se um comportamento muito semelhante em praticamente todos os tipos de experimentos.

Para valores de $\lambda < 0.4$ os resultados mantiveram-se acima do valor esperado na maioria dos experimentos, o que significa dizer que a quantidade de elementos diversos que foram incluídos pela consulta com múltiplos descritores em papéis diferentes superou os valores de λ definidos para a função objetivo proposta nesta dissertação.

No intervalo de $0.4 \leq \lambda \leq 0.6$ considerando-se todos os tipos de normalização adotados a quantidade de elementos diversos que foram incluídos pela consulta com múltiplos descritores em papéis diferentes estiveram dentro do valor esperado.

Para valores de $\lambda \geq 0.7$, os resultados mantiveram-se abaixo do valor considerado ideal na maioria dos experimentos, o que significa dizer que a quantidade de elementos diversos que foram incluídos pela consulta com múltiplos descritores em papéis diferentes foi inferior aos

valores de λ definidos para a função objetivo proposta nesta dissertação. Porém, esse comportamento também ocorreu na aplicação da R_{DIDM} em k-NDN usando um único descritor para similaridade e diversidade.

Em geral, comparando-se os experimentos acima com os resultados obtidos pelos experimentos realizados pela consulta k-NDN considerando-se apenas um descritor, o resultados obtidos foram semelhantes para todos os testes realizados. Pode-se dizer que a avaliação da diversidade incluída realmente acompanha a tendência de λ , o que permite dizer que a normalização e balanceamento é necessário. Além disso, pretende-se futuramente estudar novas formas de normalizar.

4.4 Considerações Finais

Neste capítulo foram apresentados os recursos utilizados para implementar a solução proposta, o método utilizado para validar os experimentos e o ambiente de testes utilizado para avaliar a proposta desta dissertação. Foram descritos os experimentos realizados para validar a proposta. Em seguida apresentou-se uma análise dos resultados obtidos.

Através da análise apresentada pode-se considerar que para as consultas k-NDN utilizando-se dois descritores em papéis diferentes os resultados apresentados acompanharam o valor de λ considerado como sendo o valor esperado quando comparado com os testes para apenas um descritor.

Dessa forma, espera-se que os resultados alcançados possam contribuir para a melhorar as consultas por similaridade melhorando a qualidade dos resultados apresentados ao usuário, diminuindo o ‘gap semântico’. Espera-se assim uma maior satisfação dos usuários.

Capítulo 5

CONSIDERAÇÕES FINAIS

Um dos objetivos deste trabalho foi melhorar a qualidade semântica das respostas em consulta por similaridade buscando-se apresentar resultados similares quanto ao elemento de consulta e diversos entre si, procurando atender melhor as expectativas dos usuários. Para cumprir este objetivo, este trabalho apresentou um modelo que inclui diversidade em consultas por similaridade utilizando múltiplos descritores em papéis diferentes procurando diversificar os resultados apresentados em uma consulta.

Utilizou-se o algoritmo k-NDNq para implementação da proposta alterando a função objetivo para fazer o balanceamento entre similaridade e diversidade considerando-se espaços métricos diferentes e a normalização pela maior distância, maior distância aproximada, maior distância com balanceamento FPM e maior distância aproximada com balanceamento FPM, buscando-se um equilíbrio na representatividade numérica entre os descritores.

Utilizou-se o algoritmo k-NDNq exaustivo para validar a solução. Procurou-se avaliar semanticamente os resultados e para isso os experimentos foram realizados em uma base de dados pré-classificada, para que fosse possível identificar a quantidade de classes incluídas no conjunto resultado final da consulta. Para avaliar os resultados foi proposta a função R_{DIDM} para identificar a quantidade de diversidade incluída na consulta k-NDN em relação a diversidade máxima possível de ser encontrada em k-NN e k .

Através dos testes realizados pode-se perceber que os resultados obtidos acompanharam a

tendência do valor de λ .

Espera-se assim, que os resultados deste trabalho possam melhorar a expressividade das consultas por similaridade gerando resultados que atendam melhor as expectativas do usuário.

5.1 Contribuições

Dentre as contribuições geradas por este trabalho pode-se citar:

1. Uma função que inclui diversidade utilizando múltiplos descritores em contextos diferentes, um para tratar a similaridade e outro para tratar a diversidade. Esta função pode ser aplicada em outros algoritmos existentes na literatura para realizar consultas por similaridade.
2. Métodos de normalização baseados:
 - na maior distância calculada sobre o conjunto total;
 - na maior distância com balanceamento pelo FPM;
 - na maior distância aproximada calculada para o subconjunto da consulta k-NN.
 - na maior distância aproximada com balanceamento pelo FPM;
3. Uma função R_{DIDM} para avaliar a diversidade incluída nas consultas k-NDN.

5.2 Trabalhos Futuros

Como trabalhos para complementar a proposta desta dissertação de Mestrado sugere-se:

1. Aplicar o modelo proposto em outros algoritmos existentes para verificar o comportamento da função objetivo que trata a diversidade através de mais de um descritor;
2. Considerar outros tipos de descritores para avaliar os resultados nas consultas;
3. Estudar outras maneiras para realizar a normalização e o balanceamento entre os múltiplos espaços métricos em k-NDNq;

4. Realizar experimentos em outras bases de testes que tenham uma maior aplicabilidade da solução;
5. Desenvolver uma interface para melhor visualização dos resultados obtidos.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, R.; GOLLAPUDI, S.; HALVERSON, A.; IEONG, S. Diversifying search results. In: ACM. *Proceedings of the second ACM international conference on web search and data mining*. [S.l.], 2009. p. 5–14.

AMATO, G.; RABITTI, F.; SAVINO, P.; ZEZULA, P. Region proximity in metric spaces and its use for approximate similarity search. *ACM Transactions on Information Systems (TOIS)*, ACM, v. 21, n. 2, p. 192–227, 2003.

BARROSO, R.; PONCIANO-SILVA, M.; TRAINA, A.; BUENO, R. Speeding up the combination of multiple descriptors for different boundary conditions. In: CANCELA, H.; CUADROS-VARGAS, A.; CUADROS-VARGAS, E. (Ed.). *2015 XLI Latin American Computing Conference (CLEI)*. Arequipa-Peru: CLEI, 2015. p. 82–92. ISBN 978-1-4673-9143-6. Disponível em: <<http://clei.org/clei2015/144721>>.

BARROSO, R. F.; PONCIANO-SILVA, M.; TRAINA, A. J. M.; BUENO, R. Using boundary conditions for combining multiple descriptors in similarity based queries. In: SPRINGER. *Iberoamerican Congress on Pattern Recognition*. [S.l.], 2013. p. 375–382.

BOSCH, A.; ZISSERMAN, A.; MUNOZ, X. Representing shape with a spatial pyramid kernel. In: *Proceedings of the 6th ACM International Conference on Image and Video Retrieval*. New York, NY, USA: ACM, 2007. (CIVR '07), p. 401–408. ISBN 978-1-59593-733-9. Disponível em: <<http://doi.acm.org/10.1145/1282280.1282340>>.

BUENO, R. *Tratamento de tempo e dinamicidade em dados representados em espaços métricos*. Tese (Doutorado) — Universidade de São Paulo, 2009.

BUENO, R.; KASTER, D. S.; PATERLINI, A. A.; TRAINA, A. J.; TRAINA, C. Unsupervised scaling of multi-descriptor similarity functions for medical image datasets. In: IEEE. *Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on*. [S.l.], 2009. p. 1–8.

BUENO, R.; RIBEIRO, M. X.; TRAINA, A. J.; TRAINA, C. Improving medical image retrieval through multi-descriptor similarity functions and association rules. In: IEEE. *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*. [S.l.], 2010. p. 309–314.

BUGATTI, P. H.; TRAINA, A. J. M.; TRAINA JR., C. Assessing the best integration between distance-function and image-feature to answer similarity queries. In: *Proceedings of the 2008 ACM Symposium on Applied Computing*. New York, NY,

USA: ACM, 2008. (SAC '08), p. 1225–1230. ISBN 978-1-59593-753-7. Disponível em: <<http://doi.acm.org/10.1145/1363686.1363969>>.

BUSTOS, B.; KEIM, D.; SAUPE, D.; SCHRECK, T.; VRANIC, D. Automatic selection and combination of descriptors for effective 3d similarity search. In: IEEE. *Multimedia Software Engineering, 2004. Proceedings. IEEE Sixth International Symposium on*. [S.l.], 2004. p. 514–521.

CARBONELL, J.; GOLDSTEIN, J. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In: ACM. *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 1998. p. 335–336.

CASTELLS, P.; VARGAS, S.; WANG, J. Novelty and diversity metrics for recommender systems: choice, discovery and relevance. 2011.

CIACCIA, P.; PATELLA, M.; ZEZULA, P. M-tree: An efficient access method for similarity search in metric spaces. In: *Proceedings of the 23rd International Conference on Very Large Data Bases*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997. (VLDB '97), p. 426–435. ISBN 1-55860-470-7. Disponível em: <<http://dl.acm.org/citation.cfm?id=645923.671005>>.

CLARKE, C. L.; KOLLA, M.; CORMACK, G. V.; VECHTOMOVA, O.; ASHKAN, A.; BÜTTCHER, S.; MACKINNON, I. Novelty and diversity in information retrieval evaluation. In: ACM. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2008. p. 659–666.

DATTA, R.; JOSHI, D.; LI, J.; WANG, J. Z. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (Csur)*, ACM, v. 40, n. 2, p. 5, 2008.

DIAS, R. L.; BUENO, R.; RIBEIRO, M. X. Reducing the complexity of k-nearest diverse neighbor queries in medical image datasets through fractal analysis. In: IEEE. *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. [S.l.], 2013. p. 101–106.

DROSOU, M.; PITOURA, E. Search result diversification. *ACM SIGMOD Record*, ACM, v. 39, n. 1, p. 41–47, 2010.

FALOUTSOS, C. Indexing of multimedia data. In: *Multimedia Databases in Perspective*. [S.l.]: Springer, 1997. p. 219–245.

FALOUTSOS, C.; KAMEL, I. Beyond uniformity and independence: Analysis of r-trees using the concept of fractal dimension. In: *Proceedings of the Thirteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. New York, NY, USA: ACM, 1994. (PODS '94), p. 4–13. ISBN 0-89791-642-5. Disponível em: <<http://doi.acm.org/10.1145/182591.182593>>.

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.

- GUDEWAR, A. D.; RAGHA, L. R. Ontology to improve cbir system. *International Journal of Computer Applications*, Foundation of Computer Science, v. 52, n. 21, 2012.
- KUO, C.-C.; GLOVER, F.; DHIR, K. S. Analyzing and modeling the maximum diversity problem by zero-one programming. *Decision Sciences*, Wiley Online Library, v. 24, n. 6, p. 1171–1185, 1993.
- LEUKEN, R. H. van; GARCIA, L.; OLIVARES, X.; ZWOL, R. van. Visual diversification of image search results. In: ACM. *Proceedings of the 18th international conference on World wide web*. [S.l.], 2009. p. 341–350.
- LIU, Y.; ZHANG, D.; LU, G.; MA, W. ying. *A survey of content-based image retrieval with high-level semantics*. 2007.
- RAZENTE, H. L. *Adequando consultas por similaridade para reduzir a descontinuidade semântica na recuperação de imagens por conteúdo*. Tese (Doutorado) — Universidade de São Paulo, 2009.
- RUI, Y.; HUANG, T. S.; CHANG, S.-F. Image retrieval: Current techniques, promising directions, and open issues. *Journal of visual communication and image representation*, Elsevier, v. 10, n. 1, p. 39–62, 1999.
- SANTOS, L. F.; BEDO, M. V.; PONCIANO-SILVA, M.; TRAINA, A. J.; TRAINA, C. Being similar is not enough: how to bridge usability gap through diversity in medical images. In: IEEE. *Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on*. [S.l.], 2014. p. 287–293.
- SANTOS, L. F.; OLIVEIRA, W. D.; CARVALHO, L. O.; FERREIRA, M. R.; TRAINA, A. J.; JR, C. T. Combine-and-conquer: improving the diversity in similarity search through influence sampling. In: ACM. *Proceedings of the 30th Annual ACM Symposium on Applied Computing*. [S.l.], 2015. p. 994–999.
- SANTOS, L. F.; OLIVEIRA, W. D.; FERREIRA, M. R.; CORDEIRO, R. L.; TRAINA, A. J.; JR, C. T. Evaluating the diversification of similarity query results. *Journal of Information and Data Management*, v. 4, n. 3, p. 188, 2013.
- SANTOS, L. F.; OLIVEIRA, W. D.; FERREIRA, M. R.; TRAINA, A. J.; JR, C. T. Parameter-free and domain-independent similarity search with diversity. In: ACM. *Proceedings of the 25th International Conference on Scientific and Statistical Database Management*. [S.l.], 2013. p. 5.
- SHAPIRO, L.; STOCKMAN, G. C. *Computer vision*. 2001. ed: Prentice Hall, 2001.
- SMEULDERS, A. W.; WORRING, M.; SANTINI, S.; GUPTA, A.; JAIN, R. Content-based image retrieval at the end of the early years. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 22, n. 12, p. 1349–1380, 2000.
- SMITH, J. R.; CHANG, S.-F. Visualseek: a fully automated content-based image query system. In: ACM. *Proceedings of the fourth ACM international conference on Multimedia*. [S.l.], 1997. p. 87–98.
- TORRES, R. da S.; FALCAO, A. X. Content-based image retrieval: theory and applications. *RITA*, v. 13, n. 2, p. 161–185, 2006.

TRAINA, A. J. M.; TRAINA, C.; BUENO, J. M.; CHINO, F. J. T.; AZEVEDO-MARQUES, P. Efficient content-based image retrieval through metric histograms. *World Wide Web*, Springer, v. 6, n. 2, p. 157–185, 2003.

TRAINA, C.; TRAINA, A.; FALOUTSOS, C.; SEEGER, B. Fast indexing and visualization of metric data sets using slim-trees. *IEEE Transactions on Knowledge and Data Engineering*, IEEE, v. 14, n. 2, p. 244–260, 2002.

TRAINA, C.; TRAINA, A.; FALOUTSOS, C. et al. How to improve the pruning ability of dynamic metric access methods. In: ACM. *Proceedings of the eleventh international conference on Information and knowledge management*. [S.l.], 2002. p. 219–226.

VIEIRA, M. R.; RAZENTE, H. L.; BARIONI, M. C.; HADJIELEFThERIOU, M.; SRIVASTAVA, D.; JR, C. T.; TSOTRAS, V. J. Divdb: A system for diversifying query results. *Proceedings of the VLDB Endowment*, v. 4, n. 12, p. 1395–1398, 2011.

VIEIRA, M. R.; RAZENTE, H. L.; BARIONI, M. C.; HADJIELEFThERIOU, M.; SRIVASTAVA, D.; TRAINA, C.; TSOTRAS, V. J. On query result diversification. In: IEEE. *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*. [S.l.], 2011. p. 1163–1174.

ZHANG, M.; HURLEY, N. Avoiding monotony: improving the diversity of recommendation lists. In: ACM. *Proceedings of the 2008 ACM conference on Recommender systems*. [S.l.], 2008. p. 123–130.

ZHANG, Y.; CALLAN, J.; MINKA, T. Novelty and redundancy detection in adaptive filtering. In: ACM. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2002. p. 81–88.