

UNIVERSIDADE DE SÃO PAULO ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

THIAGO KOBASHIGAWA AMORIM

Proposta de um sistema CBIR baseado em ordem

THIAGO KOBASHIGAWA AMORIM

Proposta de um sistema CBIR baseado em ordem

Versão corrigida

Dissertação apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação.

Área de concentração: Metodologia e Técnicas da Computação

Orientador: Prof. Dr. Helton Hideraldo Bíscaro

São Paulo

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha catalográfica elaborada pela Biblioteca da Escola de Artes, Ciências e Humanidades, com os dados inseridos pelo(a) autor(a)

Brenda Fontes Malheiros de Castro CRB 8-7012; Sandra Tokarevicz CRB 8-4936

```
Kobashigawa Amorim, Thiago
   Proposta de um sistema CBIR baseado em ordem /
Thiago Kobashigawa Amorim; orientador, Helton
Hideraldo Biscaro. -- São Paulo, 2022.
   68 p: il.
```

Dissertacao (Mestrado em Ciencias) - Programa de Pós-Graduação em Sistemas de Informação, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, 2022.

Versão corrigida

1. CBIR 3D. 2. Extrator espectral. 3. Malha. 4. SPHARM. 5. Ordem. I. Biscaro, Helton Hideraldo, orient. II. Título.

${\bf Agradecimentos}$

Ao meu orientador, pois sem ele o trabalho simplesmente não existiri	a.
--	----

Resumo

AMORIM, Thiago Kobashigawa. **Proposta de um sistema de CBIR baseado em ordem**. 2022. 68 f. Dissertação (Mestrado em Ciências) — Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2022.

Em computação gráfica, a recuperação de imagem por conteúdo, CBIR (*Content Based Image Retrieval*), é um sistema de banco de dados que consiste em receber uma imagem como entrada e devolver uma lista de imagens similares. As imagens usadas podem ser modelos tridimensionais, como uma malha, transformando o CBIR em um contexto 3D. Para avaliar a similaridade entre as malhas, é comum utilizar técnicas de relações de distâncias entre as malhas a fim de mapear as mais semelhantes.

Contudo, neste trabalho, medidas de distância não tiveram informação suficiente para ter uma recuperação assertiva no CBIR 3D. Feito testes com relação de ordem, foi possível ter uma precisão melhor da recuperação do conteúdo. Portanto, o objetivo deste trabalho é otimizar a fase de comparação entre as malhas utilizando uma relação de ordem ao invés de distância.

Para realizar as comparações, foram utilizadas duas técnicas de extração das características das malhas: o descritor espectral e o SPHARM (harmônicos esféricos). Para a relação da distância, foram utilizadas a distância Euclidiana e a de Manhattan. Por fim, para a relação das ordens, utilizou-se a ordem lexicográfica, ordem revlex e uma ordem sugerida, a ordem lexicográfica estendida, que obteve melhor desempenho.

Palavras-chaves: CBIR 3D. Extrator Espectral. Malha. SPHARM. Ordem.

Abstract

AMORIM, Thiago Kobashigawa. **Proposed order-based CBIR system**. 2022. 68 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2022.

In computer graphics, content-based image retrieval, CBIR (Content Based Image Retrieval), is a database system that consists of receiving an image as input and returning a list of similar images. The images used can be three-dimensional models, such as a mesh, transforming the CBIR into a 3D context. To assess the similarity between meshes, it is common to use distance relationship techniques between meshes in order to map the most similar ones.

However, in this work, distance measurements did not have enough information to have an assertive recovery in CBIR 3D. By doing tests with order relationship, it was able to have a better accuracy of content retrieval. Therefore, the objective of this work is to optimize the comparison phase between the meshes using an order relationship instead of distance.

To carry out the comparisons, it was used two techniques for extracting the characteristics of the meshes: the spectral descriptor and the SPHARM (spherical harmonics). For the distance relationship, it was used the Euclidean and Manhattan distances. Finally, for the order list, we used the lexicographical order, revlex order and an order suggested by the author, extended lexicographical order, which performed better.

Keywords: CBIR 3D. Spectral Descriptor. Mesh. SPHARM. Order.

Lista de figuras

Figura 1 – 1	Malha representativa de um polvo	17
Figura 2 – 1	Malha representativa de um óculos	17
Figura 3 - 0	Gráfico comparativo de distância entre polvo de índice 121 e cada objeto	
(da classe polvo e óculos	18
Figura 4 - '	Vetor de característica médio da classe polvo	19
Figura 5 - '	Vetor de característica médio da classe óculos	19
Figura 6 – 1	Figuras desordenadas	21
Figura 7 – 1	Figuras ordenadas	22
Figura 8 - 0	Comparação entre dois vetores pela ordem lexicográfica estendida	24
Figura 9 – 1	Detalhe da malha de um coelho	25
Figura 10 – 1	Esquema geral de um sistema CBIR	28
Figura 11 – l	Modelo exemplificativo de métricas	29
Figura 12 – I	Representação gráfica do problema algébrico	31
Figura 13 – I	Malha original à esquerda e reconstruída a partir de 200 coeficientes à	
(esquerda	32
Figura 14 – 9	Coordenadas cartesianas e coordenadas espectrais	33
Figura 15 – I	Diagrama da revisão sistemática	37
Figura 16 – I	Histograma de publicações por ano e linha de tendência	39
Figura 17 – I	Exemplos de malhas e suas classes	46
Figura 18 – 1	Descrição do fluxo de dados para o descritor espectral de $k=50$	48
Figura 19 – 1	Descrição do fluxo de dados para o descritor SPHARM	49
Figura 20 – I	Duas classes de objetos, triângulo e quadrado. Em (a) eles são desorde-	
1	nados. Em (b) eles estão ordenados	50
Figura 21 – ((a) Um novo elemento é usado como pesquisa. (b) O elemento é inserido	
1	temporariamente na lista. (c) Os 5 elementos vizinhos são retornados	50
Figura 22 – I	Exemplo de gráfico PxR	51
Figura 23 –	Gráfico da curva Precisão x Acurácia da malha 373 da base de dados	
1	utilizando o descritor espectral com 200 coeficientes	52
Figura 24 – I	Polvos x Óculos em ordem lexicográfica	53
Figura 25 – I	Polvos x Óculos em ordem lexicográfica reversa	54

Figura 26 –	Polvos x Óculos em ordem lexicográfica estendida	54
Figura 27 –	Gráfico AUC máximo dos resultados das distâncias e ordens do descritor	
	espectral, onde $k=200$ e $n=10$	55
Figura 28 –	Gráfico AUC máximo dos resultados da ordem lexicográfica estendida	
	do descritor espectral onde $k \in \{50, 100, 200\}$ e $n=10$	56
Figura 29 –	Gráfico AUC máximo da ordem lexicográfica estendida do descritor	
	espectral onde $k = 50$ e $n \in \{5, 10, 15\}$	57
Figura 30 –	Gráfico AUC máximo da ordem revlex do descritor espectral onde $\mathbf{k}=$	
	200 e n $\in \{5, 10, 15\}$	58
Figura 31 –	Gráfico AUC máximo dos resultados das distâncias e ordens do descritor	
	SPHARM onde $n = 5$	59
Figura 32 –	Gráfico AUC máximo dos resultados da ordem lexicográfica estendida	
	do descritor SPHARM onde $n \in \{5, 10, 15\} \dots$	60
Figura 33 –	Gráfico AUC máximo dos piores resultados da distância Euclidiana do	
	descritor SPHARM onde $n \in \{5, 10, 15\}$	61
Figura 34 –	Vetores médios das classes Pássaro, Alicate e Mesa	62
Figura 35 –	Gráfico AUC dos resultados da ordem lexicográfica estendida dos resul-	
	tados do descritor espectral e SPHARM onde n = $5 \dots \dots \dots$	63

Lista de algoritmos

Lista de quadros

Lista de tabelas

Tabela 1 –	Média de distância e desvio padrão entre polvo 121 <i>versus</i> sua classe e	
	polvo 121 <i>versus</i> classe óculos	18
Tabela 2 –	Publicações e técnicas usadas	42

Lista de abreviaturas e siglas

2D Bidimensional

3D Tridimensional

BD Banco de Dados

SIFT Scale-invariant feature transform/Transformação de recurso invariante

de escala

CBIR Content-based image retrieval/Recuperação de imagem por conteúdo

SVM Support Vector Machine/Máquina de vetores de suporte

3D GLCM 3D Grey Level Co-occurrence Matrices/Matrizes de coocorrência de

escala de cinza 3D

3D WT 3D Wavelet Transform/Transformada Wavelet 3D

3D GT 3D Gabor Transform/Transformada Gabor 3D

3D LBP 3D Local Binary Pattern/Padrão Binário Local 3D

VRML Virtual Reality Modeling Language / Linguagem de modelagem de reali-

dade virtual

MRMR *Minimal redundancy-maximal-relevance* / Redundância mínima-relevância

máxima

Vp Verdadeiro positivo

Vn Verdadeiro negativo

Fp Falso positivo

Vp Verdadeiro negativo

Lista de símbolos

- Φ Phi maiúsculo
- Θ Theta maiúsculo

Sumário

1	Introdução	15
1.1	Problema de pesquisa	16
1.1.1	Análise do mau desempenho do descritor espectral	16
1.2	Objetivo	19
1.3	Estrutura do documento	20
2	Conceitos fundamentais	21
2.1	Ordem	21
2.1.1	Ordem em um conjunto	21
2.1.2	Ordem Lexicográfica	22
2.1.3	Ordem lexicográfica estendida	23
2.2	Conceito de malha	24
2.2.1	Simplexo	25
2.3	CBIR 3D	27
2.4	Métricas de avaliação	28
2.5	Descritor espectral	30
2.6	SPHARM (Esféricos harmônicos)	33
3	Revisão Sistemática	36
3.1	Protocolo de Pesquisa	37
3.2	Condução e artigos obtidos	38
3.3	Publicações por ano	38
3.4	Técnicas para extração de características	36
3.5	Métricas adotadas na avaliação	42
3.6	Funções de similaridade	43
3.7	Considerações Finais	43
4	Materiais e métodos	45
4.1	Material de pesquisa	45
4.2	Método de pesquisa	46
4.2.1	Descritor espectral	47

SPHARM	47
Funções de similaridade	47
Vetores de características	48
Avaliação	49
Gráfico PxR	50
Área da curva	51
Resultados e discussões	53
Resultados com ordens para duas classes de objetos: Polvo e óculos	53
Resultados obtidos	54
Resultados do descritor espectral	55
Resultados do descritor SPHARM	58
Discussão dos resultados dos dois descritores	62
Conclusão	64
Próximos passos	64
REFERÊNCIAS	66
	Funções de similaridade Vetores de características Avaliação Gráfico PxR Área da curva Resultados e discussões Resultados com ordens para duas classes de objetos: Polvo e óculos Resultados obtidos Resultados do descritor espectral Resultados do descritor SPHARM Discussão dos resultados dos dois descritores Conclusão

1 Introdução

Com o crescente aumento e uso de imagens nos últimos anos, tem se tornado cada vez mais relevante a recuperação de imagem por conteúdo, CBIR (Content-Based Image Retrieval). Em um sistema CBIR existem quatro elementos: Técnicas para automação de extração dos metadados contendo as características primitivas da imagem; interface para obter os dados de pesquisa feita pelo usuário; métodos para comparar as similaridades/diferenças entre imagens e técnicas para indexar e armazenar de forma eficiente os metadados (WANG, 2001).

Hoje, dada uma imagem de entrada, é possível recuperar imagens relacionadas de forma eficiente de um banco de dados de grande escala. Mais especificamente na última década, novos modelos de CBIR, teorias e algoritmos foram estabelecidos (LI; YANG; MA, 2021). Profissionais de área como geografia, medicina, arquitetura, propaganda, design, moda e publicidade têm a necessidade de recuperar imagens dentro de um repositório, com o desafio de manter a eficiência (WANG, 2001). Apesar de a literatura tratar a recuperação de imagem por conteúdo, o termo "imagem" pode não necessariamente ser uma imagem 2D de fato, como neste trabalho, em que os arquivos do banco de dados são malhas. A recuperação de imagem por conteúdo consiste em receber como entrada uma imagem e retornar de saída uma ou mais imagens similares de uma base de imagens. Embora no uso de imagens bidimensionais (fotografias, quadros, mapas, entre outros) o CBIR já esteja bastante explorado, para o caso 3D a exploração é relativamente recente na literatura (BERGAMASCO; NUNES, 2019).

Uma possível representação de um objeto tridimensional é através da reconstrução da superfície do objeto via uma malha (pontos descritos por meio de coordenadas) (BERGAMASCO; NUNES, 2019). Para classificar as características de um objeto (2D ou 3D) é comum aplicar técnicas matemáticas e por fim gerar um vetor de características, que sumariza os dados do objeto em um metadado numérico, possibilitando aplicação de métricas e comparações diretas.

Uma técnica é o descritor espectral, baseado no conceito de *Compressed Sensing*, que revolucionou o campo de processamento de sinais em 2004. Com essa técnica, é possível reconstruir um sinal com um número extremamente menor do que a teoria vigente anterior permitia (CANDES; ROMBERG, 2007). Como obtido por Biscaro *et al.* (2016), para o caso

de malha representativa do ventrículo esquerdo do coração humano, o descritor espectral alcançou precisão de 85% nos melhores casos. Entretanto, em estudo em andamento, foi notado uma precisão abaixo de 17% quando o conjunto de objetos é mais heterogêneo e possui mais classes, como polvos, mesas, vasos, copos, bustos, cadeiras e entre outros. Devido a esse caso que o descritor teve mau desempenho, foi resolvido utilizar outra função de similaridade nos vetores de características. Trocou-se as comparações por meio de funções de distância para comparações por meio de ordem. Dessa forma, foi possível obter resultados muito melhores em comparação com os resultados anteriores.

Para testar a nova função de similaridade, foi utilizado além do descritor espectral outra técnica de extração das características, o SPHARM (harmônicos esféricos) (BERGA-MASCO et al., 2018), que também teve seus resultados otimizados ao se usar o sistema de ordenação. Além de usar relações de ordem como a lexicográfica, foi proposta uma ordem que leva em consideração o tamanho inteiro do vetor de características.

1.1 Problema de pesquisa

No trabalho de (BISCARO et al., 2016) foi obtido um resultado satisfatório com o descritor espectral, mas em pesquisa em andamento o descritor apresentou desempenho muito abaixo do esperado. Então surgiu a situação de desenvolver alternativas para melhorar o resultado obtido do descritor. Pensando nisso foi cogitado propor uma função de similaridade diferente na etapa de comparação do CBIR, a função de ordenação ao invés da distância. Foram feitos testes mais detalhados com malhas que o descritor espectral teve resultados ruins, em que é possível concluir que só a função distância não é suficiente para ter um resultado bom nesse caso.

1.1.1 Análise do mau desempenho do descritor espectral

Para demonstração de um caso em que a técnica atual do descritor falha, foram utilizadas duas classes de objetos bem distintas, mas para a classificação gera erro: polvo e óculos. A Figura 1 e Figura 2 demonstram as respectivas malhas desses objetos.

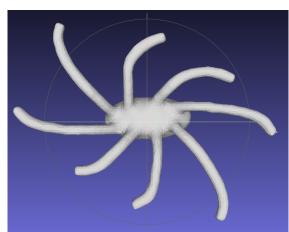
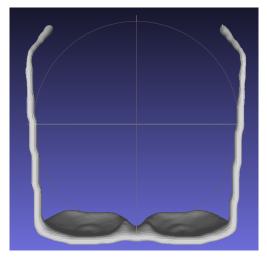


Figura 1 – Malha representativa de um polvo

Fonte - Autor



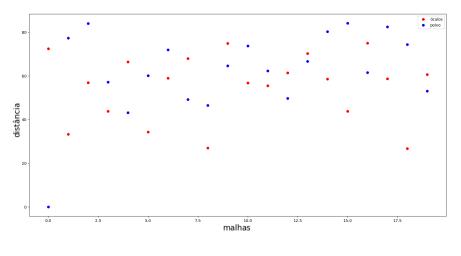


Fonte - Autor

A tomada de decisão do algoritmo é com base na distância entre os vetores de características gerado por cada malha. Para determinar de qual classe é o objeto, são realizadas comparações até encontrar a menor distância entre o vetor utilizado na busca e um vetor pré-determinado (ou seja, que já é sabido de qual classe ele pertence). Dessa forma a técnica atribui ao objeto a mesma classe do vetor de características de menor distância. Com a finalidade de saber se esse comparativo só das distâncias dos vetores é sempre válido, foi realizado o cálculo da média da distância entre um objeto e sua própria classe versus o mesmo objeto e outra classe. O objeto usado como comparativo foi o polvo da Figura 1. As classes polvo e óculos são representadas por 19 malhas distintas cada. Como demonstrado na Tabela 1, a média da distância do polvo de índice 121 é menor para a classe Óculos. Logo, o descritor iria atribuir erroneamente como Óculos a malha

que representa um polvo para 200 coeficientes. A Figura 3 ilustra como a separação entre essas duas classes não é bem definida.

Figura 3 – Gráfico comparativo de distância entre polvo de índice 121 e cada objeto da classe polvo e óculos



Fonte - Autor

Tabela 1 – Média de distância e desvio padrão entre polvo 121 *versus* sua classe e polvo 121 *versus* classe óculos

	Classe Polvo	Classe Óculos
Distância média	62,147	55,178
Desvio padrão	19,037	14,990

Fonte – Autor

Neste caso, foi encontrado um caso em que o descritor falha devido à pouca informação utilizada. Somente a distância entre os vetores de características provenientes das malhas não é informação suficiente para classificação no caso estudado. Durante essa etapa da pesquisa, também foi verificado que o descritor espectral é invariável à rotações na malha. Para ilustrar a limitação do uso apenas de métricas de distâncias, foi gerado um vetor de características médio para 20 malhas de polvos e outro vetor médio para 20 malhas de óculos. Através da média aritmética ponto a ponto dos vetores, pode-se observar como eles são similares. A Figura 4 demonstra o vetor de características médio para a classe polvo quando o coeficiente é c=200 e a Figura 5 demonstra o vetor de óculos médio com os mesmos requisitos que o polvo.

Figura 4 – Vetor de característica médio da classe polvo

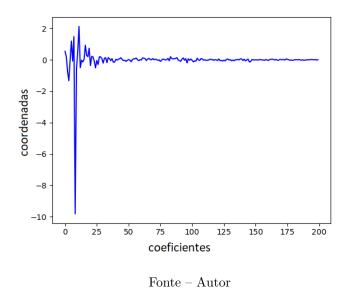
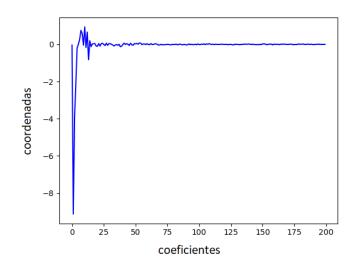


Figura 5 – Vetor de característica médio da classe óculos



Fonte – Autor

1.2 Objetivo

O objetivo deste trabalho é propor um sistema CBIR 3D baseado em uma relação de ordem proposta pelo autor. Para demonstrar os resultados, foi realizado um estudo comparativo da precisão de um CBIR usando o descritor espectral e um CBIR usando SPHARM, utilizando diferentes técnicas de função de similaridade (distância Euclidiana, Manhattan e ordenações lexicográficas, revlex e uma ordem proposta pelo autor).

1.3 Estrutura do documento

Este capítulo foi a introdução deste trabalho, em que apresentamos os tópicos principais, o problema da pesquisa e o objetivo. A disposição do restante deste documento se tratará em: Capítulo de explicação de conceitos fundamentais de matemática e processamento gráfico para compreensão desta pesquisa (Capítulo 2); descrição da revisão sistemática realizada, explorando a literatura relacionada e resultados obtidos durante a revisão para orientação desta pesquisa (Capítulo 3); materiais utilizados, metodologia e cronograma do andamento da pesquisa (Capítulo 4); resultados obtidos da pesquisa (Capítulo 5); capítulo final dedicado à conclusão obtida durante a execução deste projeto e as possibilidades de novos estudos (Capítulo 6).

2 Conceitos fundamentais

Capítulo dedicado à explicação de alguns conceitos primordiais de processamento gráfico e geometria espacial que são usados nesta pesquisa. Para a parte teórica de conceitos de geometria computacional, as definições matemáticas foram retiradas de (BISCARO, 2005).

2.1 Ordem

2.1.1 Ordem em um conjunto

Para definir o conceito de um conjunto ordenado, serão definidas regras e exemplos de ordens extraídos de (BASTOS; COSTA, 2012). Seja A um conjunto não vazio. Uma relação de ordem em um conjunto A é um subconjunto R de $A \times A$, cujo elementos satisfazem as condições:

1. Reflexiva

Para todo
$$x \in A, (x, x) \in R$$

2. Transitiva

Se
$$(x, y) \in R$$
 e $(y, z) \in R$ então $(x, z) \in R$

3. Antissimétrica

Se
$$(x, y) \in R$$
 e $(y, x) \in R$ então $x =_R y$

4. Se para quaisquer $(x, y) \in A \times A$ tivermos , $(x, y) \in R$ ou $(y, x) \in R$, então dizemos que A é um conjunto totalmente ordenado.

Se as condições 1, 2 e 3 forem satisfeitas, temos que (A,R) é um conjunto parcialmente ordenado.

A Figura 6 mostra um conjunto de elementos desordenados e a Figura 7 mostra o mesmo conjunto após passar por uma ordenação, agrupando os elementos iguais.

Figura 6 – Figuras desordenadas



Figura 7 – Figuras ordenadas



Fonte – Autor

2.1.2 Ordem Lexicográfica

Uma ordem lexicográfica é definida como a generalização da ordem alfabética dos dicionários para sequências de símbolos ordenados. Mais genericamente, dos elementos de um conjunto totalmente ordenado. Por exemplo, um conjunto de palavras (gato, cachorro, pinguim, coelho, pato, cavalo) para estar ordenado lexicograficamente precisa estar como na ordem de um dicionário: (cachorro, cavalo, coelho, gato, pato, pinguim).

Outro exemplo de ordem lexicográfica será dado com base no conjunto dos números complexos. A relação $R_1 = (a + bi, c + di) : b < d$ ou b = d e $a \le c$ é uma relação de ordem porque satisfaz todas as condições definidas anteriormente. Pode-se denominá-la como ordem lexicográfica (à direita) em \mathbb{C} . A ordem lexicográfica é uma ordem total no conjunto \mathbb{C} . Uma reta associada a R_1 pode ser imaginada interpretando os números complexos como pontos do plano cartesiano. Portanto, o conjunto dos números complexos é totalmente ordenado.

De forma genérica, a ordem lexicográfica de um conjunto de vetores finitos é definido por duas regras:

- 1. Dado dois tipos de vetores diferentes e de mesmo tamanho, $u = (u_1, u_2, ..., u_k)$ e $v = (v_1, v_2, ..., v_k)$ a ordem irá depender dos primeiros valores diferentes comparados índice a índice. Por exemplo, u < v se e somente se $u_i < v_i$, dado que i seja o primeiro valor diferente.
- 2. Se dois vetores têm tamanhos diferentes, usualmente o vetor de menor tamanho é completado com um símbolo considerado um valor menor que qualquer elemento do conjunto para cada índice faltante em relação ao vetor de maior tamanho. Uma outra variante nesse caso é simplesmente classificar o vetor de tamanho menor como menor que o vetor comparado. Essa variante é definida como shortlex order.

A seguir são citadas algumas variantes de ordens:

• Ordem lexicográfica

A ordem utilizada do dicionário e definida previamente.

• Shortlex order

Ordem característica por classificar como menor (em relação a outro vetor) a palavra que tem menos letras, ou o vetor com menos elementos.

• RevLex (Reverse Lexicographic ordering)
A ordem lexicográfica de trás pra frente.

2.1.3 Ordem lexicográfica estendida

Para este trabalho foi utilizado outro tipo de ordenação, inspirado na abordagem lexicográfica, mas que considera todas as coordenadas dos vetores. A ideia geral desta ordem é, ao comparar dois vetores u e v, usar suas coordenadas como se fossem pesos em uma escala de comparação como pode ser visto na Figura 8. A ordem lexicográfica estendida definirá o vetor v maior que u se a Equação 1 for satisfeita. Ou seja, é calculada a somatória dos elementos do vetor e o resultado do total define qual vetor é maior ou menor.

$$\left(\sum_{i=1}^{n} u_i\right) < \left(\sum_{i=1}^{n} v_i\right) \tag{1}$$

Este sistema pode gerar um empate com vetores muito diferentes, como por exemplo os vetores (1,2,3) e (3,2,1), já que ambos têm somatória igual a 6. Então foi definido um critério de desempate: Considere k a menor coordenada onde ocorreu a máxima diferença, em valor absoluto $|(u_k - v_k)|$. Se $(u_k - v_k) > 0$, o vetor u é declarado maior que v, se $(u_k - v_k) < 0$, v é declarado maior que v; Obviamente, se $\max_{i=1}^n |(u_i - v_i)| = 0$, os vetores v e v são iguais.

O critério de desempate visa dar maior peso à maior diferença entre as coordenadas, ou, em caso de empate, à menor coordenada onde ocorreu essa diferença. Por exemplo, o vetor (1,2,3) possui a mesma somatória do vetor (3,2,1). Pelo critério de desempate, o primeiro vetor é considerado menor que o segundo, pois nesse exemplo todas diferenças possuem o mesmo peso e a menor coordenada é a primeira (nesse exemplo, a ordem proposta se comporta como seria na ordem lexicográfica).

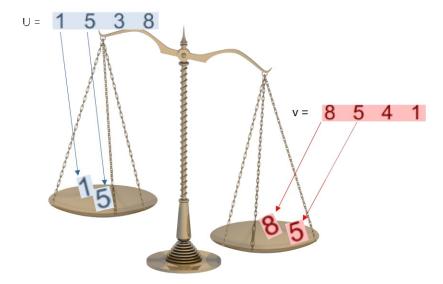


Figura 8 – Comparação entre dois vetores pela ordem lexicográfica estendida

2.2 Conceito de malha

Existem várias definições de malhas. Esta pesquisa usará o conceito de complexo simplicial (BISCARO, 2005). Uma malha é definida pela sua geometria, conectividade e propriedade. A geometria trata sobre a posição dos vértices no espaço de acordo com suas coordenadas em \mathbb{R} ; a conectividade define as adjacências entre os vértices e a propriedade tem por finalidade o aspecto visual da malha.

As malhas triangulares são as mais eficientes em relação a outros polígonos geométricos, devido à facilidade nas operações de cálculos (PENG; KIM; Jay Kuo, 2005).

A Figura 9 ilustra a malha de um coelho constituída por uma rede de vértices e triângulos.

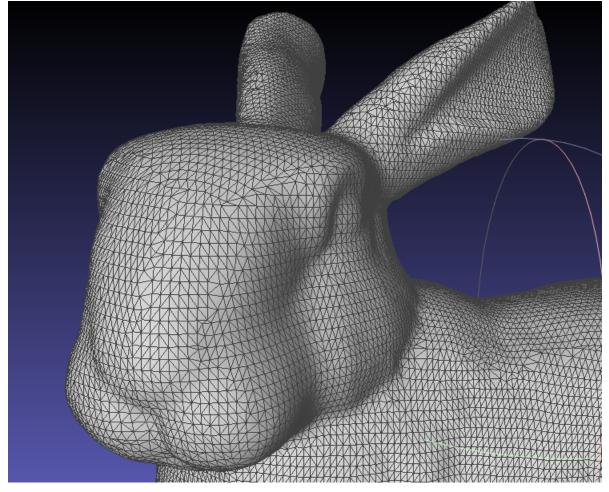


Figura 9 – Detalhe da malha de um coelho

Fonte-Autor

2.2.1 Simplexo

Um simplexo p-dimensional ou p-simplexo em \mathbb{R}^n é o fecho convexo de p+1 pontos geometricamente independentes. Um conjunto de pontos $a_0, a_1, ..., a_p \subset \mathbb{R}^n$ é geometricamente independente quando o conjunto de vetores $\{(a_1 - a_0), ..., (a_p - a_0)\}$ é linearmente independente. O fecho convexo de um conjunto de pontos é o menor conjunto convexo que contém tais pontos. Um 0-simplexo é chamado de vértice; um 1-simplexo é chamado de aresta; um 2-simplexo de triângulo; e um 3-simplexo, de tetraedro.

Considerando-se $\sigma = [a_0, a_1, ..., a_p]$ um p-simplexo, qualquer k-simplexo $0 \le k \le p$ formado por um subconjunto de $a_0, a_1, ..., a_p$ é chamado de face de σ . As faces diferentes de σ são chamadas de faces próprias, e a união das faces próprias de σ é chamada de bordo de σ e denotada por $\Delta \sigma$. O interior de um simplexo é definido como sendo $\sigma - \Delta \sigma$ e denotado por $\overset{\circ}{\sigma}$.

Um complexo simplicial K em \mathbb{R}^n é uma coleção de p-simplexos tal que:

- 1. Toda face de um simplexo de K está em K;
- A interseção de quaisquer dois simplexos de K é uma face comum a cada um deles ou é vazia;
- 3. Qualquer conjunto compacto de \mathbb{R}^n intersecta um número finito de simplexos.

A dimensão de um complexo simplicial K, denotada por dim K, é o máximo das dimensões dos simplexos de K. Se L é uma subcoleção de K tal que L é um complexo simplicial, então L é o chamado subcomplexo de K. A coleção de todos os simplexos de K com dimensão menor ou igual a n (n menor que dimensão de K) é um subcomplexo chamado n-esqueleto de K e será denotado por $K^{(n)}$.

A união de todos os simplexos de K é chamada de politopo de K (ou espaço gerado por K) e denotada por |K|. $X \subset \mathbb{R}^n$ é dito ser um poliedro se X = |K| para algum complexo simplicial finito K. O grau de um vértice $v \in K$ é definido como o número de arestas que tem v como vértice.

Considere-se K é um complexo simplicial, V o conjunto dos vértices de K. Seja |k| a coleção de todos os subconjuntos $a_0, ..., a_n$ de V, tal que os vértices $a_0, ..., a_n$ geram um simplexo de K. Tal coleção é chamada de esquema de vértices de K.

Seja σ um p-simplexo. Duas orientações de seus vértices são ditas equivalentes se elas diferem por uma permutação par. Uma permutação é denominada par se necessita de um número par de trocas para transformá-la na identidade. Se p>0, então existem duas classes de equivalência; caso contrário, apenas uma. Cada uma dessas classes é chamada de uma orientação de σ . Um simplexo orientado é um simplexo σ juntamente com uma orientação.

Considerando-se o simplexo formado pelos vértices v_0, v_1, v_2, ev_3 , a orientação $[v_0, v_1, v_2, v_3]$ será escolhida como sendo a orientação positiva e, ainda, $-\sigma$ fica definido como sendo σ com sua orientação oposta. Para um complexo simplicial K de dimensão n, um (n-1)-simplexo é um simplexo interior quando compartilhado por dois n-simplexos; caso contrário, é um simplexo de bordo. Os simplexos de dimensão menor que n-1 contidos em simplexos de bordo são também chamados de simplexo de bordo. O bordo de K é um subcomplexo $\subset K$ formado por todos os p-simplexos de bordo, p=0,1,...,n-1 de K, e cada componente conexa de |S| é chamada de componente de bordo de K. Para esta pesquisa, o conceito de malha será definido como complexos simpliciais de dimensão 2.

2.3 CBIR 3D

Um sistema CBIR é definido por recuperar um conjunto de imagens similares à imagem parâmetro enviada pelo usuário. Para realizar operações de comparações entre as imagens do banco de dados e da imagem usada como referência, é comum a transformação das imagens em vetores de características, que resumem matematicamente os atributos representativos das respectivas imagens. A medida de similaridade pode ser realizada com comparações entre os vetores. Uma medida bastante utilizada na literatura é a relação de distância Euclidiana entre os vetores. A Figura 10 detalha essas etapas.

Em Wang (2001), são destacados quatro elementos que um sistema CBIR precisa possuir:

- 1. Técnicas para automação de extração dos metadados contendo as características primitivas da imagem.
- 2. Interface para obter os dados de pesquisa feita pelo usuário.
- 3. Métodos para comparar as similaridades/diferenças entre imagens.
- 4. Técnicas para indexar e armazenar de forma eficiente os metadados.

Para este trabalho, os itens 1 e 3 são os objetos de estudo principais. A técnica para automação de extração dos metadados contendo as características primitivas da imagem será abordada na seção sobre os extratores. Os métodos para comparar as similaridades/diferenças entre imagens são com base nos vetores de características resultantes de cada objeto 3D. Em outro estudo, Bergamasco e Nunes (2019), cita duas fases do sistema CBIR: fase de extração de características e indexação; e a fase de recuperação de imagens.

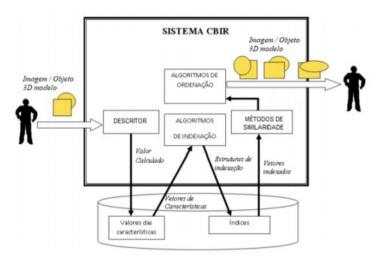


Figura 10 – Esquema geral de um sistema CBIR

Fonte - Extraído de (MARQUES, 2006)

2.4 Métricas de avaliação

Há quatro resultados possíveis em que uma amostra pode ser classificada:

- Verdadeiro positivo (Vp)
- Verdadeiro negativo (Vn)
- Falso positivo (Fp)
- Falso negativo (Fn)

Com base nos resultados da amostra, é possível extrair diversas métricas entre os quatro itens. Algumas das mais utilizadas na literatura são:

• Precisão (Pr)

Dentre todos os positivos, quantos realmente são positivos (estão corretas). É definida pela equação:

$$Pr = Vp/(Vp + Fp) \tag{2}$$

 \bullet Revocação (Rv)

São amostras que foram identificadas corretamente como pertencendo a determinada classe dentre todas as amostras que são desta classe. É definida pela equação:

$$Rv = Vp/(Vp + Fn) \tag{3}$$

• Acurácia (Ac)

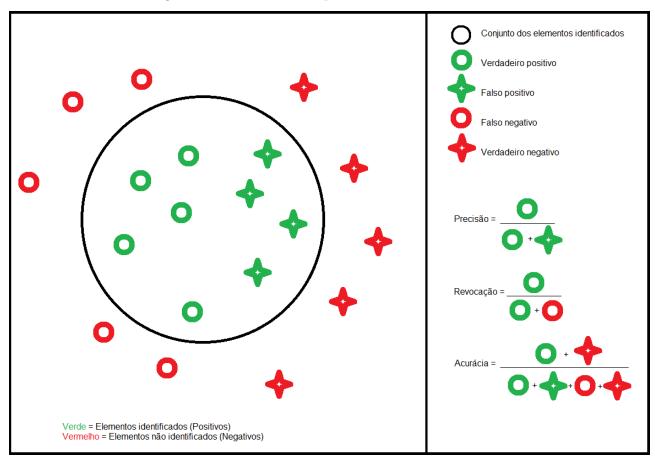
Define amostras que foram classificadas corretamente como pertencendo e não pertencendo a uma classe dentre todas as amostras. É definida pela equação:

$$Ac = (Vp + Vn)/(Vp + Vn + Fv + Fp)$$

$$\tag{4}$$

A Figura 11 ilustra a relação de cada uma das métricas definidas num universo de elementos identificados dentro de uma circunferência e não identificados fora dela.

Figura 11 – Modelo exemplificativo de métricas



Fonte-Autor

2.5 Descritor espectral

Em Candes e Romberg (2007), foi sugerido um protocolo de aquisição de dados que traduz informação analógica em digital com menos sensores que o previsto anteriormente. A mecânica sensorial é demonstrada na Equação 5, onde um sinal f(t) é armazenado usando produto interno com funções lineares.

$$yk = \langle f(t), \phi k(t) \rangle. \tag{5}$$

Dado um sinal k-esparso n-dimensional S , é preciso achar uma matriz $m\times n$ Φ_Ω e resolver o problema de otimização da Equação 6.

$$\min ||S||_{l_1}, \text{ sujeito a } \Phi_{\Omega}S = Y \tag{6}$$

Os componentes $y_1, ..., y_m$ do vetor Y são chamados de medidas do sinal S. Ω é um subconjunto de medida aleatória de tamanho $|\Omega| = m$. A teoria de compressão de sinal (CS) tem por objetivo escolher Φ_{Ω} com o menor número de medidas possíveis e ainda sim manter uma reconstrução confiável de S. Na maioria das vezes, um sinal X precisa ser reescrito como ($\Psi X = S$) em uma base apropriada, a fim de manter uma representação esparsa, onde Ψ é uma matriz de mudança de base. Então, invés de S da equação 6, teremos:

$$\Psi X = S \Leftrightarrow \Psi^* S = X$$

$$\Theta_{\Omega}S = Y$$

onde $\Theta_{\Omega} = \Phi_{\Omega} \Psi^*$

 $\frac{y}{\Phi_{\Omega}} = \frac{s}{\Phi_{\Omega}}$

Figura 12 – Representação gráfica do problema algébrico

Fonte - (SCHULZ; SILVA; VELHO, 2009)

Muitos sinais naturais como imagens, sons e malhas geométricas são esparsas ou têm uma representação espacial numa base apropriada. O algoritmo, como descrito em (BISCARO *et al.*, 2016), consiste em decompor a malha em dois conjuntos: vértices e faces.

O conjunto de vértices é decomposto em três vetores $X, Y \in Z$, cada um representando os respectivos pontos dos eixos em \mathbb{R}^3 . Contudo, os dados dos vetores não estão em representação esparsa. Para realizar uma decomposição espectral, é preciso usar uma matriz Laplaciana onde E é o conjunto de arestas de uma malha M e d_i é o número de vizinhos imediatos de um vertex v_i (a valência ou grau de v_i). Seja A a adjacência ou conectividade da matriz M, como mostrado na equação 7.

$$A_{ij} = \begin{cases} 1, \text{ se } (v_i, v_j) \in E \\ 0, \text{ caso contrário} \end{cases}$$
 (7)

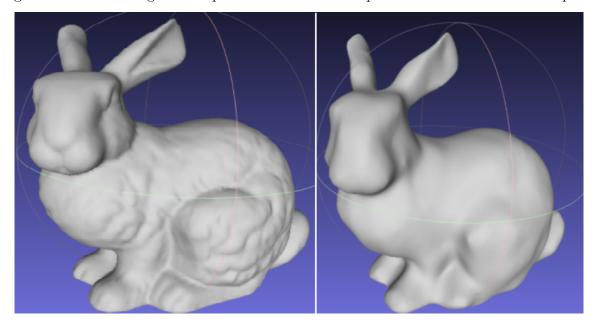
E seja D a diagonal da matriz tal que $D_{ii} = d_i$. Então, a Matriz Laplaciana em coordenadas relativas é definida como:

$$L = I - D^{-1}A \tag{8}$$

Depois, é calculado os primeiros k autovetores através de uma decomposição SVD, $L = U\Sigma V*$ onde U é uma matriz unitária $n\times k$. Σ é uma diagonal de matriz $k\times k$ e a $k\times n$ matriz unitária V* denota a conjugada transposta da matriz V $n\times k$. Cada vetor X,Y e Z é multiplicado por U para obter a representação esparsa.

Exemplificando o processo de transformação dos dados, a Figura 13 representa a malha de um coelho à esquerda e à direita representa a malha do coelho reconstruída após a transformação espectral, para k=200, onde k é o número dos coeficientes mais significantes usados.

Figura 13 – Malha original à esquerda e reconstruída a partir de 200 coeficientes à esquerda



Fonte – Autor, 2020

Para visualização da natureza de densidade das coordenadas dos vértices, a Figura 14 tem o eixo x representando o número de vértices e o eixo y representando a magnitude dos valores. Os eixos X, Y e Z das coordenadas das malhas são representados pelas cores azul, verde e vermelho, respectivamente. É possível notar que após a mudança de base, as coordenadas ficam mais esparsas, no gráfico à direita.

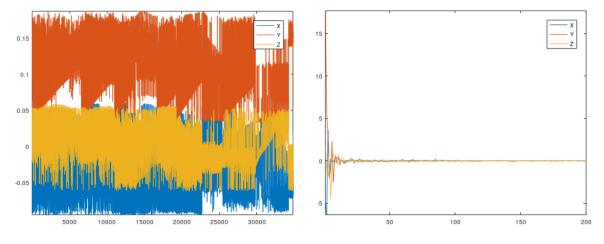


Figura 14 – Coordenadas cartesianas e coordenadas espectrais

Fonte – Autor

2.6 SPHARM (Esféricos harmônicos)

SPHARM é um conceito utilizado em áreas do conhecimento como física e engenharia. Na computação, é usado na modelagem 3D para simular efeitos de luz (BERGAMASCO, 2018). SPHARM são funções esféricas que satisfazem a equação de Laplace. Essas funções podem ser expressas como uma combinação linear de coeficientes harmônicos como mostrado na Equação 9, onde o termo $a_l m$ é uma matriz de coeficientes harmônicos complexos e o $Y_l^m(\theta, \phi)$ term são SPHARMs de grau l e ordem m para as coordenadas esféricas (θ, ϕ) . Em geral, o valor m está dentro do intervalo [-l, l] e $l \in \mathbb{N}[0, \infty]$.

$$f(\theta,\phi) = \sum_{l}^{\infty} \sum_{m=-l}^{l} a_{l} m Y_{l}^{m}(\theta,\phi)$$
(9)

O cálculo de SPHARMs pode ser dividido em várias partes. Cada fração do cálculo pode ser definida como:

1. Cálculo de harmônicos: o termo $Y_l^m(\theta,\phi)$ da Equação 9 é um harmônico de grau l e ordem m das coordenadas esféricas (θ,ϕ) . A coordenada esférica $\theta \in [0,2\pi]$ representa o ângulo polar, e a $\phi \in [0,\pi]$ representa o azimute. A equação 10 apresenta a formulação deste termo, onde N_l^m é o fator de normalização de SPHARMs, $P_l^m(\cos\phi)$ é o polinômio de Legendre associado e $e^i m \phi$ é o exponencial complexo.

$$Y_l^m(\theta,\phi) = N_l^m P_l^m(\cos\phi) e^{im\phi} \tag{10}$$

2. Fator de normalização: é um binômio que utiliza os valores de l e m como entrada (Equação 11).

$$N_l^m = \sqrt{\frac{2l + 1(l - m)!}{4\pi(l + m)!}} \tag{11}$$

3. Polinômios de Legendre associados: são soluções dos polinômios da Equação Geral de Legendre. Esses polinômios podem ser expressos como mostrado pelas Equações 12 e 13, em que l e m são as mesmas variáveis que a Equação 10 e k é uma variável de controle que começa com valor zero e usa as variáveis l e m para determinar seu valor máximo.

$$P(x)_{l}^{m} = (1 - x^{2})^{m/2} \sum_{k=0}^{\frac{lm}{2}} \frac{(2l - 2k!)}{2^{l}(lk)!k!(l - 2k - m)!} \omega$$
 (12)

$$\omega = (-1)^k x^{(l-2k-m)} \tag{13}$$

4. exponencial complexo: A equação 14 apresenta o cálculo desta variável. Este termo é calculado usando a Identidade de Euler, onde o termo i representa a parte imaginária do número complexo, ϕ é o ângulo de azimute e m é a ordem SPHARM.

$$e^{im\phi} = \cos(m\phi) + i \cdot \sin(m\phi) \tag{14}$$

5. Coeficientes harmônicos $a_l m$: esses coeficientes são projeções de $f(\theta, \phi)$ em cada função base de $Y_l^m(\theta, \phi)$ de uma esfera, como mostrado em 15.

$$a_l m = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} f(\theta, \phi) Y_l^m(\theta, \phi) \sin(d\phi dd)$$
(15)

Após calcular cada fração da Equação 9, obtém-se um vetor de tamanho N, onde N é o número de vértices da estrutura 3D. Este vetor é utilizado como vetor de características dos objetos do banco de dados.

O algoritmo que calcula SPHARM como descritor é detalhado em (BERGAMASCO, 2018) e está resumido em Algoritmo 1. Na linha 2, cada vértice v (em coordenadas cartesianas), de uma malha de entrada M, é transformado em suas coordenadas esféricas $\theta \in \phi$. Enquanto o valor de m for menor que l, o valor de $Y_l^m(\theta,\phi)$ é calculado nas linhas 7 e 8. A forma conjugada do harmônico esférico é simbolizada por $\overline{(N_l^m P_l^m(\cos\phi)e^{im\phi})}$ na linha 7. N_l^m é o cálculo binomial, $P_l^m(\cos\phi)$ é o polinômio de Legendre e $e^{im\phi}$ é a identidade de Euler. Como resultado, uma matriz M gerada contendo cada par (l, m) na equação 10.

```
Algoritmo 1 Cálculo SPHARM
Uma malha geométrica M e um valor l. Uma Matriz M com a soma de cada par (l, m).
Para cada vértice v de M
(\theta, \phi) = \text{coordenadas esféricas de } v
m=-l;
Ytemp = 0.0;
while m \leq l do
  if (m < 0) then Ytemp + = (-1)^m (N_I^m P_I^m (\cos \phi) e^{im\phi});
M[l][m] = Ytemp;
  else
    Ytemp + = (-1)^m (N_I^m P_I^m(\cos \phi) e^{im\phi});
    M[l][m] = Ytemp; \ m++; \ coef_{alm} = ((M^T*M)^-1)*(M^T*v));
```

Este capítulo tratou de abordar os conceitos fundamentais para compreensão desta pesquisa. Para o próximo capítulo, é detalhado a parte da revisão sistemática realizada na literatura da área.

3 Revisão Sistemática

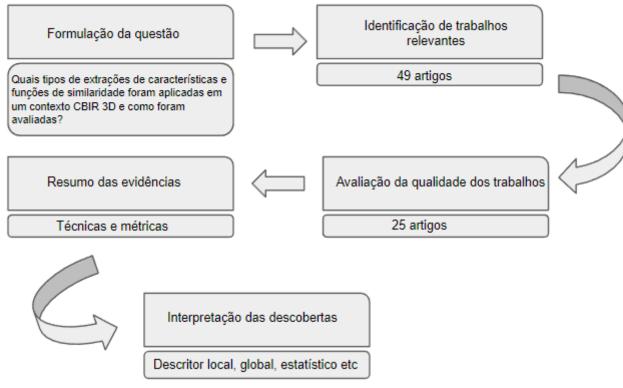
Para esta pesquisa, foi realizada uma revisão sistemática na literatura de processamento gráfico sobre extração espectral em malhas 3D. Segundo Khan *et al.* (2003): "Uma revisão ganha o adjetivo 'sistemática' se for baseada em uma questão claramente formulada, identifica estudos relevantes, avalia sua qualidade e resume as evidências por uso de metodologia explícita. É a abordagem explícita e sistemática que distingue as revisões sistemáticas de revisões exploratórias tradicionais e comentários." Além disso, a revisão é separada em cinco etapas (KHAN *et al.*, 2003):

- 1. Formulação da questão;
- 2. Identificação de trabalhos relevantes;
- 3. Avaliação da qualidade dos trabalhos;
- 4. Resumo das evidências;
- 5. Interpretação das descobertas.

A Figura 15 exibe uma visão esquemática da revisão sistemática apresentada neste trabalho.

O objetivo principal desta revisão é avaliar como os extratores espectrais funcionam e as métricas utilizadas na literatura para estimação de seus respectivos desempenhos.

Figura 15 – Diagrama da revisão sistemática



Fonte - Autor, 2020

3.1 Protocolo de Pesquisa

Para esta revisão, a questão de pesquisa formulada é:

Quais tipos de extrações de características e funções de similaridade foram aplicadas em um contexto CBIR 3D e como foram avaliadas?

A revisão foi conduzida em buscadores de artigos científicos na internet, com a listagem de fontes:

- 1. https://scholar.google.com.br/
- 2. https://ieeexplore.ieee.org/Xplore/home.jsp
- 3. https://pubmed.ncbi.nlm.nih.gov/
- 4. https://www.sciencedirect.com/
- 5. https://www.spiedigitallibrary.org/

Para a consulta nos buscadores, a palavra-chave utilizada foi: "3D CBIR"

Os critérios de aceite adotados para inclusão dos artigos nesta revisão foram:

Capítulo 3. Revisão Sistemática

38

• Uso de malha no conjunto de dados.

• Utilização de sistema CBIR ou similar.

• Artigo em idioma inglês ou português.

• Artigo disponível online.

O processo de seleção dos artigos foi com base em análise dos artigos retornados nos

buscadores pela string utilizada. Na sumarização dos resultados, foi realizada a comparação

de todos artigos utilizando uma tabela com o resumo das principais métricas e resultados

que sejam utilizadas por todos artigos selecionados.

3.2 Condução e artigos obtidos

Para esta revisão, após apuração da string utilizadas na pesquisa, foram refinados

os resultados das buscas por "3D CBIR". Houve o resultado de 49 artigos científicos que

foram filtrados pelos critérios de aceite. A quantidade de artigos retornados por cada

buscador foi 31 do Google S.; 1 do IEEE; 12 do PubMed; 4 do Science Direct e 1 do Spie.

A condução da busca é exemplificada e filtrada como demonstrado:

Fonte: Google Acadêmico

Data de busca: 03/05/2020

String utilizada: "3D CBIR"

Período considerado: A qualquer momento

Filtros utilizados: Incluir patentes e citações

3.3 Publicações por ano

Pode-se notar um aumento de publicações nos últimos anos. A Figura 16 ressalta o

aumento médio de publicações anuais.

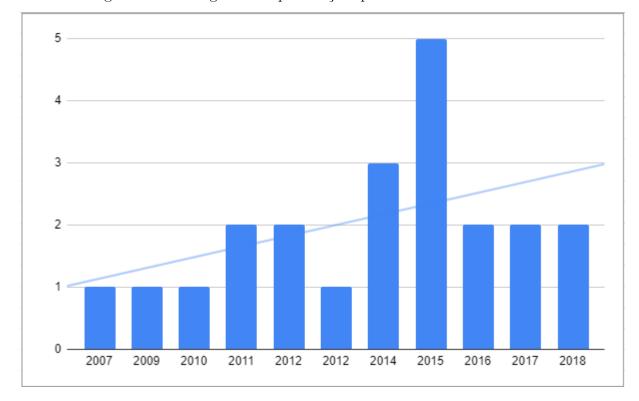


Figura 16 – Histograma de publicações por ano e linha de tendência

Fonte - Autor

3.4 Técnicas para extração de características

Na literatura da área, o CBIR aplicado a modelos 3D é muito recente, tendo nas últimas duas décadas enfoque em imagens 2D. Contudo, os resultados obtidos para o CBIR 3D são promissores (QIAN RUI HUI, 2012). Diversos estudos obtidos nesta revisão focam na área da medicina.

Em Qian Rui Hui (2012) o ambiente CBIR é usado de forma prática atrelada ao uso de aprendizado de máquina (SVM). A técnica utilizada para o CBIR 3D foi expandida do universo 2D (3D SIFT). (AK MUKHOPADHYAY S, 2017) diferencia ao utilizar junto do CBIR o CAD (Computer-Aided Diagnosis/Diagnóstico Auxiliado por Computador) para objetos representativos do pulmão humano, também utilizando SVM. A pesquisa de Kamath Disha Punjabi (2012) faz uso de SIFT aplicado diretamente em imagens 2D com diferença Gaussiana e NNS (Nearest Neighbour Search - busca do vizinho mais próximo). Aliado ao CBIR, o aprendizado de máquina tem sido utilizado em duas principais formas:

- 1. Aprendizado não-supervisionado (paradigma Bag of Word).
- 2. Aprendizado supervisionado (técnica de relevância de feedback).

A técnica BOW (Bag of Words) ou saco de palavras consiste em um dicionário visual criado previamente, funcionando como estrutura auxiliar e é responsável por relacionar cada ponto de característica (palavra) da imagem. Assim, a imagem fica representada como um histograma de ocorrências das palavras. Esta abordagem se classifica como uma abordagem estrutural (HOANG; GOUET-BRUNET, 2014).

Em relevância de feedback, são estudados os exemplos positivos e negativos selecionados pelos usuários com base no resultado do retorno, a fim de aprimorar o processo de recuperação (QIAN RUI HUI, 2012). and e YU (2007) cita uma aplicação de classificação e recuperação com foco em movimentação humana num espaço 3D. No trabalho de Gao Y. Qian (2010) o contexto de recuperação de imagens tem como foco de extração a textura do objeto 3D. Para métodos de abordagem de textura, há quatro variantes:

- 3D Grey Level Co-occurrence Matrices/Matrizes de co-ocorrência de escala de cinza 3D
- 3D Wavelet Transform/Transformada Wavelet 3D
- 3D Gabor Transform/Transformada Gabor 3D
- 3D Local Binary Pattern/Padrão Binário Local 3D

O melhor resultado obtido foi com uso da abordagem 3D LBP. Uma solução interessante de recuperação por conteúdo envolvendo CBIR 2D e 3D é o MIRAGE (XW QIAN Y, 2011), sistema desenvolvido para a informática médica, possuindo mais de 100 mil imagens 2D e 3D em seu banco de dados. Para a abordagem 3D, são utilizadas as quatro técnicas de recuperação com base em textura.

O trabalho de Qian X. Gao (2011) enfoca o uso de CBIR 3D para auxílio no diagnóstico médico, utilizando a estratégia de 3D Local Binary Pattern, que é uma abordagem de textura provinda de formas 2D. (BERGAMASCO et al., 2015b) também é voltado para medicina, utilizando agrupamento espectral para classificação dos modelos usados.

O trabalho apresentado em Bergamasco e Nunes (2015) faz uso de descritores globais e locais com dados médicos. Também demonstra o potencial de utilização desses sistemas para auxílio diagnóstico. Os descritores global e local utilizados são, respectivamente: Distance Histogram Descriptor/Descritor do Histograma de Distância (DHD) e Local Distance Histogram Descriptor/ Descritor de histograma de distância local (LDHD).

(BERGAMASCO et al., 2015a) apresenta um método novo de comparação de dois modelos, nomeado de técnica de grafos bipartidos. Utilizado em um banco de dados medicinal, houve melhora de 10% em relação à função de similaridade. Em Biscaro et al. (2016), a técnica utilizada é a de extração espectral, uma abordagem global e matemática, que visa a mudança de base da representação das malhas, de dados densos para dados espersos. A técnica é descrita detalhadamente no capítulo de Capítulo 2.

Para a área de arqueologia, Samoun et al. (2018) confecciona o ambiente de recuperação passando por uma técnica de aprendizado de máquina (máquina de vetores suporte) em que o problema de classificação se torna uma recuperação por similaridade, pois os objetos para pesquisa e consulta serão artefatos (ou partes) encontrados. Aznag et al. (2017) introduz um descritor invariante à rotações e transformadas de afinidade como coeficientes de momento e transformada de Fourier não uniforme. A diferença da técnica descrita é a dependência por parâmetro em função da forma, e não da ordem dos pontos, como é na transformada normal de Fourier.

Conforme Hoang e Gouet-Brunet (2014) detalha, as técnicas para extração de objetos 3D (3DOR - 3D Object Retrieval) podem ser definidas sob quatro aspectos:

- Baseado em estatística;
- Abordagem estrutural;
- Abordagem de mudança de base;
- Mudança de visão.

Para Hoang e Gouet-Brunet (2014) e Lavoue (2012) é utilizada, em conjunto a outras técnicas, a técnica *BOW* (*Bag of Words*) ou saco de palavras. Nessa estrutura, um dicionário visual criado previamente funciona como estrutura auxiliar e é responsável por relacionar cada ponto de característica (palavra) da imagem. Assim, a imagem fica representada como um histograma de ocorrências das palavras. Esta abordagem se classifica como uma abordagem estrutural.

Outra técnica utilizada nos artigos é a 3D Harris detector. Este detector tem por função classificar pontos como características (SIPIRAN; BUSTOS, 2011). Em relação a outros detectores esparsos, ele tem um desempenho global melhor (DUTAGACI; CHEUNG; GODIL, 2012). Em Hoang e Gouet-Brunet (2014), o 3D Harris auxilia a melhorar o seu descritor utilizando a transformada harmônica múltipla, através do operador Laplace-Beltrami.

Publicação	Técnica		
(QIAN RUI HUI, 2012)	3D SIFT		
(GAO Y. QIAN, 2010)	LBP		
(BERGAMASCO, 2018)	SPHARM		
(QIAN X. GAO, 2011)	LBP		
(BERGAMASCO et al., 2015b)	agrupamento espectral		
(BERGAMASCO et al., 2015a)	grafos bipartidos		
(BISCARO et al., 2016)	extrator espectral		
(KAMATH DISHA PUNJABI, 2012)	SIFT		
(XW QIAN Y, 2011)	3D LBP		
(AZNAG et al., 2017)	Transf. Fourier não-uniforme		
(BERGAMASCO; NUNES, 2015)	LDHD		
(Samoun <i>et al.</i> , 2018)	SVM		
(AK MUKHOPADHYAY S, 2017)	MRMR		

Tabela 2 – Publicações e técnicas usadas

Fonte - Autor

Com isso, finalmente, os pontos tridimensionais são armazenados numa representação *Bag* of Words. A Tabela 2 exibe as técnicas utilizadas em cada publicação.

3.5 Métricas adotadas na avaliação

Para todos artigos analisados, a métrica de precisão \times revocação foi utilizada para avaliação dos resultados. Em Hoang e Gouet-Brunet (2014) é usada a métrica adicional MAP (Mean Average Precision) - precisão média. Lavoue (2012) faz uso de métricas adicionais:

- 1. Nearest Neighbor (NN): Porcentagem de consultas em que a correspondência mais próxima pertence à classe consultada.
- 2. First Tier (FT: A revocação para a correspondência mais próxima (C-1), onde C é a cardinalidade da classe consultada.
- 3. The Second Tier (ST): A revocação para a correspondência mais próxima de 2(C-1).
- 4. The Discounted Cumulative Gain (DGC): Métrica que tem por objetivo refletir quão bem a recuperação do conteúdo seria vista por um ser humano.

3.6 Funções de similaridade

Na etapa de calcular similaridade no sistema CBIR, é comum o problema se resumir a comparação de vetores numéricos (vetores de características). Para a comparação são usados métodos de similaridade (MARQUES; VIEIRA, 1999).

Nesta revisão foram encontrados trabalhos que usaram funções métricas de distância como Euclidiana e Manhattan. Essas funções foram umas das comparações mais comuns nessa etapa do CBIR.

Aprendizado de máquina também é possível ser utilizado, como SVM (Máquinas de Vetores Suporte). Alguns trabalhos desta revisão (QIAN RUI HUI, 2012), (GAO Y. QIAN, 2010), (QIAN X. GAO, 2011) utilizam técnica de interseção de histograma para calcular a similaridade entre imagens.

Outra técnica é a KNN (K-Vizinhos Mais Próximos), em que a definição para a classe é por meio da comparação da distância entre o elemento a ser classificado e os grupos de classes existentes. O que tiver a menor distância será considerado do grupo do novo elemento (YU et al., 2011).

Como motivação deste trabalho, não foram encontrados artigos que usem sistemas de ordenação para a função de similaridade. Dada essa situação, considerou-se a técnica de ordenação podendo ser mais explorada na literatura.

3.7 Considerações Finais

É recorrente o uso do termo *mesh*, mas para fins diferentes, como malha tridimensional (BISCARO; LIMA, 2015), malha de um grafo (PHAM, 2018) e malha bidimensional (MADZIN, 2013). A maioria dos trabalhos tem aplicação em computação gráfica voltada para a medicina. Para descritores espectrais em malhas tridimensionais aplicadas num modelo de CBIR, com base nos resultados desta revisão, é destaque também a recente exploração na literatura.

Dos trabalhos selecionados, boa parte utiliza diversos tratamentos geométricos nas malhas 3D. Também há abordagens em que visam a análise da malha de forma local (NGUYEN; GOUET-BRUNET, 2014) ou global (BISCARO *et al.*, 2016). Houve uma parcela também que tratou de realizar uma sumarização da literatura. Na parte de

avaliação das técnicas, é consenso na literatura a utilização da métrica de precisão \times revocação no ambiente CBIR em conteúdo 2D ou 3D.

4 Materiais e métodos

Para atingir o objetivo da pesquisa, foi optado por simular ambientes de sistemas CBIR 3D em que o objeto de consulta pertence ao banco de dados original, mas que em seu teste era retirado do banco de dados. Essa prática foi adotada com todos os 380 objetos do banco de dados usado, mudando as funções de similaridade em cada CBIR. Utilizando dois descritores e variando a quantidade de coeficientes do descritor espectral, é possível gerar vetores de características que abstraem diferentes informações de cada malha. Com isso, foi possibilitado comparar não só diferentes funções de similaridades, mas o comportamento de cada descritor também.

Como o objetivo do trabalho é propor uma nova função de similaridade e uma ordenação específica, foram realizados vários testes usando diversas funções de similaridade abrangente entre distâncias e ordens. Como visto na revisão sistemática, não foram encontrados trabalhos que utilizassem o sistema de ordenação na etapa de similaridade do CBIR. Em testes prévios, pôde-se notar o desempenho melhor que o sistema de ordenação possuia em relação à sistemas de distâncias clássicos, como Euclidiana e Manhattan. Para realizar comparações válidas entre esses tipos de similaridades, foram montados cenários de testes com CBIRs idênticos que se diferenciam apenas na etapa de similaridade, e no final foi medido os diferentes retornos que esses sistemas devolviam.

4.1 Material de pesquisa

Como material de trabalho, foi utilizado um conjunto de malhas benchmark da literatura disponibilizado por (SHREC, 2007). Este repositório contém 380 objetos divididos em 19 classes, listadas a seguir:

1.	Alicate	8.	Mão
2.	Avião	9.	Mecha
3.	Cadeira	10.	Mesa
4.	Copo	11.	Óculos
5.	Dorso	12.	Pássaro
6.	Formiga	13.	Peixe
7.	Humano	14.	Polvo

15. Quadrúpede

18. Ursinho

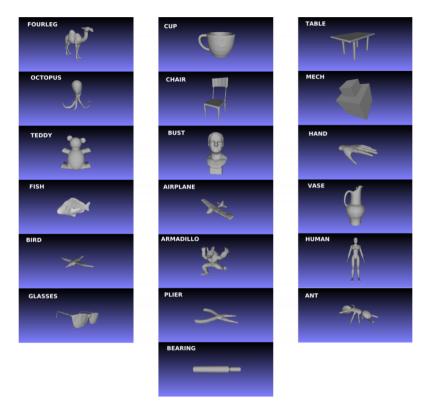
16. Rolamento

19. Vaso

17. Tatu

A Figura 17 representa uma malha por classe.

Figura 17 – Exemplos de malhas e suas classes



Fonte - Autor, 2020

Além desse banco de dados, para testes iniciais, foram utilizados objetos tridimensionais elementares: cubo e esfera. Também foi realizada manipulação nesses objetos, salientando alguns vértices, para avaliar a transformação após passar pelo extrator espectral.

4.2 Método de pesquisa

Nesta seção é detalhado como foram analisados todos resultados obtidos (Capítulo 5) e como foram criadas cada CBIR separado por descritor e por função de similaridade. Também foram abordadas as métricas usadas nos resultados e como foi comparado cada dado.

4.2.1 Descritor espectral

O descritor espectral permite configurar o número de coeficientes máximo a ser gerado pelo descritor. Com isso, optou-se por variar os valores do coeficiente k do descritor em 50, 100 e 200, obtendo 3 vetores de características de tamanho respectivo igual a k. Desse modo tem-se 3 cenários para o descritor espectral: um CBIR usando vetores de características de 50 dimensões, um CBIR usando vetores de 100 dimensões e um CBIR de 200 dimensões.

4.2.2 SPHARM

Para o caso do SPHARM, como não é utilizado coeficiente como parâmetro, os vetores de características terão tamanho variado de acordo com as dimensões da malha de entrada. Para não gerar comparações entre vetores de tamanhos diferentes, adotou-se a metodologia de truncar todos os vetores para o tamanho do menor vetor gerado. Desse modo evitou-se ajustes que seriam necessários nos casos de calcular distâncias e ordens que comparam os vetores inteiros. Assim todas comparações foram feitas entre vetores de tamanho igual. O vetor de características mínimo do banco de dados possui o número de 1343 elementos. Então para o pré-tratamento do SPHARM adotou-se o corte de todos vetores para o tamanho de 1343 elementos.

4.2.3 Funções de similaridade

Como descrito na seção 2.3, a etapa de função de similaridade pode ser explicada como um método para comparar as similaridades entre imagens a fim de obter uma saída do sistema CBIR que o resultado obtido seja semelhante ao resultado esperado. Na literatura é comum utilizar relações de distância nessa etapa, como visto na revisão sistemática (Capítulo 3). Neste trabalho, contudo, foi proposta a relação por ordem e sugeriu-se uma relação de ordem que não foi encontrada aplicação na literatura durante a revisão sistemática.

Distâncias

Foram usadas duas distâncias aplicadas comumente na literatura: Euclidiana e de Manhattan.

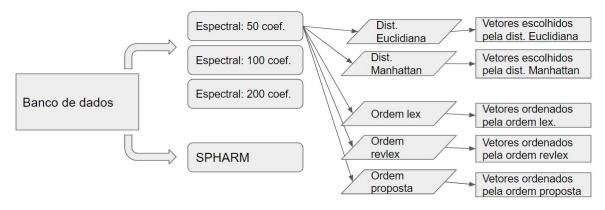
Ordens

Adotou-se três definições de ordenação: lexicográfica (ou dicionário), revlex, que é a ordem lexicográfica reversa e mais uma ordem definida pelo autor, que diferente das outras ordens usadas, considera o vetor inteiro.

4.2.4 Vetores de características

Com o banco de dados comprimido pelos descritores espectral e SPHARM, todos vetores de características foram ordenados um a um por cada função de similaridade de distância e de ordem. A Figura 18 ilustra o fluxo do tratamento que os dados passaram no caso do descritor espectral de coeficiente k=50. O mesmo fluxo aconteceu para os coeficientes k=100 e k=200. Para o caso do SPHARM, o processo foi feito apenas uma vez. A Figura 19 exibe o mesmo fluxo para esse descritor. Como cada processo de descritor resultou em cinco listas de vetores de características (cada lista contendo a descrição de uma das 380 malhas do banco de dados) e foram considerados quatro descritores, teve-se ao todo 20 listas de vetores de características.

Figura 18 – Descrição do fluxo de dados para o descritor espectral de k=50



Fonte - Autor

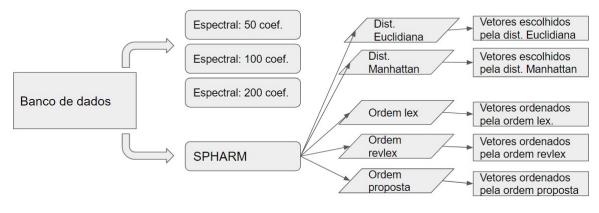


Figura 19 – Descrição do fluxo de dados para o descritor SPHARM

Fonte - Autor

4.3 Avaliação

Para avaliação do desempenho, foi utilizada uma malha do banco de dados como entrada e foram devolvidas as n malhas mais próximas na ordenação. O valor de n será 5, 10 e 15. Foi utilizada uma malha por vez. Por exemplo, primeiro foi utilizada a malha 1 das 380 e o banco de dados ficou com as 379 malhas restantes. Foram retornadas as 5 malhas mais similares, depois as 10 e depois as 15 malhas mais similares da malha 1. Para o teste com a malha 2, foi removida a malha 2 do banco de dados original, constando 379 malhas. Foram retornadas as 5 malhas mais similares da malha 2, depois as 10 e depois as 15 malhas mais similares e assim sucessivamente.

O protótipo funciona da seguinte forma: a) Todos os objetos do banco de dados são mapeados em \mathbb{R}^n e ordenados em uma linha. A situação ideal é que objetos de classes diferentes não se misturem na fila, conforme ilustrado na Figura 20. b) Quando uma busca por um objeto O é solicitada, ele é inserido temporariamente na fila e o retorno consiste nos objetos $\frac{n}{2}$ antes de O e o $\frac{n}{2}$ objetos depois de O na fila, onde n é o número de objetos desejados como retorno de busca.

A Figura 21 demonstra o processo de retornar os elementos mais semelhantes em relação a um novo elemento. Neste caso, são retornados os 5 vizinhos mais próximos, com 60% de precisão.

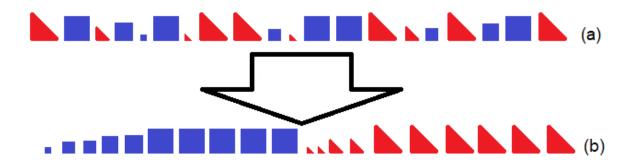


Figura 20 – Duas classes de objetos, triângulo e quadrado. Em (a) eles são desordenados. Em (b) eles estão ordenados.

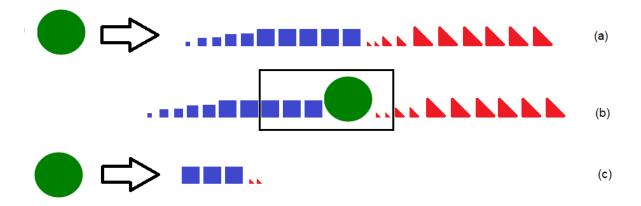


Figura 21 – (a) Um novo elemento é usado como pesquisa. (b) O elemento é inserido temporariamente na lista. (c) Os 5 elementos vizinhos são retornados.

4.3.1 Gráfico PxR

Os resultados foram avaliados pelas métricas de Precisão versus Revocação (PxR), que são as métricas mais utilizadas para avaliação de um sistema CBIR, conforme citado na revisão sistemática. A acurácia do sistema CBIR foi avaliada pela área debaixo da curva formada pelo gráfico da (PxR). Cada resultado foi avaliado para as entradas de coeficientes k=100 e 200 do descritor espectral e SPHARM, considerando as n malhas recuperadas, para n=5, 10 e 15. Os resultados foram divididos pelas classes e separados entre valores mínimo (%), máximo (%), média (%) e variação.

O cálculo de precisão pode ser feito utilizando a interseção de elementos relevantes e elementos recuperados dividido pelos elementos recuperados. Para a revocação é utilizada a interseção dos elementos relevantes e elementos recuperados divididos pelos elementos

relevantes. A Figura 22 ilustra um gráfico PXR, em que o eixo x representa os valores de revocação e o eixo y os valores de precisão. Usando como exemplo o conjunto de elementos recuperados $relevantes = \{A1, A2, A3, A4, A5\}$ e o total de elementos $recuperados = \{A2, A4, B1, A3, B4, B2, A1, B3, A5\}$, temos o primeiro elemento recuperado (A2) precisão = 1 e revocação = 1/5. O segundo elemento recuperado (A4), a precisão se mantém e a revocação = 2/5. O terceiro elemento (B1) não é um relevante, então as métricas não são computadas. Para o quarto elemento (A3), o valor de precisão atualiza para 3/4, pois para quatro elementos recuperados, três são relevantes. A revocação atualiza para 4/5, pois três dos cinco elementos relevantes foram recuperados. Assim sucessivamente até que todos elementos relevantes sejam recuperados.

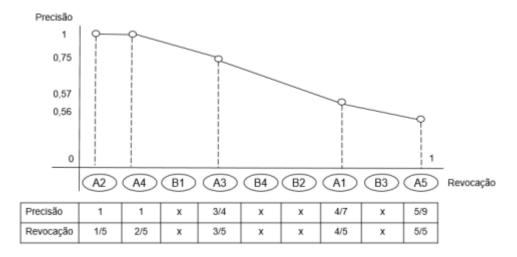


Figura 22 – Exemplo de gráfico PxR

Fonte – (BERGAMASCO, 2018)

Para ilustrar os cálculos e o retorno do CBIR, a Figura 23 contém a curva PxR para a malha identificada como número 373 (um objeto da classe Quadrúpede) do banco de dados. A configuração do CBIR utilizada foi o descritor espectral com 200 coeficientes e a ordem lexicográfica estendida como função de similaridade. Das 10 malhas recuperadas, 8 são da mesma classe do objeto usado como entrada (Quadrúpede) e 2 são da classe Mecha. Calculando a área sob a curva da figura Figura 23, tem-se uma AUC de 79%.

4.3.2 Área da curva

A área da curva da Figura 22 gera a métrica de AUC. Esse valor oscila entre 0 a 1, sendo que quanto mais próximo de 1, melhor o desempenho. O cálculo da AUC pode

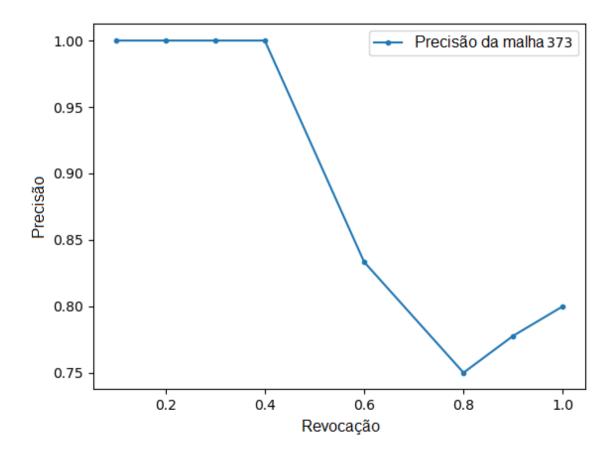


Figura 23 – Gráfico da curva Precisão x Acurácia da malha 373 da base de dados utilizando o descritor espectral com 200 coeficientes

ser obtido pela integral da Equação 16, onde p é o valor da precisão e r a revocação (BRODERSEN $et\ al.,\ 2010$).

$$AUC = \int_0^1 p(r)dr \tag{16}$$

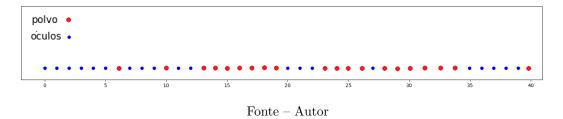
5 Resultados e discussões

Este capítulo exibe e discute os resultados obtidos durante a condução deste trabalho. Executou-se todo mecanismo de um sistema CBIR no banco de dados completo, repetindo várias etapas e mudando parâmetros, como descrito em detalhes no Capítulo 4.

5.0.1 Resultados com ordens para duas classes de objetos: Polvo e óculos

Utilizando o descritor espectral, executou-se simulações de ordens (lexicográfica, revlex e ordem lexicográfica estendida) em duas classes de 19 objetos cada do banco de dados. As classes utilizadas são do polvo e óculos, que obteram os piores resultados (12% de precisão média) quando usadas funções de similaridade de distância (Euclidiana e Manhattan). A ordem lexicográfica adotada consiste em uma relação de comparação entre vetores. Compara-se elemento a elemento entre dois vetores e é o vetor é classificado como menor, igual ou maior em relação ao outro vetor, de acordo com os elementos comparados. A Figura 24 mostra 20 malhas de polvos e 20 de óculos em ordem lexicográfica, representados pelos seus respectivos vetores de características gerados pelo descritor espectral de coeficiente 200.

Figura 24 – Polvos x Óculos em ordem lexicográfica



Como observado na Figura 24, a separação entre as classes não é trivial, dado que não há um limiar claro entre as malhas de polvo e óculos. Em outros resultados, a ordem revlex (ordem lexicográfica reversa) também não obteve melhoria considerável. A Figura 25 ilustra a ordenação.

polvo • oculos •

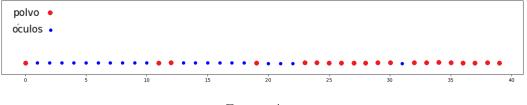
Figura 25 – Polvos x Óculos em ordem lexicográfica reversa

Fonte - Autor

Para este trabalho foi utilizado outro tipo de ordenação, baseada na lexicográfica, mas que sempre considera todos os pontos dos vetores. A nomeamos ordem lexicográfica estendida.

A Figura 26 mostra 20 malhas de polvos e 20 de óculos na ordem lexicográfica estendida, representados pelos seus respectivos vetores de características gerados pelo descritor espectral de coeficiente 200.

Figura 26 – Polvos x Óculos em ordem lexicográfica estendida



Fonte – Autor

Considerando um limitador de classes entre o elemento 20 e 25, temos 5 elementos fora da classe, aproximadamente 12% do total. Caso o usuário queira 10 elementos similares para cada uma das 40 malhas listadas, a precisão média será de 70%. Essa precisão é superior em relação à função distância Euclidiana, que obtém aproximadamente 12% de precisão.

5.1 Resultados obtidos

Esta seção contém as métricas obtidas após executar no banco de dados as compressões com os descritores e comparar os resultados das funções de similaridade de distância e ordem.

5.1.1 Resultados do descritor espectral

Para as funções de similaridade usando comparação de distâncias (Euclidiana e Manhattan), os resultados foram semelhantes, obtendo uma média de 50% de acurácia máxima para todas as classes nos casos em que o número de objetos recuperados é igual a dez. Nas funções de similaridade usando ordem, a ordem lexicográfica estendida obteve o melhor resultado em relação às demais funções de similaridade, com acurácia máxima em média de 63%, conforme ilustrado na Figura 27.

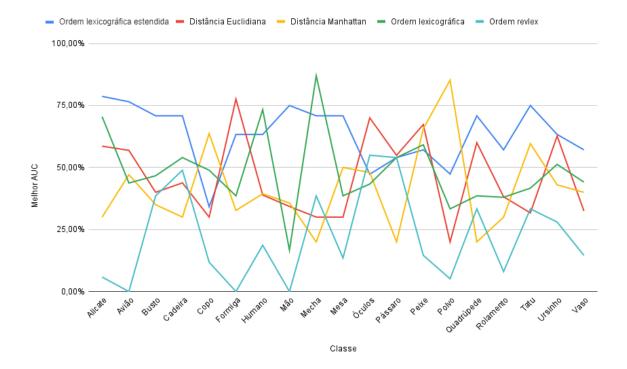


Figura 27 – Gráfico AUC máximo dos resultados das distâncias e ordens do descritor espectral, onde k=200e n=10

A Figura 28 indica que a variação do número de coeficientes entre 50, 100 e 200 no descritor espectral gerou variação nos resultados máximos da ordem lexicográfica estendida, com média de 79%, 87% e 63% respectivamente.

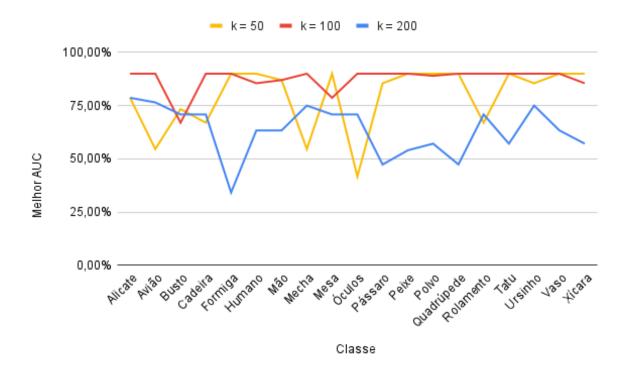


Figura 28 – Gráfico AUC máximo dos resultados da ordem lexicográfica estendida do descritor espectral onde $k \in \{50, 100, 200\}$ e n = 10

Quando o número de objetos de recuperação é pequeno, o resultado tende a ser melhor. Para os valores de recuperação testados, 5, 10 e 15, independente do valor do coeficiente de extração do descritor espectral, a maior precisão média é quando o número de objetos recuperados é igual a 5.

Descritor espectral: Melhor caso

A Figura 29 mostra o melhor resultado obtido para o descritor espectral, usando a ordem lexicográfica estendida como função de similaridade, onde o número de coeficientes para extração é igual a 50 e o número de objetos recuperados é igual a 5. Houve acurácia máxima de 100% em 15 das 19 classes testadas. Ao aumentar o número de objetos recuperados, observa-se uma ligeira queda na precisão, mas ainda mantém números elevados em relação às outras funções de similaridade.

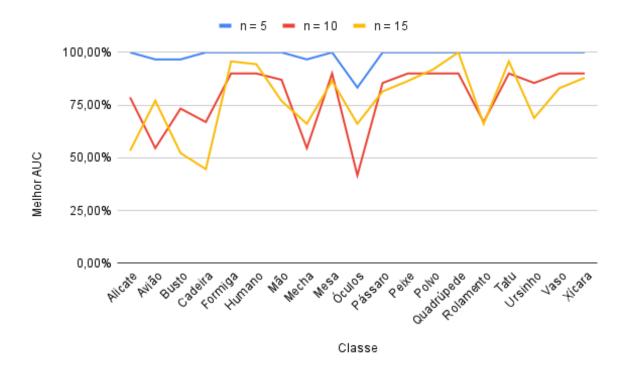


Figura 29 – Gráfico AUC máximo da ordem lexicográfica estendida do descritor espectral onde k=50 e $n\in\{5,10,15\}$

Descritor espectral: Pior caso

A Figura 30 apresenta o pior resultado obtido para o descritor espectral, utilizando o revlex como função de similaridade, onde o número de coeficientes para extração é igual a 200 e o número de objetos recuperados é igual a 5. Embora a classe Óculos tenha obtido uma precisão de 70% em um dos testes, o desempenho foi bastante ruim. 12 classes obtiveram acurácia zero e, mesmo aumentando o número de coeficientes no extrator espectral, não houve melhora significativa. Os resultados obtidos foram de 17%, 22% e 23% de acurácia média para o número de coeficientes igual a 50, 100 e 200, respectivamente.

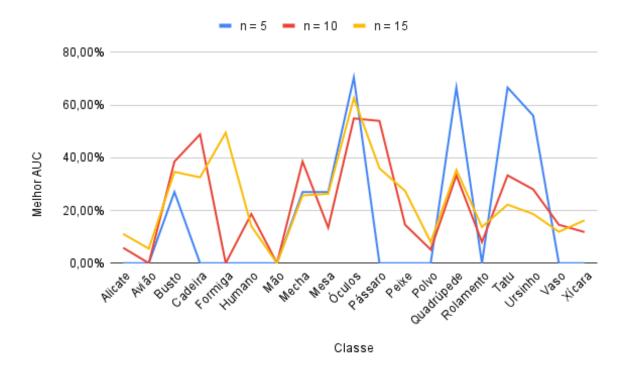


Figura 30 – Gráfico AUC máximo da ordem revlex do descritor espectral onde k=200 e $n \in \{5, 10, 15\}$

5.1.2 Resultados do descritor SPHARM

Usando a distância como função de similaridade, a distância de Manhattan e a distância euclidiana obtiveram uma precisão máxima de aproximadamente 16% nos casos em que o número de objetos recuperados é igual a 15. Esses foram seus melhores resultados. O melhor resultado para a ordem lexicográfica estendida alcançou uma média de 25% para n=5, conforme ilustrado na Figura 31.

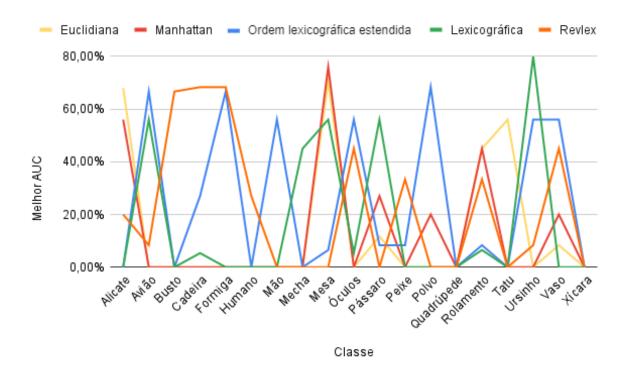


Figura 31 – Gráfico AUC máximo dos resultados das distâncias e ordens do descritor SPHARM onde n=5

Descritor SPHARM: Melhor caso

A Figura 32 mostra o melhor resultado para SPHARM, que foi obtido usando a ordem lexicográfica estendida como função de similaridade e o número de objetos recuperados é igual a 5. Houve uma acurácia máxima média de 25%.

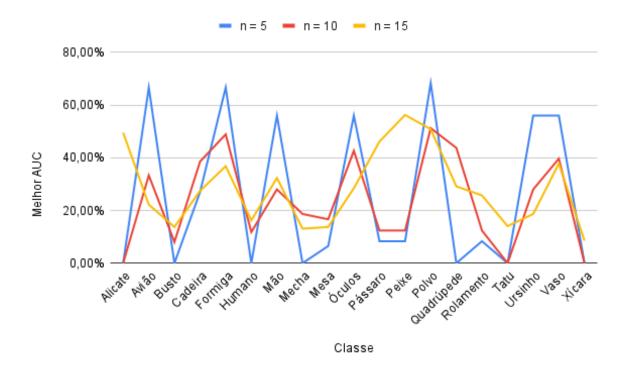


Figura 32 – Gráfico AUC máximo dos resultados da ordem lexicográfica estendida do descritor SPHARM onde $n \in \{5, 10, 15\}$

Descritor SPHARM: Pior caso

A Figura 33 apresenta o pior resultado para o descritor SPHARM, que foi obtido usando a distância euclidiana como função de similaridade. Para o pior resultado em que n=5, havia classes de 15 com precisão máxima igual a 0%.

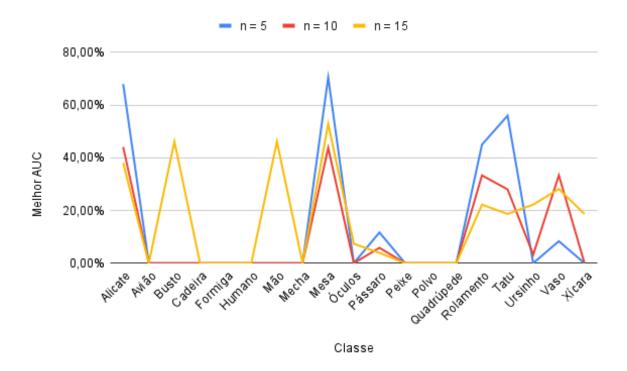


Figura 33 – Gráfico AUC máximo dos piores resultados da distância Euclidiana do descritor SPHARM onde $n \in \{5, 10, 15\}$

Para o caso de ordem lexicográfica estendida onde n = 5, três classes (mesa, alicate, pássaro) tiveram um resultado máximo de AUC médio de 4%. Analisando o vetor característico médio de cada uma das classes, percebe-se que elas não diferem claramente entre si. A imagem 34 demonstra o vetor médio de cada uma dessas três classes.

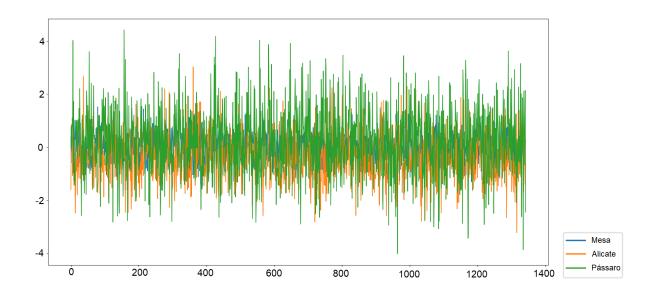


Figura 34 – Vetores médios das classes Pássaro, Alicate e Mesa.

5.1.3 Discussão dos resultados dos dois descritores

Para o conjunto de dados testado, o descritor espectral apresentou melhores resultados em relação ao descritor SPHARM. Para as funções de similaridade utilizadas, pode-se concluir que o sistema de ordem lexicográfica estendida obteve melhores resultados que a ordem lexicográfica, ordem lexicográfica reversa, distância euclidiana e distância de Manhattan, independente do descritor utilizado. Os melhores resultados espectrais e SPHARM são ilustrados em Figura 35, onde a função de similaridade é a ordem lexicográfica estendida e o número de objetos recuperados é 5. A acurácia máxima média do SPHARM foi 25% e 98% para o descritor espectral.

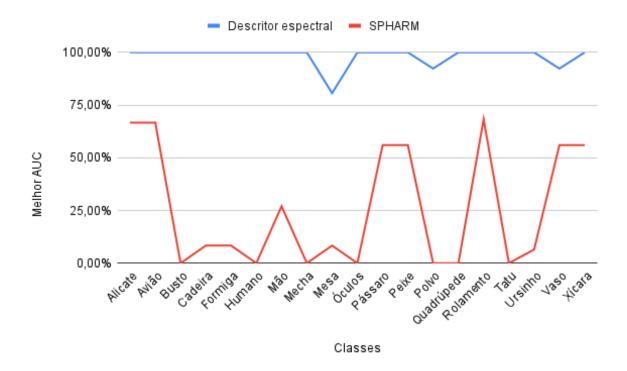


Figura 35 – Gráfico AUC dos resultados da ordem lexicográfica estendida dos resultados do descritor espectral e SPHARM onde n=5

A razão pela qual o sistema CBIR funciona melhor quando o número de objetos retornados pela pesquisa é menor é quase autoexplicativo. Quanto menor o número de objetos retornados, menor a probabilidade de erro. Obviamente, isso tem um limite e deve ser observado em comparação com outras funções de similaridade. Por exemplo, não se pode reduzir o número de objetos retornados a um único objeto, pois teríamos uma avaliação falsa de 100% de precisão para qualquer função de similaridade.

6 Conclusão

Neste trabalho foi apresentado um sistema CBIR inovador baseado em ordem em oposição aos sistemas tradicionais que usam funções de distância para avaliar a similaridade entre objetos. Devido ao mau resultado do descritor espectral em casos que as malhas possuem várias classes, obtendo precisão abaixo de 17%, em contraste do resultado de 85% obtido por Biscaro *et al.* (2016), surgiu a ideia de como aperfeiçoar o descritor para ter resultado melhor. Após realizar testes com outras funções de similaridade, foi implementada uma nova definição de ordem que obteve resultados de 63% em caso que anteriormente era 17%.

Esta nova ordem e a lexicográfica e lexicográfica reversa foram detalhadamente testadas com 380 objetos de 19 classes diferentes e comparadas, nas mesmas condições, com funções de distância tradicionais (Euclidiana e Manhattan). Para validar essa função de similaridade, foi usado também outro descritor, o SPHARM (harmônicos esféricos), que também apresentou melhoria nos resultados após implementarmos o sistema de ordenação. Resultados de distância Euclidiana apresentaram média de 13% enquanto a ordem lexicográfica estendida teve média de 25%. Independentemente do extrator de características utilizado, a ordem lexicográfica estendida obteve o melhor desempenho médio em todos casos, tendo um resultado melhor que as demais funções de similaridade, seja por distância vetorial ou ordenação. Além de avaliar uma nova função de similaridade, comparamos os dois extratores entre si.

Em particular neste conjunto de dados, pode-se notar que o extrator espectral teve melhores resultados que o SPHARM, mostrando-se um melhor extrator de características neste estudo de caso. A ordem lexicográfica estendida tem potencial para ser usada como uma função de similaridade como opção ao invés de apenas calcular a distância entre os vetores. A ordem, em resumo, é um jeito simples de priorizar características.

6.1 Próximos passos

Como trabalhos futuros, pode-se perguntar se os sistemas CBIR baseados em ordem podem ser usados em contextos diferentes das malhas geométricas, como imagens, sons e outros objetos que podem ser objetos de buscas de conteúdo. Além disso, é ressaltado que

este é um primeiro protótipo de um sistema CBIR baseado em ordens. Outras relações de ordem podem ser propostas e até mesmo sistemas que usam uma combinação de relações de ordem e funções de distância podem ser pensados. Em geral, o próprio sistema pode ser melhorado em vários aspectos.

Referências

- AK MUKHOPADHYAY S, D. A. G. M. K. N. D. Content-based image retrieval system for pulmonary nodules: Assisting radiologists in self-learning and diagnosis of lung cancer. *Journal of digital imaging*, v. 5, p. 63–77, 08 2017. Citado 2 vezes nas páginas 39 e 42.
- AND, C.-S. W. C.-H. H.; YU, M.-L. Automatic 3d cbir on kinematical human motion. In: . [S.l.]: WSEAS International Conference on Computer Engineering and Applications, 2007. Citado na página 40.
- AZNAG, K.; KANE, N.; OIRRAK, A.; BACHARI, E. E. New approach for 3d object recognition using non uniform fourier and moments coefficients. *Journal of Engineering Science and Technology*, v. 12, p. 1354–1365, 05 2017. Citado 2 vezes nas páginas 41 e 42.
- BASTOS, R.; COSTA, E. A. da. Colocando ordem nos complexos. *Colloquium Exactarum*, v. 4, p. 33–38, 2012. Citado na página 21.
- BERGAMASCO, L.; LIMA, K.; ROCHITTE, C.; NUNES, F. d. L. d. S. 3d medical objects retrieval approach using spharms descriptor and network flow as similarity measure. In: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). [S.l.: s.n.], 2018. p. 329–336. Citado na página 16.
- BERGAMASCO, L.; OLIVEIRA, H.; BÍSCARO, H.; WECHSLER, H.; NUNES, F. Using bipartite graphs for 3d cardiac model retrieval. In: . [S.l.: s.n.], 2015. v. 2015. Citado 2 vezes nas páginas 41 e 42.
- BERGAMASCO, L.; OLIVEIRA, R.; WECHSLER, H.; DAJUDA, C.; DELAMARO, M.; NUNES, F. Content-based image retrieval of 3d cardiac models to aid the diagnosis of congestive heart failure by using spectral clustering. In: . [S.l.: s.n.], 2015. Citado 2 vezes nas páginas 40 e 42.
- BERGAMASCO, L. C.; NUNES, F. Three-dimensional content-based cardiac image retrieval using global and local descriptors. In: . [S.l.]: AMIA Annu Symp Proc, 2015. v. 2015. Citado 2 vezes nas páginas 40 e 42.
- BERGAMASCO, L. C. C. Recuperação de objetos médicos 3D utilizando harmônicos esféricos e redes de fluxo. Tese (Doutorado) Escola Politécnica da Universidade de São Paulo, 2018. Citado 4 vezes nas páginas 33, 35, 42 e 51.
- BERGAMASCO, L. C.; NUNES, F. L. Intelligent retrieval and classification in three-dimensional biomedical images a systematic mapping. *Computer Science Review*, v. 31, p. 19–38, 2019. Citado 2 vezes nas páginas 15 e 27.
- BISCARO, H. H. Reconstrução a partir de pontos com a utilização de funções de Morse discretas. Tese (Doutorado) Instituto de Ciências Matemáticas e de Computação ICMC-USP, 2005. Citado 2 vezes nas páginas 21 e 24.
- BISCARO, H. H.; LIMA, J. P. Compressive representation of threedimensional models. *Journal on Interactive Systems*, v. 6, n. 1, p. 1–8, 2015. Citado na página 43.
- BISCARO, H. H.; OLIVEIRA, H.; BERGAMASCO, L. C. C.; NUNES, F. L. S. A new descriptor for retrieving 3d objects applied in congestive heart failure diagnosis. In: 2016

Referências 67

29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). [S.l.]: IEEE, 2016. p. 120–127. Citado 7 vezes nas páginas 15, 16, 31, 41, 42, 43 e 64.

- BRODERSEN, K. H.; ONG, C. S.; STEPHAN, K.; BUHMANN, J. The binormal assumption on precision-recall curves. In: . [S.l.: s.n.], 2010. p. 4263–4266. ISBN 1051-4651. Citado na página 52.
- CANDES, E.; ROMBERG, J. Sparsity and incoherence in compressive sampling. *Inverse Problems*, v. 23, p. 969–985, 2007. Citado 2 vezes nas páginas 15 e 30.
- DUTAGACI, H.; CHEUNG, C. P.; GODIL, A. Evaluation of 3d interest point detection techniques via human-generated ground truth. *The Visual Computer*, v. 28, p. 901–917, 2012. Citado na página 41.
- GAO Y. QIAN, R. H. M. L. R. C. B. B. A. C. J. R. X. Texture-based 3d image retrieval for medical applications. In: . [S.l.]: IADIS International Conference e-Health, 2010. Citado 3 vezes nas páginas 40, 42 e 43.
- HOANG, N. V.; GOUET-BRUNET, V. Manifold harmonic transform and spatial relationships for partial 3d object retrieval. In: . [S.l.: s.n.], 2014. Citado 3 vezes nas páginas 40, 41 e 42.
- KAMATH DISHA PUNJABI, T. S. D. U. S. S. M. Improving content based image retrieval using scale invariant feature transform. *International Journal of Engineering and Advanced Technology (IJEAT)*, v. 1, 2012. Citado 2 vezes nas páginas 39 e 42.
- KHAN, K. S.; KUNZ, R.; KLEIJNEN, J.; ANTES, G. Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, v. 96, p. 118–121, 2003. Citado na página 36.
- LAVOUE, G. Combination of bag-of-words descriptors for robust partial shape retrieval. *Vis Compu*, v. 28, p. 931–942, 2012. Citado 2 vezes nas páginas 41 e 42.
- LI, X.; YANG, J.; MA, J. Recent developments of content-based image retrieval (cbir). *Neurocomputing*, 01 2021. Citado na página 15.
- MADZIN, H. Information fusion of text and visual features for multi-modality medical information retrieval. Tese (Doutorado) University of Malaya, 2013. Citado na página 43.
- MARQUES, F. de Lourdes dos S. N. Introducao ao processamento de imagens medicas para auxilio ao diagnostico. Tese (Doutorado) Pontificia Universidade Catolica, 2006. Citado na página 28.
- MARQUES, O.; VIEIRA, H. *Processamento de imagens*. [S.l.]: Rio de Janeiro: Brasport, 1999. Citado na página 43.
- NGUYEN, V. H.; GOUET-BRUNET, V. A new description for scalable 3D partial object retrieval. [S.l.], 2014. Disponível em: (https://hal.archives-ouvertes.fr/hal-01126348). Citado na página 43.
- PENG, J.; KIM, C.-S.; Jay Kuo, C.-C. Technologies for 3d mesh compression: A survey. Journal of Visual Communication and Image Representation, v. 16, n. 6, p. 688 – 733, 2005. ISSN 1047-3203. Disponível em: (http://www.sciencedirect.com/science/article/pii/S1047320305000295). Citado na página 24.

Referências 68

PHAM, H. D. Computer Vision: Image Shape Geometry and Classification. 2018. Citado na página 43.

- QIAN RUI HUI, X. G. Y. 3d cbir with sparse coding for image-guided neurosurgery. Signal Processing, v. 93, p. 1673–1683, 2012. Citado 4 vezes nas páginas 39, 40, 42 e 43.
- QIAN X. GAO, M. L. R. C. B. B. Y. Content-based retrieval of 3d medical images. In: . [S.l.]: The Third International Conference on eHealth, Telemedicine, and Social Medicine, 2011. Citado 3 vezes nas páginas 40, 42 e 43.
- Samoun, L.; Fisichella, T.; Lingrand, D.; Malleus, L.; Precioso, F. An interactive content-based 3d shape retrieval system for on-site cultural heritage analysis. In: 2018 25th IEEE International Conference on Image Processing (ICIP). [S.l.: s.n.], 2018. p. 1043–1047. Citado 2 vezes nas páginas 41 e 42.
- SCHULZ, A.; SILVA, E. A. B. da; VELHO, L. Compressive sensing. In: . [S.l.]: 27 Coloquio Brasileiro de Matematica SBM, 2009. Citado na página 31.
- SHREC. Shape Retrieval Contest for CAD models 2007. 2007. Disponível em: \(\text{https://engineering.purdue.edu/PRECISE/shrec} \) Citado na página 45.
- SIPIRAN, I.; BUSTOS, B. Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visual Computer*, v. 28, p. 963–976, 2011. Citado na página 41.
- WANG, S. A Robust CBIR Approach Using Local Color Histograms. Tese (Doutorado) University of Alberta, 2001. Citado 2 vezes nas páginas 15 e 27.
- XW QIAN Y, H. R. G. The state of the art of medical imaging technology: from creation to archive and back. *The open medical informatics journal*, v. 5, p. 73–85, 2011. Citado 2 vezes nas páginas 40 e 42.
- YU, F.; LU, Z.; LUO, H.; WANG, P. *Three-Dimensional Model Analysis and Processing*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2011. ISBN 3642126502. Citado na página 43.