TÉCNICO LISBOA

RISCOS – Recuperação de Imagens usando Esboços

Luís Carlos Mendes Rosado

Dissertação para obtenção do Grau de Mestre em

Engenharia Informática e de Computadores

Orientador: Prof. Manuel João Caneira Monteiro da Fonseca

Júri

Presidente: Prof. João Paulo Marques da Silva
Orientador: Prof. Manuel João Caneira Monteiro da Fonseca
Vogal: Prof. Alfredo Manuel dos Santos Ferreira Júnior

Maio 2014

“A inteligência não ensina a aprender muitas coisas.”

Heraclito

Agradecimentos

Os meus agradecimentos são dirigidos em primeiro lugar ao meu orientador Professor Doutor Manuel João da Fonseca pelo apoio, disponibilidade e partilha de conhecimentos que transmitiu ao longo de toda a dissertação. De referir também a confiança demonstrada em mim para a realização do projecto *Crush* o qual teve uma estreita ligação com a minha dissertação. Queria ainda agradecer ao Ricardo Dias a disponibilidade demonstrada nos momentos cruciais da minha dissertação, ajudando-me a superar os problemas que foram surgindo ao longo da realização da mesma.

Agradeço também ao Eduardo, pelo companheirismo e a permanente colaboração na realização de alguns trabalhos que fomos realizando em conjunto ao longo do mestrado.

Gostaria também de agradecer à minha família pelo apoio força e motivação dado ao longo do meu percurso académico.

Por último um agradecimento muito especial à Joana, que me acompanhou desde o início da minha licenciatura, até à formação do mestrado. À Joana agradeço todo o apoio, colaboração, paciência, e ajuda que me prestou ao longo do curso, como a sua constante presença nesta etapa académica da minha vida.

Resumo

Actualmente armazenamos muitas imagens, devido por um lado, à enorme disponibilidade de dispositivos de captura (telemóveis, máquinas fotográficas) e por outro, ao crescente aparecimento diário de imagens digitais. Devido ao elevado número de imagens, torna-se uma tarefa difícil para os utilizadores, localizar as imagens pretendidas.

Existindo soluções que permitem a realização de pesquisas baseadas em texto e de imagens exemplo, pretendemos desenvolver uma solução que permita a recuperação e pesquisa de imagens *clip-art* usando esboços como interrogação. Dada a ausência desta solução, procuramos na realização do nosso trabalho obter uma representação comum entre um esboço e uma imagem para que seja possível realizar uma comparação entre ambas. Para que a nossa solução funcione, identificamos as características nas imagens *clip-art* e esboços. Seguidamente convertemo-las em descritores, onde no descritor *GeoDiff* colocamos a informação topológica e geométrica dos objectos, enquanto que no descritor de *Geometria*, apenas colocamos a informação referente à geometria do contorno global das imagens *clip-art* e esboço. Obtidos os descritores, procedemos à comparação destes, entre as imagens *clip-art* e o esboço.

Realizada a nossa solução, procedemos à avaliação desta, consistindo numa análise das imagens *clip-art* recuperadas através de esboços e imagens *clip-art* como interrogações. Obtidos os resultados, verificámos que a combinação de cada um dos algoritmos *GeoDiff* e *TopoGeo* com a *Geometria* proporcionaram bons resultados na recuperação das imagens *clip-art*.

Palavras Chave

Recuperação de imagens

Interrogações usando esboços

Detecção de contornos

Simplificação de imagens

Abstract

Nowadays we store a lot of images due to, on one hand, to the huge availability of capture devices and on the other hand, to the daily growing appearance of digital images. Due to this high number of images, it becomes a hard task to the users to locate their pretended images.

Since it exist solutions that allow the search of text as well as examples images, we pretend to develop a solution that allows the recovery and search of *clip-art* images using sketches as queries. Given the lack of this solution, we search as the achievement of our work, to obtain a common representation between a sketch and an image, so that it can be possible to perform a comparison between both. For our solution to work, we identified the characteristics in the *clip-art* images and sketches. After that, we convert them to descriptors, where in the descriptor *GeoDiff* we put the topological and geometric information of the objects, whereas in the *Geometric* descriptor, we only put the information regarding the global geometric outline of the *clip-art* images and sketches. After obtaining the descriptors, we proceed to the comparison of them, between *clip-art* images and sketches.

After achieving our solution, we proceeded to the evaluation, consisting on an analysis of the *clip-art* images recovered by the sketches and *clip-art* images as queries. After obtaining the results, we verified that the combination of the algorithms *GeoDiff* and *TopoGeo* with the *Geometric*, produced good results in the recovery of *clip-art* images.

Keywords

Recovery of images

Queries using sketches

Edge detection

Simplification of images

Índice

1 Introdução 1
1.1 Descrição do problema 1
1.2 Objectivos 2
1.3 Solução Desenvolvida 2
1.4 Contribuição e Resultados Atingidos 4
1.5 Organização da dissertação 5
2 Trabalho Relacionado 6
2.1 Recuperação de Imagens através de Esboços 6
2.2 Detecção de contornos em Imagens 14
2.3 Síntese 17
3 Arquitectura da solução 18
3.1 Arquitectura global 18
3.2 Conversão de imagem *raster* para imagem *vectorial* 20
3.2.1 Pré-Processamento 21
3.2.1.1 Combinação de Técnicas 21
3.2.1.2 Filtro de Canny 24
3.2.2 Processo de Simplificação 26
3.3 Conversão de esboço *raster* para esboço *vectorial* 28
3.3.1 Pré-Processamento 29
3.3.1.1 Filtragem detalhada 30
3.3.1.2 Filtragem Simplificada 31
3.4 Extracção do Descritor 33
3.4.1 GeoDiff 34
3.4.2 Geometria 36
3.4.3 Recuperação das imagens 37
3.4.3.1 Inserção do descritor 37
3.4.3.2 Comparação 38
3.5 Interface RISCOS 38
3.6 Síntese 41
4 Avaliação 43

4.1 Análise experimental dos filtros 43
4.1.1 Análise da escolha dos filtros 44
4.2 TopoGeo 48
4.2.1 Descrição do TopoGeo vs GeoDiff 48
4.3 Procedimento para avaliação da recuperação 50
4.3.1 Medidas de Qualidade 50
4.3.2 Resultados da recuperação usando imagens como interrogação 52
4.3.2.1 Precision & Recall 52
4.3.2.2 Mean Average Precision 54
4.3.2.3 R-Precision 57
4.3.3 Resultados da recuperação usando esboços como interrogações 59
4.3.3.1 Precision & Recall 59
4.3.3.2 Mean Average Precision 62
4.3.3.3 R-Precision 64
4.4 Discussão dos resultados 65
4.5 Síntese 66
5 Conclusões e Trabalho Futuro 68
5.1 Sumário da Dissertação 68
5.2 Conclusões Finais e Contribuições 71
5.3 Trabalho Futuro 72
6 Referências Bibliográficas 74

Lista de Figuras

Figura 2.1. Imagens com a cruz amarela não são consideradas semelhantes ao esboço. As imagens com um círculo a vermelho apesar de não se enquadrarem com o tema da pesquisa têm semelhanças com o esboço. As restantes imagens, são o resultado da semelhança existente com a imagem esboçada 7
Figura 2.2. (a) Esboço, (b) Esboço *Binarizado*, (c) *Chain Code* 7
Figura 2.3. Centróide •, Vector →, Pontos Dominantes ■ 8
Figura 2.4. Resultado das cinco imagens mais parecidas com o esboço 10
Figura 2.5. Processo de extracção de características 11
Figura 2.6. (a) Imagens e esboços simples. Aplicação do *operador Gaussiano de Laplace*, (b) Localização do objecto na Imagem 12
Figura 2.7. Efeito do parâmetro β para a detecção dos contornos 13
Figura 2.8. (a) Aplicação do *Filtro de Gabor*, (b) *Transformada de Radon* na imagem 14
Figura 2.9. Imagens (a) binarizada, (b) com o maior objecto, (c) contorno do objecto 14
Figura 2.10. Representação do histograma que contêm o número de contornos existentes em cada bloco da imagem. À direita as cinco máscaras representativas da direcção dos contornos 15
Figura 2.11. (a) Imagem Original em RGB, (b) Imagem Binarizada (c) Aplicação do Algoritmo *Moore Neighbor* 16
Figura 2.12. (a) Imagem Original a níveis de cinzento com histograma, (b) Imagem suavizada com histograma 16
Figura 3.1. Arquitectura da solução 19
Figura 3.2. Arquitectura da imagem 21
Figura 3.3. Etapas aplicadas à imagem através da aplicação *Combinação de Técnicas* 22
Figura 3.4. Imagens a processar (originais) 23
Figura 3.5. Imagens obtidas do *Filtro Gaussiano com Blur* 23
Figura 3.6. Imagens obtidas da *Binarização* 24
Figura 3.7. Imagens obtidas do *Encontra Contorno* 24
Figura 3.8. Etapas aplicadas à imagem através da aplicação do *Filtro de Canny* 24
Figura 3.9. Imagens a processar (originais) 25
Figura 3.10. Imagens obtidas do *Filtro de Canny* 25
Figura 3.11. Imagens obtidas da *Subtracção* 25
Figura 3.12. Etapas aplicadas à imagem através do Processo de Simplificação 26
Figura 3.13. Imagens provenientes da aplicação do *Filtro de Canny* após efectuado um *Threshold* 27
Figura 3.14. Imagens provenientes da aplicação do *Filtro de Canny* após efectuada uma *Operação de Fecho* 27
Figura 3.15. Imagens provenientes da aplicação da *Combinação de Técnicas* após efectuado um *Threshold* 27
Figura 3.16. Imagens provenientes da aplicação *Combinação de Técnicas* após efectuada uma *Operação de Fecho* 27
Figura 3.17. Imagens provenientes da aplicação do *Filtro de Canny* após efectuada uma *Esqueletização* 28
Figura 3.18. Imagens provenientes da aplicação *Combinação de Técnicas* após efectuada uma *Esqueletização* 28
Figura 3.19. Arquitectura do esboço 29

Figura 3.20. Etapas aplicadas ao esboço através da *Filtragem detalhada* 30
Figura 3.21. Esboços utilizados como interrogação 31
Figura 3.22. Esboços obtidos do *Filtro Gaussiano com Blur* 31
Figura 3.23. Esboços obtidos da *Binária* 31
Figura 3.24. Esboços obtidos da *Esqueletização* 31
Figura 3.25. Etapas aplicadas ao esboço através da *Filtragem Simplificada* 32
Figura 3.26. Esboços utilizados como interrogação 32
Figura 3.27. Esboços obtidos do *Filtro Gaussiano com Blur* 32
Figura 3.28. Esboços obtidos da *Binária* 33
Figura 3.29. Esboços obtidos do *Fecha Buracos* 33
Figura 3.30. Esboços obtidos do *Encontra Contornos* 33
Figura 3.31. Esboços obtidos da *Esqueletização* 33
Figura 3.32. Ilustração de uma imagem composta por objectos inclusos e adjacentes 35
Figura 3.33. Formação do grafo a partir de uma imagem 35
Figura 3.34. Processo utilizado para calcular o descritor da imagem e do esboço 36
Figura 3.35. Imagem composta apenas pelo contorno global do objecto 36
Figura 3.36. Visualização do Menu Inicial 39
Figura 3.37. Visualização do Menu *Query By Image* 40
Figura 3.38. Visualização dos Resultados após efectuada uma *Query By Image* 40
Figura 3.39. Visualização do Menu *Query By Sketch* 41
Figura 4.1. Formação do grafo usando o *TopoGeo* 49
Figura 4.2. Formação do grafo usando o *GeoDiff* 49
Figura 4.3. Gráfico da *Precision & Recall* 53
Figura 4.4. Gráfico das categorias *Precision & Recall* 54
Figura 4.5. Gráfico da *Mean Average Precision* 55
Figura 4.6. Gráfico das categorias *Mean Average Precision* 56
Figura 4.7. Gráfico da *R-Precision* 57
Figura 4.8. Gráfico das categorias *R-Precision* 58
Figura 4.9. Gráfico da *Precision & Recall* 60
Figura 4.10. Gráfico das categorias *Precision & Recall* 61
Figura 4.11. Gráfico da *Mean Average Precision* 62
Figura 4.12. Gráfico das categorias *Mean Average Precision* 63
Figura 4.13. Gráfico da *R-Precision* 64
Figura 4.14. Gráfico das categorias *R-Precision* 65

Lista de Tabelas

Tabela 4.1. Resultado da análise dos diferentes Filtros e técnicas.....45

1 Introdução

Neste capítulo descrevemos os problemas identificados nas áreas da *recuperação de imagens* e *detecção de contornos*. Apresentamos os objectivos e resumimos a nossa solução, como citamos as contribuições e os resultados atingidos. Concluímos este capítulo com o alinhamento dado à nossa dissertação.

1.1 Descrição do problema

Devido à elevada quantidade de imagens acessíveis, existe a necessidade de inovar a tecnologia relacionada com a sua pesquisa. Actualmente é possível realizar a pesquisa de imagens através de texto e de imagens de exemplo. Os utilizadores conseguem a partir de palavras-chave encontrar imagens e podem, inclusive, adicionar cor para traduzir melhor o seu desejo. É também possível realizar a pesquisa a partir de uma imagem, e através desta, encontrar imagens que sejam semelhantes à imagem de interrogação.

As pesquisas às quais os utilizadores têm acesso tornam-se difíceis por estes não conseguirem exprimir em palavras o que pretendem, ou por não terem uma imagem exemplo que identifique e exprima os seus pensamentos. Devido a isto, os resultados nem sempre são os mais desejados, apresentando-se ineficazes pelas informações que dão, e por não estarem relacionados com a pesquisa pretendida pelo utilizador.

Áreas como a recuperação de imagens e detecção de contornos em imagens, têm explorado soluções que desenvolvam e melhorem as pesquisas existentes, mas pecam por serem limitadoras e não solucionarem o problema. A limitação destas técnicas deve-se à falta de uma solução, capaz de realizar pesquisas de imagens através da interrogação de esboços.

Havendo a necessidade de explorar outras formas de pesquisa, que se enquadrem com a percepção visual do utilizador e facilitem a especificação das interrogações, explorámos uma nova solução para ultrapassar estes desafios. A solução que desenvolvemos conciliou as duas áreas, *recuperação de imagens através de esboços* e *técnicas de detecção de contornos em*

imagens, possibilitando a realização de pesquisa de imagens *clip-art* usando esboços como interrogação.

1.2 Objectivos

Como mencionado, torna-se essencial aos utilizadores efectuarem pesquisas de imagens que se enquadrem com a sua percepção visual. Devido a este facto, e motivados pela necessidade da existência de uma aplicação que faça uma pesquisa de imagens, pretendemos desenvolver uma solução que permita aos utilizadores encontrar imagens *clip-art* através de um esboço, facilitando a pesquisa das imagens *clip-art* que queremos encontrar numa colecção de imagens *clip-art*. Para atingirmos este objectivo, tornar-se-á necessário encontrar uma representação que seja comum e que possibilite a comparação entre imagens e esboços. A partir desta ideia, estudámos soluções que faziam *recuperação de imagens através de esboços* e abordagens que convertiam imagens em esboços, através da *detecção de contornos*, com o objectivo de encontrar uma representação comum.

1.3 Solução Desenvolvida

Nesta dissertação propomos uma solução que combina duas áreas, a *detecção de contornos* e a *recuperação de imagens*.

Focando-se a nossa solução numa aplicação inovadora, que permita aos utilizadores realizarem as pesquisas de imagens tal como as percepcionam visualmente, desenvolvemos uma solução baseada em dois tipos de pesquisa de imagens, relacionando-se uma em imagem exemplo e outra em esboços. Proporcionadas estas duas formas de pesquisa, o utilizador terá um maior controlo sobre a sua pesquisa, na qual poderá optar por realizar uma interrogação através de uma imagem *clip-art*, ou por um desenho.

A solução desenvolvida divide-se em duas partes principais, ou seja, à imagem aplicámos um pré-processamento, seguido de uma simplificação e conversão da imagem *raster* para um

formato *vectorial*. Posteriormente inserimos a informação extraída da imagem numa base de dados. Ao esboço efectuámos um pré-processamento, convertemo-lo para um formato *vectorial* e extraímos a informação. Concluídas as etapas tanto na imagem como no esboço, calculámos a similaridade e apresentamos os resultados.

Realizámos um pré-processamento à imagem, consistindo na ramificação de duas diferentes filtragens, numa identificámos todos os contornos presentes na imagem, enquanto na outra procurámos apenas evidenciar o contorno global da imagem. Seguidamente, realizámos uma simplificação a ambas as filtragens, focando-se esta numa combinação de técnicas que tornaram possível efectuar uma conversão da imagem *raster* para um formato *vectorial*.

Concretizada a conversão, procedemos à extracção da informação da imagem através da implementação de dois algoritmos: *GeoDiff* e *Geometria*. Em relação ao *GeoDiff*, este foi aplicado às imagens, provenientes da obtenção de todos os contornos dos objectos. Este algoritmo consistiu na obtenção de um grafo representativo das relações espaciais e diferenças geométricas entre os elementos. Por seu lado, o algoritmo de *Geometria* calcula a geometria do contorno global da imagem.

Em relação ao esboço, este também foi exposto a um pré-processamento idêntico ao da imagem, consistindo este processo em duas filtragens diferentes. Numa identificámos todos os contornos presentes no esboço, preservando o detalhe e conservando toda a informação do desenho, enquanto na outra eliminámos todo o detalhe, obtendo apenas o contorno que envolve o objecto do esboço. Seguidamente, convertemos o esboço para um formato *vectorial* e extraímos as características do mesmo. Esta extracção foi realizada similarmente à da imagem, aplicando-se o *GeoDiff*, aos esboços em que obtivemos todos os contornos, extraindo-se a topologia e a diferença geométrica dos elementos que formam os esboços. A extracção da informação geométrica do objecto, resultante da aplicação da *Geometria*, efectuou-se aos esboços em que apenas obtivemos o contorno global dos mesmos.

Concluídas as etapas na imagem e no esboço, efectuámos uma comparação entre ambos. Esta comparação foi realizada recorrendo a um cálculo de similaridade, sendo efectuado um emparelhamento entre as imagens e os esboços, que sofreram processos idênticos, ou seja, comparámos imagens em que apenas evidenciámos o contorno global, com os esboços em que apenas identificámos o contorno envolvente do objecto. Às imagens em que obtivemos todos os seus contornos, efectuámos a comparação com esboços resultantes do mesmo procedimento.

Posteriormente procedemos a uma ordenação dos valores de similaridade, e apresentamos aos nossos utilizadores os melhores resultados.

1.4 Contribuição e Resultados Atingidos

Neste trabalho, a nossa contribuição focou-se na obtenção de uma solução que permita aos utilizadores efectuarem pesquisa de imagens *clip-art* através da realização de esboços, proporcionando pesquisas mais eficazes e precisas.

Baseando-se a nossa solução no recurso de técnicas e implementação de algoritmos, procurámos com a combinação destes, obter imagens aproximadas de um esboço, para posteriormente efectuarmos uma comparação entre imagens e esboços. Posto isto, e recorrendo à área de *detecção de contornos*, procurámos efectuar às imagens e aos esboços um pré-processamento, consistindo o mesmo na combinação de técnicas que simplifiquem a imagem, resultando deste modo imagens aproximadas de um esboço. Posteriormente, e através da implementação de algoritmos de extracção de informação, obtivemos a relação espacial e geométrica das imagens e dos esboços.

Obtidas as informações necessárias relacionadas com as imagens e os esboços, recorremos à área da *recuperação de imagens* a qual nos facultou técnicas preponderantes no armazenamento da informação mas também na realização da pesquisa de imagens através de esboços.

Realizada a nossa solução, avaliámos a mesma com recurso a métricas relacionadas com a área da *recuperação da informação*. Esta avaliação, consistiu na comparação do comportamento dos algoritmos *GeoDiff* e *TopoGeo*, e na combinação de cada um destes com a *Geometria*. Efectuada a apreciação dos resultados, verificámos que o algoritmo, *GeoDiff*, consegue resultados próximos do *TopoGeo*, com a vantagem de ser menos complexo e de exigir menos poder computacional. Adicionalmente efectuamos a comparação de cada um dos algoritmos, *GeoDiff* e *TopoGeo* com a *Geometri*a, constatando que as combinações efectuadas obtiveram melhores resultados. Tal facto deve-se à combinação proporcionada pela geometria com a topologia e diferença geométrica entre objectos, face a um algoritmo que apenas utiliza a relação topológica e a diferença geométrica entre os elementos. Dados os

resultados obtidos, conseguimos provar através da implementação do algoritmo *GeoDiff*, o qual agrupa num grafo toda a informação topológica e geométrica dos objectos, que é possível obter resultados semelhantes a um algoritmo que se apresenta mais complexo e mais pesado em termos computacionais, *TopoGeo*, necessitando este, de conter à parte a informação geométrica. Combinando cada um dos algoritmos com a *Geometria*, demonstramos que a junção destes, tornou-se mais proveitoso para a nossa solução, e por conseguinte um maior número de imagens relevantes.

1.5 Organização da dissertação

Na secção seguinte, apresentamos os trabalhos estudados nas áreas de, *recuperação de imagens através de esboços* e técnicas de *detecção de contornos em imagens*. Efectuamos uma descrição e uma análise às soluções existentes, para evidenciarmos as potencialidades e falhas das técnicas utilizadas. Seguidamente, no capítulo 3, apresentamos a arquitectura da nossa solução e descrevemos as diferentes fases desenvolvidas para a nossa solução. No capítulo 4, apresentamos os métodos de avaliação utilizados e os resultados obtidos. No último capítulo, apresentamos as conclusões do trabalho fazendo uma reflexão ao trabalho desenvolvido, e propomos algum trabalho futuro.

2 Trabalho Relacionado

Neste capítulo apresentamos alguns trabalhos relacionando-os com o objectivo da nossa solução. Como descrito no capítulo anterior, o nosso objectivo focou-se na recuperação de imagens através de esboços. Existindo soluções centradas nas áreas da *detecção de contornos* e *recuperação de imagens*, procedemos à análise das mesmas, evidenciando as principais lacunas mas também realçando as contribuições essenciais para a realização da nossa solução.

2.1 Recuperação de Imagens através de Esboços

Os autores Yang Cao et al. descrevem uma técnica que converte uma imagem num esboço a partir da detecção de contornos [1]. Posteriormente extraem as características mais relevantes da imagem e comparam com o esboço desenhado pelo utilizador. A técnica escolhida para a comparação entre as imagens foi o *Chamfer Matching*, e as suas variantes *Oriented Chamfer Matching* por apresentar uma comparação de contornos eficiente. A técnica *Chamfer Matching* consiste no cálculo da distância entre o contorno da imagem e o contorno do esboço. A técnica *Oriented Chamfer Matching* é utilizada para detectar a orientação do contorno da imagem em relação ao esboço. A cada imagem é atribuída um mapa de contornos, constituída por seis tipos de orientação: de $-15°$ a $15°$, $15°$ a $45°$, $45°$ a $75°$, $75°$ a $105°$, $105°$ a $135°$, e de $135°$ a $165°$. Consoante a semelhança existente entre a direcção de contorno da imagem e o esboço, é atribuído um *hit* ao contorno correspondente, sendo utilizada a técnica *Indexable Oriented Chamfer Matching*. Esta técnica é um *array* de índices invertidos composta por todas as imagens, às quais é calculada a similaridade entre o esboço e todas as imagens.

A utilização destas técnicas permite processar uma imagem num formato semelhante ao de um esboço, sendo um factor a considerar para a nossa solução. É perceptível, a partir dos resultados produzidos pelo uso desta solução, que os mesmos nem sempre são eficientes (ver figura 2.1).

Figura 2.1. Imagens com a cruz amarela não são consideradas semelhantes ao esboço. As imagens com um círculo a vermelho apesar de não se enquadrarem com o tema da pesquisa têm semelhanças com o esboço. As restantes imagens, são o resultado da semelhança existente com a imagem esboçada

Em [2], Ricardo Santos faz a descrição de como extrair os contornos de um esboço. Começa por transformar o esboço numa imagem binária, para aperfeiçoamento dos contornos desenhados. Tendo o esboço binarizado é construída uma matriz na qual são etiquetados com caracteres diferentes os pixéis que representam o contorno, e os pixéis que representam o fundo do esboço. Obtida a matriz, é utilizada uma janela de 3 x 3, que percorrerá no sentido dos ponteiros do relógio todos os pixeis para identificar os contornos do objecto e as suas direcções. Através do auxílio da técnica de *Chain Code*, é realizada uma sequência de caracteres que formará o tipo de contorno esboçado pelo utilizador.

No nosso entender este processo poderá ser-nos útil, apesar deste trabalho ter sido realizado apenas para a comparação de formas geométricas (ver figura 2.2). A utilização da técnica *Chain Code* permite a identificação do caminho do contorno, como o reconhecimento da forma esboçada pelo utilizador.

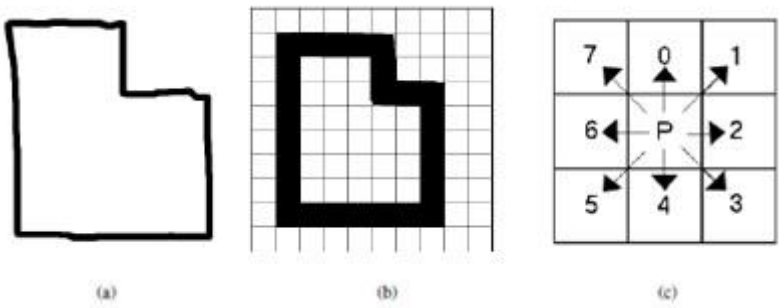


Figura 2.2. (a) Esboço, (b) Esboço *Binarizado*, (c) *Chain Code*

Os autores Horace IP et al., em [3] descrevem uma técnica que permite a partir de um esboço extrair o contorno, para posteriormente comparar com uma imagem. A técnica examina todos os pixéis que formam o objecto. Este processo começa por identificar os pontos dominantes, considerados pixéis que formam a concavidade do contorno. No caso de não existirem objectos que apresentem formas convexas, como círculos, estes são divididos em quatro partes. Para cada parte são encontrados os pixéis dominantes. Após estes pontos terem sido encontrados, calcula-se o seu centróide. A partir das coordenadas que formam o centróide, é calculado o segmento de recta que liga o centróide aos pontos dominantes do objecto. Ao ligar-se o centro de massa e dois pontos dominantes, obtém-se um triângulo (ver figura 2.3). Obtida esta forma geométrica, faz-se a sua normalização dividindo a sua área pela área total do objecto.

A seguir constrói-se um histograma em que a cada ângulo formado pelos vectores, entre o centróide e os pontos dominantes, corresponde o valor da área normalizada. Ao obter-se o histograma é utilizada uma medida de similaridade, consistindo a mesma na diferença entre o histograma do esboço, e o histograma da imagem, respectivamente.

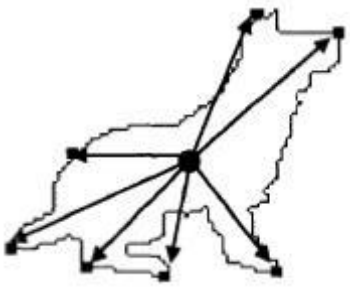


Figura 2.3. Centróide •, Vector →, Pontos Dominantes ■

Mathias Eitz et al. utilizam uma tabela de índices invertidos para pesquisar as suas imagens numa base de dados [4]. A implementação desta tabela agiliza a comparação entre as imagens e o esboço, estando as imagens organizadas por semelhança. Os autores para compararem a semelhança entre o esboço e as imagens contidas na base de dados, utilizam como medida de similaridade a *tf-idf,* (*term frequency–inverse document frequency*), correspondendo os elementos visuais da imagem a palavras. A utilização desta medida consiste na relevância de uma palavra, caso esta apareça em quantidade elevada no esboço (*high term frequency*). No

sentido inverso, e no caso de uma palavra não aparecer com elevada relevância, é considerada (*inverse document frequency*). Através dos pesos obtidos é identificada a semelhança entre o esboço e a imagem.

Folco Banfi descreve em [5], vários tipos de extracção de informação para uma imagem. O autor define como extracção automática, o cálculo da média das cores de uma imagem, realização de um histograma com as cores que compõem a imagem, ou mesmo o cálculo da covariância de uma cor de uma determinada região. Um outro tipo de extracção também referido é a análise do tamanho da região de uma imagem. O autor descreve também algumas medidas de qualidade para a análise de imagens, que permite medir a qualidade dos resultados de uma recuperação de imagens, sendo estas o *precision* e o *recall*.

A utilização destas medidas pode contribuir para o nosso trabalho. A realização de medidas que estejam relacionadas com a informação da imagem, será útil para entender a semelhança que haja entre um esboço e as imagens da base de dados. Em relação às medidas de *precision* e *recall*, estas são utilizadas nomeadamente em recuperação de informação. Devido ao nosso trabalho incidir também na extracção de informação, relativamente às imagens, estas medidas podem contribuir na averiguação se a técnica utilizada obteve os resultados esperados na semelhança entre imagens.

Konstantinos Bozas e Ebroul Izquierdo utilizam um algoritmo de *Min-Hash* para a comparação da similaridade entre uma imagem em esboço e as imagens de uma base de dados [6]. Este algoritmo consiste na análise de ambas as imagens, necessitando estas de estar binarizadas. Este facto deve-se aos valores resultantes do algoritmo, terem que estar compreendidos entre zero e um. Caso o resultado obtido pelo algoritmo seja zero, significa que os elementos em análise são considerados disjuntos. Caso o valor seja um, os elementos são iguais. Não se verificando as condições anteriores, este valor não tem significado. Na realização deste algoritmo é considerada uma janela para analisar cada bloco da imagem do esboço e da imagem real. Um índice invertido é construído para armazenar o único *min-hash* extraído da imagem real. O mesmo processo é realizado para a imagem em esboço, indo comparar o valor extraído com o da imagem, e no caso de haver semelhança, será adicionado um *hit* à imagem correspondente. Através da fórmula do algoritmo analisam-se os resultados e verifica-se a existência de similaridade.

O algoritmo *Min-Hash* é uma técnica que permite obter uma análise sobre a similaridade entre imagens, podendo contribuir para o nosso trabalho. Através dos resultados obtidos (ver figura

2.4), podemos desta forma, apresentar apenas as imagens que se assemelham com o desenho esboçado pelos nossos utilizadores.

Figura 2.4. Resultado das cinco imagens mais parecidas com o esboço

Os autores Yixin Chen et al. descrevem uma técnica que permite a pesquisa de uma imagem numa colecção de imagens [7]. Embora esta técnica não esteja directamente relacionada com o objectivo principal do trabalho, descreve uma técnica de segmentação de imagens que pode ser útil para o nosso trabalho.

Todas as imagens, incluindo a imagem de interrogação, são divididas em blocos de 4x4, em que a cada bloco são extraídas seis características. Três características estão relacionadas com a média das cores, em que é utilizado um espaço de cor LUV, correspondendo a componente L à codificação da luminância e as componentes U e V à codificação da informação da cor (crominância). As outras três características são consideradas altas frequências da transformada de *wavelet*. Estas características são obtidas pela aplicação desta, transformada na componente L da imagem. Aplicada esta transformação, decompõe-se a mesma em quatro bandas de frequências, LL (baixa baixa), LH (baixa alta), e HL (alta baixa) e HH (alta alta). A cada banda correspondem coeficientes de 2x2, sendo estas bandas utilizadas para a extracção da textura da imagem. É utilizado um algoritmo de k-médias, que permite o agrupamento dos vectores de características em classes. A cada classe são ainda acrescentadas outras duas

características, centróide e inércia, provenientes das propriedades que constituem cada região da imagem (ver figura 2.5).

A similaridade entre as duas imagens é calculada através da soma das distâncias do espaço de características entre todas as regiões diferentes que compõem a imagem.

A utilização da técnica permite a obtenção de características que formam a imagem e um esboço. Tal como descrito, agrupam-se as características em classes e assim podemos realizar uma comparação das características da imagem com as do esboço.

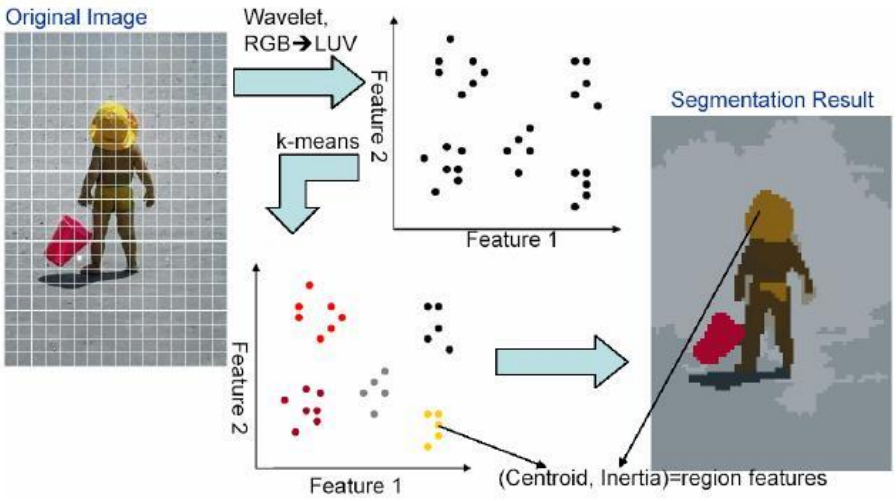


Figura 2.5. Processo de extracção de características

Em [8] John Collossome, et al. descrevem uma técnica que através de um esboço, recupera imagens e localiza objectos semelhantes a esse esboço, na imagem. Os autores começam por utilizar um operador *Gaussiano de Laplace*, para suavizar as áreas envolventes do objecto da imagem e do esboço. Em seguida constroem um histograma de gradiente, para a imagem e para o esboço, através da utilização de uma janela de n x n que percorre as imagens. A construção do histograma é realizada de acordo com a orientação dos pixéis de cada imagem. Seguidamente é calculada a distância entre os histogramas, usando a norma L^2 para encontrar as intersecções existentes entre as imagens. A localização do objecto na imagem é realizada através da técnica *RANSAC* [8]. Identificado o objecto na imagem, é criado um contorno à sua volta, usando o molde do objecto esboçado.

A utilização desta técnica demonstra que funciona para imagens a cores, localizando o objecto esboçado nas imagens. (ver figura 2.6). Pensamos que a técnica utilizada pode contribuir para obter determinados objectos que estejam nas imagens.

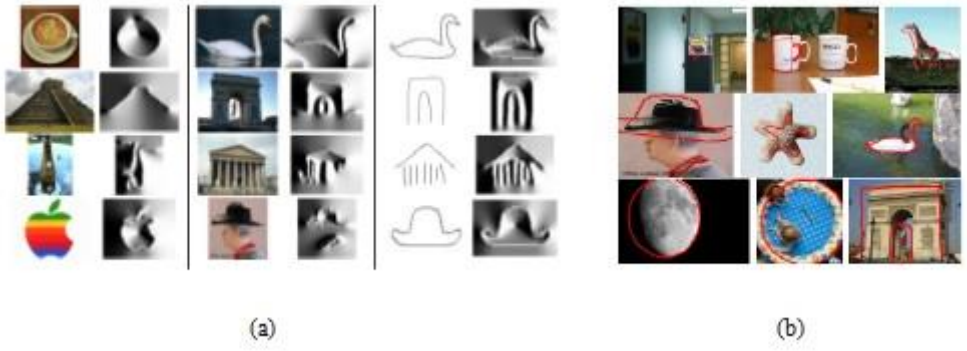


Figura 2.6. (a) Imagens e esboços simples. Aplicação do *operador Gaussiano de Laplace*, (b) Localização do objecto na Imagem

Em [9], é descrita uma técnica que converte uma imagem a cores numa imagem de contornos a preto e branco. Os autores Abdolah Chalechale, et al convertem a imagem em tons de cinzento para eliminar a saturação, permanecendo a luminosidade na imagem. A extracção dos contornos é feita através da aplicação do *operador de Canny*, com $\sigma = 1$ e uma *máscara Gaussiana* com tamanho 9. Seguidamente é calculado através da *convolução 1 D Gaussiana*, os pixéis de contorno de modo a que apenas os contornos mais relevantes sejam mantidos, de forma a melhorar a aparência destes na imagem. A utilização de um parâmetro β é preponderante para a percepção dos contornos. Quanto maior for o valor de β, menor é o número de contornos e mais perceptíveis estes são na imagem (ver figura 2.7). Obtida a imagem em contornos, é criado um círculo fatiado, analisando-se em cada fatia a direcção dos contornos. Em seguida, é utilizado a *transformada discreta de Fourier 1D*, para obter a distribuição angular espacial dos pixéis na imagem. Através da aplicação desta transformada é feita uma normalização, sendo calculada a similaridade entre a imagem convertida em contornos com a imagem esboçada, através da utilização da *distância de Manhattan*. A aplicação desta técnica, pode contribuir para a nossa solução, uma vez que extraí os contornos presentes na imagem.

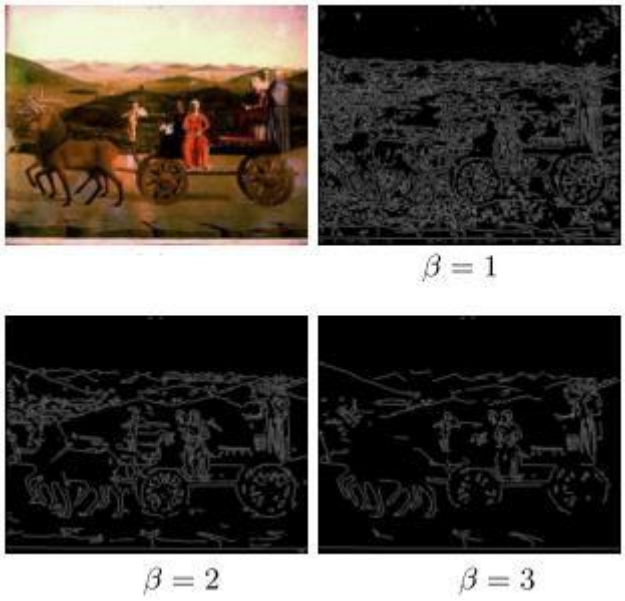


Figura 2.7. Efeito do parâmetro β para a detecção dos contornos

Em [10] é descrita uma técnica que recupera fotos de rostos com base em esboços. Os autores Hamed Galoogahi e Terence Sim utilizam um *Filtro de Gabor* para atenuar texturas grossas, como os olhos, sobrancelhas, nariz, queixo e orelhas, características que formam a cara. Seguidamente, estas texturas são modeladas através da aplicação da *Transformada de Radon*. Ao aplicar estas duas técnicas, eliminam-se as texturas finas, como rugas, sardas, sombras ou reflexos. Dada uma imagem de rosto a preto e branco, é aplicado um *Filtro de Gabor* que calcula as *magnitudes de Gabor*. Divide-se em blocos cada magnitude e constrói-se um histograma que representa um conjunto de amostras de *Radon*. Em seguida, concatenam-se todos os histogramas, originando um único histograma, representativo da forma de *Gabor*. Desta representação é obtido o rosto da imagem em forma de esboço (ver figura 2.8).

Analisando a técnica utilizada, esta é limitada para a nossa solução, pois apenas é adequada na extracção de imagens com faces.

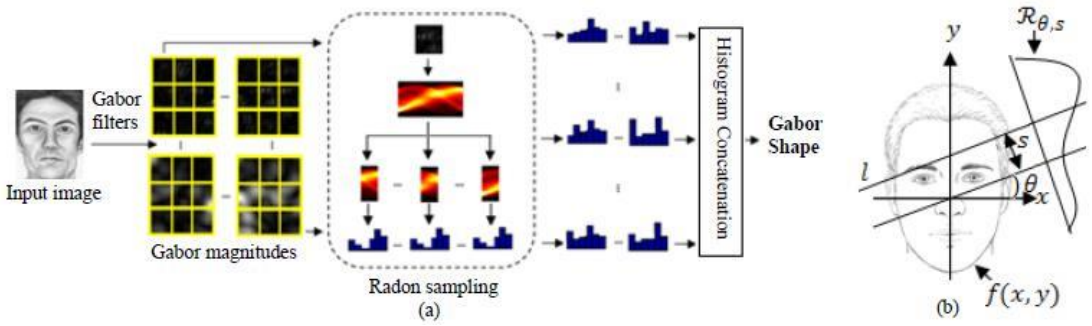


Figura 2.8. (a) Aplicação do *Filtro de Gabor*, (b) *Transformada de Radon* na imagem

2.2 Detecção de contornos em Imagens

Em [11] os autores Silva e Candeias descrevem uma técnica para a identificação de contornos numa imagem em tons de cinzento, com base num histograma. Os autores calculam a média e o desvio padrão, subtraindo estes para obter o limiar entre a região de interesse e a restante região, identificando assim os contornos.

Esta técnica permite-nos a identificação do limiar da região, obtendo assim as coordenadas dos pixéis que formam o limiar. Obtidas as posições podemos definir-lhe um contorno para posteriormente, ser comparada a sua similaridade com o esboço (ver figura 2.9).

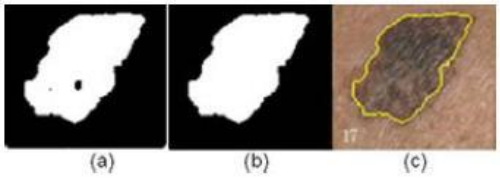


Figura 2.9. Imagens (a) binarizada, (b) com o maior objecto, (c) contorno do objecto

Os autores Mathias Eitz, et al. em [12], descrevem a técnica de extracção de contornos *Edge histogram descriptor*, que foi proposta para o estudo da norma *MPEG 7 standard*. Apesar de parecer útil para esse tipo de estudos, a mesma pode-se aplicar para o nosso caso, na medida em que ajuda a extrair os contornos de uma imagem. Esta abordagem consiste na divisão da imagem em blocos, sendo esta divisão proporcional ao tamanho da imagem. Em cada bloco

são percorridos todos os pixéis através de cinco máscaras diferentes. Estas máscaras são caracterizadas como contornos verticais, horizontais, direcção de 45 graus, 135 graus e contornos sem direcção. Cada máscara tem um tamanho de dois por dois, para posteriormente ser calculado o tipo de contorno através do uso de uma distância. Considerado o maior valor proveniente do cálculo da distância, toma-se uma decisão quanto à direcção do contorno do pixel na imagem. Posto isto, é construído para cada bloco da imagem um histograma que contém o número de direcções de cada contorno. Ao analisar-se o histograma, verifica-se qual o traçado com maior valor no histograma, sendo esse traçado, considerado o contorno mais acentuado no bloco. Esta técnica aplicada a uma imagem e a um esboço, resulta em histogramas de contornos de cada bloco. A sua utilização permite perceber a direcção dos contornos de cada bloco (ver figura 2.10). Aplicando a diferença de histogramas, conseguimos aferir a similaridade entre ambas as imagens.

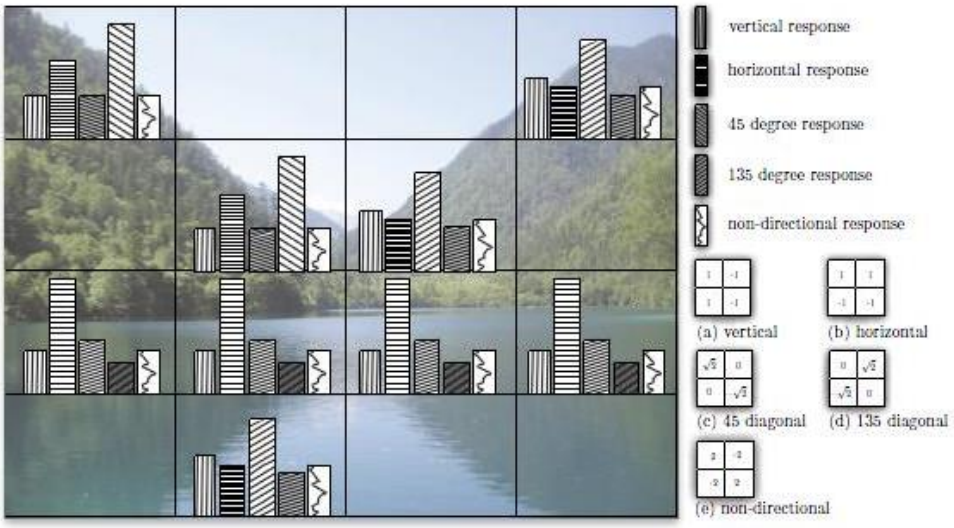


Figura 2.10. Representação do histograma que contêm o número de contornos existentes em cada bloco da imagem. À direita as cinco máscaras representativas da direcção dos contornos

O autor Hélder Matos utiliza o algoritmo *Moore Neighbor* para a extracção de contornos [13]. Este algoritmo consiste na visita de todos os pixéis que formam a imagem, tendo como objectivo, encontrar os pixéis pretos que posteriormente são usados para a identificação de contornos (ver figura 2.11).

Este algoritmo é eficiente em imagens que tenham um único objecto, ou que o mesmo não possua buracos. Caso contrário, este algoritmo não apresenta os resultados esperados.

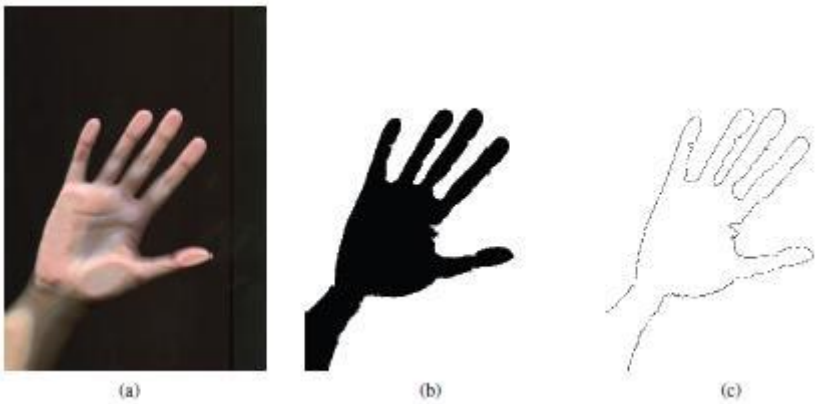


Figura 2.11. (a) Imagem Original em RGB, (b) Imagem Binarizada (c) Aplicação do Algoritmo *Moore Neighbor*

Rodrigo Zanin, Aluir Dal Poz e Érico Martins, descrevem em [14] as etapas para a concretização eficiente da extracção de contornos de uma imagem. A imagem inicialmente é submetida a um pré-processamento, utilizando um *filtro Gaussiano* para atenuar ruídos existentes na imagem, e faz-se uma separação das regiões de interesse do fundo, através de um *threshold*. Este limiar, *threshold*, é calculado a partir do valor mínimo do histograma da imagem em níveis de cinzento (ver figura 2.12). Binarizada a imagem, são calculados os seus contornos pela diferença da imagem binária de uma região, através da aplicação de um detector de contornos. Obtidos os contornos da imagem, são extraídas todas as posições dos pixéis que compõem o contorno da região. A aplicação desta técnica em relação à extracção de contornos, não contribui para a nossa solução, devido ao facto do problema tratado não ser referido com relevância na solução apresentada. Ainda assim, as técnicas utilizadas para a segmentação da imagem, podem contribuir para a remoção de ruído ou para a separação dos objectos que compõem a imagem.

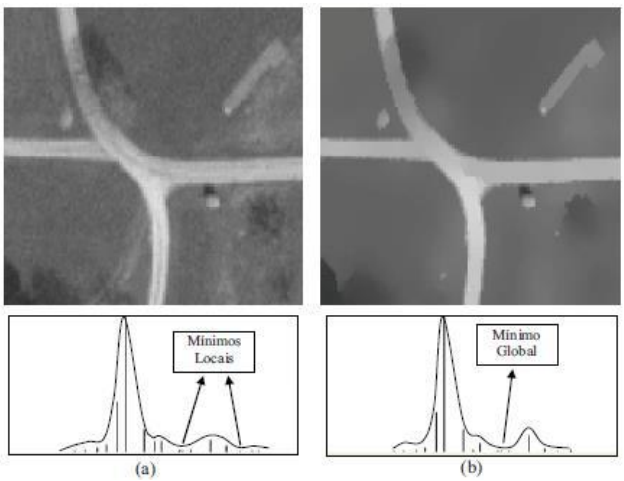


Figura 2.12. (a) Imagem Original a níveis de cinzento com histograma, (b) Imagem suavizada com histograma

2.3 Síntese

Neste capítulo descrevemos os trabalhos estudados, referentes a *recuperação de Imagens através de Esboços* e *Detecção de contornos em Imagens*. Foram analisadas soluções que consistissem em recuperar imagens através de esboços e convertessem a imagem para um formato comum a um esboço, como soluções que nos permitissem detectar contornos nas imagens. Das soluções examinadas, constatámos que existem algoritmos como o *Chain Code* que captura o contorno na imagem e que pode transformar a mesma num esboço. Averiguámos outras técnicas, como a de *RANSAC* que detecta o objecto na imagem e cria um contorno à sua volta, esboçando deste modo o objecto. Técnicas como o *Filtro de Canny* e *Filtro de Gabor* foram abordadas nas soluções, consistindo a primeira na detecção de contornos enquanto a segunda na atenuação de texturas de modo a preparar a imagem para a identificação dos contornos. Por último analisámos uma solução que utiliza o algoritmo *Moore Neighbor* sendo este limitador, na medida em que apenas funciona para imagens que não formem buracos. Analisadas as soluções, verificámos que todas contribuíram para o nosso objectivo, ou seja, tiveram em consideração a recuperação de imagens através de esboços, a partir da detecção de contornos.

3 Arquitectura da solução

Neste capítulo descrevemos toda a nossa solução, explicando todo o processo a que submetemos a imagem e o esboço. Em relação à imagem, esta foi submetida num formato *raster*, à qual realizámos um pré-processamento, consistindo esta etapa na remoção de ruído como no realçamento dos objectos mais relevantes na imagem. Efectuado o pré-processamento na imagem, procedemos a uma simplificação para eliminar pequenas imperfeições que tenham permanecido após o processo anteriormente realizado, obtendo deste modo, uma imagem convertível para um formato *vectorial*. Simplificada a imagem, convertemos a mesma para um formato comum, que nos proporcionou posteriormente a extracção do descritor espacial e geométrico dos diversos elementos que formam a imagem. Quanto ao esboço, os processos efectuados foram idênticos aos da imagem, tendo este sofrido um pré-processamento, incidindo esta etapa na remoção de imperfeições presentes no desenho. Seguidamente realizámos uma conversão do esboço para um formato comum. Efectuada a conversão realizámos uma extracção dos descritores referentes à topologia e geometria dos objectos da imagem.

Concluídos os processos de extracção de características nas imagens e no esboço, procedemos a uma comparação dos descritores respeitante a ambos, recorrendo a métricas para posteriormente apresentarmos os resultados.

3.1 Arquitectura global

Nesta secção apresentamos a arquitectura da nossa solução que se encontra dividida em duas partes principais, uma para lidar com as imagens e outra para tratar os esboços (ver figura 3.1). Em relação às imagens, primeiro realizamos um pré-processamento, e uma simplificação de forma a simplificar a imagem e obtermos uma imagem mais próxima de um esboço. A seguir, convertemo-la para um formato que nos permitiu extrair a sua informação e compará-la com a do esboço. No caso do esboço, efectuámos um pré-processamento de modo a torna-lo mais simples. Tal como na imagem, convertemo-lo para um formato que nos possibilitou a extracção das suas características.

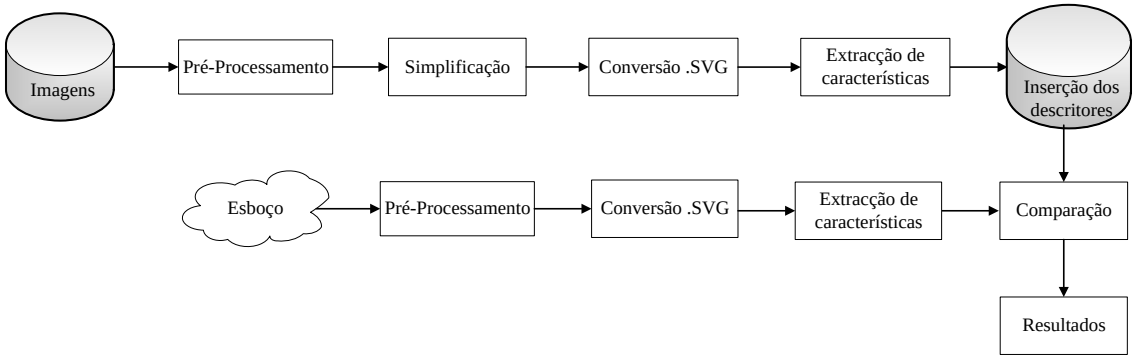


Figura 3.1. Arquitectura da solução

Começámos por realizar na imagem um pré-processamento, composto por dois processos diferentes: *Combinação de Técnicas* e *Filtro de Canny*. Ao aplicarmos o processo *Combinação de Técnicas*, obtivemos um contorno que circunscreve o objecto que compõe a imagem, sendo removido todo o detalhe presente nestes mesmos objectos. A aplicação do processo *Filtro de Canny* concedeu-nos um maior detalhe, realçando os contornos que compõem os diferentes objectos da imagem. Ultrapassados estes processos, efectuámos uma simplificação a ambas as imagens, de forma a remover pequenos objectos isolados que resultem dos processos de filtragem, bem como, possibilitou-nos proceder à conversão da mesma. Em seguida, convertemos a imagem para um formato *vectorial*, formato *Scalable Vector Graphics* (SVG)[1], que se traduziu na conversão da imagem em contornos para um formato que nos facilitasse a extracção da relação espacial e geométrica dos objectos referentes à imagem. Obtida uma imagem *vectorial*, implementámos o algoritmo *GeoDiff* aplicando-o apenas nas imagens em que preservámos toda a informação, provenientes do processo *Filtro de Canny*. Este algoritmo consistiu na construção de um grafo em árvore, resultando numa hierarquia que a cada nó corresponde um polígono da imagem, estabelecendo-se uma relação topológica e uma diferença geométrica entre os nós. Construído o grafo, extraímos o descritor para posterior comparação com a imagem esboçada pelos nossos utilizadores. No caso das imagens em que removemos todo o detalhe e apenas obtivemos o contorno envolvente do objecto, aplicámos o algoritmo de *Geometria*, que nos permitiu extrair o descritor geométrico para compararmos com os esboços dos nossos utilizadores.

[1] http://www.w3.org/Graphics/SVG/

Em relação ao esboço, realizámos um pré-processamento, consistindo esta etapa em duas filtragens distintas. Estando ambas as filtragens relacionadas com a forma como foi efectuado o tratamento à imagem, no qual preservamos o detalhe do esboço através do processo *Filtragem Detalhada* e por outro, o processo *Filtragem Simplificada* que evidenciou apenas o contorno envolvente do objecto do esboço.

Efectuado o tratamento ao esboço, convertemos este para um formato *vectorial*, tal como realizado para a imagem. A realização desta conversão tornou-se útil pois converteu-nos o esboço para um formato comum, conseguindo-se, aplicar técnicas que extraíam a informação dos objectos. Convertido o esboço extraímos a informação, através da utilização dos algoritmos *GeoDiff* e *Geometria*, onde calculamos os descritores referente ao esboço, precedidos respectivamente, da *Filtragem Detalhada* e da *Filtragem Simplificada.*

Obtidos os descritores da imagem e do esboço, procedemos a uma comparação de ambos os resultados recorrendo a métricas matemáticas. Esta comparação pressupôs um paralelismo entre os descritores relativos à topologia e diferença geométrica, obtidos da imagem com os mesmos descritores calculados do esboço, assim como uma comparação entre os descritores referentes apenas à geometria do contorno envolvente do objecto da imagem com os extraídos do esboço. Efectuada a comparação dos respectivos descritores das imagens e do esboço, apresentamos os resultados.

3.2 Conversão de imagem *raster* para imagem *vectorial*

Nesta secção apresentamos todos os processos utilizados para convertermos uma imagem *raster* em duas imagens *vectoriais* (ver figura 3.2). Considerado o pré-processamento, como uma das etapas mais relevantes no tratamento da imagem, foi através deste que recorremos a dois processos distintos, *Filtro de Canny* e *Combinação de Técnicas*. Em relação ao primeiro, este consistiu no realçar de determinados detalhes na imagem enquanto que o segundo, evidenciou o contorno que envolve o objecto. Obtidas duas imagens em contornos, tornou-se fundamental simplificarmos as imagens de forma a obtermos uma aproximação a um esboço. Para que isso fosse possível, aplicámos três técnicas diferentes às imagens. Começámos por um *threshold* de forma a minimizar o ruído proveniente da aplicação dos filtros, nomeadamente do *Filtro de Canny*. Posteriormente, uma *operação de fecho* [15] foi aplicada

para reduzirmos o número de lacunas que existiam nos contornos das imagens. Por fim, aplicámos a técnica de *esqueletização* [16], definindo-se desta forma, um contorno fino para conseguirmos converter as imagens para um formato *vectorial*.

Realizada a etapa de pré-processamento, em que obtivemos imagens em contornos, e a etapa de simplificação, na qual obtivemos imagens próximas de um esboço, efectuámos uma conversão da imagem *raster* para uma imagem *vectorial*, em formato SVG. Esta conversão foi essencial, transformando as imagens em contornos num formato decifrável, para a obtenção dos descritores de topologia e geometria dos elementos que constituem a imagem.

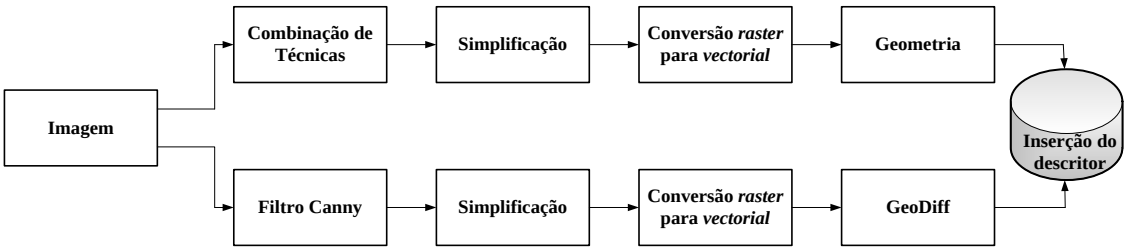


Figura 3.2. Arquitectura da imagem

3.2.1 Pré-Processamento

Considerado o pré-processamento como uma das etapas fundamentais no tratamento da imagem, estudámos e analisámos diferentes combinações de técnicas. Tal facto permitiu-nos tirar as ilações necessárias para a obtenção de uma imagem em contornos. Realizado o estudo (abordado no capítulo 4), apenas considerámos dois processos, *Combinação de Técnicas*, consistindo este na detecção do contorno que delimita o objecto na imagem, e o *Filtro de Canny*, na detecção de contornos que descrevam com mais detalhe os objectos.

3.2.1.1 Combinação de Técnicas

O processo *Combinação de Técnicas* foca-se na aplicação de um conjunto de técnicas (ver figura 3.3), que resultam na obtenção de um contorno global do objecto numa imagem.

Começámos por utilizar um *Filtro Gaussiano com Blur* [17] para remover algum ruído existente na imagem. Seguiu-se uma *Binarização* [18], para eliminar os objectos isolados ou irrelevantes da imagem. Posteriormente, usámos a técnica *Binária*, para converter a imagem para preto e branco, valores zero ou um. Por fim, aplicámos a técnica *Encontra Contornos*, para realçar os contornos presentes na imagem.

Figura 3.3. Etapas aplicadas à imagem através da aplicação *Combinação de Técnicas*

Ao recorrermos à biblioteca *ImageJ*[2], utilizámos o *Filtro Gaussiano com Blur*, consistindo este num filtro de convolução com uma função gaussiana. Através da aplicação deste filtro, obtivemos uma suavização da imagem preservando-se os contornos da mesma como se pode ver nas figuras 3.4 e 3.5. Constatámos que o mesmo, reduziu-nos o ruído presente nas imagens como o detalhe irrelevante que existia nas imagens, obtendo-se uma imagem desfocada mas sem que nela estivesse presente qualquer distorção.

[2] http://imagej.nih.gov/ij/developer/api/

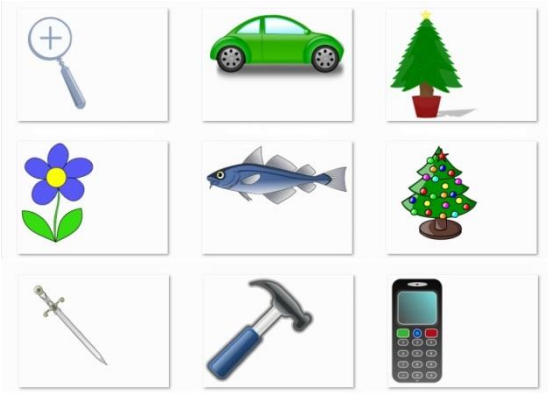


Figura 3.4. Imagens a processar (originais)

Figura 3.5. Imagens obtidas do *Filtro Gaussiano com Blur*

Seguidamente utilizámos uma *binarização*, sendo esta uma técnica intermédia entre o *Filtro Gaussiano com Blur* e o processo *binário*. Tal técnica consistiu na aplicação de um *threshold* na imagem para eliminarmos os objectos isolados ou os objectos que fossem irrelevantes na composição de um único objecto. Pretendendo obter uma imagem a preto e branco, usámos como limiar na nossa operação o valor 255, consistindo esta operação no cálculo da média da intensidade de cada pixel. Caso o valor da intensidade do pixel calculado, estivesse abaixo do limiar, convertemo-lo para valor 0, caso contrário toma o valor 255. Deste modo, obtivemos uma imagem a preto e branco, tomando todos os pixéis que formam o objecto da imagem a cor preta (ver figura 3.6).

Finalizada a técnica acima descrita, e através da biblioteca *ImageJ*, aplicámos um processo *Binário*, convertendo a imagem para o valor zero e um (preto e branco), tendo por base o histograma de toda a imagem. Os objectos que fossem relevantes na imagem foram convertidos para o valor zero, sendo os restantes substituídos pelo valor um. Tal processo foi efectuado devido aos valores dos pixéis, quando aplicada a técnica de *binarização*, serem constituídos por R=0, B=0 e G= 0 ou R= 255, B= 255 e G = 255.

A utilização da técnica *Binária* permitiu-nos a conversão dos píxeis para um único valor 0 ou 1, correspondendo este valor à camada RGB. Deste modo, obtivemos uma imagem binária e da qual foi possível, posteriormente, detectarmos os contornos mais relevantes na imagem, atenuando-se todos os restantes objectos irrelevantes que pudessem sobressair na imagem.

Convertida a imagem a preto e branco, pretendemos agora obter uma imagem em contornos, um formato próximo a um esboço. Para isso, criámos a função *Encontra Contornos* recorrendo à biblioteca *ImageJ*. Esta função é constituída por um detector de contornos,

Operador de Sobel [19], que evidencia a direcção dos contornos verticalmente e horizontalmente da imagem detectando-se o contorno global da imagem como ilustra a figura 3.7.

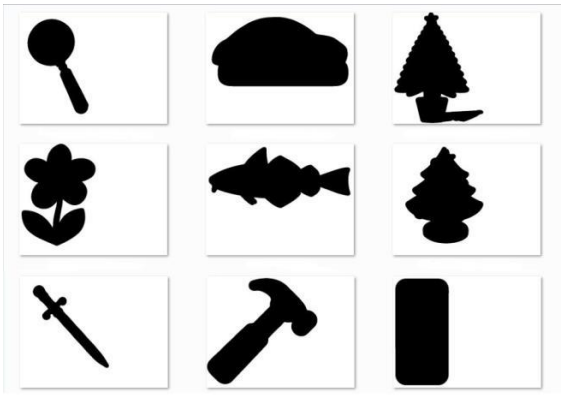


Figura 3.6. Imagens obtidas da *Binarização*

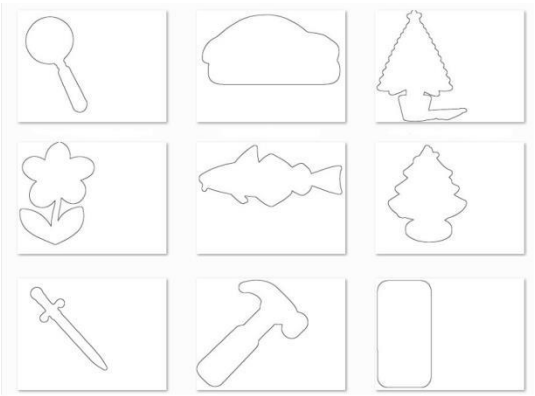


Figura 3.7. Imagens obtidas do *Encontra Contorno*

3.2.1.2 Filtro de Canny

O processo *Filtro de Canny*, consistiu na utilização de duas técnicas em conjunto (ver figura 3.8), no qual aplicámos um *Filtro de Canny* [20] directamente na imagem, sem que esta tivesse um tratamento de suavização ou remoção de ruído. Assim, foi possível detectarmos todos os contornos presentes na imagem, preservando-se todo o detalhe que constitui a imagem *raster*. Em seguida, procedemos a uma subtracção da imagem filtrada com uma imagem a branco. Tal operação permitiu realçar os contornos de uma imagem, os quais tomaram uma cor preta e o fundo branco. Desta forma, obtivemos uma filtragem, mas sem que existisse uma modificação na imagem.

Figura 3.8. Etapas aplicadas à imagem através da aplicação do *Filtro de Canny*

A utilização do *Filtro de Canny* baseou-se na detecção dos contornos da imagem, sendo preservado um maior detalhe, assim como o ruído presente na imagem. A aplicação deste filtro permitiu obtermos um maior número de objectos inerentes à imagem (ver figura 3.9 e 3.10).

Realizado este processo, utilizámos a técnica *subtracção* que consistiu numa subtracção do fundo da imagem, invertendo-se a cor dos pixéis da imagem, devido ao facto do *Filtro de Canny*, gerar um fundo a preto e os contornos a branco. De forma a facilitar-nos o processo de simplificação da imagem, invertemos a cor dos pixéis, ficando os pixéis de contorno a preto e a restante imagem (fundo) a branco (ver figura 3.11).

Figura 3.9. Imagens a processar (originais)

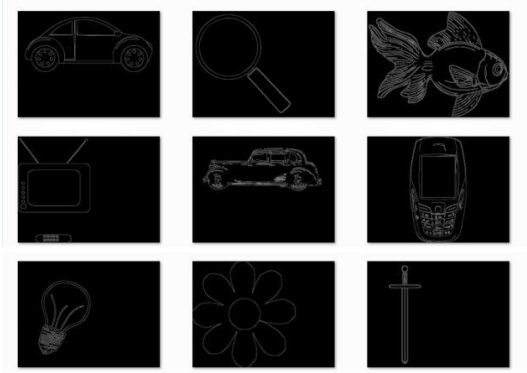


Figura 3.10. Imagens obtidas do *Filtro de Canny*

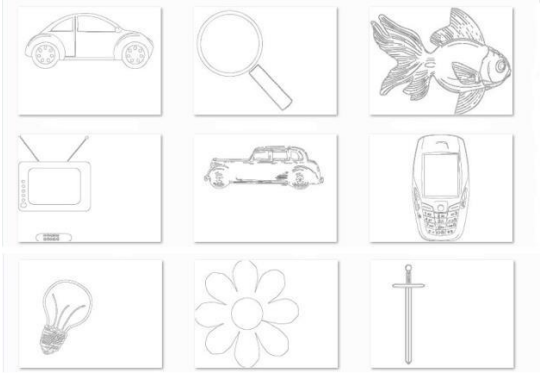


Figura 3.11. Imagens obtidas da *Subtracção*

3.2.2 Processo de Simplificação

O processo de simplificação ocorre após aplicarmos os filtros à imagem. Esta etapa consiste em reduzirmos algum ruído e pequenos buracos que ainda existem na imagem, mas também seleccionarmos os objectos fulcrais que estão presentes na mesma. A realização desta etapa prossupôs a aplicação de três operações diferentes, como ilustra a figura 3.12. Primeiro aplicámos um *threshold*, para minimizar o ruído da imagem, seguidamente aplicámos um *operador de fecho* para eliminar pequenos buracos que ainda existam na imagem, terminando o processo de simplificação com a técnica *esqueletização*, que foi necessária para a obtenção dos objectos com um contorno fino e fundamental para efectuar a passagem da imagem *raster* em *vectorial*.

Figura 3.12. Etapas aplicadas à imagem através do Processo de Simplificação

A técnica de *threshold* consistiu na utilização de um limiar muito próximo do valor máximo que um pixel pode tomar, de forma a eliminarmos todo o ruído que ainda existisse na imagem. Deste modo, todo o objecto que se encontrava acima do limiar foi considerado irrelevante, tomando a mesma cor que o fundo da imagem, ao invés dos pixéis que estavam abaixo do limiar, os quais se tornaram preponderantes para a descrição da imagem.

Após efectuado um *threshold* à imagem, realizámos uma *operação de fecho*, onde aplicámos uma erosão na imagem, após a imagem ter sofrido uma dilatação. Esta operação permitiu-nos uma suavização dos contornos, mas também a remoção de pequenos buracos ou falhas que existissem ao longo dos contornos, preenchendo-os e unindo-os, formando um único contorno. Esta técnica como o *threshold*, obtiveram resultados melhores quando aplicadas às imagens provenientes do *Filtro de Canny* (ver figura 3.13 e 3.14), devido às imagens possuírem um maior detalhe, enquanto na *Combinação de Técnicas* estes resultados não foram visíveis uma vez que as imagens contêm apenas um único contorno (ver figuras 3.15 e 3.16).

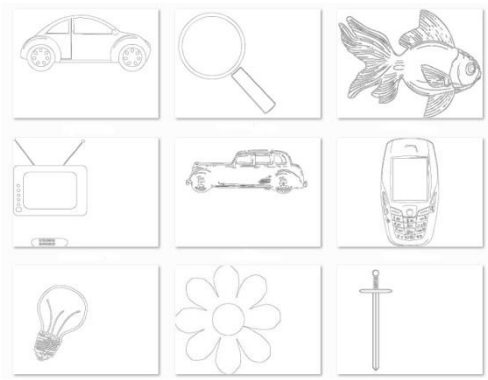


Figura 3.13. Imagens provenientes da aplicação do *Filtro de Canny* após efectuado um *Threshold*

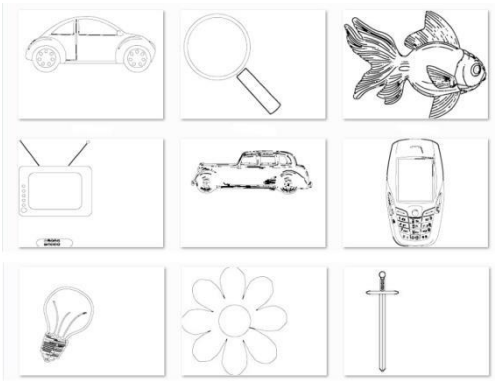


Figura 3.14. Imagens provenientes da aplicação do *Filtro de Canny* após efectuada uma *Operação de Fecho*

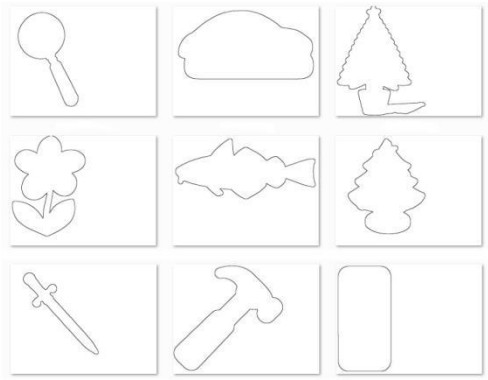


Figura 3.15. Imagens provenientes da aplicação da *Combinação de Técnicas* após efectuado um *Threshold*

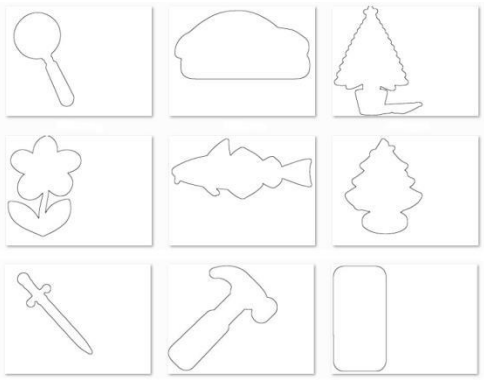


Figura 3.16. Imagens provenientes da aplicação *Combinação de Técnicas* após efectuada uma *Operação de Fecho*

Os processos efectuados anteriormente tornaram-se fundamentais para aplicarmos a *esqueletização*, tornando-se esta preponderante para a concretização da conversão da imagem num formato *raster* para um formato *vectorial*, tendo a imagem antes da conversão possuir um contorno fino. Esta técnica consistiu na transformação dos contornos numa linha com o tamanho de um pixel. Este processo preservou-nos toda a topologia do objecto, devido à forma como foi efectuado, passando a linha pelo centro do contorno (ver figura 3.17 e 3.18).

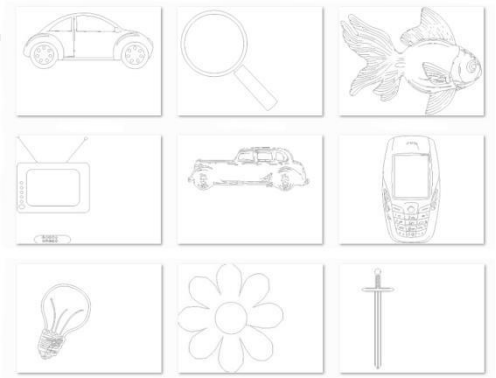


Figura 3.17. Imagens provenientes da aplicação do *Filtro de Canny* após efectuada uma *Esqueletização*

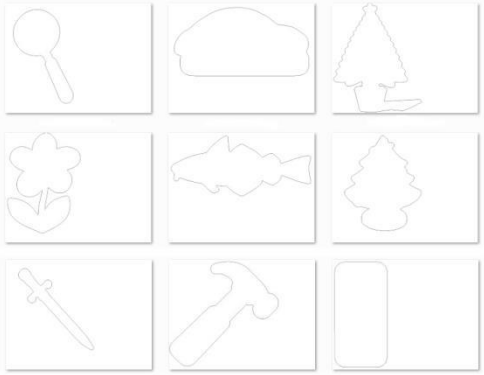


Figura 3.18. Imagens provenientes da aplicação *Combinação de Técnicas* após efectuada uma *Esqueletização*

Desta forma, e após aplicarmos esta técnica, convertemos a imagem *raster* para uma imagem *vectorial*, recorrendo à biblioteca *Batik*[3] que permitiu recriar novamente a imagem mas num formato *vectorial*. Esta conversão proporcionou uma imagem legível para a extracção dos descritores de topologia e geometria dos objectos que formam a imagem.

3.3 Conversão de esboço *raster* para esboço *vectorial*

Nesta secção expomos todos os processos realizados ao esboço, desde a forma como foi efectuado, o pré-processamento, à conversão deste para um formato *vectorial* (ver figura 3.19). O esboço foi efectuado num formato *raster*, com um sistema de cores aditivas, apesar dos traços efectuados serem a preto e branco. Devido aos factos descritos, procedemos a duas etapas, consistindo uma num pré-processamento do esboço e a outra na conversão deste para um formato *vectorial*.

[3] https://xmlgraphics.apache.org/batik/javadoc/

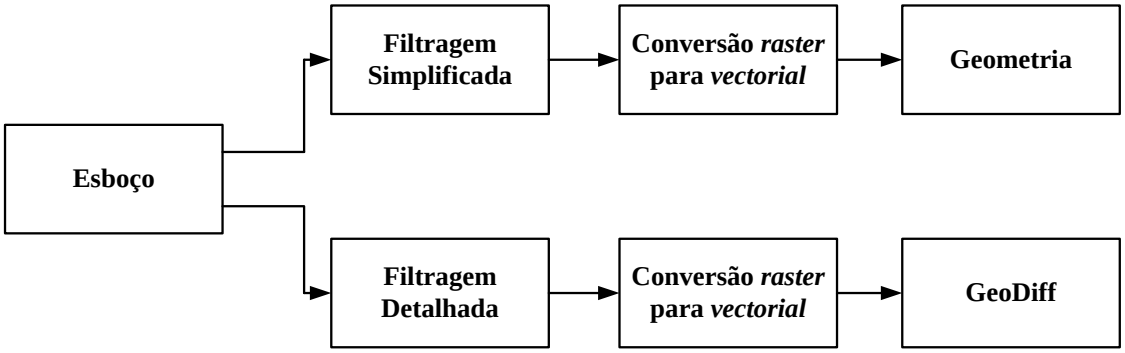


Figura 3.19. Arquitectura do esboço

O pré-processamento efectuado ao esboço divide-se em dois processos distintos, *Filtragem detalhada* e *Filtragem simplificada*. Descrevendo o primeiro processo, este teve como objectivo preservarmos todo o detalhe presente no esboço, removendo-se apenas as imperfeições que constam num desenho em forma de esboço. Em relação ao segundo processo, tivemos como foco a simplificação do esboço, ou seja, eliminámos a informação que descreve o objecto, preservando apenas o contorno envolvente do objecto. Por fim, convertemos os esboços filtrados para um formato *vectorial*, à semelhança da imagem, possibilitando posteriormente a extracção dos descritores.

3.3.1 Pré-Processamento

Nesta secção abordamos todo o processo utilizado no pré-processamento do esboço. Este processamento tornou-se similar ao efectuado na imagem, existindo algumas etapas diferentes devido às técnicas aplicadas nas imagens, quando utilizadas nos esboços, apresentarem modificações na sua topologia, diferindo da topologia inicial. Dado este facto, optámos por aplicar diferentes técnicas ao esboço mas obtendo a mesma finalidade que nas imagens. Como realizado na imagem e para existir uma coerência e conformidade no tratamento da imagem e do esboço, utilizámos dois tipos de filtragem, *Filtragem detalhada*, na qual preservámos o detalhe e a informação do esboço e *Filtragem simplificada*, em que descartámos a informação e preocupámo-nos apenas com o contorno global do esboço.

3.3.1.1 Filtragem detalhada

A *Filtragem detalhada* baseou-se num conjunto de técnicas (ver figura 3.20), que nos permitiu, ao longo da aplicação das mesmas, manter a informação presente no esboço. Inicialmente, aplicámos um *Filtro Gaussiano com Blur*, para suavizar os contornos do esboço. Seguidamente aplicámos a técnica *Binária*, convertendo o esboço para um nível de 8 bits de cores, ou seja, mantivemos o desenho a preto e branco sem o modificarmos. Esta conversão possibilitou-nos utilizar a técnica *Esqueletização*, obtendo-se um contorno fino que nos permitiu a conversão do esboço para um formato *vectorial*.

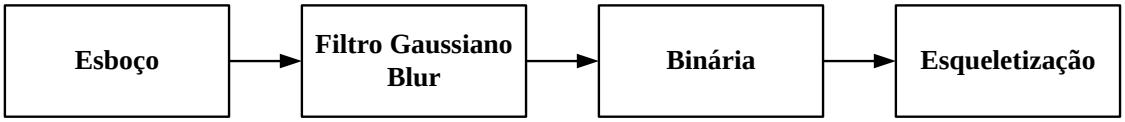


Figura 3.20. Etapas aplicadas ao esboço através da *Filtragem detalhada*

A utilização do *Filtro Gaussiano com Blur*, proveniente do uso da biblioteca *ImageJ*, suavizou-nos os contornos (ver figuras 3.21 e 3.22), eliminando-se pequenas imperfeições quando aplicada a técnica *Binária*. O recurso a ambas as técnicas suprimiu-nos aqueles traços que de uma forma ou de outra tornam o contorno defeituoso. Por outro lado, a técnica *Binária*, converteu-nos o esboço para um nível de 8 bits. Apesar de este ser desenhado a preto e branco, o esquema de cores é *true color*, nível de 24 bits. Deste modo, convertemos os níveis de cor para 8 bits mantendo o esboço a preto e branco sem que efectuasse-mos alguma modificação à sua forma geométrica (ver figura 3.23). Ao convertermos o esboço para um menor número de bits, conseguimos aplicar a técnica *Esqueletização*, também usada através da biblioteca *ImageJ*, tornando o contorno mais fino para convertermos o esboço para o formato *vectorial* (ver figura 3.24).

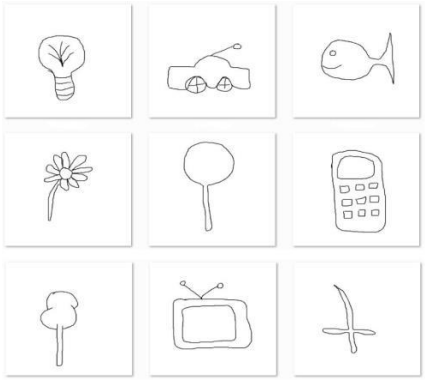


Figura 3.21. Esboços utilizados como interrogação

Figura 3.22. Esboços obtidos do *Filtro Gaussiano com Blur*

Figura 3.23. Esboços obtidos da *Binária*

Figura 3.24. Esboços obtidos da *Esqueletização*

3.3.1.2 Filtragem Simplificada

A *Filtragem Simplificada* (ver figura 3.25) consistiu em obtermos do esboço um desenho simples, ou seja, ao contrário da filtragem anteriormente descrita. Nesta, optámos por remover toda a informação presente no esboço, permanecendo o contorno exterior do objecto desenhado. Assim sendo, aplicámos um *Filtro Gaussiano com Blur*, com o propósito de suavizar os contornos do esboço. Seguidamente, aplicámos a técnica *Binária*, para obtermos o esboço num espaço de cores de 8 bits. Obtido o esboço a níveis de preto e branco, efectuámos um preenchimento dos buracos que se formaram, através da utilização da técnica *Fecha buracos*. Tendo o esboço, apenas dois tipos de cores, preto e branco, e estando os objectos todos a preto, aplicámos a técnica *Encontra contornos*, para extrairmos o contorno exterior do

objecto. Por fim, e tal como aplicado na filtragem anterior, usámos uma *Esqueletização*, para posteriormente pudermos converter o esboço para um formato *vectorial*.

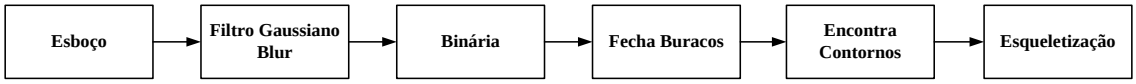


Figura 3.25. Etapas aplicadas ao esboço através da *Filtragem Simplificada*

Nesta filtragem, conseguimos simplificar mais o esboço de modo a aproximá-lo das imagens filtradas. Como tal, decidimos aplicar um *Filtro Gaussiano com Blur*, para suavizar os contornos e remover o ruído presente no esboço tal como ilustrado na figura 3.26 e 3.27. Ao usarmos a técnica *Binária*, conseguimos ocultar as imperfeições que surgiram quando se realizou um esboço. Por outro lado, esta técnica permitiu-nos unir os contornos que se encontrem próximos de si, completando o desenho e tornando-o mais realista (ver figura 3.28). A técnica *Fecha Buracos* eliminou-nos os buracos formados ou mesmo desenhados no esboço. Ao cobrirmos os buracos, obtemos uma única forma geométrica, preenchendo o objecto com a cor preta (ver figura 3.29). Assim sendo, aplicámos a técnica *Encontra Contornos*, que extraiu-nos o contorno que envolve o objecto (ver figura 3.30). Por fim, usámos a técnica *Esqueletização* (ver figura 3.31) que dispôs o esboço com um contorno fino para procedermos à conversão deste para um formato *vectorial*. Efectuadas as técnicas, procedemos à conversão de um esboço num formato *raster* para um formato *vectorial*, proporcionando a extracção das características dos objectos.

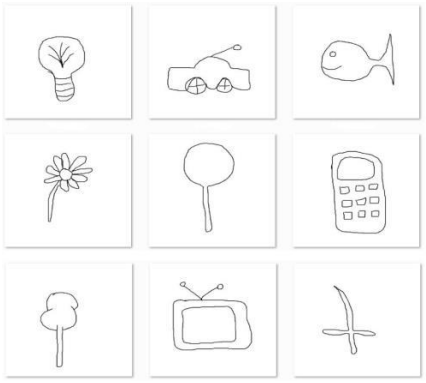


Figura 3.26. Esboços utilizados como interrogação

Figura 3.27. Esboços obtidos do *Filtro Gaussiano com Blur*

Figura 3.28. Esboços obtidos da *Binária*

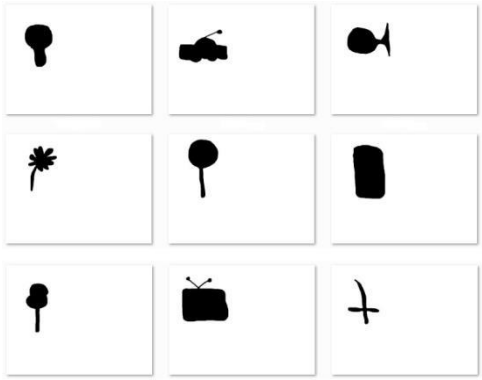


Figura 3.29. Esboços obtidos do *Fecha Buracos*

Figura 3.30. Esboços obtidos do *Encontra Contornos*

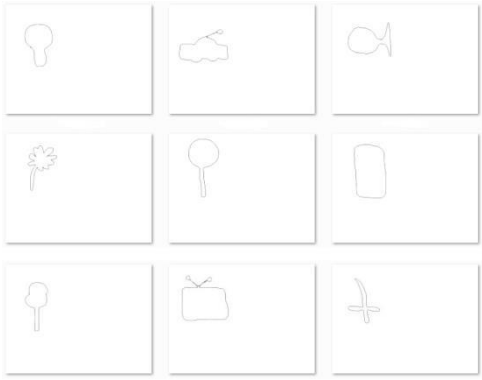


Figura 3.31. Esboços obtidos da *Esqueletização*

3.4 Extracção do Descritor

Nesta secção abordamos todo o processo que envolveu a extracção das características que compõem cada objecto na imagem e no esboço. Pretendendo realizar a extracção das características de acordo com os resultados obtidos das filtragens efectuadas, implementámos o algoritmo *GeoDiff* para as imagens e esboços onde o detalhe foi preservado, extraindo-se a informação topológica e geométrica dos objectos. O algoritmo de *Geometria* extraiu-nos apenas a geometria do contorno global do objecto nas imagens e esboços.

Em relação ao *GeoDiff*, este é uma evolução do trabalho anteriormente desenvolvido por Fonseca [21] onde os autores descrevem separadamente a relação de topologia entre os artefactos e a sua geometria. No caso do nosso algoritmo, este concilia a relação topológica e geométrica, através da construção de um grafo que descreve a organização espacial de todos

os artefactos presentes na imagem, existindo entre estes uma hierarquia. Esta hierarquia é composta por duas relações, inclusão e adjacência. A primeira é aplicada a artefactos que estejam contidos dentro de outros, e a segunda aos artefactos que estejam contidos num mesmo artefacto mas que sejam vizinhos entre si. Adicionalmente e como citado anteriormente conciliámos a informação geométrica dos artefactos, algo que não acontece no trabalho realizado anteriormente. Obtido um grafo de topologia e geometria, calculámos os descritores que constituem os artefactos a partir da matriz de adjacência do grafo. Por fim, e após termos todos os descritores calculados, obtemos o descritor que forma a imagem. Por seu lado, a *Geometria*, consistiu na obtenção de um descritor geométrico proveniente do cálculo geométrico do contorno global do objecto na imagem usando a biblioteca *Cali* [22].

3.4.1 GeoDiff

A geração do nosso grafo consistiu na relação existente entre os diferentes polígonos da imagem, obtidos pelos processos *Filtro de Canny* e *Filtragem detalhada*. O grafo é composto por um nó raiz, correspondendo este nó à imagem como um todo e por vários nós, representando cada um dos nós diferentes polígonos que formam a imagem. Na formação do grafo existe uma relação entre os nós, podendo esta ligação ser considerada de inclusão ou adjacência. Consideramos um nó de inclusão quando um dado polígono se encontra dentro de outro polígono, ou seja, todos os seus vértices têm que estar contidos no polígono que o rodeia. Os nós de adjacência são aqueles cujos vértices de dois polígonos são vizinhos entre si (ver figura 3.32). Quando um polígono é incluso de outro, consideramos que existe uma relação pai-filho. O polígono contido é filho do polígono que o contém, considerando este como o pai. Perante este caso, acrescentamos ao grafo um novo nível de profundidade, resultando deste modo, uma relação de adjacência entre os nós que estão ao mesmo nível de profundidade (ver figura 3.33).

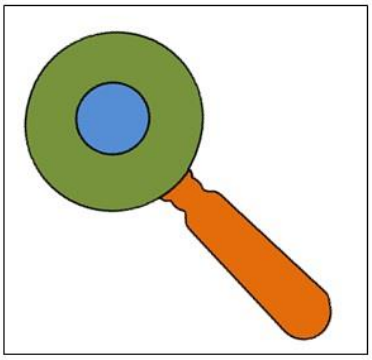


Figura 3.32. Ilustração de uma imagem composta por objectos inclusos e adjacentes

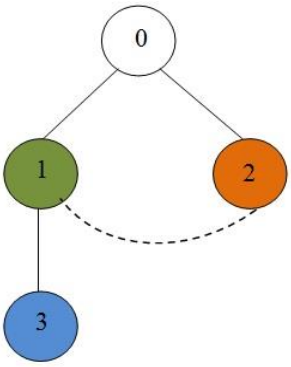


Figura 3.33. Formação do grafo a partir de uma imagem

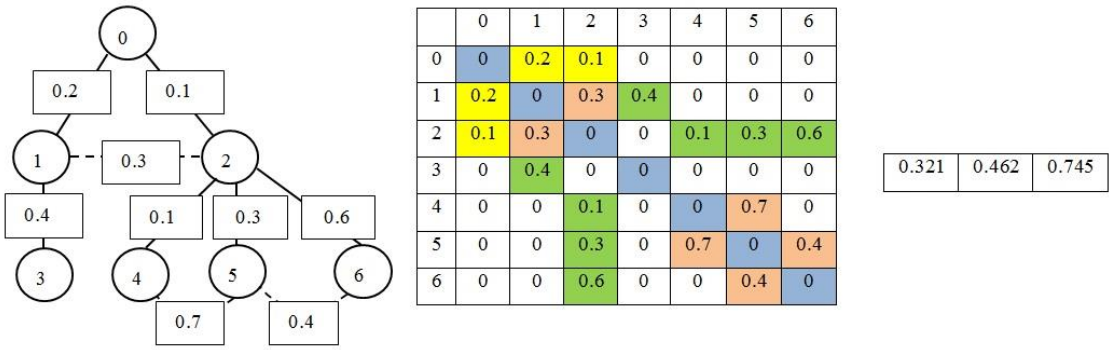
Construído o grafo e existindo uma relação entre os diferentes nós que compõem a imagem, calculamos a geometria de cada nó, recorrendo à biblioteca *Cali*. No caso do nó raiz, e sendo este o nó inicial do grafo consideramos que o seu descritor de geometria tem todas as suas componentes com o valor máximo 1. Assumimos este valor, uma vez que consideramos este, como o maior valor que qualquer nó pode obter do cálculo geométrico. Obtida a geometria dos nós, atribuímos pesos às ligações entre os nós que formam o grafo. Estes pesos são calculados de duas formas diferentes consoante o nó seja incluso ou adjacente.

Quando o nó é incluso, o cálculo efectua-se através da diferença da geometria normalizada do nó pai e do nó filho.

Em relação ao cálculo entre os nós adjacentes, este é dado pelo quociente entre a diferença da diagonal maior do nó pai e a menor distância calculada entre a geometria obtida dos nós adjacentes em causa, com a diagonal maior do nó pai.

Havendo uma relação de adjacência e inclusão e diferentes pesos geométricos nas ligações dos nós, construímos a matriz de adjacência do grafo, correspondendo o número de linhas e colunas ao número de nós adjacentes e inclusos, respectivamente. A cada posição da matriz, corresponde o valor dos pesos atribuídos às ligações.

A partir da matriz de adjacência, extraímos os descritores. Esta extracção resultou num vector, proveniente do cálculo dos valores próprios da matriz (ver figura 3.34). Posto isto, efectuámos o cálculo de similaridade entre a imagem e o esboço.



	0	1	2	3	4	5	6
0	0	0.2	0.1	0	0	0	0
1	0.2	0	0.3	0.4	0	0	0
2	0.1	0.3	0	0	0.1	0.3	0.6
3	0	0.4	0	0	0	0	0
4	0	0	0.1	0	0	0.7	0
5	0	0	0.3	0	0.7	0	0.4
6	0	0	0.6	0	0	0.4	0

0.321	0.462	0.745

Figura 3.34. Processo utilizado para calcular o descritor da imagem e do esboço

3.4.2 Geometria

Obtidas as imagens e esboços com um contorno global, aplicámos o algoritmo de *Geometria,* com recurso à biblioteca *Cali*. A utilização deste algoritmo extrai-nos o descritor geométrico do contorno global do artefacto. Esta extracção é similar a um *Convex Hull* [23], onde varremos toda a imagem e obtemos os pontos que formam o contorno envolvente do artefacto (ver figura 3.35). Obtidos os pontos, calculámos a geometria do contorno do objecto, resultando no descritor geométrico da imagem e do esboço.

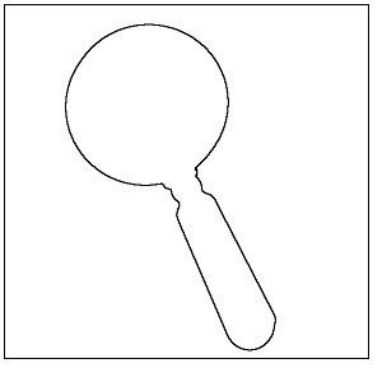


Figura 3.35. Imagem composta apenas pelo contorno global do objecto

3.4.3 Recuperação das imagens

De modo a efectuarmos uma comparação entre uma imagem e um esboço, tivemos que recorrer à *NB-Tree* [24] uma estrutura de indexação multidimensional. Esta estrutura proporciona-nos uma maior rapidez no armazenamento das imagens, utilizando técnicas de indexação que consentem o armazenamento de uma quantidade considerável de imagens para posteriormente efectuarmos a avaliação dos resultados. Esta avaliação consiste no recurso de métricas matemáticas para calcular a similaridade entre as imagens e os esboços. Este cálculo foi realizado a partir dos descritores de topologia e geometria obtidos da etapa de extracção dos descritores, efectuando-se uma comparação dos valores obtidos do esboço com os das imagens contidas na base de dados.

3.4.3.1 Inserção do descritor

O cálculo dos descritores das imagens e dos esboços foram obtidos de duas formas distintas, como descrito anteriormente. Obtivemos o descritor de topologia a partir de um grafo que representa a relação espacial entre os objectos como a diferença geométrica dos mesmos, enquanto a obtenção do descritor de geometria foi calculado através da geometria do contorno global do objecto.

A inserção dos descritores acontece após calcularmos os mesmos nas imagens. Obtidos os descritores de topologia, a partir de um grafo representativo da relação espacial e geométrica dos objectos, e o de geometria calculado através da obtenção geométrica do contorno global do objecto, foi-nos imprescindível armazenar essa informação numa base de dados. Para o efeito, usámos uma *NB-Tree*, estrutura de indexação considerada por Basgalupp [25] uma óptima ferramenta de armazenamento e pesquisa de informação. Este armazenamento foi realizado após efectuados todos os processos descritos em cima, no qual guardámos a directoria referente à imagem, o seu *id* e os valores dos descritores. Estes três atributos tornaram-se fundamentais, uma vez que o *id* identifica inequivocamente a imagem a que pertencem os valores dos descritores calculados, como o nome da directoria, permitindo a localização da imagem para posteriormente apresentarmos os resultados.

3.4.3.2 Comparação

Carregada a base de dados contendo a colecção de imagens, realizámos uma correspondência entre todas as imagens e o esboço efectuado pelos nossos utilizadores. Submetido o esboço a consulta, e efectuada a etapa de pré-processamento, calculámos os descritores de topologia e geometria. Os descritores são calculados de forma análoga aos da imagem, onde usámos um grafo de topologia para obtermos o descritor de topologia e calculámos o descritor de geometria através da geometria do contorno global do desenho. Efectuados os cálculos dos descritores de topologia e geometria do esboço, procedemos à comparação destes com os descritores guardados na base de dados. A comparação é realizada com recurso a métricas matemáticas, emparelhando os descritores provenientes das imagens com os do esboço. Este emparelhamento consiste num cálculo de similaridade entre os descritores de topologia das imagens e do esboço, e no cálculo de similaridade entre os descritores de geometria das imagens com os do esboço. Obtidos os resultados provenientes do cálculo da similaridade entre os dois descritores, realizamos uma combinação dos dois conjuntos de resultados. Desta combinação, resulta uma ordenação dos melhores valores calculados pelos dois descritores. Obtidos os melhores valores de similaridade entre os dois descritores, realizámos a correspondência entre os valores calculados e as imagens. Seguidamente apresentamos por ordem de semelhança, o número de imagens que o utilizador pretende recuperar quando realiza uma interrogação através de um esboço.

3.5 Interface RISCOS

Nesta secção explicamos a interface da nossa aplicação. Consistindo a nossa interface numa visualização simples e de fácil interpretação, apenas considerámos três operações para a nossa aplicação. Quando carregada a aplicação, o utilizador é encaminhado para o *Menu* da aplicação. Nesta página existem duas operações pelas quais o utilizador tem a opção de escolher como pretende efectuar a sua pesquisa, se através de uma imagem, *Query By Image*, ou recorrendo a um esboço, *Query By Sketch* (ver figura 3.36).

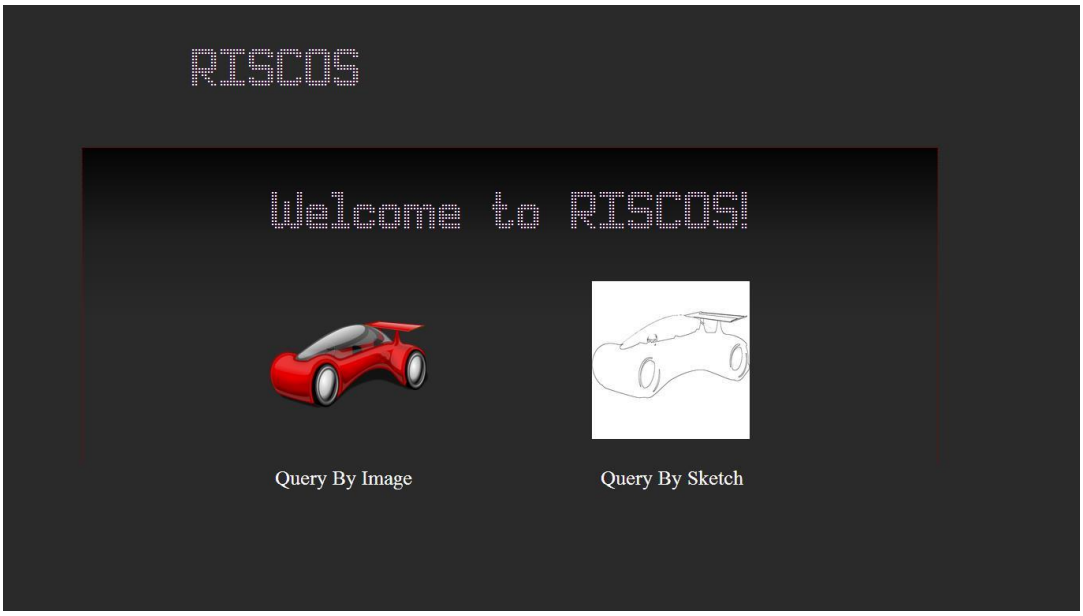


Figura 3.36. Visualização do Menu Inicial

Caso a opção seja, *Query By Image*, o utilizador é direccionado para a página que permite efectuar uma interrogação por imagem. Nesta página, são apresentadas outras opções como o *tipo de filtragem* a efectuar às imagens, o *número de imagens* que pretende recuperar e a *base de dados* que pretende utilizar para a pesquisa das imagens. Estas opções encontram-se previamente preenchidas proporcionando um maior conforto na interacção dos utilizadores com a aplicação, podendo os mesmos alterarem e escolherem as opções que melhor se adequam na realização das suas pesquisas (ver figura 3.37). Efectuada a pesquisa por imagem, o utilizador é encaminhado para uma nova página, onde apresentamos os resultados com o número de imagens indicadas no momento da pesquisa (ver figura 3.38).

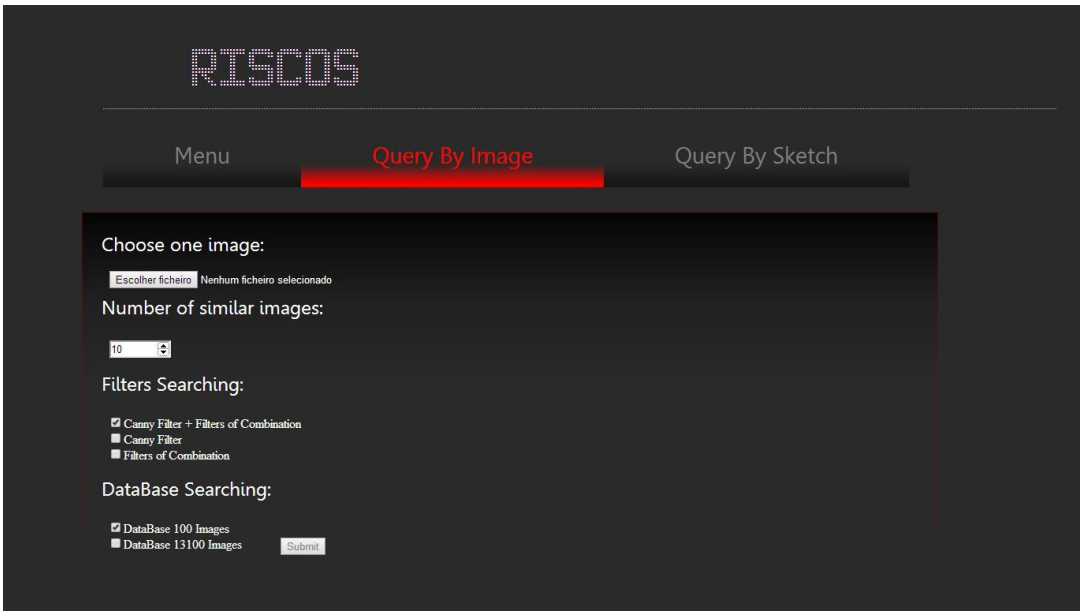


Figura 3.37. Visualização do Menu *Query By Image*

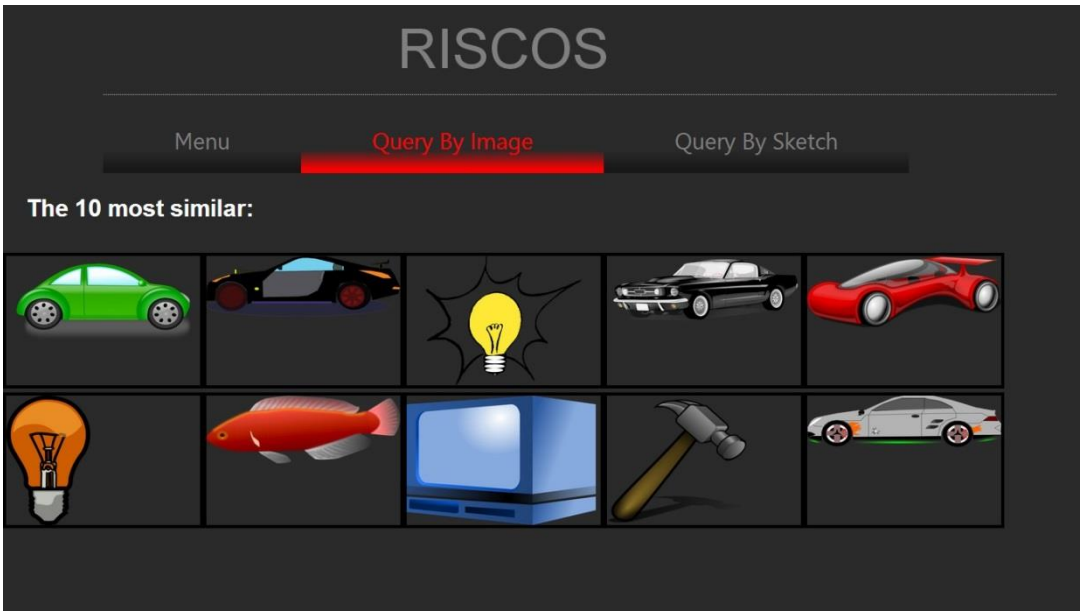


Figura 3.38. Visualização dos Resultados após efectuada uma *Query By Image*

Caso o utilizador opte por escolher a opção *Query By Sketch*, este é redireccionado para a página que permite a realização de um desenho. Nesta página, o utilizador encontra uma tela específica para efectuar o desenho, como botões auxiliares para possíveis rectificações no seu desenho. Para além destes botões, existem ainda outras opções para complementar a pesquisa (ver figura 3.39), tendo estas a mesma funcionalidade que as descritas na página *Query By*

Image. Apresentadas as páginas da nossa solução, apenas nos falta realçar o acompanhamento permanente dos três botões *Menu*, *Query By Image* e *Query By Sketch*, proporcionando aos utilizadores um maior conforto na utilização da nossa aplicação.

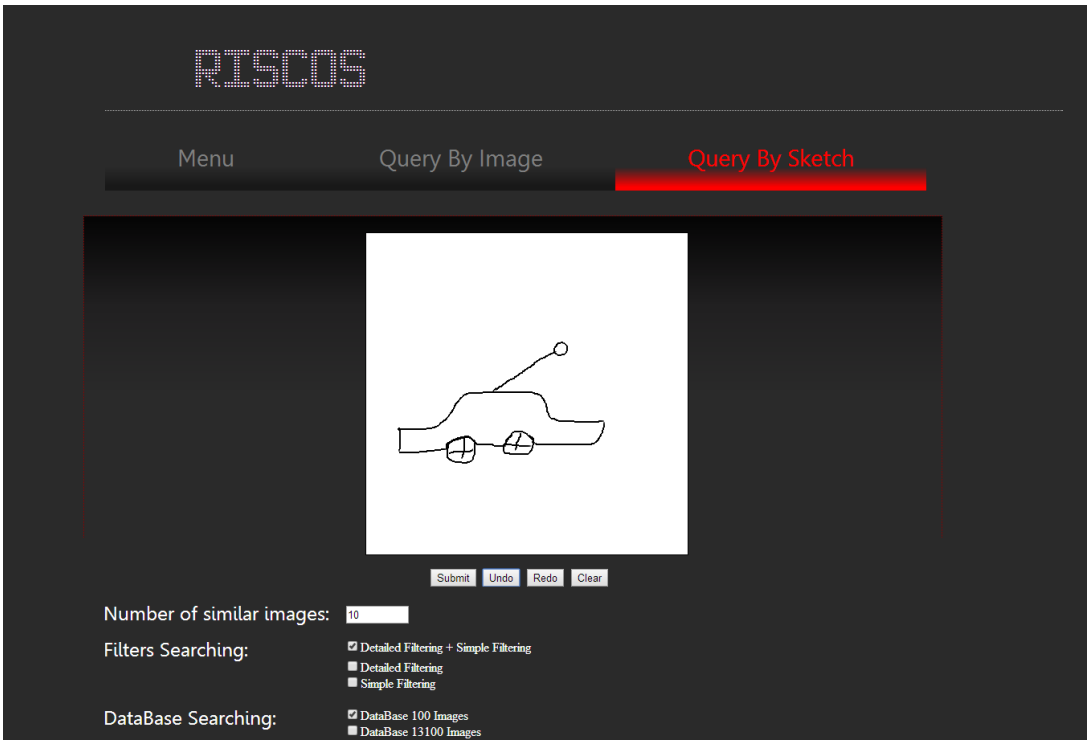


Figura 3.39. Visualização do Menu *Query By Sketch*

3.6 Síntese

Neste capítulo descrevemos a arquitectura da nossa solução. Começando pela imagem, esta sofreu um pré-processamento, consistindo o mesmo em duas filtragens: numa priorizámos o detalhe presente na imagem, enquanto na outra, preocupámo-nos com o contorno envolvente do objecto presente na imagem. Seguidamente, aplicámos uma simplificação da qual resultou, uma imagem com um contorno passível de convertê-la para um formato *vectorial*. Realizada a conversão, extraímos a informação, através dos algoritmos *GeoDiff* e *Geometria*, inserindo a informação obtida numa base de dados.

Em relação ao tratamento efectuado ao esboço, realizámos um pré-processamento com o mesmo objectivo da imagem, em que utilizámos dois tipos de filtragem: numa, preservámos o detalhe e na outra, removemo-lo e obtivemos apenas o contorno que forma o objecto da imagem. Seguidamente efectuámos uma conversão do esboço para um formato *vectorial* e extraímos as características através dos algoritmos *GeoDiff* e *Geometria*. Posteriormente, recorremos a métricas para calcular a similaridade entre as imagens e o esboço.

Concluímos este capítulo com a descrição e decisões tomadas na criação da nossa interface.

4 Avaliação

Neste capítulo abordamos todos os aspectos relevantes dos testes realizados, assim como fazemos uma apreciação das técnicas e algoritmos utilizados na nossa solução. Descrevemos a análise preliminar dos filtros efectuados às imagens, justificando as nossas decisões. Apresentamos também o algoritmo *TopoGeo*, abordado em trabalho anterior, baseando-nos neste, para a realização do nosso algoritmo *GeoDiff*. Por fim, descrevemos as métricas de qualidade, como realizamos uma avaliação para medir a qualidade da nossa solução.

4.1 Análise experimental dos filtros

Focando-se o nosso trabalho na detecção de contornos na imagem, seleccionámos um número de filtros que nos permitiu obter diferentes resultados relacionados com a detecção de contornos existentes na imagem. A utilização dos filtros permitiu-nos suavizar o ruído existente na imagem assim como evidenciar os contornos. Obtidos os primeiros resultados, após a aplicação dos filtros, e não sendo estes os desejados, combinámos diferentes técnicas de forma a obtermos uma imagem com um formato próximo de um esboço. Na tabela 4.1 é possível visualizar os testes realizados e a apreciação dos resultados.

Averiguado o *Filtro de Canny*, verificámos que este proporcionou uma suavização do ruído e uma detecção dos contornos. Por outro lado, o *Filtro de Gabor*, para além de detectar os contornos, acentuou o aparecimento das texturas que compõem a imagem bem como o seu formato geométrico. A utilização do *Filtro Laplaciano* [26] permitiu-nos evidenciar os contornos existentes na imagem, tal como o *Filtro de Sobel*. Em relação ao uso do *Filtro de Média* [27] assim como o *Filtro Gaussiano com Blur*, provocaram um desfocamento na imagem, originando a remoção de ruído presente na imagem. O *Filtro de Mediana* [27] proporcionou uma suavização da imagem, removendo o ruído presente nesta. A utilização do *Filtro de Sharpen* [28] consistiu na nitidez da imagem e acentuação dos contornos presentes.

A utilização das diferentes técnicas e operadores morfológicos teve como base de escolha a influência que estes proporcionaram a uma imagem quando aplicámos em conjunto com um

filtro. Estas técnicas demonstraram de diferentes formas a modificação dada aos contornos presentes na imagem.

4.1.1 Análise da escolha dos filtros

Filtro								Técnica					Operador Morfológico		Ordem de Execução	Apreciação
Gaussiano Blur	Laplaciano	Média	Mediana	Sharpen	Sobel	Canny	Gabor	Encontra Contornos	Fecha Buracos	Esqueletização	OutLine	Binária	Erosão	Fecho		
						x									Canny	Contornos perceptíveis e preservação do detalhe
						x						**x**	**x**		Canny Binária Erosão	Perda da informação da imagem
							x								Gabor	Aparecimento de Sombreamentos
							x	**x**							Gabor Encontra Contornos	Aparecimento de sombreamento mais saliente
	x														Laplaciano	Contornos são vísiveis
	x											**x**	**x**		Laplaciano Binária Erosão	Contornos são sumidos
	x							**x**				**x**		**x**	Laplaciano Encontra Contornos Binária Fecho	Preenchimento de pequenas regiões
		x													Média	Desfocagem da imagem
		x								**x**		**x**			Média Binária Esqueletização	Algumas regiões unem os contornos
		x							**x**			**x**			Média Binária Fecha Buracos	Preenchimento do interior do objecto

		x						**x**	**x**			**x**			Média Binária Fecha Buracos Encontra Contornos	Obtenção do contorno que envolve o objecto
			x												Mediana	Perda de alguns contornos
			x								**x**	**x**			Mediana Binária OutLine	Remoção de contornos
			x					**x**							Mediana Encontra Contornos	Preservação de detalhe na imagem
				x											Sharpen	Evidenciação de contornos das sombras
				x				**x**				**x**		**x**	Sharpen Encontra Contornos Binária Fecho	Aparecimento de algum ruído
					x										Sobel	Preservação de detalhe
					x							**x**	**x**		Sobel Binária Erosão	Perda da informação
					x			**x**							Sobel Encontra Contornos	Dilatação dos contornos
x															Gaussiano Blur	Desfocagem da imagem
x								**x**				**x**			Gaussiano Blur Binária Encontra Contornos	Evidenciação de um contorno global no objecto

Tabela 4.1. Resultado da análise dos diferentes Filtros e técnicas

Realizadas as diversas análises com diferentes combinações de filtros e técnicas (ver tabela 4.1), obtivemos resultados que nos permitiram retirar algumas conclusões.

Ao aplicarmos o *Filtro de Canny* sobre a imagem, foi evidenciado a detecção dos contornos, como a preservação do seu detalhe. Desta forma, toda a informação da imagem foi mantida. Ao combinarmos o filtro com a técnica *Binária* não verificámos qualquer transformação. Combinámos ainda o operador morfológico *Erosão*, tendo verificado que os pixéis que formavam os contornos da imagem foram eliminados, obtendo-se uma imagem preta (perda da informação na totalidade).

Recaindo o nosso objectivo sobre a detecção de contornos numa imagem, um outro tipo de filtro foi analisado, o *Filtro de Gabor*. Quando usámos este filtro foi perceptível a detecção dos contornos na imagem, mas também a enfatização das formas e texturas que compunham os diferentes objectos, assim como o realçamento de algumas sombras. Ao combinarmos a técnica *Encontra Contornos* obtivemos linhas horizontais ao longo da imagem, devido ao *Filtro de Gabor* permitir a passagem de frequências compreendidas entre um determinado limiar.

A utilização do *Filtro Laplaciano* demonstrou-nos a detecção dos contornos presentes na imagem, sendo neste caso preservado o detalhe da mesma sem que existisse ruído. Ao combinarmos o filtro com a técnica *Binária* e a técnica *Erosão*, deparámo-nos com a perda de quase toda a informação da imagem, ficando apenas pequenos pixéis sem que fossem preservadas as formas geométricas dos objectos que a compõem. Quando utilizámos a técnica *Encontra Contornos* e tendo esta como objectivo, realçar os contornos evidenciados na imagem, o resultado apresentado foi de uma maior dilatação dos mesmos. Por outro lado, formaram-se pequenos buracos no próprio contorno. Confrontados com este problema, aplicámos a técnica de *Fecho* como forma de eliminar os buracos formados. Tal técnica evidenciou a sua remoção.

O *Filtro de Média* quando aplicado na imagem possibilitou-nos uma maior suavização da imagem e remoção de ruído presente. A aplicação em conjunto com a técnica *Binária* resultou num *threshold*, verificando-se que a partir de um determinado limiar, as regiões que formavam o objecto tomaram uma cor preta, ficando as restantes com uma cor branca. Obtidos apenas dois tipos de cores na imagem, utilizámos a técnica *Fecha Buracos*, de forma a fechar algumas regiões. Deste modo, todas as regiões circunscritas pela cor preta foram fechadas obtendo-se uma única forma geométrica na imagem. Posto isto, aplicámos a técnica *Encontra Contornos* para evidenciarmos o contorno envolvente no objecto, obtendo-se uma imagem com um formato semelhante a um esboço. Uma outra técnica, *Esqueletização* foi usada, tendo sido evidenciado uma deformação do objecto. Tal facto aconteceu devido à existência de pequenas regiões que formavam o objecto, estarem preenchidas a preto e desta forma, essas regiões foram transformadas num único contorno, perdendo-se o seu formato original.

O *Filtro de Mediana* foi analisado devido à utilização desta proporcionar uma redução do ruído nas imagens, melhorando desta forma a qualidade. Este filtro, combinado com a técnica

Binária, resultou num *threshold* na imagem. Tal como noutras análises efectuadas anteriormente, permitiu-nos eliminar as sombras presentes na imagem. Ao combinarmos com a técnica *OutLine*, notámos um realçamento dos contornos provenientes da aplicação da técnica *Binária* que envolve as regiões, tendo deste modo evidenciado novos contornos na imagem inadvertidamente. Ao utilizarmos o filtro com a técnica *Encontra Contornos*, obtivemos o mesmo comportamento que a técnica *Binária*, apresentando-se, neste caso, apenas os contornos mais dilatados.

Um outro filtro foi examinado, o *Filtro de Sharpen* sendo este especificamente utilizado para a nitidez da imagem. Desta forma, foi possível controlarmos separadamente a intensidade e o espaço de atenuação do filtro nas áreas escuras e claras da imagem. Estando o nosso objectivo focado na detecção de contornos na imagem, combinámos este filtro apenas com a técnica *Encontra Contornos*. O resultado apresentado foi de uma imagem com um formato semelhante a um esboço.

A utilização do *Filtro de Sobel* consistiu na detecção de contornos na imagem, sendo neste caso, evidenciado contornos na transição de zonas mais claras para escuras e vice-versa. Combinámos este filtro com a técnica *Erosão*, tendo existindo um processo intermédio, a aplicação da técnica *Binária*, mas no qual não verificámos transformações significativas na imagem. Após utilizarmos a técnica *Erosão*, a imagem ficou sumida, tendo sido perdida toda a informação existente na mesma. Quando aliámos a técnica *Encontra Contornos* apenas com o *Filtro de Sobel*, evidenciámos não só os contornos existentes na imagem mas também o aparecimento de pequenos buracos.

A utilização do *Filtro Gaussiano com Blur* originou uma desfocagem da imagem conseguindo preservar os contornos desta. Este filtro, quando combinado com a técnica *Binária*, removeu pequenas regiões isoladas no objecto. Ao combinarmos ainda com a técnica *Encontra Contornos*, obtivemos uma imagem com um formato bastante semelhante a um esboço.

Efectuada a análise aos diversos filtros, verificámos que o filtro que proporcionou uma melhor suavização da imagem reduzindo o seu ruído, foi o *Filtro Gaussiano com Blur*. Este filtro demonstrou-nos uma maior preservação dos contornos nas imagens. Outros filtros também apresentaram bons resultados na conservação dos contornos como foi o caso dos *Filtros de Mediana* e o de *Média*.

Efectuado a análise dos diferentes filtros e técnicas às imagens, concluímos que o *Filtro de Canny*, preservou os detalhes da imagem, obtendo-se os contornos da mesma. Por outro lado, a combinação do *Filtro Gaussiano com Blur* com as técnicas *Binária* e *Encontra Contornos* obteve um contorno que envolve o objecto da imagem.

Em suma, e após concretizada a análise com as diferentes técnicas, determinámos que a implementação de dois processos distintos na imagem enriquece a nossa solução. Desta forma, obtivemos uma imagem em contornos e uma imagem com um contorno global do objecto.

4.2 TopoGeo

Nesta secção pretendemos apresentar as diferenças existentes entre dois algoritmos. Como referido anteriormente, o nosso algoritmo *GeoDiff* é uma evolução do algoritmo *TopoGeo*, desenvolvido em trabalho anterior por Fonseca [21]. Ambos incidem na extracção de um descritor de topologia e geometria dos objectos na imagem, diferindo os algoritmos na sua implementação. Em relação ao *TopoGeo*, é utilizado um grafo que descreve a topologia dos diferentes elementos da imagem, contendo à parte uma descrição geométrica destes. Por seu lado, o *GeoDiff* na construção do grafo, agrupa a informação topológica e geométrica dos elementos da imagem, tornando-se menos pesado em termos computacionais e complexo em relação ao algoritmo desenvolvido anteriormente.

4.2.1 Descrição do TopoGeo vs GeoDiff

O algoritmo *TopoGeo* descreve as imagens usando um grafo de topologia e geometria, formado por nós inclusos e adjacentes, os quais representam um elemento diferente na imagem. O grafo é constituído por uma hierarquia pai-filho, nós inclusos, resultando por conseguinte, uma relação de adjacência entre os nós do mesmo nível de profundidade, nós adjacentes, tal como efectuámos para o *GeoDiff*. Existindo diferenças de implementação entre os dois algoritmos, e começando pelo *TopoGeo*, este na formação do grafo é constituído por ligações entre os nós, sendo atribuído por omissão às ligações verticais, ou seja, relação pai-

filho, o valor geométrico 1, sendo apenas calculada a geometria nas ligações entre os nós adjacentes. Esta geometria é dada por um vector adicional com 11 nós, correspondendo a cada um destes, uma característica geométrica. A cada um dos nós do grafo, que representam um elemento da imagem, é realizada uma correspondência para os 11 nós adicionais, atribuindo-se um peso igual ao valor dessa característica geométrica para esse elemento (ver figura 4.1). No caso do *GeoDiff*, calculamos o peso geométrico entre o nó raiz do grafo (pai) e os nós que descendem deste (filhos), como calculamos para as restantes ligações a diferença geométrica entre os nós, nós pai-filho e nós adjacentes. O resultado obtido é atribuído às ligações formadas entre os nós (ver figura 4.2).

Apresentadas as diferenças de implementação entre os algoritmos, *TopoGeo* e *GeoDiff*, evidencia-se que o *TopoGeo*, apresenta uma maior complexidade e um maior peso em termos computacionais, em relação ao *GeoDiff*, devido à forma como descreve a geometria dos elementos na imagem. Existindo esta complexidade, optámos por, no *GeoDiff* descrever a geometria dos elementos na imagem, recorrendo a um cálculo geométrico em vez de utilizar um vector adicional para o efeito. Desta forma, o nosso algoritmo agrupa a informação topológica e geométrica no grafo, reduzindo a complexidade e o peso em termos computacionais do mesmo, obtendo à mesma a topologia e a geometria dos elementos da imagem.

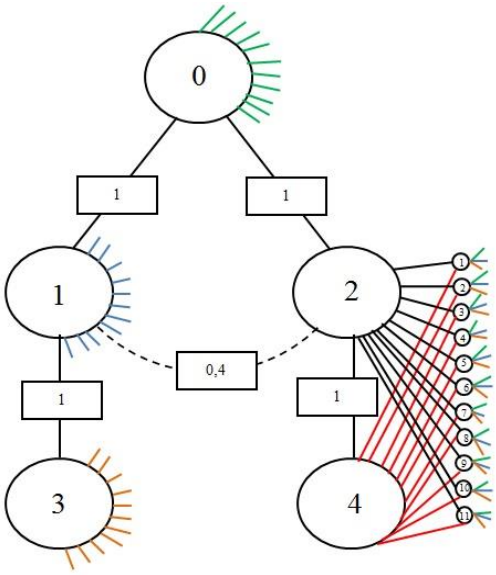


Figura 4.1. Formação do grafo usando o *TopoGeo*

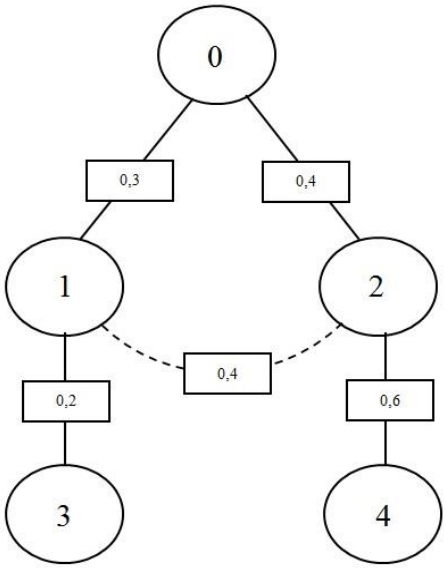


Figura 4.2. Formação do grafo usando o *GeoDiff*

4.3 Procedimento para avaliação da recuperação

Os testes efectuados consistiram na avaliação do desempenho da nossa solução em comparação com outro algoritmo já existente. As pesquisas foram realizadas usando imagens e esboços como interrogações, e uma base de dados controlada obtida do *OpenClipart*[4], composta por dez categorias: *Bulbs*, *Car*, *Fish*, *Flower*, *Hammer*, *Magnifying Glass*, *Phone*, *Sword*, *Tree* e *TV*, tendo cada categoria 10 imagens, num total de 100. Utilizámos ainda outra base de dados com aproximadamente 13000 imagens também obtida do *OpenClipart*, que incluía as 100 imagens da base de dados anterior.

A realização dos testes teve como objectivo averiguarmos o desempenho de dois algoritmos, o *TopoGeo* e o *GeoDiff*, bem como avaliarmos o comportamento destes quando combinados com o descritor adicional de *Geometria* do contorno. Para cada teste usámos como métricas, a *Precision & Recall* [29], o *Mean Average Precision* [30] e o *R-Precision* [31].

4.3.1 Medidas de Qualidade

Para avaliar o desempenho da nossa solução contabilizámos o número de imagens relevantes e não relevantes recuperadas, assim como a ordenação dos resultados, usando imagens como interrogação. Como medidas de desempenho usámos a *Precision & Recall*, que são das métricas mais utilizadas na área de recuperação da informação. A *Precision* centra-se nos resultados recuperados quantos foram relevantes e a *Recall* centra-se em todos os resultados relevantes quantos foram recuperados. Estas medidas podem ser calculadas através das equações 4.1 e 4.2.

$$\text{precision} = \frac{|\text{documentos relevantes} \cap \text{documentos recuperados}|}{|\text{documentos recuperados}|} \qquad (4.1)$$

[4] http://openclipart.org/

$$\text{recall} = \frac{|\text{documentos relevantes} \cap \text{documentos recuperados}|}{|\text{documentos relevantes}|} \qquad (4.2)$$

Também utilizámos outras duas medidas para complementar a análise dos resultados, sendo estas a *Mean Average Precision* (MAP) e a *R-Precision*. Tais medidas, avaliam o desempenho da nossa solução tendo em conta a ordenação dos resultados. Para analisarmos a qualidade dos resultados na base de dados de grandes dimensões, usámos a *R-Precision* (R_p) (equação 4.3) que avalia a relevância dos resultados até ao fim de *n* resultados de uma lista ordenada:

$$\mathrm{R_p} = \frac{\mathrm{r}}{\mathrm{n}} \qquad (4.3)$$

sendo *n*, o número de resultados retornados para conseguirmos ter *r* resultados considerados relevantes.

Em relação ao *Average Precision* [32] (equação 4.4) este é dado pelo valor médio das pesquisas realizadas,

$$\mathrm{AveP} = \frac{\sum_{\mathrm{n=1}}^{\mathrm{N}} \mathrm{R_p} \times \mathrm{rel(n)}}{\mathrm{r_q}} \qquad (4.4)$$

Onde, r_q corresponde ao número total de resultados considerados relevantes na nossa interrogação; *N* o número de resultados recuperados e *rel(n)* (equação 4.5) uma função binária dada pela relevância do n^{th} resultado.

$$\mathrm{rel(n)} = \begin{cases} 1 & \text{se o } \mathrm{n^{th}} \text{ resultado é relevante} \\ 0 & \text{caso contrário} \end{cases} \qquad (4.5)$$

Consistindo o *Mean Average Precision* (equação 4.6) no valor médio da *Average Precision*,

$$\mathrm{MAP} = \frac{\sum_{n=1}^{N} \mathrm{AveP}(n)}{N} \qquad (4.6)$$

Onde *N* corresponde ao número de resultados recuperados.

4.3.2 Resultados da recuperação usando imagens como interrogação

Para as métricas *Precision & Recall*, utilizámos a base de dados com 100 imagens e usámos todas as imagens como interrogações. No caso do *Mean Average Precision*, apenas considerámos os 10 primeiros resultados de cada interrogação, para o cálculo. Em relação à métrica *R-Precision*, utilizámos a base de dados 13000 imagens, em que usámos como interrogações duas imagens de cada uma das dez categorias, obtendo-se no total 20 interrogações. Considerámos para este teste um *R*=3 para as 100 imagens recuperadas. Em todos os testes realizados fizemos uma avaliação do desempenho dos algoritmos *GeoDiff*, *TopoGeo*, *GeoDiff* combinado com *Geometria* e *TopoGeo* combinado com *Geometria*, como uma análise dos resultados por cada categoria.

4.3.2.1 Precision & Recall

A medida *Precision & Recall*, permitiu-nos comparar e analisar os resultados obtidos ao usarmos os algoritmos *GeoDiff* e o *TopoGeo* assim como a combinação de cada um deles, com o algoritmo de *Geometria*. A nossa análise foi realizada de duas formas. Primeiro estudámos o comportamento geral dos algoritmos e depois fizemos uma análise mais detalhada para cada categoria da nossa base de dados. Em termos gerais constatámos que quer o *GeoDiff* combinado com a *Geometria*, quer o *TopoGeo* em conjunto com a *Geometria*, atingiram resultados significativamente superiores (aproximadamente 10% a mais na precisão), em relação aos algoritmos *GeoDiff* e *TopoGeo* (ver figura 4.3).

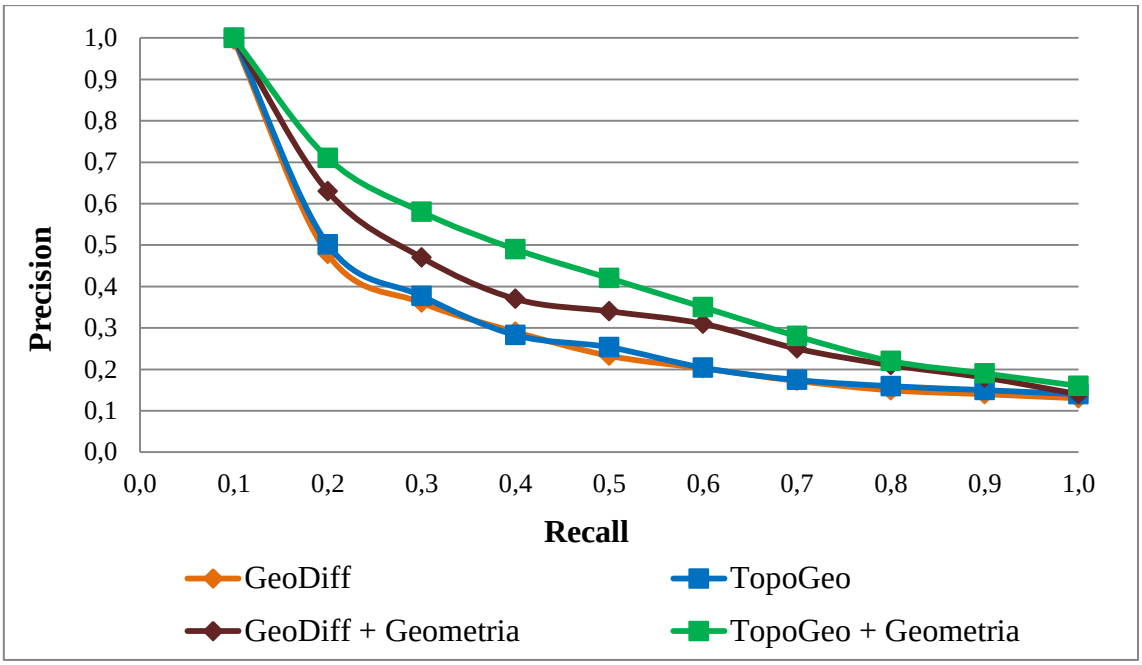


Figura 4.3. Gráfico da *Precision & Recall*

Analisando os resultados obtidos para cada categoria, verificámos que os algoritmos obtiveram diferentes resultados. No caso do *GeoDiff* concluímos que as categorias de imagens *Fish* e *Phone* obtiveram uma precisão superior às outras categorias, resultando num maior número de imagens relevantes recuperadas. A categoria *Fish* apresentou uma precisão de 42%, enquanto a categoria *Phone* obteve uma precisão de 39%, como ilustra a figura 4.4. O *TopoGeo*, por seu lado apresentou na categoria *Fish*, uma precisão de 38% e na categoria *Sword*, 40% de precisão como se constata pela figura 4.4.

Em relação à combinação do *GeoDiff* com a *Geometria*, verificámos que esta apresentou melhores resultados em relação ao *GeoDiff*. Verificámos que em todas as categorias, os resultados obtidos foram superiores, exceptuando-se a categoria *Flower*, a qual obteve uma precisão abaixo do *GeoDiff*. Analisados os resultados apresentados pela figura 4.4, constatámos que a categoria *Fish* e *Sword* obteve a melhor precisão com 55% e 54%, respectivamente. No caso da combinação da *Geometria* com o *TopoGeo*, realçamos as categorias *Magnifying Glass* e *Fish*, tendo estas obtido 74% e 52% de precisão, respectivamente (ver figura 4.4).

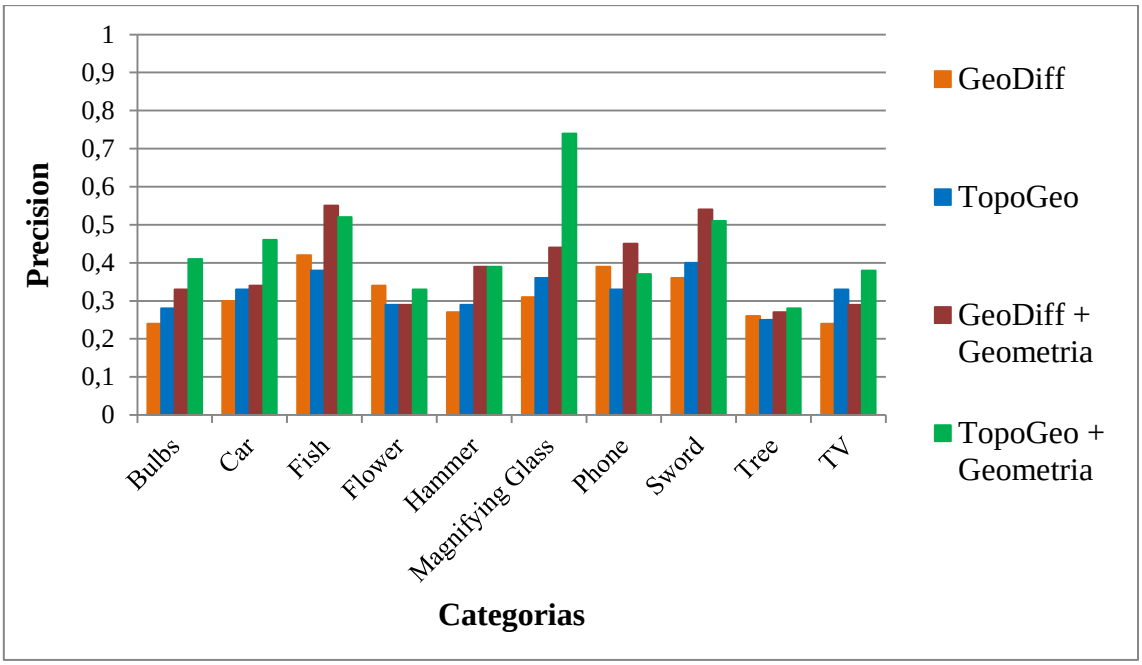


Figura 4.4. Gráfico das categorias *Precision & Recall*

Em suma, e após verificarmos os resultados obtidos pela métrica *Precision & Recall*, constatámos que o *GeoDiff* e o *TopoGeo* obtiveram resultados muito semelhantes, não se evidenciando a superioridade do algoritmo mais complexo e mais pesado em termos computacionais. Quando combinado cada um destes com a *Geometria*, os resultados obtidos foram superiores. Tal facto deve-se à combinação de um descritor que obtém a informação da topologia e diferenças geométricas dos objectos, com um descritor que calcula a geometria do contorno global do objecto na imagem. Esta combinação, não só enriquece a solução como melhorou os resultados, proporcionando um aumento de 10% da precisão em relação à utilização de um descritor que apenas descreve a topologia e as diferenças geométricas dos objectos.

4.3.2.2 Mean Average Precision

Analisados os resultados da *Precision & Recall*, averiguámos posteriormente os resultados tendo em consideração as 10 primeiras imagens recuperadas para cada interrogação. Para tal, utilizámos a métrica *Mean Average Precision* em que comparámos na generalidade os

resultados obtidos através do *GeoDiff* e *TopoGeo*, bem como a combinação de cada um destes com a *Geometria*. Realizámos também uma comparação dos resultados com recurso às categorias. Obtidos os resultados e começando a análise pelos algoritmos *GeoDiff* e *TopoGeo*, verificámos que estes apresentaram um comportamento semelhante, os quais obtiveram uma precisão média de 20% (ver figura 4.5). Quando combinados os algoritmos com a *Geometria*, estes melhoraram os seus resultados, constatando-se para o *TopoGeo*, um aumento na precisão média de 12%, enquanto o *GeoDiff* apenas progrediu em 6% a precisão média, tal como ilustra a figura 4.5.

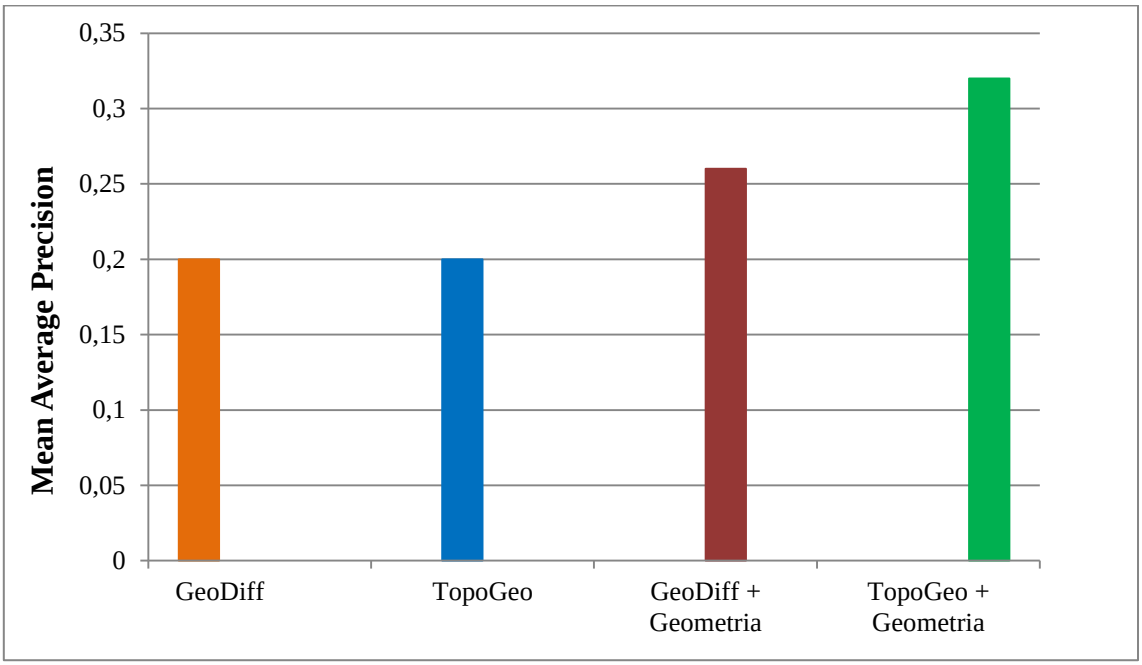


Figura 4.5. Gráfico da *Mean Average Precision*

Em relação à análise efectuada às categorias (ver figura 4.6), verificámos que em alguns casos os resultados foram melhores no *TopoGeo* e outros no *GeoDiff*, evidenciando-se ainda assim, uma superioridade quando utilizada a *Geometria*. Obtidos os resultados, averiguámos que a categoria que obteve melhor precisão quando utilizado apenas o *TopoGeo* foi a *Sword*, com uma precisão média de 31%. Por seu lado, o *GeoDiff* obteve duas categorias com melhores resultados, sendo estas a *Fish* e a *Phone*, ambas com uma precisão média de 28%. Quando analisados os resultados das categorias utilizando a *Geometria* combinada com o *GeoDiff* e a mesma aliada ao *TopoGeo*, os resultados apresentados foram melhores face aos resultados

analisados e descritos anteriormente. A utilização da *Geometria* demonstrou-se vantajosa, tornando os resultados ainda melhores quando combinada com o *TopoGeo*, obtendo em cinco categorias uma precisão média acima dos 30%. Destas categorias, destacou-se a *Magnifying Glass*, que obteve uma precisão média de 55%, seguindo-se a *Sword* com 43%. Em relação ao *GeoDiff*, os resultados obtidos também melhoraram ao aliarmos a *Geometria*, constatando-se pelos resultados apresentados nas diferentes categorias. Mencionamos a categoria *Sword*, que apresentou uma precisão média de 46%, e a categoria *Fish* com uma precisão média de 40%, tal como ilustra a figura 4.6.

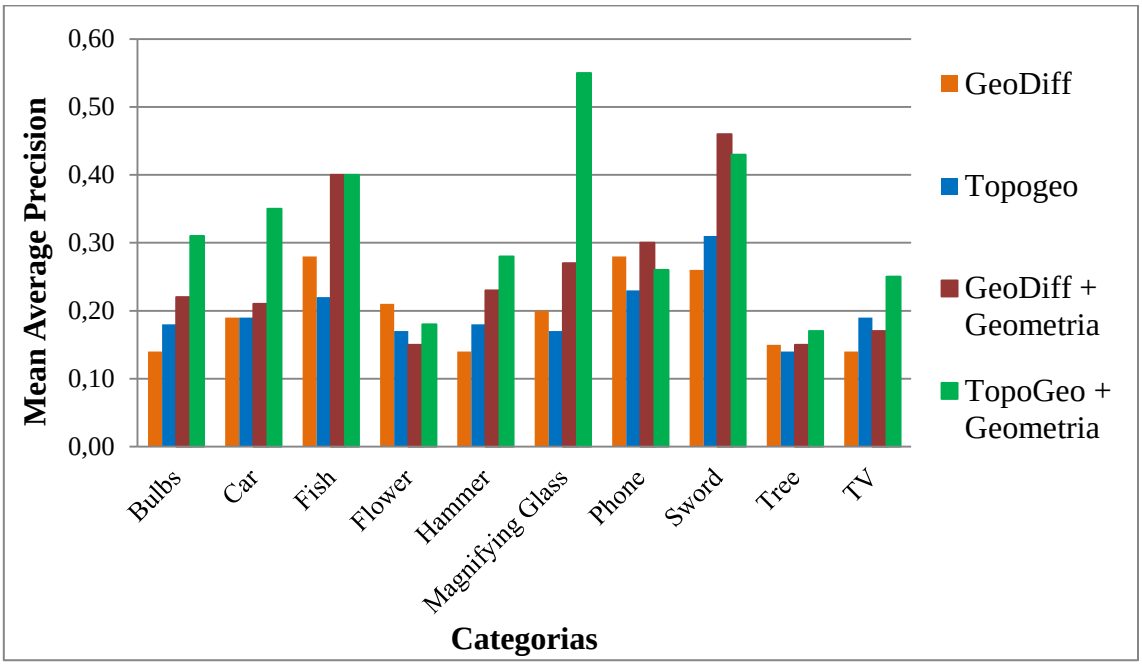


Figura 4.6. Gráfico das categorias *Mean Average Precision*

Em suma, e após efectuada a análise do comportamento na globalidade dos algoritmos e posteriormente com as categorias, constatou-se que os resultados obtidos pelo *GeoDiff* e *TopoGeo* foram idênticos, destacando-se pela positiva, quando combinados cada um com o descritor de *Geometria*. Esta combinação proporcionou uma maior recuperação das imagens nas 10 primeiras posições de cada interrogação. Tal facto deve-se à combinação do descritor que descreve a topologia e diferenças geométricas dos objectos com o descritor que calcula a geometria do contorno global de um objecto

4.3.2.3 R-Precision

Realizada a análise através das métricas *Precision & Recall* e o *Mean Average Precision*, considerámos a métrica *R-Precision* para averiguar os resultados com uma base de dados composta por aproximadamente 13000 imagens. Efectuado o teste, fomos averiguar os resultados tendo em conta os algoritmos utilizados e as categorias das imagens. Obtidos os resultados correspondentes a cada algoritmo, constatámos que o *TopoGeo* obteve o melhor resultado com um *R-Precision* de 11%. Por seu lado, o *GeoDiff* obteve um *R-Precision* de 6% enquanto o *GeoDiff* combinado com a *Geometria* e o *TopoGeo* combinado com a *Geometria* obtiveram ambos um *R-Precision* de 3% (ver figura 4.7). Após obtidos estes resultados, concluímos que os algoritmos sem a combinação da *Geometria* tiveram um comportamento muito melhor. Tal deveu-se ao facto de existirem na base de dados muitas imagens que possuíam uma geometria muito idêntica às imagens submetidas como interrogações.

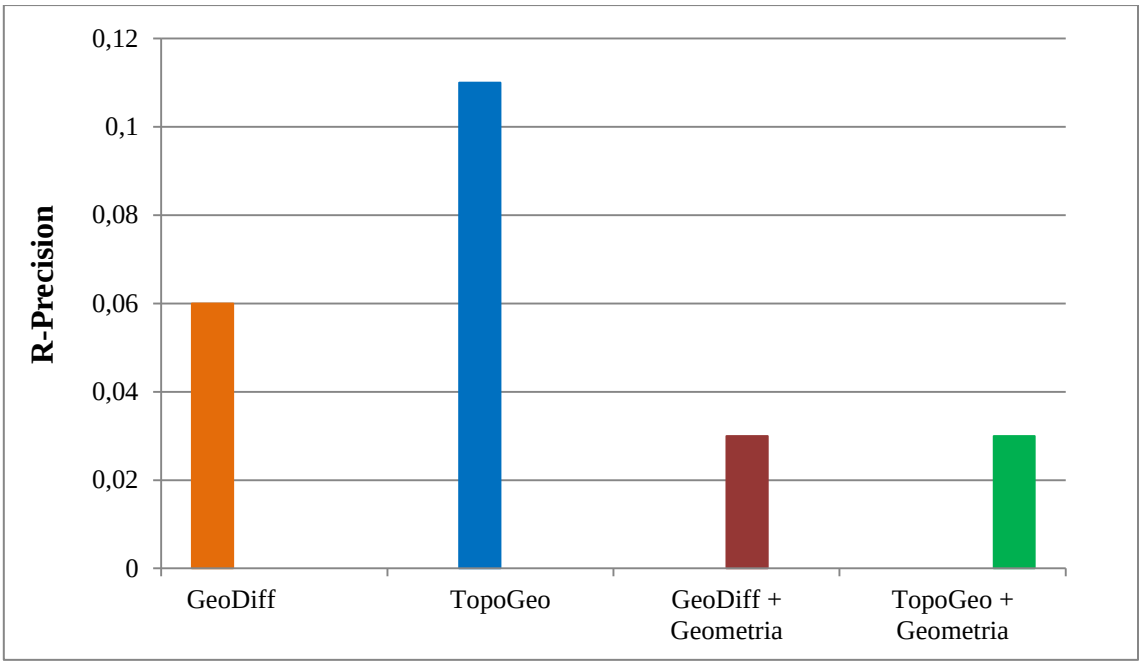


Figura 4.7. Gráfico da *R-Precision*

Em relação aos resultados obtidos pelas diferentes categorias (ver figura 4.8), verificámos que no caso do *TopoGeo*, as categorias que se evidenciaram foram a *Bulbs* e a *Phone* obtendo um *R-Precision* de 51% e 26%, respectivamente. No *GeoDiff* a categoria que se destacou foi a *Phone* obtendo um *R-Precision* de 29%, tal como no *GeoDiff*, combinado com a *Geometria* e o *TopoGeo* com a *Geometria*, tendo estes obtido um *R-Precision* de 8% e 7%, respectivamente, como demonstra a figura 4.8. Obtidos estes resultados, aferimos que a categoria que melhores resultados recuperou foi a *Phone*, para a base de dados não controlada.

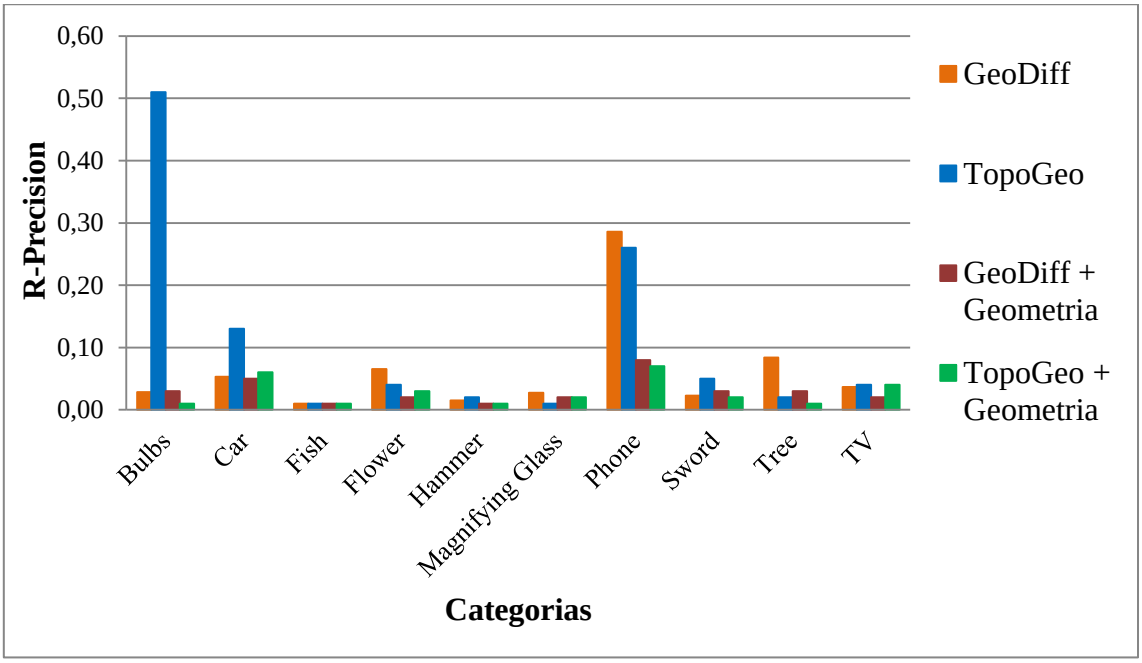


Figura 4.8. Gráfico das categorias *R-Precision*

Analisados os resultados, concluímos que estes demonstraram-se diferentes dos anteriores, ou seja, a combinação de cada um dos algoritmos *GeoDiff* e *TopoGeo* com a *Geometria*, mostraram-se inferiores à utilização de apenas um dos algoritmos. Tal facto deve-se ao elevado número de imagens na base de dados apresentarem uma geometria muito próxima das imagens a interrogações. Consistindo o descritor de *Geometria* no cálculo da geometria do contorno global do objecto na imagem, este apresentou alguma influência na combinação com cada um dos algoritmos, *GeoDiff* e *TopoGeo*, em que apenas descrevem a topologia e geometria dos objectos na imagem. Por outro lado, ao utilizarmos apenas os algoritmos sem a combinação da *Geometria*, verificámos que o *TopoGeo* obteve melhores resultados que o

GeoDiff, constatando-se a preponderância da complexidade da extracção das características geométricas dos objectos efectuada pelo *TopoGeo*.

4.3.3 Resultados da recuperação usando esboços como interrogações

Para a realização destes testes começámos por recolher um conjunto de esboços desenhados pelos utilizadores, para cada uma das 10 categorias de imagens. A cada utilizador foi solicitado que efectuasse um esboço de cada categoria, sem que estes tivessem conhecimento das imagens que estavam na base de dados. Assim sendo, recolhemos de cada utilizador 10 esboços, que posteriormente submetemos como interrogações à base de dados controlada, composta por 100 imagens e à outra com aproximadamente 13000 imagens. Em relação à base de dados formada por 100 imagens, utilizámos as métricas *Precision & Recall*, e o *Mean Average Precision*, através dos quais recuperámos 100 imagens e analisámos a relevância das mesmas. Quanto à métrica *R-Precision*, recorremos à base de dados composta pelas 13000 imagens, utilizando um R=3, tendo também nesta avaliação apenas recuperado 100 imagens.

A recuperação das imagens, usando esboços como interrogações, teve por base a análise dos algoritmos, *GeoDiff*, *TopoGeo* e a combinação de cada um destes com a *Geometria*. Após obtidos os resultados, fazemos uma comparação dos diferentes algoritmos no global e para as diferentes categorias que compõem a nossa base de dados.

4.3.3.1 Precision & Recall

Realizados os testes usando esboços como interrogações, averiguámos os resultados obtidos recorrendo à métrica *Precision & Recall*. Analisando o gráfico da figura 4.9, no qual apresentamos a comparação dos diferentes algoritmos, *GeoDiff*, *TopoGeo* e a combinação destes com a *Geometria*, verificámos que os resultados obtidos foram melhores quando se combina com a *Geometria*. Tal facto demonstrou que a precisão melhorou 6%. A utilização da combinação do *GeoDiff* com a *Geometria* tornou-se uma vantagem para a nossa solução quando utilizados esboços como interrogações, visto que os resultados obtidos foram

melhores quando da utilização apenas do *GeoDiff*. Esta combinação permitiu conciliar a informação extraída em grafo da topologia e diferença geométrica com a informação obtida do cálculo da geometria do objecto, tornando desta forma, os resultados mais precisos.

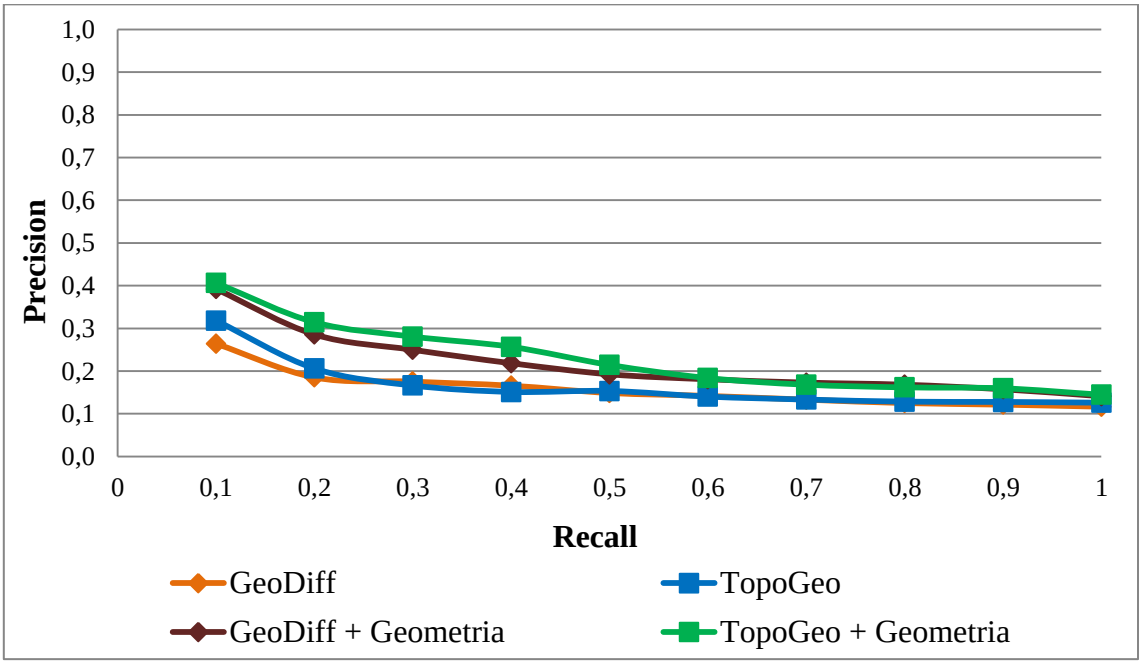


Figura 4.9. Gráfico da *Precision & Recall*

Analisando os resultados obtidos para cada categoria (ver figura 4.10), verificámos que no *GeoDiff* a categoria *Hammer* obteve uma precisão acima de 20%. Em relação ao *TopoGeo*, as categorias que sobressaíram foram a *Magnifying Glass* obtendo uma precisão de 25%, assim como a *Sword*, e a *Hammer* tendo estas obtido uma precisão de 22% e 23%, respectivamente, como ilustra a figura 4.10. Relacionando estes resultados com os obtidos através do *GeoDiff* e o *TopoGeo*, combinados com a *Geometria*, verificámos que os resultados foram superiores. Tal facto constatou-se através dos resultados demonstrados pelas diferentes categorias, existindo no caso do *GeoDiff* com *Geometria* (ver figura 4.10), uma categoria que obteve mais que o dobro da precisão obtida pelo *GeoDiff*, caso da *Magnifying Glass*, assim como as categorias *Phone* e *Sword*, estas com um aumento da precisão de 11% e 15%, respectivamente. No caso do *TopoGeo*, combinado com a *Geometria*, os resultados obtidos também foram superiores face à utilização apenas do *TopoGeo*, verificando-se uma precisão acima dos 50% na categoria *Magnifying Glass* (a qual melhorou em 27%) e a *Sword* com um

aumento de aproximadamente 8%. Verificou-se também que as categorias *Phone* e *Tree* melhoraram em 15% e 6% de precisão, respectivamente.

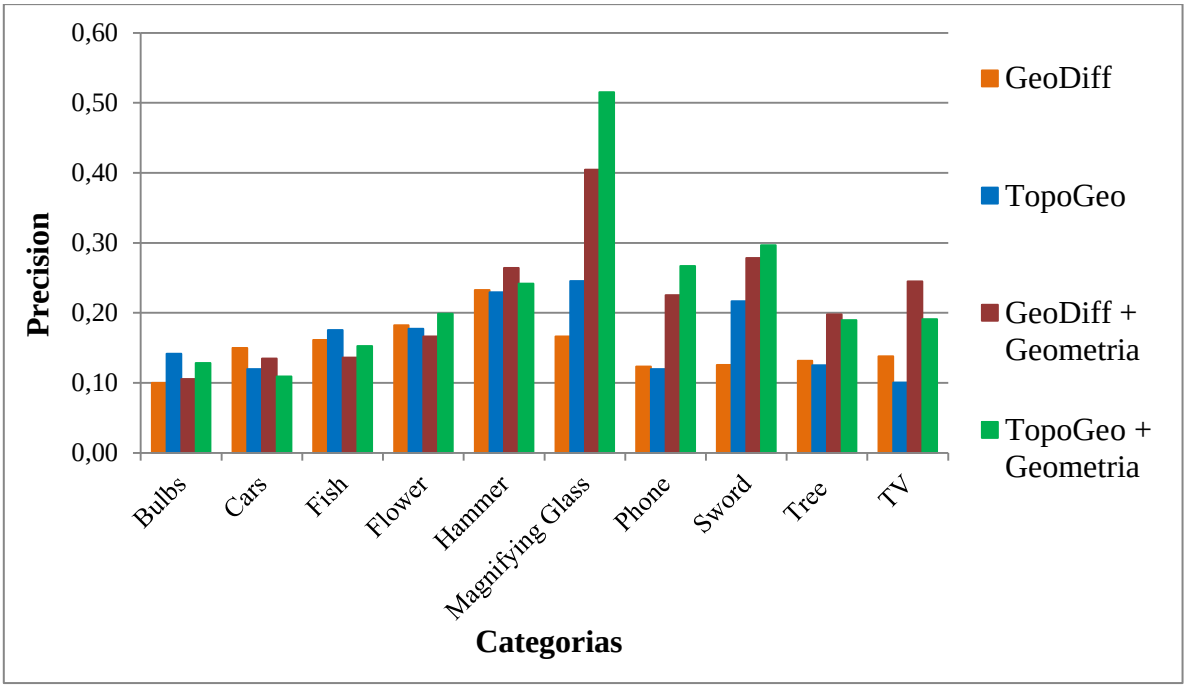


Figura 4.10. Gráfico das categorias *Precision & Recall*

Concretizados os resultados referentes ao *Precision & Recall*, foi perceptível que ao combinarmos a *Geometria* com o *TopoGeo* e o *GeoDiff* melhoraram a precisão, tendo esse valor aumentado em 6%. Quanto aos resultados obtidos pelas diferentes categorias, estes também foram melhores, verificando-se em 6 categorias um aumento da precisão, tendo uma categoria atingido um aumento de 20% a 30%. Dados os resultados, concluímos que a combinação do cálculo da geometria do contorno global do objecto, com a descrição da topologia e diferença geométrica dos objectos, obtiveram um maior número de imagens semelhantes aos esboços desenhados.

4.3.3.2 Mean Average Precision

Analisados os esboços realizados pelos nossos utilizadores, e averiguados os mesmos através da métrica *Precision & Recall*, fomos analisá-los novamente, mas agora para calcular a *Mean Average Precision*. Para isso, considerámos apenas os 10 primeiros resultados obtidos, provenientes das interrogações efectuadas através de esboços. Começando a análise pelo comportamento demonstrado pelos algoritmos *TopoGeo* e *GeoDiff* (ver figura 4.11), verificámos que ambos apresentaram uma precisão média de 5%. Por seu lado, ao combinarmos estes algoritmos com a *Geometria*, a precisão média melhorou em 4% no caso do *GeoDiff* e 6% no *TopoGeo* (ver figura 4.11). Obtidos os resultados e analisado o comportamento dos algoritmos na globalidade, concluímos que a utilização da *Geometria* nas imagens melhorou e enriqueceu a nossa solução, resultando num maior aparecimento de imagens relevantes nas 10 primeiras posições de cada interrogação.

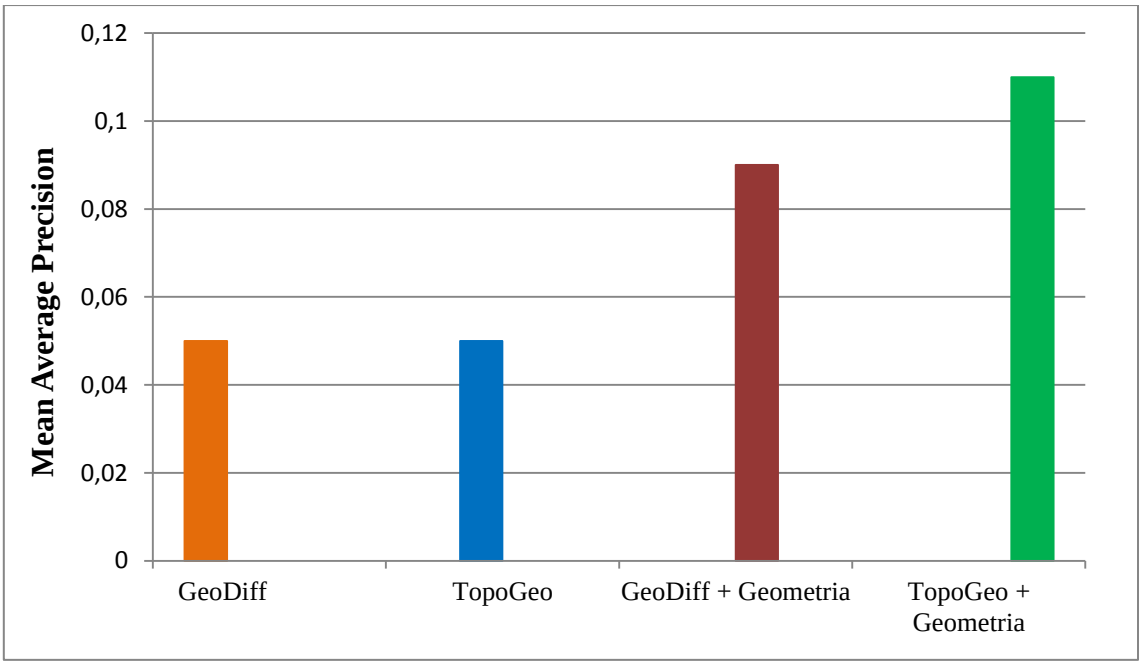


Figura 4.11. Gráfico da *Mean Average Precision*

Realizada a análise do comportamento dos algoritmos aos esboços, fomos averiguar os resultados obtidos mas pelas diferentes categorias. Neste caso, verificámos que existiram

quatro categorias que sobressaíram em relação às restantes, *Magnifying Glass*, *Phone*, *Sword* e *TV*, obtendo o *GeoDiff* combinado com a *Geometria* 21%, 10%, 15% e 13%, respectivamente (ver figura 4.12). Por seu lado o *TopoGeo* com *Geometria*, obteve 32%, 19%, 19% e 10%, respectivamente. O *GeoDiff* e o *TopoGeo*, os melhores resultados que obtiveram foram de 9% e 10% para a categoria *Hammer*, respectivamente. Após verificados estes resultados como ilustrado na figura 4.12 e como dito anteriormente, inferimos que a utilização da *Geometria* foi uma mais-valia na nossa solução, demonstrando mais uma vez a superioridade dos resultados obtidos quando a usámos. A utilização desta permitiu-nos obter um maior número de imagens relevantes nas primeiras posições, em relação à utilização apenas de um único algoritmo, casos do *GeoDiff* e *TopoGeo*.

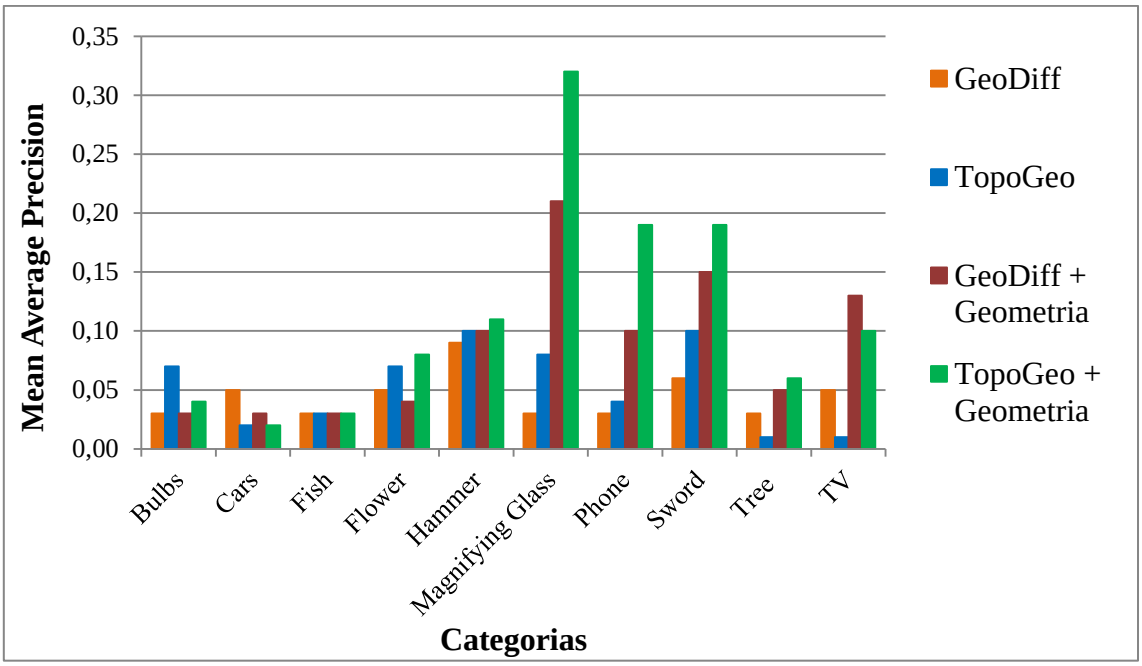


Figura 4.12. Gráfico das categorias *Mean Average Precision*

Em suma, verificámos que os algoritmos *GeoDiff* e *TopoGeo* obtiveram um comportamento idêntico na recuperação de imagens. Quando combinado cada um destes com a *Geometria*, os resultados obtidos foram superiores. Tal facto, e como descrito em casos anteriores, deveu-se à preponderância que o descritor de geometria apresentou na recuperação de imagens através do contorno global do objecto.

4.3.3.3 R-Precision

Realizados os testes através da interrogação de esboços à base de dados composta por 100 imagens, averiguámos em seguida os resultados obtidos mas utilizando a base de dados com 13000 imagens. Esta análise foi efectuada através da métrica *R-Precision*, na qual considerámos um R = 3, tendo nós, constatado pela figura 4.13 que o *R-Precision* obtido por todos os algoritmos foi de 1%.

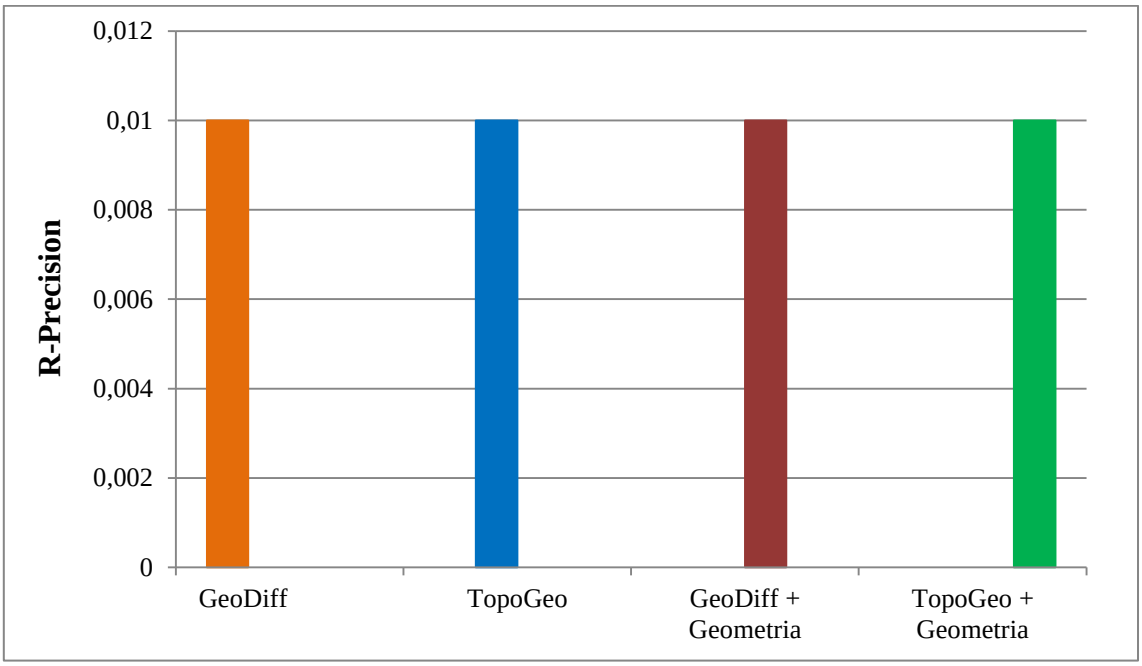


Figura 4.13. Gráfico da *R-Precision*

Concretizada a análise na globalidade aos diferentes algoritmos, averiguámos os resultados obtidos nas diferentes categorias. Recolhidos os resultados das diferentes categorias (ver figura 4.14), verificámos que a categoria *Flower* obteve um maior número de *R-Precision*, destacando-se entre estes, o *TopoGeo* com *Geometria*, obtendo um *R-Precision* de 5%, seguindo-se o *GeoDiff* combinado com *Geometria* com um *R-Precision* próximo dos 5%. Em relação ao *GeoDiff* o melhor resultado obtido foi também pela categoria *Flower* com um *R-Precision* próximo dos 3%, enquanto o *TopoGeo* obteve o maior *R-Precision* próximo dos 6% através da categoria *Sword.*

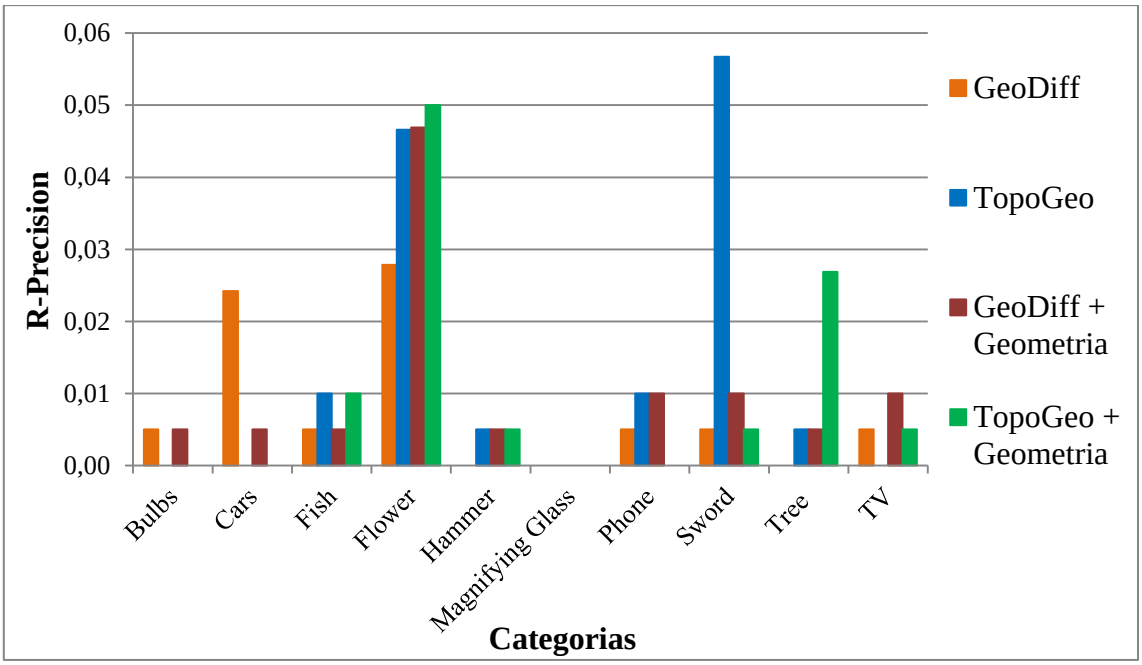


Figura 4.14. Gráfico das categorias *R-Precision*

Analisados os resultados dos algoritmos na generalidade, verificámos que todos os algoritmos apresentaram o mesmo *R-Precision*. Tal facto deve-se ao reduzido número de imagens recuperadas para as 100 primeiras posições através de esboços como interrogações.

Em relação aos resultados obtidos pelas diferentes categorias, concluímos que a utilização do *GeoDiff* e do *TopoGeo*, combinado com a *Geometria*, apresentou melhores resultados, face ao *GeoDiff* e *TopoGeo* em 4 categorias (*Flower*, *Hammer*, *Tree* e *TV*). Dados estes factos, podemos afirmar que a utilização da *Geometria* combinada com os diferentes algoritmos *GeoDiff* e *TopoGeo* obteve melhores resultados, e por conseguinte, uma maior recuperação de imagens relevantes na nossa solução.

4.4 Discussão dos resultados

Concretizada a análise dos resultados obtidos pelas imagens e esboços como interrogações, concluímos que o algoritmo *GeoDiff*, apesar de ser menos complexo face ao *TopoGeo*, nalguns casos obteve um comportamento semelhante ao *TopoGeo* (ver figura 4.3 e 4.11). Verificámos também que a nossa solução obteve melhores resultados quando combinámos

dois descritores. Posto isto, e obtidos os resultados pelas imagens e esboços como interrogações, verificámos que estes foram melhores após recorrermos à *Geometria*. Tal deve-se, ao facto do *GeoDiff* e do *TopoGeo* extraírem a relação espacial e geométrica dos objectos na imagem, proporcionando resultados mais proveitosos para as imagens em que apenas preservámos o detalhe da mesma. Por seu lado, a *Geometria* ao calcular apenas a geometria do contorno global do objecto, enriquece a nossa solução. Dados estes factos, concluímos que a combinação do descritor *GeoDiff*, que descreve a organização espacial e as diferenças de geometria entre os elementos da imagem com o descritor de *Geometria*, que calcula a geometria do contorno global do objecto, apresentaram um maior número de imagens recuperadas, e por conseguinte um enriquecimento da nossa solução (ver figura 4.3 e 4.11).

Como apresentado anteriormente, através da comparação efectuada entre o *GeoDiff* e o *TopoGeo*, concluímos que a implementação do nosso algoritmo é menos complexa e menos pesada computacionalmente, tratando a geometria dos objectos de modo diferente face ao *TopoGeo*. Dados estes factos, constatámos que nalguns casos, quer nas imagens como interrogações, quer nos esboços como interrogações, este obteve resultados muito próximos do *TopoGeo* (ver figura 4.3 e 4.11).

Em suma, concluímos que o *GeoDiff*, ao apresentar uma menor complexidade e um menor peso computacional face ao *TopoGeo*, obteve resultados semelhantes. Quando combinado dois descritores, *GeoDiff* e *Geometria*, obtivemos um maior número de imagens relevantes e por conseguinte um enriquecimento da nossa solução.

4.5 Síntese

Neste capítulo descrevemos a avaliação efectuada à nossa solução. Começámos por analisar os filtros utilizados na imagem, como evidenciámos as principais diferenças entre o *TopoGeo e GeoDiff.* Seguidamente realizámos uma avaliação dos resultados obtidos, utilizando imagens e esboços como interrogações. A avaliação teve por base a utilização de uma base de dados controlada, composta por 10 categorias, cada uma com 10 imagens, e outra constituída por aproximadamente 13000 imagens. No caso da base de dados controlada, as medidas usadas para averiguar os resultados foram a *Precision & Recall* e o *Mean Average Precision*. Quando utilizada a base de dados com as 13000 imagens, utilizámos a medida *R-Precision*.

Nas avaliações realizadas, apresentámos os resultados obtidos para os algoritmos *TopoGeo* e *GeoDiff* e a combinação de cada um destes com a *Geometria*, como fizemos uma apreciação dos resultados obtidos por cada categoria. Em suma, esta avaliação permitiu-nos concluir que a combinação com a *Geometria* enriqueceu a solução, apresentando melhores resultados e por conseguinte, um maior número de imagens recuperadas pelas imagens e esboços como interrogações.

5 Conclusões e Trabalho Futuro

Neste capítulo descrevemos sumariamente o conteúdo da dissertação, assim como fazemos uma reflexão sobre o trabalho realizado. Apresentamos ainda alguns problemas que surgiram durante a realização da nossa solução e que devem ser abordados e considerados num futuro próximo, de forma a melhorar o trabalho efectuado.

5.1 Sumário da Dissertação

Nesta dissertação propusemos uma solução capaz de melhorar e enriquecer as pesquisas de imagens, usando esboços ou imagens como interrogações.

No capítulo 2, abordámos os trabalhos efectuados nas áreas de *recuperação de imagens através de esboços* e *técnicas de detecção de contornos em imagens*. Descrevemos e analisámos as soluções existentes, evidenciando as suas potencialidades e falhas.

No capítulo 3, descrevemos a nossa solução. Começámos por apresentar a arquitectura global da solução, na qual submetemos a imagem a um pré-processamento e a uma simplificação, seguindo-se uma conversão para um formato *vectorial*. Posteriormente extraímos a informação e inserimo-la numa base de dados. No caso do esboço, este sofreu um pré-processamento, seguindo-se uma conversão para um formato *vectorial*. Tal como na imagem, extraímos as características relacionadas com a topologia e a geometria. Através de métricas, calculámos a similaridade entre o esboço e a imagem, apresentando os resultados desta comparação. Seguidamente, detalhamos os processos efectuados à imagem. Um pré-processamento foi aplicado, consistindo este em dois processos, *Filtro de Canny* e *Combinação de técnicas*. O primeiro processo teve por base a utilização de um *Filtro de Canny*, no qual preservámos o detalhe e obtivemos uma imagem em contornos. Dados os contornos encontrarem-se a branco e o fundo a preto, recorremos à técnica de *subtracção*, invertendo as cores da imagem. O segundo processo consistiu na obtenção de uma imagem em contornos mas que nesta, todo o detalhe presente fosse removido, à excepção do contorno global (mais externo). Desta forma, aplicámos um conjunto de técnicas como o *Filtro Gaussiano com Blur*, obtendo como resultado uma suavização da imagem. Seguidamente

efectuámos uma *binarização* para remover os objectos relevantes na imagem, convertendo esta para binário através da técnica *binária*, obtendo-se uma imagem a preto e branco. Os pixéis relevantes tomaram o valor 0, enquanto os restantes o valor 1. Obtida uma imagem a preto e branco aplicámos a técnica *Encontra Contornos*, tendo como resultado uma imagem com um contorno global. Obtidas as imagens em contornos, através da aplicação dos dois tipos de filtragem, efectuámos uma *Simplificação*, consistindo a mesma na combinação de três técnicas: *threshold, fecho* e *esqueletização*. Esta etapa resume-se na aplicação de um *threshold* e um *fecho* para eliminar possíveis traços irrelevantes, para assim pudermos aplicar a *esqueletização* que proporcionou um contorno mais fino na imagem, passível de ser convertido para um formato *vectorial*. Tendo nós a imagem em contornos finos, convertemos a mesma para um formato *vectorial*.

Descritos os processos aplicados à imagem, descrevemos como os mesmos foram efectuados ao esboço. Começámos por realizar um pré-processamento, consistindo este na obtenção de dois esboços, um mais detalhado enquanto no outro eliminámos todo o detalhe e apenas preservámos o contorno envolvente do objecto do esboço. Realizámos uma *Filtragem Detalhada*, na qual utilizámos um *Filtro Gaussiano com Blur* para suavizar os contornos do esboço. Seguiu-se a aplicação da técnica *Binária*, que converteu o esboço para um nível de 8 bits. Esta conversão preservou-nos o detalhe e permitiu-nos aplicar a técnica *Esqueletização* para assim podermos converter o esboço para um formato *vectorial*. A *Filtragem Simplificada* baseou-se na obtenção de um esboço com um contorno global. Desta forma, começámos por aplicar um *Filtro Gaussiano com Blur* e a técnica *Binária*, tendo estes o mesmo efeito, quando usados na filtragem anterior. Seguiu-se a aplicação da técnica *Fecha Buracos*, que removeu os buracos provenientes da realização do esboço. Obtido um esboço com os objectos preenchidos a preto, detectámos o contorno global, através da utilização da técnica *Encontra Contornos*. Por último efectuámos uma *Esqueletização* consistindo este processo na obtenção de um esboço com um contorno fino, o qual permitiu que possamos realizar uma conversão para um formato *vectorial*.

Efectuado o pré-processamento à imagem e ao esboço, extraímos as características de ambos, através dos algoritmos *GeoDiff*, o qual aplicámos apenas à imagem e esboço em que preservámos o detalhe e a *Geometria*, onde apenas aplicámos à imagem e esboço em que obtivemos o contorno global do objecto. O algoritmo *GeoDiff* consistiu na construção de um grafo que descreve a organização espacial e as diferenças geométricas entre os elementos da imagem. Constituída a imagem por objectos, fizemos corresponder a cada nó do grafo um

objecto da imagem, existindo assim uma hierarquia pai-filho. Os objectos que se encontrem inclusos de outros na imagem, têm uma relação pai-filho, enquanto os que são adjacentes entre si, encontram-se ao mesmo nível de profundidade no grafo. Para cada ligação é calculada a diferença de geometria entre os objectos, gerando-se assim uma matriz de geometria. Obtida a matriz, calculámos os descritores da imagem e do esboço. Em relação ao algoritmo de *Geometria*, este baseou-se no cálculo geométrico do contorno global do objecto obtido a partir da imagem, resultando esse cálculo, no descritor geométrico.

Obtidos os descritores da imagem e do esboço, calculámos a similaridade entre ambos, por meio de métricas, e apresentámos os resultados.

Mencionados os processos realizados na nossa solução, descrevemos a interacção que os utilizadores têm com a nossa solução. A realização da nossa interface centrou-se na simplicidade e facilidade de manuseamento. Assim sendo, começámos com a apresentação de um menu atractivo e de fácil compreensão, no qual identificámos explicitamente que pesquisas os utilizadores podem efectuar quando utilizada a nossa aplicação. Os utilizadores ao realizarem a pesquisa através de imagens de interrogação ou através de esboços deparam-se com os campos necessários na realização da pesquisa preenchidos. Os campos preenchidos, possibilitam aos utilizadores optarem apenas por submeter uma imagem a interrogação, caso a pesquisa seja feita através da interrogação de uma imagem exemplo, ou através de esboços tendo assim que fazer um desenho e utilizá-lo como interrogação.

No capítulo 4, e após descrita a nossa solução, começámos por apresentar os testes efectuados a imagens e esboços com diferentes filtros combinados com diferentes técnicas. Seguidamente recorremos a medidas de qualidade que nos permitiram averiguar os resultados obtidos. Utilizámos uma base de dados controlada, formada por 10 categorias, cada uma com 10 imagens e outra com aproximadamente 13000 imagens. Para avaliarmos a nossa solução, utilizámos a base de dados controlada para averiguarmos os resultados quando recorremos às medidas *Precision & Recall* e ao *Mean Average Precision*. A base de dados composta por 13000 imagens foi utilizada para analisarmos os resultados quando usámos a medida *R-Precision*. Procedemos também a uma comparação entre os algoritmos *GeoDiff* e *Topogeo*, os quais foram tidos em consideração na avaliação dos nossos resultados.

5.2 Conclusões Finais e Contribuições

No capítulo 4, apresentámos os resultados experimentais efectuados à nossa solução dos quais retirámos as devidas conclusões. Existindo pesquisa de imagens através de imagens exemplo, criámos uma pesquisa de imagens mas através de esboços. Como forma de enriquecer estas pesquisas, optámos pela aplicação de duas filtragens, as quais consistiram na preservação do detalhe e outra na simplificação das imagens. Assim sendo, realizámos diversos testes, resumindo-se os mesmos na análise do comportamento da nossa solução quando recorremos ao *TopoGeo* ou por outro lado ao *GeoDiff* como analisámos a combinação de cada um destes com a *Geometria*. Averiguámos também os resultados apresentados pelas diferentes categorias. Posto isto, podemos afirmar que o *GeoDiff* apresentou resultados semelhantes ao *TopoGeo*, quando não utilizada a informação geométrica do contorno global do objecto, concluindo-se deste modo que um algoritmo menos pesado em termos computacionais e menos complexo, *GeoDiff*, teve o mesmo comportamento que o algoritmo *TopoGeo*, como ilustra a figura 4.3. Quando combinado cada um dos algoritmos com a *Geometria*, verificou-se que os resultados obtidos foram melhores, apresentando um maior número de imagens relevantes (ver figura 4.3). A contribuição que o algoritmo *GeoDiff* oferece é uma menor complexidade e menor peso em termos computacionais na extracção da informação topológica e geométrica dos elementos na imagem. Por outro lado, o algoritmo de *Geometria*, ao calcular a geometria do contorno global do objecto na imagem, enriqueceu a nossa solução. A combinação de ambos os descritores proporcionou um maior número de imagens relevantes, melhorando os resultados da nossa solução.

Outro aspecto a realçar, relaciona-se com os processos de filtragem utilizados na nossa solução. Existindo dois tipos de filtragem, em que numa preservamos o detalhe e a informação pertencente à imagem, e na outra descartamos a informação dando ênfase ao contorno que forma o objecto, podemos assegurar que este processamento enriqueceu as pesquisas. Deste modo, obtivemos uma maior precisão dos resultados quando submetemos esboços a interrogação.

Analisados os resultados, e tal como referido no capítulo 2, a combinação das técnicas *recuperação de imagens através de esboços* e *detecção de contornos em imagens* quando combinadas, apresentaram resultados promissores.

Um contributo essencial para a nossa solução foi a arquitectura da mesma, a qual combinou todos os processos e proporcionou-nos os resultados a que nos propusemos. Referimo-nos às etapas de pré-processamento, as quais nos permitiram tratar convenientemente as imagens e os esboços de modo a obtermos os mesmos em contornos. Posteriormente aplicámos uma conversão às imagens e ao esboço, tendo dessa forma um formato comum a ambos e assim podermos utilizar o *GeoDiff* para extraímos as características essenciais e a *Geometria* para a extracção das informações relacionadas com a geometria do contorno global do objecto. Esta combinação tornou a nossa solução mais enriquecedora, uma vez que a recuperação de imagens melhorou, como verificado pela análise dos resultados, recuperando um maior número de imagens em comparação com o uso apenas do *GeoDiff*.

5.3 Trabalho Futuro

Analisada a nossa solução e averiguados os resultados obtidos, concluímos que a mesma apresentou algumas condicionantes, as quais podem ser abordadas num futuro próximo. Apesar de não termos dado enfâse ao desenvolvimento da interface, esta poderá ser melhorada para uma possível publicação *online* da solução de recuperação de imagens. Tendo em consideração este aspecto, uma melhoria da usabilidade deve ser ponderada, proporcionando aos utilizadores uma maior interactividade e motivação no uso da solução. Em relação ao trabalho a que nos propusemos, apesar dos resultados serem bons, consideramos que estes ainda podem ser melhorados. Começando pelo tratamento realizado às imagens, este utilizou duas alternativas de filtragem: numa preservámos o detalhe e na outra eliminámos a informação obtendo apenas um contorno. A realização do processamento às imagens e aos esboços descartou as cores mas preservou a topologia das mesmas. Existindo esta condicionante na nossa solução, um dos caminhos a explorar no futuro seria a utilização da cor. Combinando a nossa solução com esta alternativa, enriquecer-se-iam as pesquisas, considerando-se imagens com a mesma topologia mas também com a mesma coloração. Realizada esta solução, os utilizadores teriam duas opções de pesquisa, uma usando esboços monocromáticos e outra através de esboços a cores.

Um outro aspecto que pode valorizar a solução será a implementação de um novo algoritmo que seja mais eficaz na extracção de características, proporcionando um melhoramento na

detecção da informação presente nas imagens, tornando deste modo o cálculo da similaridade mais preciso.

6 Referências Bibliográficas

[1] Y. Cao, C. Wang, L. Zhang e L. Zhang, "Edgel Index for Large-Scale Sketch-based Image Search," em *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[2] R. Santos, "Reconhecimento de Esboços de Formas Geométricas Contidas em Fluxogramas," em *S.W.I.B - Simpósios SBSC, WebMedia, IHC e SBBD*, Brasil, Outubro 2010.

[3] H. Ip, A. Cheng, W. Wong e J. Feng, "Affine-invariant Sketch-based Retrieval of Images," em *Proceeding CGI '01 Proceedings of the International Conference on Computer Graphics*, USA, 2001.

[4] M. Eitz, R. Richter, H. Kristian, T. Boubekeur e M. Alexa, "Sketch-Based Shape Retrieval," em *ACM Transactions on Graphics* , 2012.

[5] F. Banfi, "Content-Based Image Retrieval Using Hand-Drawn Sketches and Local Features: a Study on Visual Dissimilarity," St. Paul, Fribourg, 2000.

[6] K. Bozas e E. Izquierdo, "Large Scale Sketch Based Image Retrieval Using Patch Hashing," em *ISVC* , 2012.

[7] Y. Chen, V. Roussev, G. R. I. G. e Y. Gao, "Content-Based Image Retrieval for Digital Forensics," em *In Proceedings of the First International Conference on Digital Forensics*, 2005.

[8] R. Hu, B. Mark e J. Collossome, "Gradient Field Descriptor For Sketch Based Retrieval and Localization," em *In Proceedings of Intl. Conf. on Image Proc. (ICIP)*, Hong Kong, Setembro 2010.

[9] A. Mertins, A. Chalechale e G. Naghdy, "Sketch-Based Image Retrieval Using Angular Partitioning," em *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology*, 2003.

[10] H. Galoogahi e T. Sim, “Face Photo Retrieval by Sketch Example,” em *Proceedings of the 20th ACM international conference on Multimedia - MM '12*, Nova Iorque - USA, 2012.

[11] M. Silva e A. Candeias, “Um Novo Algoritmo para Detecção de Contorno de Melanomas em Imagens de Pele,” em *Sibgrapi - Workshop of Undergraduate Work*, Brasil, Outubro 2008.

[12] M. Eitz, K. Hildebrand, T. B. Boubekeu e M. Alexa, “An evaluation of descriptors for large-scale image retrieval from sketched feature lines,” em *Computers & Graphics*, 2009.

[13] H. Matos, “Reconhecimento Biométrico Baseado na Geometria da Mão,” em *Tese de mestrado integrado em Engenharia Electrotécnica e de Computadores*, Porto, 2010.

[14] R. Zanin, A. Dal Poz e É. Martins, “Esqueletização na Modelagem e Extração de Cruzamentos Simples de Rodovias,” em *VI Workshop de Visão Computacional*, 2010.

[15] Z. Fang, M. Yulei e Z. Junpeng, “Medical Image Processing Based on Mathematical Morphology,” em *The 2nd International Conference on Computer Application and System Modeling*, 2012.

[16] P. Tarábek, “Morphology Image Pre-Processing for Thinning Algorithms,” em *Control and Management Systems*, 2007.

[17] A. Pires, “Processamento de Imagem para Detecção de Chamas: Estudo e estabelecimento de bases para o desenvolvimento de um sistema de detecção de chamas de baixo custo por meio de processamento de imagem,” Universidade do Estado de Santa Catarina , 2012.

[18] M. V. Shah, M. A. Choudhary e P. K. Tewari, “River extraction from satellite image,” em *International Journal of Computer Science Issues*, 2011.

[19] F. Silva e F. Alves, “Aplicação de Técnicas de Processamento de Imagens Digitais em Imagens Geradas por Ultra-Som,” em *8º Encontro Regional de Matemática Aplicada e Computacional*, Universidade Federal do Rio Grande do Norte, 2008.

[20] J. Canny, “A computational approach to edge detection,” em *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 1986.

[21] M. Fonseca, “Sketch-Based Retrieval in Large Sets of Drawings,” em *Dissertation for the degree of Doctor of Philosophy in Information Systems and Computer Engineering*, Instituto Superior Técnico, 2004.

[22] M. Fonseca e J. Jorge, “CALI: Uma Biblioteca de Componentes para Interfaces Caligráficas,” em *X Encontro Português de Computação Gráfica*, Portugal, 1999.

[23] J. Boyce e D. Dobkin, “Finding External Polygons,” em *SIAM Journal on Computing*, Fevereiro, 1985.

[24] M. Fonseca e J. Jorge, “NB-Tree: An Indexing Structure for Content-Based Retrieval in Large Databases,” Portugal, 2003.

[25] M. Basgalupp, “LEGAL-Tree: um algoritmo genético multi-objetivo para indução de árvores de decisão,” em *Tese de Douturado em Ciências de Computação e Matemática Computacional*, São Carlos, 2010.

[26] M. Sharifi, M. Fathy e M. Mahmoudi, “A Classified and Comparative Study of Edge Detection Algorithms,” em *Information Technology: Coding and Computing, 2002. Proceedings. International Conference on*, 2002.

[27] N. Pássaro e N. Faria, “Filtragem de Imagens Fixas,” em *Faculdade de Engenharia da Universidade do Porto*, Porto, 2006.

[28] T. Mahmoud e S. Marshall, “Edge-Detected Guided Morphological Filter for Image Sharpening,” em *EURASIP Journal on Image and Video Processing*, 2008.

[29] D. Inkpen, “Information Retrieval on the Internet,” em *Ph.D*, University of Toronto, Canada.

[30] R. Jain, S. Nathawat, R. Bhargava e G. Purohit, “Efficient Information Retrieval for Ambiguous Words,” em *International Journal of Advanced Research in Computer and Communication Engineering*, 2013.

[31] S. Kim, T. Baldwin e M. Kan, “Evaluating N-gram based Evaluation Metrics for Automatic Keyphrase Extraction,” em *Proceedings of the 23rd International Conference on Computacional Linguistics*, Beijing, 2010.

[32] K. Kishida, “Property of Average Precision and its Generalization: An Examination of Evaluation Indicator for Information Retrieval Experiments,” em *Faculty of Cultural Information Resources, Surgudadai University*, Japan, 2005.