

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MINERAÇÃO VISUAL DE IMAGENS ALIADA A
CONSULTAS PELOS K-VIZINHOS DIVERSOS
MAIS PRÓXIMOS: FLEXIBILIZANDO E
MAXIMIZANDO O ENTENDIMENTO DE
CONSULTAS POR CONTEÚDO DE IMAGENS**

RAFAEL LOOSLI DIAS

ORIENTADORA: PROFA. DRA. MARCELA XAVIER RIBEIRO

São Carlos – SP

Fevereiro/2014

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MINERAÇÃO VISUAL DE IMAGENS ALIADA A
CONSULTAS PELOS K-VIZINHOS DIVERSOS
MAIS PRÓXIMOS: FLEXIBILIZANDO E
MAXIMIZANDO O ENTENDIMENTO DE
CONSULTAS POR CONTEÚDO DE IMAGENS**

RAFAEL LOOSLI DIAS

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

Orientadora: Profa. Dra. Marcela Xavier Ribeiro

São Carlos – SP

Fevereiro/2014

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

D541mv

Dias, Rafael Loosli.

Mineração visual de imagens aliada a consultas pelos k-vizinhos diversos mais próximos : flexibilizando e maximizando o entendimento de consultas por conteúdo de imagens / Rafael Loosli Dias. -- São Carlos : UFSCar, 2014. 92 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2013.

1. Data mining (Mineração de dados). 2. Recuperação de imagem. 3. *Gap* semântico. 4. Similaridade. 5. Diversidade. 6. Mineração visual de dados. I. Título.

CDD: 005.741 (20^a)


Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

**“Mineração visual de imagens aliada a consultas
pelos k-vizinhos diversos mais próximos:
flexibilizando e maximizando o entendimento de
consultas por conteúdo de imagens”**

Rafael Loosli Dias

Dissertação de Mestrado 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

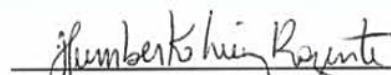
Membros da Banca:



Prof. Dra. Marcela Xavier Ribeiro
(Orientadora - DC/UFSCar)



Prof. Dr. Renato Bueno
(DC/UFSCar)



Prof. Dr. Humberto Luiz Razente
(UFU)

São Carlos
Agosto/2013

A todos aqueles que, diretamente ou indiretamente, permitiram a conclusão deste trabalho.

AGRADECIMENTOS

Aos meus pais Martha e Guilherme e minha irmã Luiza, por participarem desta etapa de minha vida e serem exemplos de pessoas a serem seguidas.

À Simone, por estar sempre ao meu lado e compartilhar os momentos de alegria e de dificuldade na trajetória deste projeto.

Aos meus parentes que, apesar da distância, sempre apoiaram e incentivaram meus estudos.

À minha orientadora Marcela, por me ensinar, direcionar e auxiliar em todas as etapas deste projeto de mestrado.

Ao prof. Renato, pelo apoio e o levantamento de ótimas ideias adicionadas neste trabalho.

Aos meus amigos de mestrado - Elias, Du, Papotti, César, Carlos, Diego, Vitinho, Mirela, Marcos e Francisco - pelo companheirismo nas atividades realizadas dentro e fora do laboratório de estudos.

Aos meus amigos de São Carlos - Apoio, Vitor, Pedrosa, Batman, Robin, Pizi, Kim, Mister e Uísque - pelos momentos de diversão vivenciados.

Aos meus amigos de Campinas - Artur, Thiago, Glauber, Giulia e Maria - pela amizade e presença em outros momentos importantes de minha vida.

Ao Departamento de Computação da Universidade Federal de São Carlos (DC - UFSCar), pelo espaço proporcionado para estudo e auxílio na apresentação dos resultados obtidos.

À CAPES pelo auxílio financeiro.

É sábio o homem que põs em si tudo que leva à felicidade ou dela se aproxima.

Sócrates

RESUMO

Sistemas de recuperação de imagens por conteúdo (do Inglês, Content-Based Image Retrieval - CBIR) utilizam informações visuais de cor, forma e textura para representar as imagens em vetores de características. A representação numérica encontrada para as imagens é utilizada na execução da consulta através de uma métrica que avalie a distância entre os vetores. Em geral, existe uma inconsistência entre a percepção do ser humano na avaliação de similaridade entre imagens se comparada com a computada por sistemas CBIR, sendo esta descontinuidade denominada Gap Semântico. Adicionar um fator de diversidade na consulta tem-se mostrado como uma maneira de superar este problema, permitindo que o usuário especifique o grau de dissimilaridade entre as imagens resultantes e altere o resultado da consulta. Adicionar diversidade em consulta, no entanto, requer alto custo computacional e a redução das possibilidades de conjuntos para resposta é de difícil entendimento para o usuário. Este trabalho de mestrado propôs a utilização de técnicas de Mineração Visual de Dados (MVD) aplicadas sobre consultas em sistemas CBIR, melhorando a interpretabilidade da medida de similaridade e diversidade, assim como a relevância do resultado obtido. O usuário passa a exercer um papel ativo na consulta por conteúdo de imagens, permitindo que o mesmo dirija o processo, aproximando o resultado ao esperado pela cognição humana e reduzindo o gap semântico.

Palavras-chave: Recuperação de Imagens por Conteúdo, Gap Semântico, Similaridade, Diversidade, Mineração Visual de Dados, Mineração Visual de Imagens

ABSTRACT

Content-Based Image Retrieval systems use visual information like color, shape and texture to represent images in feature vectors. The numerical representation found for the images is used in query execution through a metric to evaluate the distance between vectors. In general, there is an inconsistency in the evaluation of similarity between images according to human perception and the results computed by CBIR systems, which is called Semantic Gap. One way to overcome this problem is by the addition of a diversity factor in query execution, allowing the user to specify a degree of dissimilarity between the resulting images and changing the query result. Adding diversity in consultation, however, requires high computational cost and the reduction of possible subsets to be analyzed is a difficult task to be understood by the user. This masters degree thesis aims to make use of Visual Data Mining techniques applied to queries in CBIR systems, improving the interpretability of the measure of similarity and diversity, as well as the relevance of the result according to the judgment and prior knowledge of the user. The user takes an active role in the retrieval of images by their content, guiding its result and, consequently, reducing the Semantic Gap. Additionally, a better understanding of the diversity and similarity factors involved in the query is supported by visualization and interaction techniques.

Keywords: Content-Based Image Retrieval, Semantic Gap, Similarity, Diversity, Visual Data Mining

LISTA DE FIGURAS

2.1	Exemplo de consulta utilizando Range Query.	22
2.2	Exemplo de consulta utilizando k NN.	23
2.3	Pontos equidistantes situados a uma distância r do objeto O , de acordo com as métricas L_1 , L_2 e L_∞	24
2.4	Gráfico de Revocação X Precisão.	27
3.1	Imagem retirada de (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Ilustração das cinco etapas envolvidas no Processo de Descoberta de Conhecimento em Base de Dados (KDD).	33
3.2	Imagem adaptada de (KEIM et al., 2002). Técnicas de Mineração Visual de Dados.	35
3.3	Imagem retirada de (CLEVELAND, 1979). Técnica de Visualização <i>Scatterplots</i> para dados gerados artificialmente e aproximação de seus valores para uma curva.	37
3.4	Imagem retirada de (INSELBERG; DIMSDALE, 1990). Técnica de visualização por coordenadas paralelas.	38
3.5	Imagem retirada de (ROBERTSON; MACKINLAY; CARD, 1991). Visualização proposta pela técnica Cone-Tree.	39
3.6	Imagem retirada de (BIER et al., 1993). Exemplo de utilização da ferramenta <i>Magic Lenses</i> juntamente com o recurso de colorir o objeto selecionado.	40
3.7	Imagem retirada de (BIER et al., 1993). Exemplo de utilização da ferramenta <i>Magic Lenses</i> onde são utilizadas duas "lentes" para filtrar a área desejada.	41
3.8	Imagem retirada de (BRANDOLI et al., 2010). Ilustração do sistema proposto.	43
3.9	Imagem retirada de (BRANDOLI et al., 2010). Exemplo de visualização de 7 classes de imagens distintas, sendo que 5 delas apresentam boa identificação e 2 encontram-se misturadas.	44

3.10	Imagem retirada de (ELER et al., 2009). Ilustração de uma possível projeção gerada pela ferramenta <i>PEx-Image</i> , a qual permite verificar a árvore filogenética e visualizar imagens pertencentes a uma dada evolução (ramificação).	45
3.11	Imagem adaptada de (JOHNSON; SHNEIDERMAN, 1991). Exemplo da aplicação da técnica <i>nested tree-map</i> (a) e da <i>tree-map</i> convencional (b).	47
4.1	Passos do método proposto para a consulta pelos vizinhos diversos mais próximos (<i>k</i> -NDN), considerando o espaço <i>N</i> -Dimensional.	52
4.2	Passos do método proposto para a consulta pelos vizinhos diversos mais próximos (<i>k</i> -NDN), considerando a utilização das técnicas de MVD.	54
4.3	Técnicas de visualização e interação empregadas no trabalho desenvolvido. . .	55
4.4	Exemplo de movimentação de instâncias na projeção inicial.	59
4.5	Exemplo de visualização do identificador ou prévia de imagem relativa à instância requisitada pelo usuário.	60
4.6	Técnicas de visualização e interação utilizadas na projeção final.	61
4.7	Exploração da consulta <i>k</i> -NDN aumentando o valor do <i>trade-off</i> entre similaridade e diversidade.	62
4.8	Exemplo de aplicação do trabalho desenvolvido.	62
5.1	Imagem de consulta utilizada para o estudo de caso do conjunto de imagens Balan704.	67
5.2	Resultados para o conjunto de imagens Balan704 alterando-se a compensação entre 0 e 0,9. Imagens do conjunto de resposta pertencentes à categoria Cabeça Axial, com exceção da assinalada em destaque (Cabeça Sagital).	68
5.3	Abrangência da consulta pelos vizinhos diversos mais próximos (<i>k</i> NDN) no estudo de caso para o conjunto Balan704.	69
5.4	Visualização de imagens relativas a instâncias no estudo de caso: (a) Cabeça Axial e (b) Cabeça Sagital.	70
5.5	Visualização de centróides pertencentes a cada categoria do conjunto Balan704.	71
5.6	Imagem de consulta utilizada para o estudo de caso do conjunto de imagens CT_Pulmão_ROIs.	72

5.7	Resultados para o conjunto de imagens CT_Pulmão_ROIs alterando-se a compensação entre 0 e 0,9. Imagens no conjunto de resposta pertencentes à categoria Consolidação, com exceção das assinaladas em destaque (Favo de Mel).	73
5.8	Exploração da vizinhança da consulta no estudo de caso para o conjunto CT_Pulmão_ROIs.	74
5.9	Imagem de consulta utilizada para o estudo de caso do conjunto de imagens FullFaces.	75
5.10	Resultados para $\lambda = 0$ considerando o espaço N-Dimensional para o conjunto FullFaces.	75
5.11	Projeção inicial e visualização de imagens do conjunto FullFaces.	76
5.12	Projeção inicial após distorções e manipulações para o conjunto FullFaces.	77
5.13	Resultados para $\lambda = 0$ considerando o espaço de projeção para o conjunto FullFaces.	77
5.14	Gráfico de Revocação e Precisão para comparação dos resultados obtidos para o conjunto FullFaces no experimento realizado.	78
5.15	Resultados comparativos do espaço N-Dimensional e de projeção para o conjunto de imagens FullFaces. A compensação é alterada no intervalo entre 0 e 0,9. Em cada resultado, são evidenciadas as imagens retornadas para o espaço N-Dimensional (à esquerda) e de projeção (à direita).	79

LISTA DE TABELAS

2.1	Levantamento, em porcentagem, de revocação e precisão.	27
5.1	Categorias presentes no conjunto de imagens Balan704.	67
5.2	Categorias presentes no conjunto de imagens CT_Pulmão_ROIs.	72
5.3	Levantamento, em porcentagem, de revocação e precisão para o espaço N-Dimensional do conjunto FullFaces.	78
5.4	Levantamento, em porcentagem, de revocação e precisão para o espaço de projeção do conjunto FullFaces.	78
5.5	Resultados para o conjunto Balan704 considerando o algoritmo sequencial. . .	80
5.6	Resultados para o conjunto Balan704 considerando o algoritmo concorrente. . .	80
5.7	Resultados para o conjunto CT_Pulmão_ROIs considerando o algoritmo sequencial.	81
5.8	Resultados para o conjunto CT_Pulmão_ROIs considerando o algoritmo concorrente.	81
5.9	Resultados para o conjunto FullFaces considerando o algoritmo sequencial. . .	81
5.10	Resultados para o conjunto FullFaces considerando o algoritmo concorrente. . .	81
5.11	Resultados do ganho computacional entre o algoritmo sequencial e concorrente para o conjunto FullFaces. Proporção dada pelo custo do algoritmo sequencial pelo concorrente.	82

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO	15
1.1 Considerações Iniciais	15
1.2 Motivação	16
1.3 Organização do Trabalho	17
1.4 Considerações Finais	17
I Conceitos e Embasamento Teórico	18
CAPÍTULO 2 – CONSULTA POR CONTEÚDO DE IMAGENS (CBIR)	19
2.1 Considerações Iniciais	19
2.2 Representação de Imagens	20
2.2.1 Extração de Características de Imagens	20
2.3 Consulta por Similaridade	22
2.4 Funções de Distância	23
2.5 Indexação de Imagens	24
2.6 Avaliação dos Resultados	25
2.6.1 Revocação e Precisão	25
2.7 Descontinuidade Semântica (<i>Gap Semântico</i>)	28
2.8 Fator Diversidade	29
2.9 Considerações Finais	31

CAPÍTULO 3 – MINERAÇÃO VISUAL DE DADOS	32
3.1 Considerações Iniciais	32
3.2 Técnicas de Mineração Visual de Dados	34
3.2.1 Classificação pela Informação a ser Visualizada	34
3.2.2 Classificação por Técnicas de Visualização	36
3.2.3 Classificação por Técnicas de Interação e Distorção	38
3.3 Mineração Visual de Imagens	41
3.3.1 Visualização Baseada em Transformações Geométricas	41
3.3.2 Visualização Baseada em Hierarquia	44
3.3.3 Avaliação da Qualidade da Visualização	47
3.4 Considerações Finais	48
II Trabalho Desenvolvido	49
CAPÍTULO 4 – TRABALHO DESENVOLVIDO	50
4.1 Considerações Iniciais	50
4.2 Detalhamento do Trabalho Desenvolvido	51
4.2.1 Dimensão Fractal	55
4.2.2 <i>k</i> -Nearest Diverse Neighbor Query (<i>k</i> -NDNq)	57
4.2.3 Técnicas de Mineração Visual de Dados	59
4.3 Limitações do Trabalho Desenvolvido	63
4.4 Considerações Finais	64
CAPÍTULO 5 – RESULTADOS OBTIDOS	65
5.1 Considerações Iniciais	65
5.2 Estudos de Caso	66
5.2.1 Estudo de Caso 1 (Balan704 Dataset)	66

5.2.2	Estudo de Caso 2 (CT_Pulmão_ROIs Dataset)	71
5.2.3	Estudo de Caso 3 (FullFaces Dataset)	75
5.3	Custo Computacional	79
5.4	Considerações Finais	82
III Conclusões e Trabalhos Futuros		83
CAPÍTULO 6 – CONCLUSÕES E TRABALHOS FUTUROS		84
6.1	Considerações Iniciais	84
6.2	Principais Contribuições	85
6.3	Trabalhos Futuros	87
6.4	Considerações Finais	87
REFERÊNCIAS		88
GLOSSÁRIO		93

Capítulo 1

INTRODUÇÃO

Este capítulo apresenta a introdução para este trabalho de mestrado, sendo dividido da seguinte maneira: na Seção 1.1 são dadas as considerações iniciais, também contextualizando o trabalho realizado. A Seção 1.2 apresenta a motivação para a realização deste projeto, indicando a contribuição do trabalho em questão. A Seção 1.3 levanta a organização da monografia e, finalmente, na Seção 1.4 são dadas as considerações finais deste capítulo.

1.1 Considerações Iniciais

A crescente geração de dados e o armazenamento de informações são alguns dos principais motivos para a existência e aplicação de técnicas computacionais. Algoritmos de Mineração de Dados (MD) são amplamente utilizados para a extração de conhecimento e descoberta de padrões em bases com grande quantidade de dados.

Um grande problema de trabalhar em um cenário em que os dados são compostos por imagens é a complexidade presente nas mesmas. Imagens possuem diversas características visuais relacionadas a cor, forma e textura. Para que possam ser trabalhadas computacionalmente, é preciso encontrar uma representação numérica, sendo este um processo que acaba, muitas vezes, se distanciando da percepção humana. Esta diferença entre a representação encontrada por métodos computacionais e a cognição humana é denominada Gap Semântico.

Sistemas de Recuperação de Imagens por Conteúdo (*Content-Based Image retrieval systems* - CBIR) permitem que sejam realizadas consultas e, como resultado, são retornadas as imagens com maior grau de similaridade em relação à imagem de consulta. Para verificar a similaridade entre as imagens da base e a imagem de consulta, características visuais referentes a cor, forma e/ou textura são extraídas e armazenadas em uma estrutura denominada vetor de ca-

racterísticas. A similaridade entre imagens é então verificada através de uma métrica estipulada, avaliando a distância entre os vetores levantados.

A Mineração Visual de Dados (MVD) permite que, além do caráter automático proporcionado pelos algoritmos de MD convencionais, sejam adicionados recursos visuais e interativos, aproximando os resultados da cognição humana. Desta forma, o usuário passa a atuar ativamente no processo de extração de conhecimento e no descobrimento de padrões em bases de imagens.

1.2 Motivação

Os sistemas CBIR possuem o Gap Semântico como uma das suas principais limitações. Imagens possuem natureza complexa, sendo difícil avaliar a similaridade entre as mesmas devido ao grande número de atributos a serem analisados. Desta forma, a representação numérica obtida através de características visuais automaticamente extraídas proporciona um resultado que, muitas vezes, difere do esperado pela cognição humana.

Adicionar um fator de diversidade nas consultas por similaridade tem-se mostrado como uma tentativa de reduzir os problemas gerados pelo Gap Semântico. A partir deste princípio, é considerada determinada dissimilaridade entre as imagens do conjunto resultante, adicionando uma relação não somente entre a imagem consulta e as demais imagens, mas também entre os elementos do subconjunto resultante. O valor considerado na relação entre similaridade e diversidade pode ser fornecido pelo usuário, permitindo que o resultado esteja mais próximo de sua expectativa.

Utilizar o *feedback* e o conhecimento do usuário permite que o mesmo deixe de ser um agente passivo na consulta por conteúdo de imagens, atuando de forma ativa para a obtenção de um resultado mais promissor. Técnicas de MVD permitem a utilização do poder computacional para avaliar as imagens de forma mais objetiva e ágil e a interação do usuário para possibilitar a melhora da semântica do resultado. Empregar técnicas de visualização e interação permitem não só que o resultado seja dirigido de acordo com julgamentos e conhecimento prévio do usuário com relação ao conjunto de dados, mas podem também proporcionar melhor compreensão de fatores considerados no resultado.

Em diversas áreas do conhecimento, incluindo o campo de atuação da medicina, imagens são constantemente utilizadas para armazenar diferentes informações referentes ao estado momentâneo de dado contexto. Desta forma, mostra-se conveniente utilizar métodos automáticos para verificar relacionamentos entre características apresentadas em imagens, além de padrões

encontrados nas mesmas. A adição de recursos de visualização e interação para o usuário, neste contexto, permite que os resultados encontrados por métodos automáticos se aproximem da expectativa do usuário, promovendo também uma melhor compreensão dos fatores envolvidos na obtenção dos mesmos.

1.3 Organização do Trabalho

Este trabalho encontra-se organizado da seguinte maneira: No Capítulo 2, são dados os conceitos acerca dos sistema CBIR e o fator de diversidade na consulta, os quais, dada uma imagem de consulta, permitem encontrar as imagens mais similares respeitando a diversidade empregada no subconjunto resultante; O Capítulo 3 apresenta o conceito de MVD, além das técnicas envolvidas pelo mesmo; No Capítulo 4 são discutidos os conceitos e o desenvolvimento deste trabalho de mestrado; No Capítulo 5 são ilustrados os resultados obtidos através da metodologia desenvolvida, evidenciando as vantagens de utilização da mesma; e, no Capítulo 6 é dada a conclusão acerca deste trabalho de mestrado, indicando também possibilidades de trabalhos futuros.

1.4 Considerações Finais

Neste capítulo foi apresentada uma introdução para a importância da utilização dos recursos de diversidade e MVD em consultas de imagens por conteúdo. O problema do Gap Semântico pode ser reduzido pela participação ativa do usuário na tarefa de exploração do conhecimento. Logo, este trabalho utiliza técnicas de MVD para proporcionar melhor eficácia da busca por objetos similares e diversos em bases de imagens, permitindo a participação ativa do usuário na obtenção dos resultados e uma melhor compreensão dos fatores envolvidos nos mesmos.

Parte I

Conceitos e Embasamento Teórico

Capítulo 2

CONSULTA POR CONTEÚDO DE IMAGENS (CBIR)

Os sistemas de consulta por conteúdo de imagens (Content-Based Image Retrieval - CBIR) permitem que imagens sejam relacionadas de acordo com a similaridade verificada entre as mesmas. Na Seção 2.1 deste capítulo serão discutidas as considerações iniciais acerca do funcionamento dos sistemas CBIR. A Seção 2.2 apresenta técnicas para obtenção da representação numérica de imagens. Na Seção 2.3 são discutidas as abordagens utilizadas na consulta por similaridade entre imagens. A Seção 2.4 define os conceitos acerca de funções de distância, sendo estas necessárias para a verificação da similaridade entre imagens. Na Seção 2.5 são discutidos os conceitos acerca da indexação de imagens. A Seção 2.6 demonstra técnicas de medida para a avaliação de resultados obtidos na tarefa de busca por similaridade. A Seção 2.7 aborda o conceito da descontinuidade semântica (Gap Semântico). Na Seção 2.8 é explicado o contexto em que o fator diversidade é adicionado em sistemas CBIR e, na Seção 2.9, são dadas as considerações finais deste capítulo.

2.1 Considerações Iniciais

Em diversas áreas do conhecimento, a quantidade de imagens geradas e armazenadas é crescente, gerando maior demanda de ferramentas que permitem a análise automática das mesmas.

Imagens são consideradas dados complexos e, em geral, para a sua manipulação, precisam ser transformadas em informações mais simples. Os sistemas CBIR realizam consultas em uma base de imagens com base no critério de semelhança. Esses sistemas envolvem um conjunto de métodos que processam as imagens a fim de obter a representação das mesmas em vetores de características. Esses vetores são utilizados, no lugar das imagens, para a execução de consultas. Em geral, os sistemas CBIR envolvem o pré-processamento da imagem, a extração de

características, a avaliação de similaridade entre as imagens e técnicas de indexação de imagens. O resultado de uma consulta em um sistema CBIR é um conjunto de imagens ordenadas de acordo com sua similaridade à imagem de consulta.

Neste trabalho de mestrado, os sistemas de CBIR têm sua aplicabilidade expandida, onde técnicas de Mineração Visual de Dados (MVD) são aplicadas ao resultado da consulta, permitindo ao usuário ter mais informações sobre a distribuição das imagens resultantes, ao invés de simplesmente a informação de similaridade. Assim, o sistema CBIR é um dos núcleos deste trabalho de mestrado e, neste capítulo, são detalhados os principais tópicos referentes ao mesmo.

2.2 Representação de Imagens

Para que as imagens possam ser trabalhadas computacionalmente, é necessário encontrar uma representação numérica que sintetize o conteúdo das mesmas, levando-se em consideração características visuais específicas, como cor, forma, e textura.

De um modo geral, a atividade de representação numérica da imagem envolve: (a) pré-processamento da imagem; (b) segmentação da imagem; e (c) extração de características da imagem. Os passos (a) e (b) são opcionais.

Pré-Processamento

A principal característica da atividade de pré-processamento da imagem é a aplicação de técnicas computacionais para simplificar a representação da imagem com o intuito de reduzir o ruído e/ou para realçar características de interesse na imagem. Filtro da mediana e aplicação de transformadas de *Fourier* ou *Wavelet* são exemplos de técnicas empregadas na atividade de pré-processamento da imagem.

Segmentação

O processo de segmentação de imagens consiste na utilização de métodos computacionais para separar regiões ou objetos que possuam informações de interesse em uma determinada imagem. Desta forma, informações não relevantes podem ser removidas, mantendo-se somente características de interesse para o contexto da aplicação.

2.2.1 Extração de Características de Imagens

A tarefa de extração de características da imagem é feita pela aplicação de algoritmos de processamento de imagem para a extração de características visuais de cor, forma e textura de

imagem. Através destas, é obtido uma representação numérica da imagem, sendo esta denominada vetor de características. Esse vetor é então utilizado em qualquer manipulação desejada entre imagens, viabilizando a consulta por conteúdo de imagens (*Content-Based Image Retrieval* - CBIR), mineração de imagens e a indexação de imagens.

Extração de Características Baseada em Cores

Muitas técnicas encontradas para a manipulação de imagens digitais utilizam como base histogramas. O histograma é uma abordagem que realiza um mapeamento numérico através da quantização e da análise da imagem em determinados níveis de cores.

Trabalhos são encontrados na literatura na verificação de similaridade (distância) entre duas imagens através da utilização de histograma de cores (HAFNER et al., 1995), (KO; LEE; BYUN, 2000). Tal constatação é devido ao fato de que: (a) histogramas de cores são computacionalmente triviais para calcular; (b) pequenas mudanças do ângulo de visão da câmera tendem a não afetar os valores de histogramas de cores; e (c) objetos diferentes frequentemente possuem histogramas de cores distintos (PASS; ZABIH; MILLER, 1996).

Extração de Características Baseada em Forma

A extração das formas/contornos presentes em imagens é uma das possibilidades encontradas na atividade de extração de características. O modelo de contorno ativos *Snakes* (KASS; WITKIN; TERZOPOULOS, 1988) e os momentos de *Zernike* (KHOTANZAD; HONG, 1990) são algumas das alternativas para encontrar uma representação da imagem a partir da detecção de contornos.

Extração de Características Baseada em Textura

O conceito de textura possui ampla discussão em sua definição. Generalizando, "texturas podem ser modeladas como uma coleção de objetos primitivos similares, não necessariamente idênticos, denominados *textels*"(KUMAR; PANG, 2002).

Dentre várias abordagens para a extração de características de imagens utilizando o conceito de textura, existem as técnicas baseadas nos filtros de *Gabor* (MANJUNATH; MA, 1996) e nas matrizes de co-ocorrência (*Spatial Gray Level Dependence*, SGLD) (HARALICK; SHANMUGAM; DINSTEN, 1973).

A escolha de uma ou mais técnicas de extração de características está diretamente relacionada com o contexto em que o conjunto de imagens está inserido, viabilizando encontrar uma representação numérica que atenda as características visuais específicas deste contexto. No cenário de imagens médicas, por exemplo, deseja-se encontrar características que viabilizem a

representação de anomalias em pacientes, como calcificações, células atípicas, entre outros.

2.3 Consulta por Similaridade

Ao se trabalhar com imagens, o critério por similaridade é mais adequado do que o conceito de igualdade, visto que a consulta por igualdade tem pouca utilidade na prática. Existe a possibilidade de serem encontradas imagens idênticas em uma dada base, porém estes são casos em que há duplicidade de dados. No caso hipotético em que duas imagens representativas de um mesmo objeto são tiradas em momentos diferentes, as mesmas, em geral, não são idênticas, uma vez que diferem em determinados atributos visuais.

O critério usado em CBIR é a similaridade. Desta forma, é possível fazer uma consulta verificando quais imagens possuem determinadas características mais semelhantes à imagem de busca. Essa tarefa pode ser por abrangência ou vizinhança. Tais técnicas são detalhadas a seguir.

A consulta por abrangência (*Range Query* - RQ) permite que, dada uma imagem de consulta (O) e um raio de tamanho r , todas as imagens encontradas a uma distância (relativa à imagem de consulta) igual ou inferior a r são recuperadas. Assim, é definido na consulta a proximidade exigida por uma distância máxima, sendo recuperadas todas as imagens que satisfazem a restrição imposta. Uma situação hipotética para a consulta por abrangência é ilustrada na Figura 2.1, onde, a partir de dada imagem central (O) e de uma distância máxima (r), são encontradas como semelhantes quatro imagens (em azul) definidas por uma distância inferior à imposta previamente.

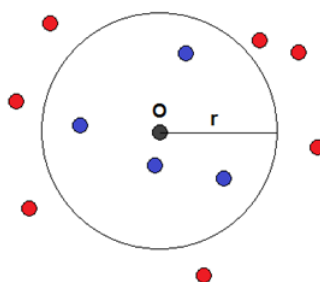


Figura 2.1: Exemplo de consulta utilizando Range Query.

A consulta por vizinhança (*k-Nearest Neighbor* - kNN) possibilita encontrar, através de uma imagem de consulta e um número inteiro k , as k imagens mais semelhantes (mais próximas) à imagem de consulta. Logo, nesta abordagem são levantadas todas as k imagens mais próximas/semelhantes da imagem de consulta, independentemente da distância em que estas

podem estar localizadas. Uma situação hipotética com a utilização da consulta por vizinhança é ilustrada na Figura 2.2, onde, a partir de uma imagem dada como entrada (O) e da definição $k = 4$, são encontradas as quatro imagens mais próximas (em azul).

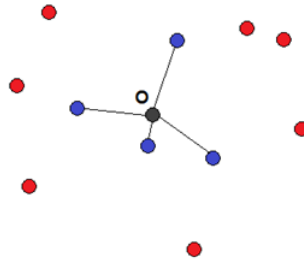


Figura 2.2: Exemplo de consulta utilizando k NN.

Ambos métodos descritos são utilizados na consulta por similaridade, já que propõem abordagens distintas. Na RQ tem-se um maior desempenho computacional por já estar definido o raio de abrangência, no entanto, esse tipo de consulta necessita do conhecimento prévio da distribuição das imagens para determinar um raio de abrangência adequado. Já na consulta por vizinhança, tem-se somente a necessidade de declarar o número de imagens mais similares a serem encontradas, descartando a necessidade de conhecimento prévio do espaço em que as imagens estão distribuídas para a determinação da distância.

Neste trabalho são usadas ambas as abordagens para avaliação, possibilitando que técnicas de visualização sejam também incorporadas ao processo. A visualização dos resultados é fornecida para que o usuário possa ter melhor noção de ambas as aplicações na análise de imagens semelhantes, evidenciando-se o entendimento das vantagens e desvantagens de cada uma.

2.4 Funções de Distância

A similaridade entre duas imagens é mensurada pelo uso de funções de dissimilaridade, também chamadas de funções de distância. Desta forma, a semelhança entre duas imagens passa a ser representada pela distância entre as mesmas dentro de um espaço definido.

No processo de verificação de semelhança entre imagens, as funções de distância são aplicadas nos vetores de características obtidos no processo de extração de características. Tal representação dos atributos da imagem passa a ser determinante para o posicionamento de cada objeto da base em um local específico do espaço.

Existem várias métricas para a aplicação do cálculo da distância entre imagens, sendo que as mais conhecidas fazem parte da família *Minkowski* (L_p), as quais são definidas na Equação

2.1.

$$\left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2.1)$$

Existem três ocorrências bem conhecidas da família L_p . No caso em que $p = 1$, temos a denominada Distância de Manhattan; No caso de $p = 2$, temos a Distância Euclidiana; e, no caso do cálculo para o limite de L_p quando p tende ao infinito, temos a métrica L_∞ . A Figura 2.3 ilustra pontos equidistantes em relação ao ponto central (O) segundo as métricas da família *Minkowski*.

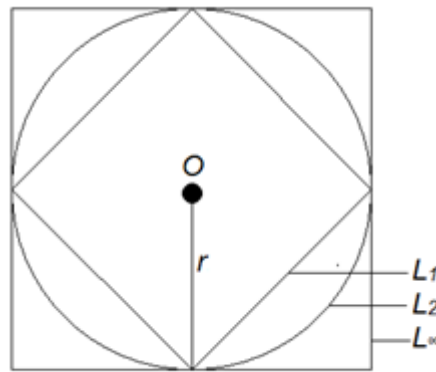


Figura 2.3: Pontos equidistantes situados a uma distância r do objeto O , de acordo com as métricas L_1 , L_2 e L_∞ .

2.5 Indexação de Imagens

Para melhorar a eficiência das operações realizadas em uma base de dados, são utilizadas estruturas de indexação, também chamadas de métodos de acesso. Tais estruturas de indexação também são utilizadas em bases de dados complexos, como imagens, por exemplo. O objetivo é reduzir o número de cálculos de distância necessários para a realização da consulta. Os métodos de acesso são classificados em:

- **Métodos de Acesso Espaciais (MAE):** Estruturas que armazenam informações relativas ao posicionamento de determinado objeto em um espaço N -dimensional, onde N é o número de dimensões do espaço em questão. Contudo, devido a necessidade de armazenar as coordenadas de cada objeto da base, tais métodos são eficientes para um número limitado de dimensões, normalmente não ultrapassando 10 ou 20 (ZHOU; PU, 2002). Exemplos

de MAE encontrados na literatura são: *R-Tree* (GUTTMAN, 1984), *R*- Tree* (BECKMANN et al., 1990), *SR-Tree* (KATAYAMA; SATOH, 1997).

- Métodos de Acesso Métricos (MAM): Estruturas que guardam somente a informação da distância entre os objetos presentes na base (CHANG; CHEN; LEE, 2002). Exemplos de MAM encontrados na literatura são: *M-Tree* (CIACCIA; PATELLA; ZEZULA, 1997), *Slim-Tree* (TRAINA JR. et al., 2000), *DF-Tree* (TRAINA JR. et al., 2002).

Na utilização de um MAM, necessariamente a função de distância escolhida deve ser métrica. Uma função de distância d é considerada métrica se, para qualquer objeto X, Y e Z forem respeitadas as seguintes propriedades (CHANG; CHEN; LEE, 2002):

Auto-similaridade: $d(X, X) = 0$;

Positividade: $d(X, Y) \geq 0$;

Simetria: $d(X, Y) = d(Y, X)$;

Desigualdade Triangular: $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

2.6 Avaliação dos Resultados

Para assegurar a eficácia da consulta, é necessário verificar se a mesma trouxe resultados satisfatórios. Em outras palavras, é averiguada a relação dos objetos significantes na consulta, também denominados elementos relevantes, com os resultados obtidos e esperados.

Existem diferentes meios de mensurar a eficiência de uma dada consulta. Nesta seção, é levada em conta uma abordagem de âmbito estatístico, através da utilização de medidas de precisão e revocação.

2.6.1 Revocação e Precisão

Uma das principais abordagens para a verificação da eficácia da consulta é baseada na amostragem do resultado obtido. Desta forma, algumas métricas são utilizadas para melhor mensurar esta atividade, sendo, na maioria das vezes, empregados os conceitos de revocação (*recall*) e precisão (*precision*) (BAEZA-YATES; RIBEIRO-NETO, 1999).

O termo revocação está relacionado à fração que representa o subconjunto de elementos relevantes obtidos em relação ao conjunto total de elementos relevantes pertinentes à consulta.

Assim, neste projeto, a revocação indica a quantidade de imagens relevantes provenientes da consulta (I_{RC}) em relação à quantidade total de imagens relevantes existentes na base (I_R).

$$Revocação = \frac{I_{RC}}{I_R} \quad (2.2)$$

A precisão é definida como a fração que representa o subconjunto de elementos relevantes em relação ao conjunto total obtido na consulta. No contexto deste projeto, a precisão indica a fração de imagens relevantes (I_{RC}) pela quantidade total de imagens obtidas na consulta (I_C).

$$Precisão = \frac{I_{RC}}{I_C} \quad (2.3)$$

Um resultado mais claro da avaliação é obtido através da análise do gráfico de revocação e precisão. Para exemplificar a construção do mesmo (BAEZA-YATES; RIBEIRO-NETO, 1999), é considerada a realização de determinada consulta C em uma base onde existem dez elementos relevantes (E_R).

$$E_R = e_4, e_{13}, e_{18}, e_{23}, e_{27}, e_{31}, e_{52}, e_{55}, e_{56}, e_{68} \quad (2.4)$$

$$C = \begin{pmatrix} 1.e_{18}\bullet & 6.e_{42} & 11.e_{72} \\ 2.e_3 & 7.e_{14} & 12.e_{36} \\ 3.e_4\bullet & 8.e_{30} & 13.e_{11} \\ 4.e_{50} & 9.e_9 & 14.e_{67} \\ 5.e_{52}\bullet & 10.e_{27}\bullet & 15.e_{56}\bullet \end{pmatrix} \quad (2.5)$$

A consulta C possui como retorno os objetos em uma dada ordem ou classificação e os elementos marcados por \bullet pertencem ao conjunto dos elementos relevantes.

Ao examinar o primeiro objeto da sequência ($e_{18} \bullet$), é verificado que o mesmo pertence ao conjunto de elementos relevantes, tendo uma precisão de 100% e a revocação igual a 10%, já que, até o momento, somente um elemento relevante foi encontrado em um total de dez. O segundo elemento encontrado na consulta (e_3), não é relevante, então é avaliado o terceiro elemento ($e_4 \bullet$) ocasionando em uma precisão de aproximadamente 66% (dois objetos relevantes no total de três) e a revocação em 20% (dois elementos relevantes do total de dez). A análise prossegue até o momento em que todos os elementos retornados são analisados, resultando nos valores encontrados na Tabela 2.1.

Revocação (%)	Precisão (%)
10	100
20	66,6
30	60
40	40
50	33,3

Tabela 2.1: Levantamento, em porcentagem, de revocação e precisão.

É então considerado que a precisão para valores de revocação superiores a 50% tendem a zero, visto que a nem todos os elementos relevantes foram encontrados na consulta *C*. Logo, o gráfico de revocação e precisão pode ser construído (Figura 2.4), levando-se em conta que é feita uma interpolação para o caso em que o nível de revocação é igual a 0%.

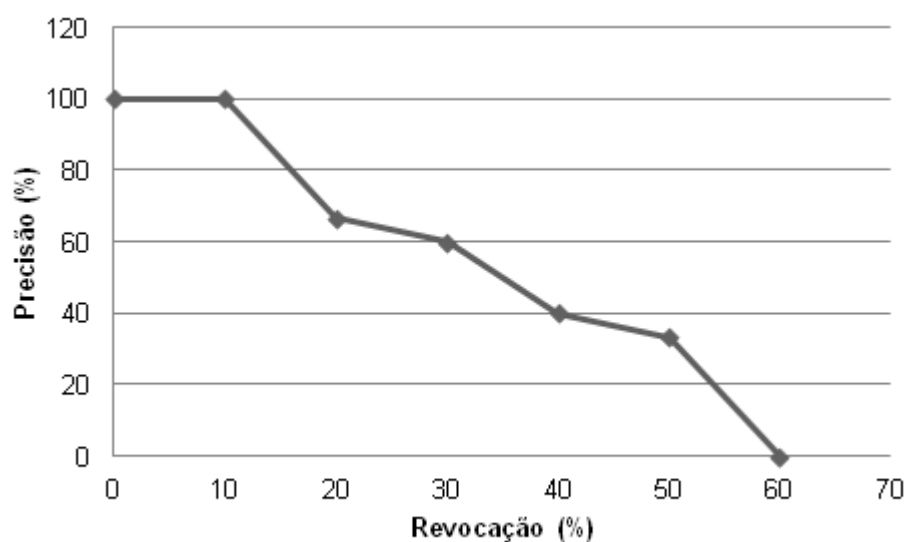


Figura 2.4: Gráfico de Revocação X Precisão.

Para obter um resultado mais satisfatório entre a precisão e a revocação, normalmente são realizadas diversas consultas e, através destas, é feita uma média dos valores encontrados para cada intervalo.

Em termos práticos, a condição ideal esperada seria que tanto a revocação quanto a precisão apresentassem um valor igual a 1 (100%). Neste cenário, todos os elementos relevantes esperados foram obtidos na consulta, sendo que a mesma também não obteve nenhum outro elemento além desses.

2.7 Descontinuidade Semântica (*Gap Semântico*)

Segundo (LIU et al., 2007), existem três níveis de consulta em um sistema CBIR, sendo estes:

Nível 1: Recuperação por características primitivas, como cor, textura, forma ou a localização que elementos da imagem se encontram. Um exemplo de consulta para esse nível poderia ser: "encontre imagens como essa";

Nível 2: Recuperação de determinados objetos pela identificação de características derivadas, avaliando um certo grau de inferência lógica. Um exemplo de consulta desse nível poderia ser: "encontre imagens que possuem flores";

Nível 3: Recuperação por atributos abstratos, envolvendo uma quantidade significativa de fatores relacionados a objetos ou da cena em que a imagem se contextualiza. Esse nível inclui a recuperação de eventos específicos, imagens com certa significância emocional ou religiosa, por exemplo. Um exemplo de consulta para esse nível poderia ser: "encontre imagens de pessoas alegres".

Os níveis 2 e 3 são denominados como recuperação semântica de imagens e a lacuna existente entre os níveis 1 e 2 denominada como *gap semântico*. Em outras palavras, o *gap semântico* é a discrepância entre características de baixo nível (cor, textura, forma, etc) e características de alto nível (próximas à cognição humana, podendo possuir várias interpretações intermediárias, como palavras, sentimentos, etc) (EAKINS; GRAHAM; FRANKLIN, 1999).

Para reduzir a descontinuidade semântica na consulta por conteúdo de imagens, são propostas cinco abordagens (LIU et al., 2007), sendo:

- Utilização de ontologias para definir características de alto nível do usuário;
- Utilização de métodos de aprendizagem supervisionados ou não supervisionados para associar características de baixo nível com conceitos envolvidos na consulta;
- Adição de técnicas de realimentação de relevância para, de forma iterativa, aprender as expectativas do usuário e adequar os resultados obtidos;
- Geração de modelos (*templates*) semânticos para suportar características de alto nível;
- Utilização de informações textuais e imagens providas da Web, levando-se em conta a associação de descrições relativas às imagens, como textos ou *hyperlinks*.

Em (RAZENTE, 2009) é verificado que as abordagens baseadas em aprendizagem de máquina e em realimentações de relevância tem demonstrado resultados mais satisfatórios para a aproximação das características de alto nível e, conseqüentemente, reduzir a descontinuidade semântica. No caso de realimentação de consulta, é proposto técnicas como:

Avaliação da distribuição de pesos: Associar um peso mais elevado para dimensões mais representativas das categorias das imagens dadas como relevantes;

Movimentação do centro de consulta (*Query Point Movement - QPM*): Mudar a centro de consulta para uma região mais próxima das imagens consideradas relevantes retornadas na consulta anterior.

2.8 Fator Diversidade

A adição do fator de diversidade em aplicações que possuem grandes conjuntos de dados tem sido uma tentativa de melhorar a análise exploratória e, conseqüentemente, reduzir a descontinuidade semântica de sistemas baseados em recuperação por conteúdo de imagens. Neste sentido, não apenas a similaridade entre a imagem de consulta com as demais imagens da base é considerada, mas também certa dissimilaridade entre imagens no subconjunto resultante.

Em alguns casos, conjuntos de dados podem conter imagens com grande grau de similaridade entre si. Neste cenário, é possível que uma consulta por vizinhança (*kNN query*) retorne um subconjunto muito homogêneo de imagens, especialmente quando os dados estão distribuídos em *clusters* (JAIN; SARDA; HARITSA, 2004). No entanto, um resultado com alta homogeneidade não é, muitas vezes, esperado pelo usuário. Uma solução para este problema pode ser obtida através da adição de um fator de diversidade nos resultados de consultas por vizinhança, considerando assim uma certa dissimilaridade entre as imagens no subconjunto resultante.

Para ilustrar a importância do fator de diversidade, considerando-se a análise de imagens médicas em uma situação hipotética, uma imagem de consulta possivelmente tem mais semelhança com um grupo de imagens normais (negativo). No entanto, a mesma possui uma pequena semelhança em relação a um grupo com determinada anomalia (positivo). Neste caso, se somente similaridade fosse considerada, unicamente imagens do grupo normal seriam retornadas. No entanto, se certo fator de diversidade fosse aplicado, imagens do grupo anormal também poderiam ser obtidas, possibilitando, neste caso, uma melhor análise do diagnóstico da imagem de consulta.

Aplicar diversidade em consultas por vizinhança, no entanto, requer alto custo computacio-

nal, pois exige um grande número de cálculos de distância para verificar a dissimilaridade entre diferentes subconjuntos de imagens.

Para superar o alto custo computacional, algoritmos aleatórios ou gulosos podem ser utilizados na viabilização da adição do fator de diversidade nos resultados de consultas por vizinhança. A verificação aleatória de conjuntos candidatos pode, eventualmente, encontrar a melhor solução (ótima global). No entanto, este fato não é garantido e, além disso, os resultados obtidos podem se distanciar do esperado pelo usuário. Algoritmos gulosos, em contrapartida, reduzem o número de possíveis candidatos a fim de encontrar uma solução ótima local, a qual é esperada ser igual à ótima global ou, pelo menos, suficientemente perto da mesma (WEISS, 1986). Além disso, a seleção dos subconjuntos de possíveis candidatos em algoritmos gulosos é uma tarefa que pode ser guiada pelo usuário, facilitando a compreensão de quais imagens foram consideradas neste processo.

Devido ao elevado custo na aplicação de diversidade em consultas por vizinhança, aproximações foram elaboradas para que o número de cálculos de distância seja reduzido. Diferentes métodos são citados em (VIEIRA et al., 2011a), como *Swap*, *BSwap*, *Maximal Marginal Relevance*, *Motley*, *Max-Sum Dispersion*, *Clustering*, *Greed Marginal Contribution* e *GRASP with Neighbor Expansion*, a fim de utilizar uma abordagem que priorize a velocidade ou a qualidade dos resultados.

Em (VIEIRA et al., 2011a), o usuário pode manipular um valor de compensação (*trade-off*) entre diversidade e similaridade, além de outros parâmetros utilizados na consulta. No entanto, mesmo que o usuário possa participar mais ativamente do processo de consulta, a compreensão de como o fator de diversidade está sendo aplicado em todo o conjunto de dados é ainda de difícil entendimento. Elementos aleatórios ou ponderados são escolhidos para reduzir a quantidade possíveis candidatos (conjunto de imagens) a serem analisados, mas a compreensão de quais conjuntos foram considerados no processo ou as restrições de similaridade e diversidade empregadas não são atividades triviais para o usuário.

Em (JAIN; SARDA; HARITSA, 2004), o fator de diversidade é calculado com base no coeficiente de *Gower*. Um limiar de diversidade mínimo é fixado e é verificado se a diversidade (dissimilaridade) calculada entre os elementos do subconjunto candidato satisfaz o valor estipulado. Se o limiar é baixo o suficiente, o algoritmo encontra a solução sem muitas iterações, caso contrário, são realizadas quantas iterações forem necessárias para encontrar um subconjunto que atenda ao limite mínimo de diversidade.

2.9 Considerações Finais

Neste capítulo foram levantadas técnicas necessárias para o desenvolvimento de sistemas CBIR. A descontinuidade semântica (*Semantic Gap*) surge, neste contexto, pela dificuldade em obter uma representação numérica das características visuais que atenda as mesmas expectativas do usuário. O conceito de diversidade é aplicado na consulta por vizinhança, aliando-se técnica de visualização e interação para obtenção de resultados mais satisfatórios para o usuário. Por fim, neste trabalho é feita uma avaliação do resultado obtido, usando as medidas de precisão e revocação a fim de quantificar a contribuição deste projeto em diferentes contextos de sua aplicação.

Capítulo 3

MINERAÇÃO VISUAL DE DADOS

A Mineração Visual de Dados (MVD) tem como intuito proporcionar uma maneira inovadora de lidar com o constante fluxo de informações produzidas pela humanidade. O principal objetivo desta atividade é promover a relação entre algoritmos tradicionais de Mineração de Dados (MD) e a visualização de dados, proporcionando uma facilitação na exploração de informações. Na Seção 3.1 são dadas as considerações iniciais acerca do conceito de MVD, assim como as vantagens de sua utilização neste projeto de mestrado. A Seção 3.2 especifica as diferentes técnicas na aplicação da MVD. A Seção 3.3 ilustra a contextualização da MVD voltado para o contexto imagens e a Seção 3.4 trata das considerações finais deste capítulo.

3.1 Considerações Iniciais

Milhares de dados são produzidos diariamente, porém a atividade de transformação dos mesmos em conhecimento mostra-se uma tarefa cada vez mais complexa. O Processo de Descoberta de Conhecimento em Base de Dados (*Knowledge Discovery in Databases - KDD*) (FAYYAD et al., 1996) promove um auxílio nesta atividade, possibilitando desde um pré-processamento dos dados até a interpretação de padrões encontrados. Desta forma, o KDD permite a exploração das seguintes atividades:

- Verificação: Verificação da hipótese inicial proposta;
- Descobrimto: Padrões encontrados pelo próprio sistema, dividindo-se em duas linhas:
 - Predição: Padrões encontrados pelo sistema com o objetivo de prever o comportamento de uma entidade baseada em informações de sua história;

Descrição: Padrões encontrados pelo próprio sistema com o propósito de descrever os dados de forma a facilitar a compreensão do mesmo pelo usuário.

Os procedimentos envolvidos no processo iterativo de aplicação do KDD são ilustrados na Figura 3.1 e classificados nas seguintes etapas:

- Seleção: Seleção ou foco de um determinado subconjunto de dados relevantes para o processo de extração de conhecimento;
- Pré-processamento: Procedimentos para a remoção de ruídos, dados com informações corrompidas e tomada de outras soluções para complicações que possam comprometer a consistência das informações;
- Transformação: Aplicação de transformações nos dados para atender os objetivos do KDD designados previamente ou para reduzir a complexidade e facilitar a manipulação dos mesmos;
- Mineração de Dados: Etapa em que é aplicado o algoritmo de extração de mineração de dados, tendo como objetivo a extração de conhecimento;
- Interpretação: Interpretação dos resultados obtidos na iteração do KDD. Caso os objetivos tenham sido atingidos, o processo é finalizado e o conhecimento útil é obtido, caso contrário é retornado para uma etapa anterior e uma nova iteração é realizada.

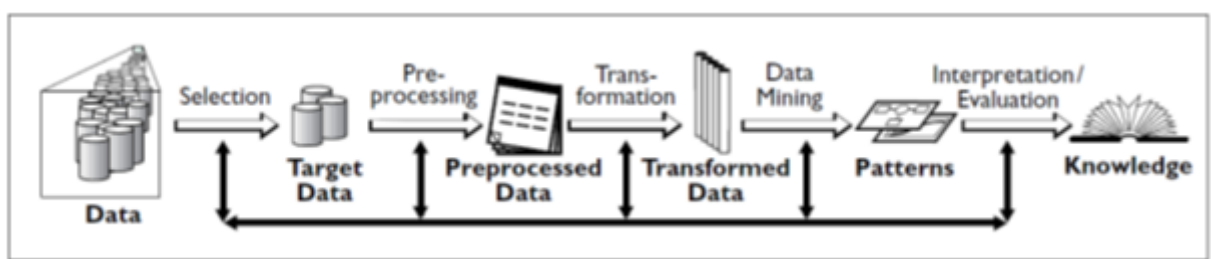


Figura 3.1: Imagem retirada de (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Ilustração das cinco etapas envolvidas no Processo de Descoberta de Conhecimento em Base de Dados (KDD).

Algoritmos tradicionais de MD mostram-se importantes no processo de extração do conhecimento, proporcionando a automatização na extração do conhecimento. Todavia, em casos em que os dados apresentam maiores complexidades, como imagens, a simples aplicação de um mecanismo de MD mostra-se computacionalmente inviável.

A representação numérica de imagens não é uma tarefa trivial, acarretando em um distanciamento da cognição humana e, conseqüentemente, gerando resultados insatisfatórios no

processo de extração de conhecimento. Logo, técnicas voltadas a visualização e interação provenientes da MVD podem ser utilizadas ao longo das etapas do KDD, permitindo que as informações oriundas das imagens sejam trabalhadas, compreendidas e manipuladas pelo usuário, facilitando o processo de extração do conhecimento.

A união de algoritmos de MD juntamente com técnicas de MVD permite criar um processo de extração de conhecimento que une a eficiência computacional com a percepção humana. Desta maneira, a extração de conhecimento de conjunto de imagens pode ser obtido tendo o usuário como participante ativo neste exercício.

A MVD propõe não somente a exibição de informações provenientes do processo de MD tradicional, mas também possibilita que o usuário dê diretivas que possam facilitar no processo de reconhecimento de padrões. Assim, o usuário consegue interagir, de acordo com sua percepção, direcionando o mecanismo de mineração para as características que julgar mais relevantes.

O mecanismo de MVD provê um potencial de melhoria no processo de MD especialmente para (MÜLLER; SCHUMANN, 2002):

- Bases com grande quantidade de informação;
- Em casos que pouco se sabe em relação aos dados e os objetivos da exploração são vagos;
- Em casos que há pouca homogeneidade e grande comprometimento dos dados devido a, por exemplo, presença de ruídos.

3.2 Técnicas de Mineração Visual de Dados

As técnicas de MVD podem ser diferenciadas quanto à informação a ser visualizada, à técnica utilizada para a visualização e às técnicas de interação e distorção empregadas (KEIM et al., 2002). Logo, com a combinação desses três elementos, é possível obter uma abordagem específica para o problema de MVD em questão, apresentada na Figura 3.2, levando-se em consideração as técnicas de interação e distorção aplicadas (eixo X), a informação a ser visualizada (eixo Y) e a técnica de visualização (eixo Z).

3.2.1 Classificação pela Informação a ser Visualizada

Milhares de informações são geradas e armazenadas diariamente, sendo então associadas a objetos em uma base. Tais objetos possuem um dado número de atributos com diferentes valores para sua identificação.

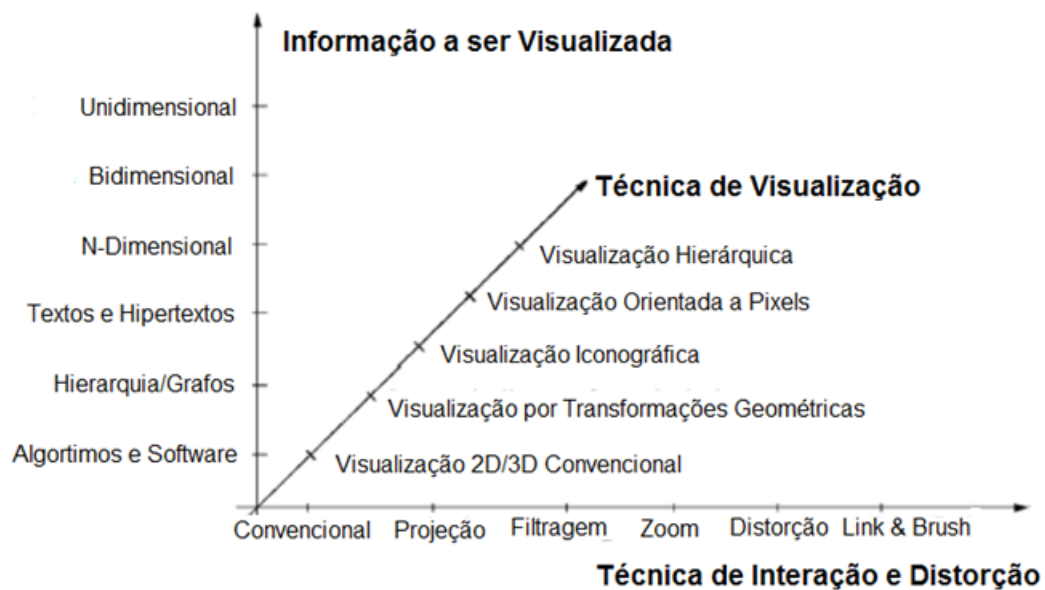


Figura 3.2: Imagem adaptada de (KEIM et al., 2002). Técnicas de Mineração Visual de Dados.

Objetos podem ser avaliados em diferentes dimensionalidades dependendo do número de atributos e do tipo de informações encontradas na base. Logo, existem diversas disposições conhecidas para os dados, estando estas relacionadas a Dados Unidimensionais, Dados Bidimensionais, N-Dimensionais, Textos e Hipertextos, Grafos ou Algoritmos e Software.

Dependendo do problema em questão, os dados persistentes na base são tratados de forma a apresentarem N atributos interessantes, onde N é um número inteiro positivo, para a tarefa de análise e extração de dados. Desta forma, o problema em questão passa a apresentar N dimensões, sendo este valor igual ao número de atributos levantados para cada objeto.

O mapeamento de dados que apresentam um grande número de dimensões em um dispositivo de visualização tradicional de duas dimensões não é uma tarefa simples. São necessárias técnicas de projeção que possibilitem representar devidamente a informação contida na base de dados além proporcionar ao usuário um ambiente adequado para interpretação favorecida pela cognição humana.

Imagens, as quais foram abordadas nesse estudo, normalmente apresentam diversos atributos relativos a características visuais. Em casos em que imagens são representadas em escala de cinza, por exemplo, cada nível corresponde a uma dimensão determinada pela frequência em que ocorre na imagem.

3.2.2 Classificação por Técnicas de Visualização

A visualização da informação é o principal procedimento no processo de MVD. A simples representação textual não permite, muitas vezes, o total aproveitamento dos recursos providos pela cognição humana.

A forma como os dados e, possivelmente, padrões são representados é fundamental para proporcionar ao usuário um bom entendimento do contexto em questão.

Não obstante, existem diversas formas de visualização, sendo necessário avaliar para cada caso específico a representação direcionada para a melhor compreensão do contexto em que os dados estão inseridos. Esta seção levanta aspectos sobre a Visualização de Transformações Geométricas, Visualização por Coordenadas Paralelas e Visualização Hierárquica.

Visualização de Transformações Geométricas

Em alguns casos, mostra-se viável a aplicação de transformações geométricas para que os dados possam ser visualizados de forma mais simplificada, facilitando o processo de extração do conhecimento. Dados multidimensionais, por exemplo, necessitam, muitas vezes, dependendo da dimensionalidade do problema em questão, ser trabalhados para encontrar uma representação viável a um dispositivo bidimensional tradicional de exibição. A Figura 3.3 mostra um exemplo de aplicação da técnica de visualização *Scatterplots*.

A técnica de Gráfico de Dispersão (*Scatterplots*) é uma das técnicas mais antigas para a projeção de dados de grandes dimensões para um contexto bidimensional. A partir das N dimensões, são geradas $N * (N - 1) / 2$ pares de projeções paralelas, sendo que cada par denota a relação das informações contidas entre duas dimensões (WARD, 1994). A técnica de visualização por *Scatterplots* facilita na exploração dos dados por meios de (ANDREWS, 1972):

- Clusterização: Dada a função de mapeamento, é possível fazer a verificação de objetos que encontram-se próximos segundo a distância (Euclidiana) em que os mesmos se encontram;
- Testes para verificação de significância de determinados valores: Com a distribuição dos objetos no plano cartesiano, certos objetos podem ser verificados como *outliers* (não significantes para o problema em questão) e *inliers* (significantes para o contexto), dependendo da distância que estes se encontram em relação aos demais objetos visualizados. Para esta atividade é normalmente utilizada uma função para verificar o nível de significância desses objetos, analisando, por exemplo, a média, a variância ou o desvio padrão das distâncias.

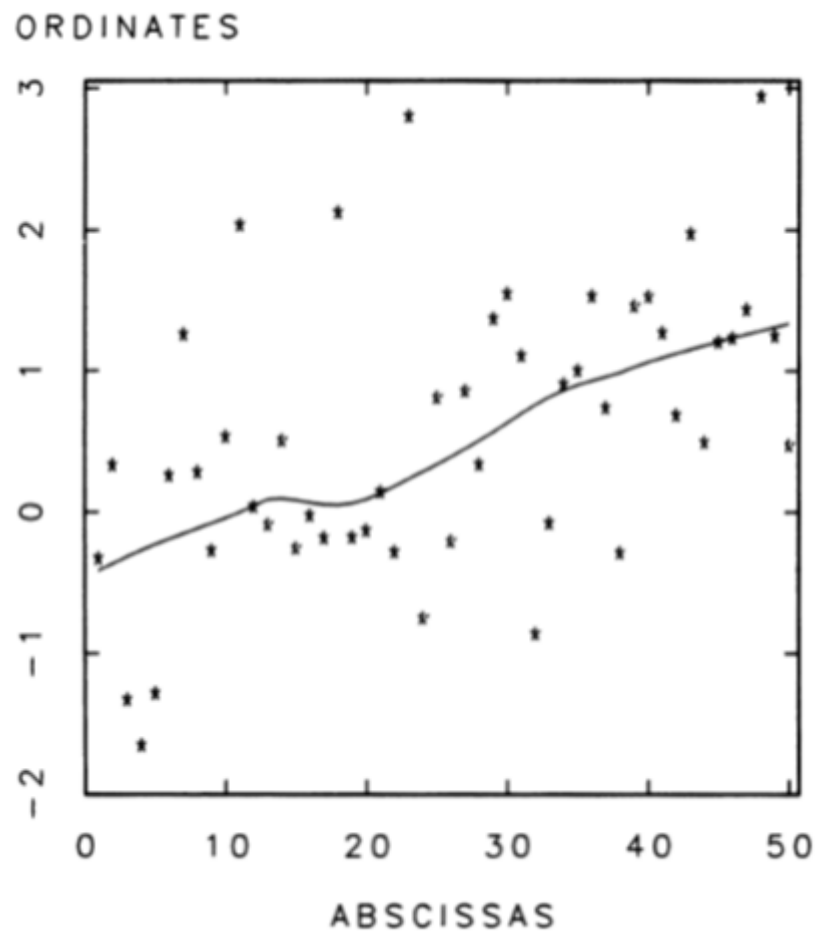


Figura 3.3: Imagem retirada de (CLEVELAND, 1979). Técnica de Visualização *Scatterplots* para dados gerados artificialmente e aproximação de seus valores para uma curva.

Visualização por Coordenadas Paralelas

A visualização por coordenadas paralelas parte do princípio de representar as N dimensões em uma visualização bidimensional, utilizando-se N eixos paralelos com uma mesma orientação positiva e equidistantes entre si (INSELBERG; DIMSDALE, 1990).

Os valores são mapeados através de arcos de curvas, normalmente sendo estes segmentos de reta, que interceptam os eixos no ponto que corresponde ao seu valor em cada uma das dimensões consideradas. Um exemplo de visualização utilizando coordenadas paralelas é ilustrado na Figura 3.4, onde são avaliados n atributos, os eixos verticais X_n representam as dimensões e a linha poligonal representa a relação entre dimensões baseada no valor C_n para dado atributo.

Visualização Hierárquica

Uma das diferentes maneiras para se visualizar dados N -Dimensionais é através da representação em forma hierárquica dos atributos. Para tanto, o espaço N -dimensional é subdividido e

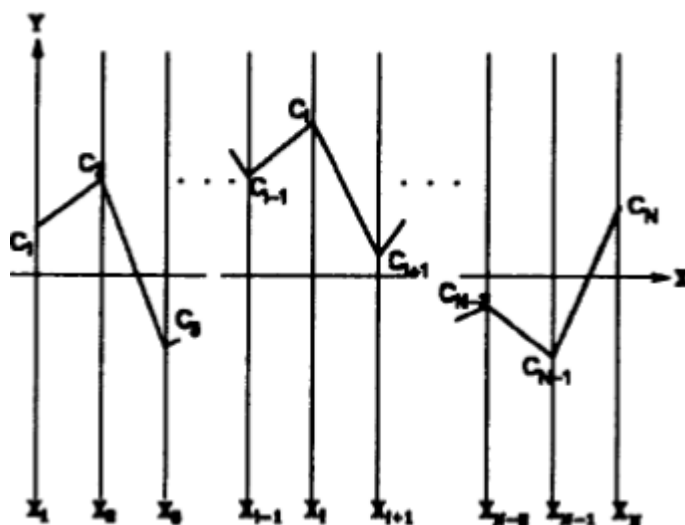


Figura 3.4: Imagem retirada de (INSELBERG; DIMSDALE, 1990). Técnica de visualização por coordenadas paralelas.

os subespaços resultantes são dispostos de forma hierárquica, projetando um sistema de coordenadas dentro de um outro sistema de coordenadas mais amplo.

Em (LEBLANC; WARD; WITTELS, 1990), é proposta a exibição dos dados em um sistema de coordenadas bidimensional, mapeando-se os valores correspondentes em uma janela retangular. Posteriormente, a cada necessidade de inserção de novas dimensões, o espaço é utilizado como uma visualização virtual e é feita uma divisão hierárquica, subdividindo-se o espaço corrente em um novo sistema de coordenadas.

Em (ROBERTSON; MACKINLAY; CARD, 1991) é proposto uma visualização hierárquica denominada *Cone-Tree*. A ferramenta desenvolvida permite visualizar os dados dispostos em uma hierarquia de árvores, através de uma simbologia *top-down*, onde os nós filhos encontram-se abaixo dos nós pais. Para evitar dificuldades na sobreposição de nós, é proposto que os diferentes níveis da árvore estejam dispostos em uma projeção tridimensional em formato de cone, possibilitando a interação do usuário através do cursor. A Figura 3.5 ilustra um exemplo da técnica *Cone-Tree*.

3.2.3 Classificação por Técnicas de Interação e Distorção

A interatividade é a chave que possibilita o usuário fazer parte da atividade exploratória na tarefa de MVD. As técnicas de interação permitem que o usuário analise e avalie os dados de forma manipular informações de acordo com o objetivo exploratório.

As técnicas de distorção mostram-se como uma ferramenta para focar a visualização de

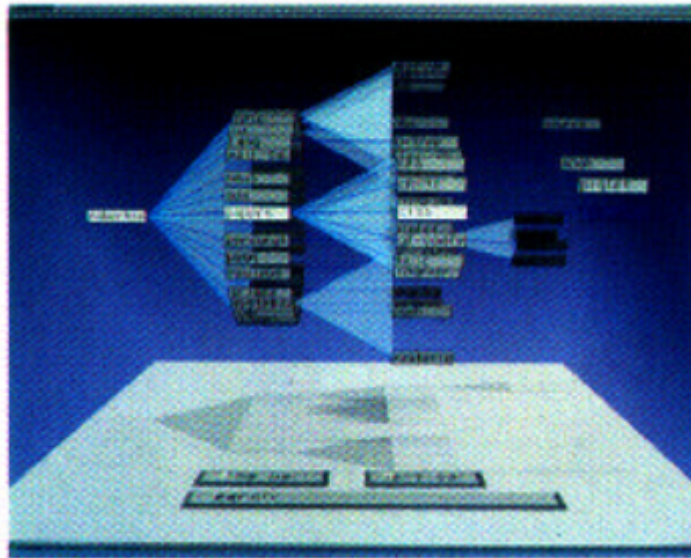


Figura 3.5: Imagem retirada de (ROBERTSON; MACKINLAY; CARD, 1991). Visualização proposta pela técnica Cone-Tree.

um determinado subconjunto de dados, mas respeitando a relação entre os mesmos. Logo, a principal ideia desta técnica é a mostrar com maior detalhamento uma certa porção desejada dos objetos, enquanto os demais são representados com um nível de detalhamento inferior.

Algumas das técnicas de Interação e Distorção são: Projeções Dinâmicas, Filtros Iterativos e Zoom Interativo.

Projeções Dinâmicas

Em casos onde se mostra necessária a visualização de um grande número de atributos e muitos valores associados a estes, a visualização de toda a informação torna-se complicada. A ideia da utilização de projeções dinâmicas é sanar este tipo de problema, permitindo que a projeção utilizada seja alterada dinamicamente ao longo do tempo para facilitar a exploração de dados multidimensionais. Desta forma, é selecionado um subconjunto de dados a ser visualizado com ênfase, sem comprometer os demais dados.

A projeção pode ser alterada de forma automática devido a algum fator identificado pelo sistema, como a possibilidade de uma projeção mais viável para a visualização da informação. Existe também a possibilidade de desencadear a mudança de projeção pela requisição do usuário, através da seleção de um determinado subconjunto de dados, por exemplo.

Em (ROBERTSON; MACKINLAY; CARD, 1991), é proposto que o usuário possa selecionar um dos nós pertencentes à *Cone-Tree* construída e, logo em seguida, a projeção é alterada através do rotacionamento da árvore, facilitando a visualização da hierarquia de nós desejada.

Filtros Iterativos

Outra possibilidade de abordagem em um cenário de alta dimensionalidade com grandes quantidades de objetos é a utilização de filtros dinâmicos e iterativos. Assim, somente um subconjunto de dados é visualizado ou destacado, permitindo então a concentração de esforços empregados na tarefa de exploração de padrões em somente um subconjunto específico.

Existem duas formas principais na aplicação de filtros iterativos. Existe a possibilidade da seleção direta de um subconjunto de dados (*browsing*) ou pela especificação das propriedades do subconjunto desejado (*querying*). Enquanto a técnica de *browsing* mostra-se mais intuitiva, visto que o subconjunto de dados é primeiramente visualizado para posteriormente ser selecionado, a técnica de *querying* possui maior eficiência pelo fato de que não necessita de visualização prévia.

Um exemplo de aplicação visual utilizando filtro dinâmico através de *browsing* pode ser encontrado em (BIER et al., 1993). A ferramenta *Magic Lenses* é constituída por um polígono em formato de retângulo para a seleção da área e objetos desejados, juntamente como um outro recurso para a destacar o objeto relevante para a filtragem, podendo evidenciar diferentes atributos ou valores.

A Figura 3.6 ilustra um exemplo de utilização da ferramenta *Magic Lenses* juntamente com o recurso de colorir o objeto selecionado. Em (a) a situação inicial onde existem seis objetos; em (b) o usuário utilizou a ferramenta preenchendo o objeto desejado com a cor verde e em (c) o usuário utilizou a ferramenta e selecionou um contorno de cor azul para filtrar o objeto desejado.

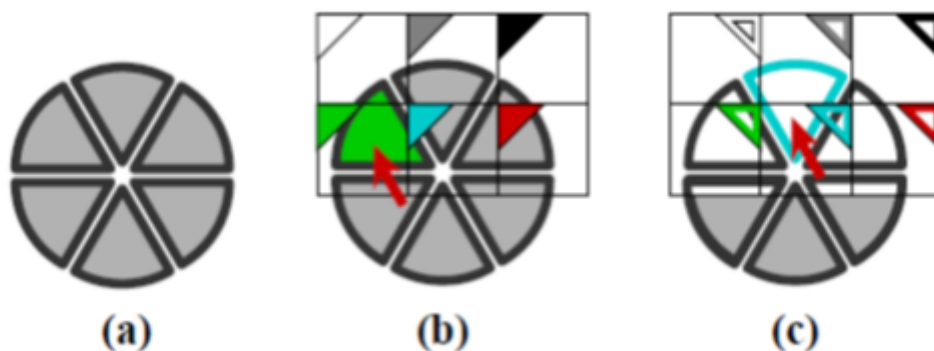


Figura 3.6: Imagem retirada de (BIER et al., 1993). Exemplo de utilização da ferramenta *Magic Lenses* juntamente com o recurso de colorir o objeto selecionado.

A Figura 3.7 demonstra uma outra aplicação da ferramenta *Magic Lenses*, onde são utilizadas duas "lentes" para filtrar a área desejada. São empregados duas janelas para filtragem, sendo

uma para adicionar sombra ao contorno da imagem e outra para exibir o conteúdo em branco e preto.

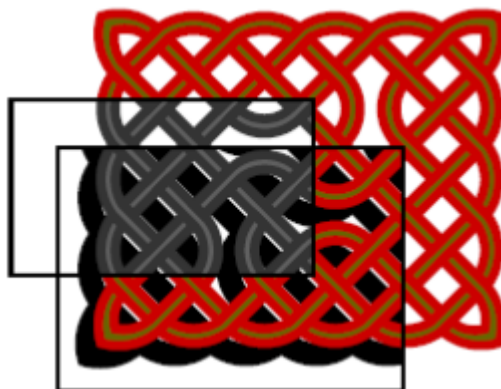


Figura 3.7: Imagem retirada de (BIER et al., 1993). Exemplo de utilização da ferramenta *Magic Lenses* onde são utilizadas duas "lentes" para filtrar a área desejada.

Zoom Interativo

Quando diversos dados são manipulados para visualização, mostra-se viável tratá-los de forma a abranger uma quantidade inferior de informações visuais, através da exibição do resumo dos mesmos. Caso haja necessidade de maior detalhamento, é viabilizada a opção de disponibilizar os dados em diferentes resoluções.

Obter um maior nível de detalhamento não significa apenas a visualização de objetos antes ocultos, mas também possibilita que sejam disponibilizadas mais informações relativas ao objeto, como uma descrição mais minuciosa de atributos ou valores, por exemplo.

3.3 Mineração Visual de Imagens

Como levantado na Seção 3.2.2 existem diversos recursos para a visualização de informação multidimensional em um *display* bidimensional convencional. Contudo, as principais abordagens encontradas em um cenário onde os objetos de estudo são imagens consistem na utilização de Transformações Geométricas ou de Visualizações Hierárquicas.

3.3.1 Visualização Baseada em Transformações Geométricas

A visualização de dados multidimensionais através da utilização de transformações geométricas implica em realizar um mapeamento da informação em um espaço visual de dimensionalidade inferior, mantendo, na medida do possível, as relações de distância definidas no

espaço original. O subespaço de projeção é, normalmente, unidimensional, bidimensional ou tridimensional para facilitar a compreensão humana.

Dentre as possibilidades na abordagem da visualização por transformações geométricas, existem as classes de transformações lineares, não lineares e metodologias híbridas que utilizam ambas abordagens.

- **Transformações lineares:** Permitem projetar informações multidimensionais em uma nova base ortogonal de menor dimensão. Uma técnica de mapeamento linear amplamente conhecida é a *Principal Component Analysis (PCA)* (WOLD; ESBENSEN; GELADI, 1987).
- **Transformações não lineares:** Tentam minimizar uma função de perda de informação, denominada função de *stress*, causada pela redução da dimensionalidade. Majoritariamente, tal função baseia-se na medida de dissimilaridade entre elementos na dimensão original e no novo espaço de projeção. Um exemplo de uma técnica não-linear é *Multidimensional Scaling (MDS)* (BRANDOLI et al., 2010) e *Sammon Mapping* (SAMMON, 1969).
- **Híbrido de mapeamento:** Usa vantagens das técnicas lineares e não-lineares. Métodos conhecidos são *Least-Square Projection (LSP)* (PAULOVICH et al., 2008), *Fast-Map* (FALOUTSOS; LIN, 1995).

Existem diversas vantagens e desvantagens relativas a cada metodologia de transformação, mas, geralmente, transformações lineares realizam a projeção através da solução de um sistema linear, proporcionando um cálculo mais rápido do mapeamento. Técnicas baseadas em transformações não lineares tentam minimizar a função de perda de informação (função de *stress*), permitindo que a nova distribuição mantenha relacionamentos entre instâncias o mais próximo possível do original. Técnicas de mapeamento híbrido são usadas para obter vantagens de ambas as abordagens.

Técnicas de mapeamento, no entanto, fazem uso de uma única transformação global. Neste sentido, se dada instância tiver sua posição alterada ou outras informações forem adicionadas, toda a projeção é recalculada. Desta forma, tal dependência global impossibilita que seja feita uma exploração visual e interativa adequada, prejudicando a ação do usuário no processo de descoberta de conhecimento. No entanto, são encontrados métodos que utilizam projeções locais na literatura, os quais fornecem mapeamentos locais descontinuados na projeção. Logo, através destas metodologias, alterações realizadas na distribuição dos dados não causam impacto sobre toda a projeção, mas somente em uma porção da mesma.

As metodologias PLP (PAULOVICH et al., 2011) e LAMP (JOIA et al., 2011) propõem mapeamentos locais que, através de um subconjunto amostral, é realizada uma projeção inicial para orientar o posicionamento dos demais objetos. A técnica PLP utiliza força bruta para posicionar o subconjunto de amostras, distribuindo os objetos restantes são projetadas utilizando-se operadores laplacianos obtidos a partir de grafos desconexos da vizinhanças. No entanto, alterações na projeção, exigidas pelo usuário, impactam em atualizações dos grafos, aumentando o custo computacional. A metodologia LAMP, no entanto, faz uso de amostras (Pontos de Controle) e sua localização no espaço de projeção inicial, construindo, posteriormente, um conjunto de mapeamentos ortogonais afins, um para cada instância a ser projetada. Assim, tais Pontos de Controle podem ser manipulados para a distorção do espaço de projeção inicial, ocasionando no reposicionamento das instâncias vinculadas a sua vizinhança.

Em (BRANDOLI et al., 2010) é proposto um sistema onde as imagens são mapeadas em um espaço bidimensional de acordo com as características extraídas. Através de um processo iterativo e com a participação ativa do usuário, é avaliado se o resultado obtido é suficiente, exigindo, caso necessário, a alteração da técnica empregada na extração das características das imagens. A Figura 3.8 ilustra o sistema proposto, sendo este distribuído nas etapas de: (1) extração do vetor de características das imagens; (2) projeção dos vetores em um espaço bidimensional; (3) cálculo da silhueta (relação de distancias inter-grupo e intra-grupo) para verificar a qualidade da projeção; e (4) mudança dos parâmetros de extração de características de imagens. A Figura 3.9 demonstra um exemplo da aplicação do sistema, levando-se em conta um cenário com 70 imagens distribuídas igualmente em 7 classes distintas, sendo que 5 delas apresentam boa identificação e 2 encontram-se misturadas.

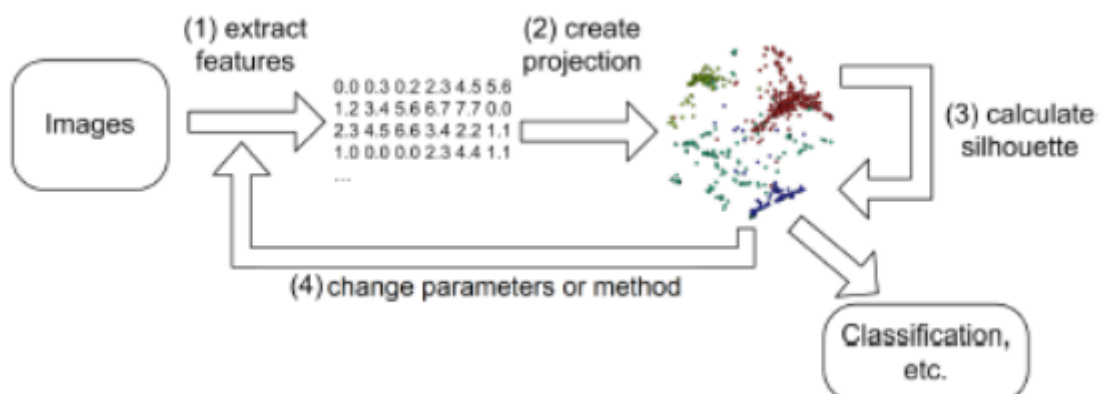


Figura 3.8: Imagem retirada de (BRANDOLI et al., 2010). Ilustração do sistema proposto.

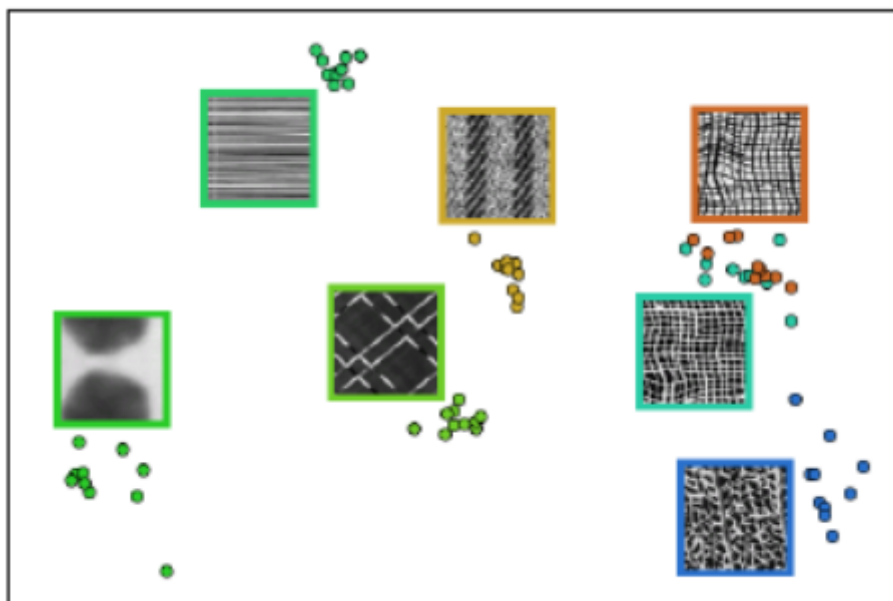


Figura 3.9: Imagem retirada de (BRANDOLI et al., 2010). Exemplo de visualização de 7 classes de imagens distintas, sendo que 5 delas apresentam boa identificação e 2 encontram-se misturadas.

3.3.2 Visualização Baseada em Hierarquia

Uma possibilidade de visualização de imagens é disponibilizá-las de forma hierárquica. Em outras palavras, realizar uma projeção das imagens através da representação encontrada na etapa de extração de características, levando-se em conta uma hierarquia estabelecida pela similaridade entre as mesmas. Desta forma, a construção dessas relações de similaridade pode assumir uma estrutura semelhante a de uma árvore, a qual pode ser estruturada segundo diferentes metodologias.

Segundo (NGUYEN; HUANG, 2002) a visualização em estrutura de árvore pode ser classificada em duas técnicas principais, sendo estas definidas como *Connection* e *Enclosure*.

Connection

A árvore é construída através da utilização de um diagrama denominado *node-link*. Um conjunto de arestas são desenhadas no diagrama para estabelecer um vínculo entre nós pais e nós filhos, sendo estas responsáveis pela indicação do relacionamento entre os mesmos.

A vantagem de utilização desta técnica de visualização de estrutura em árvore é a fácil interpretação humana, visto que a estrutura de relacionamento entre os nós é trivialmente reconhecida e definida pelas arestas.

Dentre os exemplos encontrados na literatura para a construção da hierarquia em árvores

pela utilização do diagrama *node-link* são: *hyperbolic-tree* (LAMPING; RAO; PIROLI, 1995) e *Cone-Tree* (ROBERTSON; MACKINLAY; CARD, 1991).

Uma das abordagens encontradas na literatura para a construção da árvore de similaridade consiste na utilização do método *Neighbor Joining* (NJ) (SAITOU; NEI, 1987). A ideia consiste em, através do cálculo da matriz de distância, empregar a construção de uma árvore filogenética. Uma árvore filogenética tem o propósito de expressar relações de similaridade evolucionária, aplicando o princípio da evolução mínima. A evolução mínima refere-se a um conceito baseado na análise de diferentes topologias da árvore filogenética na tentativa de minimizar a quantidade total de mudanças evolucionárias.

O método de NJ pressupõe a construção de uma árvore sem raiz definida, onde cada par de objetos mapeados são unidos a um ancestral hipotético em comum. Desta forma, os nós folhas passam a ser representativos de objetos; nós internos representativos de objetos ancestrais hipotéticos; e o tamanho das arestas correspondente à similaridade entre os objetos, sejam esses nós-folha ou nós internos. A Figura 3.10 ilustra um exemplo produzido pela ferramenta *PEx-Image* (*Projection Explorer for Images*) (ELER et al., 2009), o qual também possibilita múltiplas visualizações do espaço de projeção de imagens.

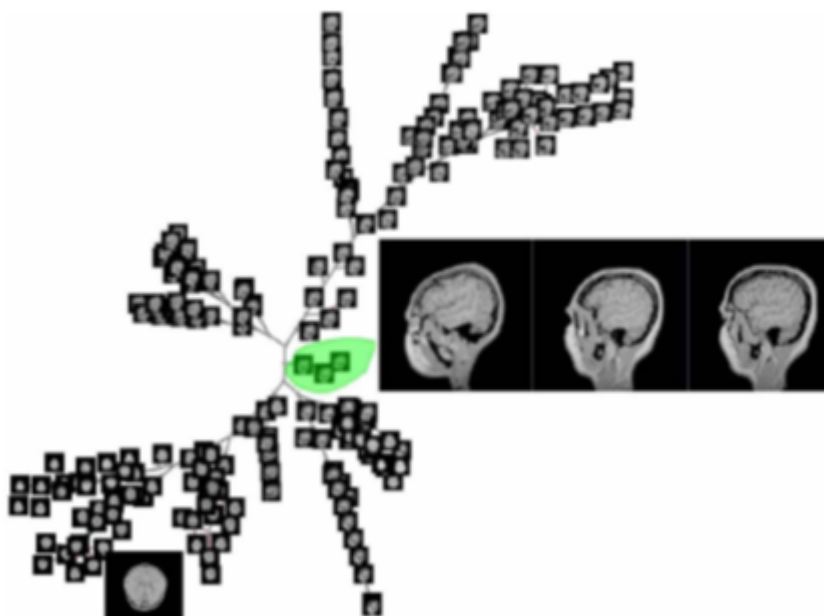


Figura 3.10: Imagem retirada de (ELER et al., 2009). Ilustração de uma possível projeção gerada pela ferramenta *PEx-Image*, a qual permite verificar a árvore filogenética e visualizar imagens pertencentes a uma dada evolução (ramificação).

Levando-se em conta os procedimentos levantados em (CUADROS et al., 2007), a construção da árvore filogenética pelo método NJ, consiste em, inicialmente, construir uma árvore em formato de estrela, onde as N folhas são conectadas a um nó inicial. Juntamente, é construída

a matriz de distância D , onde cada elemento (i, j) representa a distância (similaridade) calculada entre os objetos i e j .

Em cada iteração do processo de construção da árvore, é verificado o menor valor para a soma de dois ramos que relacionam os nós i e j (Equação 3.1). Desta forma, é inserido um nó interno X , com i e j sendo os nós filhos e o novo nó conectado como um ancestral em comum de i e j .

$$S_{ij} = \frac{1}{2(N-2)} \sum_{k \neq i, j} (D_{ik} + D_{jk}) + \frac{D_{ij}}{2} + \frac{1}{N-2} \sum_{k, l \neq i, j}^{k < l} D_{kl} \quad (3.1)$$

Com a inserção do novo nó, a distância entre i e X e a distância entre j e X são calculadas (Equações 3.2 e 3.3).

$$D_{iX} = \frac{D_{ij} + \frac{\sum_{k \neq j} D_{ik}}{N-2} - \frac{\sum_{k \neq i} D_{jk}}{N-2}}{2} \quad (3.2)$$

$$D_{jX} = \frac{D_{ij} + \frac{\sum_{k \neq i} D_{jk}}{N-2} - \frac{\sum_{k \neq j} D_{ik}}{N-2}}{2} \quad (3.3)$$

A matriz de distância D é então atualizada, substituindo os objetos i e j por X e recalculando a nova distância de X para os demais nós Y que constam na matriz (Equação 3.4).

$$D_{XY} = \frac{D_{iY} + D_{jY}}{2} \quad (3.4)$$

Os passos de união de nós são realizados até o momento em que existirem apenas dois nós na matriz de distância D .

Enclosure

Uma possibilidade de representação da hierarquia entre imagens é através da utilização do conceito de *enclosure*. Assim, objetos que encontram-se em um nível mais elevado da hierarquia englobam os demais objetos de níveis inferiores da hierarquia. Desta forma, a cada nova inserção de uma imagem na hierarquia, a área de visualização destinada para um nível superior é subdividida para comportar o novo objeto inserido.

Um exemplo de aplicação da técnica de *enclosure* para a representação hierárquica de imagens é vista na metodologia *tree-maps* (JOHNSON; SHNEIDERMAN, 1991). A abordagem consiste em mapear cada nó em uma área retangular, sendo esta subdividida a cada nova requisição de

inserção de um nó filho. O tamanho da área utilizada por cada nó está relacionada diretamente com a importância do conteúdo do novo objeto inserido na hierarquia. A Figura 3.11 (a) apresenta a utilização da abordagem *nested tree-maps*, indicando uma visualização explícita para a hierarquia e a Figura 3.11 (b) ilustra a utilização da *tree-map* convencional, tendo a estrutura da árvore implícita e exibindo somente os nós folha para melhor proveito da área de projeção.

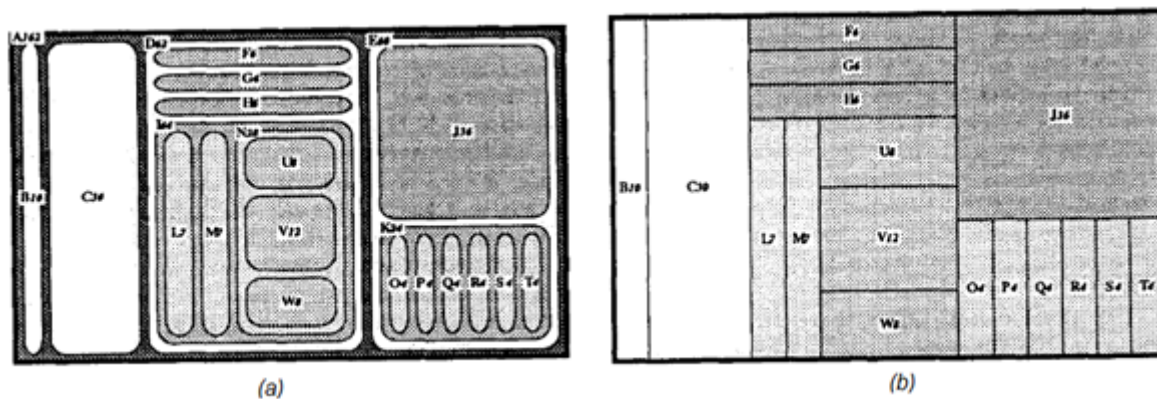


Figura 3.11: Imagem adaptada de (JOHNSON; SHNEIDERMAN, 1991). Exemplo da aplicação da técnica *nested tree-map* (a) e da *tree-map* convencional (b).

3.3.3 Avaliação da Qualidade da Visualização

Apesar do principal objetivo da etapa de visualização da informação ser de caráter subjetivo, certas métricas podem ser utilizadas para avaliar a qualidade da visualização.

Considerando-se a utilização de transformações geométricas para a projeção de imagens em um espaço de dimensionalidade reduzida, certos agrupamentos (*clusters*) podem ser identificados. Logo, duas medidas de distância podem ser utilizadas para verificar a distribuição dessas projeções no espaço, sendo estas definidas por relações intra-grupo e inter-grupo. As relações intra-grupo estão associadas às distâncias entre elementos pertencentes a um mesmo grupo (*cluster*), enquanto as relações inter-grupo avaliam a distância entre elementos de grupos diferentes.

Uma métrica que calcula a correlação entre as relações intra-grupo e inter-grupo é definida como silhueta (KAUFMAN; ROUSSEEUW, 2009). Por definição, a silhueta é dada como a delimitação de cada grupo (*cluster*) no espaço de projeção, possibilitando verificar quais objetos pertencem a dado grupo e quais objetos não possuem uma boa relação ao grupo devido a sua distância. Logo a silhueta média (S_{med}) da projeção é dada por:

$$S_{med} = \frac{1}{N} \sum_{i=1}^N \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (3.5)$$

Onde $a(i)$ representa a média da distância de dado objeto i relativo ao grupo A a qual pertence (relação intra-grupo); $b(i)$ a distância mínima entre o objeto i e os demais agrupamentos, com exceção do grupo A (relação inter-grupo); e N é o número de elementos da projeção.

3.4 Considerações Finais

Neste capítulo, foi apresentado o conceito de Mineração Visual de Imagens e a importância do mesmo para o problema em questão. Foram também levantadas as técnicas que envolvem a MVD e sua aplicação dentro do contexto de imagens em que este projeto de mestrado é proposto. Foram apresentadas as técnicas de visualização e interação que permitem proporcionar um melhor aproveitamento da participação ativa do usuário na tarefa de extração de conhecimento. Adicionalmente, foi discutido como pode ser feita a avaliação da qualidade da(s) técnica(s) de visualização baseadas em projeção.

Parte II

Trabalho Desenvolvido

Capítulo 4

TRABALHO DESENVOLVIDO

O trabalho desenvolvido neste projeto de mestrado permite a redução da descontinuidade semântica (Gap Semântico) de sistemas de Consulta por Conteúdo de Imagem (CBIR) através da utilização do fator de diversidade e técnicas de Mineração Visual de Dados (MVD). A MVD é usada para prover maior interação do usuário na consulta, permitindo também um melhor entendimento dos parâmetros de similaridade e diversidade utilizados. O usuário, pela metodologia proposta, passa a ter um papel ativo na consulta, agindo de acordo com seu julgamento e conhecimento prévio, aproximando os resultados para o esperado. Na Seção 4.1 deste capítulo são dadas as considerações iniciais acerca do trabalho desenvolvido. Na Seção 4.2 é discutido detalhadamente o método proposto e os conceitos envolvidos no mesmo. Na Seção 4.3 são enumeradas algumas das limitações encontradas para o trabalho desenvolvido e, na Seção 4.4, são dadas as considerações finais deste capítulo.

4.1 Considerações Iniciais

Neste capítulo são apresentadas as técnicas desenvolvidas neste projeto de mestrado, as quais propõem a redução da descontinuidade semântica (Gap Semântico) em sistemas CBIR. Um novo método para realizar consultas de k vizinhos diversos mais próximos (*k-Nearest Diverse Neighbors*, k -NDN) foi proposto, utilizando a análise da dimensão fractal do conjunto de dados e adicionando de técnicas de MVD. Logo, é esperado que os resultados estejam mais próximos do esperado pela cognição humana, visto que o usuário passa a atuar ativamente na consulta, fornecendo conhecimento prévio para melhorar a semântica dos resultados.

O método proposto realiza um corte inteligente do espaço onde é realizada a consulta, baseando-se na dimensão fractal do conjunto de imagens. Desta forma, o subconjunto de dados

é fixado em um primeiro momento, possibilitando uma melhor compreensão dos resultados e permitindo que os mesmos sejam avaliados através de um algoritmo concorrente, diminuindo o tempo de execução da consulta.

Através de técnicas de Projeção Multidimensional, uma projeção intermediária de determinadas instâncias do conjunto de imagens (Pontos de Controle) é dada, juntamente com a projeção da imagem de consulta. Neste momento, o usuário pode manipular tais instâncias e direcionar os resultados do sistema CBIR de acordo com seu conhecimento prévio sobre o conjunto de dados, aproximando instâncias/categorias que apresentam algum grau de similaridade de acordo com o julgamento empregado.

É apresentado o método proposto que foi chamado de *Visual-interactive k-NDNq Method* (Vik), o qual permite que o usuário refaça a consulta através da manipulação/distorção de informações apresentadas em uma projeção inicial, sendo esta utilizada para refazer a consulta com o uso da nova distribuição das imagens. Isso torna o ambiente CBIR mais flexível possibilitando que o mesmo retorne resultados mais próximos do que é desejado pelo usuário.

Uma projeção final é apresentada e técnicas de interação provenientes da MVD são adicionadas para promover uma melhor interatividade na exploração das informações projetadas. São também fornecidos artifícios para evidenciar aspectos relacionados aos resultados do sistema CBIR, a distribuição do conjunto de dados e as imagens que foram consideradas no cálculo da relação entre similaridade diversidade na consulta pelos k vizinhos diversos mais próximos.

Resultados desconsiderando as técnicas de MVD empregadas no método proposto são também fornecidos para o usuário comparar as vantagens do trabalho desenvolvido.

4.2 Detalhamento do Trabalho Desenvolvido

Aplicar um fator de diversidade na consulta por similaridade, como levantando na Seção 2.8, pode auxiliar na redução da descontinuidade semântica, diminuindo a discrepância dos resultados obtidos em relação ao esperado pelo usuário. Contudo, adicionar diversificação na consulta exige um alto custo computacional, inviabilizando a análise de todas as possibilidades de resultados dentro do conjunto de imagens.

Avaliar o impacto de parâmetros que relacionam similaridade e diversidade na consulta é de difícil compreensão para o usuário. Além disso, o conhecimento e julgamento do usuário não são utilizados neste processo para que o mesmo possa compreender e guiar e orientar a consulta para obter um resultado mais satisfatório.

Neste trabalho de mestrado, foi desenvolvida a metodologia denominada ViK. A mesma não só propõe uma forma mais intuitiva e menos custosa para execução de consultas por vizinhos diversos mais próximos (k NDN), mas também possibilita que o usuário participe mais ativamente durante o processo da consulta e exploração dos resultados. Através do conhecimento prévio e julgamento do usuário, a consulta pode ser refeita para atingir melhores resultados.

Uma visão geral do método desenvolvido, desconsiderando as técnicas de MVD empregadas, é apresentada na Figura 4.1. O conceito de dimensão intrínseca do conjunto de dados é utilizado para reduzir o custo da adição do fator de diversidade em consultas por vizinhança. Os resultados dessa abordagem consideram os conceitos de similaridade e diversidade aplicados no espaço N -Dimensional, o qual é representado pela cardinalidade N do vetor de características obtido na etapa de extração de características.

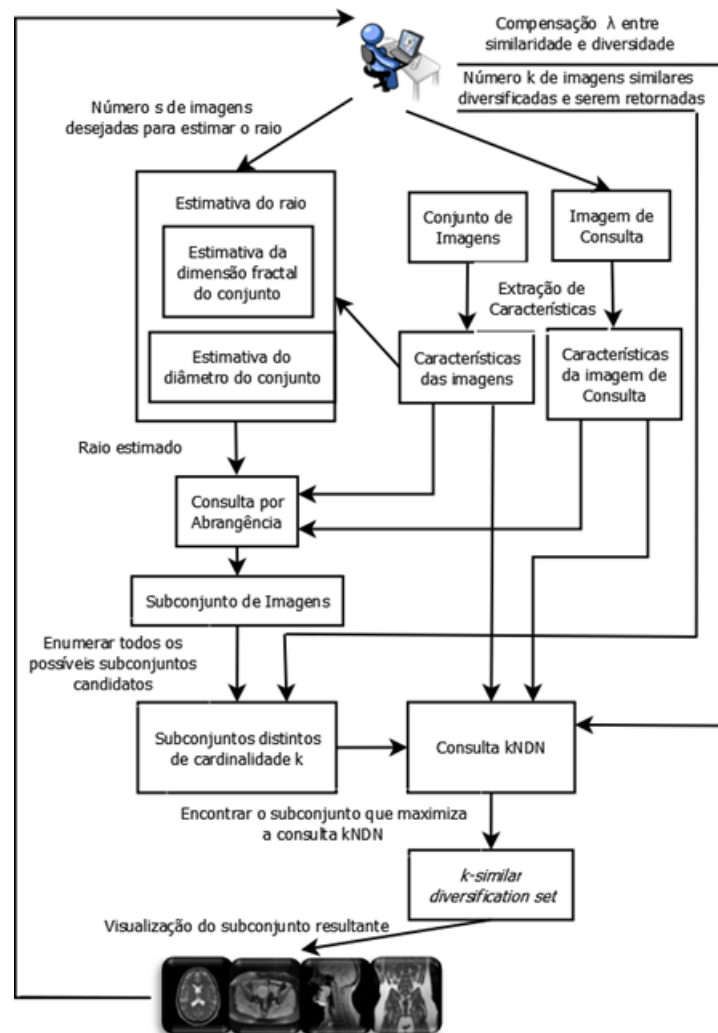


Figura 4.1: Passos do método proposto para a consulta pelos vizinhos diversos mais próximos (k -NDN), considerando o espaço N -Dimensional.

O usuário fornece a imagem de consulta (ver Figura 4.1); o número s de imagens desejadas

a serem estimadas pelo raio na análise da dimensão fractal do conjunto de dados; a relação (*trade-off*) λ entre os fatores de similaridade e diversidade; e o número k de imagens similares diversificadas a serem encontradas. A análise da dimensão fractal do conjunto de imagens é utilizada para estimar um valor para o raio, o qual, através de uma consulta por abrangência (*Range query*), retorna o subconjunto de imagens que serão consideradas na avaliação dos conjuntos de candidatos para a consulta k -NDN. Para cada subconjunto candidato, é calculada, através de uma função de pontuação, a relação entre similaridade e diversidade. O subconjunto que maximiza a função, denominado *k-similar diversification set*, é encontrado e uma visualização das imagens correspondentes é dada ao usuário.

Os resultados obtidos na abordagem N -Dimensional são utilizados para analisar as vantagens da adição de técnicas de MVD em sistemas CBIR. O principal objetivo da metodologia proposta é fornecer maior flexibilidade para o usuário, permitindo que o mesmo direcione os resultados de acordo com o seu julgamento e conhecimento prévio do conjunto de imagens. Adicionalmente, o método proposto fornece uma melhor compreensão do relacionamento da imagem de consulta com as demais imagens do conjunto, fornecendo artifícios para que o entendimento do resultado obtido pelo sistema CBIR e outros fatores que este abrange, como a relação entre similaridade e diversidade. Facilitando a compreensão dos resultados e fatores envolvidos na consulta k -NDN e colocando o usuário como um agente ativo neste processo, o sistema desenvolvido tem como principal funcionalidade a redução da descontinuidade semântica.

Diferentemente do uso de todas as características extraídas do conjunto de imagens, como levantado na Figura 4.1, as técnicas de MVD utilizam informações do espaço de projeção manipulado pelo usuário para verificar o subconjunto de instâncias que melhor responde a consulta k -NDN. Desta forma, os procedimentos realizados e os resultados encontrados passam a utilizar informações oriundas do espaço projetado e distorcido pelo usuário. Uma visão geral do método proposto com a adição das técnicas de MVD é ilustrado na Figura 4.2.

A visualização utilizada no trabalho desenvolvido é baseada na projeção LAMP (JOIA et al., 2011), visto que a mesma possui características que possibilitam maior interatividade (Seção 3.3.1). Desta forma, amostras (Pontos de Controle) representativas de cada categoria do conjunto de imagens são previamente selecionadas e projetadas pela técnica *Fast-Map* (FALOUTSOS; LIN, 1995), juntamente com a projeção da imagem de consulta. O usuário, nesta etapa, tem a liberdade de manipular os Pontos de Controle e a imagem de consulta projetados, ajustando e posicionando instâncias por sua similaridade de acordo com seu julgamento. Finalmente, uma projeção final é dada utilizando como referência a distorção do espaço inicial

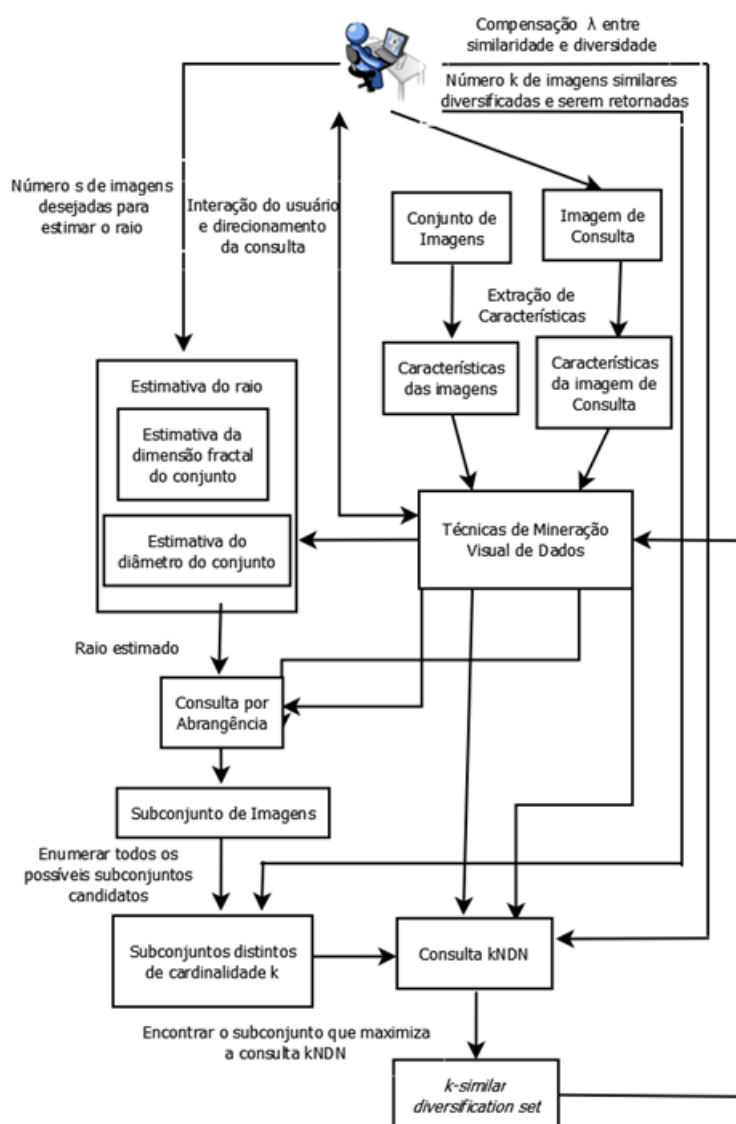


Figura 4.2: Passos do método proposto para a consulta pelos vizinhos diversos mais próximos (k -NDN), considerando a utilização das técnicas de MVD.

de projeção. É então verificado qual subconjunto candidato melhor responde as necessidade da consulta k -NDN de acordo com as informações oriundas da projeção final. Outras técnicas de MVD são também adicionadas nessa etapa para aproveitar o caráter exploratório do espaço de projeção e dos resultados obtidos. A Figura 4.3 ilustra a visualização, interação e obtenção dos resultados guiados pelo usuário segundo a metodologia proposta.

Maior detalhamento sobre a utilização dos conceitos relacionados à análise da dimensão fractal, da consulta por vizinhos diversos mais próximos (k -Nearest Diverse Neighbor Query (k -NDNq)) e das técnicas de MVD empregadas na metodologia ViK são dadas nas próximas subseções.

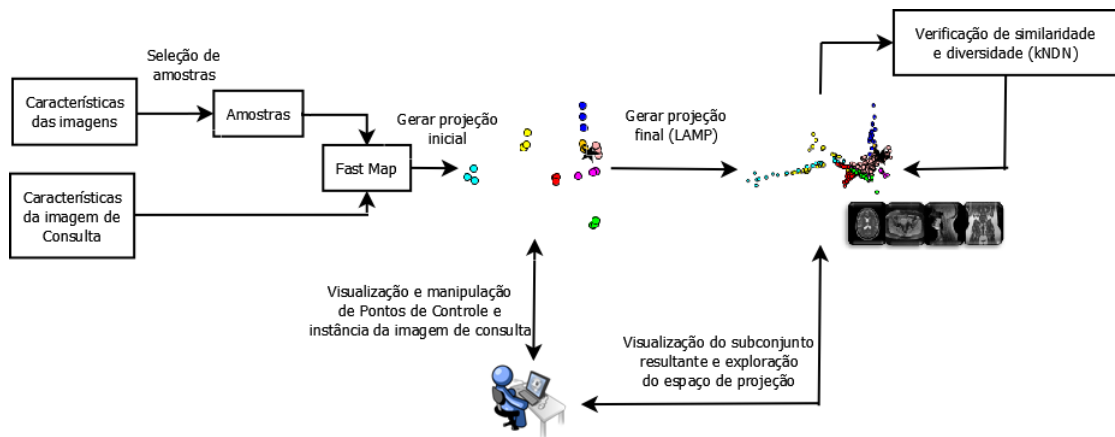


Figura 4.3: Técnicas de visualização e interação empregadas no trabalho desenvolvido.

4.2.1 Dimensão Fractal

A análise da dimensão fractal foi utilizada para reduzir a quantidade de imagens e, consequentemente, o número de possíveis combinações a serem verificadas na consulta k -NDN.

A partir de um determinado número s de elementos, a análise da dimensão fractal do conjunto de imagens pode ser utilizada para estimar um valor para o raio que, a partir de uma consulta por abrangência (*Range query*), retornaria a mesma quantidade s de elementos de uma consulta por vizinhança (kNN) (ARANTES et al., 2003). Um limiar para a cardinalidade do subconjunto de imagens a ser avaliado na consulta k -NDN é estipulado, respeitando, por exemplo, limitações de hardware. Desta forma, o usuário pode, dentro da limitação estabelecida, manipular e escolher um valor de acordo com suas necessidades e conhecimento prévio do conjunto de imagens, possibilitando que o mesmo tenha um papel ativo no estabelecimentos de quais imagens serão analisadas nos resultados respeitando os fatores de similaridade e diversidade.

A dimensão fractal utilizada neste trabalho é uma aproximação da dimensão fractal correlacional calculada de acordo com o descrito em (ARANTES et al., 2003). A dimensão fractal correlacional do conjunto de dados representa a sua dimensionalidade intrínseca e possui a propriedade de aproximar a distribuição das distâncias entre elementos no conjunto de dados (FALOUTSOS; LIN, 1995). O raio r é estimado com base na dimensão fractal de correlação do conjunto de dados D ; no número de objetos N do conjunto de dados; no diâmetro R do conjunto de dados; e no número s de elementos desejados a serem encontrados (ARANTES et al., 2003).

$$r = R \cdot \exp\left(\frac{\log(s(s-1)) - \log(N(N-1))}{D}\right) \quad (4.1)$$

Calcular o diâmetro R do conjunto de dados necessita um alto custo computacional, visto

que, para encontrar a solução ótima global, todas as distâncias entre cada instância deve ser calculada e avaliada. Para reduzir o custo desta tarefa, foi utilizada uma metodologia similar em relação a seleção de "objetos pivot" para a projeção *FastMap* (FALOUTSOS; LIN, 1995). A técnica é baseada na seleção de um elemento arbitrário e é encontrado o elemento do conjunto com maior distância em relação a este (Algoritmo 5.1).

```

public static double estimateDataSetDiameter(double [][]
    values, int maxIterations) {
    double distance;
    double maxDistance = Double.MAX_VALUE * (-1);

    int globalMaxIndex = values.length - 1; int
        localMaxIndex = 0;
    int count = 0;
    do {
        for (int i = 0; i < values.length; i++) {
            distance = CBIR.Distance.EuclideanDistance(
                values[globalMaxIndex], values[i]);
            if (distance > maxDistance) {
                maxDistance = distance;
                localMaxIndex = i;
            }
        }
        globalMaxIndex = localMaxIndex;
        count++;
    } while (count < maxIterations);

    return maxDistance;
}

```

Algoritmo 5.1: Cálculo do diâmetro aproximado para um conjunto de dados.

Segundo o algoritmo para estimar o diâmetro do conjunto de dados (Algoritmo 5.1), são dadas como entrada as informações referentes ao conjunto (*values*) e o número máximo de interações desejadas para a estimativa (*maxIterations*). É ajustado para um valor mínimo a distância encontrada e armazenado o índice referente ao elemento mais distante. Para cada iteração, o elemento referente a maior distância (*globalMaxIndex*) é utilizado para verificar, para

cada elemento do conjunto, aquele que é mais distante (*localMaxIndex*). Caso a distância entre o elemento sendo analisado e o elemento mais distante seja maior do que a distância máxima encontrada (*maxDistance*), a variável de maior distância é atualizada e o índice do elemento mais distante localmente (*localMaxIndex*) passa a ser o elemento corrente. Ao fim da comparação de todos elementos do conjunto, o índice do elemento mais distante (*globalMaxIndex*) é atualizado e uma nova iteração é realizada, caso não tenha excedido o número de iterações especificadas como parâmetro (*maxIterations*). Por fim, o algoritmo tem como retorno a distância máxima (*globalMaxDistance*) estimada entre os elementos do conjunto.

A aproximação da dimensão intrínseca, calculada pela dimensão fractal, neste trabalho, permite reduzir a complexidade do espaço original (N-Dimensional) para um espaço de menor dimensionalidade. Contudo, é importante ressaltar que a dimensão encontrada se difere da dimensão de fato utilizada na projeção e representação das imagens pelas técnicas de MVD empregadas. Dessa forma, é esperado que haja uma perda razoável de informações nessa redução de complexidade dimensional, mas, através das técnicas de MVD e da interação do usuário no processo da consulta, é possível adicionar o conhecimento provindo da cognição humana e obter resultados mais satisfatórios.

4.2.2 *k*-Nearest Diverse Neighbor Query (*k*-NDNq)

Para enumerar cada subconjunto possível analisando-se os elementos retornados pelo corte do espaço promovido pelo raio estimado, é necessária uma lógica combinatória. Para um conjunto de dados com u elementos e uma cardinalidade v desejada, o número de combinações possíveis é dado por:

$$C_{u,v} = \frac{u!}{v! \cdot (u-v)!} \quad (4.2)$$

Para enumerar todas os possíveis conjuntos de combinações candidatos, a técnica proposta em (KNUTH, 2006) foi utilizada.

Devido ao grande número de possíveis conjuntos de combinações a serem listados e analisados (e.g: para um conjunto de dados com 1000 imagens e um subconjunto de cardinalidade desejada igual a 5, temos 8.250.291.250.200 possíveis combinações distintas), um número máximo de combinações deve ser definido. O usuário pode manipular o número de elementos a ser considerado nesta tarefa combinatória, levando-se em considerações a limitação imposta, por exemplo, por questões de recursos computacionais. O valor escolhido, o qual é dado como entrada para a etapa de análise fractal do conjunto de dados, permite com que seja reduzido o

número de elementos a serem considerados no levantamento dos subconjuntos candidatos.

É realizada uma consulta por abrangência (*Range query*) utilizando o raio retornado pela etapa de análise da dimensão fractal, reduzindo o espaço e a quantidade de elementos a serem analisados. Utilizar a consulta por abrangência, neste caso, é mais viável do que uma simples consulta kNN, visto que o raio não precisa ser recalculado em uma nova consulta, já que leva em consideração a distribuição de todo o conjunto de imagens (Equação 4.1). Finalmente, é calculada uma função de pontuação para cada possível combinação de elementos selecionados pela consulta por abrangência (*Range query*).

A função de pontuação proposta em (VIEIRA et al., 2011b) foi utilizada para encontrar o *k-similar diversification set*. Dado $S = \{s_1, \dots, s_n\}$ um conjunto de n elementos; uma compensação (*trade-off*) λ , $0 \leq \lambda \leq 1$, a qual especifica a relação entre similaridade e diversidade; e um número k de elementos desejados para retorno da consulta ($k \leq n$), o conjunto que melhor atende a consulta k -NN (*k-similar diversification set*) é dado por:

$$R = \operatorname{argmax}_{S' \subseteq S, k=|S'|} \Gamma(q, S') \quad (4.3)$$

Onde a função de pontuação para cada subconjunto candidato S' de imagens é dado por:

$$\Gamma(q, S') = (k-1)(1-\lambda) \cdot \operatorname{sim}(q, S') + 2\lambda \cdot \operatorname{div}(S') \quad (4.4)$$

Em que a Similaridade é representada por uma função que verifica a distância entre a imagem de consulta e as demais imagens do subconjunto,

$$\operatorname{sim}(q, S') = \sum_{i=1}^k \delta_{\operatorname{sim}}(q, s_i), s_i \in S' \quad (4.5)$$

E a Diversidade verificada através de uma função que calcula a distância entre todas as imagens do subconjunto.

$$\operatorname{div}(S') = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \delta_{\operatorname{div}}(s_i, s_j), s_i, s_j \in S' \quad (4.6)$$

No método proposto, S é o subconjunto de imagens obtido pela consulta por abrangência (*Range query*) através do raio estimado pela análise da dimensão fractal do conjunto de dados. S' é o subconjunto de imagem que representa uma possível combinação de k elementos de S , com base no algoritmo proposto em (KNUTH, 2006).

Finalmente, considerando a proposta de solução para o espaço N-dimensional (Figura 4.1), as imagens correspondentes ao *k-similar diversification set* são dadas ao usuário.

No caso da abordagem em que são utilizadas as técnicas de MVD (Figura 4.2), os mesmos procedimentos são realizados, apresentando como diferença a utilização das informações do espaço de projeção distorcido pelo usuário ao invés de todas as características extraídas para cada instância do conjunto de imagens.

4.2.3 Técnicas de Mineração Visual de Dados

Para utilizar o conhecimento do usuário através da visualização e interação com as instâncias de conjuntos de imagens, a técnica de Projeção Multidimensional LAMP (JOIA et al., 2011) foi considerada. É dada uma projeção inicial com algumas amostras de instâncias do conjunto de imagens (Pontos de Controle), juntamente com a instanciação da imagem de consulta, para que o usuário possa manipular o posicionamento de tais instâncias e direcionar o resultado da consulta *k*-NDN. A Figura 4.4 demonstra um exemplo de manipulação de instâncias na projeção inicial do método proposto, na qual, a partir da projeção inicial dada (Figura 4.4a), o usuário manipula o posicionamento das instâncias (incluindo a imagem de consulta) de acordo com o seu julgamento sobre a similaridade entre imagens/categorias do conjunto de dados (Figura 4.4b).

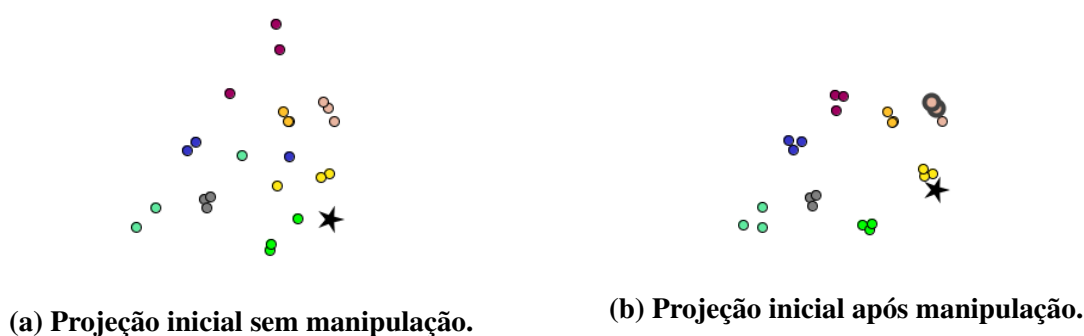


Figura 4.4: Exemplo de movimentação de instâncias na projeção inicial.

Para auxiliar a atividade de distorção e manipulação do espaço de projeção inicial, o usuário pode requisitar a visualização do identificador ou uma prévia da imagem relativa à instância projetada. Apesar de serem ilustrados casos em que as imagens estão previamente classificadas, este mecanismo de visualização e interação pode também permitir a análise de imagens não classificadas, possibilitando que o usuário faça essa classificação manualmente pela alteração no posicionamento das instâncias. A Figura 4.5 ilustra um exemplo das possíveis visualizações de informações relativas às instâncias da projeção.

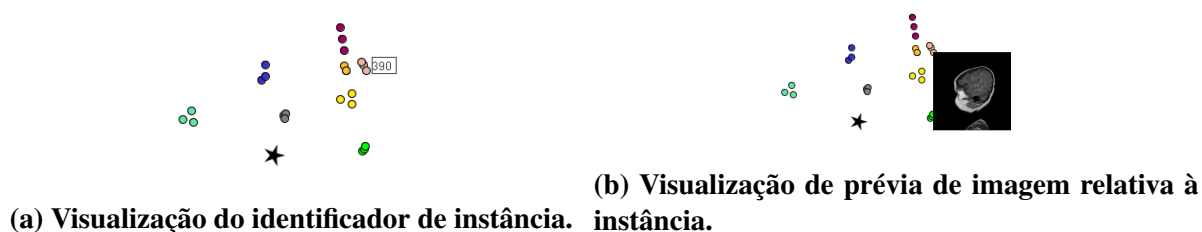


Figura 4.5: Exemplo de visualização do identificador ou prévia de imagem relativa à instância requisitada pelo usuário.

Utilizando as informações oriundas do espaço de projeção inicial distorcido e a porcentagem de vizinhos desejada, assim como na técnica LAMP (JOIA et al., 2011), uma visualização do espaço de projeção final é fornecida ao usuário. As informações relativas ao posicionamento de cada instância na projeção final são utilizadas para, finalmente, encontrar o subconjunto de imagens que satisfaz a consulta k -NDN, assim como ilustrado na Figura 4.2 da Seção 4.2.

O resultado obtido pelo direcionamento do usuário na utilização das técnicas de MVD é exibido, juntamente com o resultado que seria obtido considerando o espaço N-Dimensional (Figura 4.1) representado pelas características extraídas para as imagens do conjunto. Para evidenciar as vantagens da orientação do usuário na obtenção de resultados da consulta k -NDN, artifícios adicionais são fornecidos para evidenciar fatores envolvidos nesse processo (Figura 4.6). Logo, é dada a possibilidade de visualizar, na projeção final, informações referentes a imagem de dada instância (Figura 4.6a); centroides de cada agrupamento/categoria (Figura 4.6b); Pontos de Controle utilizados para a projeção final (Figura 4.6c); raio de corte do conjunto utilizado na consulta k -NDN (Figura 4.6d); evidenciação do resultado obtido considerando o espaço de projeção (Figura 4.6e); e evidenciação do resultado obtido considerando o espaço N-Dimensional (Figura 4.6f).

Além de reduzir o Gap Semântico pela participação ativa do usuário na consulta, o trabalho desenvolvido permite uma melhor compreensão do impacto da adição do fator de diversidade na consulta k -NDN. Desta forma, é possível verificar quais imagens foram consideradas para o levantamento dos subconjuntos candidatos ao (k -similar diversification set), além da proximidade (similaridade) em que estas se encontram.

Um exemplo de verificação do impacto do aumento da diversidade e dos elementos considerados pelo raio da consulta k -NDN é ilustrado na Figura 4.7, alterando-se os valores do *trade-off* entre similaridade e diversidade de $\lambda = 0,15$ (Figura 4.7a) e $\lambda = 0,55$ (4.7b). É requisitado pelo usuário a visualização do raio de abrangência da consulta e os resultados considerando os espaços de projeção e N-Dimensional. O usuário pode também visualizar imagens relativas a instâncias específicas da projeção, facilitando a compreensão de quais imagens foram

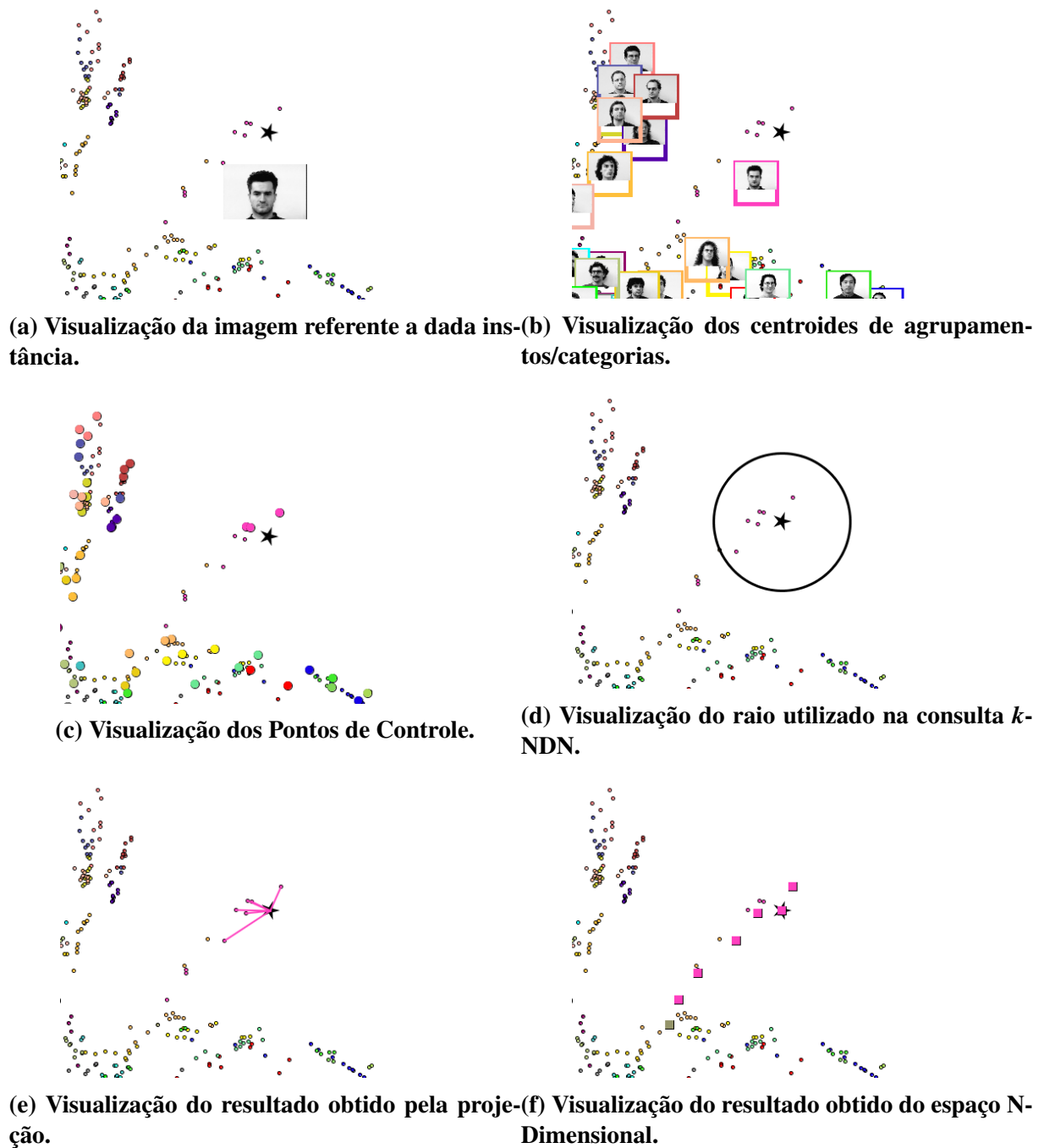


Figura 4.6: Técnicas de visualização e interação utilizadas na projeção final.

consideradas para realização da consulta.

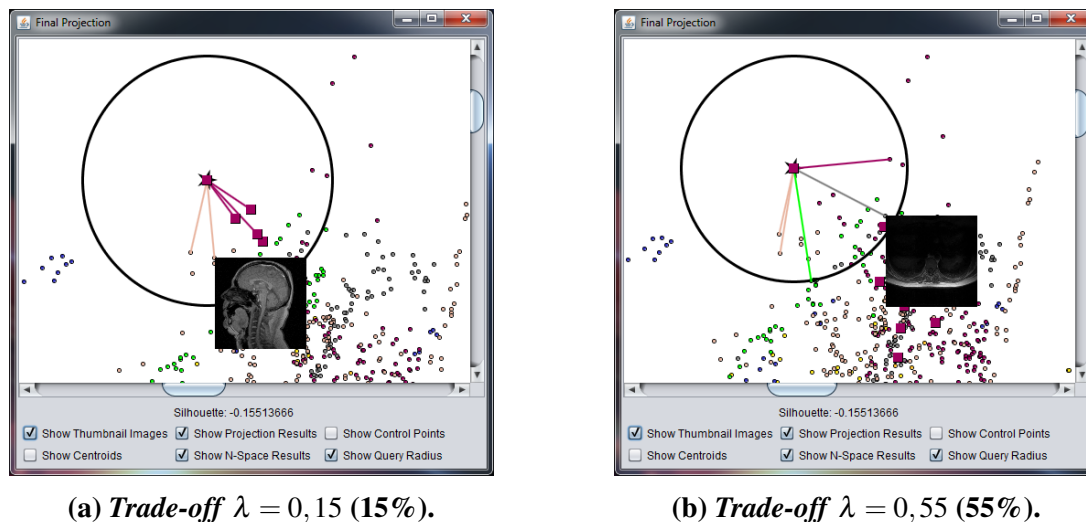


Figura 4.7: Exploração da consulta k -NDN aumentando o valor do *trade-off* entre similaridade e diversidade.

Um exemplo da aplicação do trabalho desenvolvido com a utilização de técnicas de MVD (4.2) é ilustrado na Figura 4.8.

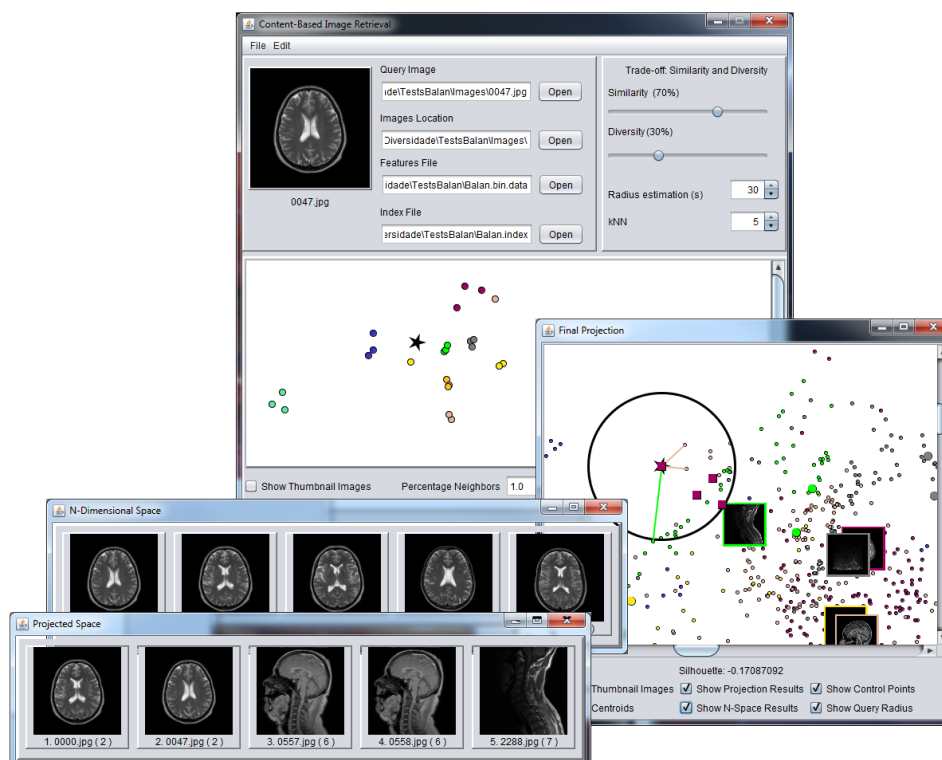


Figura 4.8: Exemplo de aplicação do trabalho desenvolvido.

4.3 Limitações do Trabalho Desenvolvido

Ao longo do desenvolvimento do método demonstrado neste trabalho de mestrado, algumas limitações puderam ser observadas.

Redução da complexidade

A utilização aproximação da dimensão intrínseca, calculada através da dimensão fractal, permite que a complexidade da consulta por conteúdo de imagens, adicionando-se o fato de diversidade, seja reduzido. Contudo, ao projetar tais informações em um espaço bidimensional (espaço de projeção), existe uma perda razoável de informações, visto que parte da representatividade de cada imagem é perdida. Tal evidenciação fica clara ao obter-se um valor aproximado para a dimensão fractal do conjunto de imagens superior a $D = 2$.

Apesar das dificuldades encontradas, ao utilizar o conhecimento do usuário, é esperado que o comprometimento das informações representativas das imagens seja reduzido. Isso ocorre devido ao fato de que é dada a possibilidade do usuário interagir e distorcer o espaço inicial de projeção, adicionando informações mais pertinentes e direcionando os resultados de acordo com o seu julgamento.

Escolha da dimensionalidade da projeção

Na tarefa de redução da dimensionalidade para o contexto em que o conjunto de imagens se encontra, foram avaliadas as diferentes possibilidades do espaço final de projeção, tendo este uma dimensionalidade igual a d . Contudo, para que seja feito um maior aproveitamento da capacidade de percepção do usuário, a dimensionalidade escolhida deve ser $d = 1, 2, 3$.

Caso a dimensionalidade do espaço final escolhido fosse $d = 1$, haveria uma grande perda de informações, já que toda a representação obtida no processo de extrações de características ficaria reduzida a somente uma dimensão. Dessa forma, não seria possível obter uma boa distinção e representatividade de cada imagem do conjunto.

Se fosse escolhida a dimensionalidade $d = 3$, uma maior quantidade de informações poderia ser exibida, reduzindo a perda de informações. Desta forma, a visualização considerando-se um espaço desta dimensionalidade seria conveniente para a compreensão do usuário. Não obstante, a tarefa de manipulação e distorção do espaço é dificultada, visto que recursos teriam que ser adicionados para movimentação vertical, horizontal e também em profundidade, tornando-se uma atividade de grande complexidade quando envolvido um grande número de instâncias na projeção.

Devido às características apontadas para $d = 1$ e $d = 3$, foi escolhido o espaço bidimensi-

onal ($d = 2$) para a projeção das informações relativas às imagens. Dessa forma, foi possível projetar mais informações do que para $d = 1$ e ainda assim manter uma maior facilidade de manipulação das projeções se comparado ao espaço tridimensional ($d = 3$).

Representação de grande quantidade de categorias

Na projeção inicial, são dadas instâncias relativas a cada categoria do conjunto de imagens, sendo estas denominadas "Pontos de Controle". Dessa forma, cada categoria fica representada por um pequeno subconjunto de amostras com o intuito de permitir a manipulação do usuário, permitindo que este determine a proximidade entre as instâncias de acordo com seu julgamento.

Um problema surge no cenário em que uma grande quantidade de categorias é observado. Tal fato é encontrado devido a dificuldade de distinção entre cada categoria, a qual, neste trabalho, foi feita através da utilização de cores distintas. Dessa forma, a diferença visual entre as projeções de diferentes categorias passa a ser menor, aproximando categorias distintas a uma mesma faixa de cor. Com isso, a tarefa de manipulação de categorias feitas pelo usuário se torna mais complexa.

Para resolver o problema da projeção de grande quantidades de categorias, poderiam ser levados em consideração outras técnicas como a restrição de categorias mais próximas à imagem de consulta, ou até mesmo a seleção prévia de categorias desejadas na avaliação feitas pelo próprio usuário, por exemplo.

4.4 Considerações Finais

Neste capítulo foi descrito como este projeto de mestrado foi conduzido. É apresentado o método ViK, pelo qual o usuário passa a ter um papel ativo na consulta por conteúdo de imagem, direcionando o resultado de acordo com seu julgamento e conhecimento prévio sobre o conjunto de dados, sendo este possível através da adição de técnicas de MVD. Adicionalmente, é viabilizado uma melhor compreensão da aplicação do fator de diversidade em consultas pelos vizinhos diversos mais próximos (k -NDN), permitindo melhor entendimento sobre quais imagens foram consideradas no processo. A seleção de possíveis candidatos ao subconjunto que atende mais adequadamente à consulta k -NDN é feita através da análise da dimensão fractal do conjunto de dados, permitindo encontrar um raio que reduza o espaço em que a consulta é aplicada. Para avaliar as vantagens da utilização das técnicas de MVD envolvidas, os resultados relativos ao espaço N-Dimensional, representado pelas características extraídas do conjunto de imagens, e ao espaço de projeção são dados para comparação.

Capítulo 5

RESULTADOS OBTIDOS

Para avaliar a aplicabilidade do método proposto, neste capítulo são demonstrados os estudos de caso elaborados com diferentes conjuntos de dados. É demonstrado não somente as vantagens obtidas pela proposta da adição do fator de diversidade em consultas por vizinhança (kNN), mas também os benefícios adicionais pelo uso de técnicas de Mineração Visual de Dados (MVD). Na Seção 5.1 deste capítulo são discutidas as considerações iniciais acerca dos experimentos e resultados obtidos para a metodologia proposta. Na Seção 5.2 são levantados os diferentes estudos de caso avaliados. A Seção 5.3 apresenta os resultados levando-se em consideração o custo computacional observado nos experimentos realizados e, na Seção 5.4, são dadas as considerações finais e importância dos resultados apresentados neste capítulo.

5.1 Considerações Iniciais

Para avaliar os benefícios e ganhos obtidos com a metodologia proposta neste trabalho de mestrado, foram levantadas diferentes situações e contextos em que podem ser aplicadas as técnicas desenvolvidas. Desta forma, é demonstrado não somente as possíveis vantagens da abordagem demonstrada, mas também a extensibilidade e possibilidade de aplicação das mesmas em conjuntos de imagens com temas variados.

Neste capítulo, são apresentados os artifícios utilizados para avaliação das técnicas apresentadas neste trabalho de mestrado, assim como os resultados obtidos. São demonstrados aspectos relacionados à qualidade e compreensão do usuário com relação à consulta e o conjunto de dados, assim como o custo e o ganho computacional obtidos pela metodologia desenvolvida.

5.2 Estudos de Caso

Nos experimentos realizados foram sugeridos diferentes conjuntos de imagens. Como principal objetivo, os esforços foram principalmente direcionados à avaliação do impacto para a área de imagens médicas.

Proporcionar uma melhor compreensão e maior participação do usuário em consultas por conteúdo de imagens, no contexto de imagens médicas, pode permitir melhores resultados no diagnóstico de determinada imagem de consulta. Adicionalmente, pode-se também ser avaliada a relação entre diferentes categorias de anomalias entre imagens, permitindo o questionamento de possíveis proximidades entre as mesmas.

Para demonstrar a extensibilidade do trabalho desenvolvido, foi também levado em consideração um outro contexto de imagens, sendo este representado por um conjunto de imagens de faces humanas. Desta forma, é demonstrado que é também possível obter resultados mais vantajosos pelas técnicas propostas neste projeto de mestrado.

Nas demais subseções, são dados os estudos de caso relacionados a dois conjuntos de imagens médicas: Balan704 Dataset (seção 5.2.1) e ROIs Dataset (5.2.2) e ao conjunto de faces FullFaces Dataset (5.2.3). São avaliadas diferentes situações em que a metodologia proposta pode auxiliar na compreensão dos resultados, como vantagens das restrições impostas pela redução do espaço original do conjunto de imagens e os ganhos obtidos com a adição de técnicas de MVD.

5.2.1 Estudo de Caso 1 (Balan704 Dataset)

Para avaliar os resultados do trabalho desenvolvido para a área médica, neste estudo de caso, é levado em conta o conjunto de imagens Balan704. Este fora obtido por contribuição do Hospital das Clínicas de Ribeirão Preto (HCRP) da Universidade de São Paulo (USP).

O conjunto Balan704 é composto por 704 imagens de ressonância magnética (RM), divididas em oito categorias distintas. Cada categoria representa um subconjunto de imagens de diferentes sessões do corpo humano, como detalhado na Tabela 5.1.

A representação numérica utilizada para as imagens desse conjunto foi obtida segundo o apontado em (BALAN, 2007). Nessa abordagem, as imagens foram divididas em cinco regiões de segmentação, sendo que, para cada região, foram extraídas seis características, como a massa, o centro de massa e o nível de cinza médio. Segundo (BALAN, 2007), tais características são as mais representativas para esse conjunto específico de dados.









1. Angiograma	
2. Pélvis Axial	
3. Cabeça Axial	
4. Cabeça Sagital	
5. Abdômen Coronal	
6. Cabeça Coronal	
7. Espinha Sagital	
8. Abdômen Axial	

Tabela 5.1: Categorias presentes no conjunto de imagens Balan704.

Neste estudo de caso, desejou-se avaliar as vantagens da metodologia proposta para a avaliação de determinada anomalia em uma região específica do corpo humano. Desta forma, é esperado que os resultados não só sejam condizentes com a necessidade do especialista, mas também que as técnicas de MVD utilizadas facilitem compreender os resultados obtidos e a relação entre outras imagens no conjunto de dados.

Como experimento, foi considerada a imagem ilustrada na Figura 5.1, pretendendo-se analisar uma possível anomalia relacionada com a região da cabeça. Logo, é esperado que, ao aumentar a proporção de diversidade na consulta, imagens mais diversificadas sejam encontradas. Todavia, devido à restrição imposta inicialmente pela redução do conjunto de imagens, uma similaridade mínima é imposta na consulta.



Figura 5.1: Imagem de consulta utilizada para o estudo de caso do conjunto de imagens Balan704.

Na Figura 5.2, são dados os resultados obtidos considerando a abordagem que utiliza o espaço de características (Figura 4.1 da Seção 4.2). Na Figura 5.2a, pode-se notar que as imagens retornadas pertencem a mesma categoria (Cabeça Axial), visto que não foi imposto nenhum fator de diversidade e, desta forma, a consulta foi feita apenas considerando a similaridade entre as imagens. Ao considerar uma pequena proporção de diversidade na consulta (Figura 5.2b), as

imagens obtidas possuem maior diversidade entre si, porém ainda se restringem a mesma categoria. Adicionando-se um fator de diversidade moderado (Figura 5.2c), imagens ainda mais diversas foram obtidas para a categoria Cabeça Axial. Não obstante, ao ser imposta uma alta diversidade na consulta (Figura 5.2d), uma imagem de outra categoria (Cabeça Sagital) foi obtida (em destaque).

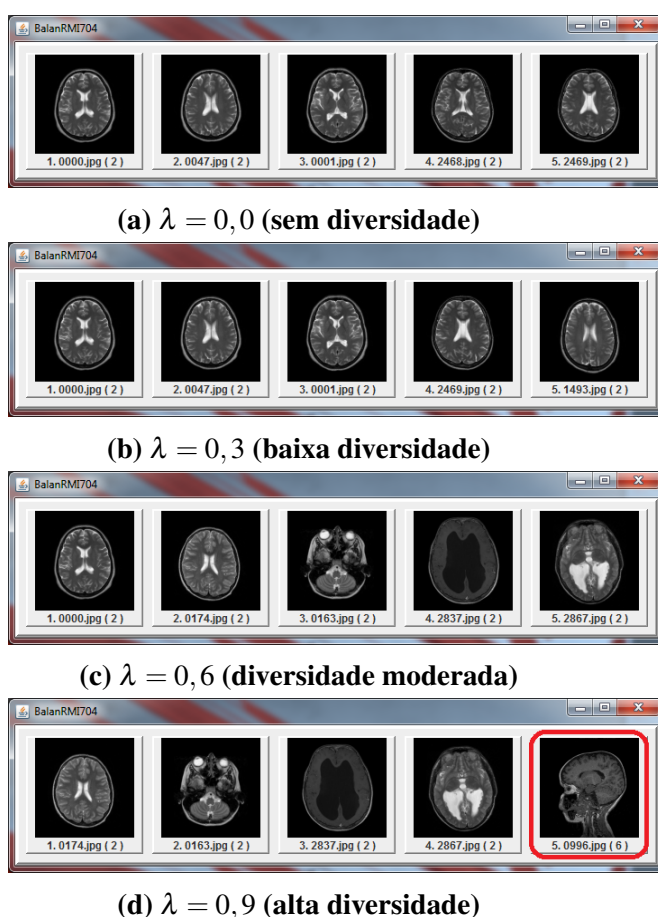


Figura 5.2: Resultados para o conjunto de imagens Balan704 alterando-se a compensação entre 0 e 0,9. Imagens do conjunto de resposta pertencentes à categoria Cabeça Axial, com exceção da assinalada em destaque (Cabeça Sagital).

A partir dos resultados, pode-se observar que somente imagens de uma mesma região do corpo humano foram obtidas. Apesar da imagem de uma categoria diferente (Cabeça Sagital) ser encontrada ao considerar um grau elevado de diversidade, a mesma é viável para a análise de anomalias na sessão avaliada do corpo humano. Ter uma imagem de um ângulo/corte diferente para comparação pode, em determinados casos, proporcionar uma melhor compreensão de dado diagnóstico para o contexto em que esse estudo de caso foi realizado.

Se não fosse dada uma restrição para a diversidade através do raio estabelecido pela análise fractal do conjunto de dados, imagens de outras categorias mais diversas poderiam ser retornadas. Dessa forma, imagens de outros grupos, como Abdômen Axial ou Coluna Sagital não

seriam interessantes para avaliar diagnósticos voltados a anomalias na região da cabeça (DIAS; BUENO; RIBEIRO, 2013).

Contudo, apesar do auxílio no diagnóstico de problemas como nesse estudo de caso, o usuário ainda não tem um entendimento claro da distribuição/similaridade entre as diferentes categorias do conjunto de imagens.

Para complementar a compreensão do usuário, as técnicas de MVD apresentadas neste trabalho de mestrado podem ser utilizadas. Logo, através da visualização da projeção final, desconsiderando-se qualquer distorção do espaço projetado inicialmente, pode-se verificar quais imagens encontram-se mais próximas à imagem de consulta, como ilustrado na Figura 5.3.

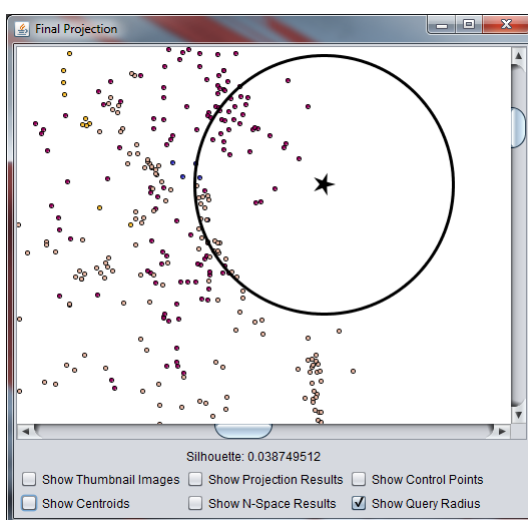
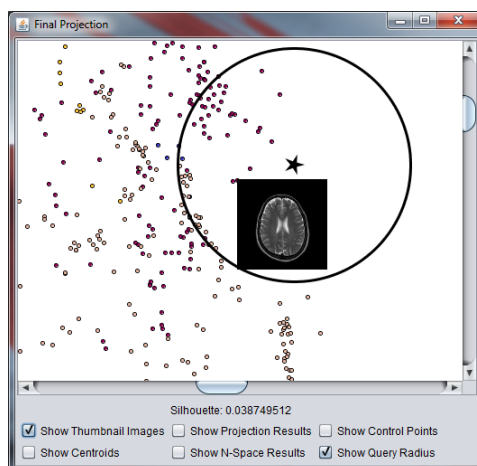


Figura 5.3: Abrangência da consulta pelos vizinhos diversos mais próximos (kNDN) no estudo de caso para o conjunto Balan704.

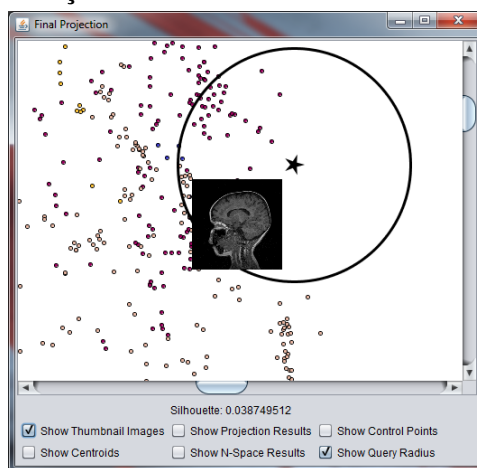
Pela visualização da projeção final, é identificado que existem instâncias que estão relacionadas majoritariamente com duas categorias distintas de imagens. O usuário pode então requisitar a visualização de imagens relativas a instâncias específicas, como demonstrado na Figura 5.4.

Como identificado na Figura 5.4, é verificado que as categorias com maiores números de instâncias estão relacionadas aos subconjuntos de imagens de Cabeça Axial e Cabeça Sagital. É possível então notar que, considerando a projeção do conjunto de dados realizada neste estudo de caso, imagens relativas a tais categorias estão majoritariamente presentes nos subconjuntos avaliados na consulta.

Adicionalmente, é possível requisitar a visualização das centróides relativas a cada categoria na projeção final. Desta forma, pode ser avaliada a proximidade em que cada categoria se



(a) Visualização de instância relativa à categoria Cabeça Axial.



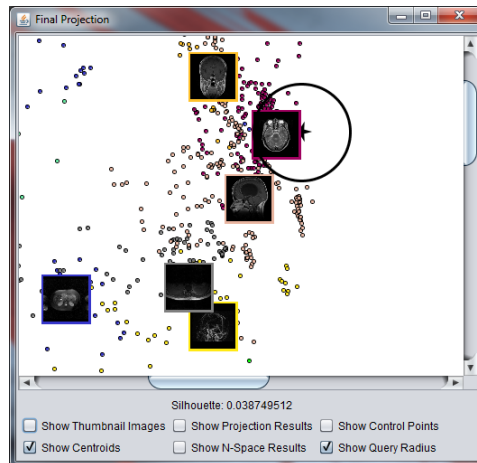
(b) Visualização de instância relativa à categoria Cabeça Sagital.

Figura 5.4: Visualização de imagens relativas a instâncias no estudo de caso: (a) Cabeça Axial e (b) Cabeça Sagital.

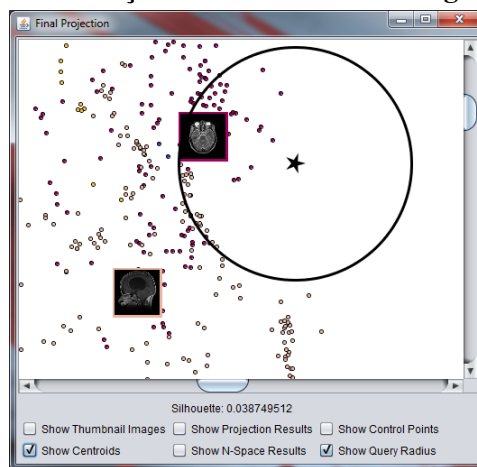
encontra na distribuição. A Figura 5.5 demonstra, respectivamente, a visualização da projeção final para este estudo de caso (Figura 5.5a) e a visualização com maior proximidade da área de consulta e das instâncias relativas às categorias Cabeça Axial e Cabeça Sagital (Figura 5.5b).

Pela visualização da Figura 5.5a, é possível verificar que, caso o raio de seleção de imagens para a consulta (estimado pela análise fractal do conjunto de dados) sofresse um pequeno aumento, imagens de outra categoria (Cabeça Coronal) poderiam ser retornadas. Tal categoria poderia também auxiliar na avaliação de anomalias nessa região do corpo humano. Adicionalmente, é verificado que as demais categorias encontram-se distantes da imagem de consulta, indicando que a consideração de somente imagens mais próximas à consulta, dadas pela restrição do raio estimado, se mostra viável.

Pelos resultados obtidos neste estudo de caso, através das técnicas de MVD utilizadas neste



(a) Visualização das centróides de categorias.



(b) Zoom da visualização de centróides.

Figura 5.5: Visualização de centróides pertencentes a cada categoria do conjunto Balan704.

trabalho de mestrado, é possível verificar que o corte do espaço original e dos conjuntos de imagens a serem analisados mostra-se viável para a avaliação da imagem de consulta e da compreensão da relação entre as categorias de imagens. Tal redução permite que a diversidade seja considerada dentro do limite de similaridade, possibilitando que, majoritariamente, imagens relativas a uma mesma região do corpo sejam retornadas, o que auxilia na avaliação de anomalias neste contexto.

5.2.2 Estudo de Caso 2 (CT_Pulmão_ROIs Dataset)

Ainda avaliando os benefícios dos métodos desenvolvidos para o contexto da área médica, é considerado, neste estudo de caso, o conjunto de imagens CT_Pulmão_ROIs. O mesmo foi obtido por contribuição do Hospital das Clínicas de Ribeirão Preto (HCRP) da Universidade de São Paulo (USP).

O conjunto CT_Pulmão_ROIs apresenta uma coleção de 3.257 imagens que contêm regiões de interesse (Regions of Interest - ROIs) de tomografias digitalizadas de imagens de pulmão. O conjunto é dividido em seis diferentes categorias, onde cinco apresentam padrões anormais distintos e uma composta por imagens com observação de padrões normais. A Tabela 5.2 ilustra as diferentes categorias presentes no conjunto CT_Pulmão_ROIs.







1.Consolidação	
2.Enfisema	
3.Vidro Fosco	
4.Favo de Mel	
5.Espessamento	
6.Normal	

Tabela 5.2: Categorias presentes no conjunto de imagens CT_Pulmão_ROIs.

Diversas doenças pulmonares estão relacionadas a características determinadas pela presença de alterações conhecidas na estrutura pulmonar. Especialistas de domínio chamam estas alterações de achados radiológicos (*radiological findings*) (PONCIANO-SILVA et al., 2009).

Para este conjunto de dados, características visuais relacionadas à textura (Haralick) (HARALICK; SHANMUGAM; DINSTEIN, 1973) foram utilizadas para obter a representação numérica das imagens.

Neste estudo de caso, ainda visando avaliar as contribuições deste trabalho para área médica, demonstra-se a possibilidade de serem analisados relacionamentos entre diferentes diagnósticos (anomalias) de categorias distintas.

Na consulta, foi utilizada uma imagem previamente classificada na categoria Consolidação (Figura 5.6). Logo, é esperado que, com o aumento do fator de diversidade na consulta, imagens com maior dissimilaridade entre si sejam retornadas.

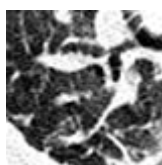
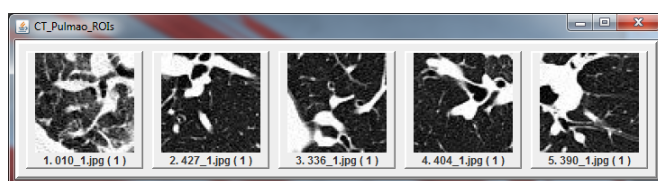


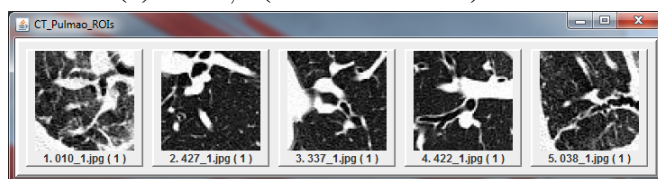
Figura 5.6: Imagem de consulta utilizada para o estudo de caso do conjunto de imagens CT_Pulmão_ROIs.

Os resultados obtidos para este estudo de caso são ilustrados na Figura 5.7. É considerada a

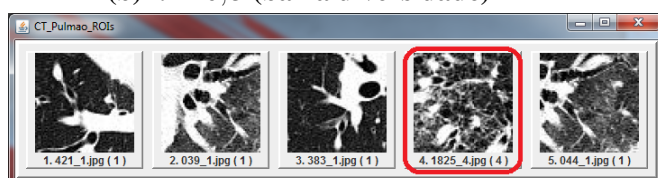
compensação entre diversidade e similaridade para $\lambda = 0$ (Figura 5.7a), $\lambda = 0,3$ (Figura 5.7b), $\lambda = 0,6$ (Figura 5.7c), $\lambda = 0,9$ (Figura 5.7d).



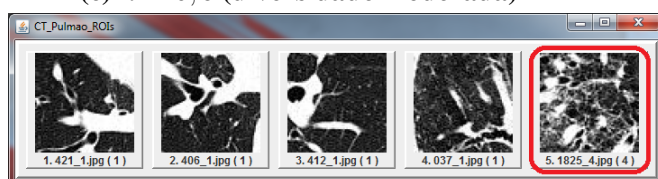
(a) $\lambda = 0,0$ (sem diversidade)



(b) $\lambda = 0,3$ (baixa diversidade)



(c) $\lambda = 0,6$ (diversidade moderada)



(d) $\lambda = 0,9$ (alta diversidade)

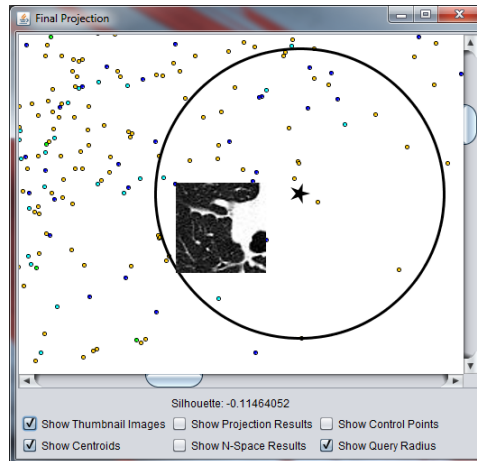
Figura 5.7: Resultados para o conjunto de imagens CT_Pulmão_ROIs alterando-se a compensação entre 0 e 0,9. Imagens no conjunto de resposta pertencentes à categoria Consolidação, com exceção das assinaladas em destaque (Favo de Mel).

Os resultados representados na Figura 5.7 demonstram que, até um baixo fator de diversidade, somente imagens da categoria Consolidação são retornadas. Contudo, ao aumentar a diversidade para um nível moderado ou alto, uma imagem representativa de outra categoria (Favo de Mel) fora encontrada (em destaque).

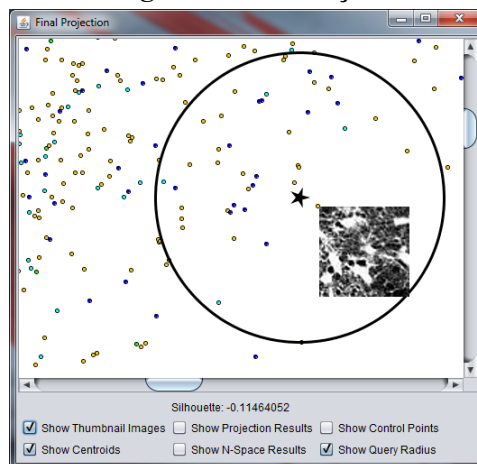
Tal descoberta contendo uma imagem de outra categoria ao aumentar o grau de diversidade pode indicar que a imagem de consulta pode possuir características de ambas anomalias (categorias). Logo, é possível questionar se a imagem de consulta previamente classificada como Consolidação possui também fatores relacionados a outras anomalias. Adicionalmente, devido a esta proximidade entre categorias, pode ser estudada a possibilidade da existência de uma correlação entre as respectivas anomalias (DIAS; BUENO; RIBEIRO, 2013).

A fim de complementar as conclusões do estudo demonstrado, é possível visualizar a projeção das instâncias relativas ao conjunto de dados, levando-se em consideração as técnicas de

MVD apresentadas. Dessa maneira, tais proximidades entre imagens podem ser avaliadas nas projeções ilustradas na Figura 5.8. É observado que, majoritariamente, são encontradas instâncias relativas a imagens da categoria Consolidação (Figura 5.8a) e Favo de Mel (5.8b) na vizinhança da imagem de consulta.



(a) Verificação de proximidade a imagem pertencente à categoria Consolidação.



(b) Verificação de proximidade a imagem pertencente à categoria Favo de Mel.

Figura 5.8: Exploração da vizinhança da consulta no estudo de caso para o conjunto CT_Pulmão_ROIs.

Neste estudo, foi exemplificado que a metodologia permite o questionamento de dado diagnóstico previamente classificado dentre outras possíveis anomalias. Adicionalmente, é viabilizada a inferência e estudo de possíveis correlações entre diferentes anomalias em conjuntos de imagens médicas.

5.2.3 Estudo de Caso 3 (FullFaces Dataset)

Para demonstrar a extensibilidade da metodologia apresentada neste trabalho de mestrado, um outro conjunto de imagens foi utilizado, o qual não possui ligações com a área médica. Desta forma, o conjunto FullFaces foi utilizado para elaborar o estudo de caso onde o objetivo foi avaliar o método proposto no contexto de imagens de faces humanas.

O conjunto FullFaces possui 300 imagens de faces, sendo estas obtidas de 30 indivíduos distintos e em dez diferentes posicionamentos de suas cabeças. As categorias definidas para este conjunto representam cada indivíduo analisado.

A representação numérica das imagens desse conjunto foram obtidas através do levantamento do histograma em escala de cinza. Neste estudo de caso, é mostrado como as técnicas de MVD podem ser utilizadas para obter resultados mais satisfatórios, utilizando o conhecimento do usuário para tal.

Para avaliação e comparação dos resultados, a imagem ilustrada na Figura 5.9 foi utilizada para consulta.



Figura 5.9: Imagem de consulta utilizada para o estudo de caso do conjunto de imagens FullFaces.

Como ponto de partida do experimento, foi verificado quais seriam os resultados obtidos se fosse considerado a abordagem que utiliza o espaço de características original (N-Dimensional). Desta forma, considerando apenas a similaridade na consulta (compensação entre diversidade e similaridade $\lambda = 0$), foram obtidos os resultados mostrados na Figura 5.10.

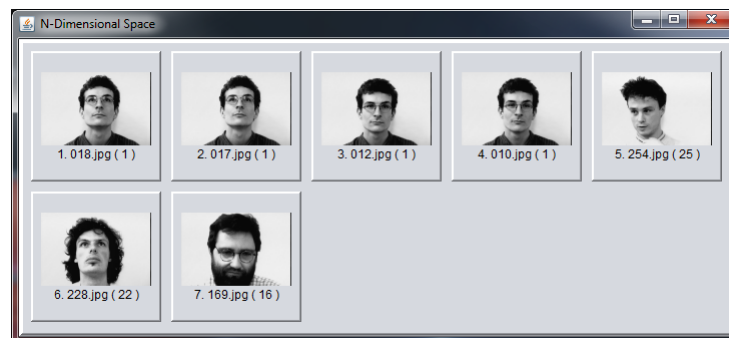


Figura 5.10: Resultados para $\lambda = 0$ considerando o espaço N-Dimensional para o conjunto FullFaces.

Avaliando-se o resultado obtido ao considerar o espaço de características original, pode-se notar que as três primeiras imagens pertencem ao mesmo indivíduo (categoria) da imagem de consulta. Contudo, as demais imagens representam três outras categorias distintas.

Para avaliar e compreender a proximidade entre a imagem de consulta e as demais categorias encontradas no resultado, é dada a primeira projeção (Figura 5.11). O usuário pode então requisitar a visualização das imagens relativas a determinadas instâncias, verificando que, como esperado, imagens de diferentes categorias encontram-se próximas. Tal constatação comprova que características relativas ao histograma em escala de cinza não são suficientes para representar adequadamente o conjunto de imagens considerado neste estudo de caso.

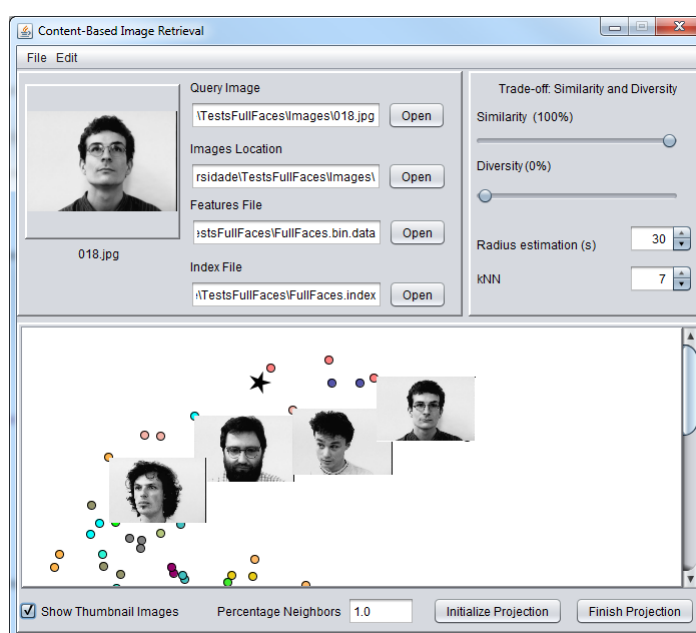


Figura 5.11: Projeção inicial e visualização de imagens do conjunto FullFaces.

Contudo, através da manipulação de instâncias e distorção do espaço de projeção inicial, é possível agrupar imagens e categorias e reposicioná-las em regiões de acordo com o conhecimento e julgamento do usuário. Logo, é possível agrupar as imagens de mesma categoria (reduzir as distâncias intra-grupo) e separar imagens de categorias distintas (aumentar as distâncias inter-grupo). A partir dos resultados encontrados, considerando-se o espaço N-Dimensional (Figura 5.10), as imagens relativas a categoria da imagem de consulta e das demais categorias não desejadas são reposicionadas de acordo com o demonstrado na Figura 5.12.

Depois de interagir com a projeção inicial e distorcer o espaço, a projeção final é gerada. Os resultados considerando o espaço final de projeção são dados na Figura 5.13.

Avaliando-se os resultados da Figura 5.13, é possível notar que todas as imagens obtidas estão condizentes com a mesma categoria da imagem de consulta. Desta forma, considerando-

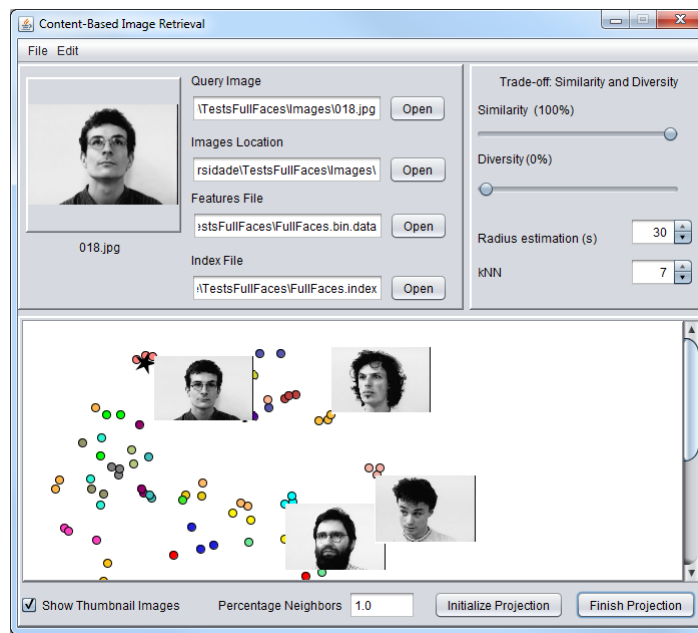


Figura 5.12: Projeção inicial após distorções e manipulações para o conjunto FullFaces.

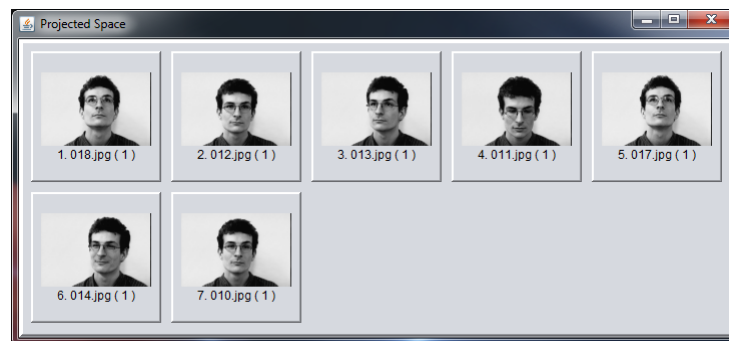


Figura 5.13: Resultados para $\lambda = 0$ considerando o espaço de projeção para o conjunto FullFaces.

se apenas a similaridade na consulta, as técnicas de MVD aplicadas puderam obter resultados mais satisfatórios, como demonstrado nos valores de precisão e revocação dados nas Tabelas 5.3 e 5.4.

A partir dos dados referentes à consulta dada como exemplo levantados nas Tabelas 5.3 e 5.4, pode ser traçado o gráfico comparativo de precisão e revocação para ambas abordagens (Figura 5.14). Através da visualização e comparação dos resultados, nota-se que, pela projeção, os resultados através das técnicas de MVD apresentaram precisão de 100% até a revocação de 70%, enquanto os resultados para o espaço N-Dimensional decaem para 57,14% ao atingir uma revocação de 40%.

Adicionalmente, pode ser verificada a diferença em adicionar níveis mais elevados de diversidade nas abordagens utilizando os espaços N-Dimensional e projetado. A comparação pode ser feita visualizando-se a projeção final, onde o usuário pode evidenciar as instâncias

N-Dimensional	
Revocação (%)	Precisão (%)
10	100
20	100
30	100
40	57,14

Tabela 5.3: Levantamento, em porcentagem, de revocação e precisão para o espaço N-Dimensional do conjunto FullFaces.

Espaço de Projeção	
Revocação (%)	Precisão (%)
10	100
20	100
30	100
40	100
50	100
60	100
70	100

Tabela 5.4: Levantamento, em porcentagem, de revocação e precisão para o espaço de projeção do conjunto FullFaces.

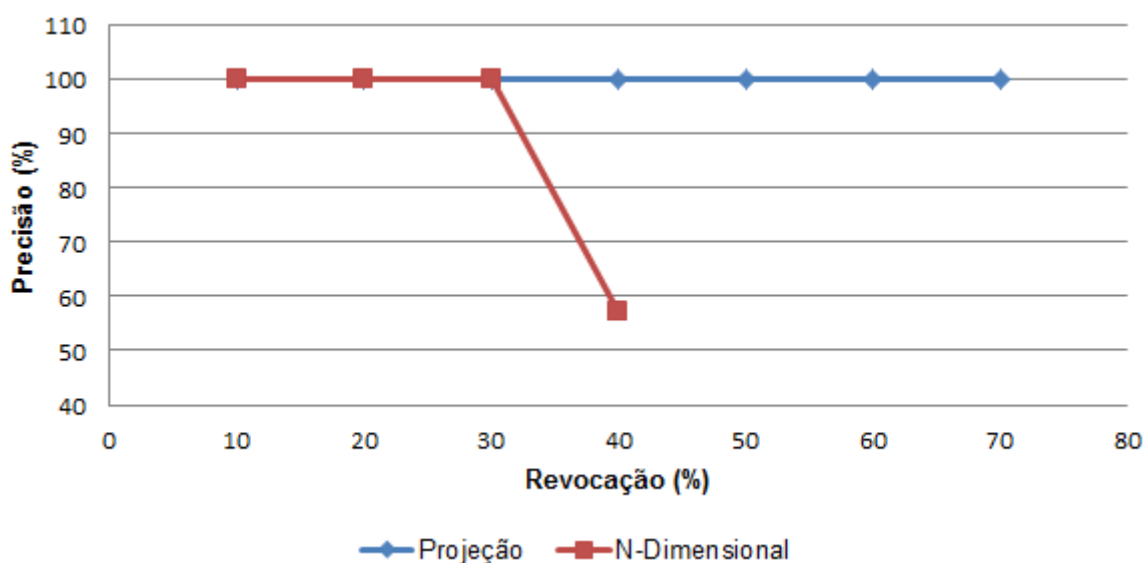


Figura 5.14: Gráfico de Revocação e Precisão para comparação dos resultados obtidos para o conjunto FullFaces no experimento realizado.

resultantes de ambas abordagens. Na Figura 5.15, é dada a visualização e os resultados para a compensação entre diversidade e similaridade $\lambda = 0$ (Figura 5.15a), $\lambda = 0,3$ (Figura 5.15b), $\lambda = 0,6$ (Figura 5.15c) e $\lambda = 0,9$ (Figura 5.15d). Os resultados considerando o espaço N-Dimensional e de projeção encontram-se, respectivamente, à esquerda (evidenciados por uma projeção quadrada) e à direita da visualização (evidenciados por uma reta a partir da projeção

da imagem de consulta). É também evidenciado o corte das imagens analisadas para o resultado através do raio desenhado, possuindo em seu centro a imagem de consulta.

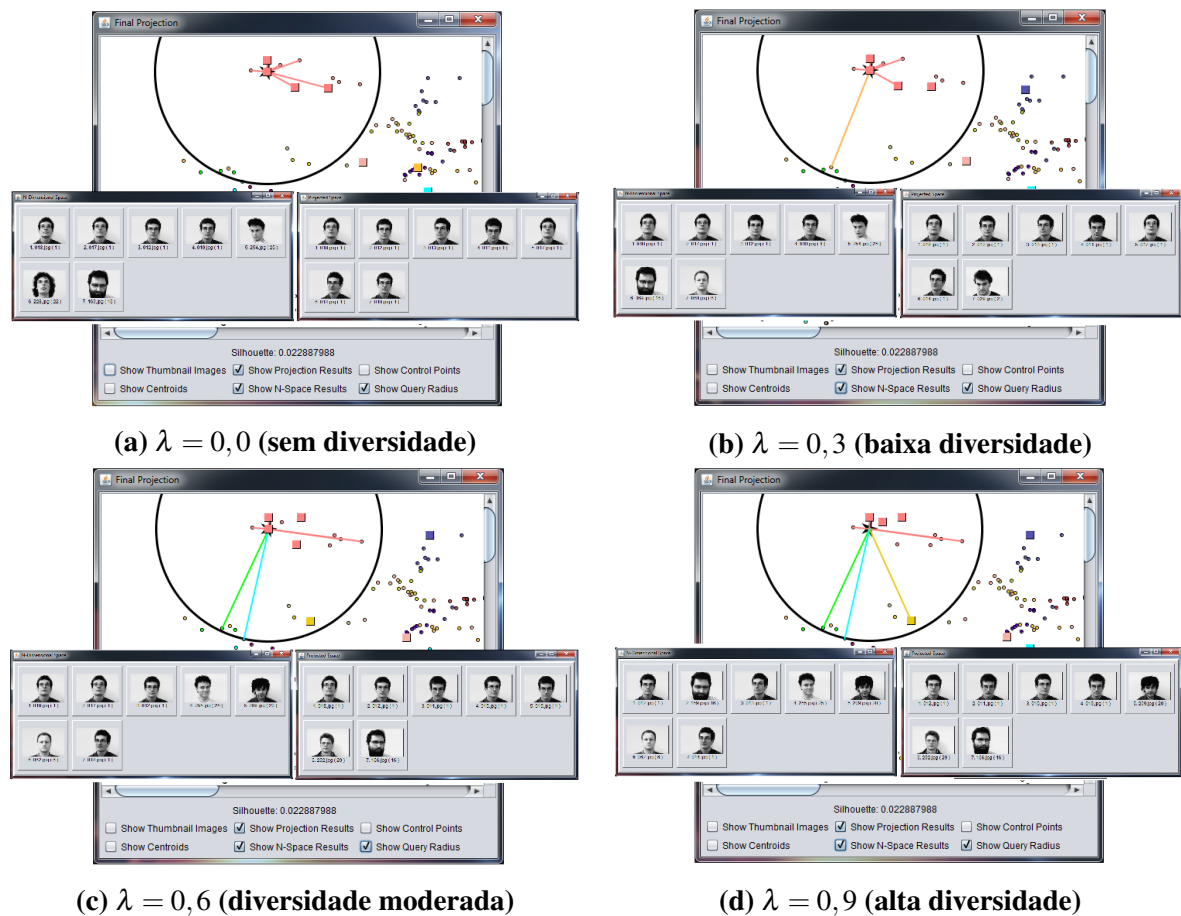


Figura 5.15: Resultados comparativos do espaço N-Dimensional e de projeção para o conjunto de imagens FullFaces. A compensação é alterada no intervalo entre 0 e 0,9. Em cada resultado, são evidenciadas as imagens retornadas para o espaço N-Dimensional (à esquerda) e de projeção (à direita).

Através dos resultados expressados na Figura 5.15 pode-se notar que, além de permitir a melhor compreensão dos resultados, as técnicas de MVD empregadas neste trabalho permitem que os resultados estejam restritos às proximidades da imagem de consulta. Em outras palavras, os resultados considerando o espaço N-Dimensional se encontram mais dispersos e possuem pouca relação entre si, ao ponto em que os resultados da projeção respeitam a vizinhança da imagem de consulta.

5.3 Custo Computacional

A fim de avaliar o custo computacional envolvido na consulta dos vizinhos diversos mais próximos (kNDN), foi levantado o tempo necessário para encontrar o subconjunto resultante

(*k*-similar diversification set). Adicionalmente, foi avaliado o ganho computacional com a utilização do algoritmo concorrente, sendo este possível pela seleção prévia das imagens obtidas pela análise da dimensão fractal do conjunto.

Nos experimentos realizados, os subconjuntos candidatos foram formados considerando as 50 imagens mais próximas à imagem de consulta, com a compensação entre similaridade e diversidade $\lambda = 0,5$. Foi utilizada uma máquina com processador Intel Core Q720 (1,60 GHz) de quatro núcleos (oito virtuais) e 8,00GB de memória RAM para o levantamentos dos dados.

Os resultados obtidos para o conjunto de imagens Balan704 encontram-se nas Tabelas 5.5 (algoritmo sequencial) e 5.6 (algoritmo concorrente). Para cada consulta kNDN, no intervalo de $k = 1$ a $k = 5$, foram realizadas dez execuções com o intuito de obter a média e o desvio padrão do tempo exigido pelas mesmas. É também detalhado o número de combinações distintas avaliadas em cada consulta.

Balan704 Sequencial			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	0,2	0,6324
2-NDNq	1225	7,7	11,3534
3-NDNq	19600	139,8	12,5059
4-NDNq	230300	2163,7	83,6740
5-NDNq	2118760	28981,7	298,3551

Tabela 5.5: Resultados para o conjunto Balan704 considerando o algoritmo sequencial.

Balan704 Concorrente			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	1,6	1,5776
2-NDNq	1225	9,5	18,1613
3-NDNq	19600	68,4	21,9301
4-NDNq	230300	911,8	22,6166
5-NDNq	2118760	14227,1	658,9507

Tabela 5.6: Resultados para o conjunto Balan704 considerando o algoritmo concorrente.

Analogamente aos experimentos realizados para o conjunto Balan704, foram também avaliados os resultados para os conjuntos CT_Pulmão_ROIs (Tabelas 5.7 e 5.8) e FullFaces (Tabelas 5.9 e 5.10).

Variações nos custos entre os conjuntos de imagens avaliados são encontradas devido a diferenças de dimensionalidade de cada cenário. Para o conjunto Balan704, foram consideradas 30 características para a representação de cada imagem no conjunto, para o conjunto CT_Pulmão_ROIs 140 e FullFaces 256. Tais variações implicam em maior complexidade no cálculo da distância/similaridade entre imagens.

CT_Pulmão_ROIs Sequencial			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	0,2	0,6324
2-NDNq	1225	17,2	12,3270
3-NDNq	19600	413,5	22,3171
4-NDNq	230300	7214,2	61,3528
5-NDNq	2118760	97080	1244,1262

Tabela 5.7: Resultados para o conjunto CT_Pulmão_ROIs considerando o algoritmo sequencial.

CT_Pulmão_ROIs Concorrente			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	3,3	2,8693
2-NDNq	1225	11,3	15,0114
3-NDNq	19600	185,7	18,2638
4-NDNq	230300	2802,2	41,7952
5-NDNq	2118760	39245,1	2165,1382

Tabela 5.8: Resultados para o conjunto CT_Pulmão_ROIs considerando o algoritmo concorrente.

FullFaces Sequencial			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	0,3	0,9486
2-NDNq	1225	29,1	16,0447
3-NDNq	19600	720,5	38,6846
4-NDNq	230300	12710,4	151,7345
5-NDNq	2118760	172219,8	1341,8148

Tabela 5.9: Resultados para o conjunto FullFaces considerando o algoritmo sequencial.

FullFaces Concorrente			
Imagens a Retornar	Combinações Distintas	Média (ms)	Desvio Padrão
1-NDNq	50	4	3,7118
2-NDNq	1225	17,7	20,4126
3-NDNq	19600	295,8	23,3037
4-NDNq	230300	4700,2	58,1201
5-NDNq	2118760	63636,7	1200,9126

Tabela 5.10: Resultados para o conjunto FullFaces considerando o algoritmo concorrente.

Avaliando-se os resultados obtidos para os três conjuntos de imagens, é possível perceber que existe um ganho computacional considerável na utilização do algoritmo concorrente ao ser aumentado o número de imagens a serem retornadas. Com um número mais elevado de combinações a serem analisadas, o uso de múltiplos fluxos de execução (*Threads*) permitem que as mesmas sejam processadas paralelamente, reduzindo o tempo necessário para a verificação de todos os candidatos. Um quadro comparativo da relação de ganho na proporção $\frac{\text{Custo do Algoritmo Sequencial}}{\text{Custo do Algoritmo Concorrente}}$ obtido para o conjunto FullFaces é demonstrado na Tabela 5.11.

Comparação do Ganho Computacional		
Imagens a Retornar	Combinações Distintas	Proporção
1-NDNq	50	0,075
2-NDNq	1225	1,6440
3-NDNq	19600	2,4357
4-NDNq	230300	2,7042
5-NDNq	2118760	2,7062

Tabela 5.11: Resultados do ganho computacional entre o algoritmo sequencial e concorrente para o conjunto FullFaces. Proporção dada pelo custo do algoritmo sequencial pelo concorrente.

A partir dos resultados de ganho demonstrados na Tabela 5.11, pode-se perceber que, para a consulta 5-NDNq no conjunto FullFaces, o custo para o algoritmo sequencial é, aproximadamente, 2,70 vezes maior do que o concorrente.

5.4 Considerações Finais

Neste capítulo foram apresentados os experimentos e estudos de caso utilizados para a validação do trabalho desenvolvido para este projeto de mestrado. Foram avaliados diferentes conjuntos de imagens, sendo principalmente voltados para a análise de bases de imagens médicas, mas mostrando a extensibilidade das técnicas para outros contextos. Foram dadas as vantagens obtidas pela utilização das técnicas propostas, sendo estas principalmente relacionadas a melhor compreensão dos resultados obtidos pela consulta por conteúdo de imagens; melhor entendimento do impacto da adição do fator de diversidade na consulta e no conjunto de imagens; melhor compreensão do relacionamento entre diferentes categorias de um conjunto de imagens; e menor custo computacional devido à redução do conjunto de imagens e uso de algoritmo concorrente na adição do fator de diversidade na consulta.

Parte III

Conclusões e Trabalhos Futuros

Capítulo 6

CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho de mestrado foi apresentado uma nova abordagem para viabilizar a adição de diversidade em consultas por conteúdo de imagens. Foram também aplicadas técnicas de Mineração Visual de Dados (MVD) para obter resultados mais próximos a cognição humana e melhorar a compreensão dos mesmos. Na Seção 6.1 são dadas as considerações iniciais deste capítulo. Na Seção 6.2 são detalhadas as principais contribuições alcançadas no trabalho desenvolvido. A Seção 6.3 enumera as possibilidades de trabalhos futuros acerca do que foi alcançado com este estudo e na Seção 6.4 são dadas as considerações finais deste capítulo.

6.1 Considerações Iniciais

Consultas por conteúdo de imagens mostram-se como complemento para consultas textuais. Em determinadas situações, existe maior interesse nas características visuais presentes nas imagens do que simplesmente marcações textuais referentes as mesmas.

A manipulação de informações relativas ao conteúdo visual de imagens, no entanto, é de natureza complexa. Definir como a informação deve ser extraída para a devida representação das informações visuais de imagens, muitas vezes, se distancia do esperado pela cognição humana.

Sistemas de consulta por conteúdo de imagens (*Content-Based Image Retrieval Systems - CBIR Systems*) utilizam a representação numérica de imagens para verificar a similaridade entre as mesmas. A verificação da similaridade é realizada através de uma métrica, como a distância Euclidiana, para medir a proximidade em que estes vetores de características se encontram.

Os sistemas CBIR verificam a similaridade entre os vetores de características através de duas abordagens: pela análise da vizinhança (*k-Nearest Neighbor query*) e por abrangência

(*Range query*). A consulta por vizinhança retorna os k elementos mais próximos à imagem de consulta, enquanto a consulta por abrangência retorna todos os elementos encontrados dentro do raio especificado.

Como descrito anteriormente, encontrar a representação numérica para imagens nem sempre atende às expectativas do usuário. Essa discrepância entre a interpretação visual humana e a representação numérica é denominada na literatura como descontinuidade semântica (*Semantic Gap*).

Um dos artifícios encontrado na literatura para reduzir o problema da descontinuidade semântica é a dado pela inserção de um fator de diversidade na consulta. Logo, além da similaridade verificada entre a imagem de consulta e as demais imagens do conjunto, certa dissimilaridade no subconjunto resultante é também considerada.

Aplicar diversidade em um grande conjunto de dados, no entanto, requer um alto custo computacional. Um grande número de cálculos de distância tem que ser computado para verificar a dissimilaridade entre todas imagens do conjunto, mostrando-se inviável a verificação da diversidade para todo o conjunto de imagens.

Outra possibilidade de artifício para redução da descontinuidade semântica é utilizar o conhecimento do usuário para auxiliar no processo da consulta. Assim, técnicas de Mineração Visual de Dados (MVD) permitem não só que o usuário entenda melhor os resultados, mas também interaja durante todo o processo de consulta por conteúdo de imagens.

Neste projeto de mestrado, foi empregada uma nova abordagem para viabilizar a adição do fator de diversidade em consultas por vizinhança em sistemas CBIR. Um corte inteligente é feito no número de imagens a serem avaliadas nessa etapa, restringindo-se a elementos mais próximos a imagem de consulta. Adicionalmente, são incorporadas técnicas de MVD para permitir que o usuário direcione o processo de consulta de acordo com o seu julgamento, tornando os resultados obtidos mais próximos a cognição humana.

6.2 Principais Contribuições

Neste projeto de mestrado, é apresentada uma nova metodologia para a adição de diversidade em consultas por conteúdo de imagens. Desta forma, é possível encontrar resultados mais condizentes com as necessidades do usuário e reduzir a complexidade computacional na obtenção de resultados. Adicionalmente, foram utilizados recursos de visualização e interação para proporcionar um papel mais ativo do usuário nas etapas da consulta, possibilitando também

uma melhor compreensão dos resultados e relacionamentos entre imagens e categorias.

A redução do espaço de consulta é feita utilizando-se a análise da dimensão fractal do conjunto de imagens para estimar um valor que, seguido de uma consulta por abrangência, se aproxima dos resultados que seriam obtidos através de uma consulta por vizinhança. O corte do espaço proporciona não somente uma restrição de similaridade na consulta, mas também que os subconjuntos a serem avaliados sejam levantados previamente, permitindo a utilização de um algoritmo concorrente. Assim, os mesmos resultados podem ser obtidos com um custo computacional inferior em relação à abordagem sequencial (Seção 5.3).

Técnicas de MVD foram também utilizadas para utilizar o conhecimento e julgamento do usuário a fim de obter resultados mais próximos ao esperado pela cognição humana. A técnica FastMap (FALOUTSOS; LIN, 1995) foi utilizada para proporcionar uma primeira projeção com algumas instâncias, sendo estas denominadas Pontos de Controle. O usuário pode então manipular os Pontos de Controle e reposicioná-los de acordo com seu julgamento, distorcendo o espaço de projeção. A partir das informações da projeção inicial, a técnica LAMP (JOIA et al., 2011) foi utilizada para posicionar as demais instâncias do conjunto de imagens. Desta forma, os resultados são exibidos ao usuário e técnicas de visualização e interação podem ser utilizadas para melhor compreensão dos resultados e da relação entre as instancias/categorias projetadas. Adicionalmente, é possível comparar os resultados com a abordagem considerando o espaço N-Dimensional, além de verificar o subespaço em que as imagens foram consideradas para a consulta ao adicionar a diversidade.

Foi demonstrado na Seção 5.2 que a metodologia empregada neste projeto de mestrado provê bons resultados não somente para a área de atuação médica, mas é também extensível a outros contextos. O corte do espaço original do conjunto de imagens pode proporcionar a seleção de instâncias mais próximas à imagem de consulta, permitindo que a restrição de similaridade considere apenas imagens relacionadas ao diagnóstico de interesse (Seção 5.2.1). Através das técnicas de MVD, é também viabilizado o questionamento de um diagnóstico previamente classificado, podendo este estar próximo a outros não avaliados anteriormente (Seção 5.2.2). Finalmente, em um outro contexto, como o de imagens de faces humanas, é possível utilizar o conhecimento e julgamento do usuário para obter resultados mais satisfatórios (Seção 5.2.3).

Parte dos resultados demonstrados neste trabalho foram apresentados no evento *The 26th IEEE International Symposium on Computer-Based Medical Systems (CBMS 2013)*, como *Full Paper* intitulado *Reducing the Complexity of k-Nearest Diverse Neighbor Queries in Medical Image Datasets through Fractal Analysis*.

6.3 Trabalhos Futuros

Os principais pontos levantados como trabalhos futuros para a pesquisa mencionada neste projeto de mestrado são:

- *Feedback* de especialistas para avaliação da qualidade dos resultados e aplicabilidade deste trabalho para a área médica;
- Avaliar a viabilidade do custo computacional da metodologia desenvolvida para diferentes aplicações da área médica, como, por exemplo, para âmbito operacional ou estudo;
- Avaliar a seleção prévia de imagens para consulta priorizando outros fatores, como diversidade ou novidade;
- Levantamento de outras técnicas de visualização e interação a fim de facilitar a manipulação dos pontos de controle (projeção inicial) caso o número de categorias seja muito elevado.

6.4 Considerações Finais

Neste capítulo foram apresentadas as conclusões acerca do trabalho desenvolvido neste projeto de mestrado. Foram levantadas as principais dificuldades encontradas em sistemas de consulta por conteúdo de imagens (CBIR), as quais dão origem ao problema da descontinuidade semântica. Para solucionar o problema, foi proposta uma maneira inovadora para viabilizar a utilização do fator de diversidade na consulta por vizinhança. Foram também utilizadas técnicas de MVD para proporcionar um papel mais ativo do usuário na descoberta de conhecimento, aproximando os resultados do esperado pela cognição humana e permitindo uma melhor compreensão dos fatores envolvidos nos mesmos. Por fim, são levantadas possibilidades de trabalhos futuros.

REFERÊNCIAS

- ANDREWS, D. F. Plots of high-dimensional data. *Biometrics*, International Biometric Society, v. 28, n. 1, p. pp. 125–136, 1972. ISSN 0006341.
- ARANTES, A.; VIEIRA, M.; TRAINA, A.; TRAINA JR., C. The fractal dimension making similarity queries more efficient. In: *Proceedings of the II ACM SIGKDD Workshop on Fractals, Power Laws and Other Next Generation Data Mining Tools*. [S.l.: s.n.], 2003. p. 12–17.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. *Modern information retrieval*. [S.l.]: ACM press New York., 1999.
- BALAN, A. *Métodos Adaptativos de Segmentação Aplicados à Recuperação de Imagens por Conteúdo*. 2007. Tese (Doutorado) — ICMC-USP, 2007.
- BECKMANN, N.; KRIEGEL, H.-P.; SCHNEIDER, R.; SEEGER, B. The r*-tree: an efficient and robust access method for points and rectangles. In: *Proceedings of the 1990 ACM SIGMOD international conference on Management of data*. New York, NY, USA: ACM, 1990. (SIGMOD '90), p. 322–331. ISBN 0-89791-365-5.
- BIER, E. A.; STONE, M. C.; PIER, K.; BUXTON, W.; DEROSE, T. D. Toolglass and magic lenses: the see-through interface. In: *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1993. (SIGGRAPH '93), p. 73–80. ISBN 0-89791-601-8.
- BRANDOLI, B.; ELER, D.; PAULOVICH, F.; MINGHIM, R.; BATISTA, J. Visual data exploration to feature space definition. In: *Graphics, Patterns and Images (SIBGRAPI), 2010 23rd SIBGRAPI Conference on*. [S.l.: s.n.], 2010. p. 32–39.
- CHANG, S.-K.; CHEN, Z.; LEE, S.-Y. *Recent Advances in Visual Information Systems: 5th International Conference, VISUAL 2002 Hsin Chu, Taiwan, March 11-13, 2002. Proceedings*. [S.l.]: Springer, 2002.
- CIACCIA, P.; PATELLA, M.; ZEZULA, P. M-tree: an efficient access method for similarity search in metric spaces. In: *Proceedings of the Twenty-third International Conference on Very Large Data Bases, Athens, Greece, 26-29 August 1997*. [S.l.: s.n.], 1997. p. 426.
- CLEVELAND, W. S. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, v. 74, n. 368, p. 829–836, 1979.

- CUADROS, A. M.; PAULOVICH, F. V.; MINGHIM, R.; TELLES, G. P. Point placement by phylogenetic trees and its application to visual analysis of document collections. In: *IEEE Symposium on Visual Technology 2007 (VAST 2007)*. [S.l.: s.n.], 2007. p. 99–106.
- 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: *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. [S.l.: s.n.], 2013. p. 101–106.
- EAKINS, J.; GRAHAM, M.; FRANKLIN, T. Content-based image retrieval. *Library and Information Briefings*, v. 85, p. 1–15, 1999.
- ELER, D. M.; NAKAZAKI, M. Y.; PAULOVICH, F. V.; SANTOS, D. P.; ANDERY, G. F.; OLIVEIRA, M. C. F.; NETO, J. B.; MINGHIM, R. Visual analysis of image collections. *The Visual Computer*, Springer, v. 25, n. 10, p. 923–937, 2009.
- FALOUTSOS, C.; LIN, K.-I. Fastmap: a fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 24, n. 2, p. 163–174, maio 1995. ISSN 0163-5808.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. The kdd process for extracting useful knowledge from volumes of data. *Commun. ACM*, ACM, New York, NY, USA, v. 39, n. 11, p. 27–34, nov. 1996. ISSN 0001-0782.
- FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. et al. Knowledge discovery and data mining: Towards a unifying framework. In: *KDD*. [S.l.: s.n.], 1996. v. 96, p. 82–88.
- GUTTMAN, A. R-trees: a dynamic index structure for spatial searching. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 14, n. 2, p. 47–57, jun. 1984. ISSN 0163-5808.
- HAFNER, J.; SAWHNEY, H. S.; EQUITZ, W.; FLICKNER, M.; NIBLACK, W. Efficient color histogram indexing for quadratic form distance functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, IEEE, v. 17, n. 7, p. 729–736, 1995.
- HARALICK, R.; SHANMUGAM, K.; DINSTEIN, I. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3, n. 6, p. 610–621, 1973. ISSN 0018-9472.
- INSELBERG, A.; DIMSDALE, B. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In: *Proceedings of the 1st conference on Visualization*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1990. (VIS '90), p. 361–378. ISBN 0-8186-2083-8.
- JAIN, A.; SARDA, P.; HARITSA, J. R. Providing diversity in k-nearest neighbor query results. In: *Advances in Knowledge Discovery and Data Mining*. [S.l.]: Springer, 2004. p. 404–413.
- JOHNSON, B.; SHNEIDERMAN, B. Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In: *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on*. [S.l.: s.n.], 1991. p. 284–291.
- JOIA, P.; PAULOVICH, F. V.; COIMBRA, D.; CUMINATO, J. A.; NONATO, L. G. Local affine multidimensional projection. *Visualization and Computer Graphics, IEEE Transactions on*, IEEE, v. 17, n. 12, p. 2563–2571, 2011.

- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. *International journal of computer vision*, Springer, v. 1, n. 4, p. 321–331, 1988.
- KATAYAMA, N.; SATOH, S. The sr-tree: an index structure for high-dimensional nearest neighbor queries. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 26, n. 2, p. 369–380, jun. 1997. ISSN 0163-5808.
- KAUFMAN, L.; ROUSSEEUW, P. J. *Finding groups in data: an introduction to cluster analysis*. [S.l.]: Wiley, 2009.
- KEIM, D. A. et al. Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics*, v. 8, n. 1, p. 1–8, 2002.
- KHOTANZAD, A.; HONG, Y. H. Invariant image recognition by zernike moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, IEEE, v. 12, n. 5, p. 489–497, 1990.
- KNUTH, D. *The art of computer programming*. [S.l.]: addison-Wesley, 2006.
- KO, B.; LEE, H.-S.; BYUN, H. Image retrieval using flexible image subblocks. In: *Proceedings of the 2000 ACM symposium on Applied computing - Volume 2*. New York, NY, USA: ACM, 2000. (SAC '00), p. 574–578. ISBN 1-58113-240-9.
- KUMAR, A.; PANG, G. K. Defect detection in textured materials using gabor filters. *Industry Applications, IEEE Transactions on*, IEEE, v. 38, n. 2, p. 425–440, 2002.
- LAMPING, J.; RAO, R.; PIROLI, P. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995. (CHI '95), p. 401–408. ISBN 0-201-84705-1.
- LEBLANC, J.; WARD, M. O.; WITTELS, N. Exploring n-dimensional databases. In: *Proceedings of the 1st conference on Visualization '90*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1990. (VIS '90), p. 230–237. ISBN 0-8186-2083-8.
- LIU, Y.; ZHANG, D.; LU, G.; MA, W.-Y. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, v. 40, n. 1, p. 262 – 282, 2007. ISSN 0031-3203.
- MANJUNATH, B. S.; MA, W.-Y. Texture features for browsing and retrieval of image data. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, IEEE, v. 18, n. 8, p. 837–842, 1996.
- MÜLLER, W.; SCHUMANN, H. Visual data mining. *NORSIGD Info*, Citeseer, v. 2, p. 2002, 2002.
- NGUYEN, Q. V.; HUANG, M. L. A space-optimized tree visualization. In: IEEE. *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*. [S.l.], 2002. p. 85–92.
- PASS, G.; ZABIH, R.; MILLER, J. Comparing images using color coherence vectors. In: *Proceedings of the fourth ACM international conference on Multimedia*. New York, NY, USA: ACM, 1996. (MULTIMEDIA '96), p. 65–73. ISBN 0-89791-871-1.

- PAULOVICH, F. V.; ELER, D. M.; POCO, J.; BOTHA, C. P.; MINGHIM, R.; NONATO, L. G. Piece wise laplacian-based projection for interactive data exploration and organization. In: WILEY ONLINE LIBRARY. *Computer Graphics Forum*. [S.l.], 2011. v. 30, n. 3, p. 1091–1100.
- PAULOVICH, F. V.; NONATO, L. G.; MINGHIM, R.; LEVKOWITZ, H. Least square projection: A fast high-precision multidimensional projection technique and its application to document mapping. *Visualization and Computer Graphics, IEEE Transactions on*, IEEE, v. 14, n. 3, p. 564–575, 2008.
- PONCIANO-SILVA, M.; TRAINA, A.; AZEVEDO-MARQUES, P.; FELIPE, J.; TRAINA JR., C. Including the perceptual parameter to tune the retrieval ability of pulmonary cbir systems. In: *Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on*. [S.l.: s.n.], 2009. p. 1–8. ISSN 1063-7125.
- RAZENTE, H. L. *Adequando consultas por similaridade para reduzir a descontinuidade semântica na recuperação de imagens por conteúdo*. Tese (Doutorado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2009.
- ROBERTSON, G. G.; MACKINLAY, J. D.; CARD, S. K. Cone trees: animated 3d visualizations of hierarchical information. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 1991. (CHI '91), p. 189–194. ISBN 0-89791-383-3.
- SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution, SMOE*, v. 4, n. 4, p. 406–425, 1987.
- SAMMON, J. A nonlinear mapping for data structure analysis. *Computers, IEEE Transactions on*, C-18, n. 5, p. 401–409, 1969. ISSN 0018-9340.
- TRAINA JR., C.; TRAINA, A.; FILHO, R. S.; FALOUTSOS, C. How to improve the pruning ability of dynamic metric access methods. In: *Proceedings of the eleventh international conference on Information and knowledge management*. New York, NY, USA: ACM, 2002. (CIKM '02), p. 219–226. ISBN 1-58113-492-4.
- TRAINA JR., C.; TRAINA, A.; SEEGER, B.; FALOUTSOS, C. Slim-trees: High performance metric trees minimizing overlap between nodes. In: ZANIOLO, C.; LOCKEMANN, P.; SCHOLL, M.; GRUST, T. (Ed.). *Advances in Database Technology EDBT 2000*. [S.l.]: Springer Berlin Heidelberg, 2000, (Lecture Notes in Computer Science, v. 1777). p. 51–65. ISBN 978-3-540-67227-2.
- VIEIRA, M.; RAZENTE, H.; BARIONI, M.; HADJIELEFTHERIOU, M.; SRIVASTAVA, D.; TRAINA JR., C.; TSOTRAS, V. On query result diversification. In: *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*. [S.l.: s.n.], 2011. p. 1163–1174. ISSN 1063-6382.
- VIEIRA, M. R.; RAZENTE, H. L.; BARIONI, M. C.; HADJIELEFTHERIOU, M.; SRIVASTAVA, D.; TRAINA JR., C.; TSOTRAS, V. J. Divdb: A system for diversifying query results. *Proceedings of the VLDB Endowment*, v. 4, n. 12, p. 4, 2011.

WARD, M. O. Xmdvtool: integrating multiple methods for visualizing multivariate data. In: *Proceedings of the conference on Visualization '94*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1994. (VIS '94), p. 326–333. ISBN 0-7803-2521-4.

WEISS, M. *Data Structures and Algorithm Analysis in C*. [S.l.]: Addison-Wesley Publishing Company, 1986.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, v. 2, n. 13, p. 37 – 52, 1987. ISSN 0169-7439. Proceedings of the Multivariate Statistical Workshop for Geologists and Geochemists.

ZHOU, X.; PU, P. *Visual and multimedia information management*. [S.l.]: Springer, 2002.

GLOSSÁRIO

CBIR – *Content-Based Image Retrieval / Recuperação de Imagens por Conteúdo*

KDD – *Knowledge Discovery in Databases / Processo de Descoberta de Conhecimento em Bases de Dados*

LAMP – *Local Affine Multidimensional Projection*

MAE – *Método de Acesso Espacial*

MAM – *Método de Acesso Métrico*

MD – *Mineração de Dados*

MVD – *Mineração Visual de Dados*

NJ – *Neighbor Joining*

ROIs – *Regions of Interest / Regiões de Interesse*

RQ – *Range Query / Consulta pro Abrangência*

Vik – *Visual-interactive k-NDNq Method*

kNDN – *k-Nearest Diverse Neighbor / K Vizinhos Diversos Mais Próximos*

kNN – *k-Nearest Neighbors / k Vizinhos Mais Próximos*