

Técnicas de Visualização para Sistemas de Recuperação de Imagens por Conteúdo

Este exemplar corresponde à redação final da
Dissertação devidamente corrigida e defendida
por Sheila Maricela Pinto Cáceres e aprovada
pela Banca Examinadora.

Campinas, 15 de Outubro de 2010.



Prof. Dr. Ricardo da Silva Torres
Instituto de Computação - Unicamp
(Orientador)

Dissertação apresentada ao Instituto de Com-
putação, UNICAMP, como requisito parcial
para a obtenção do título de Mestre em Ciência
da Computação.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Bibliotecária: Silvania Renata de Jesus Ribeiro Cirilo – CRB8 / 6592

Cáceres, Sheila Maricela Pinto

C113t Técnicas de visualização para sistemas de recuperação de imagens por conteúdo/Sheila Maricela Pinto Cáceres-- Campinas, [S.P. : s.n.], 2010.

Orientador : Ricardo da Silva Torres

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação.

1.Visualização - Técnicas. 2.Imagens - Recuperação.
3.Recuperação da informação. 4.Visualização de informação. I. Torres, Ricardo da Silva. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Título em inglês: Visualization techniques for content-based image retrieval

Palavras-chave em inglês (Keywords): 1.Visualization – Techniques. 2.Image – Retrieval. 3. Information retrieval. 4.Information visualization.

Área de concentração: Ciência da Computação

Titulação: Mestre em Ciência da Computação

Banca examinadora: Prof. Dr. Celmar Guimarães da Silva (FT– UNICAMP)
Profª. Dra. Maria Cecília Calani Baranaukas (IC - UNICAMP)
Profª. Dra. Vânia Paula de Almeida Neris (UFSCar)
Profª. Dra. Anamaria Gomide (IC – UNICAMP)

Data da defesa: 15/10/2010

Programa de Pós-Graduação: Mestrado em Ciência da Computação

TERMO DE APROVAÇÃO

Dissertação Defendida e Aprovada em 15 de Outubro de 2010, pela Banca examinadora composta pelos Professores Doutores:

Celmar Guimarães da Silva

Prof. Dr. Celmar Guimarães da Silva
Faculdade de Tecnologia / UNICAMP

Maria Cecília Calani Baranauskas

Prof^a. Dr^a. Maria Cecília Calani Baranauskas
IC / UNICAMP

Ricardo Torres

Prof. Dr. Ricardo da Silva Torres
IC / UNICAMP

Técnicas de Visualização para Sistemas de Recuperação de Imagens por Conteúdo

Sheila Maricela Pinto Cáceres¹

Outubro de 2010

Banca Examinadora:

- Prof. Dr. Ricardo da Silva Torres
Instituto de Computação - Unicamp (Orientador)
- Prof. Dr. Celmar Guimarães da Silva
Faculdade de Tecnologia - Unicamp
- Prof^a. Dr^a. Maria Cecília Calani Baranauskas
Instituto de Computação - Unicamp
- Prof^a. Dr^a. Vânia Paula de Almeida Neris
Universidade Federal de São Carlos (UFSCar)
- Prof^a. Dr^a. Anamaria Gomide
Instituto de Computação - Unicamp

¹Suporte financeiro de: Bolsa do CNPq (processo 135526/2008-6) 2008-2009, Bolsa da FAPESP (processo 2009/04732-0) 2009-2010.

Resumo

Um sistema de Recuperação de Imagens por Conteúdo (CBIR) oferece mecanismos necessários para busca e recuperação de imagens baseando-se em propriedades visuais como cor, textura, forma, etc. Em um processo de busca de imagens, a apresentação de resultados é um componente essencial, na medida em que a obtenção desses resultados é o motivo da existência do sistema. Conseqüentemente, o uso de técnicas de visualização apropriadas pode determinar o sucesso ou o fracasso de um sistema CBIR. Técnicas de visualização são valiosas ferramentas na exploração de grandes quantidades de dados, como coleções de imagens. Contudo, técnicas para visualizar imagens retornadas por sistemas CBIR têm sido pobremente exploradas.

Este trabalho apresenta um estudo comparativo e avaliação de várias técnicas de visualização para sistemas CBIR. Como resultado desse estudo, propõe-se um conjunto de técnicas originais que tentam suprir algumas das limitações identificadas em métodos da literatura. Dentre as características das técnicas propostas, destacam-se o enfoque baseado no centro e o uso de técnicas de agrupamento de dados para representar a similaridade intrínseca entre as imagens retornadas. Resultados experimentais mostram que os métodos propostos superam outras estratégias de visualização, considerando-se diversos critérios, como adequação para mostrar resultados em sistemas CBIR, quantidade de informação oferecida, satisfação de usuário, etc.

As principais contribuições deste trabalho são: (i) estudo comparativo e análise de sete técnicas de visualização, quatro delas existentes na literatura e três técnicas novas propostas; (ii) avaliação de duas técnicas da literatura nunca antes avaliadas: anéis concêntricos e espiral; (iii) especificação e implementação de três novas técnicas de visualização baseadas em agrupamento; (iv) especificação e implementação de um framework para desenvolvimento de novas estruturas visuais para sistemas CBIR no qual foram implementadas as técnicas de visualização estudadas.

Abstract

A Content-Based Image Retrieval (CBIR) system offers mechanisms needed to search and retrieve images based on visual properties such as color, texture, shape, etc. In an image search process, the presentation of results is an essential component as the retrieval of relevant images is the reason of the system existence. Consequently, the use of appropriate visualization techniques may determine the success of a CBIR system. Visualization techniques are valuable tools for the exploration of a great quantity of data, such as images collections. However, techniques for visualizing images in CBIR systems have been poorly explored.

This work presents a comparative study of several visualization techniques for CBIR systems. As a result of this study, several original techniques were proposed trying to fulfill some of the absent characteristics in existing methods, such as the central-based focus and the use of clustering approaches to represent the intrinsic similarity between retrieved images. Experimental results show that the proposed methods overcome other visualization strategies by considering several criteria such as adaptation to show CBIR results, information load, user satisfaction, etc.

The main contributions of this work are: (i) comparative study and analysis of seven visualization techniques, four of them from the literature and three new ones; (ii) validation of two techniques never evaluated before: concentric rings and spiral; (iii) specification and implementation of three new techniques of visualization based on clustering; (iv) specification and implementation of a framework for developing new visual structures for content-based image retrieval systems. The studied techniques were implemented by using this framework.

Agradecimentos

Agradeço a Deus por guiar meus passos a cada etapa da minha vida, pois sem Ele, nada disto teria sido possível.

Ao meu orientador, Professor Dr. Ricardo da Silva Torres, por me guiar durante o desenvolvimento desta pesquisa e pela paciência com as minhas falhas e dúvidas.

Agradeço à professora Cecília Baranauskas pela ajuda na especificação e realização dos testes, assim como a Vânia e Roberto pela valiosa contribuição com conselhos e dicas.

Agradeço aos usuários que participaram voluntariamente dos testes e sem os quais não teria sido possível validar as técnicas estudadas.

Ao Roberto Pinho pela ajuda com o método por ele proposto.

Ao conjunto de pessoas dos laboratórios LIS e RECOD, pelo convívio amigo e pela disponibilidade em me ajudar.

Ao Instituto de Computação e os funcionários que nele trabalham, pela boa disposição para me atender sempre com muita rapidez e eficiência.

Ao CNPq (processo 135526/2008-6) e à FAPESP (processo 2009/04732-0) pelo apoio financeiro.

Agradeço aos meus pais Mario Pinto Díaz e Melina Cáceres Paredes pelo constante exemplo e por me cuidar e acompanhar no decorrer da vida. Agradeço às minhas irmãs Anahí, Janett e Flor pelo apoio incondicional dados quando precisei, mesmo à distância. Estendo este agradecimento a todos os meus familiares, em especial à minha tia Pastora cujas orações e pensamentos me acompanharam.

Agradeço aos meus grandes amigos do mestrado: Aline, André, Lehilton, Thiago F. e Thiago Q. por compartilhar comigo inúmeros momentos gratos, permitindo-me descobrir dia-a-dia a cultura brasileira. Agradeço especialmente ao Fabinho Lofredo pela ajuda incondicional e companheirismo brindado. Agradeço também ao Tom pela ajuda com a língua.

Não posso esquecer das pessoas que me impulsionaram e me animaram a vir ao Brasil para cumprir este desafio. Agradeço em especial ao meu professor Alfredo Paz Valderrama por me animar a fazer sempre aquilo que desejasse, me inspirando a concretizar este trabalho. Agradeço também ao professor Ernesto Cuadros pelo incentivo a realizar atividades de pesquisa no Brasil.

Finalmente agradeço a este lindo país, o Brasil, por me receber de braços abertos. Aqui aprendi, conheci, cresci e amadureci muito em diversas facetas da minha vida e ainda continuo neste crescimento pessoal. Estendo este agradecimento às muitas pessoas anônimas que me receberam ou que de alguma maneira me ajudaram desde a minha chegada ao Brasil.

Sumário

Resumo	vii
Abstract	ix
Agradecimentos	xi
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	2
1.3 Contribuições	3
1.4 Estrutura da Dissertação	4
2 Trabalhos e Conceitos Relacionados	5
2.1 Considerações iniciais	5
2.2 Recuperação de imagens por conteúdo	5
2.2.1 Arquitetura típica de um sistema CBIR	6
2.2.2 Sistemas CBIR	7
2.3 Visualização de informação	7
2.3.1 Processo de transformação de dados em representações visuais	9
2.3.2 Processo de transformação de imagens em representações visuais	11
2.4 Agrupamento de dados	12
2.4.1 Abordagem aglomerativa	12
2.4.2 Abordagem divisiva ou particional	14
2.4.3 Abordagem DAHC	15
2.5 Técnicas para a visualização de resultados de sistemas CBIR	16
2.5.1 Técnicas de exibição simples	16
2.5.2 Técnicas com enfoque central	18
2.5.3 Técnicas baseadas em agrupamento	21
2.5.4 Outras Técnicas	25

2.5.5	Resumo de sistemas CBIR segundo a técnica de visualização usada	26
2.6	Avaliação na área de Interação Humano-Computador	27
2.7	Considerações finais	30
3	Técnicas de Visualização Propostas	31
3.1	Clustering Set	33
3.1.1	<i>Design</i> do espaço	36
3.1.2	Uso de agrupamento	39
3.1.3	Agrupando por hierarquia	42
3.2	Clustering Set - Espiral	44
3.3	Clustering Set - Anéis Concêntricos	45
3.4	<i>Framework</i> para técnicas de visualização	46
3.5	Considerações finais	48
4	Avaliação	49
4.1	Considerações iniciais	49
4.2	Primeiro experimento	49
4.2.1	Projeto do experimento	50
4.2.2	Desenvolvimento dos testes	54
4.2.3	Resultados	57
4.2.4	Análise	69
4.3	Segundo experimento	73
4.3.1	Projeto do experimento	77
4.3.2	Desenvolvimento dos testes	79
4.3.3	Resultados	81
4.3.4	Análise	87
4.4	Terceiro experimento	90
4.4.1	Projeto do experimento	95
4.4.2	Desenvolvimento dos Testes	95
4.4.3	Resultados	96
4.4.4	Análise	103
4.5	Considerações finais	105
5	Conclusões	107
5.1	Conclusões	107
5.2	Contribuições	110
5.3	Trabalhos futuros	110

Lista de Tabelas

2.1	Quadro de técnicas de agrupamento sobre conjuntos de imagens	24
2.2	Quadro de técnicas de visualização usando agrupamento	25
2.3	Quadro de técnicas de visualização dos sistemas CBIR da literatura	27
3.1	Quadro de abordagens de visualização	32
4.1	Quadro de usuários participantes.	52
4.2	Quadro de resultados da turma de graduação MC750	63
4.3	Quadro de resultados da turma de posgraduação MO825B	64
4.4	Quadro de resultados da turma de posgraduação MO812R	65
4.5	Quadro de entendimento do segundo experimento	85
4.6	Quadro de entendimento do segundo experimento	100

Lista de Figuras

2.1	Arquitetura de um sistema de recuperação de imagens por conteúdo [1]. . . .	6
2.2	Marcha da força armada de Napoleão.	8
2.3	Modelo do processo de visualização de informação [2].	9
2.4	Interface do sistema FilmFinder [3].	10
2.5	Interface do sistema Google Earth [4].	11
2.6	Adaptação do modelo do processo de visualização de informação para sistemas CBIR [5].	12
2.7	Abordagens Aglomerativa e Divisiva	14
2.8	Interface do sistema QBIC [6].	17
2.9	Interface do sistema Amore [7].	17
2.10	Estrutura visual Grid.	18
2.11	Interface do sistema PRISM [8].	18
2.12	Estrutura visual Anéis Concêntricos.	19
2.13	Estrutura visual Espiral.	20
2.14	Estrutura visual Ring [9].	21
2.15	Interface proposta em [10].	22
2.16	Visualização baseada em clusters proposta em [11].	22
2.17	Estrutura visual Incremental Board [12].	26
3.1	Estrutura Clustering Set	35
3.2	Investigações de local	36
3.3	Exemplo de uma estrutura circular com $m = 15$	37
3.4	Exemplo de uma estrutura composta com $m = 6$ no primeiro nível e $m = 20$ no segundo nível.	37
3.5	Exemplo de uma estrutura composta usando Anéis concêntricos.	38
3.6	Exemplo de uma estrutura composta usando Espirais.	38
3.7	Estrutura para $k = 6$	41
3.8	Estrutura para $k = 60$	41
3.9	Estrutura para $k = 500$	42

3.10	Agrupamento hierárquico.	43
3.11	Modelo resumido do processo CBIR com uso de abordagens de agrupamento.	44
3.12	Estrutura visual Clustering Set - Espiral.	45
3.13	Estrutura visual Clustering Set - Anéis Concêntricos.	46
3.14	Espiral de Anéis Concêntricos.	47
3.15	Espiral de Espirais.	47
3.16	Estrutura visual Espiral de Circles.	48
4.1	Experimento 1 - Técnica proposta Clustering Set usada no experimento.	55
4.2	Experimento 1 - Desenvolvimento do teste com usuários.	56
4.3	Experimento 1 - Teste com usuários - Grid.	56
4.4	Experimento 1 - Teste com usuários - Anéis Concêntricos.	56
4.5	Experimento 1 - Teste com usuários - Espiral.	56
4.6	Experimento 1 - Teste com usuários - Incremental Board.	56
4.7	Experimento 1 - Teste com usuários - Clustering Set.	56
4.8	Experimento 1 - Perfil dos usuários da disciplina MC750	57
4.9	Experimento 1 - Perfil dos usuários da disciplina MO825B	58
4.10	Experimento 1 - Perfil dos usuários da disciplina MO812R	58
4.11	Experimento 1 - Familiaridade com computadores	59
4.12	Experimento 1 - Frequência de uso do computadores	59
4.13	Experimento 1 - Familiaridade com máquinas de buscas Web.	59
4.14	Experimento 1 - Frequência com que usa as máquinas de buscas.	59
4.15	Experimento 1 - Familiaridade com sistemas de busca de imagens.	60
4.16	Experimento 1 - Frequência com que usa sistemas de busca de imagens.	60
4.17	Experimento 1 - Familiaridade com sistemas de busca de imagens por conteúdo.	60
4.18	Experimento 1 - Frequência com que usa sistemas de busca de imagens por conteúdo.	60
4.19	Experimento 1 - Perfil de usuário	61
4.20	Experimento 1 - Avaliação de MC750	63
4.21	Experimento 1 - Avaliação de MO825.	64
4.22	Experimento 1 - Avaliação de MO812R	65
4.23	Experimento 1 - Satisfação com a disposição de imagens.	66
4.24	Experimento 1 - Grau de entendimento da distribuição das imagens.	67
4.25	Experimento 1 - Facilidade de se encontrar a imagem de consulta.	67
4.26	Experimento 1 - Facilidade de se encontrar as imagens mais relevantes.	68
4.27	Experimento 1 - Facilidade de identificar as imagens menos relevantes.	68
4.28	Experimento 1 - Avaliação das cinco estruturas visuais por critério.	69
4.29	Experimento 2: Grid.	74

4.30	Experimento 2: Grid Modificado.	74
4.31	Experimento 2: Anéis Concêntricos.	75
4.32	Experimento 2: Espiral.	75
4.33	Experimento 2: Incremental Board.	76
4.34	Experimento 2: Clustering Set.	76
4.35	Experimento 2: Clustering Set - Anéis Concêntricos.	77
4.36	Experimento 2: Clustering Set - Espiral.	77
4.37	Experimento 2 - Desenvolvimento do teste com usuários.	80
4.38	Experimento 2 - Desenvolvimento do teste com usuários.	80
4.39	Experimento 2 - Gênero.	82
4.40	Experimento 2 - Escolaridade.	82
4.41	Experimento 2 - Familiaridade com computadores.	82
4.42	Experimento 2 - Frequência de uso de computadores.	82
4.43	Experimento 2 - Familiaridade com máquinas de buscas Web.	83
4.44	Experimento 2 - Frequência com que usa as máquinas de buscas.	83
4.45	Experimento 2 - Familiaridade com sistemas de busca de imagens.	83
4.46	Experimento 2 - Frequência com que usa sistemas de busca de imagens.	83
4.47	Experimento 2 - Familiaridade com sistemas de busca de imagens por conteúdo.	84
4.48	Experimento 2 - Frequência com que usa sistemas de busca de imagens por conteúdo.	84
4.49	Experimento 2 - Entendimento.	85
4.50	Experimento 2 - Grau de satisfação.	86
4.51	Experimento 2 - Quantidade de informação.	86
4.52	Experimento 2 - Ordenação das estruturas baseadas na adequação para visualização em sistemas CBIR.	87
4.53	Ordem das imagens.	88
4.54	Outra ordem das imagens.	88
4.55	Experimento 3: Grid.	91
4.56	Experimento 3: Grid Modificado.	91
4.57	Experimento 3: Anéis Concêntricos.	92
4.58	Experimento 3: Espiral.	92
4.59	Experimento 3: Incremental Board.	93
4.60	Experimento 3: Clustering Set.	93
4.61	Experimento 3: Clustering Set - Anéis Concêntricos.	94
4.62	Experimento 3: Clustering Set - Espiral.	94
4.63	Experimento 3 - Gênero	97
4.64	Experimento 3 - Escolaridade	97
4.65	Experimento 3 - Principal finalidade com que usa o computador	97

4.66	Experimento 3 - Familiaridade com computadores.	98
4.67	Experimento 3 - Frequência de uso de computadores.	98
4.68	Experimento 3 - Familiaridade com máquinas de buscas Web.	98
4.69	Experimento 3 - Frequência com que usa as máquinas de buscas	98
4.70	Experimento 3 - Familiaridade com sistemas de busca de imagens	99
4.71	Experimento 3 - Frequência com que usa sistemas de busca de imagens . . .	99
4.72	Experimento 3 - Familiaridade com sistemas de busca de imagens por conteúdo	99
4.73	Experimento 3 - Frequência com que usa sistemas de busca de imagens por conteúdo	99
4.74	Experimento 3 - Grau de Entendimento.	100
4.75	Experimento 3 - Grau de Satisfação.	101
4.76	Experimento 3 - Quantidade de informação.	102
4.77	Experimento 3 - Ordenação das estruturas baseadas na adequação para visu- alização em sistemas CBIR.	102

Capítulo 1

Introdução

1.1 Contextualização

Dados diversos fatores como o acelerado crescimento da Internet e os avanços nas tecnologias associadas a dispositivos para obtenção e armazenamento de informação visual, o número de imagens em formato digital tem aumentado exponencialmente, gerando grandes coleções de imagens. Estas coleções precisam ser eficientemente organizadas e administradas para que suas imagens possam ser recuperadas e conseqüentemente usadas.

Diversas técnicas de recuperação de imagens têm sido apresentadas na literatura. Uma das primeiras abordagens foi a recuperação a partir de anotações textuais, segundo a qual cada imagem é associada a uma ou mais palavras e a consulta consiste basicamente na busca de imagens que tenham sido anotadas usando-se termos definidos pelo usuário. Atualmente, essa técnica é a mais popular, sendo usada pelos conhecidos motores de busca Google [13], Yahoo [14], Altavista [15], etc. Uma das principais limitações desse método está relacionada à subjetividade da anotação, dado que depende de diversos fatores, como o ponto de vista, o estado de ânimo e a interpretação subjetiva do(s) usuário(s) encarregado(s) de anotar textualmente cada imagem. Além disso, outro problema associado é o custo e o tempo necessário para fazer anotações em uma grande coleção de imagens. Algumas abordagens têm sido propostas para realizar anotações automaticamente [16, 17], porém, ainda apresentam resultados imprecisos.

Na década de 1990, foi apresentada outra abordagem que tem obtido uma importante relevância no contexto da recuperação de imagens devido às novas possibilidades oferecidas em contraste àquelas encontradas em sistemas de busca de imagens baseada em descrições textuais. Esta proposta foi denominada “*Content-Based Image Retrieval*” (Recuperação de imagens por conteúdo) [1, 18–21], doravante referida como CBIR. Os sistemas CBIR tentam recuperar as imagens similares a padrões de consulta definidos pelo usuário (por exemplo, uma

imagem de consulta). As consultas podem também ser representadas por outras imagens, características visuais específicas desejadas, esboços do usuário, etc. Estes sistemas têm sido aplicados com sucesso em distintos domínios, como astronomia, botânica, biodiversidade [22], mineralogia, medicina [23] e bibliotecas digitais [24, 25].

Em um sistema CBIR, a busca é baseada em propriedades visuais das imagens como cor [26, 27], textura [28], forma [29–31], relações espaciais entre objetos, etc. Cada uma destas propriedades é caracterizada por descritores específicos. Um descritor pode ser definido [21] por uma tupla composta por : (i) um algoritmo de extração de características baseado em técnicas de processamento de imagens, que codifica as propriedades visuais da imagem em um vetor de características; e (ii) uma função de similaridade que calcula a similaridade entre duas imagens como uma função de distância entre seus vetores de características correspondentes. No domínio de recuperação de imagens por conteúdo, um descritor é considerado melhor do que outro se sua utilização resulta em um número maior de imagens relevantes retornadas para uma dada consulta.

1.2 Motivação

A grande maioria dos sistemas CBIR tem sido desenvolvidos no ambiente acadêmico. Lamentavelmente, na maior parte desses protótipos, não se teve o devido interesse nem a preocupação suficiente pela apresentação dos resultados na interface. Na maior parte dos casos, somente é permitida uma forma simples de exploração e, algumas vezes, nenhuma forma de manipulação dos dados recuperados é oferecida. Isso representa um problema crítico, considerando que esta etapa refere-se à parte final de um processo de busca e que este pode determinar o sucesso ou fracasso de um sistema CBIR.

A interface de um sistema CBIR deve disponibilizar ao usuário os recursos necessários para suprir as suas necessidades. A partir dessa interface, o usuário pode especificar uma consulta indicando parâmetros (por exemplo, imagem de consulta, número de imagens que devem ser retornadas). Posteriormente, a consulta será processada sobre o universo de imagens e um conjunto organizado (representado por uma tabela de dados) de imagens relevantes será devolvido segundo uma medida de similaridade [32, 33]. Este conjunto de imagens devolvidas precisa ser mapeado coerentemente em uma forma visual dentro do espaço disponível para a visualização do usuário com ajuda de diversas características gráficas, como posição, tamanho, cor, forma e marcas visuais adicionais de cada imagem e do conjunto como um todo. Estas propriedades gráficas formam uma estrutura visual. Finalmente, o usuário poderá interagir com essa estrutura para extrair informações de interesse tornando uma representação estática em uma visão da estrutura. Este processo de mapeamento dos dados em formas visuais que suportam interação humana é conhecido como visualização [2].

A visualização merece especial cuidado dado que uma técnica de visualização adequada potencialmente proverá maior quantidade de informação em menor tempo. Por um lado, não adianta apresentar uma complexa estrutura de visualização se o usuário fica confundido e até desconfortável em interagir com o sistema de busca. Por outro lado, também não é conveniente apresentar uma estrutura tão simples que o usuário não consiga encontrar o que está procurando.

Na procura por uma técnica de visualização adequada para os usuários, podem-se também utilizar mecanismos de agrupamento para representar o universo de imagens. Estes mecanismos permitem agrupar os resultados obtidos em grupos coerentes de elementos nos quais imagens similares estejam localizadas em lugares fisicamente próximos [10, 11, 32, 34].

Uma característica desejável no processo de visualização é o *Relevance Feedback* (realimentação de relevância). Realimentação de relevância é uma técnica de modificação de consulta que tenta capturar as necessidades do usuário a partir do refinamento da consulta [35, 35–40]. Assim, após um número de iterações, a interface mostrará um conjunto de imagens relevantes para o usuário.

Este projeto de pesquisa está centrado na avaliação, especificação e desenvolvimento de técnicas de visualização (existentes e propostas) para os resultados de sistemas CBIR. Foram estudadas, implementadas e avaliadas algumas técnicas de visualização. Considera-se que estas técnicas, se usadas corretamente, são valiosas ferramentas de análise e exploração de grandes coleções de imagens.

1.3 Contribuições

As principais contribuições deste trabalho são:

- Estudo comparativo e análise de sete técnicas de visualização; quatro delas existentes na literatura e três novas técnicas propostas.
- Validação de duas técnicas da literatura nunca antes avaliadas [5]. Essas técnicas provêm uma distribuição com enfoque central baseando-se respectivamente em uma estrutura espiral e em um conjunto de anéis concêntricos.
- Especificação e implementação de três novas técnicas de visualização: uma técnica de visualização baseada em agrupamento, denominada Clustering Set; uma técnica híbrida de visualização baseada na técnica Clustering Set e em outra técnica recentemente proposta denominada Espiral [5]; e uma técnica híbrida também baseada na técnica Clustering Set e em outra técnica da literatura denominada Anéis Concêntricos [5].

- Especificação e implementação de um *framework* para desenvolvimento de novas estruturas visuais para sistemas de recuperação de imagens por conteúdo no qual foram implementadas as técnicas de visualização estudadas.

1.4 Estrutura da Dissertação

O presente documento está organizado em 5 capítulos:

- O **Capítulo 2** apresenta os conceitos e os trabalhos relacionados, descrevendo vantagens e desvantagens de algumas abordagens propostas na literatura;
- O **Capítulo 3** apresenta os métodos propostos;
- O **Capítulo 4** apresenta os experimentos desenvolvidos e a validação das técnicas estudadas.
- Finalmente, o **Capítulo 5** apresenta as conclusões do trabalho, assim como as principais contribuições e possíveis trabalhos futuros.

Capítulo 2

Trabalhos e Conceitos Relacionados

2.1 Considerações iniciais

Este capítulo aborda o estado da arte sobre visualização de imagens em sistemas CBIR, destacando os principais trabalhos da área e os conceitos relacionados para o desenvolvimento deste projeto. Primeiramente, apresenta-se o processo de recuperação de imagens por conteúdo - CBIR (seção 2.2) e mostra-se uma arquitetura típica de um sistema CBIR. Posteriormente, aborda-se a visualização de informação; especificamente o processo de transformação de dados (dados de maneira geral e no caso de imagens) em representações visuais (seção 2.3). A seguir, são descritas algumas abordagens de agrupamento (seção 2.4) de dados a serem usadas em capítulo posteriores. Após isso, realiza-se um estudo sobre as técnicas existentes para visualizar os resultados de uma consulta CBIR agrupando-as em 3 categorias detalhadas na seção 2.5. Finalmente apresentam-se brevemente as técnicas de avaliação na área de interação humano-computador (seção 2.6) que guiarão o desenvolvimento dos experimentos no Capítulo 4.

2.2 Recuperação de imagens por conteúdo

Um sistema de recuperação de imagens por conteúdo ou sistema CBIR é centrado na noção de similaridade de imagens. Dado um banco com um grande número de imagens, o usuário deseja recuperar as imagens mais similares a um padrão de consulta (normalmente uma imagem definida como exemplo). O processo de recuperação é baseado na comparação das imagens a partir do uso de descritores [1].

2.2.1 Arquitetura típica de um sistema CBIR

A Figura 2.1 mostra a arquitetura típica de um sistema CBIR proposta em [1]. Existem dois subsistemas principais: inserção de dados e processamento de consultas.

- **Inserção de dados:** representado pelos módulos tracejados. É responsável por extrair os vetores de características das imagens e armazená-los na base de imagens. O processo é realizado uma única vez para cada imagem e para cada descritor. Com este processo são obtidos valores numéricos que formam os vetores de características que representam o conteúdo das imagens da coleção.
- **Processamento de consultas:** representado pelos módulos restantes. A interface permite ao usuário especificar uma consulta utilizando-se um padrão de consulta e visualizar as imagens recuperadas. Este módulo extrai o vetor de características do padrão de consulta e aplica uma função de distância para avaliar a similaridade entre a imagem de consulta e as imagens da base. Posteriormente, esse módulo ordena as imagens da base de acordo com a similaridade e retorna as mais similares para o módulo de interface.

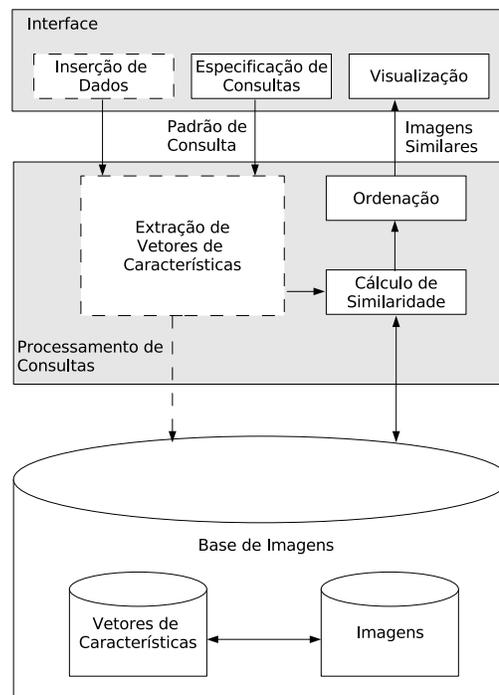


Figura 2.1: Arquitetura de um sistema de recuperação de imagens por conteúdo [1].

2.2.2 Sistemas CBIR

Nestas últimas duas décadas, vários sistemas CBIR têm sido desenvolvidos. Um dos primeiros e mais difundidos foi o QBIC (*Query By Image Content*) [6]. Ele estabeleceu as bases para alguns outros sistemas, como Virage [41]. O QBIC é um sistema comercial proposto em 1993 pela IBM [6]. Desde então, foram feitas numerosas atualizações ao produto [42, 43]. Ele permite que os usuários definam graficamente consultas baseadas em propriedades visuais (cor, forma, textura); também permite realizar consultas baseadas em imagens de exemplo, *sketch* de usuário e padrões de cor ou textura. No último caso, a cor desejada é ajustada por *sliders* de movimento. A função de similaridade está baseada em distância euclidiana e a indexação no uso de estruturas do tipo R*-trees.

O Amore (*Advanced Multimedia Oriented Retrieval Engine*) [7, 44] é um motor de busca na *World Wide Web* que permite a recuperação de imagens baseando-se em palavras-chave, em uma imagem de exemplo, ou na combinação das duas. O Amore permite a extração de cor e forma. O usuário pode definir a importância de propriedades visuais nos resultados e, até mesmo, escolher se os resultados devem ser visual ou semanticamente similares à consulta dada. Para mostrar os resultados, esses sistemas usam a forma de *grid de thumbnails*.

O PRISM (*Perceptually-Relevant Image Search Machine*) [8] é um protótipo para CBIR cuja interface permite refinar consultas usando a técnica de realimentação de relevância. Sua principal desvantagem é que a visão do usuário é bastante limitada, dado que poucas imagens são apresentadas por vez. Essas imagens são mostradas em uma “linha” horizontal chamada *filmstrip*.

O ImageGrouper [45, 46] propõe um novo conceito de *Query-by-Groups* considerando um ou mais conjuntos de imagens como a unidade básica da consulta, permitindo, além disso, adicionar grupos negativos de imagens que não devem ser recuperadas pelo sistema.

Para mais detalhes sobre os sistemas CBIR existentes, podem-se consultar os trabalhos de Veltkamp [47] e Venters [48], os quais realizaram um levantamento de um grande número de sistemas CBIR existentes até 2002.

2.3 Visualização de informação

Gráficos e imagens usadas do jeito adequado podem ser importantes ferramentas para expressar o conhecimento escondido em grandes volumes de dados. Um exemplo clássico da efetividade da transmissão de informação por meio de gráficos é o histórico “*Carte figurative des pertes successives en hommes de l’Armée Française dans la campagne de Russie 1812-1813*” (Figura 2.2 ²) de Charles Joseph Minard, que mostra o terrível destino da força

²Imagem de domínio público disponível em <http://en.wikipedia.org/wiki/File:Minard.png>. Último acesso em 08/2010.

permite manipular a visualização ou os dados ocultos para explorar tais mudanças [50].

Os tipos de dados a serem visualizados [51, 52] podem ser:

- unidimensional: documentos, código fonte, lista sequencial;
- bidimensional: mapas, grades;
- tridimensional: objetos físicos;
- multi-Dimensional: dados com atributos múltiplos;
- temporal: dados que variam ao longo do tempo;
- árvore/hierárquica: sistemas de arquivos, gráfico de organização;
- rede: dados que representam relações entre objetos.

2.3.1 Processo de transformação de dados em representações visuais

A informação inicialmente encontra-se em estado bruto e precisa ser processada, organizada e representada de um jeito visual compreensível aos sentidos do usuário. Card et al. [2] propuseram um modelo, mostrado na Figura 2.3, para representar este processo.

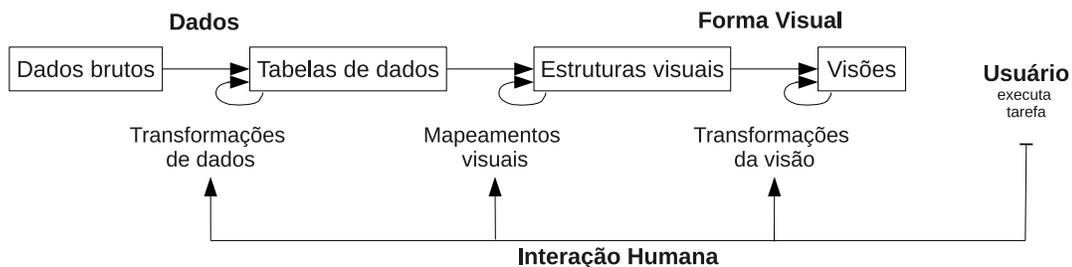


Figura 2.3: Modelo do processo de visualização de informação [2].

Segundo esse modelo, os dados brutos precisam passar por um processo de transformação. Esse processo organiza as inter-relações desses dados de maneira similar à organização de uma tabela em um banco de dados relacional, gerando assim uma tabela de dados [53]. Uma vez organizados em uma tabela de dados, é preciso representar os dados visualmente. Com tal fim, os dados são mapeados em estruturas visuais. Uma estrutura visual é uma representação visual dos dados por meio de marcas e propriedades gráficas para codificar informação. É

importante que o mapeamento preserve os dados. Um mapeamento é expressivo se todos os dados em uma tabela de dados são também expressados na estrutura visual, sem adicionar nem omitir informação. Considera-se que um mapeamento é mais efetivo quando é mais rápido de ser interpretado [2]. Uma vez que os dados tenham sido representados por meio de estruturas, é possível permitir ao usuário interagir com uma estrutura para extrair dela mais informação de interesse. Esse processo consiste na criação de novas visões da estrutura visual de acordo com a necessidade do usuário e é chamado transformação da visão. Card et al. [2] apontam três tipos de transformações da visão:

- Investigações de local: Interagindo sobre o local de um dado, pode-se revelar informação adicional. Um exemplo desta transformação é a técnica de detalhes sob demanda. Esta técnica expande um ou mais objetos na visão revelando mais informações. Essa técnica foi adotada no programa *Film finder* [3]. Neste programa, cada quadrado representa um filme. Quando o usuário se aproxima de um filme em particular, abre-se uma janela que exibe os detalhes do filme escolhido (Figura 2.4).

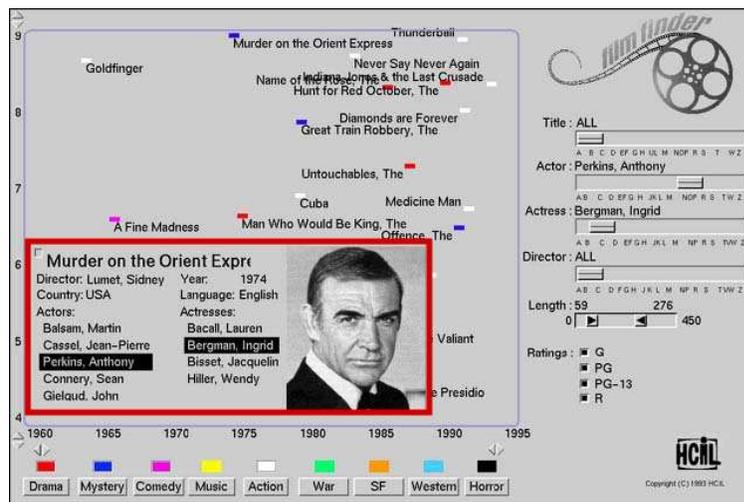


Figura 2.4: Interface do sistema FilmFinder [3].

- Controles de ponto de vista: Consiste em um conjunto de técnicas como aproximação e afastamento (*zoom*), deslocamento (*pan*), seleção de uma parte a ser examinada (*clip*), e visão geral + detalhes (*overview + details*). Um exemplo deste tipo de transformações encontra-se no conhecido Google Earth (Figura 2.5).

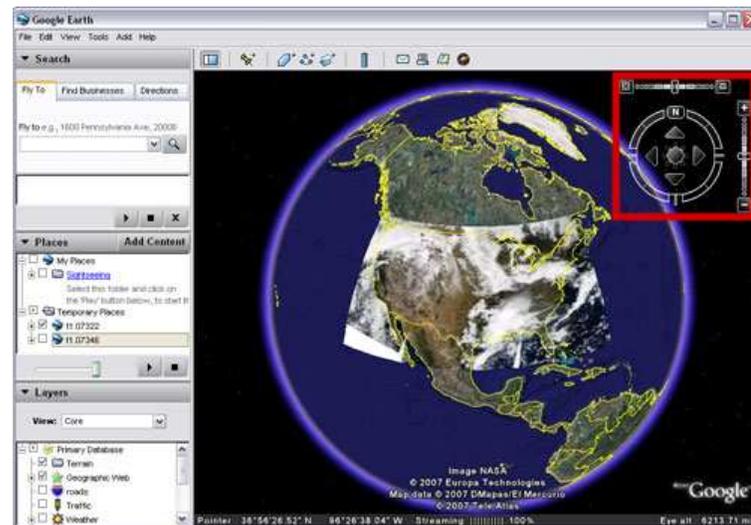


Figura 2.5: Interface do sistema Google Earth [4].

- Distorção: É uma transformação visual que modifica a estrutura visual para criar visões de foco e contexto. O panorama (*overview*) e ao mesmo tempo o detalhe são combinados em uma mesma estrutura visual. A distorção é considerada efetiva se o usuário puder perceber a estrutura visual original não distorcida através da distorção. [2].

2.3.2 Processo de transformação de imagens em representações visuais

Baseados no modelo do processo de visualização de informação proposto por [2], foi feita uma adaptação para CBIR em [5]. Esse modelo define o processo de visualização dos resultados de um sistema CBIR a partir de transformações sucessivas que ocorrem a partir dos dados brutos até as representações visuais mediante três transformações: transformação dos dados (TD), mapeamento visual (MV) e transformação da visão (TV).

Em um processo CBIR, os dados brutos estão representados pelo banco de dados de imagens. O conteúdo das imagens é extraído em vetores de características por meio de algoritmos de processamento de imagens. Esta fase gera uma tabela de dados composta por cada imagem e pelo vetor de características correspondente (T1). Posteriormente, quando o usuário define uma imagem de consulta (IC), um vetor de características da nova imagem (FIC) é extraído automaticamente e computa-se a distância entre o FIC e os vetores de características armazenados em T1, gerando uma segunda Tabela de dados (T2). Após esse processo, o usuário pode selecionar o número de imagens que deseja recuperar, o que gera

uma terceira tabela de dados (T3), subconjunto de T2. Posteriormente, a tabela de dados T3 é mapeada em diversas estruturas visuais (MV). Finalmente, o usuário pode interagir com a interface usando diversas transformações da visão, como as detalhadas na seção 2.3.1.

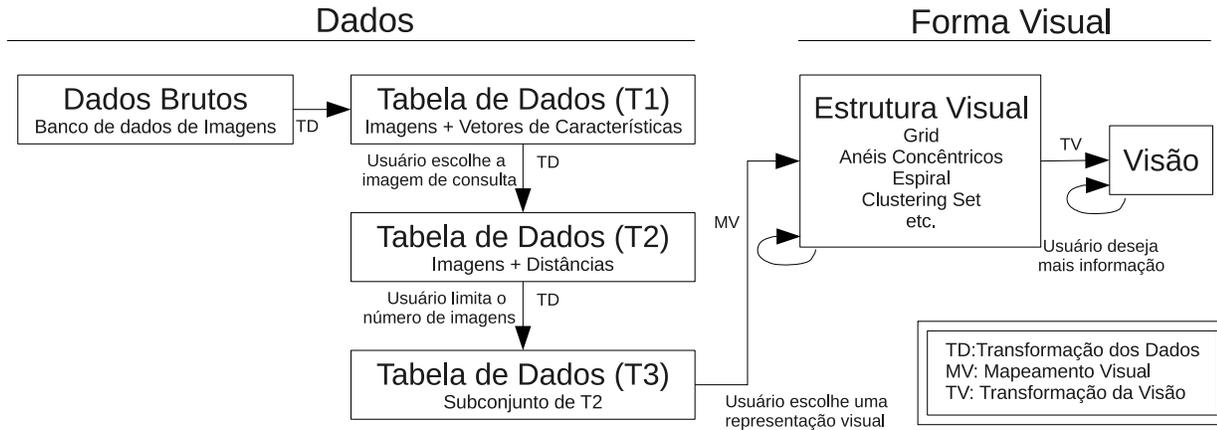


Figura 2.6: Adaptação do modelo do processo de visualização de informação para sistemas CBIR [5].

2.4 Agrupamento de dados

Agrupamento (*clustering*) pode ser definido como a classificação não supervisionada de padrões (observações, itens de dados, ou vetores de características) em grupos (*clusters*) [54]. Este tipo de classificação não requer categorias pré-definidas e os dados não precisam pertencer a um grupo previamente, como acontece na classificação supervisionada. Nessa abordagem, o problema é descobrir a estrutura dos dados, se existir alguma [55]. Deve-se, portanto, averiguar a existência de grupos e a quantidade deles, tentando maximizar a similaridade dos elementos dentro de um grupo (*intra-cluster*) e minimizar a similaridade dos elementos que pertencem a grupos diferentes (*inter-cluster*). A seguir são apresentadas algumas técnicas de agrupamento.

2.4.1 Abordagem aglomerativa

O enfoque aglomerativo utiliza uma abordagem *bottom-up*, na qual cada elemento do conjunto é inicialmente associado a um grupo distinto e novos grupos são formados pela união dos grupos existentes até que algum critério de parada seja satisfeito [54]. Essa união ocorre de acordo com alguma medida que forneça a informação sobre quais deles estão mais

próximos uns dos outros (Figura 2.7).

Algumas técnicas aglomerativas são descritas a seguir:

Single link

Single link [54], também chamado *nearest neighbour*, é um método que consiste em juntar, a cada passo, os dois grupos que estão mais próximos em um dado momento. Calcula-se a distância entre dois grupos como a distância dos dois elementos mais próximos que pertencem a grupos distintos. Os dois grupos mais próximos são unidos. Considere $C = \{c_1, c_2, \dots, c_k\}$, um conjunto com k grupos. Para $c_i \in C$, $i = 1, 2, \dots, k$, denota-se x_{c_i} um elemento $x \in c_i$. Também denota-se por $d(x_{c_i}, x_{c_j})$ a distância entre dois elementos x_{c_i} e x_{c_j} . A equação a seguir mostra a distância $D(c_i, c_j)$ entre os grupos c_i e c_j .

$$D(c_i, c_j) = \text{Min}\{d(x_{c_i}, x_{c_j})\} \quad (2.1)$$

Em agrupamentos hierárquicos, a cada passo, são unidos os dois grupos c_i e c_j que minimizem $D(c_i, c_j)$. A complexidade de tempo desse método é $O(n^2 \log n)$, porém é possível fazer uma otimização, por meio do uso de uma Árvore Geradora Mínima (MST - *Minimum Spanning Tree*) dos dados, o que poderia ser construído em $O(n \log^2 n)$. A complexidade de espaço requerido por este algoritmo é de ordem quadrática, $O(n^2)$, já que é preciso armazenar uma matriz de similaridade de tamanho $n \times n$.

Complete link

Complete link [54], também chamado *farthest neighbor*, é um método de agrupamento no qual a distância entre grupos é definida como a distância entre o par mais distante de objetos de grupos distintos. Usando a mesma notação do método anterior, define-se a distância entre dois grupos como:

$$D(c_i, c_j) = \text{Max}\{d(x_{c_i}, x_{c_j})\} \quad (2.2)$$

Em agrupamentos hierárquicos, a cada passo, são unidos os dois grupos c_i e c_j cuja distância minimiza $D(c_i, c_j)$. A complexidade de tempo desse método é $O(n^2 \log n)$. A complexidade em espaço é a mesma que o algoritmo anterior, $O(n^2)$, já que também é preciso armazenar uma matriz de similaridade.

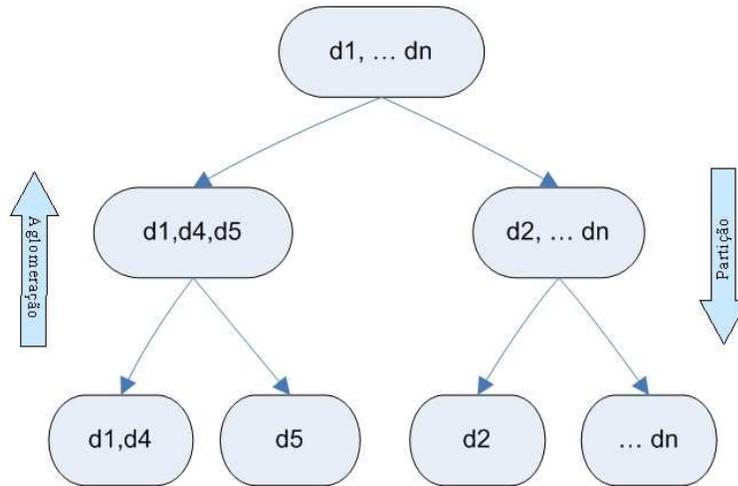


Figura 2.7: Abordagens Aglomerativa e Divisiva

2.4.2 Abordagem divisiva ou particional

Ao contrário da abordagem aglomerativa, o enfoque divisivo utiliza uma abordagem *top-down*. No início, tem-se um único grupo contendo todos os elementos do conjunto. A cada passo, são efetuadas divisões, formando novos grupos de tamanhos menores, até que seja satisfeito algum critério de parada [54] (Figura 2.7).

Alguns algoritmos que particionam ou dividem o espaço e geram agrupamentos são discutidos a seguir:

1. k-means

K-means [56, 57] é um conhecido algoritmo para agrupamento que fornece uma maneira simples de classificar os dados. Porém, o número de grupos k deve ser fixado a priori. Os centros (doravante centroides) iniciais de cada grupo também precisam ser especificados (usualmente elementos aleatórios são escolhidos). Cada dimensão do centroide será representada pela média aritmética na mesma dimensão de todos os elementos do conjunto a ser agrupado. Portanto, o centroide pode ser um ponto fictício fora do conjunto de elementos existentes.

O algoritmo é descrito a seguir:

Algoritmo K-means

- 1: Definir o número de clusters k .
 - 2: Escolher k centroides iniciais (prossivelmente aleatórios).
 - 3: **Enquanto** os centroides mudam:
 - 4: Cada elemento é associado ao centróide mais próximo.
 - 5: Considerando a nova associação dos elementos, cada centróide é recomputado.
 - 6: **fim enquanto**
-

A complexidade de tempo do algoritmo k-means é $O(nkt)$, onde n é o número de elementos do conjunto de dados, k é o número de grupos definido a priori e t é o número de iterações do algoritmo até se estabilizar (até que os centroides não mudem mais). A complexidade de espaço é $O(k)$.

2. k-medoids

Este método é similar ao anterior. Em contraste ao k-means, o k-medoids escolhe como centros elementos existentes. Um medoide é um elemento pertencente ao grupo que representa esse conjunto. A distância média do medoide a todos os elementos do grupo é mínima. Qualquer outro elemento no grupo tem uma distância média maior ou igual. Uma vez que este método usa um elemento próprio do conjunto, ele é mais robusto que o k-means com relação aos *outliers*. O algoritmo k-medoids mais usado é o *Partitioning Around Medoids* (PAM) [58].

2.4.3 Abordagem DAHC

O *Divisive-Agglomerative Hierarchical Clustering* [59] é uma técnica híbrida de agrupamento que combina as características das estratégias particional e aglomerativa para obter uma estrutura hierárquica baseada somente na função de similaridade pré-definida. Essa combinação gera boas soluções com menor número de operações computacionais, permitindo uma redução considerável no tempo de processamento da consulta. Essa redução é possível porque não é preciso fazer uma varredura de todas as imagens. É necessário apenas comparar o elemento de consulta com os elementos representativos de cada grupo encontrados na estrutura hierárquica (árvore).

2.5 Técnicas para a visualização de resultados de sistemas CBIR

Simultaneamente ao rápido desenvolvimento de técnicas usadas em CBIR, aumentou velozmente o número de sistemas CBIR. Lamentavelmente, o interesse pela última etapa do processo CBIR, a visualização dos resultados de uma consulta, não cresceu na mesma proporção. Esse fato é preocupante dado que não adianta, por exemplo, usar descritores mais eficazes ou estruturas mais complicadas de indexação se o usuário final não souber como interagir com o sistema ou compreender os resultados mostrados pelo sistema de busca. A apresentação dos resultados da busca é talvez um dos fatores mais importantes na aceitação e popularidade de um sistema de recuperação de imagens [18].

Cada sistema CBIR usa distintos mecanismos para mostrar os resultados de uma busca. Após estudar as distintas técnicas usadas nesses sistemas, propõe-se a seguinte categorização dessas técnicas: técnicas de exibição simples, técnicas com enfoque central e técnicas baseadas em agrupamento. Essas técnicas são descritas a seguir.

2.5.1 Técnicas de exibição simples

Enquadra-se neste conjunto qualquer técnica que não faz nenhuma mudança física na imagem com exceção da mudança de tamanho, concentrando-se na simples disposição do conjunto ordenado de imagens numa linha ou grade. Este é o conjunto mais amplamente usado para visualização de resultados de sistemas CBIR. Porém, essas técnicas utilizam um enfoque pouco eficiente, porque não consideram relações intrínsecas entre as imagens, exibindo imagens com diferentes graus de similaridade (baseada em conteúdo) na mesma distância física ao padrão de consulta. Se uma coleção é suficientemente pequena, essa técnica usualmente funciona bem, mas em uma coleção grande, a navegabilidade do usuário é severamente afetada.

1. *Grid*

Esta é a técnica tradicional e mais comum para a visualização de resultados de sistemas CBIR provavelmente devido à facilidade de implementação e à grande familiaridade dos usuários. As imagens são dispostas de uma maneira matricial, na qual a primeira posição situa-se no canto superior esquerdo e é ocupada pelo padrão de consulta e as posições sucessivas são ocupadas pelo conjunto ordenado de imagens relevantes distribuídas da esquerda para direita e de cima para baixo. Mesmo sendo o mais usado, este enfoque é ineficiente, pois não reflete a ordem de similaridade das imagens, já que mostra imagens com diferentes graus de similaridade a uma mesma distância física ao

padrão de consulta. Entre os sistemas CBIR que utilizam essas técnicas estão [6, 7, 41]. As Figuras 2.8 e 2.9 mostram alguns exemplos de sistemas CBIR cuja interface de resposta utiliza esta técnica.

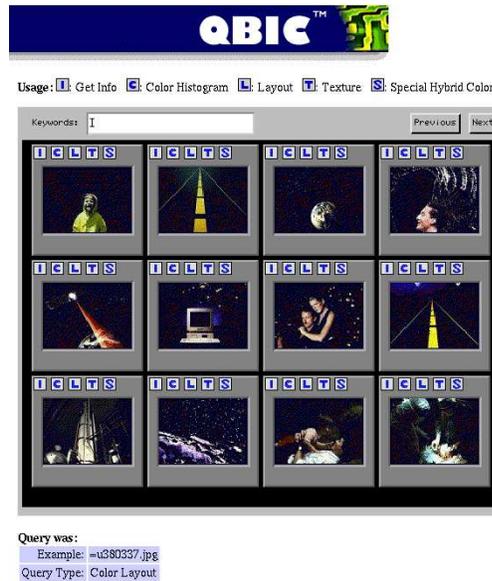


Figura 2.8: Interface do sistema QBIC [6].



Figura 2.9: Interface do sistema Amore [7].

No caso do sistema Amore, uma outra forma de visualizar os resultados ainda foi proposta: Uma *scatter plot*, com valores de cor e forma nos eixos, ou em uma parede de perspectiva [44].

A Figura 2.10 mostra um protótipo da estrutura implementada usada nos testes experimentais detalhados no Capítulo 4.



Figura 2.10: Estrutura visual Grid.

2. *Filmstrip*

Esta abordagem foi proposta no desenvolvimento do PRISM [8]. Utiliza um “*filmstrip*” (Figura 2.11) para mostrar as imagens. Um “*filmstrip*” é uma sequência de imagens localizadas ao longo de uma linha. A principal desvantagem desta técnica é a pouca informação oferecida ao usuário por vez; situação especialmente ruim em grandes bancos de dados. Em contraposição, existe uma ampla área de trabalho que permite a interação do usuário mediante a escolha de exemplos positivos.



Figura 2.11: Interface do sistema PRISM [8].

2.5.2 Técnicas com enfoque central

Dadas as deficiências do conjunto de técnicas anterior, surgiram algumas abordagens [11, 34, 60] cujo objetivo é localizar o centro da estrutura no centro da área de visualização.

Nesse lugar é posta a imagem de consulta, facilitando, desse modo, a exploração das imagens ao seu redor. Em alguns casos, contempla-se a variação do tamanho e cores das imagens, sendo estas menores e mais escuras à medida que se afastam do centro. Na área de visualização de informação, este conjunto de técnicas também são conhecidas como técnicas radiais [61]

Algumas técnicas que pertencem a este grupo foram incluídas nos experimentos realizados, como será visto no Capítulo 4.

1. Anéis concêntricos

Esta abordagem foi proposta em [5]. As imagens são distribuídas sobre uma série de anéis concêntricos. Cada anel é o círculo, definido em coordenadas polares, formado pelo conjunto de pontos com raio $r = k$, onde k é uma constante. A constante k incrementa-se para cada anel. Este incremento diminui a cada anel, ou seja, anéis sucessivos estarão mais próximos conforme k é incrementado.

A Figura 2.12 ilustra a técnica de anéis concêntricos.

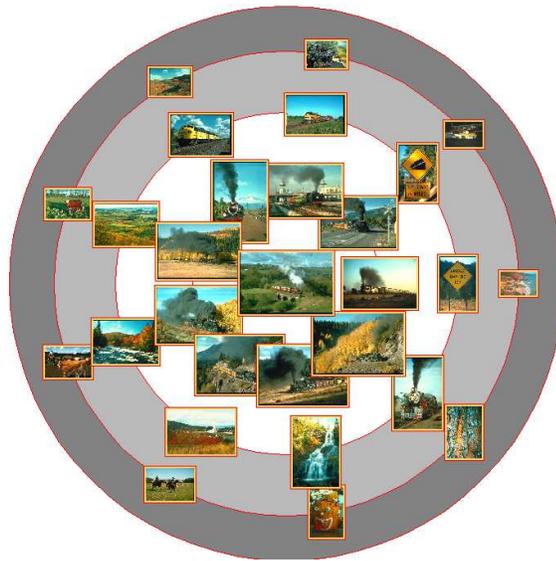


Figura 2.12: Estrutura visual Anéis Concêntricos.

2. Espiral

Esta representação foi inicialmente proposta em [5]. Diversas curvas espirais foram analisadas com o intuito de achar aquela que melhor encaixe com as características desejadas: a curva espiral se aproxima de si mesma à medida que se afasta do centro.

A espiral mais conhecida é a espiral arquimedeaniana, a qual deve seu nome ao matemático

Arquimedes quem a estudou. Pode ser representada pela equação polar:

$$r = k\Theta^a \quad (2.3)$$

onde r é a distância radial, Θ é o ângulo polar, k é uma constante, e a é uma constante que determina quão rapidamente a espiral se enrolará. Esta equação definiu as bases para a família de espirais Arquimedeanas que respeitam a mesma fórmula variando apenas os parâmetros.

A espiral original de Arquimedes foi proposta com $a = 1$ e, neste caso, voltas sucessivas da espiral têm uma distância de separação constante e, portanto, foi descartada. As espirais Hiperbólica e de Fermat são tipos específicos da família de espirais Arquimedeanas. A espiral hiperbólica considera $a = -1$. As espirais hiperbólica e logarítmica foram descartadas porque se afastam rapidamente da origem. A espiral de Fermat, também conhecida como espiral parabólica, considera $a = 0,5$ então a sua equação polar poderia ser descrita como

$$r^2 = k^2\Theta \quad (2.4)$$

Selecionamos a espiral de Fermat porque voltas sucessivas da espiral estão cada vez mais próximas à medida que se afastam do centro. A Figura 2.13 exibe um protótipo da estrutura implementada.

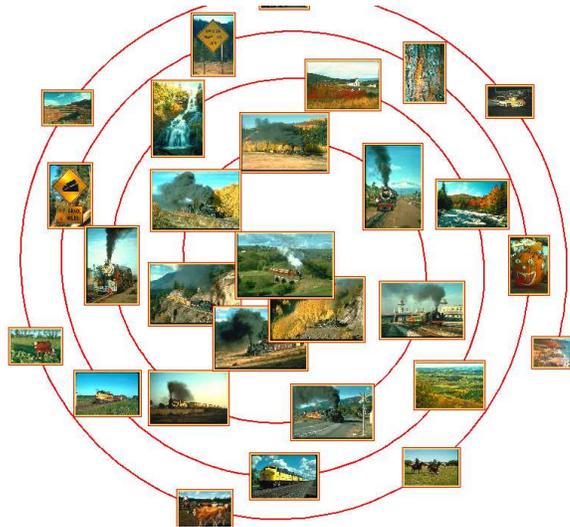


Figura 2.13: Estrutura visual Espiral.

3. Ring

Em [9] foi estudada uma outra técnica similar à abordagem mostrada na seção 1. As

imagens são organizadas e apresentadas ao usuário em forma de anéis concêntricos não visíveis. Além da estrutura, existem um conjunto de controles (anéis visíveis ao redor da estrutura) que fornecem opções de busca para navegar pelo conjunto de imagens (ver Figura 2.14).

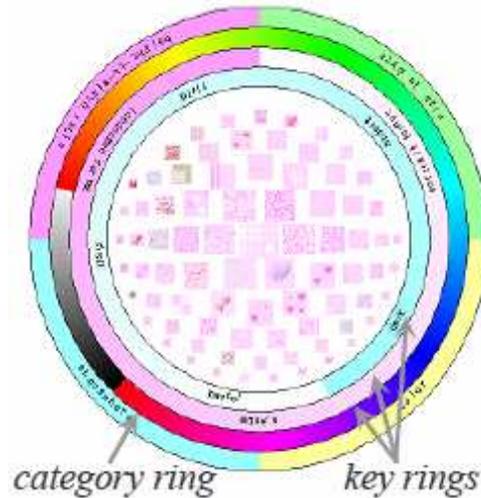


Figura 2.14: Estrutura visual Ring [9].

2.5.3 Técnicas baseadas em agrupamento

Em um sistema CBIR tradicional, as imagens retornadas são organizadas em ordem de similaridade à consulta dada [6, 8, 41, 62]. Geralmente a disposição destas imagens na interface não considera as relações intrínsecas entre as imagens. Para mostrar um conjunto mais coerente de imagens, estas poderiam estar próximas das suas imagens similares. Isto pode ser feito empregando técnicas de agrupamento (*clustering*).

Assim, as técnicas para CBIR baseadas em agrupamento são aquelas em que as imagens semelhantes são organizadas dentro do mesmo grupo. A localização de cada grupo dependerá da similaridade do grupo em relação à consulta.

Alguns trabalhos sobre CBIR encontrados na literatura (ver Tabela 2.1 e Tabela 2.2 na coluna Red. Dimens.) reduzem o número de dimensões do espaço multidimensional no qual estão representadas as imagens.

A principal vantagem deste conjunto de técnicas é fornecer ao usuário uma melhor visualização e entendimento do universo de imagens disponível. Em contraste, o principal problema desta abordagem é a possível sobreposição de imagens (*overlapping*). É importante lidar com este problema dado que imagens muito pequenas não são visíveis (por exemplo, Chen et al. [32] mostra a coleção como um conjunto de pontos) e imagens grandes aumentam o grau de

sobreposição, dificultando a identificação de possíveis imagens de interesse.

Neste contexto, alguns trabalhos têm sido desenvolvidos [10, 32, 34, 63]. As Figuras 2.15 e 2.16 mostram exemplos desta abordagem.



Figura 2.15: Interface proposta em [10].

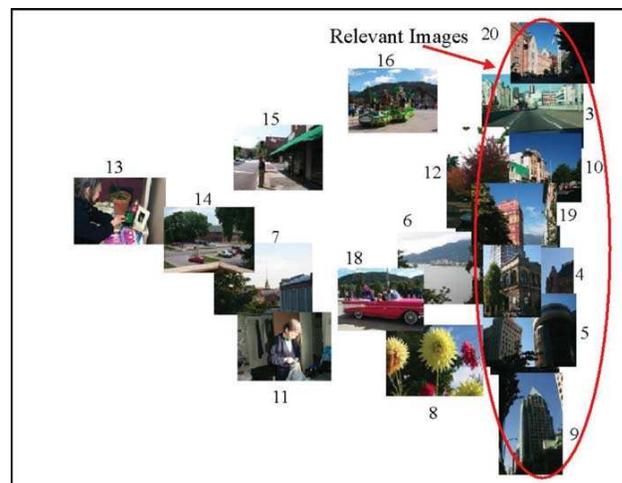


Figura 2.16: Visualização baseada em clusters proposta em [11].

A Tabela 2.1 apresenta um conjunto de trabalhos que usam técnicas de agrupamento

sobre coleções de imagens. Nesta tabela, as características exploradas por coluna são: Trabalho (citação), Descrição, Agrupamento (técnica de agrupamento utilizada), Red. Dimens. (técnica de redução de dimensionalidade empregada ou a ausência dela) e finalmente R.F. - *relevance feedback* (presença ou não de técnicas de realimentação de relevância).

Trabalho	Descrição	Agrupamento	Red. Dimens.	R.F.
Bathia [64]	Cria uma estrutura de dados hierárquica baseada em agrupamento.	Hierárquica	-	Sim
Chen et al. [65]	Tenta agrupar imagens semanticamente similares supondo que imagens semanticamente similares encontram-se agrupadas em algum espaço de características.	Graph-partitioning - Ncuts.	-	Sim
Deng Cai et al. [66]	Agrupar imagens de páginas web considerando a representação visual das imagens, a representação textual extraída do código html e a informação dos enlaces (links).	Spectral clustering, k-means	-	Não
Zheng et al. [67]	Propõe-se um novo método chamado LPC.	LPC	LPP	Não
Rocha et al. [59]	Técnica híbrida sobre agrupamento hierárquico.	Qualquer método de clustering divisivo e aglomerativo	-	Não

Tabela 2.1: Trabalhos correlatos de técnicas de agrupamento aplicadas a conjuntos de imagens. O agrupamento não é usado na etapa de visualização.

Siglas: LPC: Locality Preserving Clustering; LPP: Locality preserving projections.

A Tabela 2.2 apresenta um resumo dos principais trabalhos de CBIR que utilizam técnicas de agrupamento para a visualização dos resultados. Nela as características exploradas

são: Trabalho (citação), Agrupamento (técnica de agrupamento utilizado), Red. Dimens. (técnica utilizada para reduzir o espaço multi-dimensional), R.F. - *relevance feedback* (realimentação de relevância), Sobreposição (*overlapping*).

Trabalho	Agrupamento	Red. Dimens.	R.F.	Sobreposição
Nguyen et al [10]	Graph-ISOMAP	MDS	Sim	Sim
chen et al. [32]	Graph-PFNETs	PFNETs	Não	Não
Moghaddam et al. [11, 34]	PCA Splat	PCA	Não	Sim
Bartolini et al. [63]	Hierárquica	MDS	-	-
Laaksonen et al. [68]	SOM	SOM	Sim	Não
Rodden et al. [69]	SOM	MDS	-	Sim

Tabela 2.2: Trabalhos correlatos de técnicas de visualização que utilizam agrupamento em sistemas CBIR.

Siglas: PCA: Principal Component Analysis; MDS: Multi-Dimensional Scaling; MMM: Markov Model Mediator; PFNETs: PathFinder Nets; SOM: Self-Organizing Maps.

2.5.4 Outras Técnicas

Incremental Board

A abordagem Incremental Board foi proposta em [12]. Esta técnica foi inicialmente criada para visualização de conjuntos de textos e estendida para imagens. É uma técnica incremental (ideal para conjuntos de dados que estão em constante mudança) que mantém uma disposição coerente de elementos tentando preservar as similaridades de um conjunto de dados multidimensional distribuídos numa grade 2D. Itens similares são colocados juntos, enquanto vizinhos não similares são afastados. Assim, similarmente às técnicas baseadas em agrupamento, esta técnica preserva visualmente as relações intrínsecas entre as imagens. A Figura 2.17 ilustra esta técnica.



Figura 2.17: Estrutura visual Incremental Board [12].

2.5.5 Resumo de sistemas CBIR segundo a técnica de visualização usada

A Tabela 2.3 apresenta um conjunto de sistemas CBIR encontrados na literatura e o respectivo método de visualização utilizado por cada um para mostrar os resultados de uma consulta dada. O uso de relevance feedback também é considerado. Cabe ressaltar a grande preferência dos desenvolvedores pelo uso da tradicional grade de imagens possivelmente pela facilidade de implementação e pela familiaridade com o usuário. Porém isso não significa que este método seja o mais adequado ou confortável para os usuários finais. Testes experimentais com usuários, apresentados no Capítulo 4, tratam destas questões.

Sistemas CBIR	Técnicas de visualização				R.F.
	Grade	Centered	Clustering	3D	
Amore [7]	x				
Chabot [70]	x				
Diogenes [71]	x				Não
ImageGrouper [45, 46]	x				
Mars [72]	x				Sim
3DMars [72]				x	Sim
MetaSEEK [73]	x				
NETRA [74]	x				Sim
Photobook [75]	x				
PIBE [63]			x		
PicSOM [76]	x		x		
Prism [8]	x				
QBIC [6, 42, 43]	x				
RETIN [77–79]	x				
RING [9]		x			
Virage [41]	x				
Viper [80]	x				
VisualSEEK [81]	x				

Tabela 2.3: Técnicas de visualização dos sistemas CBIR da literatura

2.6 Avaliação na área de Interação Humano-Computador

Segundo [82], idealmente, a avaliação deve ocorrer durante o ciclo de vida do *design* e seus resultados devem ser utilizados para melhorias gradativas da interface. É claro que não se pode pretender efetuar extensivos testes experimentais durante todo o processo de *design*, mas técnicas informais e analíticas devem ser utilizadas. Portanto, não se pode efetuar a avaliação somente no final do processo de desenvolvimento de qualquer sistema. Deve-se ter claro também que exaustivos testes não garantem perfeição. Sempre resta um grau de incerteza e, portanto, o planejamento deve prever a realização de avaliação e o reparo de problemas contínuos.

O objetivo de uma avaliação é conhecer o que os usuários precisam e os problemas que experimentam [82]. Assim, dependendo do caso, podem-se usar distintos métodos de avaliação, classificados a seguir:

- **Inspeção de Usabilidade:** É um conjunto de métodos nos quais avaliadores inspecionam ou examinam aspectos relacionados à usabilidade de uma interface de usuário. Os avaliadores podem ser especialistas em usabilidade, consultores de desenvolvimento de software, especialistas em um determinado padrão de interface, usuários finais, etc. Dentre os métodos de inspeção, destacam-se:
 - **Avaliação Heurística:** A inspeção é feita baseada em uma pequena lista de heurísticas de usabilidade.
 - **Revisão de *Guidelines*:** A inspeção se baseia em uma lista de aproximadamente 1000 *guidelines*.
 - **Inspeção de Consistência:** Verifica-se a consistência quanto à terminologia, cores, layout, formatos, etc.
 - **Percurso cognitivo:** O avaliador simula o usuário percorrendo a interface e executando tarefas típicas.
- **Testes de Usabilidade:** São métodos de avaliação centrados no usuário para verificar a facilidade de uso de uma interface. Incluem métodos experimentais ou empíricos, métodos observacionais e técnicas de questionamento.

DECIDE é um *framework* que visa guiar o processo de avaliação de um software por meio de um conjunto de passos a serem seguidos. O seu nome deve-se às iniciais do conjunto de passos recomendados para uma avaliação, na língua original na qual foi proposto (Inglês). Estes são:

- ***Determine the goals* (Determinar os objetivos):** Para fazer uma avaliação relevante, o primeiro passo é reconhecer os objetivos da avaliação para proceder com a formulação de questões apropriadas que permitam conseguir o que é esperado do experimento. As perguntas a seguir devem ser respondidas.
 - Quais são os objetivos de alto nível da avaliação? Os objetivos devem guiar a avaliação. Desta forma, determinar quais são estes objetivos é o primeiro passo ao planejar a avaliação. Estes objetivos guiarão a escolha do paradigma de avaliação que será utilizado no estudo.
 - Quem quer a avaliação e por quê?
- ***Explore the questions* (Explorar as questões):** Devem ser identificadas as questões que ajudem a satisfazer os objetivos identificados no item anterior.

- *Choose the evaluation paradigm and techniques* (**Escolher o paradigma de avaliação e técnicas**): Este passo consiste em escolher o paradigma de avaliação e as técnicas a serem usadas. O paradigma de avaliação determinará os tipos de técnicas que são usadas. Questões práticas e éticas também devem ser levadas em conta, pois não adianta escolher as técnicas mais apropriadas se não puderem ser utilizadas por serem muito caras, por tomarem muito tempo, ou por requererem um equipamento não disponível.
- *Identify the practical issues* (**Identificar as questões práticas**): Existem muitas questões práticas a considerar quando se faz qualquer tipo de avaliação e é importante identificá-las antes de começar qualquer avaliação. Algumas questões a serem consideradas incluem usuários, facilidades e equipamento, calendários e orçamentos e competência dos avaliadores. Dependendo da disponibilidade dos recursos, os compromissos podem envolver a adoção ou substituição de técnicas.
- *Decide how to deal with the ethical issues* (**Decidir como lidar com questões éticas**): Algumas organizações proveem códigos de ética que devem ser seguidos pelos seus membros, especialmente se as suas atividades envolvem outras pessoas. Por exemplo, o anonimato dos participantes deve ser preservado. De modo similar, não deve ser possível identificar indivíduos na redação dos relatórios. Em geral as seguintes *guidelines* ajudarão a garantir que as avaliações são feitas eticamente e que os passos adequados para proteger os direitos dos usuários são levados em consideração.
 - Deve-se dizer aos participantes qual é o objetivo do estudo e o que exatamente se espera. Devem-se resumir, em linhas gerais, o processo de avaliação, o tempo que durará o experimento, o tipo de dados que será coletado e como esses dados serão analisados.
 - Deve-se explicar que dados demográficos, financeiros, de saúde e outras informações sensíveis são confidenciais. Deve-se manter a anonimidade também se áudio ou vídeo são utilizados.
 - Deve-se informar aos usuários que são livres para interromper a avaliação em qualquer momento se sentirem desconforto com o procedimento.
 - Se é possível, deve-se pagar aos usuários, porque isto cria uma relação formal.
 - Devem-se evitar citações que possam revelar a identidade de uma pessoa. Se alguma citação precisa ser reportada, podem-se substituir palavras que possam revelar a fonte com palavras representativas entre colchetes como [x].
 - Deve-se pedir permissão ao usuário para citação se for preciso, prometer anonimato e oferecer uma cópia do relatório com resultados obtidos antes de ser distribuído.

Uma regra geral a ser considerada quando se faz avaliação é *fazer aos outros somente aquilo com o que você não se importaria se fosse feito com você*.

- *Evaluate, interpret, and present the data* (**Avaliar interpretar e apresentar os dados**): Deve-se decidir também que dados coletar, como analisá-los e como apresentá-los. Algumas questões surgem nesta etapa como: os dados serão tratados estatisticamente? Se existir dados qualitativos, como serão analisados e representados?

Confiabilidade: A técnica é confiável? Confiabilidade ou consistência significa que a técnica produz os mesmos resultados em ocasiões separadas sob as mesmas circunstâncias. Diferentes processos de avaliação têm diferentes graus de confiabilidade.

Validade: A técnica de avaliação medirá o que se deseja que meça? A validade analisa se a técnica de avaliação mede o que deveria medir. A validade envolve a técnica e a forma na qual é desenvolvida.

Escopo: Os resultados podem ser generalizados? O escopo refere-se a quão generalizáveis podem ser as descobertas da pesquisa.

2.7 Considerações finais

Neste capítulo foram apresentados os conceitos básicos e os trabalhos relacionados que marcam o estado da arte deste trabalho. Apresentou-se primeiramente a definição de CBIR, mostrando-se a arquitetura típica desses sistemas e alguns exemplos de sistemas CBIR existentes. Posteriormente, introduziu-se a área de Visualização de Informação, mostrando os principais passos do processo de visualização de informação, tanto em aplicações genéricas como aplicado a sistemas CBIR. Depois, introduziram-se algumas abordagens para agrupamento de dados. Em seguida, estudaram-se as técnicas de visualização de resultados existentes, categorizadas em três grupos. Finalmente, a última parte do capítulo aborda conceitos de Interação humano-computador (IHC), com foco na etapa de avaliação. O próximo capítulo apresenta um conjunto de métodos propostos baseados no estudo das limitações dos métodos existentes e em sugestões de usuários que colaboraram com testes experimentais desenvolvidos em etapas iniciais desta pesquisa.

Capítulo 3

Técnicas de Visualização Propostas

Dadas as técnicas estudadas na seção 2.5, analisaram-se as vantagens e desvantagens dos grupos propostos para, finalmente, propor uma técnica de visualização que supra algumas das características ausentes dos métodos achados na literatura.

A Tabela 3.1 resume os três grupos de técnicas estudados na seção 2.5. Como resultado desse estudo, apresenta-se uma nova técnica denominada Clustering Set, cujo objetivo é aproveitar as vantagens dos diversos métodos na procura de uma técnica mais robusta para visualização em sistemas CBIR. Após analisar os resultados do primeiro experimento, a técnica Clustering Set apresentou resultados promissores (seção 4.2) e os usuários avaliadores recomendaram a criação de uma técnica híbrida, a partir do método proposto (Clustering Set) em conjunto com outro método que aproveite melhor o espaço disponível. Assim, decidiu-se juntar as vantagens do Clustering Set com outros métodos estudados. Escolheram-se dois métodos com enfoque central, devido à fácil adoção deles na estrutura proposta e porque esses métodos demonstraram superioridade e boa aceitação por parte dos usuários, como será mostrado no Capítulo 4. Dessa maneira, foram desenvolvidos dois métodos híbridos denominados Clustering Set - Espiral e Clustering Set - Anéis Concêntricos. A seguir, essas três técnicas propostas serão descritas.

Técnica	Referência	Vantagem	Desvantagem
Técnicas de exibição simples	[6, 7, 41]	<ul style="list-style-type: none"> • Fácil implementação. • Familiaridade dos usuários. 	<ul style="list-style-type: none"> • Visão do usuário reduzida. • Não considera relações intrínsecas entre imagens.
Técnicas com enfoque central	[5, 11, 34, 60]	<ul style="list-style-type: none"> • Distribuição estruturada do universo de imagens. • Navegabilidade e exploração mais confortáveis dado o conjunto estruturado de imagens. 	<ul style="list-style-type: none"> • Imagens similares não são exibidas em lugares próximos.
Técnicas baseadas em agrupamento	[10, 32, 34]	<ul style="list-style-type: none"> • Distribuição coerente do conjunto de imagens na qual imagens similares situam-se próximas. • Melhor entendimento do conjunto de imagens como um todo e não só em relação à imagem de consulta. 	<ul style="list-style-type: none"> • Sobreposição (na maior parte dos casos).
Clustering Set: Proposta baseada nas abordagens centrais e de agrupamento.	Original	<ul style="list-style-type: none"> • Distribuição estruturada do conjunto de imagens com técnicas de enfoque central. • Distribuição coerente do conjunto mantendo relações intrínsecas entre as imagens com técnicas de agrupamento. • Maior entendimento e facilidade de navegabilidade e exploração do conjunto de imagens. 	<ul style="list-style-type: none"> • Não adequada para mostrar a uma ordenação sequencial do conjunto de imagens.

Tabela 3.1: Abordagens de visualização.

3.1 Clustering Set

Como foi mostrado na seção 2.5, cada técnica tem vantagens e desvantagens. Por um lado, as técnicas com enfoque no centro oferecem uma boa navegabilidade e exploração de uma maneira organizada sobre o conjunto de imagens. Contudo, essas técnicas geralmente não mantêm imagens similares próximas umas das outras exceto às imagens próximas à imagem de consulta. Por outro lado, técnicas baseadas em agrupamento provêem uma distribuição coerente de imagens, exibindo imagens similares próximas umas das outras. Este último conjunto de técnicas usualmente foca-se em uma grande quantidade de informação sem, em muitos casos, tomar cuidado de prover um *design* adequado para distribuir os grupos de imagens; o que normalmente causa confusão para os usuários.

O objetivo do método proposto é juntar os aspectos positivos dessas duas abordagens. A nova técnica distribui as imagens ao redor do centro enquanto preserva relações intrínsecas entre imagens similares. Oferece a exploração confortável e a boa navegabilidade das técnicas com enfoque no centro, combinada com a informação de similaridade visual entre as imagens oferecida pelas técnicas baseadas em *clustering*.

Clustering Set (Figura 3.1) é uma contribuição original deste trabalho. Esta técnica distribui imagens similares em conjuntos de imagens seguindo um enfoque central procurando evitar a sobreposição.

Resumidamente a técnica consiste em colocar a imagem de consulta no centro da interface e distribuir agrupamentos contendo as imagens mais similares em torno da imagem central, em forma circular. Da mesma maneira, cada agrupamento é representado como um conjunto de imagens distribuídas sobre um círculo não visível. No centro do círculo situa-se a imagem mais representativa de cada agrupamento. Essa estrutura pode ser representada por uma árvore, cuja raiz seria a imagem de consulta, as imagens representativas dos grupos seriam os filhos da raiz e as imagens dentro dos grupos seriam os filhos destas imagens representativas. Este método é flexível e diversos algoritmos de *clustering* podem ser usados na determinação dos clusters (grupos).

Em resumo, os passos seguidos pelo algoritmo são:

1. A imagem de consulta é colocada no centro do *display*.
2. Um algoritmo de *clustering* é usado para agrupar as imagens considerando a similaridade quanto ao conteúdo visual. Cada imagem é representada por um vetor de características. A distância entre duas imagens é dada por uma medida de similaridade entre estes vetores de características. O algoritmo de agrupamento utiliza estas distâncias para formar os grupos.
3. Deve-se escolher uma imagem representativa para cada agrupamento. Dependendo

do algoritmo de agrupamento, a imagem representativa de cada cluster pode ser o elemento mais próximo ao centroide (k-means), o elemento medoide do *cluster* (k-medoids), etc. No caso de k-means, um centroide é representado como um vetor de características. Cada dimensão do centroide foi computada como a média da respectiva dimensão do vetor de características de cada imagem do cluster. Para escolher a imagem representativa em k-means, se computa a distância de todas as imagens ao centroide. Escolhe-se aquela que minimiza esta distância.

4. Posteriormente, ordenam-se os elementos representativos de cada cluster por similaridade com a imagem de consulta. Elementos menos distantes à imagem de consulta são considerados mais relevantes. Isto fornece a ordem de relevância dos clusters para posterior visualização.
5. Uma vez obtida a ordem dos clusters, as imagens representativas são colocadas ao redor da imagem de consulta seguindo uma trajetória circular de mais relevante a menos relevante. Cada imagem representativa é associada com a imagem de consulta por meio de uma aresta. Imagens representativas mais relevantes terão uma aresta maior com cor mais chamativa.
6. A seguir, ordenam-se os elementos dentro de cada grupo ao respeito do elemento representativo do mesmo jeito em que se ordenaram as imagens representativas ao respeito da imagem de consulta.
7. Uma vez ordenadas, as imagens de cada *cluster* são colocadas ao redor da respectiva imagem representativa seguindo também uma trajetória circular. Estas diminuem de tamanho a medida que são menos relevantes.
8. As imagens dentro de cada grupo possuem uma cor de borda que varia do amarelo claro ao verde escuro. Conforme as imagens ficam menos relevantes, o tamanho é menor e a cor maenos chamativa.
9. Visualmente os grupos também diminuem de tamanho conforme a relevância com respeito do padrão de consulta diminui.

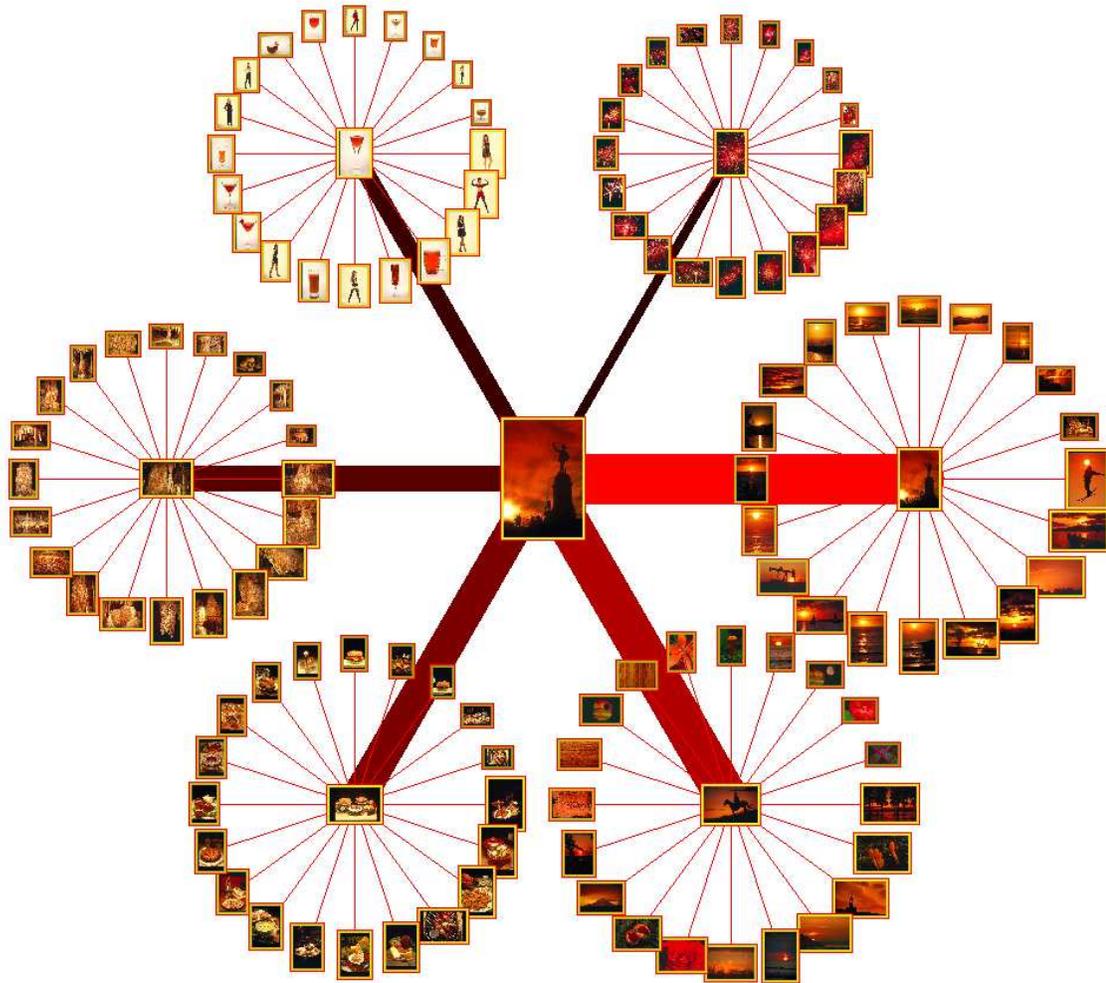


Figura 3.1: Estrutura Clustering Set

Cabe destacar que foram adicionadas transformações da visão para permitir navegar pelo conjunto de imagens. Investigações de local foram usadas para ampliar a quantidade de informação oferecida ao se aproximar de uma imagem. Nesse caso, exibe-se uma janela contendo a imagem em tamanho real, o identificador e o nome da imagem, como mostrado na Figura 3.2. No caso em que o tamanho da imagem seja grande demais, redimensiona-se a imagem para ser visualizada por inteiro no *display*.

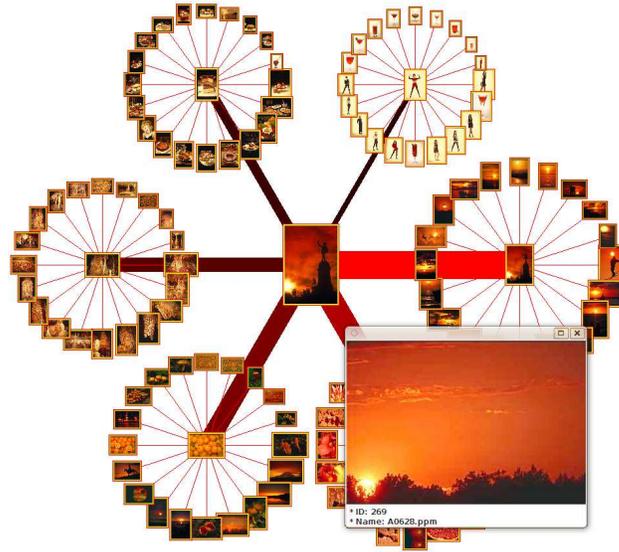


Figura 3.2: Investigações de local

Maiores detalhes são descritos a seguir.

3.1.1 *Design* do espaço

Em um sistema CBIR, após a realização de uma consulta a partir de um padrão de consulta (neste caso, uma imagem ou desenho do usuário), mostra-se ao usuário um conjunto de imagens resultantes. Deseja-se fornecer então, um espaço coerente contendo uma estrutura visual entendível e explorável que suporte o conjunto de imagens devolvidas. Para facilitar a exploração, é importante permitir ao usuário a interação com o conjunto de imagens possibilitando a criação de novas visões a partir das estruturas visuais apresentadas. Além disso, o usuário pode ter a possibilidade de refinar a sua consulta mediante a escolha de alguma imagem de interesse ao navegar através do universo de imagens. Assim, um dos desafios desta pesquisa é achar uma boa representação deste espaço. Deve-se tentar maximizar a quantidade de informação útil fornecida ao usuário no menor tempo possível.

Considere um elemento e_i da estrutura visual associado a um ponto p_i de coordenadas (x_i, y_i) com $x_i, y_i \in \mathbb{Z}$. Como o método é baseado em formas circulares, define-se um espaço em coordenadas polares. A base da estrutura é uma circunferência descrita por raio $r = a$ onde a é uma constante. Sabemos que a circunferência possui 2π radianos ou 360° ; assim, dependendo do número de elementos m a serem colocados na circunferência, o ângulo φ entre dois elementos contíguos, p_1 e p_2 , será dado por $\varphi = 2\pi/m$. Assim, podemos observar que cada elemento representado na estrutura é equidistante dos elementos contíguos a ele. O comprimento do arco definido entre dois elementos contíguos será dado por $\varphi \times r$

Assim, obtem-se uma estrutura visual cujos m elementos podem ser células com imagens ou com outros elementos. A Figura 3.3 mostra um exemplo desta estrutura para $m = 15$ onde os elementos finais são imagens.

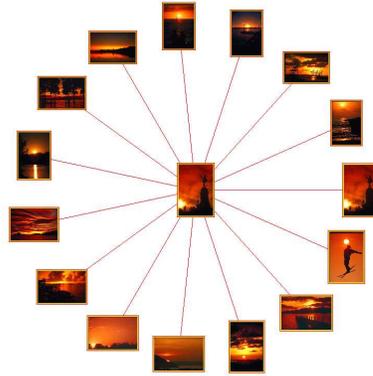


Figura 3.3: Exemplo de uma estrutura circular com $m = 15$.

Os elementos dessa estrutura, além de serem imagens, também podem ser novas instâncias da mesma estrutura visual. A seguir, a Figura 3.4 apresenta um exemplo desta estrutura para $m = 6$ no primeiro nível e para $m = 20$ no segundo nível.

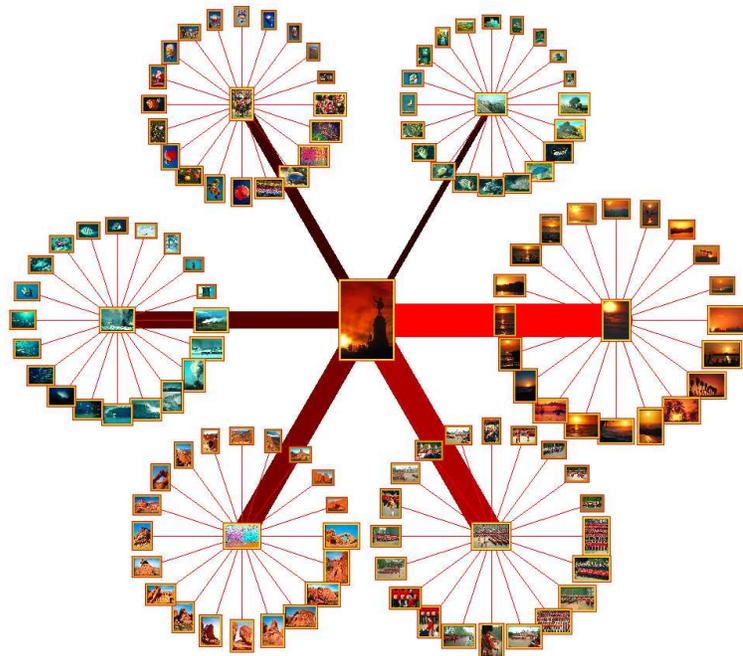


Figura 3.4: Exemplo de uma estrutura composta com $m = 6$ no primeiro nível e $m = 20$ no segundo nível.

Estas novas instâncias das estruturas circulares, no segundo nível, podem ser substituídas por diversas estruturas (tal como esta estrutura ou como as estudadas anteriormente). Este fato permite uma grande flexibilidade e a possibilidade de gerar estruturas híbridas.

As Figuras 3.5 e 3.6 apresentam alguns exemplos desta substituição no segundo nível.

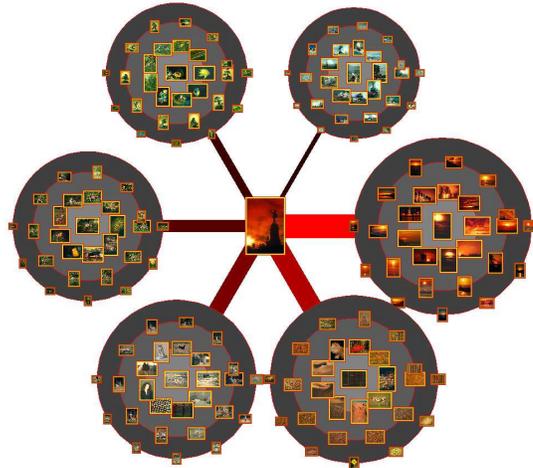


Figura 3.5: Exemplo de uma estrutura composta usando Anéis concêntricos.

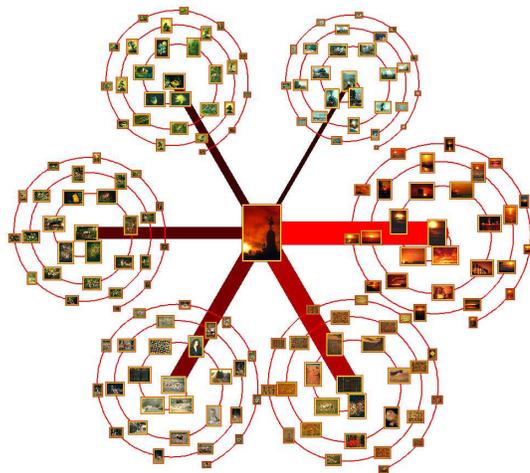


Figura 3.6: Exemplo de uma estrutura composta usando Espirais.

Esta representação pode ser estendida a qualquer quantidade de níveis e o arcabouço desenvolvido para implementar as técnicas de visualização permite criar esse tipo de estruturas compostas. Nas estruturas testadas, somente consideram-se dois níveis porque um terceiro mostraria informação quase imperceptível para o olho humano. Porém, a funcionalidade de

zoom foi implementada para todas as estruturas e experimentos futuros poderão explorar estruturas compostas a vários níveis. Além disso, foi desenvolvida uma outra funcionalidade, que será detalhada mais adiante (seção 3.1.3), para refinar a consulta, a qual incrementa as possibilidades de exploração do conjunto de imagens fornecendo outras visões da estrutura visual.

3.1.2 Uso de agrupamento

Para oferecer ao usuário uma estrutura coerente, além de distribuir os dados em uma estrutura organizada, é importante, preservar as relações intrínsecas entre as imagens do conjunto. Assim, usaram-se técnicas de agrupamento de dados.

Introduz-se aqui a notação usada para explicar com mais detalhes a técnica proposta:

- k é o número de grupos usados na técnica de agrupamento.
- c é um grupo
- $C = \{c_1, c_2, \dots, c_k\}$ é um conjunto de k grupos onde $c_i \in C$, $i = 1, 2, \dots, k$
- x é um elemento do conjunto de dados.
- $X = \{x_1, x_2, \dots, x_n\}$ é o conjunto dos n elementos a serem agrupados.
- x_{c_i} é um elemento $x \in c_i$
- s_{c_i} é o elemento representativo do grupo c_i

Inicialmente, têm-se n elementos a serem agrupados. Nesse ponto, é possível usar qualquer técnica de agrupamento (o que fornece uma abordagem flexível) com enfoque particional que possa produzir um número dado de k grupos. A técnica escolhida de agrupamento deve fornecer também a possibilidade de encontrar um elemento representativo dentro de cada grupo (não necessariamente deve ser parte do conjunto de dados). Desse modo, a comparação entre *clusters* se reduz à comparação entre seus elementos representativos. A redução no tempo de processamento é importante porque o usuário lidará com o tempo de resposta do sistema. Se este tempo for superior a 0,1 segundos, será preciso fornecer algum sinal que informe que o sistema esta processando a consulta. [2].

Podem ser usadas diversas abordagens, como as descritas na seção 2.4, que fornecem as características desejadas, como k-means ou k-medoids. Assim, nesta implementação, usa-se k-means [56] devido a alguns fatores como a sua complexidade computacional $O(nkt)$ (descrita

na seção 2.4.2) e ao baixo espaço ocupado em disco, comparado com outras técnicas usadas em abordagens hierárquicas.

Um algoritmo de agrupamento é um método de classificação não supervisionada, no qual cada elemento x é classificado como pertencente a um grupo c . Mas, no contexto CBIR, deseja-se, além disso, obter os grupos mais parecidos com um padrão de consulta dado. Assim, é feita uma ordenação dos elementos representativos (média dos elementos de cada grupo). O elemento representativo usado em k-means é o centroide dos elementos (média aritmética). Porém, visualmente precisa-se de uma imagem representativa existente no grupo. Assim, usa-se como imagem representativa aquela imagem dentro do grupo que seja mais próxima do centroide.

Como foi visto, o tempo de processamento de k-means é diretamente proporcional ao número de clusters, k ; portanto, deve-se procurar um valor de k adequado. Um k muito pequeno fará com que o tempo de processamento seja curto mas, em contraste, particionará o conjunto de imagens em poucos agrupamentos, o que ocasionará que as imagens mais similares à consulta do usuário provavelmente estejam somente dentro de um ou dois grupos e a maior parte do espaço disponibilize imagens irrelevantes para o usuário. Se k for muito grande, o tempo de processamento aumentará, porém os clusters serão mais específicos e somente imagens muito próximas farão parte de um grupo. Cabe destacar que dentro de cada agrupamento, são mostradas l imagens. A primeira imagem mostrada é a imagem representativa e as restantes $l - 1$ imagens são as mais próximas ao elemento representativo (centroide, medoide, etc) onde l é um parâmetro que deve ser escolhido levando-se em conta diversos fatores. Um l muito pequeno não mostrará informação suficiente e possivelmente muito espaço fique sem uso. Por outro lado, um l muito grande ocasionará sobreposição de imagens. Na Figura 3.1, por exemplo, considerou-se $l = 21$. Nas Figuras 3.5 e 3.6 considerou-se $l = 28$, dado que essas estruturas permitem um melhor aproveitamento do espaço.

Visualmente, como foi descrito na seção anterior, serão distribuídos 6 grupos no espaço disponível. Porém esses 6 grupos devem representar grupos de interesse para o usuário. Portanto, calculam-se k grupos e os 6 grupos mais próximos à consulta do usuário são mostrados no *display* de saída. As Figuras 3.7, 3.8 e 3.9 mostram o impacto da escolha de k sobre o conjunto resultante. Pode-se notar que os grupos da Figura 3.8 são mais interessantes e valiosos para o usuário que os que aparecem na Figura 3.7 devido ao maior grau de detalhe sobre imagens de interesse. Porém a Figura 3.9 mostra grupos muito específicos e alguns agrupamentos têm poucos elementos, restando espaço livre. Pal et al. [83] procuraram valores para um k adequado e propõem uma aproximação de $k = \sqrt{n}$. Sobre o conjunto de 3.906 imagens usadas, segundo esta formula, k seria 62,5. Este valor confirma a análise anterior na qual $k = 60$ (3.8) fornece uma boa visualização do conjunto de imagens ao usuário.

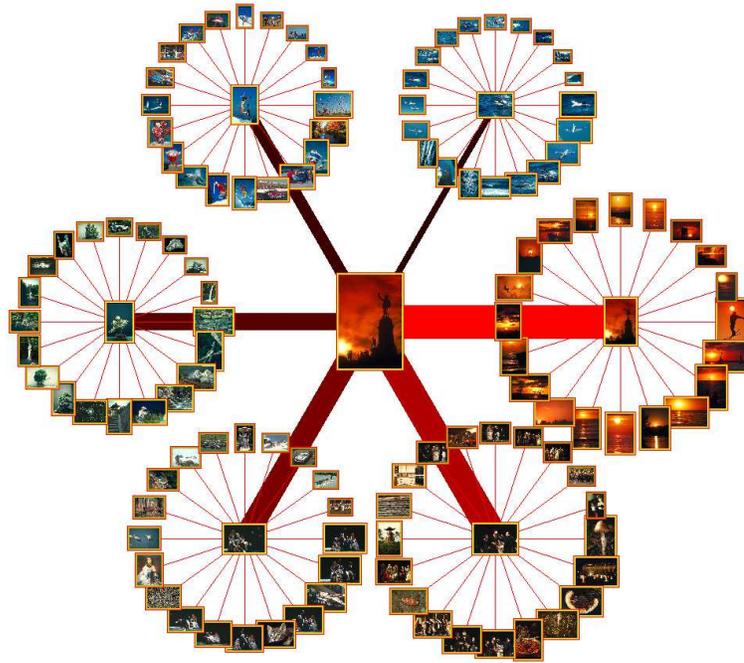


Figura 3.7: Estrutura para $k = 6$.

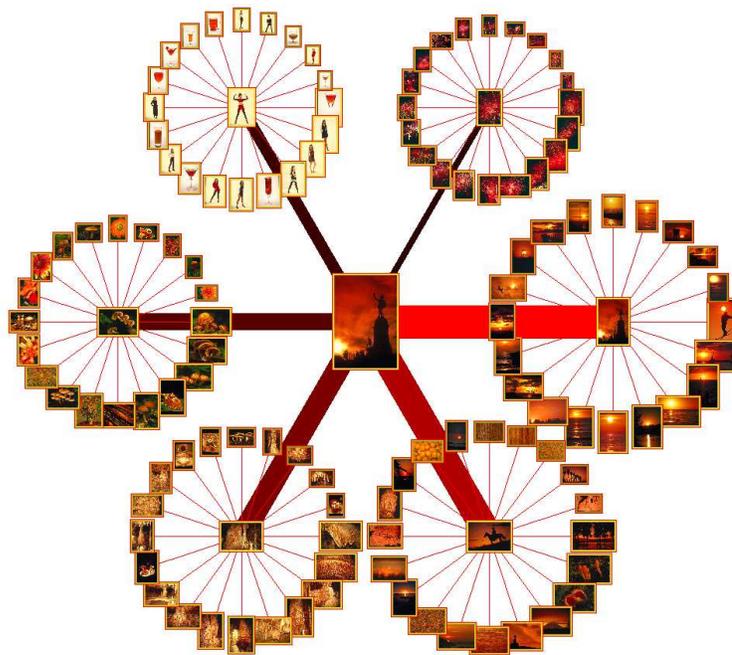


Figura 3.8: Estrutura para $k = 60$.

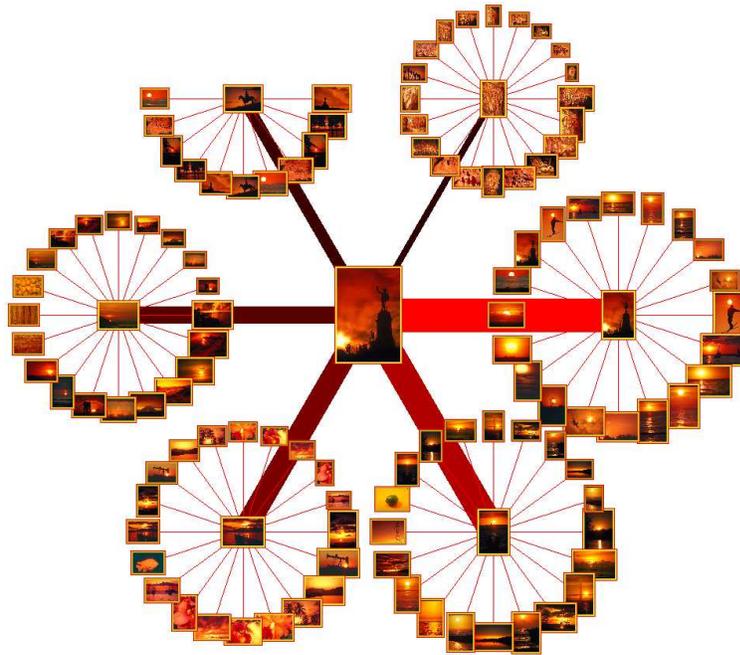


Figura 3.9: Estrutura para $k = 500$.

O conjunto inicial de centroides é estabelecido aleatoriamente, com exceção de um deles: a imagem de consulta. Nas sucessivas iterações do algoritmo de agrupamento os centroides vão mudando com exceção da imagem de consulta que fica fixada como centroide. Isto garante que ao final do algoritmo a imagem de consulta seja um dos centroides e as imagens mais próximas a ela sejam mostradas com ela. De outro jeito a imagem de consulta poderia ficar longe dos centroides e o resultado mostrado poderia ser menos preciso.

3.1.3 Agrupando por hierarquia

Após obter as estruturas vistas até o momento, pode-se selecionar uma imagem de interesse dentro de um grupo. Isso repetirá o processo e uma nova instância da estrutura visual será mostrada. Nesta segunda instância, a imagem de consulta será a imagem selecionada e o algoritmo de agrupamento usará somente os elementos existentes dentro do grupo ao qual pertence a imagem selecionada. Isto permitirá ao usuário refinar a sua consulta e explorar os grupos de imagens. As Figuras 3.10 e 3.11 ilustram este processo.

No primeiro nível, usa-se o algoritmo de agrupamento com $k = 60$ por obter melhor resultado, conforme explicado anteriormente. Com isso, são devolvidos 60 grupos, dos quais os 6 mais próximos à consulta do usuário são representados no espaço, como foi explicado na seção anterior. Diferentemente ao primeiro nível, no segundo nível, particiona-se o espaço em 6 partes ($k = 6$) dado que essa é a totalidade de elementos mostrados no espaço. Isso foi feito

porque cada novo nível origina-se da interação com o usuário e é justamente o usuário que refina a sua consulta e escolhe um grupo. Portanto, as imagens dentro do grupo escolhido pelo usuário são de interesse dele. Podem-se estender vários níveis até o usuário ficar satisfeito, ou até que a estrutura convirja quando os elementos de interesse existentes já foram mostrados.

A Figura 3.10 representa o processo de agrupamento hierárquico como uma árvore, onde cada nó representa um grupo (caixa) e cada linha horizontal representa um nível da hierarquia. O grupo de cor cinza indica que foi escolhida uma imagem dentro daquele grupo. As setas que saem dela indicam que uma nova estrutura com $k = 6$ foi criada a partir das imagens contidas no grupo.

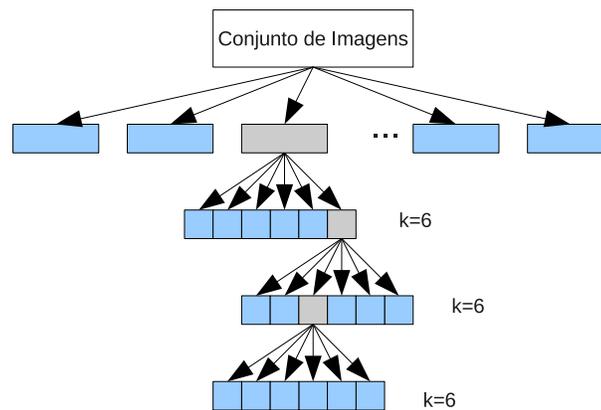


Figura 3.10: Agrupamento hierárquico.

A Figura 3.11 apresenta brevemente o processo de um sistema CBIR com a utilização de abordagens de agrupamento usadas neste trabalho. Primeiro, o usuário realiza uma consulta. Em seguida, características de baixo nível são extraídas das imagens com ajuda de descritores. Posteriormente, realiza-se um processo de agrupamento sobre o conjunto de vetores de características extraídos e um resultado é apresentado ao usuário com ajuda de alguma técnica descrita neste capítulo. Se o usuário não fica satisfeito com o conjunto devolvido, o processo se repete e uma segunda etapa de agrupamento é desenvolvida sobre o conjunto de imagens pertencentes ao grupo que possui a imagem que o usuário selecionou como relevante. Enquanto o usuário não estiver satisfeito, este processo se repete.

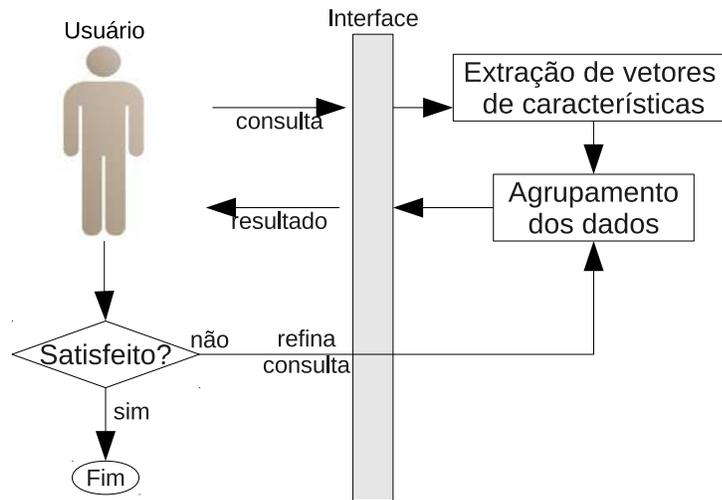


Figura 3.11: Modelo resumido do processo CBIR com uso de abordagens de agrupamento.

3.2 Clustering Set - Espiral

Esta técnica combina as técnicas Clustering Set e Espiral. Este método é similar ao método anterior quanto à posição da imagem de consulta e à disposição dos clusters. Porém desta vez, cada cluster é representado por uma estrutura espiral. O primeiro grupo é o mais relevante e portanto é desenhado com o maior tamanho e na direita da imagem de consulta. Grupos sucessivos terão conjuntos de imagens menos relevantes. Assim, o tamanho dos grupos diminui à medida que diminui a relevância de cada um. É importante destacar também a utilização de recursos visuais como cor e tamanho. Os centroides de cada grupo foram localizados no centro e têm o maior tamanho dentre as imagens do grupo. Imagens sucessivas são reduzidas em tamanho e escurecidas à medida que se afastam do centro. Existem ligações entre a imagem de consulta e o centro de cada cluster. Estas ligações são representadas por meio de arestas cuja cor varia de vermelho a preto dependendo da relevância do grupo em questão.

A Figura 3.12 mostra esta técnica.

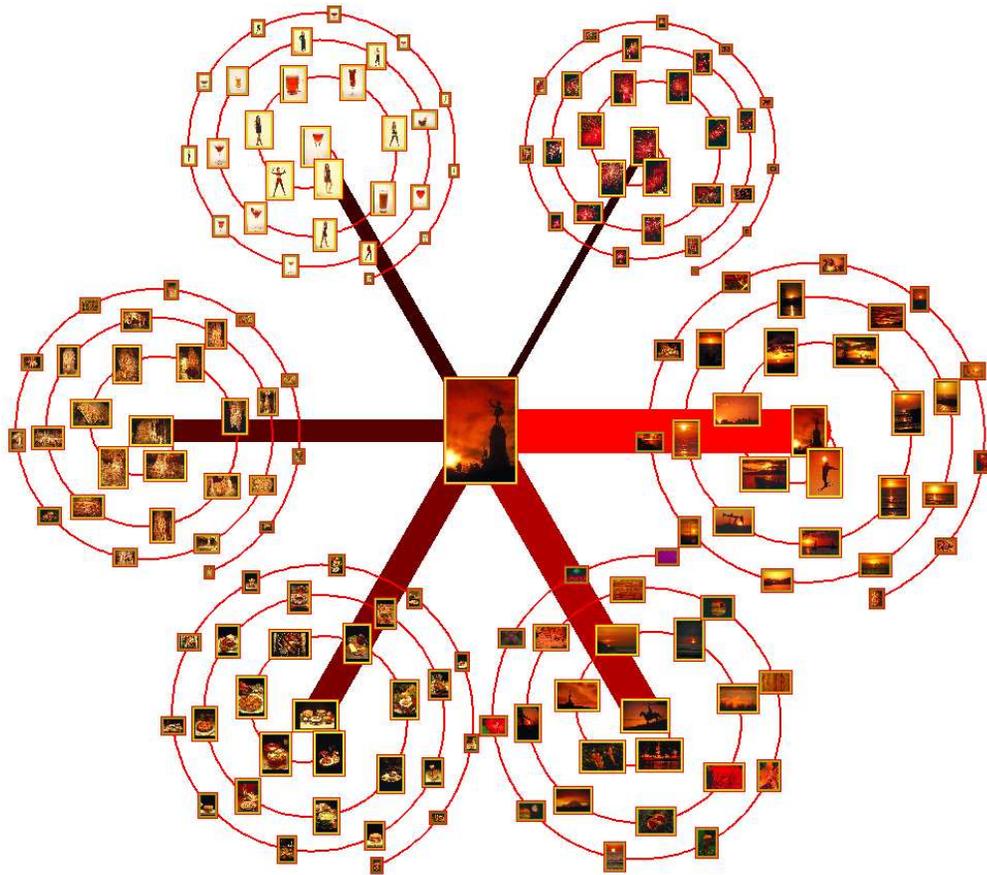


Figura 3.12: Estrutura visual Clustering Set - Espiral.

3.3 Clustering Set - Anéis Concêntricos

Este método é baseado nas mesmas ideias que os dois métodos predecessores. O método combina a técnica proposta Clustering Set com a técnica de Anéis Concêntricos. A disposição da imagem de consulta e a posição dos clusters são similares aos métodos anteriores, porém cada cluster é representado por uma estrutura de anéis concêntricos, como pode ser visto na Figura 3.13. Dentro de cada grupo, as imagens vão reduzindo de tamanho e escurecendo o tom à medida que se afastam do centro. Na estrutura geral, cada grupo também vai diminuindo de tamanho à medida que o grupo é menos relevante e a aresta que liga cada centroide com a imagem de consulta fica mais escura e estreita refletindo a relevância do grupo.

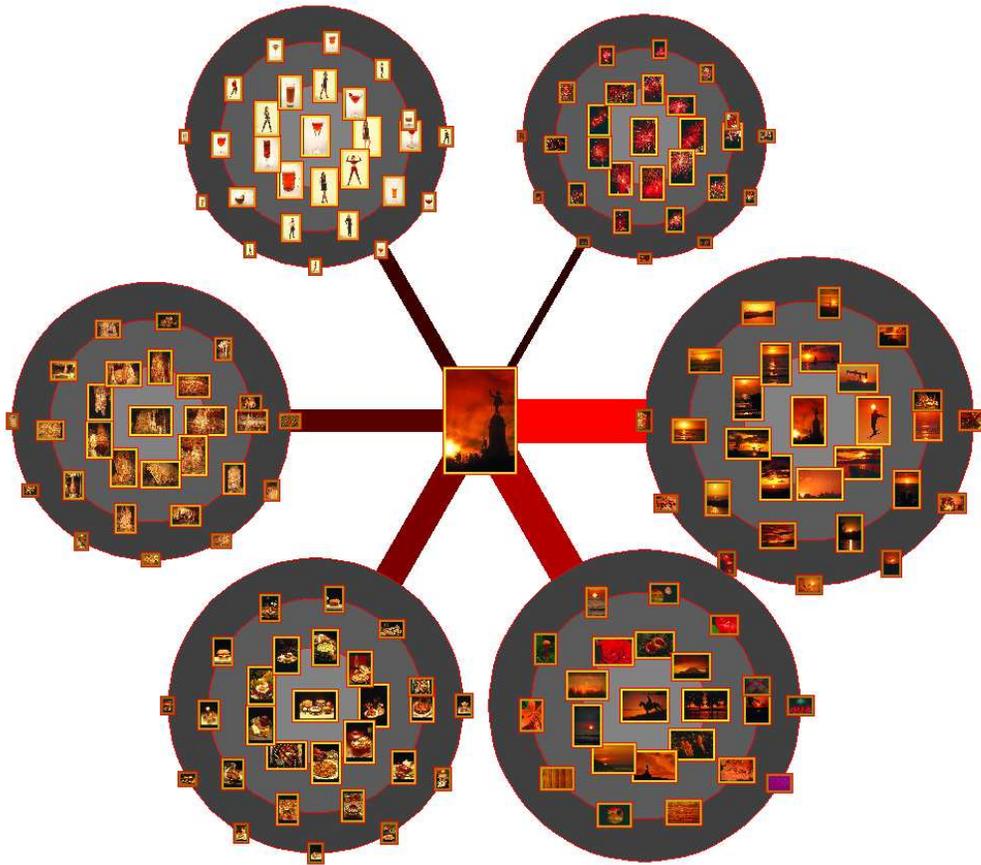


Figura 3.13: Estrutura visual Clustering Set - Anéis Concêntricos.

3.4 *Framework* para técnicas de visualização

Implementou-se um *framework* para suportar os métodos estudados e propostos. Esse *framework* é flexível para se adaptar a qualquer técnica ou combinação de técnicas nele implementadas. Considera-se que cada elemento da estrutura pode ser um elemento gráfico. Este elemento pode ser um simples nó (célula com imagem) ou uma estrutura completa anidada. Para suportar esta característica, usou-se o padrão de projeto *Composite* [84].

Devido a esta última característica, o *framework* permite a criação de estruturas fractais mediante a composição sucessiva de estruturas anidadas dentro de outras estruturas.

As Figuras 3.14, 3.15, 3.16 mostram alguns exemplos das possíveis extensões do *framework* a serem exploradas em trabalhos futuros. Nelas, os elementos encontram-se agrupados em *clusters* ordenados por relevância em relação à imagem de consulta.

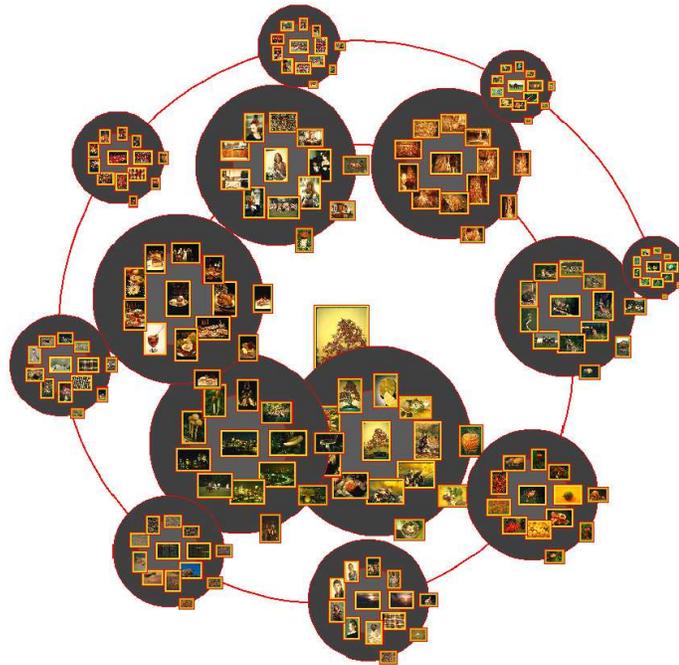


Figura 3.14: Espiral de Anéis Concêntricos.

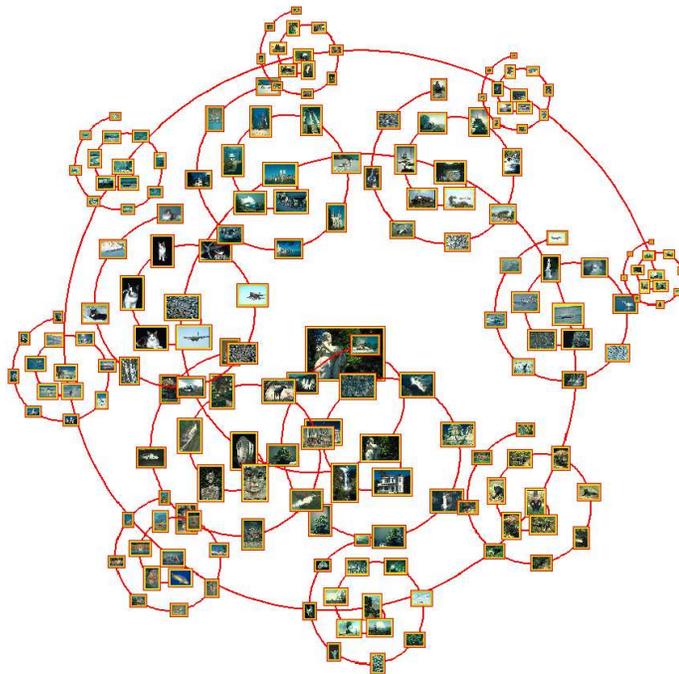


Figura 3.15: Espiral de Espirais.

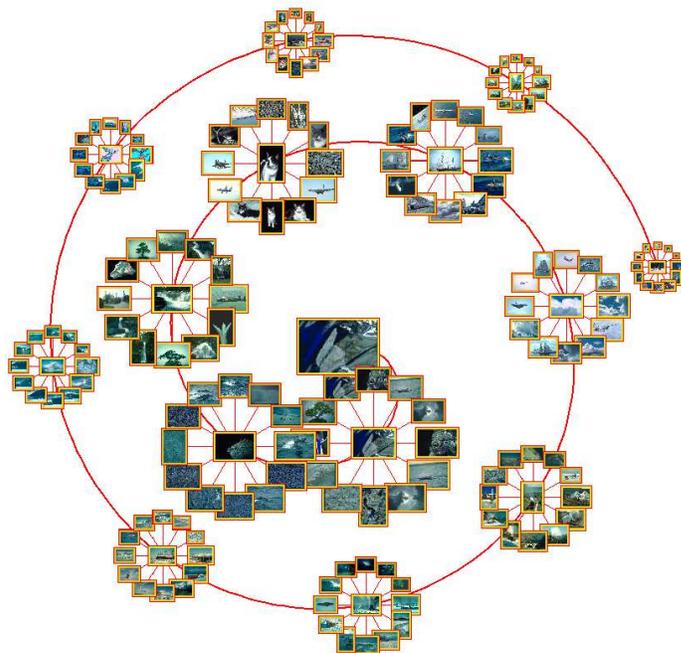


Figura 3.16: Estrutura visual Espiral de Circles.

3.5 Considerações finais

Neste capítulo, foram apresentados três métodos de visualização de resultados para sistemas CBIR baseados em clusterização que tentam suprir algumas limitações identificadas em métodos da literatura. O próximo capítulo apresenta a avaliação experimental dos métodos estudados da literatura e dos métodos propostos.

Capítulo 4

Avaliação

4.1 Considerações iniciais

Dado que é o usuário o que finalmente interage com qualquer sistema, é primordial fazer experimentos com usuários para avaliar um sistema. A aceitação ou rejeição do sistema pelos usuários determinará o seu sucesso. Assim, para avaliar as técnicas de visualização estudadas, foram executados três testes durante o desenvolvimento do projeto com o intuito de guiar o desenvolvimento da pesquisa e descobrir mais informação sobre métodos de visualização para CBIR desde etapas iniciais. Para dirigir o desenvolvimento dos experimentos usou-se o framework DECIDE apresentado na seção 2.6.

4.2 Primeiro experimento

Como foi visto na seção 2.6, a avaliação deve ocorrer durante o ciclo de vida do *design* e seus resultados devem ser utilizados para melhorias gradativas da interface. Assim, este experimento foi realizado após uma primeira etapa do projeto com o intuito de guiar seus próximos passos. Foram avaliadas cinco técnicas de visualização estudadas nesta primeira etapa, visando descobrir técnicas promissoras para as quais dirigir os próximos esforços. Quatro delas foram técnicas existentes na literatura (duas delas nunca antes avaliadas) e uma delas foi proposta após um estudo inicial das vantagens e desvantagens das técnicas estudadas. As técnicas avaliadas neste experimento foram:

- Grid;
- Anéis Concêntricos;
- Espiral;

- Incremental Board;
- Clustering Set.

4.2.1 Projeto do experimento

Nos experimentos realizados, planejaram-se cuidadosamente os passos a serem seguidos conforme o framework DECIDE. Estes passos são descritos a seguir:

1. Determinar os objetivos

- Quais são os objetivos de alto nível da avaliação?
 - Comparar os métodos estudados (os quais seguem diferentes abordagens) para identificar vantagens e desvantagens de cada um, possibilitando levantar aspectos-chaves a serem considerados em passos futuros.
 - Identificar o grau de satisfação dos usuários.
 - Identificar o grau de entendimento das estruturas visuais e a distribuição das imagens.
 - Determinar a facilidade para achar imagens nas estruturas segundo os usuários.
 - Solicitar sugestões para melhorar as estruturas avaliadas e para desenvolver eventuais estruturas novas.
 - Avaliar dois métodos da literatura nunca antes testados (Espiral e Anéis Concêntricos).
 - Avaliar a técnica proposta de visualização e consequentemente validar a hipótese de que a combinação das técnicas com enfoque central e técnicas de agrupamento oferece uma estrutura visual mais eficiente que mistura as vantagens destas duas abordagens.
- Quem quer a avaliação e por quê?

Este estudo objetiva determinar características que poderiam ser úteis para qualquer pessoa interessada em projetar um sistema CBIR. Acredita-se que este trabalho constitui uma contribuição valiosa para o desenvolvimento de futuros sistemas CBIR.

2. Explorar as questões específicas a serem respondidas

Dados os objetivos anteriores, exploraram-se um conjunto de questões que ajudaram a cumprir estes objetivos. Como o objetivo principal é a comparação de métodos,

repetiu-se um subconjunto de perguntas para cada método apresentado. As perguntas exploradas foram:

- Qual é o grau de satisfação frente a cada estrutura?
- Qual é o grau de entendimento frente a cada estrutura ?
- Quão fácil é localizar a imagem de consulta, as imagens mais relevantes e as imagens menos relevantes?
- Qual estrutura é mais adequada para visualizar os resultados de um sistema CBIR?
- Como as estruturas propostas poderiam ser melhoradas?
- Como novas estruturas poderiam ser criadas?

3. Escolher o paradigma de avaliação e as técnicas

De acordo com o padrão ISO 9341, usabilidade é a extensão para a qual um produto pode ser usado por usuários específicos para conseguir objetivos específicos com efetividade, eficiência e satisfação no contexto específico de uso. Neste caso, deseja-se avaliar critérios de usabilidade no contexto de visualização para sistemas CBIR portanto usaram-se testes de Usabilidade.

Nos testes de usabilidade usam-se técnicas de questionamento. As técnicas de questionamento consistem em fazer perguntas ao usuário para descobrir se o sistema está de acordo com as suas necessidades e possivelmente revelar dados interessantes que não teriam sido considerados pelo *designer*. As principais técnicas de questionamento são entrevistas e questionários [85]. As entrevistas e os questionários são técnicas estabelecidas usadas com diversos fins como avaliação ou captura de requisitos de software. Neste caso, em testes de usabilidade e estudos correlatos, estas técnicas são usadas para perguntar sobre fatos, comportamento, crenças, atitudes, etc [86]. As entrevistas coletam geralmente dados qualitativos enquanto os questionários tendem a coletar dados quantitativos. Neste experimento preferiu-se o uso de questionários devido a sua capacidade de atingir um grande número de pessoas o que se reflete em significância estatística como será visto nos resultados.

Existem vários questionários de avaliação de usabilidade conhecidos, como o Quis [87], o SUMI [88], entre outros. Cada questionário é útil para propósitos diferentes. Dado que cada sistema tem um contexto específico e situações particulares, cada um requer uma maneira diferente de avaliação. Elaboraram-se dois questionários que ajudassem ao processo de avaliação, os quais serão descritos no próximo passo.

Maiores detalhes serão tratados na seção 4.2.2.

4. Identificar as questões práticas

Participantes

Para que os resultados fossem mais significativos, procurou-se um conjunto de colaboradores que tenham conhecimento prévio sobre aspectos da visualização em interfaces. Com tal fim, desenvolveu-se a avaliação com a colaboração de alunos voluntários de três disciplinas oferecidas no Instituto de Computação da Unicamp. Duas destas disciplinas estavam relacionadas à área de IHC (Interação Humano-Computador) e a outra (amostra menor de usuários) também tinha conhecimento do tema em menor proporção. A Tabela 4.1 apresenta o número de usuários por cada uma das três disciplinas.

Disciplina	Nome	Tamanho/amostra
MC750	Construção de Interfaces Homem-Computador	23
MO825B	Semiótica da Interação Humano-Artefato Digital	11
MO812R	Bibliotecas Digitais	4
TOTAL		38

Tabela 4.1: Quadro de usuários participantes.

MC750: Este foi o grupo mais numeroso. Dos 23 alunos que participaram no experimento só 2 deles pertencem à pós-graduação. Quase todos eles são estudantes dos últimos anos de graduação no Instituto de Computação da UNICAMP.

MO825B: Dos 11 alunos que participaram no experimento, 4 deles ainda pertencem a graduação. Os demais alunos estão no programa de pós-graduação.

MO812R: Grupo pequeno de alunos da pós-graduação.

Equipamento

Cada experimento foi realizado em distintos laboratórios do Instituto de Computação da UNICAMP dependendo da disponibilidade e do horário em que foi feito. Os laboratórios estão equipados com 30 computadores em média. Assim, cada participante teve acesso a um computador e teve a liberdade para escolher o sistema operacional a usar.

Restrições de tempo e orçamento

O experimento durou aproximadamente 20 minutos mas não houve restrição de tempo e cada usuário foi livre para acabar o experimento quando achou conveniente. Alguns usuários ficaram mais de 20 minutos preenchendo cuidadosamente as últimas questões que eram perguntas abertas sobre sugestões para a pesquisa em desenvolvimento. Os usuários aceitaram participar voluntariamente sem receber pagamento.

Material

Foi redigido um conjunto de documentos tanto informativos quanto de avaliação. Na elaboração do formulário de avaliação, tentou-se dirimir ao máximo as dúvidas dos avaliadores sobre as técnicas em questão. Assim, as perguntas respondidas pelos usuários forneceriam a informação precisa para os avaliadores.

Antes de começar a avaliação, cada participante recebeu um conjunto de documentos. Estes documentos são:

- Instruções aos usuários: Detalharam-se os passos do experimento e as principais questões éticas.
- Termo de Consentimento Livre e Esclarecido (TCLE): Trata-se de um documento obrigatório em toda pesquisa que envolve usuários. Nele contemplam-se as principais questões do experimento. Esse documento é assinado pelo usuário em sinal de conformidade com as condições do experimento.
- Formulário de Perfil de usuário: Permite capturar o perfil do usuário no contexto requerido. Desejava-se saber a expertise do usuário sobre critérios específicos. Portanto, além de dados básicos como escolaridade e sexo, perguntou-se pela familiaridade e frequência com que costumam usar o computador, realizar buscas, realizar buscas de imagens e buscas de imagens por conteúdo (CBIR). Os resultados desses formulários são mostrados na seção 4.2.3. As questões foram formuladas segundo a escala de Likert [89] mediante a qual o usuário tem a possibilidade de avaliar cada pergunta com um número de 1 até 5, sendo 5, o critério mais alto e 1, o mais baixo para cada questão.
- Formulário de Avaliação: É o documento mais importante do experimento. Nele o usuário emitiu as suas opiniões sobre cada estrutura apresentada. Este questionário avaliou 5 critérios para cada uma das 5 técnicas de visualização avaliadas. Cada critério foi refletido em uma pergunta, totalizando 25 questões. Cada uma destas questões utilizou o formato da escala Likert [89] na qual cada critério é mensurado em uma escala de 1 até 5. Finalmente, além das 25 questões, adicionaram-se mais três: Uma pergunta para ordenar as técnicas de avaliação conforme a opinião do usuário, e outras duas questões abertas nas quais os participantes puderam escrever sugestões para melhorar as técnicas avaliadas e para propor alguma eventual estrutura nova. Este tipo de questão é útil para obter informações ou sugestões, as quais seriam difíceis de conseguir com questões fechadas. As opiniões podem fornecer informação importante que o avaliador poderia não ter percebido.

Cada participante teve acesso a este conjunto de documentos impressos.

5. Decidir como lidar com questões éticas

Nos documentos de Instruções aos Usuários e no TCLE que foram entregues a cada usuário, explicaram-se claramente as questões éticas. Além disso, falou-se verbalmente sobre os principais detalhes éticos ao início de cada experimento.

6. Avaliar, interpretar e apresentar os dados

Finalmente após escolher os paradigmas e técnicas adequadas para responder as perguntas formuladas que satisfazem os objetivos estabelecidos, é preciso realizar a avaliação, apresentação dos dados e interpretá-los. Este processo é discutido com maior detalhe nas seções 4.2.2 e 4.2.3.

4.2.2 Desenvolvimento dos testes

Foi realizado o mesmo experimento em três etapas com três conjuntos de usuários distintos. O objeto a ser avaliado foi uma apresentação com um conjunto de slides, nos quais se tinha uma imagem por cada slide. Cada imagem é uma captura de tela dos métodos implementados, uma para cada estrutura visual a ser avaliada. O procedimento do experimento consistiu nas seguintes etapas:

1. Cada usuário escolheu um computador para fazer o experimento. Decidiu-se usar um computador para cada participante para dar total liberdade aos participantes para avançar e retroceder na apresentação contendo as estruturas avaliadas quantas vezes fosse preciso até que cada um tivesse certeza das respostas a serem dadas no questionário de avaliação. Além disso, o fato de ser possível navegar entre as imagens mais de uma vez permitiu aos usuários perceber detalhes de cada estrutura que possivelmente não teriam sido enxergados se as imagens tivessem sido mostradas uma vez só.
2. Em seguida, entregou-se o conjunto de documentos descritos no quarto passo do *framework* DECIDE e explicou-se verbalmente os principais detalhes do experimento, reforçando assim o documento de instruções aos usuários.
3. Posteriormente, os usuários leram e assinaram o Termo de Consentimento Livre e Esclarecido.
4. Seguidamente pediu-se o preenchimento do formulário de perfil de usuário.
5. Após isso foi feito o *download* do arquivo da apresentação localizado em um endereço indicado pelo avaliador. Dado que as diversas máquinas dos diferentes laboratórios têm uma desigualdade de software e hardware e, além disso, a preferência dos usuários pode acarretar a eleição de distintas plataformas, com distintos softwares instalados

nelas, previu-se o fornecimento do mesmo arquivo com diferentes formatos, caso não se achasse disponível o software adequado. As Figuras 2.10, 2.12, 2.13, 2.17, 4.1 foram agrupadas em um mesmo arquivo que compõe a apresentação em slides que foi avaliada.

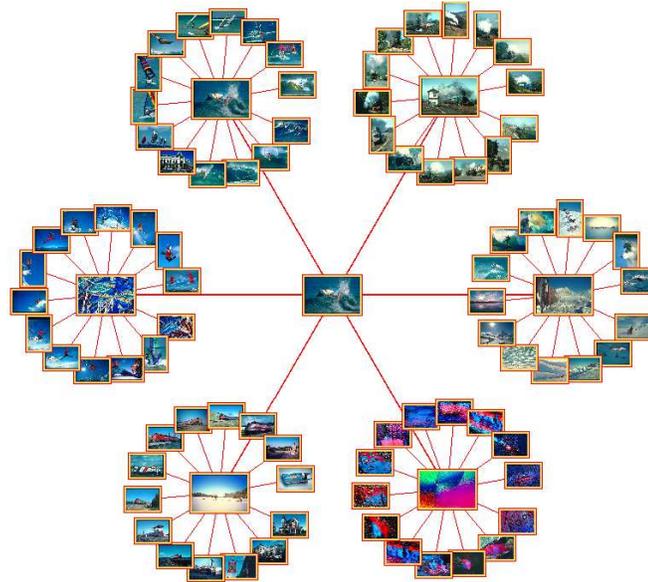


Figura 4.1: Experimento 1 - Técnica proposta Clustering Set usada no experimento.

6. Cada usuário examinou as imagens cuidadosamente ao mesmo tempo que preenchia o questionário de avaliação. Solicitou-se especial atenção a diversos detalhes como a disposição das imagens, a quantidade de informação contida na tela. Pediu-se também que se tentasse identificar a imagem de consulta, as imagens mais relevantes e as menos relevantes com o objetivo que os usuários encontrem possíveis vantagens e desvantagens de cada estrutura visual perante cada tarefa solicitada.

Dado o grande número de usuários, não foi possível usar técnicas como o *Thinking-aloud* [90], muito útil em testes de usabilidade. Mas, em contraste, foram adicionadas ao questionário de avaliação duas perguntas abertas.

Cabe destacar que não se controlou o tempo dispensado por cada participante na avaliação.

As Figuras 4.2, 4.3, 4.4, 4.5, 4.6 e 4.7 apresentam os usuários em uma sessão de avaliação.



Figura 4.2: Experimento 1 - Desenvolvimento do teste com usuários.



Figura 4.3: Experimento 1 - Teste com usuários - Grid.



Figura 4.4: Experimento 1 - Teste com usuários - Anéis Concêntricos.



Figura 4.5: Experimento 1 - Teste com usuários - Espiral.

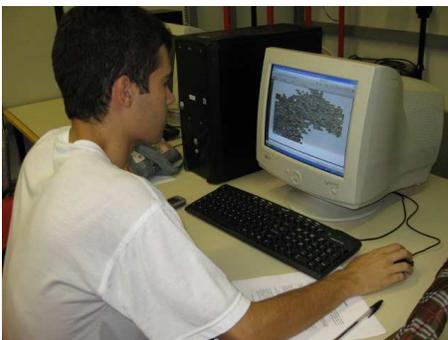


Figura 4.6: Experimento 1 - Teste com usuários - Incremental Board.

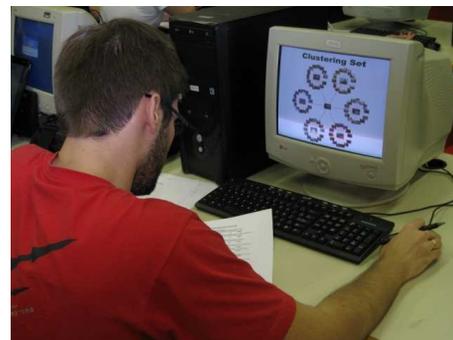


Figura 4.7: Experimento 1 - Teste com usuários - Clustering Set.

4.2.3 Resultados

Após ter avaliado as três turmas, a informação obtida foi processada e os resultados obtidos são descritos a seguir.

1. Perfil de Usuário

Como foi mencionado, o experimento foi feito com a colaboração dos alunos de três disciplinas. Os participantes foram divididos por faixa etária, sendo que 76,3% do total de usuários encontraram-se na faixa etária entre 20 a 25 anos, e o restante, 23,7% de usuários, teve entre 26 e 35 anos. Como foi dito na seção 4, todos os participantes cursam ensino superior, dos quais 68,4% cursam estudos de pós-graduação e 31,6% estudam graduação. A experiência dos participantes foi estabelecida de acordo com a familiaridade e frequência no desenvolvimento de tarefas descritas nas seções a seguir. Nestas seções será mostrado um conjunto de gráficos que sintetizam o perfil dos usuários dividido por turmas e consolidado.

Perfil de usuário por turma

As Figuras 4.8, 4.9 e 4.10 representam um resumo do perfil dos usuários por turma. Cada barra representa um critério do perfil (descritos na legenda) e o tamanho da barra representa a média das respostas para cada critério do conjunto de usuários de cada turma.

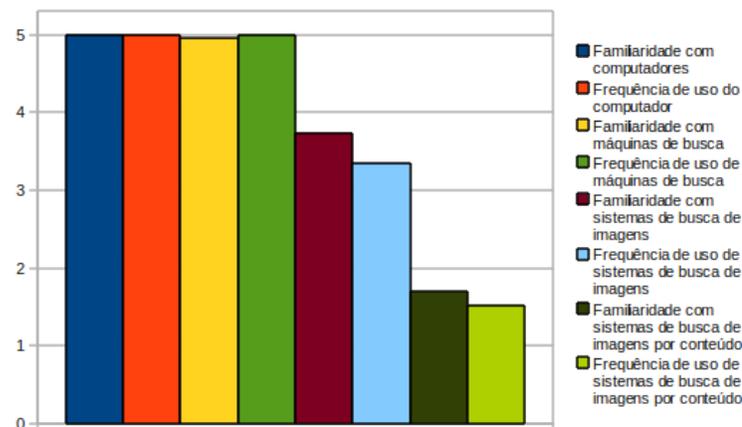


Figura 4.8: Experimento 1 - Perfil dos usuários da disciplina MC750

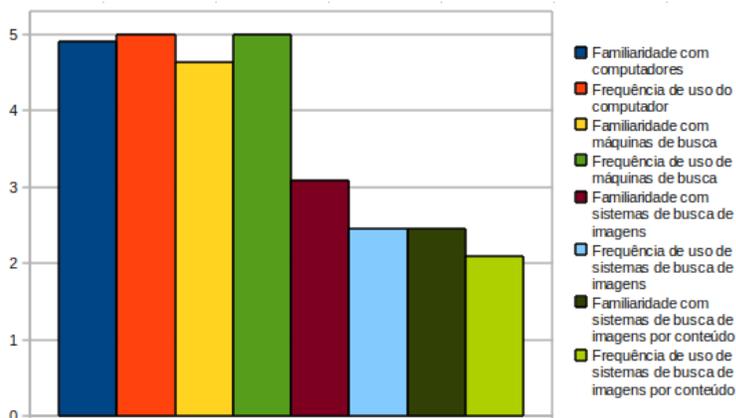


Figura 4.9: Experimento 1 - Perfil dos usuários da disciplina MO825B

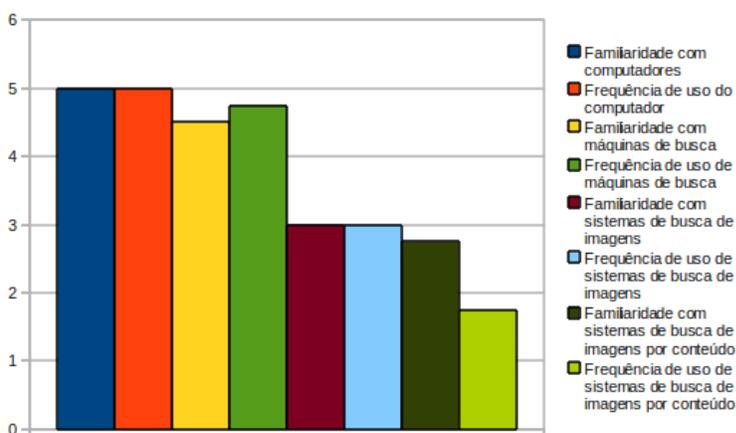


Figura 4.10: Experimento 1 - Perfil dos usuários da disciplina MO812R

Perfil de usuário consolidado

As Figuras 4.11, 4.12, 4.13, 4.14, 4.15, 4.16, 4.17 e 4.18 apresentam a avaliação feita pelo conjunto completo dos usuários para cada critério.

No eixo horizontal encontra-se a escala de avaliação (1-5) e no eixo vertical o número de usuários que avaliaram o critério em determinada escala. Se a distribuição das barras se concentra nos valores altos da escala, então os usuários consideram-se mais experientes no critério em questão.

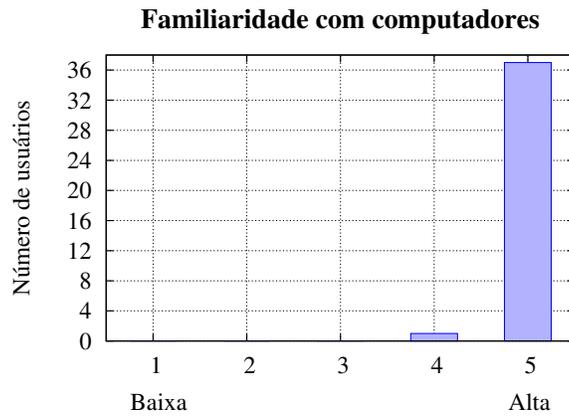


Figura 4.11: Experimento 1 - Familiaridade com computadores

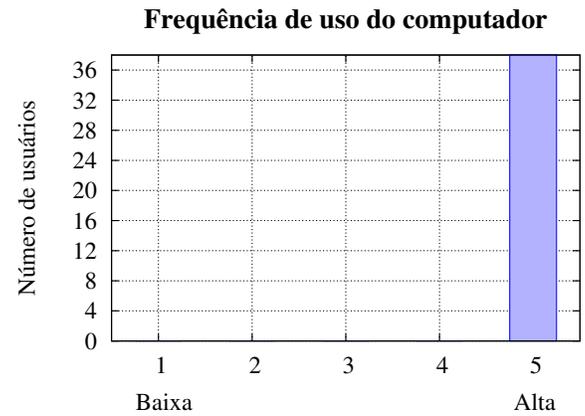


Figura 4.12: Experimento 1 - Frequência de uso do computadores

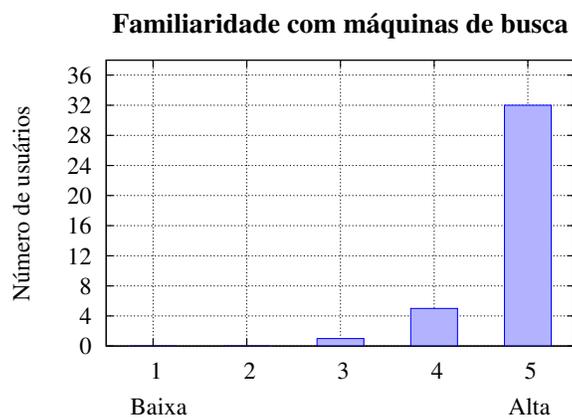


Figura 4.13: Experimento 1 - Familiaridade com máquinas de buscas Web.

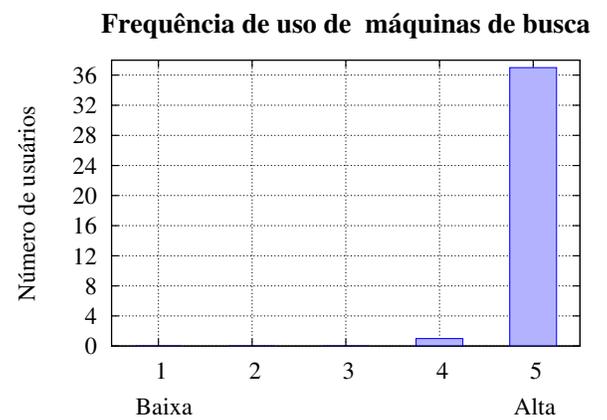


Figura 4.14: Experimento 1 - Frequência com que usa as máquinas de buscas.

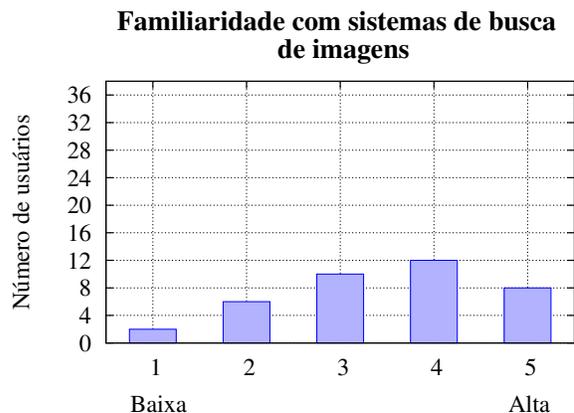


Figura 4.15: Experimento 1 - Familiaridade com sistemas de busca de imagens.

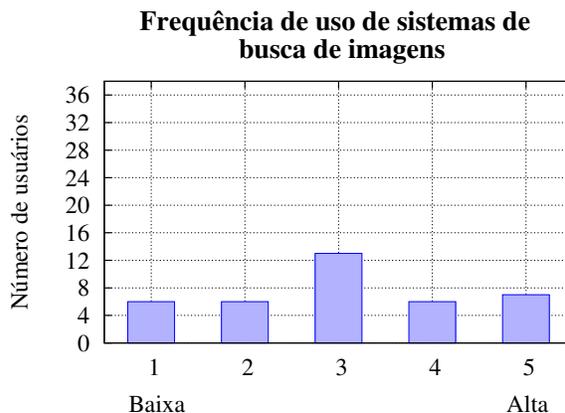


Figura 4.16: Experimento 1 - Frequência com que usa sistemas de busca de imagens.

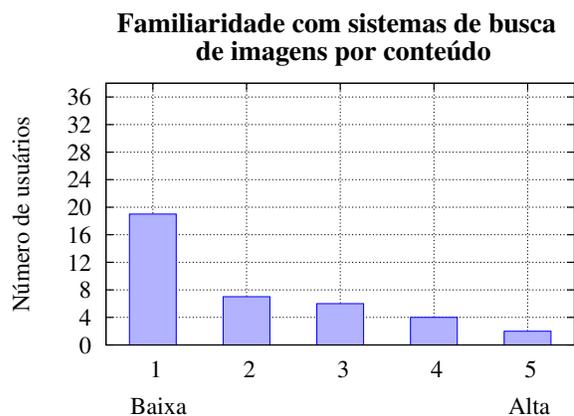


Figura 4.17: Experimento 1 - Familiaridade com sistemas de busca de imagens por conteúdo.

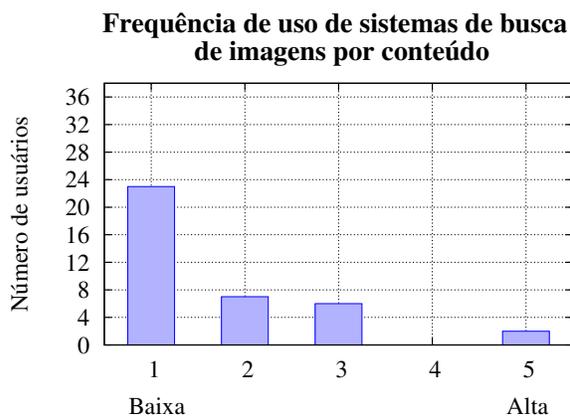


Figura 4.18: Experimento 1 - Frequência com que usa sistemas de busca de imagens por conteúdo.

A Figura 4.19 representa o perfil consolidado para os 38 usuários que participaram deste experimento. O tamanho de cada barra representa a média das respostas do conjunto de usuários para cada critério avaliado descrito na legenda.

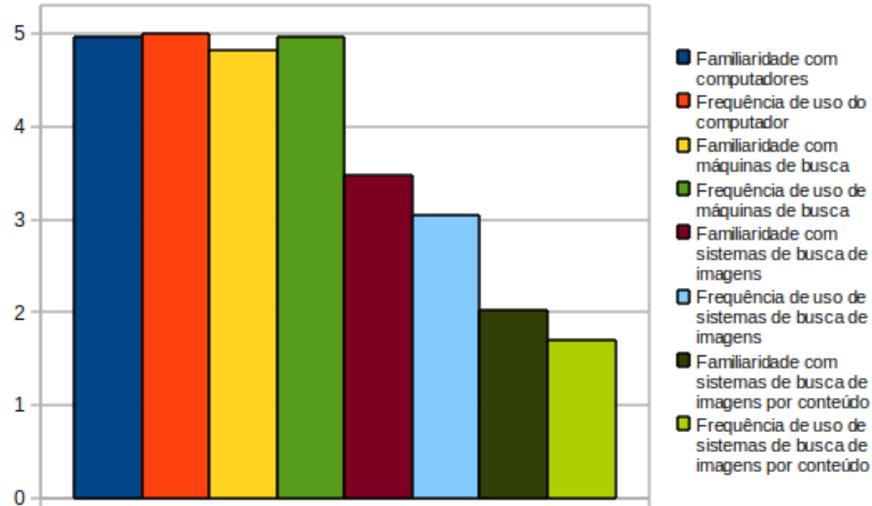


Figura 4.19: Experimento 1 - Perfil de usuário

2. Avaliação das Estruturas Visuais

Como foi mencionado, o experimento foi feito com a colaboração dos alunos de três disciplinas.

Dado um conjunto de dados x , neste caso o conjunto de respostas para uma questão (avaliação de um critério específico para uma técnica de visualização na escala de 1 a 5), x_i representa o valor i do conjunto x e pode ser tratada como uma variável aleatória. Considera-se a média amostral μ do conjunto x definida pela equação 4.1.

$$\mu = \frac{1}{n} \sum_{x=1}^n x_i \quad (4.1)$$

E o desvio padrão σ das amostras em x é definido segundo a equação 4.2:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{n - 1}} \quad (4.2)$$

Com o intuito de analisar os dados dos experimentos e avaliar as técnicas de visualização, é importante o uso de técnicas estatísticas. Como é sabido, a distribuição dos dados na natureza é aproximadamente Normal. Portanto, espera-se que a maior parte de x se encontre em um intervalo adjacente a μ . O tamanho do intervalo varia dependendo da quantidade de dados que deseja que esteja dentro dele. Se o intervalo cobrisse todos as amostras x (100% dos dados), isso poderia cobrir os *outliers* e possivelmente não ajudaria a descobrir onde se encontra a maior concentração dos dados o qual nos impediria achar resultados estatisticamente significativos. Geralmente, um intervalo

que possua 95% dos dados dentro dele é aceitável e usualmente utilizado. Para conseguir um intervalo de confiança de 95%, é preciso usar um fator z da Tabela para Distribuição Normal Padronizada. Neste caso $z = 1,96$ para um intervalo de 95% de confiança. Finalmente para achar o intervalo de confiança, achamos ω

$$\omega = z \frac{\sigma}{\sqrt{n}} \quad (4.3)$$

Logo, o intervalo que tem 95% dos dados esta entre $\mu - \omega$ e $\mu + \omega$.

O tamanho deste intervalo e o grau de sobreposição com outros intervalos podem determinar se os resultados foram estatisticamente significativos ou não.

As Tabelas 4.2, 4.3 e 4.4 apresentam os dados obtidos para cada conjunto de participantes. As colunas ($\mu \pm \omega$) representam o intervalo de confiança de cada critério sobre cada uma das estruturas avaliadas.

As Figuras 4.20, 4.21 e 4.22 mostram a média μ para cada critério avaliado, considerando cada estrutura visual.

Etapa 1 - Disciplina MC750

	Satisfação		Entendimento		Img Consulta		+ Relevantes		- Relevantes	
	$\mu \pm \omega$	σ								
Grid	3,870 ± 0,432	1,058	4,435 ± 0,386	0,945	4,217 ± 0,426	1,043	2,739 ± 0,496	1,214	2,739 ± 0,606	1,484
Anéis Conc.	2,826 ± 0,546	1,337	3,870 ± 0,466	1,140	3,783 ± 0,426	1,043	3,391 ± 0,457	1,118	3,870 ± 0,482	1,180
Espiral	2,522 ± 0,576	1,410	3,174 ± 0,546	1,337	3,348 ± 0,502	1,229	3,130 ± 0,449	1,100	3,478 ± 0,549	1,344
Inc. Board	1,826 ± 0,383	0,937	1,826 ± 0,421	1,029	1,435 ± 0,297	0,728	2,174 ± 0,546	1,337	2,273 ± 0,550	1,316
Clust. Set	4,348 ± 0,362	0,885	4,696 ± 0,228	0,559	4,522 ± 0,299	0,730	4,826 ± 0,158	0,388	4,174 ± 0,518	1,267

Tabela 4.2: Turma MC750

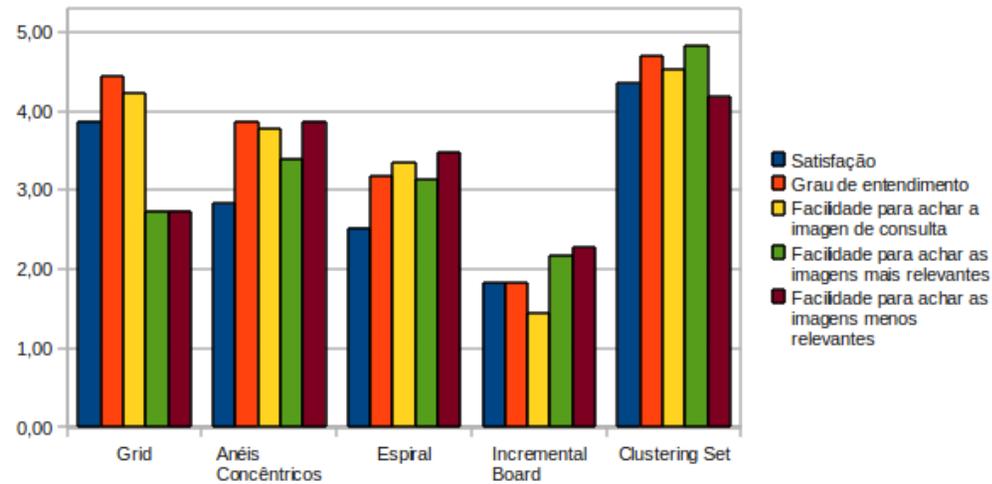


Figura 4.20: Experimento 1 - Avaliação de MC750

	Satisfação		Entendimento		Img Consulta		+ Relevantes		- Relevantes	
	$\mu \pm \omega$	σ								
Grid	3,364 ± 0,760	1,286	4,000 ± 0,748	1,265	3,455 ± 0,809	1,368	3,545 ± 0,612	1,036	3,727 ± 0,652	1,104
Anéis Conc.	3,727 ± 0,652	1,104	3,909 ± 0,768	1,300	4,455 ± 0,485	0,820	3,909 ± 0,722	1,221	4,000 ± 0,647	1,095
Espiral	3,000 ± 0,793	1,342	3,727 ± 0,752	1,272	4,182 ± 0,785	1,328	3,818 ± 0,690	1,168	3,818 ± 0,785	1,328
Inc. Board	1,818 ± 0,637	1,079	1,818 ± 0,785	1,328	1,455 ± 0,485	0,820	2,455 ± 0,929	1,572	2,636 ± 0,963	1,629
Clust. Set	4,182 ± 0,516	0,874	4,636 ± 0,398	0,674	4,364 ± 0,546	0,924	4,818 ± 0,356	0,603	4,300 ± 0,657	1,059

Tabela 4.3: Turma MO825B

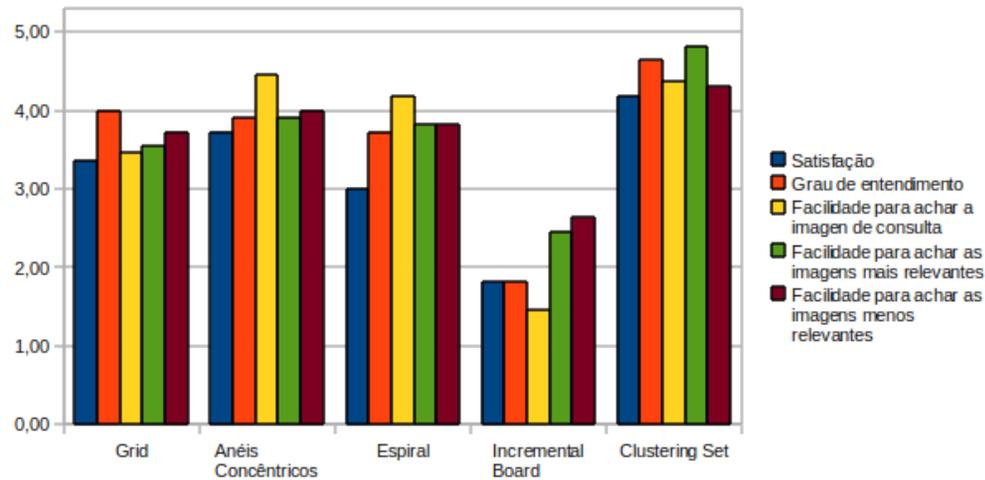


Figura 4.21: Experimento 1 - Avaliação de MO825.

Etapa 3 - Disciplina MO812R

	Satisfação		Entendimento		Img Consulta		+ Relevantes		- Relevantes	
	$\mu \pm \omega$	σ								
Grid	3,250 ± 0,938	0,957	3,750 ± 1,470	1,500	4,250 ± 1,470	1,500	4,250 ± 0,938	0,957	3,750 ± 0,938	0,957
Anéis Conc.	3,000 ± 1,386	1,414	4,250 ± 0,938	0,957	3,750 ± 1,233	1,258	4,250 ± 0,938	0,957	4,000 ± 1,132	1,155
Espiral	3,500 ± 1,877	1,915	4,750 ± 0,490	0,500	4,000 ± 1,386	1,414	3,750 ± 1,470	1,500	3,750 ± 1,470	1,500
Inc. Board	1,750 ± 0,938	0,957	2,000 ± 1,132	1,155	1,000 ± 0,000	0,000	2,500 ± 0,566	0,577	2,000 ± 1,132	1,155
Clust. Set	5,000 ± 0,000	0,000	4,500 ± 0,566	0,577	4,750 ± 0,490	0,500	4,750 ± 0,490	0,500	3,000 ± 1,600	1,633

Tabela 4.4: Turma MO812R

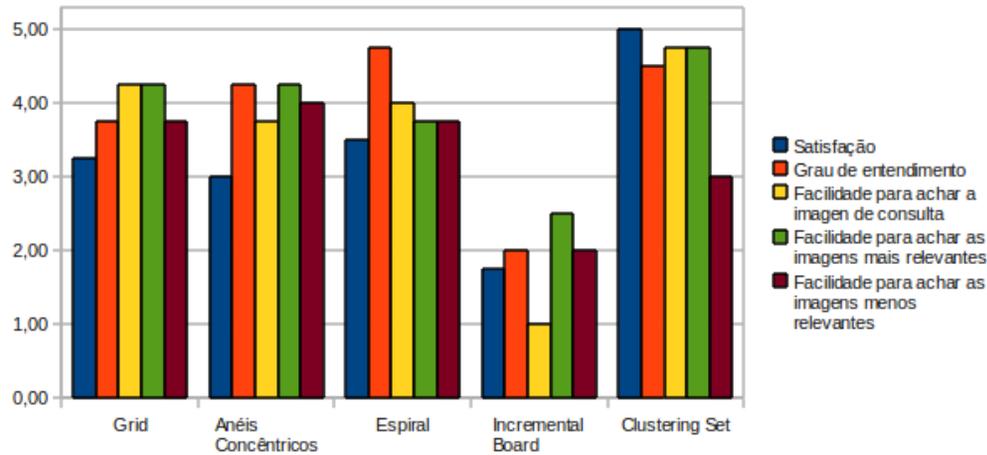


Figura 4.22: Experimento 1 - Avaliação de MO812R

3. Experimento Consolidado

Esta seção apresenta os resultados finais levando-se em conta os 38 participantes do experimento para cada critério avaliado.

As Figuras 4.23, 4.24, 4.25, 4.26 e 4.27 apresentam gráficos de intervalo de confiança onde o 95% dos dados coletados para cada critério avaliado oscila dentro de cada intervalo (linha vertical azul). O centro de cada intervalo é a média das amostras e está representado pelo símbolo \oplus . Esta média esta dada pela equação $\mu = \frac{\sum_{i=1}^n x_i}{n}$, onde n representa o número total de dados (quantidade de participantes do experimento) e x_i representa o valor entre 1 e 5 que cada usuário atribuiu a cada critério avaliado. O resultado desta fórmula, apresenta-se ao lado direito do símbolo em menção. O símbolo \otimes representa a mediana dos dados (50% dos dados são menores ou iguais a ele).

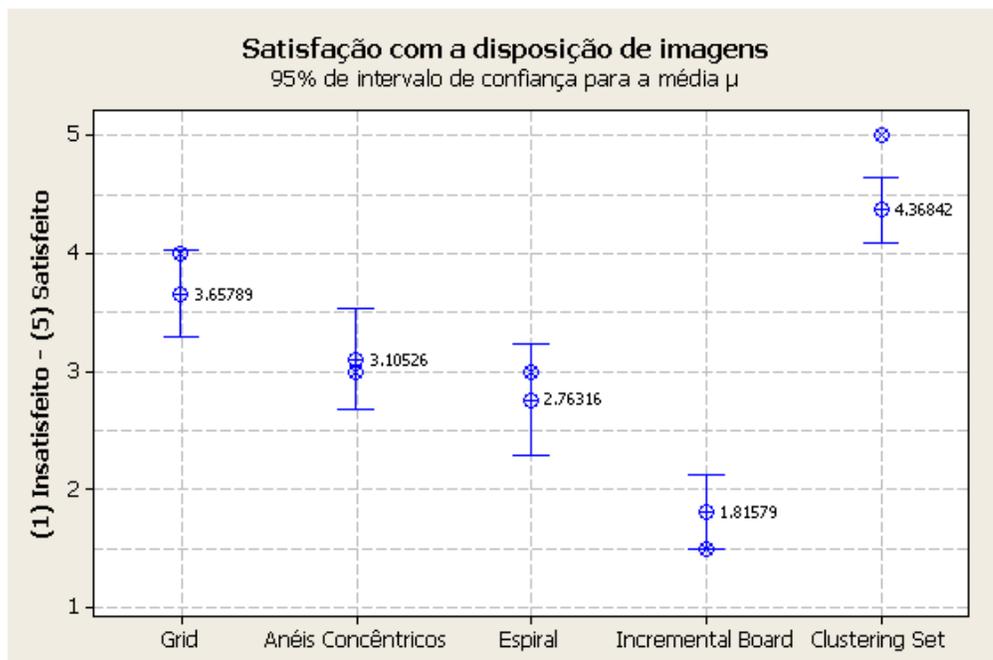


Figura 4.23: Experimento 1 - Satisfação com a disposição de imagens.

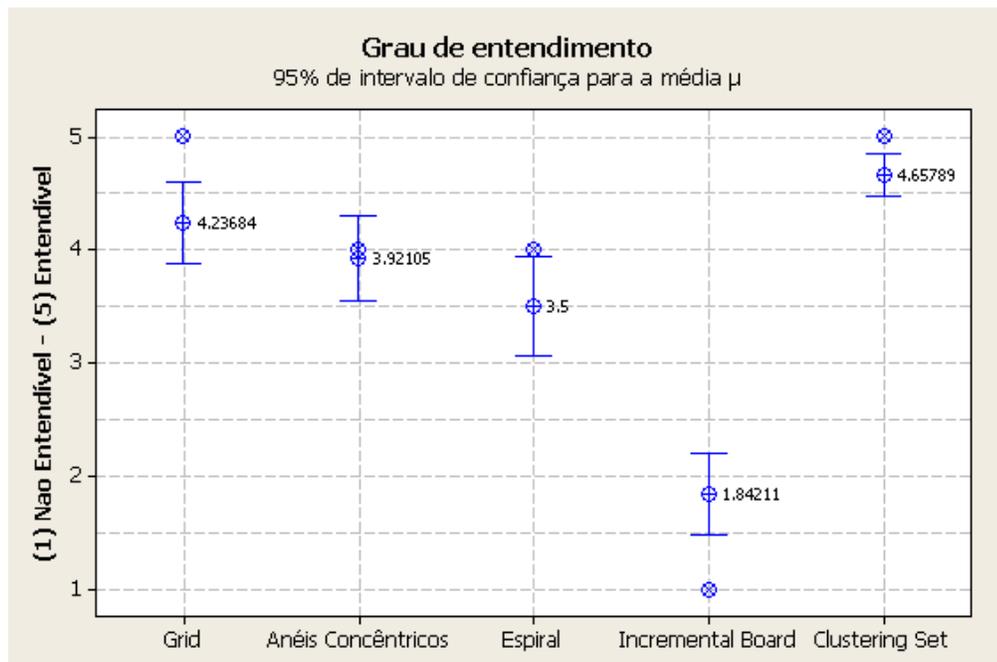


Figura 4.24: Experimento 1 - Grau de entendimento da distribuição das imagens.

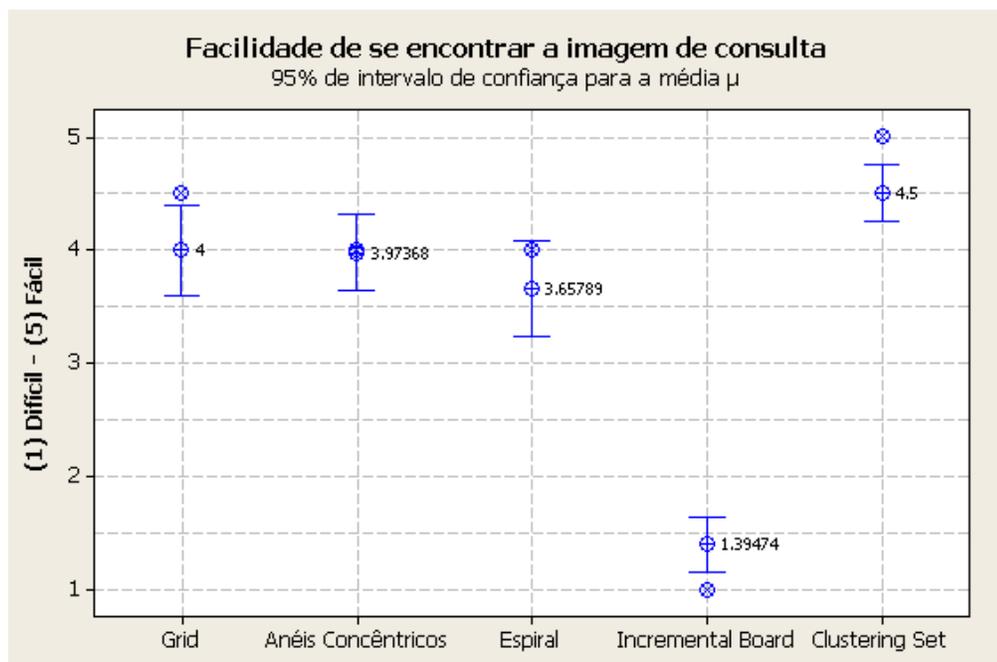


Figura 4.25: Experimento 1 - Facilidade de se encontrar a imagem de consulta.

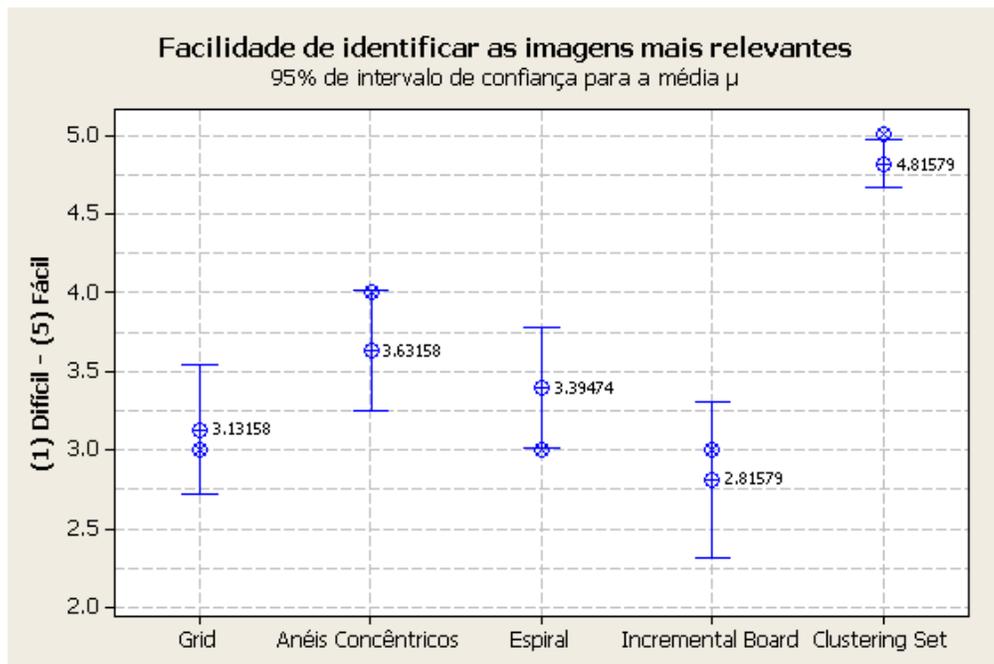


Figura 4.26: Experimento 1 - Facilidade de se encontrar as imagens mais relevantes.

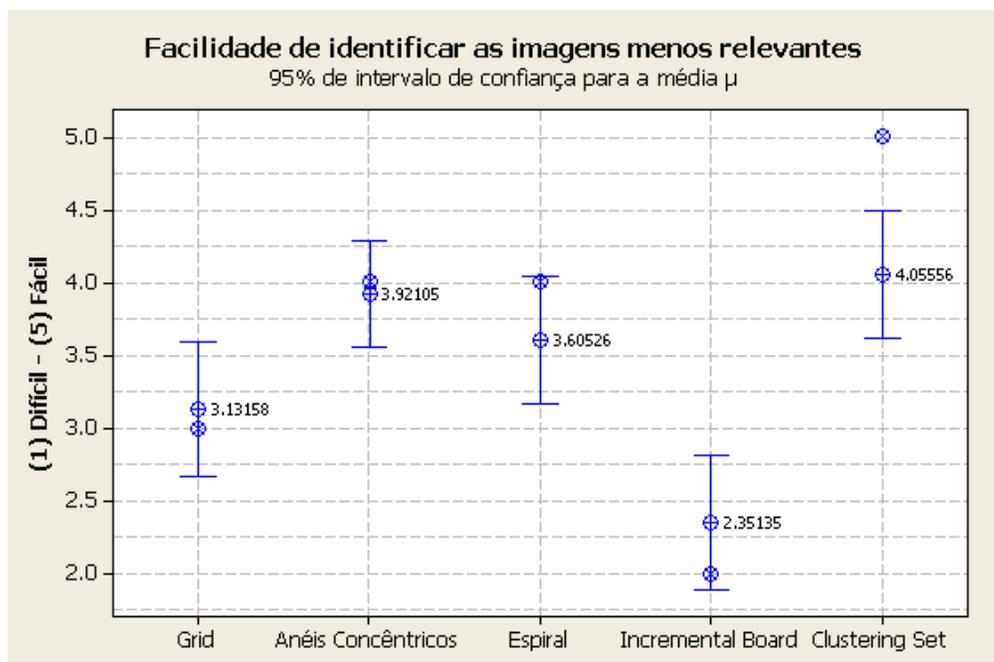


Figura 4.27: Experimento 1 - Facilidade de identificar as imagens menos relevantes.

A Figura 4.28 ilustra as estruturas avaliadas; nela, os números da escala (1-5) representam as médias dos valores alocados pelos usuários para cada critério e para cada estrutura. Essa figura mostra um resumo dos gráficos vistos anteriormente.

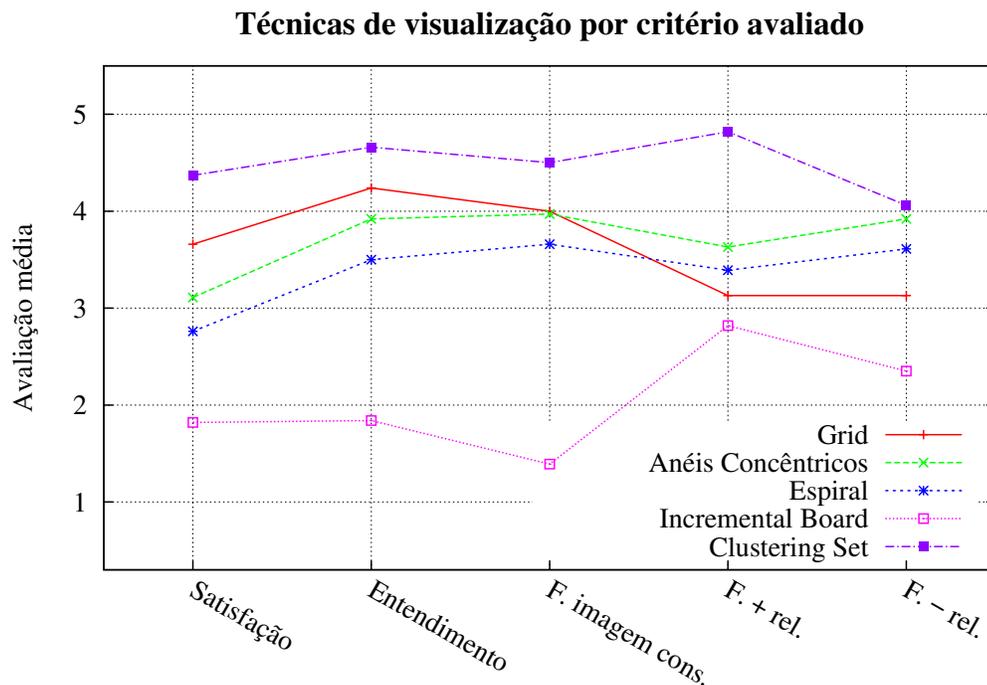


Figura 4.28: Experimento 1 - Avaliação das cinco estruturas visuais por critério.

4.2.4 Análise

Cabe destacar que a análise da maior parte de resultados foi feita em função dos gráficos de intervalos de confiança. Cabe destacar que se dois intervalos de confiança se sobrepõem, não pode se afirmar estatisticamente que a avaliação de um deles seja melhor que o outro.

- Notou-se que quase todos os participantes estão familiarizados com o uso do computador; porém, na procura de imagens, o conjunto de usuários tem menos experiência. Diante da pergunta sobre buscas CBIR responderam que têm uma familiaridade e frequência de utilização muito baixa.
- As técnicas com enfoque central (anéis concêntricos e espiral), foram em geral bem avaliadas nos critérios que implicam achar as imagens mais relevantes e menos relevantes, o que mostraria que estas técnicas fornecem uma estrutura organizada onde o usuário

é capaz de reconhecer a localização dos elementos, o que demonstra o entendimento adequado que esse tipo de estruturas provê.

- O método Incremental Board foi o mais confuso para os usuários (como é mostrado na Figura 4.28) devido principalmente à falta de uma sinalização adequada sobre a imagem de consulta e ao tamanho reduzido das imagens. Isso pode ser justificado pelo fato de que este método inicialmente foi proposto para a visualização de conjuntos de elementos e não para a visualização de elementos resultantes de uma consulta. Por tal motivo, nos experimentos seguintes foram feitas algumas modificações em cor para ressaltar relevância em relação da imagem de consulta.

Esse fato ressaltou também a necessidade da utilização da característica de zoom sobre esta estrutura visual, principalmente sobre os métodos que fornecem uma maior quantidade de imagens na tela ao mesmo tempo. A seguir são apresentados algumas opiniões dos usuários ao respeito desta técnica.

- *O Incremental Board é extremamente difícil de visualizar, talvez funcione com menor número de imagens.*
- *O Incremental Board tem muitas imagens; fica difícil encontrar o que se procura. Talvez funcione melhor com um pouco menos de imagens.*
- *No Incremental Board não se consegue ver as imagens.*

- De acordo com a Figura 4.28, o método melhor avaliado foi o Clustering Set (linha roxa) o qual obteve a ponderação média melhor para todos os critérios avaliados. Porém somente a média não representa resultados estatisticamente significativos. Assim, analisando os gráficos de intervalo de confiança, pode-se concluir que o método Clustering Set foi o que forneceu ao usuário a maior satisfação (Figura 4.23) e facilidade para achar as imagens mais relevantes (Figura 4.26). Pode-se notar também, que o Clustering Set foi o método mais entendível (Figura 4.24) em empate com o já conhecido e conseqüentemente já entendido método Grid. Finalmente, quanto a facilidade para achar a imagem de consulta (Figura 4.25) e facilidade para achar as imagens menos relevantes (Figura 4.27), o Clustering set tecnicamente empatou (sobreposição dos intervalos de confiança) na melhor avaliação junto com outros métodos. Portanto, considera-se que este método é uma importante contribuição para a visualização de resultados de um sistema de busca de imagens.
- Considerando que o método proposto Clustering Set foi bem sucedido, analisou-se a possível causa deste sucesso segundo as respostas e opiniões dos usuários. Inicialmente pensou-se que o motivo fosse o uso de algoritmos de agrupamento nessa técnica (dado

que mantem relações intrínsecas entre as imagens mostrando imagens similares próximas) mas a técnica com pior avaliação (Incremental Board) também preserva as relações intrínsecas entre as imagens. Porém, a técnica Incremental Board, não mostra separação entre grupos similares de imagens nem é uma estrutura com enfoque central. Assim, também observou-se que outras técnicas com boa aceitação foram anéis concêntricos e espiral, as quais permitem uma distribuição organizada e entendível do conjunto de imagens. Além disso, alguns usuários opinaram que a distribuição por categorias ou grupos mostrou-se útil e organizada. Portanto, pode-se concluir que a hipótese está correta (a combinação das técnicas com enfoque central e técnicas de agrupamento fornece uma estrutura visual mais eficiente que mistura as vantagens destas duas abordagens - página 50). Essa nova estrutura facilitou a cognição do conjunto de imagens e considera-se adequado e recomendável para a visualização de resultados de um sistema de busca de imagens.

- Verificou-se muito interesse dos usuários em relação ao método proposto Clustering Set. Esse método foi mencionado em grande parte das sugestões dos usuários, apontando a criação de um método mais robusto, produto da combinação do *Clustering Set* com algum outro método estudado, como Anéis Concêntricos, Espiral ou Grid. Esse interesse no Clustering Set seria outro indicativo do bom aceite do método. A seguir algumas dessas sugestões para melhorar esta técnica são apresentadas:

- *Acredito que o Clustering Set poderia ser usado em conjunto com anéis concêntricos.*
- *Usar o Clustering Set com a disposição em Grid em cada cluster para aproveitar melhor o espaço.*
- *Deve-se imitar a ideia de cluster mais com anéis.*
- *Talvez uma união do Clustering Set com Anéis Concêntricos.*
- *Utilizar uma estrutura que seja um merge entre Espiral e Clustering Set, onde fique mais claro dentro de cada conjunto similar qual é a ordem de similaridade.*
- *Talvez uma mescla entre Grid e Clustering Set.*

- As perguntas subjetivas (abertas) foram muito enriquecedoras, permitindo aos usuários expressar ideias e opiniões sobre vários aspectos como interatividade com o sistema.

- *A ideia de clusters visuais é muito interessante mas pode gerar perda de informações devido ao tamanho das imagens. Se for possível que as imagens sejam ampliadas temporariamente com um honering do mouse, ela seria bem satisfatória.*
- *O layout em espiral tende a perder seu sentido uma vez que ele vai contra a ordem natural de busca de informação na tela.*
- *Um Clustering Set com relevância de cada cluster. As imagens mais relevantes poderiam estar maiores como no clustering set.*

- Comprovou-se uma grande familiaridade dos usuários com o método Grid, provavelmente por ser o mais conhecido atualmente. Segundo o critério *grau de entendimento*, este método foi bem avaliado e foi em média o segundo melhor depois do método Clustering Set. Cabe ressaltar que, por outro lado, em relação ao critério de facilidade de se achar imagens mais relevantes e menos relevantes, este método obteve uma pontuação baixa.
- A utilização de cores no método Anéis Concêntricos ajudou no entendimento da distribuição das imagens. Sugeriu-se aplicar essa ideia no método Espiral.

Talvez uso de cores de fundo na proposta Espiral, tal como foi feito em Anéis Concêntricos, poderia ajudar a organizar a espiral.

- Nota-se que o método Clustering Set tem a menor dispersão dos dados em comparação com outras técnicas, ou seja, teve medidas de desvio padrão menores e consequentemente intervalos de confiança menores. O que significa que as opiniões dos usuários em geral concordaram ao respeito dele.
- Nota-se também que a pergunta que ocasionou maior divergência sobre os dados para todas as estruturas foi a facilidade por achar imagens menos relevantes, o que mostra a dificuldade do usuário por reconhecer imagens pouco relevantes. Uma possível solução seria o uso de marcas visuais para ajudar o processo de cognição, usando, por exemplo, cores claras sob as imagens relevantes e cores opacas sob as imagens menos relevantes.

4.3 Segundo experimento

Uma vez analisados os resultados e sugestões do primeiro experimento, observou-se que a estrutura proposta, Clustering Set, foi a melhor avaliada em todos os critérios questionados (em três dos quais obteve-se a melhor avaliação em empate com outros métodos). Observou-se também que as sugestões de diversos usuários apontaram a criação de um método híbrido que misture as vantagens de Clustering Set com outros métodos avaliados. Portanto, foram propostos dois métodos adicionais híbridos baseados nesta técnica. Estes métodos combinaram o método proposto com outros métodos que obtiveram boa avaliação no primeiro experimento, como anéis concêntricos e espiral.

Adicionou-se uma série de mudanças visuais sobre as estruturas, tentando incrementar a facilidade para entendê-las. Dentre as mudanças adicionadas destacam-se variação de cor e variação de tamanho nas células de algumas estruturas, com o objetivo de torná-las mais intuitivas ao usuário.

Para cada imagem dentro de uma estrutura visual, tem-se uma célula retangular. Esta célula é levemente maior que a imagem correspondente e possui uma cor e uma posição de acordo com o grau de similaridade da imagem correspondente à imagem de consulta e/ou às imagens semelhantes que se encontram próximas. A imagem situa-se no meio desta célula deixando aparecer uma borda colorida ao redor.

Neste experimento, com o intuito de avaliar somente as estruturas e não os algoritmos CBIR de baixo nível, as imagens foram tiradas mantendo-se a célula associada.

Para as estruturas cujas células mudaram de cor, a célula que contém a imagem de consulta manteve o tom amarelo e células consecutivas foram associadas a tons mais avermelhados e escuros, gerando uma gama de cores.

A seguir, detalham-se as estruturas avaliadas e as mudanças feitas com relação ao primeiro experimento.

- Grid: Esta estrutura foi a mesma que foi usada no primeiro experimento, mas desta vez sem imagens, como pode ser visto na Figura 4.29.
- Grid Modificado: Estrutura similar à anterior, com diferença nas cores das células, como pode ser visto na Figura 4.30.
- Anéis Concêntricos: Adicionou-se variação de cor (Figura 4.31).
- Espiral: As cores também foram modificadas (Figura 4.32).
- Incremental Board: Neste método adicionaram-se cores para conseguir distinguir a imagem de consulta e algumas imagens relevantes. (Figura 4.33).

- Clustering Set: Em relação ao primeiro experimento, alterou-se a cor das células dentro dos clusters, e incrementou-se a espessura das arestas que unem a imagem de consulta com a imagem representativa de cada grupo de acordo com a relevância dos grupos. Quanto mais relevante for um grupo (próximo à imagem de consulta), a aresta respectiva será mais espessa (ver Figura 4.34).
- Clustering Set - Espiral: Esta técnica nasceu após as sugestões do primeiro experimento na procura de um método mais robusto e útil para o usuário. As arestas conservam as modificações feitas para a estrutura Clustering Set. Cada agrupamento está representado pela estrutura de Anéis Concêntricos com modificações de cor nas células dos grupos para indicar relevância (ver Figura 4.35).
- Clustering Set - Anéis Concêntricos: Técnica híbrida proposta detalhada na seção 3.3. Similarmente ao método anterior, este método considerou as sugestões dos usuários do primeiro experimento. Cada agrupamento utiliza a estrutura Espiral com modificações de cor nas células dos grupos para indicar relevância (ver Figura 4.36).

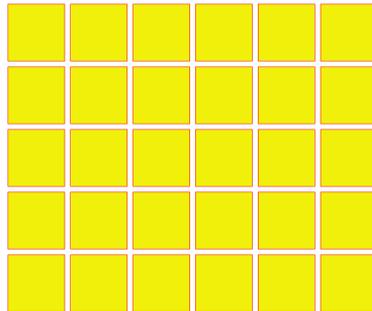


Figura 4.29: Experimento 2: Grid.

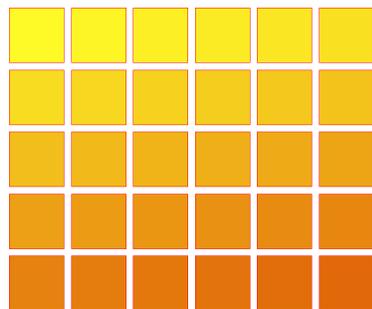


Figura 4.30: Experimento 2: Grid Modificado.

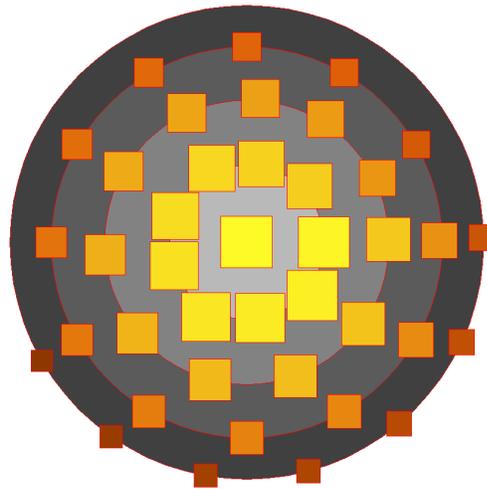


Figura 4.31: Experimento 2: Anéis Concêntricos.

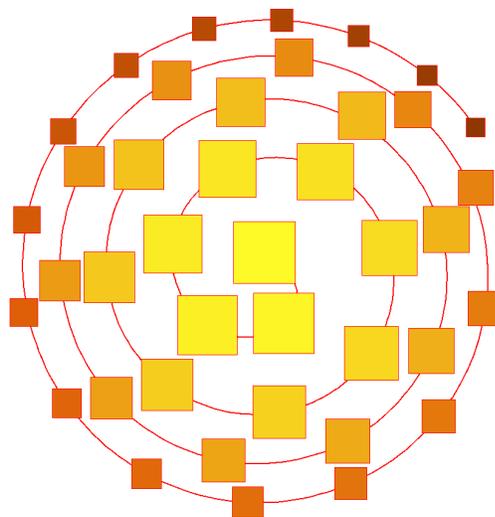


Figura 4.32: Experimento 2: Espiral.

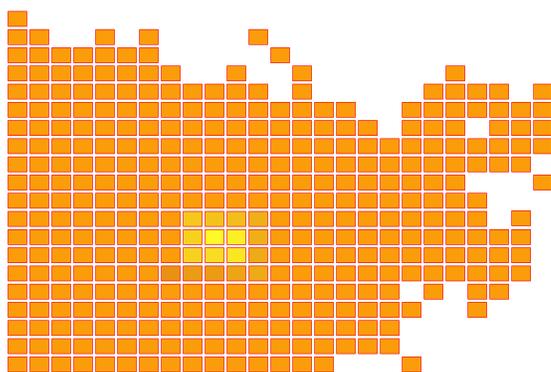


Figura 4.33: Experimento 2: Incremental Board.

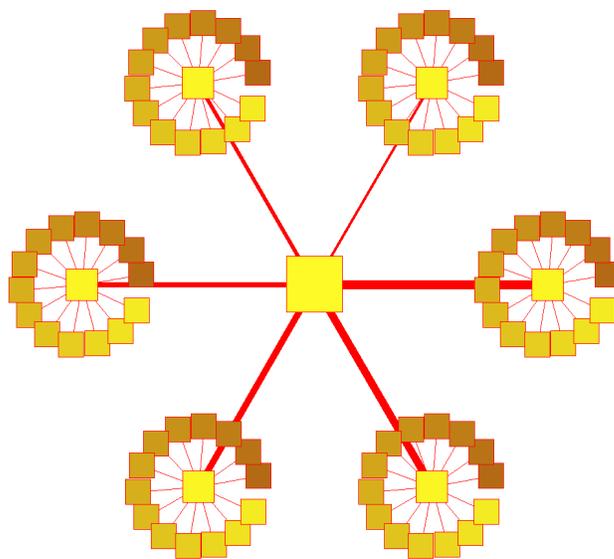


Figura 4.34: Experimento 2: Clustering Set.

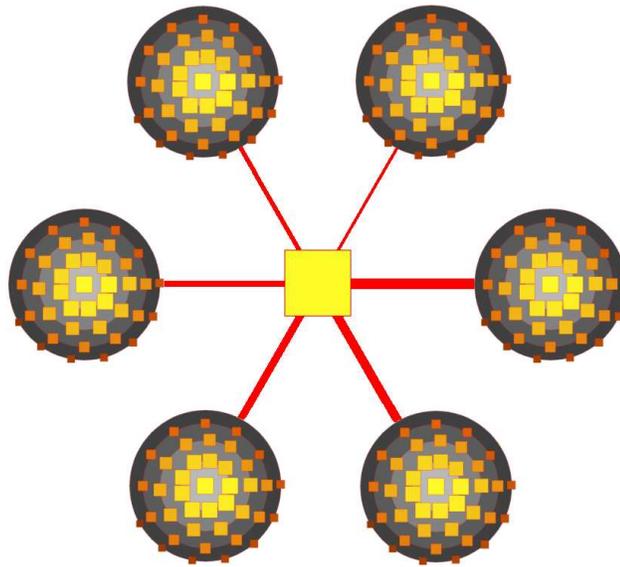


Figura 4.35: Experimento 2: Clustering Set - Anéis Concêntricos.

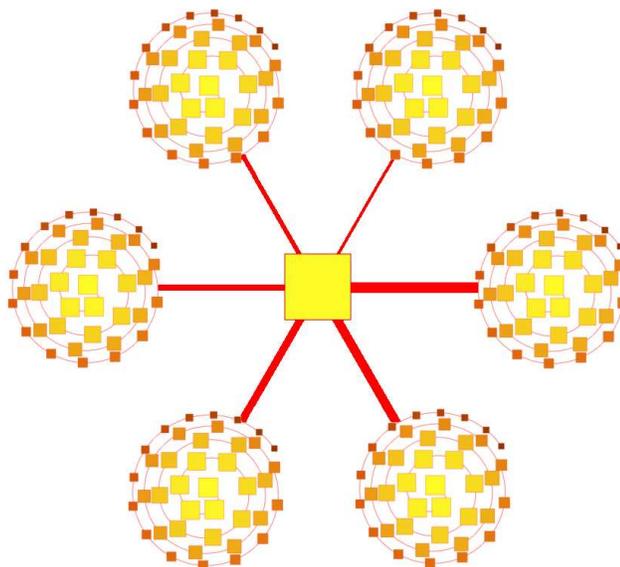


Figura 4.36: Experimento 2: Clustering Set - Espiral.

4.3.1 Projeto do experimento

Seguindo a metodologia usada previamente, detalham-se a seguir os passos do framework DECIDE.

1. Determinar os objetivos

- Quais são os objetivos de alto nível da avaliação?
Mantiveram-se os mesmos objetivos de alto nível usados no primeiro experimento (seção 4.2.1) e adicionaram-se os objetivos a seguir:
 - Avaliar a técnica híbrida Clustering Set - Anéis Concêntricos.
 - Avaliar a técnica híbrida Clustering Set - Espiral.
- Quem quer a avaliação e por quê?
Ao igual que foi detalhado no primeiro experimento, este experimento pode ser útil para qualquer pessoa interessada em projetar um sistema CBIR.

2. Explorar as questões específicas a serem respondidas

Para garantir que o usuário entendeu a estrutura, alguns pontos foram considerados como:

- É possível entender a distribuição das imagens? Para conhecer o real entendimento do usuário sobre as estruturas, exploraram-se perguntas como: o usuário é capaz de achar a imagem de consulta, a terceira imagem mais relevante, a imagem menos relevante e a terceira imagem menos relevante de cada estrutura?
- Qual é o grau de satisfação frente a cada estrutura?
- A estrutura fornece informação suficiente ao usuário?
- Qual estrutura é mais adequada para visualizar os resultados de um sistema CBIR?
- Como as estruturas propostas poderiam ser melhoradas?
- Como novas estruturas poderiam ser criadas?

3. Escolher o paradigma de avaliação e as técnicas

Assim, como no experimento anterior, desejava-se avaliar critérios de usabilidade, portanto usou-se o método de avaliação Teste de Usabilidade. O instrumento de avaliação consistiu no uso de questionários em papel. Maiores detalhes serão tratados na Seção 4.3.2.

4. Identificar as questões práticas

Participantes

A avaliação desenvolveu-se com a colaboração de 18 alunos voluntários de uma disciplina oferecida no Instituto de Computação da Unicamp. Esta atividade não implicou nenhuma penalidade ou ganho sobre a disciplina precedente.

Equipamento

O experimento foi realizado em laboratórios do Instituto de Computação da Unicamp em horário específico. Cada participante teve acesso a um computador. O usuário teve a liberdade para escolher o sistema operacional a usar.

Material

Usaram-se os mesmos quatro documentos que foram usados no Experimento 1, adaptados às necessidades deste experimento.

5. Decidir como lidar com questões éticas

Nos documentos de Instruções aos Usuários e no TCLE que foram entregues a cada usuário, explicam-se claramente as questões éticas. Além disso, falou-se sobre os principais detalhes éticos ao começo de cada experimento.

6. Avaliar, interpretar e apresentar os dados

A avaliação e apresentação dos dados será discutida na seção 4.3.3.

4.3.2 Desenvolvimento dos testes

Os testes foram realizados em laboratório. O questionário de avaliação teve as estruturas visuais impressas. Considerando que o papel não reflete todas as características visuais que um monitor de computador por exemplo pode mostrar, disponibilizou-se um arquivo com as estruturas visuais a serem avaliadas. As imagens mostradas neste arquivo representam as estruturas visuais sem imagens. As estruturas estão formadas por células vazias, onde cada célula representa uma imagem. A própria imagem de consulta é considerada ao mesmo tempo como a imagem mais relevante do conjunto devolvido.

O procedimento do experimento consistiu em:

1. Cada usuário ocupou um computador.
2. Em seguida, entregou-se o conjunto de documentos em papel anteriormente descritos e explicou-se verbalmente os principais detalhes do experimento reforçando assim o documento de instruções aos usuários.
3. Posteriormente, os usuários leram e assinaram o Termo de Consentimento Livre e Esclarecido.
4. Uma vez tendo aceitado as condições do TCLE, os usuários leram o documento de instruções aos usuários.
5. Os usuários preencheram o documento de perfil de usuário.

6. Seguidamente pediu-se aos usuários entrar em um site onde se encontrava uma apresentação com slides, contendo uma imagem por slide. Cada imagem representava uma estrutura visual a ser avaliada.
7. Finalmente os usuários avaliaram as estruturas mediante o formulário de avaliação. Neste documento encontravam-se as estruturas visuais impressas. Pediu-se especial cuidado em diversas características e recursos visuais que apoiam cada estrutura, como cor, tamanho, espessura, etc.

As Figuras 4.37 e 4.37 mostram aos usuários durante o desenvolvimento desse experimento.



Figura 4.37: Experimento 2 - Desenvolvimento do teste com usuários.



Figura 4.38: Experimento 2 - Desenvolvimento do teste com usuários.

Foram feitas duas perguntas abertas em cujas respostas os usuários puderam expressar o que pensam sobre as estruturas visuais apresentadas e dar sugestões sobre novas possíveis estruturas.

Como foi dito, o foco deste experimento não foi avaliar a efetividade das técnicas de recuperação, e sim a estrutura visual apresentada. Portanto, para evitar que os usuários desviassem a atenção da estrutura, as imagens da estrutura foram substituídas por células vazias.

4.3.3 Resultados

Esta seção apresenta os resultados experimentais obtidos.

1. Perfil de Usuário

Este experimento contou com a participação de 18 usuários voluntários. Os participantes foram divididos por faixa etária sendo que 63,9% do total de usuários encontram-se na faixa etária entre 20 a 25 anos, e o restante 36,1% de usuários têm mais de 25 anos. Deles, 72,2% são de sexo masculino, e 27,8%, feminino (ver Figura 4.39). Todos os participantes cursam ensino superior dos quais 72,2% cursam estudos de pós-graduação e 27,8% estudam graduação (ver Figura 4.40). Quanto à finalidade com que usam o computador, os participantes indicaram que, em geral, usam o computador para vários objetivos como estudo, pesquisa, trabalho, diversão, comunicação, etc.

Similarmente ao primeiro experimento, no formulário de perfil, além de dados básicos como escolaridade perguntou-se pela familiaridade e frequência com que costumam usar o computador, realizar buscas, realizar buscas de imagens e buscas de imagens por conteúdo (CBIR). A experiência dos participantes foi estabelecida de acordo com a familiaridade e frequência com que realizam esses critérios. A maior parte de usuários está familiarizada com computadores em geral, mas quando as perguntas questionam pontos cada vez mais específicos, o conjunto de usuários admite ter menor familiaridade e frequência com o critério que está sendo considerado.

As Figuras 4.41, 4.42, 4.43, 4.44, 4.45, 4.46, 4.47 e 4.48 mostram diversos critérios que definem o perfil dos usuários. Os usuários avaliaram cada critério na escala de 1 até 5 (eixo horizontal). Se a distribuição das barras concentra-se nos valores altos da escala, então os usuários consideram que são mais experientes no critério em questão. Se a distribuição concentra-se no meio, os usuários serão regularmente experientes e se a distribuição concentra-se nos valores baixos, os usuários terão pouca experiência.

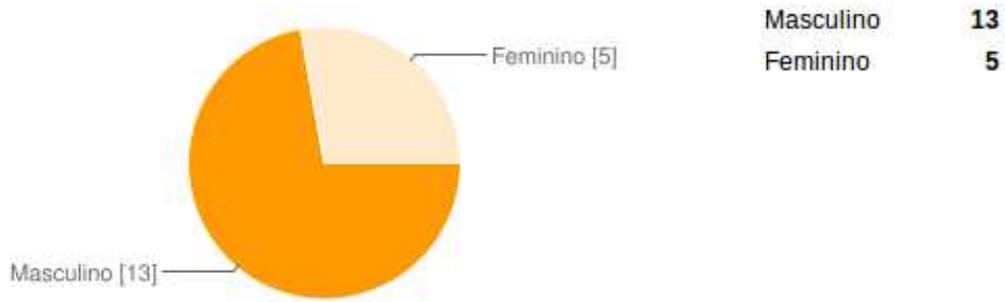


Figura 4.39: Experimento 2 - Gênero.

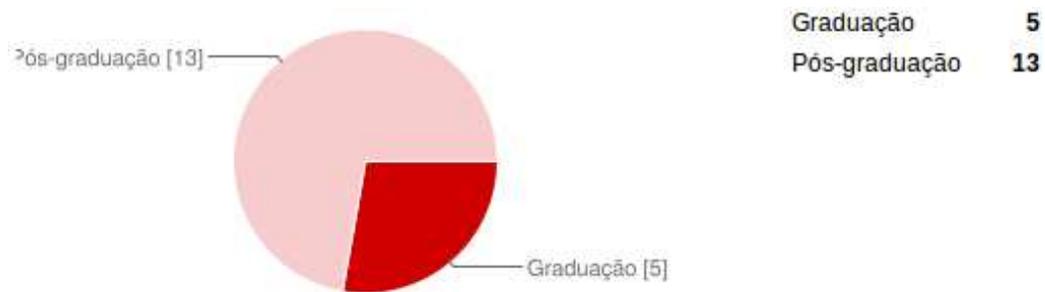


Figura 4.40: Experimento 2 - Escolaridade.

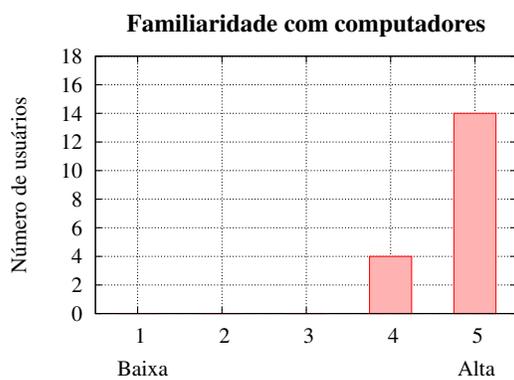


Figura 4.41: Experimento 2 - Familiaridade com computadores.

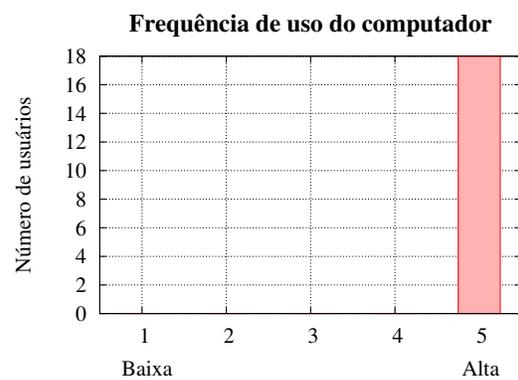


Figura 4.42: Experimento 2 - Frequência de uso de computadores.

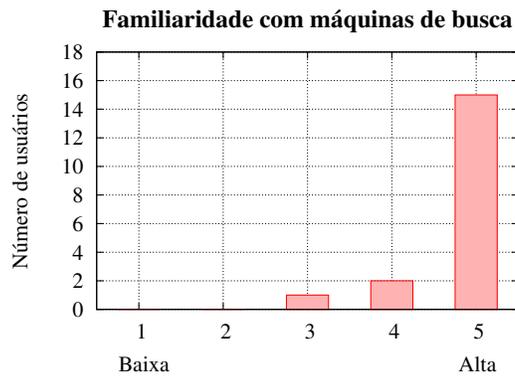


Figura 4.43: Experimento 2 - Familiaridade com máquinas de buscas Web.

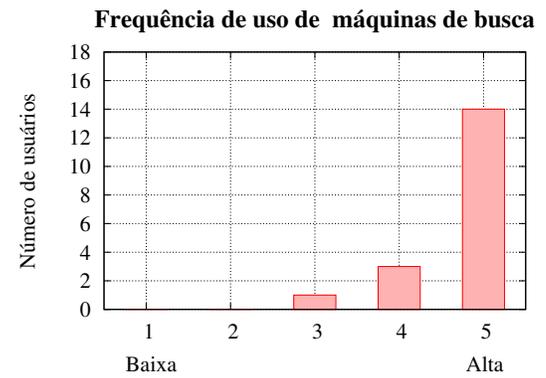


Figura 4.44: Experimento 2 - Frequência com que usa as máquinas de buscas.

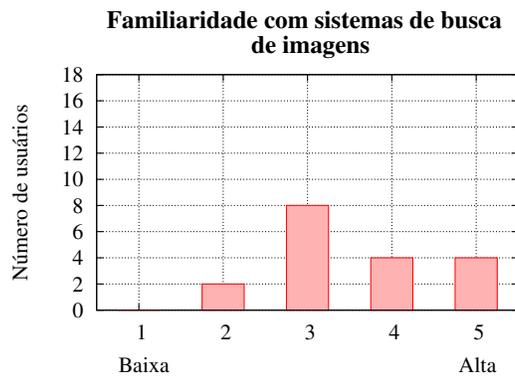


Figura 4.45: Experimento 2 - Familiaridade com sistemas de busca de imagens.

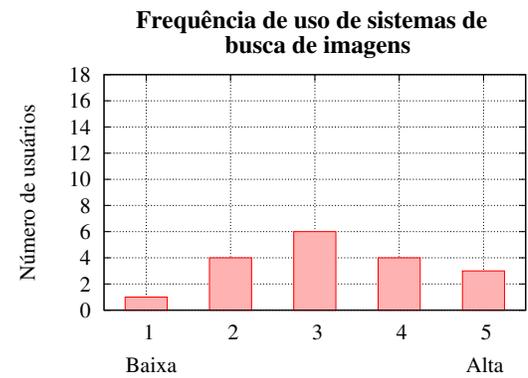


Figura 4.46: Experimento 2 - Frequência com que usa sistemas de busca de imagens.

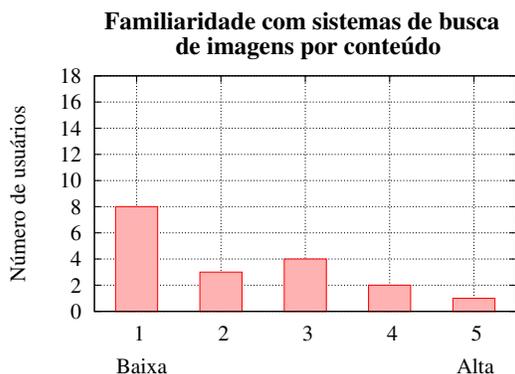


Figura 4.47: Experimento 2 - Familiaridade com sistemas de busca de imagens por conteúdo.

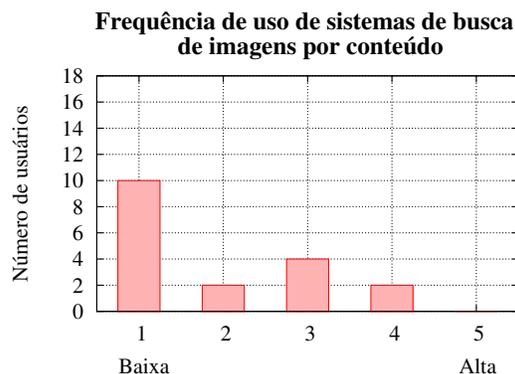


Figura 4.48: Experimento 2 - Frequência com que usa sistemas de busca de imagens por conteúdo.

2. Avaliação das Estruturas

A seguir é apresentado o conjunto de resultados obtidos deste experimento para diversos critérios, considerando-se as oito estruturas avaliadas. Um conjunto de perguntas foi feito sobre diversos aspectos para cada técnica avaliada. Primeiramente pediu-se para achar a célula que conteria a imagem mais relevante, a terceira mais relevante, a menos relevante e a terceira menos relevante, com o intuito de se descobrir o grau de entendimento que cada usuário teve sobre cada estrutura avaliada.

A Tabela 4.5 exibe a percentagem de respostas corretas para cada célula que foi procurada dentro de cada técnica de visualização avaliada. A Figura 4.49 apresenta os mesmos resultados graficamente.

	Grid	Grid com Cores	Anéis Concêntricos	Espiral
mais rel.	66,67%	61,11%	94,44%	94,44%
3º mais rel.	66,67%	61,11%	44,44%	77,78%
menos rel.	83,33%	77,78%	61,11%	83,33%
3º menos rel.	66,67%	61,11%	55,56%	83,33%

	Incremental Board	Clustering Set	ClusteringSet - Anéis Concêntricos	ClusteringSet - Espiral
mais rel.	88,89%	88,89%	94,44%	94,44%
3º mais rel.	66,67%	5,56%	11,76%	11,76%
menos rel.	0,00%	38,89%	29,41%	41,18%
3º menos rel.	0,00%	22,22%	11,76%	23,53%

Tabela 4.5: Experimento 2 - Entendimento.

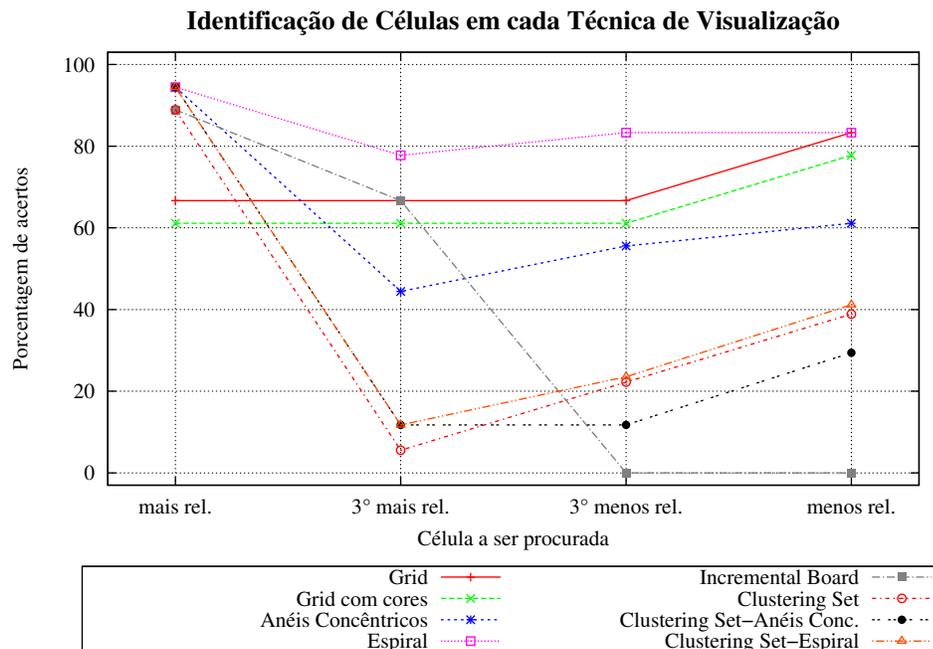


Figura 4.49: Experimento 2 - Entendimento.

As Figuras 4.50, 4.51 e 4.52 mostram respectivamente o grau de satisfação, a quantidade de informação e a ordenação das estruturas por adequação para representar o resultado

de sistemas CBIR segundo a opinião dos usuários.

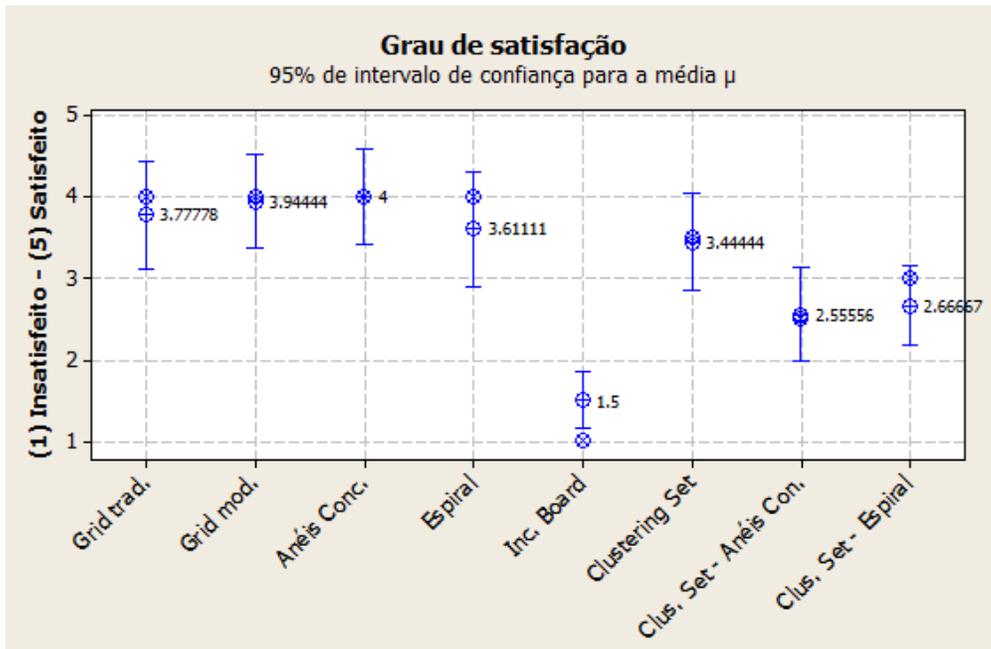


Figura 4.50: Experimento 2 - Grau de satisfação.

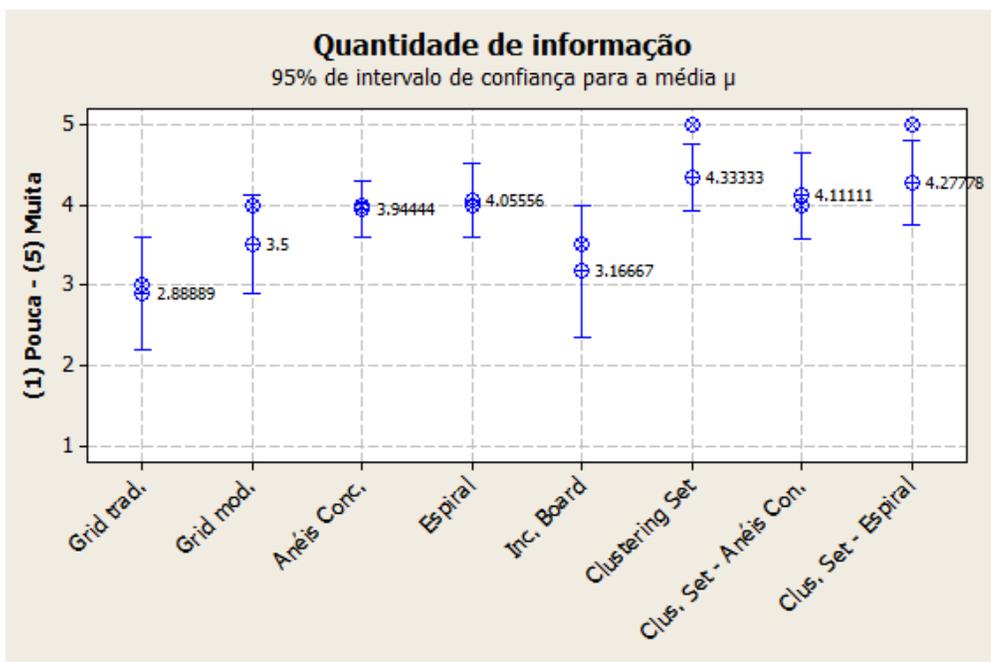


Figura 4.51: Experimento 2 - Quantidade de informação.

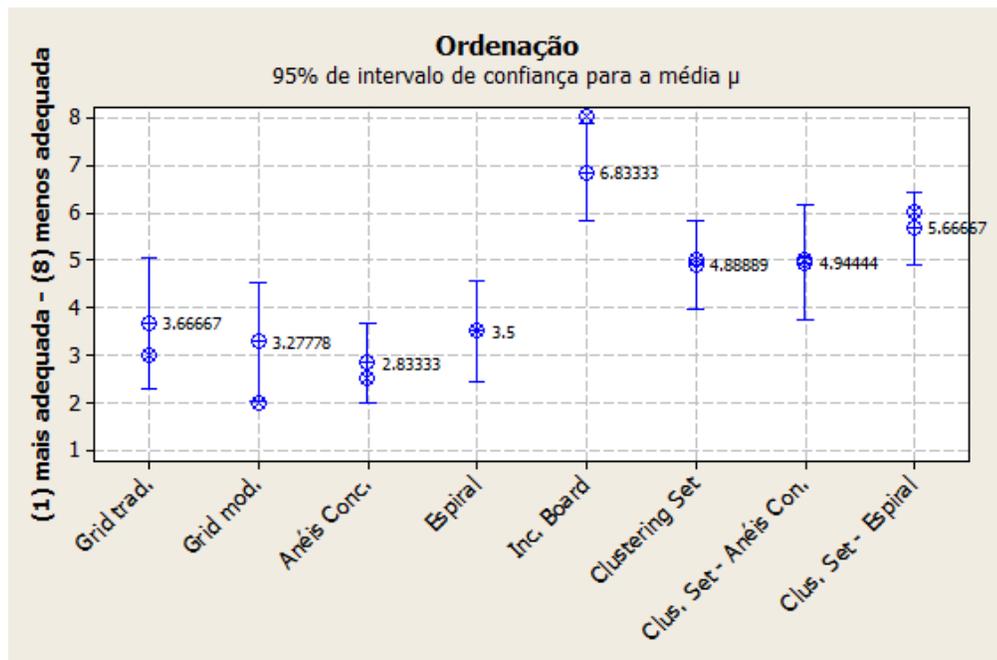


Figura 4.52: Experimento 2 - Ordenação das estruturas baseadas na adequação para visualização em sistemas CBIR.

4.3.4 Análise

- Em geral, os usuários estão familiarizados e usam frequentemente o computador e máquinas de busca. Com relação à busca de imagens, a maior parte dos usuários considera que tem uma familiaridade e frequência de uso média. Finalmente, com respeito ao uso de sistemas CBIR, a maior parte de usuários admitiu ter pouca familiaridade e frequência de uso.
- As estruturas foram disponibilizadas no computador, mas considerando-se que o questionário de avaliação devia ser preenchido manualmente em papel, muitos usuários não usaram as estruturas disponíveis online no grau esperado. Dessa forma, alguns usuários não perceberam algumas marcas visuais que por sua vez eram difíceis de enxergar no papel, como a variação da espessura das arestas que unem a imagem de consulta com as imagens representativas. Os resultados mostraram que o grau de entendimento (inferido da percentagem de acertos para localizar algumas células) das estruturas com alta quantidade de informação (Clustering Set e estruturas híbridas) foi baixo em geral, em contraste aos resultados do primeiro experimento. Estes fatos levaram à modificação de algumas marcas visuais como cor, tamanho e largura para incrementar o grau do

entendimento dessas estruturas. Estas modificações foram avaliadas em um terceiro experimento detalhado na próxima seção (seção 4.4).

- Adicionou-se um método denominado Grid Modificado. Este método é parecido ao método Grid, com a única diferença de apresentar cores que variam conforme a relevância em relação à consulta. Os resultados mostram um empate técnico entre ambos métodos.
- Quanto ao entendimento, para os três métodos propostos, a célula com a imagem mais relevante teve alta percentagem de acerto, porém a terceira célula mais relevante, a célula menos relevante e a terceira célula menos relevante tiveram uma baixa percentagem de acerto. Isto se refletiu no grau de satisfação e na adequação para visualizar resultados em sistemas CBIR, nos quais estes métodos foram avaliados com pontuação baixa. Além do formulário ter sido preenchido em papel, possivelmente as estruturas não foram suficientemente intuitivas para o usuário.
- Um problema detectado foi que muitos usuários fizeram um seguimento não esperado nas imagens. Esperava-se que os usuários considerassem mais relevantes todas as imagens do primeiro agrupamento e depois as imagens do segundo e assim por diante (Figura 4.53). Na prática, uma importante quantidade deles fez o sequenciamento (ordenação) das imagens pelos centros dos grupos como mostrado na Figura 3.1. Isso pode ser justificado pelo fato de que esta técnica sugere um percurso por grupos e não sequencial. Em trabalhos futuros, podem-se abordar estes fatores.

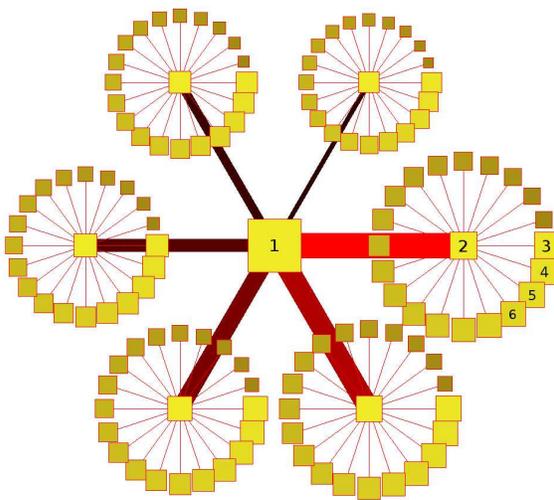


Figura 4.53: Ordem das imagens.

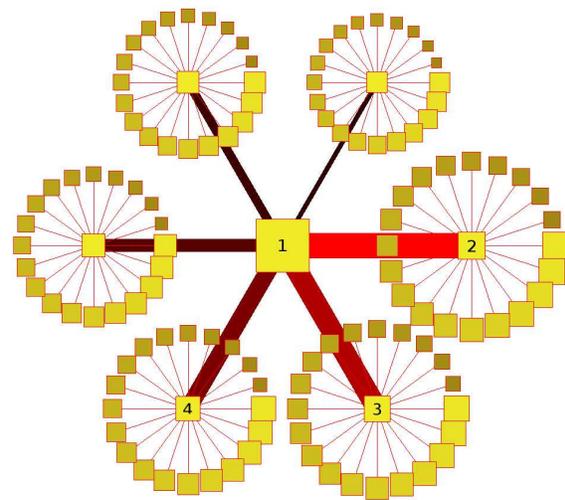


Figura 4.54: Outra ordem das imagens.

- Similarmente ao primeiro experimento, a estrutura Grid mostrou ter uma boa aceitação devido aos usuários a conhecerem *a priori* e estarem acostumados a ela, como pode se apreciar na sugestão a seguir:

Levar em consideração as estruturas com as quais os usuários já estão familiarizados. Poderia ser uma nova estrutura, mas com traços das estruturas tradicionais de hoje.

- Sobre a pergunta para a criação de novas estruturas, algumas ideias indicam o uso de estrutura 3D ou o uso de árvores.

- *Algum tipo de estrutura 3D.*
- *Estrutura em árvore pode ser analisada.*

- Uma sugestão que merece destaque diz respeito à utilização de Clustering Sets dentro de outros Clustering Sets. Como foi explicado no Capítulo 3 este ponto foi implementado de duas formas possíveis: (i) O arcabouço desenvolvido permite a criação de qualquer quantidade de níveis de Clustering Sets vistos na tela num momento dado. Porém, como as imagens ficariam muito pequenas com muitos níveis, nestes experimentos a estrutura permaneceu da forma apresentada (com dois níveis). (ii) Pode-se também explorar as imagens dentro de um grupo hierarquicamente. Ao se clicar sobre uma imagem dentro de um grupo, uma nova instância da mesma técnica é mostrada, fazendo um novo agrupamento, desta vez sobre o conjunto de imagens do grupo selecionado.

Clustering Set de Clustering sets navegáveis.

- Uma outra sugestão a tomar em conta foi que nem sempre muitos dados na tela representam uma boa escolha. Deve-se tomar cuidado para não mostrar muitas informações que possam confundir ao usuário no lugar de ajudar a visualizar o conjunto de imagens.

Não tentar codificar informações demais na estrutura. “Menos pode ser mais”.

4.4 Terceiro experimento

No segundo experimento verificou-se que as marcas visuais dos métodos propostos, como espessura das arestas, foram pouco claras para indicar uma ordenação dos grupos. Vale lembrar que o grupo mais próximo à imagem de consulta foi ligado pela aresta mais espessa. Essa espessura vai diminuindo à medida que os clusters ficam menos relevantes.

Portanto realizou-se este terceiro experimento mudando algumas características visuais (detalhadas a seguir) das estruturas e levando em consideração algumas sugestões dos usuários feitas no segundo experimento.

Similarmente ao segundo experimento, as imagens foram tiradas e conservou-se somente uma célula vazia no lugar. Nas estruturas propostas, incrementou-se a espessura das arestas que unem a imagem central com as imagens representativas de cada grupo, mantendo a relação espessura-relevância (grupos mais relevantes possuem uma aresta mais espessa). Do mesmo jeito, variou-se a cor das arestas do vermelho ao preto conforme diminua a relevância dos grupos.

Na estrutura Clustering Set mudou-se o tamanho das células dentro de cada grupo conforme uma ordenação interna de cada grupo, a imagem maior é a mais próxima ao elemento representativo do grupo (centroide, medoide, etc) e o tamanho vai diminuindo conforme a similaridade com o elemento representativo diminua.

No segundo experimento a gama de cores das células variou da cor amarela à cor vermelha escuro. Porém amarelo e vermelho são consideradas cores quentes o que indicaria maior relevância tanto para o amarelo como para o vermelho (por tal motivo um usuário marcou as imagens relevantes e irrelevantes em posição invertida). Neste terceiro experimento mudou-se a gama de cores do amarelo - vermelho escuro para o amarelo-verde escuro, dado que o verde é uma cor fria.

Mostra-se a seguir o resultado dessas mudanças sobre as mesmas oito estruturas avaliadas no segundo experimento:

- Grid: Figura 4.55;
- Grid modificada: Figura 4.56;
- Anéis Concêntricos: Figura 4.57;
- Espiral: Figura 4.58;
- Incremental Board: Figura 4.59;
- Clustering Set: Figura 4.60;
- Clustering Set - Anéis Concêntricos: Figura 4.61;

- Clustering Set - Espiral: Figura 4.62.

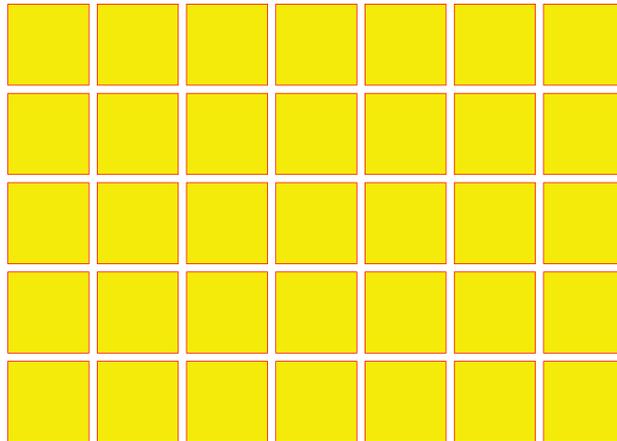


Figura 4.55: Experimento 3: Grid.

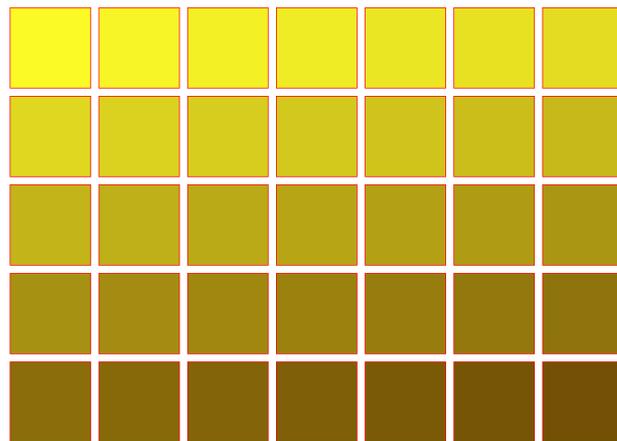


Figura 4.56: Experimento 3: Grid Modificado.

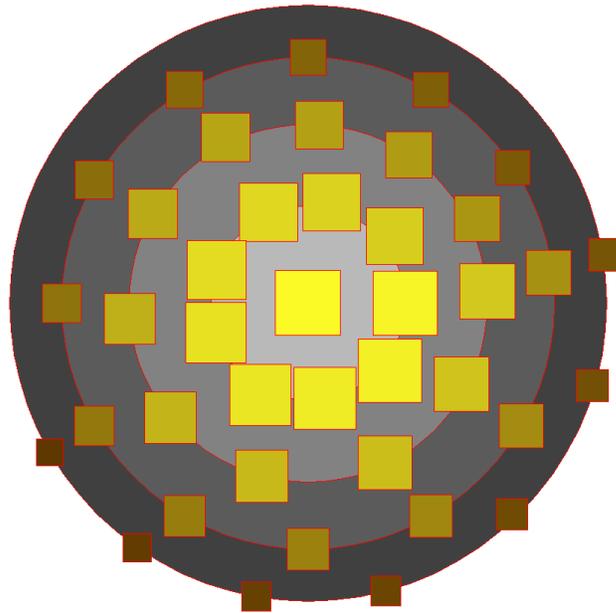


Figura 4.57: Experimento 3: Anéis Concêntricos.

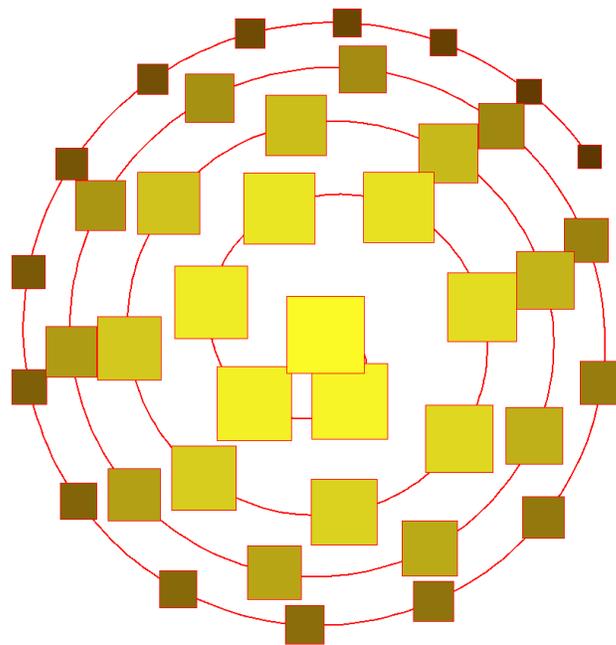


Figura 4.58: Experimento 3: Espiral.

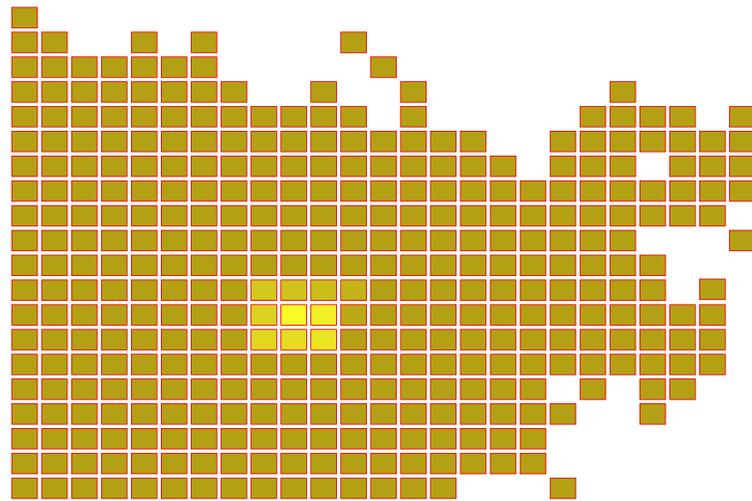


Figura 4.59: Experimento 3: Incremental Board.

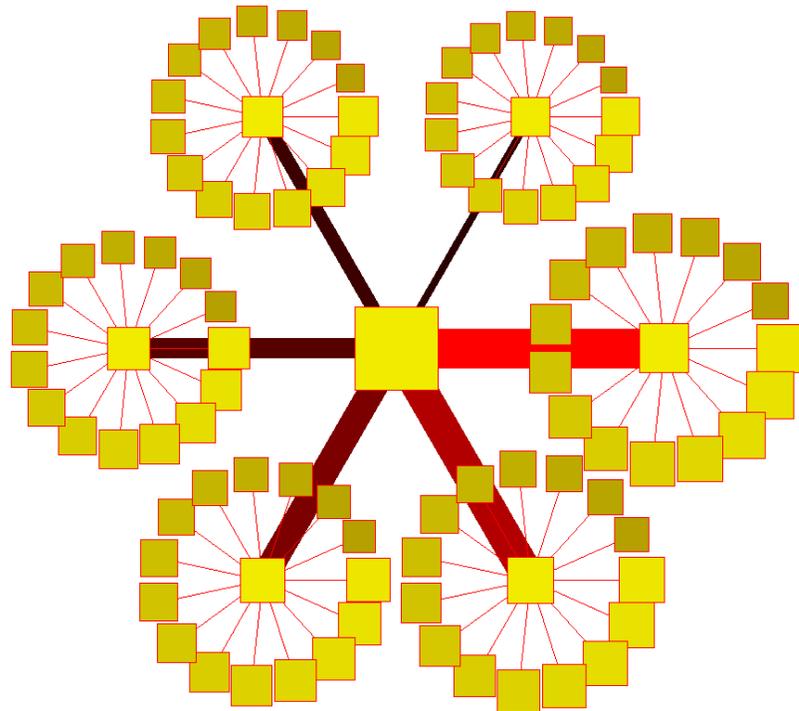


Figura 4.60: Experimento 3: Clustering Set.

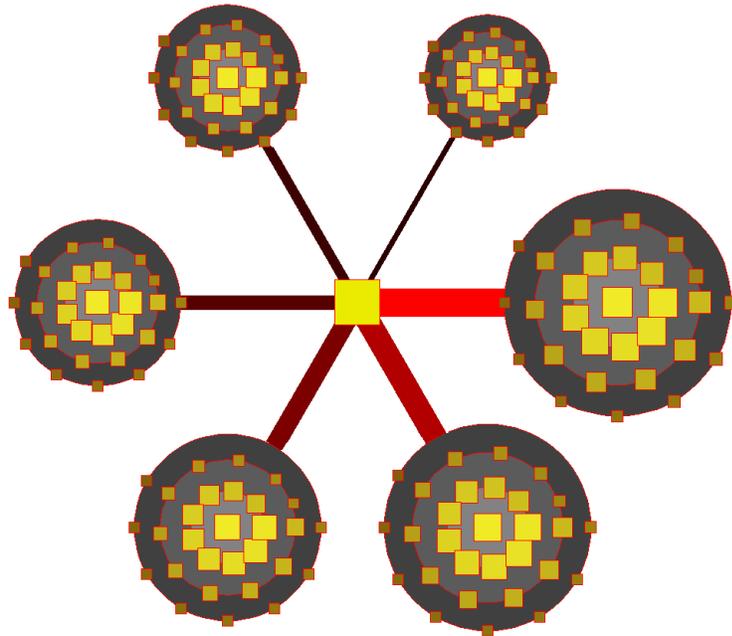


Figura 4.61: Experimento 3: Clustering Set - Anéis Concêntricos.

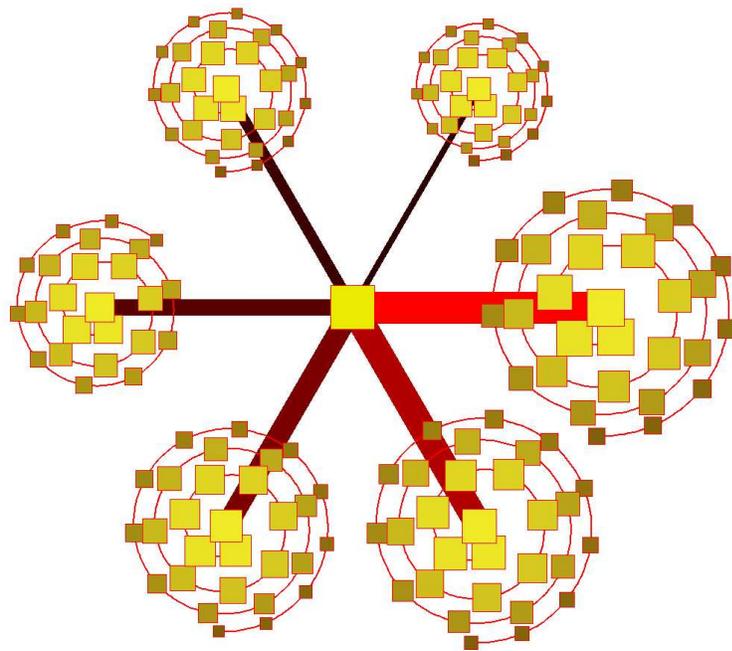


Figura 4.62: Experimento 3: Clustering Set - Espiral.

4.4.1 Projeto do experimento

O projeto experimental seguiu o mesmo planejamento adotado no segundo experimento. As estruturas avaliadas são as mesmas que foram avaliadas no segundo experimento (mantendo a ausência das imagens), modificadas de acordo com os resultados e as sugestões feitas pelos usuários. Desta vez os questionários de perfil e de avaliação foram respondidos no computador com o intuito de melhorar a percepção dos usuários aproveitando os recursos visuais fornecidos pela tela de computador. Quase todos os participantes foram alunos da pós-graduação do Instituto de Computação e foram convidados para fazer o experimento, com exceção de um usuário que se ofereceu voluntariamente. O número de usuários participantes foi de 20 pessoas. No questionário de avaliação foi feita uma série de perguntas sobre cada estrutura visual. As imagens mostradas nele representam as estruturas visuais sem imagens. A própria imagem de consulta é considerada ao mesmo tempo como a imagem mais relevante do conjunto devolvido. Similarmente ao primeiro e segundo experimentos, mantiveram-se duas perguntas abertas, as quais tiveram o intuito de obter sugestões dos usuários para melhorar as estruturas propostas e para propor eventuais estruturas novas.

4.4.2 Desenvolvimento dos Testes

O questionário de avaliação foi uma apresentação sobre a qual cada usuário fez algumas marcações.

O procedimento do experimento consistiu nos seguintes passos:

1. Cada usuário ocupou um computador, dado que era preciso preencher os formulários e porque as estruturas não estavam impressas e so podia ser visualizadas no computador.
2. Em seguida, entregaram-se dois documentos, o TCLE e o documento de instruções de usuários e explicou-se verbalmente os principais detalhes do experimento reforçando assim o documento de instruções aos usuários.
3. Posteriormente, os usuários leram e assinaram o Termo de Consentimento Livre e Esclarecido.
4. Uma vez tendo aceitado as condições do TCLE, os usuários leram o documento de instruções.
5. Em seguida, os usuários entraram em um site onde se encontravam os documentos do formulário de perfil de usuário e de avaliação.
6. Os usuários preencheram o formulário de perfil de usuário online.

7. Finalmente os usuários avaliaram as estruturas mediante o formulário de avaliação. Pediu-se especial cuidado em diversas características e recursos visuais que apoiam cada estrutura como cor, tamanho, largura, etc.

Foram mantidas as duas perguntas abertas do experimento anterior onde cada usuário pôde expressar a opinião sobre as estruturas visuais apresentadas e dar sugestões sobre novas possíveis estruturas.

4.4.3 Resultados

Esta seção apresenta os resultados experimentais obtidos.

1. Perfil de Usuário

Do total de participantes (20 pessoas), 45% deles (9 pessoas) encontram-se na faixa etária entre 22 a 25 anos, 50% (10 pessoas) se encontram entre 26 e 29 anos e 5% (1 pessoa) tem mais de 30 anos. Destes 20 participantes, 90% deles são de gênero masculino e 10 % de gênero feminino. Todos os participantes cursam ensino superior, dos quais 95% estudam pós-graduação em computação e 5% estudam graduação na física.

Como pode se observar, em contraste com os primeiros dois experimentos, desta vez o conjunto de usuários, em média, concentra-se em uma faixa etária maior devido principalmente a que quase a totalidade deles pertencem ao programa de pós-graduação.

Os critérios avaliados no formulário de perfil dos experimentos anteriores foram mantidos neste experimento. Usou-se a mesma escala de 1 (baixo) até 5 (alto) para avaliar esses critérios. Similarmente aos dois primeiros experimentos, a maior parte dos usuários está familiarizada com computadores, mas quando as perguntas questionam pontos cada vez mais específicos, o conjunto de usuários admite ter baixa familiaridade e/ou frequência no tópico abordado.

As Figuras 4.63 e 4.64 mostram o gênero e a escolaridade dos participantes. A Figura 4.65 mostra a principal finalidade com que os usuários usam o computador. As Figuras 4.66, 4.67, 4.65, 4.68, 4.69, 4.70, 4.71, 4.72 4.73 apresentam cada critério avaliado no perfil de usuário. Similarmente aos experimentos anteriores, o eixo vertical representa o número de usuários que avaliaram o critério na escala (1-5) indicada no eixo horizontal. Se a distribuição das barras se concentra nos valores altos da escala horizontal, os usuários consideram-se mais experientes e se a distribuição se concentra nos valores baixos do eixo horizontal, os usuários consideram que têm pouca experiência no critério em questão.

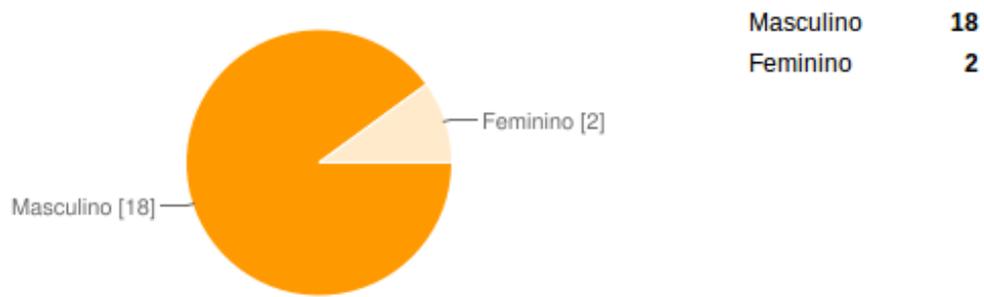


Figura 4.63: Experimento 3 - Gênero

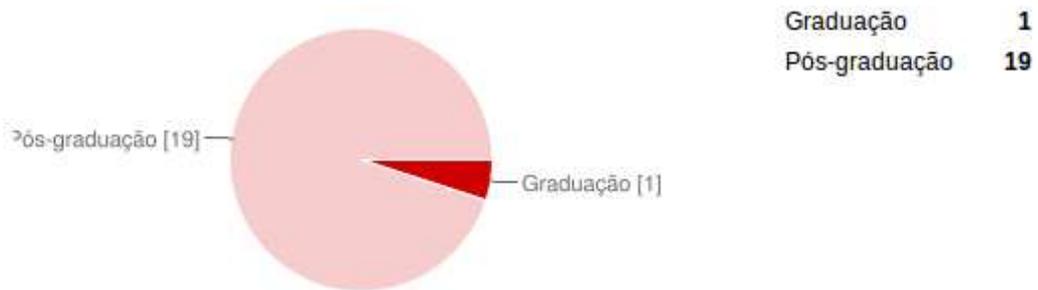


Figura 4.64: Experimento 3 - Escolaridade

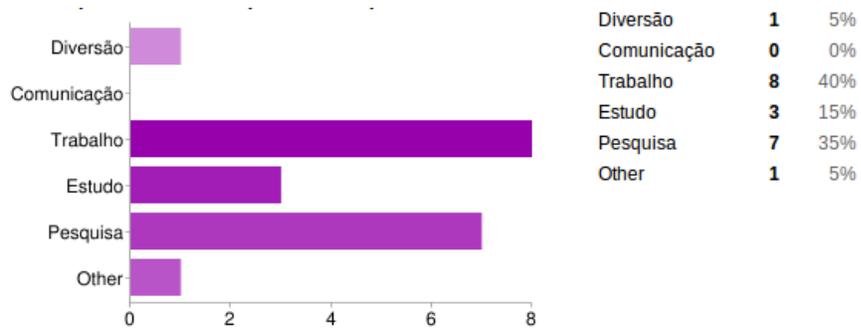


Figura 4.65: Experimento 3 - Principal finalidade com que usa o computador

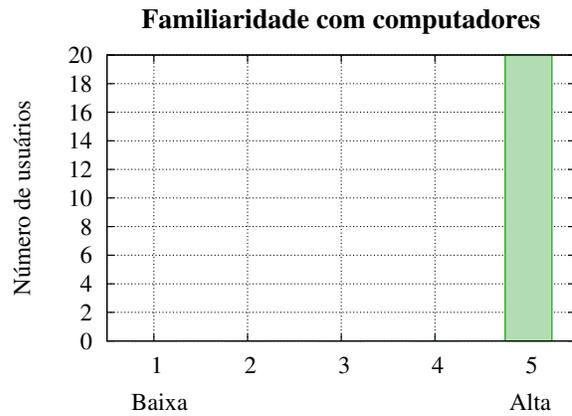


Figura 4.66: Experimento 3 - Familiaridade com computadores.

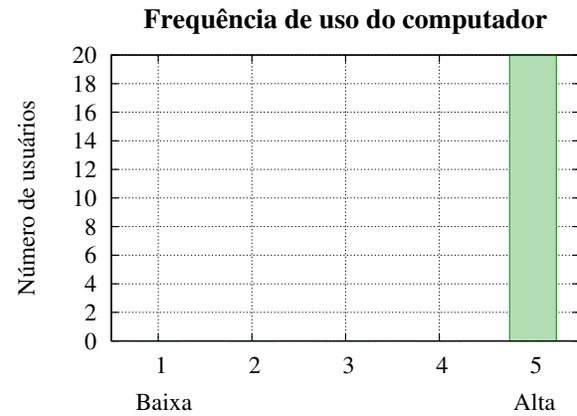


Figura 4.67: Experimento 3 - Frequência de uso de computadores.

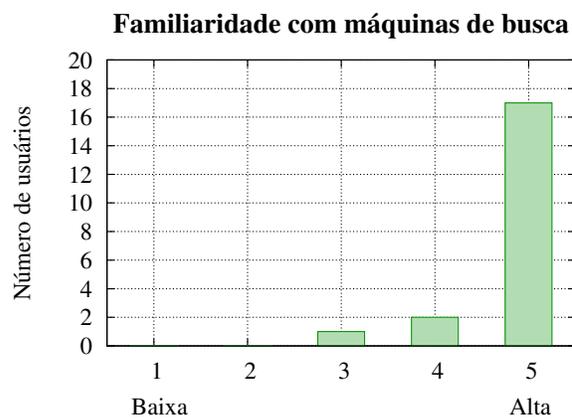


Figura 4.68: Experimento 3 - Familiaridade com máquinas de buscas Web.

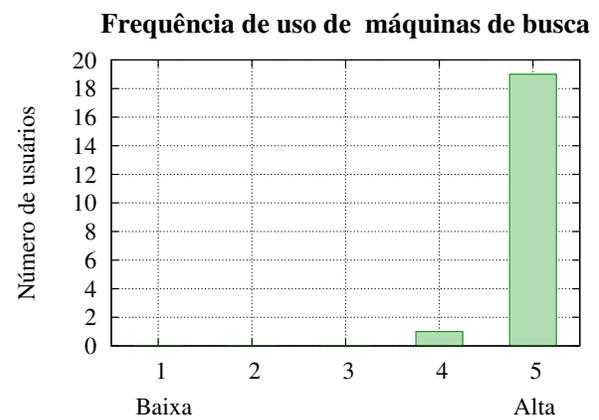


Figura 4.69: Experimento 3 - Frequência com que usa as máquinas de buscas

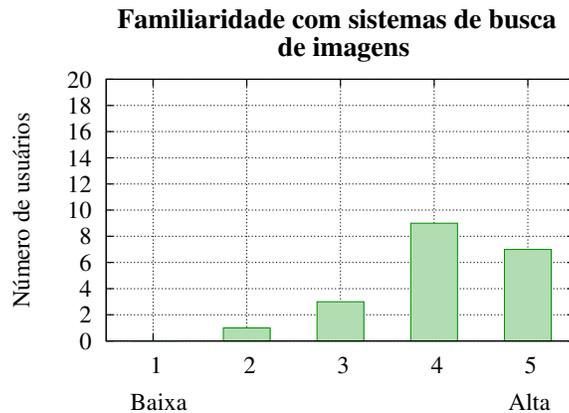


Figura 4.70: Experimento 3 - Familiaridade com sistemas de busca de imagens

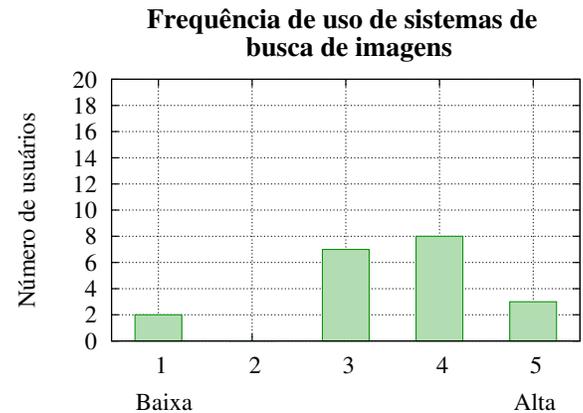


Figura 4.71: Experimento 3 - Frequência com que usa sistemas de busca de imagens

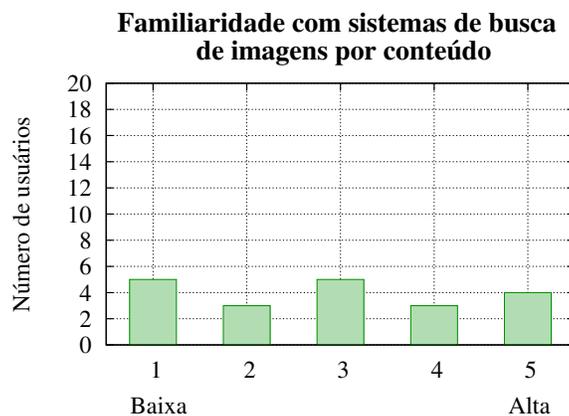


Figura 4.72: Experimento 3 - Familiaridade com sistemas de busca de imagens por conteúdo

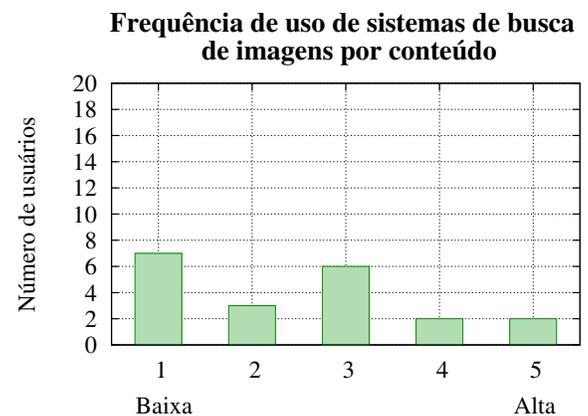


Figura 4.73: Experimento 3 - Frequência com que usa sistemas de busca de imagens por conteúdo

2. Avaliação das Estruturas

A seguir é apresentado o conjunto de resultados obtidos deste experimento para diversos critérios. Avaliaram-se oito estruturas visuais.

Cabe lembrar que as células mudaram de cor com respeito ao segundo experimento e incrementou-se a espessura das arestas nas técnicas de visualização propostas para dar realce à relevância dos grupos.

	Grid	Grid com Cores	Anéis Concêntricos	Espiral
mais rel.	90,0%	85,0%	100,0%	100,0%
3° mais rel.	80,0%	70,0%	65,0%	95,0%
menos rel.	95,0%	95,0%	85,0%	90,0%
3° menos rel.	80,00%	85,0%	80,0%	90,0%

	Incremental Board	Clustering Set	ClusteringSet - Anéis Concêntricos	ClusteringSet - Espiral
mais rel.	80,0%	95,0%	95,0%	95,0%
3° mais rel.	75,0%	35,0%	50,00%	60,0%
menos rel.	0,00%	80,0%	45,0%	75,0%
3° menos rel.	0,00%	80,0%	50,0%	75,0%

Tabela 4.6: Experimento 2 - Entendimento.

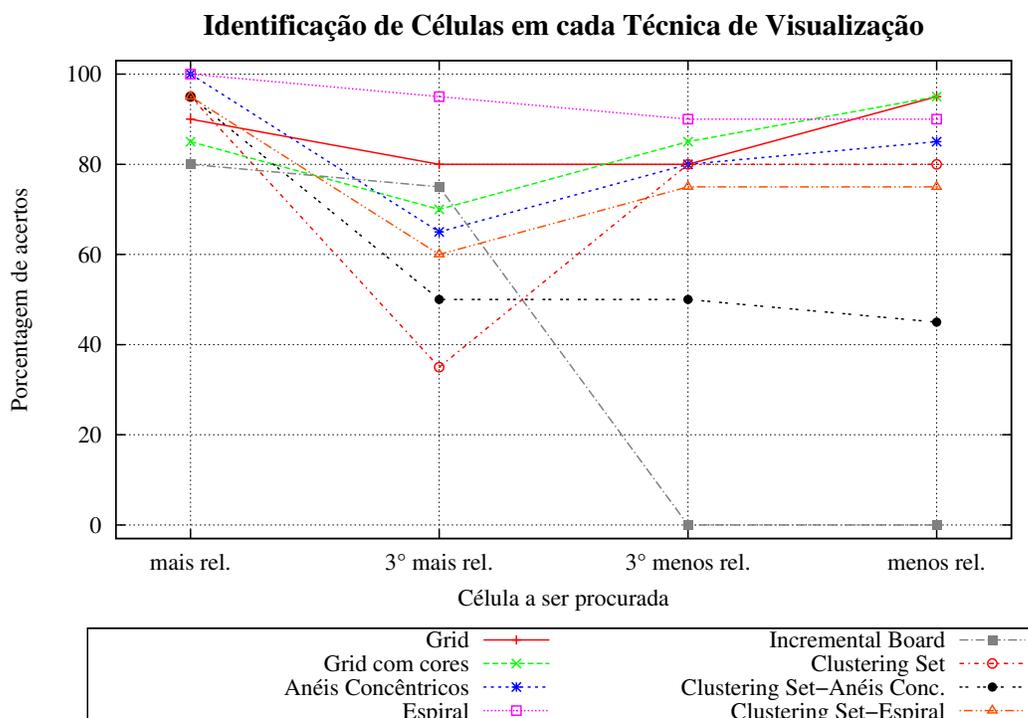


Figura 4.74: Experimento 3 - Grau de Entendimento.

Similarmente ao segundo experimento, pediu-se para achar um conjunto de células com o intuito de obter dados sobre quão entediante é cada estrutura. A Tabela 4.6 apresenta a porcentagem de usuários que achou a célula esperada para cada critério e por cada estrutura visual. A Figura 4.74 mostra estes resultados em forma gráfica.

As Figuras 4.75, 4.76, 4.77 apresentam os gráficos de intervalo de confiança para o grau de satisfação, quantidade de informação fornecida e ordenação das estruturas conforme adequação para visualização de resultados em sistemas CBIR. Cada intervalo representa 95% dos dados amostrados. A média dos valores de cada estrutura encontra-se no centro de cada intervalo e está representada pelo símbolo \oplus seguido do respectivo valor numérico ao lado direito. A mediana de cada intervalo também pode ser identificada pelo símbolo \otimes .

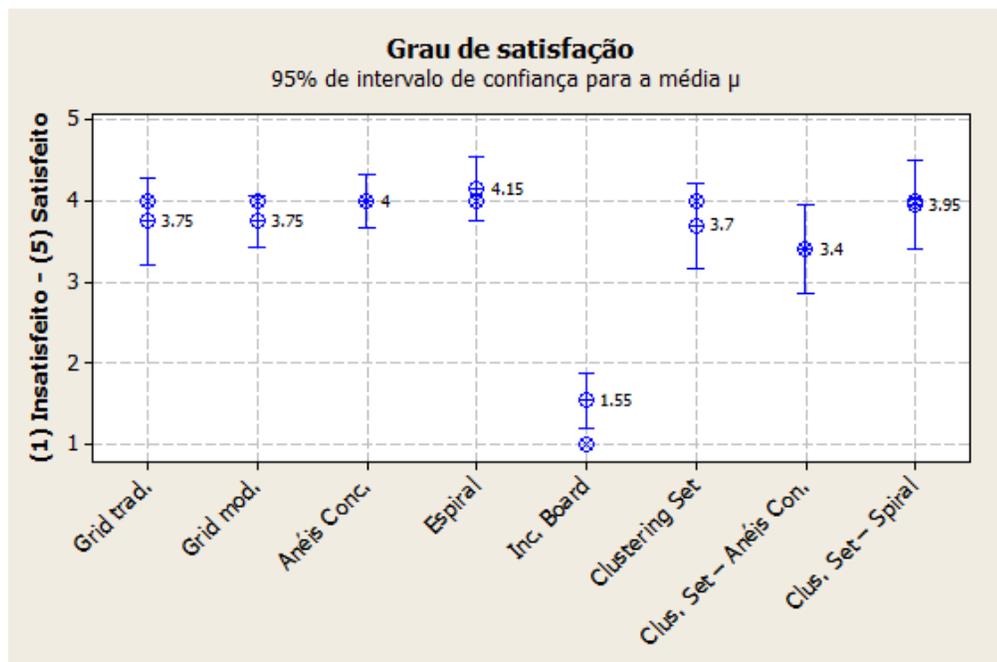


Figura 4.75: Experimento 3 - Grau de Satisfação.

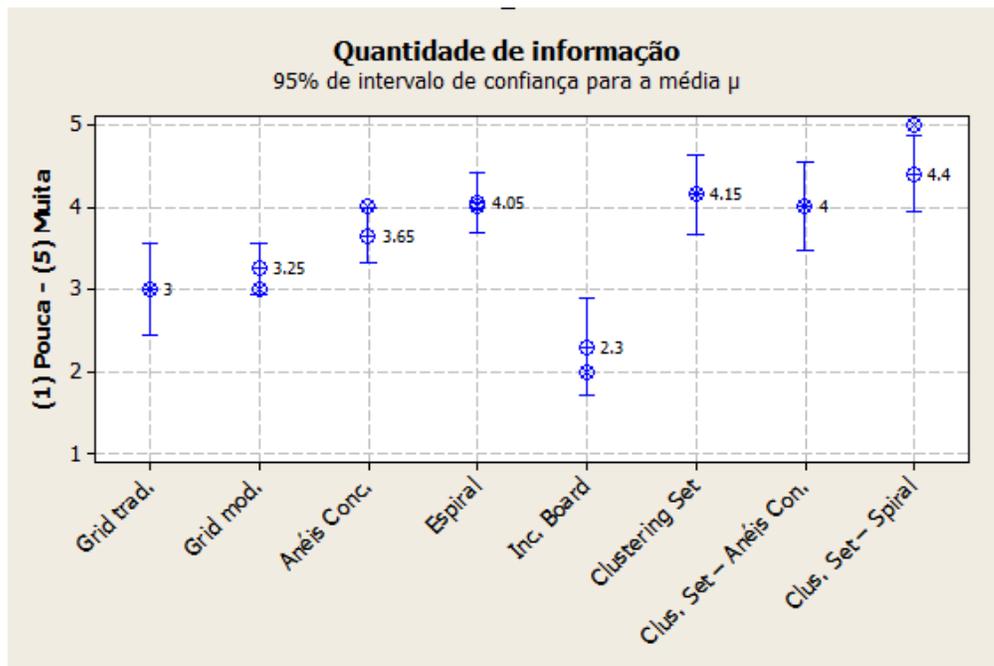


Figura 4.76: Experimento 3 - Quantidade de informação.

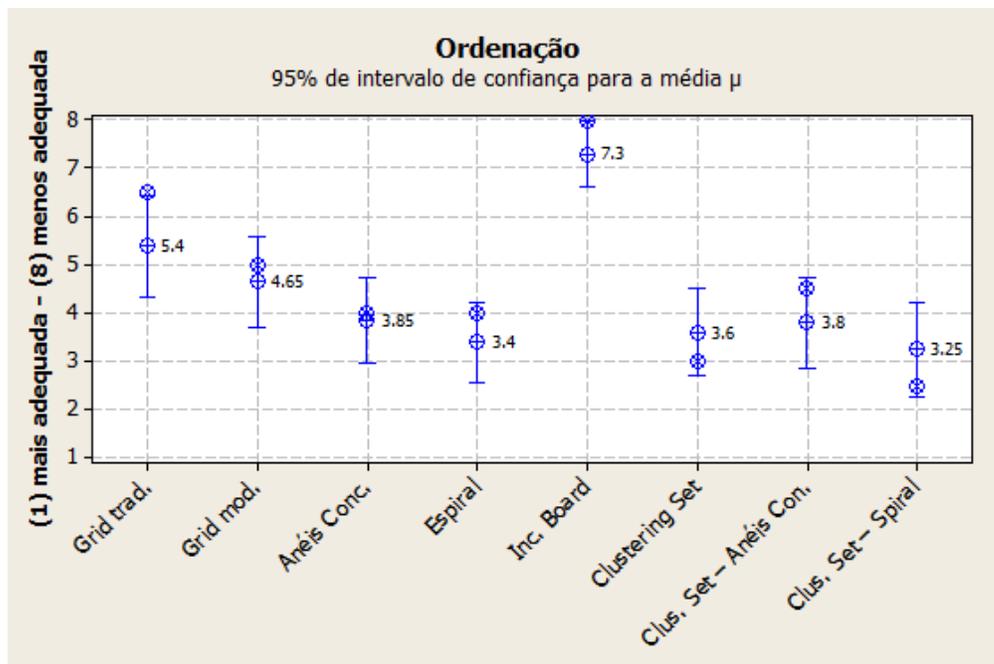


Figura 4.77: Experimento 3 - Ordenação das estruturas baseadas na adequação para visualização em sistemas CBIR.

4.4.4 Análise

- Após as modificações visuais feitas entre o segundo e o terceiro experimento, o impacto nos resultados é marcante. O grau de entendimento para as técnicas propostas (as quais não tiveram bom grau de entendimento no segundo experimento), melhorou consideravelmente (ver Figuras 4.49 e 4.74). Com relação à adequação para seu uso em visualização de sistemas CBIR, a preferência dos usuários por uma das técnicas propostas (Clustering Set - Espiral) melhorou em relação ao segundo experimento, como pode ser visto nos respectivos intervalos de confiança das Figuras 4.52 e 4.77.
- Quando foi pedido aos usuários que ordenassem as estruturas de acordo com a sua preferência (Figura 4.77), pode-se afirmar que o método pior avaliado foi o método Incremental Board. Observa-se também que o método proposto Clustering Set-Espiral e o método Espiral foram melhor avaliados em comparação ao método Grid tradicional. Estes métodos obtiveram um empate técnico com o restante dos métodos.
- Quanto ao Grau de satisfação (Figura 4.75), houve um empate técnico entre os métodos avaliados, com exceção do método Incremental Board, que obteve baixa avaliação.
- Sobre as estruturas Grid, embora os resultados mostrem que os usuários acham que esta estrutura é uma das menos adequadas para visualizar os resultados em sistemas CBIR (Figura 4.77) e consideram que esta estrutura mostra pouca informação comparada com as outras estruturas avaliadas (Figura 4.76), o grau de entendimento (Figura 4.74) e o nível de satisfação (Figura 4.75) foram elevados, aparentemente, devido principalmente a grande familiaridade dos usuários por este método como mostra a opinião do usuário a seguir.

Algumas das estruturas não são sugestivas. O ideal é que as estruturas utilizadas sejam as mais próximas possíveis dos modelos comuns da internet, como grids, já que é mais intuitivo aos usuários comuns.

- Ao respeito dos métodos Grid Tradicional e Grid Modificado, os resultados mostram um empate técnico entre ambos métodos (Figuras 4.74, 4.75, 4.76, 4.77). Cabe ressaltar que também obteve-se um empate técnico para estes métodos no segundo experimento.
- Cabe ressaltar que a estrutura que forneceu maior facilidade para achar células sequenciais (Figura 4.74), em média, foi a estrutura Espiral, sendo a que apresentou maior facilidade para achar a imagem mais relevante, a terceira imagem mais relevante e a terceira imagem menos relevante. Porém, as estruturas grid obtiveram a maior porcentagem de acertos ao tentar achar as imagens menos relevantes.

- O método Incremental Board teve a pior avaliação em todos os critérios considerados (so ou em empate) com exceção do subcritério para escolher a terceira imagem mais relevante na avaliação do grau de entendimento.
- Na escolha do método mais adequado para a visualização de resultados de sistemas CBIR (Figura 4.77), o método híbrido Clustering Set - Espiral foi bem avaliado, em empate técnico com outros métodos. Este resultado é coerente com resultados anteriores dado que o método Clustering Set foi o melhor avaliado no primeiro experimento para todos os critérios avaliados (so ou em empate técnico com outros métodos) e o método Espiral obteve alto grau de entendimento em questões relacionadas à percepção de ordem no segundo e terceiro experimento. Acredita-se que a distribuição organizada do método Clustering Set fora dos grupos (extra-cluster) combinado com a clareza da técnica espiral para representar as imagens intra-cluster, oferece a opção mais eficiente para mostrar resultados de sistemas CBIR.
- Algumas sugestões foram feitas sobre os métodos novos propostos. A última delas sugere a navegação dentro dos grupos nas técnicas propostas. Essa modificação foi considerada e detalhada no Capítulo 3.

- *Considerando uma quantidade muito grande de informações, acho que a estrutura mais adequada seria a Clustering set com o tamanho de cada cluster variando conforme a relevância do resultado. Esta é uma estrutura interessante porque pode esboçar as relações entre as imagens dentro de um determinado cluster. No entanto, acredito que um problema de visualizar a relevância entre as imagens de um mesmo cluster pode surgir quando existe uma quantidade de imagens muito grande em um mesmo cluster. Parece-me que o controle da granularidade do cluster é crítico.*
- *Para os clustering sets, acredito que o espaçamento entre cada cluster deveria ser o mesmo (as imagens das bordas dos dois clusters maiores quase se misturam).*
- *Uma sugestão seria a possibilidade de interação, por exemplo, nas técnicas baseadas em clustering set, pode-se permitir a focalização em algum dos clusters através de click ou movimentar a estrutura com o movimento do mouse.*

- Sobre estas possíveis estruturas novas, algumas sugestões apontam a utilização de uma árvore ou grafo como mostrado a seguir.

- *Acredito que seria bom testar uma estrutura em árvore ou grafo. A raiz ou nó inicial seria a imagem original.*
- *Uma disposição em árvore submersa, onde as imagens seriam mostradas como se a melhor imagem fosse puxada para o fundo de um lago, porém sem diminuir drasticamente seu tamanho. Suas folhas seriam mostradas em diferentes tamanhos de acordo com sua relevância.*
- *Uso de estruturas 3D.*
- *As estruturas poderiam ser melhoradas dispondo as imagens em uma estrutura de relevância apresentando da esquerda para direita e de cima para baixo as imagens mais relevantes, distribuindo ao redor e no sentido desta diagonal as imagens mais similares. Este é interessante pois a leitura dos padrões ocidentais é orientada desta forma.*

- Outras sugestões apontam o uso de transformações da visão e/ou formas de interatividade. Essas sugestões encontram-se disponíveis no protótipo funcional. A primeira delas é um tipo de transformação da visão denominado investigação de local e foi apresentado na seção 2.3.1.

- *Apenas para navegação, acredito que, ao se passar o mouse sobre uma imagem, um zoom poderia auxiliar a visualização da imagem de interesse naquele momento.*
- *A possibilidade de interação é uma boa ideia.*
- *Deve-se focar no tamanho da imagem e no campo visual.*

4.5 Considerações finais

Neste capítulo foram apresentados os três experimentos feitos com usuários para avaliar as técnicas de visualização estudadas em distintas etapas do trabalho. Apresentou-se o planejamento, procedimento, resultados obtidos e a análise correspondente para cada experimento desenvolvido. O próximo capítulo apresenta as conclusões finais desta dissertação.

Capítulo 5

Conclusões

5.1 Conclusões

Ao lidar com uma grande quantidade de imagens, como por exemplo o resultado de uma consulta CBIR, os métodos de visualização tornam-se ferramentas indispensáveis para extrair informação útil. Esses métodos auxiliarão ao usuário na exploração de grandes quantidades de informação fornecendo um universo entendível e coerente do conjunto de imagens.

Este trabalho realizou um estudo comparativo de diversas técnicas de visualização procurando achar as características chaves necessárias para uma correta cognição de um grande conjunto de imagens. Após uma primeira etapa deste estudo, revelaram-se três hipóteses: **(i)** Uma técnica suportada por uma estrutura com enfoque central proveria uma distribuição mais entendível do conjunto de dados. **(ii)** O uso de técnicas que mantenham as relações intrínsecas entre as imagens ajuda na visualização de imagens dado que fornece um universo mais coerente de imagens. Isto permitiria visualizar não somente imagens similares ao padrão de consulta, mas também seria possível preservar as relações entre as imagens do conjunto (conjunto coerente como um todo), mostrando-se imagens similares próximas no espaço físico. **(iii)** A criação de uma técnica que combine as vantagens das técnicas dos pontos (i) e (ii) forneceria uma abordagem mais eficiente na visualização do universo de imagens. Esta técnica combinaria a disposição estruturada das técnicas com enfoque central em conjunto com a coerência fornecida pelas técnicas que mantêm relações intrínsecas entre imagens. Para validar estas hipóteses, implementaram-se e posteriormente avaliaram-se diversas técnicas. Especificamente para validar a hipótese (iii), foi proposta, implementada e avaliada uma técnica de visualização denominada Clustering Set.

Considerou-se adequado realizar testes experimentais com usuários durante o ciclo de desenvolvimento deste trabalho. Estes testes foram úteis para guiar os passos da pesquisa e orientar o seu rumo. Assim, foram realizadas três avaliações.

O primeiro experimento teve como objetivo avaliar um conjunto das técnicas de visualização estudadas, visando descobrir técnicas promissoras para as quais foram dirigidos os seguintes passos da pesquisa. Dentre os principais resultados, destaca-se que os métodos com enfoque central obtiveram boas qualificações validando assim a hipótese (i). Observou-se também, que o método Clustering Set teve as melhores avaliações (so ou em empate) apresentando uma grande aceitação e preferência dos usuários participantes o que validaria a hipótese (iii) e confirmaria novamente a hipótese (i) dado que esta técnica possui enfoque no centro e fornece bom entendimento. Observou-se também, que o método Incremental Board, sendo um método que mantém as relações intrínsecas entre as imagens, foi o método pior avaliado. Segundo as opiniões dos usuários os principais motivos da baixa avaliação foram o reduzido tamanho das imagens que permite ver uma maior quantidade delas (muita informação) e a não sinalização das imagens relevantes e irrelevantes. Portanto considera-se que a hipótese (ii) (técnicas que mantenham as relações intrínsecas entre as imagens ajuda a visualização de imagens) nem sempre é verdade. Técnicas que mantêm relações intrínsecas entre as imagens podem ajudar o processo de visualização somente em combinação com estruturas visuais que permitam apreciar e aproveitar a principal vantagem de manter relações intrínsecas entre imagens (universo coerente de imagens). As técnicas que mantêm relações intrínsecas entre imagens podem ser usadas em conjunto com outras técnicas que forneçam um espaço confortável como as técnicas com enfoque central (como foi mencionado anteriormente). Os usuários apontaram também, nas suas sugestões, a criação de outras técnicas híbridas baseadas no método Clustering Set proposto. Assim, outras duas técnicas foram propostas se baseando nesse método. Ambas foram técnicas híbridas que misturavam as técnicas melhor avaliadas no primeiro experimento.

Posteriormente desenvolveu-se um segundo experimento, no qual as imagens foram substituídas por células vazias para avaliar as estruturas visuais e não os algoritmos de baixo nível do processo CBIR aplicado às imagens. Neste segundo experimento, as estruturas melhor avaliadas, considerando a média, foram Anéis concêntricos e Espiral. As técnicas propostas não foram tão bem sucedidas como no primeiro experimento. Para descobrir os possíveis motivos desse fato, discutiram-se as principais diferenças entre os dois experimentos, encontrando: a retirada das imagens e a utilização de papel como médio para avaliar as estruturas. O papel diminui significativamente a qualidade dos objetos representados, em comparação, por exemplo, a uma tela de computador. Ele oferece ao usuário menos recursos, como menor quantidade de cores e impossibilidade de ampliar as imagens. Além disso, a retirada das imagens possivelmente acarretou que os usuários não percebessem a distribuição por categorias ou grupos, onde a similaridade das imagens dentro de cada grupo é alta e a similaridade entre grupos distintos é baixa. Isto indicaria que as estruturas propostas não eram tão intuitivas como se tinha pensado. Para separar o real efeito do uso do papel, decidiu-se realizar um terceiro experimento no qual o meio de avaliação fosse o computador.

Antes da realização do terceiro experimento, analisaram-se também as técnicas de melhor desempenho no segundo experimento, como Anéis concêntricos e Espiral. Anéis concêntricos é a única estrutura que usa cores no fundo e não somente ao redor das imagens. Imagens menos relevantes se encontram na região mais escura da estrutura. Isto reforçou a importância de marcas visuais para auxiliar no *design* de estruturas. Portanto, adicionaram-se um conjunto de mudanças visuais relativas a cor, tamanho, espessura das arestas, etc, a fim de melhorar o grau de entendimento do usuário sobre o conjunto de estruturas.

Finalmente um terceiro experimento foi efetuado com a realimentação obtida pelos experimentos anteriores. As estruturas melhor avaliadas, em média, foram a estrutura híbrida Clustering Set - Espiral e a estrutura Espiral.

Neste experimento, as técnicas propostas obtiveram uma melhora significativa em relação ao segundo experimento, porém os resultados ainda diferem levemente dos resultados do primeiro experimento. Um fator a considerar é a diferença de perguntas entre o primeiro e os demais experimentos. No primeiro experimento se perguntou a opinião do usuário sobre os diversos fatores avaliados como entendimento, satisfação, etc, enquanto que para o segundo e terceiro experimento fizeram-se um conjunto de perguntas mais específicas para avaliar o entendimento das estruturas (localização de células específicas), como pedir achar a terceira célula mais relevante. Porém é importante considerar que as diversas estruturas têm diferentes formas de mostrar o conjunto de imagens. As técnicas Grid, Anéis Concêntricos e Espiral mostram um conjunto ordenado de imagens, onde é mais trivial achar qualquer imagem no conjunto. As técnicas baseadas em agrupamento com enfoque central, ou seja as técnicas propostas, têm uma distribuição baseada em grupos, portanto possivelmente as perguntas adequadas para um experimento futuro seriam sobre a relevância dos grupos e não das imagens. E o método Incremental Board, tem uma distribuição em grade mantendo as relações intrínsecas entre as imagens sem enfoque central. Neste último método, inicialmente não existia nenhuma marca que diferenciasse a imagem mais relevante nem as imagens irrelevantes. Assim, adicionou-se uma modificação na cor da célula da imagem mais relevante, e as mais próximas a ela. Porém ainda é difícil achar as imagens menos relevantes. Um eventual futuro experimento deverá considerar estes fatores.

Recomenda-se a estrutura híbrida Clustering Set - Espiral para a visualização de resultados de sistemas CBIR. Esta técnica respeita as relações intrínsecas entre as imagens agrupando imagens parecidas em grupos distribuídos segundo a trajetória de uma curva espiral. Também respeita o enfoque baseado no centro dando comodidade ao usuário para explorar o universo de imagens de uma forma mais intuitiva e natural. Outra vantagem desta estrutura híbrida é a ausência de sobreposição, fator importante para a correta exploração dos resultados de um sistema CBIR.

Cabe destacar que nos experimentos se avaliaram somente as estruturas visuais num contexto estático, porém o contexto real é dinâmico permitindo a interação com o sistema.

Além disso existem diversos fatores que podem afetar os valores dos resultados como o estado de ânimos dos usuários a experiência prévia de cada um, o perfil, etc. A análise estatística é imprescindível na realização de experimentos, mas é difícil conseguir medir todos os fatores que intervêm no real contexto de uma avaliação com usuários.

5.2 Contribuições

As principais contribuições deste trabalho são:

- Estudo comparativo e análise de sete técnicas de visualização, quatro delas existentes na literatura e três novas técnicas propostas.
- Validação de duas técnicas da literatura nunca antes avaliadas [5]. Essas técnicas proveem uma distribuição com enfoque central, baseando-se respectivamente em uma estrutura espiral e um conjunto de anéis concêntricos.
- Especificação e implementação de três novas técnicas de visualização: Uma técnica de visualização baseada em agrupamento denominada Clustering Set; uma técnica híbrida de visualização baseada na técnica Clustering Set e outra técnica recentemente proposta denominada Espiral [5]; e uma técnica híbrida também baseada na técnica Clustering Set e em outra técnica da literatura denominada Anéis Concêntricos [5].
- Especificação e implementação de um *framework* para desenvolvimento de novas estruturas visuais para sistemas de recuperação de imagens por conteúdo, no qual foram implementadas as técnicas de visualização estudadas.

5.3 Trabalhos futuros

Um próximo passo desta pesquisa seria a avaliação com um protótipo que permita a interação com o usuário. Este protótipo já se encontra funcional, porém precisam ser adicionados mecanismos de *relevance feedback* considerando diversos descritores. O usuário seria capaz de marcar as imagens relevantes e não relevantes e o sistema retornaria um novo conjunto de resultados a serem distribuídos na estrutura visual.

Sugere-se o estudo de outras técnicas de agrupamento para melhorar o método proposto. Experimentos de desempenho sobre estas técnicas de agrupamento podem ser conduzidos para avaliar o seu desempenho sobre grandes coleções de imagens.

Pode-se aprofundar também o estudo de estruturas fractais para a criação de técnicas novas. Planeja-se fazer algumas mudanças à estrutura Clustering set (variando a espessura

e cor das arestas em todos os níveis) para torna-la uma estrutura fractal. Após isso, poderia se avaliar, em futuros experimentos, se essa nova variação representa uma melhora sobre o método Clustering Set aqui proposto.

Podem-se acoplar as técnicas aqui estudadas com a ferramenta EVA [91]. Eva é uma ferramenta que possui um amplo conjunto de descritores para extrair características primitivas das imagens. Acoplando essa ferramenta poderia-se enriquecer o processo CBIR possibilitando a melhoria da qualidade das respostas ao usuário. Pode-se também fornecer ao usuário a possibilidade de escolher o descritor ou conjunto de descritores a serem usados no processo CBIR desde a interface, e assim, fariam-se também avaliações com usuários com foco nos descritores e algoritmos de baixo nível

O presente trabalho também pode-se acoplar com o RFCore [92]. RFCore, é um arcabouço genérico para atividades de realimentação de relevância na manipulação de objetos digitais. Poderiam se usar as estruturas visuais aqui fornecidas para apresentar e explorar o conjunto de objetos digitais possibilitando a sua manipulação mediante diversos recursos como a característica de *drag and drop*, já implementada no protótipo atual.

Pode-se estender as técnicas de visualização aqui estudadas para a visualização e exploração de outro tipo de objetos digitais como vídeo. Cada imagem representaria um vídeo e ao escolher um vídeo poderia se estender a informação de esse determinado elemento (*Details on demand*) sendo mostrada uma janela com um reprodutor de vídeo, os frames do vídeo e um conjunto de frames representativos que resumam o vídeo. Poderiam-se explorar outras formas de navegação entre vídeos ou entre os frames representativos de um mesmo vídeo.

Podem-se explorar e avaliar diversas combinações das técnicas estudadas aproveitando-se da flexibilidade do framework desenvolvido como visto na seção 3.4.

Referências Bibliográficas

- [1] TORRES, R. da S.; FALCÃO, A. X. Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada*, v. 13, n. 2, p. 161–185, 2006.
- [2] CARD, S. K.; MACKINLAY, J. D.; SHNEIDERMAN, B. (Ed.). *Readings in information visualization: using vision to think*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999. ISBN 1-55860-533-9.
- [3] AHLBERG, C.; TRUVÉ, S. Tight coupling: guiding user actions in a direct manipulation retrieval system. In: *HCI '95: Proceedings of the HCI'95 conference on People and computers X*. New York, NY, USA: Cambridge University Press, 1995. p. 305–321. ISBN 0-521-56729-7.
- [4] Google Earth. Accessed on 08/2008. [Http://www.ssec.wisc.edu/media/spotlight/google-earth.htm](http://www.ssec.wisc.edu/media/spotlight/google-earth.htm).
- [5] TORRES, R. da S. et al. Visual structures for image browsing. In: *CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management*. New York, NY, USA: ACM, 2003. p. 49–55. ISBN 1-58113-723-0.
- [6] FLICKNER, M. et al. Query by image and video content: The QBIC system. *IEEE Computer*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 28, n. 9, p. 23–32, 1995. ISSN 0018-9162.
- [7] MUKHERJEA, S.; HIRATA, K.; HARA, Y. Amore: A world wide web image retrieval engine. *World Wide Web*, Kluwer Academic Publishers, Hingham, MA, USA, v. 2, n. 3, p. 115–132, 1999. ISSN 1386-145X.
- [8] MAYRON, L. M. et al. A forward-looking user interface for cbir and cfir systems. In: *ISM '06: Proceedings of the Eighth IEEE International Symposium on Multimedia*. Washington, DC, USA: [s.n.], 2006. p. 779–780. ISBN 0-7695-2746-9.

- [9] KAJIYAMA, T.; KAND, N.; SATOH, S. Examination and enhancement of a ring-structured graphical search interface based on usability testing. In: *SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: [s.n.], 2005. p. 623–624. ISBN 1-59593-034-5.
- [10] NGUYEN, G. P.; WORRING, M. Interactive access to large image collections using similarity-based visualization. *Journal of Visual Languages and Computing*, Academic Press, Inc., Orlando, FL, USA, v. 19, n. 2, p. 203–224, 2008. ISSN 1045-926X.
- [11] MOGHADDAM, B. et al. Pdh: a human-centric interface for image libraries. In: *Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on*. [S.l.: s.n.], 2002. v. 1, p. 901–904 vol.1.
- [12] PINHO, R. *Espaço incremental para a mineração visual de conjuntos dinâmicos de documentos*. Tese (Doutorado) — Doutorado em Ciências da Computação e Matemática Computacional, São Carlos, SP, 2009.
- [13] GOOGLE. *Google Image Search*. Accessed on 02/2009. [Http://images.google.com/](http://images.google.com/).
- [14] YAHOO. *Yahoo! Image Search*. Accessed on 02/2009. [Http://images.search.yahoo.com/](http://images.search.yahoo.com/).
- [15] ALTAVISTA. Accessed on 02/2009. [Http://www.altavista.com/image/](http://www.altavista.com/image/).
- [16] CHANG, H.-T. Web image retrieval systems with automatic web image annotating techniques. *WSEAS Transactions on Information Science and Applications*, World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, v. 5, n. 8, p. 1313–1322, 2008. ISSN 1790-0832.
- [17] JEON, J.; LAVRENKO, V.; MANMATHA, R. Automatic image annotation and retrieval using cross-media relevance models. In: *SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. New York, NY, USA: ACM, 2003. p. 119–126. ISBN 1-58113-646-3.
- [18] DATTA, R. et al. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computer Survey*, ACM, New York, NY, USA, v. 40, n. 2, p. 1–60, 2008. ISSN 0360-0300.
- [19] SMEULDERS, A. W. M. et al. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Washington, DC, USA, v. 22, n. 12, p. 1349–1380, 2000.

- [20] VASCONCELOS, N. From pixels to semantic spaces: Advances in content-based image retrieval. *Computer*, v. 40, n. 7, p. 20–26, July 2007. ISSN 0018-9162.
- [21] TORRES, R. da S. et al. Recuperação de imagens: Desafios e novos rumos. In: *Seminário Integrado de Software e Hardware (SEMISH)*. [S.l.: s.n.], 2008.
- [22] TORRES, R. da S. *Ambiente de Gerenciamento de Imagens e Dados Espaciais para Desenvolvimento de Aplicações em Biodiversidade*. Tese (Doutorado) — Instituto de Computação, Unicamp, Campinas, SP, October 2004.
- [23] OLIVEIRA, M. C.; CIRNE, W.; MARQUES, P. M. de A. Towards applying content-based image retrieval in the clinical routine. *Future Generation Computer Systems*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 23, n. 3, p. 466–474, 2007. ISSN 0167-739X.
- [24] BREITENEDER, C.; EIDENBERGER, H. Content-based image retrieval and digital libraries. In: *Kyoto Int. Conference on Digital Libraries*. [S.l.: s.n.], 2000. p. 67–74.
- [25] CHRISTEL, M. G.; CONESCU, R. M. Addressing the challenge of visual information access from digital image and video libraries. In: *JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*. New York, NY, USA: ACM, 2005. p. 69–78. ISBN 1-58113-876-8.
- [26] WANG, S.; WANG, S. *A robust CBIR approach using local color histograms*. [S.l.], 2001.
- [27] ALMEIDA, J. et al. Making colors worth more than a thousand words. In: *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*. Fortaleza, Ceará, Brazil: [s.n.], 2008. p. 1180–1186. ISBN 978-1-59593-753-7.
- [28] HUANG, P. W.; DAI, S. K. Image retrieval by texture similarity. *Pattern Recognition*, v. 36, n. 3, p. 665–679, 2003. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/B6V14-4619FY6-1/2/a5b5b49b0a315bd61edfb849d9430662>>.
- [29] LIN, H.-J. et al. A study of shape-based image retrieval. *International Conference on Distributed Computing Systems Workshops*, IEEE Computer Society, Los Alamitos, CA, USA, v. 1, p. 118–123, 2004.
- [30] AMORES, J. et al. Boosting contextual information in content-based image retrieval. In: *MIR '04: Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*. New York, NY, USA: ACM, 2004. p. 31–38. ISBN 1-58113-940-3.

- [31] SIVIC, J.; ZISSERMAN, A. Video google: A text retrieval approach to object matching in videos. In: *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*. Washington, DC, USA: IEEE Computer Society, 2003. p. 1470. ISBN 0-7695-1950-4.
- [32] CHEN, C.; GAGAUDAKIS, G.; ROSIN, P. Content-based image visualization. In: *IV '00: Proceedings of the International Conference on Information Visualisation*. Washington, DC, USA: [s.n.], 2000. p. 13. ISBN 0-7695-0743-3.
- [33] QIU, G.; MORRIS, J.; FAN, X. Visual guided navigation for image retrieval. *Pattern Recognition*, Elsevier Science Inc., New York, NY, USA, v. 40, n. 6, p. 1711–1721, 2007. ISSN 0031-3203.
- [34] MOGHADDAM, B. et al. Visualization and user-modeling for browsing personal photo libraries. *International Journal of Computer Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 56, n. 1-2, p. 109–130, 2004. ISSN 0920-5691.
- [35] RUI, Y. et al. Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 8, p. 644–655, 1998.
- [36] ZHOU, X. S.; HUANG, T. S. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, v. 8, n. 6, p. 536–544, April 2003. Disponível em: <<http://dx.doi.org/10.1007/s00530-002-0070-3>>.
- [37] SANTINI, S.; GUPTA, A.; JAIN, R. Emergent semantics through interaction in image databases. *IEEE Transactions on Knowledge and Data Engineering*, v. 13, n. 3, p. 337–351, May/June 2001. ISSN 1041-4347.
- [38] KAMPFE, T. et al. Indi - intelligent database navigation by interactive and intuitive content-based image retrieval. In: *Proceedings of International Conference on Image Processing*. [S.l.: s.n.], 2002. v. 3, p. 921–924 vol.3. ISSN 1522-4880.
- [39] LUO, J.; NASCIMENTO, M. A. Content-based sub-image retrieval using relevance feedback. In: *MMDB '04: Proceedings of the 2nd ACM international workshop on Multimedia databases*. New York, NY, USA: [s.n.], 2004. p. 2–9. ISBN 1-58113-975-6.
- [40] WONG, K.-M.; CHEUNG, K.-W.; PO, L.-M. Mirror: an interactive content based image retrieval system. In: *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*. [S.l.: s.n.], 2005. p. 1541–1544 Vol. 2.

- [41] BACH, J. et al. Virage image search engine: an open framework for image management. In: *Proceedings of the SPIE, Storage and Retrieval for Image and Video Databases IV*. San Jose, CA: [s.n.], 1996. p. 76–87.
- [42] NIBLACK, W. et al. Updates to the qbic system. In: *Proceedings of Storage and Retrieval for Image and Video Databases VI*. San Jose, California, USA: SPIE, 1997. p. 150–161.
- [43] PETKOVIC, D. et al. Recent applications of ibm’s query by image content (qbic). In: *SAC ’96: Proceedings of the 1996 ACM symposium on Applied Computing*. New York, NY, USA: ACM, 1996. p. 2–6. ISBN 0-89791-820-7.
- [44] MUKHERJEA, S.; HIRATA, K.; HARA, Y. Towards a multimedia world-wide web information retrieval engine. In: *Selected papers from the sixth international conference on World Wide Web*. Essex, UK: Elsevier Science Publishers Ltd., 1997. p. 1181–1191.
- [45] M. MANOLA L., H. T. N. Imagegroup: a group-oriented user interface for content-based image retrieval and digital image arrangement. *Journal of Visual Languages & Computing*, v. 14, p. 363–386, August 2003. Disponível em: <<http://www.ingentaconnect.com/content/els/1045926x/2003/00000014/00000004/art00034>>.
- [46] HUANG, T. et al. Learning in content-based image retrieval. In: *ICDL ’02: Proceedings of the 2nd International Conference on Development and Learning*. Washington, DC, USA: IEEE Computer Society, 2002. p. 155–162. ISBN 0-7695-1459-6.
- [47] VELTKAMP, R. C.; TANASE, M. *Content-Based Image Retrieval Systems: A Survey*. [S.l.], 2002.
- [48] VENTERS, C. C.; COOPER, D. M. *A Review of Content-Based Image Retrieval Systems*. [S.l.], 2000.
- [49] TUFTE, E. R. *The Visual Display of Quantitative Information*. Second. [S.l.]: Graphics Press, 2001. ISBN 0961392142.
- [50] Information Visualization Research Group at University of California, I. *Information Visualization Research*. Accessed on 08/2010. [Http://www.isr.uci.edu/research-visualization.html](http://www.isr.uci.edu/research-visualization.html).
- [51] SHNEIDERMAN, B. The eyes have it: A task by data type taxonomy for information visualizations. In: *VL ’96: Proceedings of the 1996 IEEE Symposium on Visual Languages*. Washington, DC, USA: IEEE Computer Society, 1996. p. 336. ISBN 0-8186-7508-X.

- [52] KEIM, D. A. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 8, n. 1, p. 1–8, 2002. ISSN 1077-2626.
- [53] SILVA, C. G. aes da. *Exploração de bases de dados de ambientes de Educação a Distância por meio de ferramentas de consulta apoiadas por Visualização de Informação*. Tese (Doutorado) — Intitute of Computing, UNICAMP., 2006.
- [54] JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM Computing Surveys*, ACM, New York, NY, USA, v. 31, n. 3, p. 264–323, 1999. ISSN 0360-0300.
- [55] KUNCHEVA, L. I. *Combining Pattern Classifiers: Methods and Algorithms*. [S.l.]: Wiley-Interscience, 2004. ISBN 0471210781.
- [56] MACQUEEN, J. B. Some methods for classification and analysis of multivariate observations. In: BERKELEY UNIVERSITY OF CALIFORNIA PRESS, .-. (Ed.). *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. [S.l.: s.n.], 1967.
- [57] JAIN, A. K. Data clustering: 50 years beyond k-means. *Pattern Recogn. Lett.*, Elsevier Science Inc., New York, NY, USA, v. 31, n. 8, p. 651–666, 2010. ISSN 0167-8655.
- [58] THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition, Third Edition*. Orlando, FL, USA: Academic Press, Inc., 2006. ISBN 0123695317.
- [59] ROCHA, A. et al. Efficient and flexible cluster-and-search for cbir. In: *ACIVS '08: Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems*. [S.l.: s.n.], 2008. p. 77–88. ISBN 978-3-540-88457-6.
- [60] GUERIN-DUGUE, A.; AYACHE, S.; BERRUT, C. Image retrieval: a first step for a human centered approach. In: *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*. [S.l.: s.n.], 2003. v. 1, p. 21–25 Vol.1.
- [61] DRAPER, G. M.; LIVNAT, Y.; RIESENFELD, R. F. A survey of radial methods for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 15, p. 759–776, September 2009. ISSN 1077-2626. Disponível em: <<http://portal.acm.org/citation.cfm?id=1591897.1592278>>.

- [62] FLANK, S. et al. Photofile: A digital library for image retrieval. In: SOCIETY, I. C. (Ed.). *Proceedings of International Conference on Multimedia Computing and Systems*. Washington, DC: [s.n.], 1995. p. 292–295.
- [63] BARTOLINI, I.; CIACCIA, P.; PATELLA, M. The pipe personalizable image browsing engine. In: *CVDB '04: Proceedings of the 1st international workshop on Computer vision meets databases*. New York, NY, USA: [s.n.], 2004. p. 43–50. ISBN 1-58113-917-9.
- [64] BHATIA, S. Hierarchical clustering for image databases. In: *Electro Information Technology, 2005 IEEE International Conference on*. [S.l.: s.n.], 2005. p. 6 pp.–6.
- [65] CHEN, Y.; WANG, J. Z.; KROVETZ, R. CLUE: Cluster-based retrieval of images by unsupervised learning. *IEEE Transactions on Image Processing*, v. 14, p. 1187–1201, 2004. Disponível em: <citeseer.comp.nus.edu.sg/691021.html>.
- [66] CAI, D. et al. Hierarchical clustering of www image search results using visual, textual and link information. In: *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*. New York, NY, USA: ACM, 2004. p. 952–959. ISBN 1-58113-893-8.
- [67] ZHENG, X. et al. Locality preserving clustering for image database. In: *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*. New York, NY, USA: ACM, 2004. p. 885–891. ISBN 1-58113-893-8.
- [68] LAAKSONEN, J.; KOSKELA, M.; OJA, E. Self-organising maps as a relevance feedback technique in content-based image retrieval. *Pattern Analysis and Applications*, v. 4, p. 140–152, Jul 2001.
- [69] RODDEN, K. et al. Does organisation by similarity assist image browsing? In: *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 2001. p. 190–197. ISBN 1-58113-327-8.
- [70] OGLE, V. E.; STONEBRAKER, M. Chabot: Retrieval from a relational database of images. *Computer*, IEEE Computer Society, Los Alamitos, CA, USA, v. 28, p. 40–48, 1995. ISSN 0018-9162.
- [71] ASLANDOGAN, Y. A.; YU, C. T. Multiple evidence combination in image retrieval: Diogenes searches for people on the web. In: *SIGIR '00: Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: ACM, 2000. p. 88–95. ISBN 1-58113-226-3.

- [72] ORTEGA, M. et al. Supporting similarity queries in mars. In: *MULTIMEDIA '97: Proceedings of the fifth ACM international conference on Multimedia*. New York, NY, USA: ACM, 1997. p. 403–413. ISBN 0-89791-991-2.
- [73] BEIGI, M. et al. Metaseek: A content-based meta-search engine for images. In: *Proceedings of the SPIE Storage and Retrieval for Image and Video Databases*. [S.l.: s.n.], 1997. p. 118–128.
- [74] MA, W.-Y.; MANJUNATH, B. S. Netra: a toolbox for navigating large image databases. *Multimedia Systems*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, v. 7, n. 3, p. 184–198, 1999. ISSN 0942-4962.
- [75] PENTLAND, A.; PICARD, R. W.; SCLAROFF, S. Photobook: content-based manipulation of image databases. *International Journal of Computer Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 18, n. 3, p. 233–254, 1996. ISSN 0920-5691.
- [76] LAAKSONEN, J. et al. Picsom—content-based image retrieval with self-organizing maps. *Pattern Recognition Letters*, Elsevier Science Inc., New York, NY, USA, v. 21, n. 13-14, p. 1199–1207, 2000. ISSN 0167-8655.
- [77] FOURNIER, J. et al. Retin: A content-based image indexing and retrieval system. *Pattern Analysis and Applications*, Springer London, v. 4, p. 153–173, 2001. ISSN 1433-7541 (Print) 1433-755X (Online).
- [78] CORD, M. et al. Interactive exploration for image retrieval. *EURASIP Journal on Applied Signal Processing*, Hindawi Publishing Corp., New York, NY, United States, v. 2005, p. 2173–2186, 2005. ISSN 1110-8657.
- [79] GONY, J. et al. Retin: a smart interactive digital media retrieval system. In: *CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval*. New York, NY, USA: ACM, 2007. p. 93–96. ISBN 978-1-59593-733-9.
- [80] SQUIRE, D. M. et al. Content-based query of image databases: inspirations from text retrieval. *Pattern Recognition Letters*, v. 21, n. 13-14, p. 1193–1198, 2000. ISSN 0167-8655. Selected Papers from The 11th Scandinavian Conference on Image. Disponível em: <<http://www.sciencedirect.com/science/article/B6V15-48CPF5G-5T/2/4f79ba3c43823bd5439cd4e682865e86>>.
- [81] SMITH, J. R.; CHANG, S.-F. Visualseek: a fully automated content-based image query system. In: *MULTIMEDIA '96: Proceedings of the fourth ACM international conference on Multimedia*. New York, NY, USA: ACM, 1996. p. 87–98. ISBN 0-89791-871-1.

- [82] ROCHA, H. V.; BARANAUSKAS, M. C. C. *Design e Avaliação de Interfaces Humano-Computador*. Campinas, SP, Brasil: EdUnicamp, 2003.
- [83] PAL, N.; BEZDEK, J. On cluster validity for the fuzzy c-means model. In: *Fuzzy Systems, IEEE Transactions on*. [S.l.: s.n.], 1995.
- [84] GAMMA RICHARD HELM, R. J. E.; VLISSIDES, J. *Design Patterns: Elements of Reusable Object-Oriented Software*. [S.l.]: Addison Wesley, 1994.
- [85] DIX, A. et al. *Human Computer Interaction*. 3rd ed.. ed. Harlow, England: Pearson/Prentice-Hall, 2004. ISBN 0-13-046109-1.
- [86] PREECE, J.; ROGERS, Y.; SHARP, H. *Interaction Design*. New York, NY, USA: John Wiley & Sons, Inc., 2002. ISBN 0471492787.
- [87] QUESTIONNAIRE for User Interaction Satisfaction. Accessed on 08/2010. [Http://lap.umd.edu/QUIS](http://lap.umd.edu/QUIS).
- [88] SOFTWARE Usability Measurement Inventory. Accessed on 08/2010. [Http://sumi.ucc.ie/](http://sumi.ucc.ie/).
- [89] Web Center for social research methods. *Likert Scaling*. Accessed on 10/2010. [Http://www.socialresearchmethods.net/kb/scallik.php](http://www.socialresearchmethods.net/kb/scallik.php).
- [90] LEWIS, C. *Using the "thinking-aloud" method in cognitive interface design*. [S.l.], 1982.
- [91] PENATTI, O. A.; TORRES, R. da S. Eva: an evaluation tool for comparing descriptors in content-based image retrieval tasks. In: *MIR '10: Proceedings of the international conference on Multimedia information retrieval*. New York, NY, USA: ACM, 2010. p. 413–416. ISBN 978-1-60558-815-5.
- [92] CALUMBY, R. T. *Recuperação Multimodal de Imagens com Realimentação de Relevância Baseada em Programação Genética*. Dissertação (Mestrado) — Universidade Estadual de Campinas, 2010.