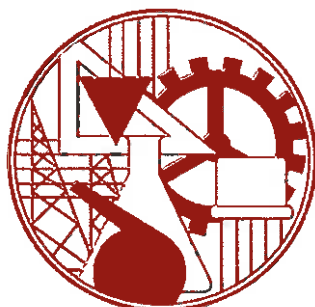


# **Instituto Superior de Engenharia de Lisboa**

Área Departamental de Engenharia de Electrónica e Telecomunicações e de Computadores



## **Aplicação para Pesquisa de Clip Arts em Dispositivos Móveis Baseada em Esboços**

**ANA SOFIA BRÁS CARRILHO MOURATO**

**Trabalho Final de Mestrado para Obtenção do Grau de Mestre em  
Engenharia de Redes de Comunicação e Multimédia**

**Orientador:**

Professor Doutor Rui Jesus

**Júri:**

**Presidente:**

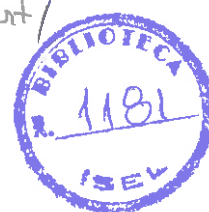
Professor Doutor Arnaldo Abrantes

**Vogais:**

Professor Doutor João Ferreira

Professor Doutor Rui Jesus

Lisboa  
Dezembro, 2013



## Instituto Superior de Engenharia de Lisboa

Área Departamental de Engenharia de Electrónica e Telecomunicações e de Computadores



### Aplicação para Pesquisa de Clip Arts em Dispositivos Móveis Baseada em Esboços

ANA SOFIA BRÁS CARRILHO MOURATO

Trabalho Final de Mestrado para Obtenção do Grau de Mestre em  
Engenharia de Redes de Comunicação e Multimédia

**Orientador:**

Professor Doutor Rui Jesus

**Júri:**

**Presidente:**

Professor Doutor Arnaldo Abrantes

**Vogais:**

Professor Doutor João Ferreira

Professor Doutor Rui Jesus

Lisboa  
Dezembro, 2013

**Aplicação para Pesquisa de Clip Arts em Dispositivos Móveis Baseada em Esboços**

Copyright © Ana Sofia Brás Carrilho Mourato, Instituto Superior de Engenharia de Lisboa

O Instituto Superior de Engenharia de Lisboa tem o direito de arquivar e publicar esta dissertação, seja através de exemplares impressos, reproduzidos em papel, de forma digital, ou por qualquer outro meio conhecido, podendo divulgá-la através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

*Para os meus pais*

# Agradecimentos

Em primeiro lugar, gostaria de agradecer ao meu orientador, ao Professor Rui Jesus, pelo apoio incansável, disponibilidade e motivação demonstrados ao longo de todo o desenvolvimento desta dissertação. Foi um trabalho árduo, mas vale sempre a pena trabalhar e lutar quando no fim conseguimos cumprir os nossos objetivos. Agradeço ao projeto CRUSH (Clip art Retrieval using Sketches), no qual esta dissertação se insere, com referência PTDC/EIA-EIA/108077/2008, financiado pela Fundação para a Ciência e Tecnologia, que financiou a compra de um *tablet Samsung* para a realização desta dissertação. Agradeço ainda todo o apoio do Pedro Martins, aluno da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (FCT/UNL), que tendo desenvolvido uma dissertação também inserida no projeto CRUSH, se prontificou sempre para disponibilizar informações e conteúdos, que contribuíram significativamente para o sucesso no desenvolvimento desta dissertação. Quero ainda agradecer a todos os que de alguma forma me acompanharam neste percurso, em especial aos meus colegas de curso, pelo companheirismo e interesse sempre demonstrados.

Por fim mas não menos importante, quero agradecer a toda a minha família, pelo apoio incondicional e compreensão que sempre me prestaram. À minha mãe, que me acompanhou sempre com carinho e atenção, ao longo de todas as etapas da minha formação académica, deitando-se tarde muitas vezes só para me acompanhar enquanto trabalhava. Ao meu pai, pelo orgulho que sei que sente neste momento e por toda a motivação que sempre me transmitiu. À minha irmã, por me encorajar a ser melhor, por todo o carinho e cumplicidade que partilhamos. Ao meu pequeno sobrinho Tiago, que todos os dias me dá alegria e força para continuar a lutar por um futuro melhor.

Com o apoio de todos os que aqui mencionei e eventualmente outros que não foram referidos, consegui chegar ao fim de mais uma etapa na minha vida.

# Resumo

---

Existe atualmente um crescimento exponencial na diversidade de imagens, entre elas *clip arts*. Quando um utilizador procura uma imagem *clip art* para inserir nos seus documentos, pretende fazê-lo de forma rápida e eficaz. De uma forma geral, os *clip arts* podem ser encontrados através da utilização de palavras-chave, categorias ou, se estes métodos se apresentarem ineficientes, percorrendo manualmente várias listas de imagens, podendo tornar-se uma tarefa desgastante e monótona para o utilizador. Como resultado, surgiram os sistemas CBIR (Content-Based Image Retrieval), que podem ser uma solução mais eficiente, com técnicas que podem resolver alguns dos maiores problemas na pesquisa de imagens. Em alguns casos, pode ser mais fácil procurar uma imagem através do desenho de algumas formas básicas produzidas pela imaginação, do que associar descrições textuais que muitas vezes não conseguem captar a essência da imagem [1]. Esta é uma forma de dar relevância à interpretação visual do conteúdo da imagem para indexação e recuperação.

Nesta dissertação, pretende-se apresentar uma solução que permite aos utilizadores pesquisar *clip arts* através dos seus próprios esboços, com base no conteúdo das imagens. Como tal, pretende-se fazer o estudo e o desenvolvimento de uma interface de pesquisa para dispositivos móveis, nomeadamente para *tablets* com sistema operativo Android. Este trabalho inclui ainda a avaliação da aplicação com utilizadores.

**Palavras-chave:** Pesquisa baseada em esboço, Imagens *Clip art*, Imagens vetoriais, Imagens *raster*, Aplicações para Android

# Abstract

---

There is now an exponential growth in diversity of images, including clip arts. When a user searches for a clip art image to insert in its documents, he wants to do it quickly and efficiently. Typically, clip arts can be found through the use of keywords, categories or, if these methods present themselves inefficient, manually through several lists of images, making it a exhausting and tedious task for the user. As a result, emerged CBIR systems (Content-Based Image Retrieval) that may be a more efficient solution using techniques that can solve some of the biggest problems in the image retrieval. In some instances, it may be easier to search for an image by drawing some basic shapes produced by the imagination, than to associate a textual description that can not always capture the essence of the image [1]. This is a way to give importance to visual interpretation of image content for indexing and retrieval.

In this dissertation, we intend to present a solution that allows users to search clip arts through their own sketches, based on the content of images. As such, it is intended to study and develop a search interface for mobile devices, especially for tablets with Android operating system. This work also includes evaluation of application with users.

**Keywords:** Search based on sketch, Clip art images, Vector images, Raster images, Android applications

# Acrónimos e Abreviaturas

<b>API</b>	Application Programming Interface
<b>CBIR</b>	Content-Based Image Retrieval
<b>CPU</b>	Central Processing Unit
<b>HCI</b>	Human-Computer Interaction
<b>HSB</b>	Hue, Saturation, Brightness
<b>HSV</b>	Hue, Saturation, Value
<b>IP</b>	Internet Protocol
<b>JAR</b>	Java ARchive
<b>JPEG</b>	Joint Photographic Experts Group
<b>MPEG</b>	Moving Picture Experts Group
<b>mHCI</b>	mobile Human-Computer Interaction
<b>PCA</b>	Principal Component Analysis
<b>PNG</b>	Portable Network Graphics
<b>QbE</b>	Query by Example
<b>QbIC</b>	Query by Image Content
<b>QbS</b>	Query by Sketching
<b>RGB</b>	Red, Green, Blue
<b>SVG</b>	Scalable Vector Graphics



# Conteúdo

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>1</b>
1.1	MOTIVAÇÃO .....	2
1.2	DESCRIÇÃO E CONTEXTO .....	3
1.3	SOLUÇÃO APRESENTADA .....	5
1.3.1	<i>Recuperação de Imagem baseada no Conteúdo.....</i>	<i>6</i>
1.3.2	<i>Principais Contribuições Previstas .....</i>	<i>7</i>
1.4	ESTRUTURA DA DISSERTAÇÃO .....	8
<b>2</b>	<b>TRABALHO RELACIONADO .....</b>	<b>11</b>
2.1	INTRODUÇÃO .....	11
2.2	CARACTERÍSTICAS DAS IMAGENS CLIP ART .....	13
2.2.1	<i>Imagens Raster.....</i>	<i>13</i>
2.2.2	<i>Imagens Vetoriais .....</i>	<i>14</i>
2.3	APLICAÇÕES RELACIONADAS .....	15
2.3.1	<i>Indagare.....</i>	<i>16</i>
2.3.2	<i>Retrievr .....</i>	<i>16</i>
2.3.3	<i>Electronic Cocktail Napkin.....</i>	<i>17</i>
2.3.4	<i>Google Goggles .....</i>	<i>18</i>
2.3.5	<i>Google Images .....</i>	<i>19</i>
2.3.6	<i>Microsoft Word e Microsoft Power Point.....</i>	<i>19</i>
2.4	INTERFACES DE QUERY .....	20
2.4.1	<i>Pesquisa por Texto ou Categoria (Query by Keywords) .....</i>	<i>20</i>
2.4.2	<i>Pesquisa por Imagens (Query by Example).....</i>	<i>21</i>
2.4.3	<i>Pesquisa por Esboço (Query by Sketch).....</i>	<i>22</i>
2.5	FERRAMENTAS DE DESENHO EM DISPOSITIVOS MÓVEIS.....	22
2.5.1	<i>Sketch.....</i>	<i>23</i>
2.5.2	<i>Doodledroid.....</i>	<i>23</i>
2.5.3	<i>Paperless.....</i>	<i>24</i>

2.5.4	<i>SketchBook Pro for Tablets</i> .....	24
<b>3</b>	<b>INTERFACE DE PESQUISA PARA TABLET</b> .....	<b>27</b>
3.1	NOTA INTRODUTÓRIA.....	27
3.2	PRIMEIRA ANÁLISE.....	28
3.2.1	<i>Menu da Aplicação</i> .....	29
3.2.2	<i>Área de Desenho</i> .....	33
3.2.3	<i>Visualização de Resultados</i> .....	34
3.2.4	<i>Características de Usabilidade</i> .....	36
3.3	IMPLEMENTAÇÃO DA INTERFACE.....	37
3.3.1	<i>Especificação da Plataforma</i> .....	37
3.3.2	<i>Desenvolvimento</i> .....	38
3.4	AVALIAÇÃO DA INTERFACE UTILIZADOR .....	41
<b>4</b>	<b>SISTEMA DE RECUPERAÇÃO DE IMAGEM</b> .....	<b>43</b>
4.1	SISTEMA DESENVOLVIDO.....	43
4.1.1	<i>Processamento da Base de Dados</i> .....	44
4.1.2	<i>Visão Geral do Sistema em Execução</i> .....	45
4.2	COMUNICAÇÃO ENTRE O DISPOSITIVO MÓVEL E O SISTEMA DE RECUPERAÇÃO DE IMAGEM	47
<b>5</b>	<b>EXTRAÇÃO DE INFORMAÇÃO EM CLIP ARTS</b> .....	<b>49</b>
5.1	PRÉ-PROCESSAMENTO .....	49
5.2	DESCRITOR MOMENTOS DE COR.....	51
5.3	DESCRITOR TOPO GEO .....	52
5.4	NORMALIZAÇÃO DE DESCRITORES .....	54
5.5	CÁLCULO DE DISTÂNCIAS .....	55
5.6	DESCRITORES COMBINADOS .....	56
<b>6</b>	<b>TESTES E RESULTADOS EXPERIMENTAIS</b> .....	<b>59</b>
6.1	DATASET .....	59
6.1.1	<i>Bases de Dados de Clip Arts</i> .....	60
6.1.2	<i>Esboços Utilizados para Teste</i> .....	61
6.2	MÉTODOS DE AVALIAÇÃO .....	62
6.3	TESTES PARA QUERY BY SKETCH .....	64
6.3.1	<i>Base de Dados com 100 Imagens</i> .....	64
6.3.2	<i>Base de Dados com 12955 Imagens</i> .....	67

6.3.3	<i>Desempenho dos Descritores para Esboço</i> .....	68
6.4	TESTES PARA QUERY BY IMAGE.....	70
6.4.1	<i>Base de Dados com 100 Imagens</i> .....	71
6.4.2	<i>Base de Dados com 12955 Imagens</i> .....	72
6.4.3	<i>Desempenho dos Descritores para Imagem</i> .....	75
6.5	TESTES DE USABILIDADE .....	76
6.5.1	<i>Método</i> .....	77
6.5.2	<i>Participantes</i> .....	78
6.5.3	<i>Resultados</i> .....	78
6.5.4	<i>Síntese</i> .....	82
6.6	CONCLUSÕES SOBRE OS RESULTADOS OBTIDOS.....	83
7	CONCLUSÃO E TRABALHO FUTURO.....	86
8	ANEXOS.....	94
8.1	QUESTIONÁRIO .....	94
8.2	RESULTADOS .....	100
8.3	GRÁFICOS SÍNTESE.....	105

# Lista de Figuras

Figura 1.1: Exemplo ilustrativo da pesquisa de clip arts na aplicação Microsoft Power Point, apresentando uma falha inerente à pesquisa por palavra-chave/ categoria.....	2
Figura 1.2: Imagem ilustrativa da forma como os humanos fazem esboços e reconhecem objetos. ....	4
Figura 1.3: Diagrama de blocos, representando a estrutura correspondente à solução apresentada.....	6
Figura 2.1: Exemplos de imagens clip art retiradas do site <i>openclipart.org</i> . ....	13
Figura 2.2: Exemplo ilustrativo, apresentando a diferença entre uma imagem do tipo raster e outra do tipo vetorial, quando aumentadas. ....	15
Figura 2.3: Imagem ilustrativa da aplicação Indagare. ....	16
Figura 2.4: Imagens ilustrativas da aplicação Retrieve. ....	17
Figura 2.5: Imagem ilustrativa da aplicação Electronic Cocktail Napkin. ....	18
Figura 2.6: Imagens ilustrativas da aplicação Google Goggles. ....	18
Figura 2.7: Imagens ilustrativas da aplicação Google Images. ....	19
Figura 2.8: Exemplos de pesquisa de imagens clip art nas aplicações Microsoft Word e Microsoft Power Point. ....	20
Figura 2.9: Imagens ilustrativas da aplicação Skitch. ....	23
Figura 2.10: Imagens ilustrativas da aplicação Doodledroid. ....	24
Figura 2.11: Imagens ilustrativas da aplicação Paperless. ....	24
Figura 2.12: Imagens ilustrativas da aplicação SketchBook Pro for Tablets. ....	25
Figura 3.1: Imagem descritiva do menu Help da interface de pesquisa desenvolvida. ....	29
Figura 3.2: Imagem descritiva do menu Color e respetiva paleta de cores disponibilizada na interface de pesquisa desenvolvida. ....	30
Figura 3.3: Imagem descritiva do menu Shapes e respetivas formas geométricas disponibilizadas na interface de pesquisa desenvolvida. ....	31
Figura 3.4: Imagem descritiva dos menus Draw e Erase e respetiva conjugação entre um e outro, referentes à interface de pesquisa desenvolvida. ....	31
Figura 3.5: Imagem descritiva do menu Clear da interface de pesquisa desenvolvida. ....	32

Figura 3.6: Imagem descritiva do menu Search na interface de pesquisa desenvolvida. ....	33
Figura 3.7: Imagem descritiva da utilização da aplicação no <i>tablet</i> , com a seleção de cores diferentes, várias formas geométricas, desenho livre e utilização da ferramenta para apagar pequenas áreas.....	34
Figura 3.8: A figura (a) mostra um esboço feito pelo utilizador e os resultados obtidos depois da pesquisa. A figura (b) mostra os resultados obtidos, quando o utilizador faz uma nova pesquisa através da seleção de uma das imagens apresentadas na barra lateral cinzenta. ....	36
Figura 3.9: A figura (a) apresenta uma mensagem a informar o utilizador que não tem conexão à rede e portanto não pode naquele momento efetuar pesquisas, podendo mesmo assim utilizar a aplicação para desenho. Por fim, a figura (b) mostra o estado da aplicação quando não há ligação à Internet, desabilitando a opção "Search". ....	37
Figura 3.10: Imagem demonstrativa da criação de um novo projeto Android no Eclipse.....	39
Figura 3.11: Imagem descritiva da forma como deve ser executada uma aplicação Android utilizando o Eclipse. Caso se queira que a aplicação seja executada num dispositivo móvel real, é necessário que este seja previamente ligado ao computador. Caso contrário, a aplicação pode ser executada num emulador no próprio Eclipse. ....	40
Figura 3.12: Imagem ilustrativa do ícone que é gerado pela aplicação Android e criado no Desktop do tablet. ....	40
Figura 4.1: Representação ilustrativa de um documento XML contendo os vetores de descritores associados a todas as imagens da base de dados. Neste caso específico, a ilustração representa os descritores extraídos com a utilização do descritor Momentos de Cor.....	45
Figura 4.2: Modelo representativo das principais operações realizadas no processo de recuperação de imagens clip art, para o sistema que foi desenvolvido. ....	46
Figura 4.3: Modelo genérico de comunicação entre cliente e servidor utilizando sockets.....	47
Figura 5.1: Representação da execução do pré-processamento nas imagens. A figura (a) representa a imagem raster original e a figura (b) representa a imagem depois de removido o espaço vazio à sua volta. A figura (c) mostra o redimensionamento aplicado à imagem e a figura (b) apresenta a imagem final a ser utilizada, depois de aplicado o filtro <i>blur</i> .....	50
Figura 5.2: Representação ilustrativa da divisão da imagem em blocos, quando é utilizado o descritor Momentos de Cor.....	52

Figura 5.3: Representação ilustrativa da transformação de uma imagem no formato raster (PNG) para o formato vetorial (SVG), retirando a componente cor.....	53
Figura 6.1: Imagem ilustrativa, com alguns exemplos de imagens clip art pertencentes a cada uma das categorias utilizadas (11 categorias no total).....	61
Figura 6.2: Imagem ilustrativa dos esboços que foram utilizados para testes. Estes esboços são referidos nos testes como "Esboço Detalhado".....	62
Figura 6.3: Imagem ilustrativa dos esboços que foram utilizados para testes. Estes esboços são refeidos nos testes como "Esboço Simples".....	62
Figura 6.4: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Detalhado" para a base de dados de 100 imagens.....	66
Figura 6.5: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Simples" para a base de dados de 100 imagens.....	66
Figura 6.6: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Detalhado" para a base de dados de 12955 imagens.....	68
Figura 6.7: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Simples" para a base de dados de 12955 imagens.....	68
Figura 6.8: Gráfico com a estatística global de desempenho das categorias, para os testes realizados com Query by Sketch.....	69
Figura 6.9: Gráfico com a estatística de desempenho global de cada um dos descritores, para os testes realizados com Query by Sketch. ....	70
Figura 6.10: Gráfico com os resultados globais dos testes realizados com um conjunto de imagens exemplo para a base de dados de 100 imagens. ....	72
Figura 6.11: Gráfico com os resultados globais dos testes realizados com um conjunto de imagens exemplo para a base de dados de 12955 imagens. ....	73
Figura 6.12: Imagem ilustrativa, apresentando formas muito diferentes em imagens clip art pertencendo à mesma categoria (e.g. "Tree"). ....	74
Figura 6.13: Imagens ilustrativas de clip arts pertencentes à categoria "Car", mostrando a disparidade entre formas e cores.....	74
Figura 6.14: Imagens ilustrativas de clip arts pertencentes à categoria "Magnifying Glass", mostrando a semelhança entre formas e cores.....	75
Figura 6.15: Gráfico com a estatística global de desempenho das categorias, para os testes realizados através de Query by Image, para as duas bases de dados.....	75
Figura 6.16: Gráfico com a estatística de desempenho global de cada um dos descritores, para os testes realizados com Query by Image. ....	76

Figura 6.17: Resultados obtidos com a questão “Considero que a interface tem as ferramentas essenciais de desenho”.....	79
Figura 6.18: Resultados obtidos com a questão “A aplicação é intuitiva e fácil de utilizar”..	79
Figura 6.19: Resultados obtidos com a questão “O aspeto visual da interface é apelativo”..	80
Figura 6.20: Como classifica os resultados obtidos pela query? (realizar o esboço de uma lâmpada).....	81
Figura 6.21: Como classifica os resultados obtidos pela query? (realizar o esboço de um carro) .....	81
Figura 6.22: Imagens clip art representando alguns casos onde a cor não é um fator discriminativo na imagem.....	84
Figura 8.1: Como classifica os resultados obtidos pela query? (realizar o esboço de uma lâmpada).....	105
Figura 8.2: Como classifica os resultados obtidos pela query? (refinar a pesquisa com a imagem de uma lâmpada) .....	105
Figura 8.3: Como classifica os resultados obtidos pela query? (realizar o esboço de um carro) .....	106
Figura 8.4: Como classifica os resultados obtidos pela query? (refinar a pesquisa com a imagem de uma lâmpada) .....	106
Figura 8.5: Considero que a interface tem as ferramentas essenciais de desenho. ....	106
Figura 8.6: A aplicação é intuitiva e fácil de utilizar. ....	107
Figura 8.7: O aspeto visual da interface é apelativo. ....	107
Figura 8.8: Eu utilizaria esta aplicação para complementar os meus conteúdos. ....	107
Figura 8.9: Na sua opinião, a apresentação de resultados na barra lateral, quando a query é o esboço, é adequada?.....	108
Figura 8.10: Na sua opinião, é adequada a apresentação de um novo ecrã, quando a pesquisa é feita através de um dos clip arts retornados? .....	108

# Lista de Tabelas

Tabela 6.1: Resultados obtidos com os testes utilizando Query by Sketch para a base de dados pequena (100 imagens). Os valores apresentados correspondem à medida R-Precision com $N = 3$ .....	65
Tabela 6.2: Resultados obtidos com os testes utilizando Query by Sketch para a base de dados grande (12955 imagens). Os valores apresentados correspondem à medida R-Precision com $N = 3$ .....	67
Tabela 6.3: Resultados obtidos com os testes utilizando Query by Image para a base de dados pequena (100 imagens). Os valores apresentados correspondem à medida R-Precision com $N = 3$ .....	71
Tabela 6.4: Resultados obtidos com os testes utilizando Query by Image para a base de dados grande (12955 imagens). Os valores apresentados correspondem à medida R-Precision com $N = 3$ .....	73
Tabela 8.1: Tabela exemplificativa de respostas às perguntas introdutórias do questionário. ....	100
Tabela 8.2: Tabela com exemplos de respostas dos utilizadores, ao realizar as tarefas 1 e 2. ....	100
Tabela 8.3: Tabela exemplificativa de respostas dos utilizadores à tarefa 3, questões b), c) e d). ....	101
Tabela 8.4: Tabela exemplificativa das respostas dos utilizadores às questões a) e b), pertencentes à tarefa 4.....	101
Tabela 8.5: Tabela exemplificativa das respostas dos utilizadores às questões para avaliação da interface.....	102
Tabela 8.6: Tabela exemplificativa das respostas dos utilizadores, às questões finais relacionadas com o teste de usabilidade realizado.....	103
Tabela 8.7: Exemplos de observações finais dadas pelos utilizadores no fim de realizar o teste de usabilidade com a aplicação.....	104



# 1 Introdução

*Clip arts* são coleções de imagens, geralmente utilizados para complementar documentos e outros conteúdos, para uso pessoal e/ ou comercial. No entanto, nem sempre é fácil para o utilizador encontrar uma imagem que se enquadre no contexto pretendido.

A grande quantidade de imagens *clip art*, seja na Internet ou em coleções adquiridas pelos utilizadores, provoca alguma dificuldade na organização das imagens por contextos, categorias, ou mesmo palavras-chave. Alguns estudos têm mostrado que a pesquisa baseada em palavras-chave (*query by keywords*), limita significativamente a expressividade dos utilizadores, prejudicando assim a eficácia da pesquisa [2]. Por outro lado, grande parte da meta-informação associada às imagens é produzida manualmente, o que torna este processo moroso, complexo e nem sempre eficaz [3]. Para além disso, os contextos ou descrições manuais são feitas de forma subjetiva, provocando muitas vezes um desequilíbrio na organização das imagens. Uma solução mais adequada, deve ter em conta informação extraída automaticamente do conteúdo das imagens [1]. Neste contexto, tem vindo a ser desenvolvido algum trabalho relativamente à pesquisa de imagens, através da utilização de imagens como exemplo (*query by example*) [4], [5].

O objetivo desta dissertação, consiste em construir diferentes estratégias de pesquisa, de forma a solucionar algumas das dificuldades que os utilizadores sentem quando procuram *clip arts*. Uma vez que a associação de contextos ou categorias nem sempre é um processo simples e eficaz para os utilizadores, é importante analisar diferentes abordagens, que dêem liberdade suficiente aos utilizadores, não os restringindo a descrições textuais. No decorrer deste estudo, pretende-se desenvolver uma interface para *tablets*, que permita aos utilizadores pesquisar *clip arts* através dos seus próprios esboços. Esta solução tem como objetivo oferecer uma ferramenta de pesquisa intuitiva, de modo a que não seja necessário perder demasiado tempo à procura de uma imagem *clip art*.

## 1.1 Motivação

A grande quantidade de imagens *clip art* torna a sua recuperação cada vez mais difícil, uma vez que estas podem ser contextualizadas de forma muito distinta, dependendo de quem as produz ou publica. Como tal, duas pessoas diferentes podem classificar a mesma imagem de formas distintas, seja na descrição ou na categoria onde a enquadram. Para além disso, existe ainda a questão temporal, tendo em conta que a classificação manual das imagens pode ser uma tarefa demorada.

Na Figura 1.1 é apresentado um exemplo, que mostra como pode ser difícil para um utilizador encontrar um *clip art*. No exemplo captado, utilizando a aplicação Microsoft Power Point, é possível perceber que apesar de existirem flores na base de dados, o utilizador não consegue obter nenhuma imagem desta categoria através da *keyword* "Flower", dado que as imagens de flores estão inseridas no contexto "Plants". Esta é uma das dificuldades que surge frequentemente no dia-a-dia dos utilizadores, que procuram por imagens *clip art*.



Figura 1.1: Exemplo ilustrativo da pesquisa de clip arts na aplicação Microsoft Power Point, apresentando uma falha inerente à pesquisa por palavra-chave/ categoria.

Neste sentido, existem algumas soluções que permitem ao utilizador efetuar pesquisas através de imagens exemplo (*query by example*) [5], analisando algumas propriedades no conteúdo da imagem [6]. Porém, esta solução nem sempre é eficiente, dado que o utilizador necessita de ter previamente uma imagem que represente aquilo que procura. Deste modo, propõe-se uma diferente abordagem, que consiste na utilização de esboços feitos pelo utilizador para realizar a pesquisa. A utilização de desenhos simples e representativos, pode ser a melhor forma de o utilizador analisar e explorar os conteúdos visuais que procura.

O desenvolvimento desta dissertação visa o estudo de diferentes estratégias, como a utilização de esboços para pesquisar imagens, utilizando uma interface de pesquisa num dispositivo móvel. Os dispositivos móveis (*smartphones* e *tablets*) oferecem cada vez mais, grande sensibilidade e capacidade no que diz respeito ao desenho, permitindo aos utilizadores realizar esboços à mão livre de forma ágil e intuitiva (em particular os dispositivos com interação *pen based*). Nesta interface, é oferecido ao utilizador um conjunto de ferramentas de desenho e pesquisa, que permitem ao utilizador pesquisar *clip arts* através dos seus esboços e posteriormente através de imagens exemplo. Com este tipo de sistema, o utilizador pode progressivamente alcançar uma aprendizagem que lhe permite saber como deve desenhar e que componentes são mais ou menos eficazes no desenho, interpretando os resultados obtidos em cada iteração.

## 1.2 Descrição e Contexto

Atualmente, existem várias tecnologias no que respeita à recuperação de imagens digitais. Porém, nem sempre os métodos existentes correspondem às necessidades dos utilizadores, que muitas vezes não sabem exatamente aquilo que procuram. É neste contexto que surgem os sistemas CBIR [4], que permitem organizar e recuperar imagens digitais através do seu conteúdo [6].

Como é referido no estudo apresentado em [7] os humanos não são, de uma forma geral, artistas fiéis. Como tal, é comum que as representações de objetos sejam feitas de forma demasiado simples ou demasiado exagerada. Dado que as pessoas representam o mesmo objeto, utilizando diversos graus de realismo e diferentes estilos de desenho, é necessário que seja utilizado um *dataset* de grandes dimensões, que se adeque aos mais variados tipos de esboços. Desta forma, é importante referir que o trabalho desenvolvido nesta área, apesar de ser já extenso, necessita ainda de ser refinado no que diz respeito à interpretação de esboços feitos pelo utilizador. A Figura 1.2 ilustra a forma como, de uma forma geral, os humanos desenham e interpretam os diferentes tipos de objetos [7].



Figura 1.2: Imagem ilustrativa da forma como os humanos fazem esboços e reconhecem objetos.

Os *clip arts* são imagens do tipo *raster* ou vetorial [8], o que permite utilizar várias tecnologias quando se pretende analisar o conteúdo das imagens. As imagens *raster* permitem extrair características como cor e textura [9] e as imagens vetoriais permitem extrair informação como topologia e geometria [10].

No desenvolvimento desta dissertação, pretende-se estudar métodos e soluções que permitam aos utilizadores pesquisar *clip arts* através de esboços, utilizando dispositivos móveis. Pretende-se desenvolver um sistema, onde a recuperação é totalmente baseada no conteúdo das imagens. Para além disso, pretende-se ainda analisar o desempenho da conjugação de dois métodos de pesquisa diferentes, em busca de uma solução melhor, que pode ajudar o utilizador a obter melhores resultados. Esta conjugação consiste na utilização de pesquisa através de esboço (*query by sketch* – mais facilidade para o utilizador) e através de imagem exemplo (*query by example* – melhor desempenho para o sistema). Em suma, o utilizador deve iniciar a sua pesquisa com a realização de um esboço com algumas formas e cores simples, que conduzam a alguns resultados idênticos ao que foi desenhado. Posteriormente, o utilizador pode utilizar uma das imagens retornadas como imagem exemplo, realizando uma segunda pesquisa de modo a obter melhores resultados, uma vez que uma imagem tende a ser mais completa do que um esboço.

Esta dissertação insere-se no âmbito do projeto CRUSH (*Clip art Retrieval using Sketches*), que está a ser desenvolvido pelo INESC ID (Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa), do Instituto Superior Técnico da Universidade Técnica de Lisboa e pela Faculdade de Ciências e Tecnologias da Universidade Nova de Lisboa. O projecto CRUSH tem como objectivo criar uma plataforma de pesquisa de *clip arts* através de imagens vectoriais e *raster* numa colecção de grande dimensão, tendo o apoio financeiro da Fundação para a Ciência e Tecnologia.

O trabalho apresentado em [36], foi realizado também no âmbito do projeto CRUSH, de modo a estudar, implementar e avaliar diferentes descritores e a combinação entre eles,

quando se pretende realizar a pesquisa de imagens *clip art* utilizando *query by example*. Para além disso, uma das propostas de trabalho futuro era o desenvolvimento de uma interface de pesquisa que permitisse pesquisar os *clip arts* através de esboços feitos pelo utilizador. Como tal, o trabalho desenvolvido no âmbito desta dissertação pretende dar continuidade a esta proposta, onde se pretende desenvolver uma interface de pesquisa que permita procurar *clip arts* através de esboços. Para além disso, nesta dissertação pretende-se não só utilizar o esboço como *input* de pesquisa, mas também imagens exemplo, avaliando as capacidades entre a composição de *query by sketch* e *query by example*.

### 1.3 Solução Apresentada

Na elaboração desta dissertação, pretende-se efetuar um estudo sobre a melhor forma de encontrar o *clip art* pretendido, através de esboços feitos pelo utilizador. Entende-se que em certos momentos, é mais favorável para um utilizador realizar um esboço do tipo de imagem que procura, em vez de se ver obrigado a pensar em contextos ou descrições textuais, ou em que categoria se enquadra esse tipo de imagem. Nem sempre é fácil para o utilizador descrever uma imagem, pois é provável que dois utilizadores diferentes contextualizem a mesma imagem de formas distintas. É neste sentido, que se pretende valorizar as diferentes perceções de cada utilizador, dando-lhe a possibilidade de construir progressivamente um conteúdo visual, que represente a imagem que procura. É previsível que em alguns casos, o utilizador se consiga exprimir melhor através do desenho e da sua perceção visual, do que através de palavras, contextos ou categorias. *"No entanto, a recuperação de resultados significativos na percepção pode ser difícil, pois os utilizadores geralmente desenharam esboços de uma forma abstrata, que são geometricamente distantes das fotografias ou modelos reais"* [7].

Em resumo, pretende-se construir uma interface de pesquisa direcionada à recuperação de imagens *clip art*, baseada em *query by sketch* e *query by example*. A utilização de *query by example*, utilizada após uma primeira pesquisa através de esboço, revela-se uma potencial melhoria na definição dos resultados. Esta aplicação utiliza uma arquitetura cliente-servidor, uma vez que o sistema é composto por dois componentes: a interface de pesquisa para *clip arts* (cliente) e o sistema de recuperação de imagem (servidor). No *tablet* será implementada

a aplicação cliente, que é utilizada como interface de pesquisa, oferecendo todas as ferramentas necessárias para a realização do esboço e da pesquisa. Num computador externo, é executado o servidor, que realiza todo o processamento de imagem e devolve os resultados à interface de pesquisa por ordem de semelhança. Dado que a interação com o utilizador é um fator relevante, no desenvolvimento de aplicações para dispositivos móveis, é necessário ter em conta metodologias centradas no utilizador [11], [12].

Na Figura 1.3, é apresentada uma visão geral da solução proposta, que demonstra os seus principais componentes: Cliente – interface de pesquisa instalada num *tablet*; Servidor – sistema de recuperação de imagem. A interface de pesquisa desenvolvida é direcionada para dispositivos com sistema operativo Android.

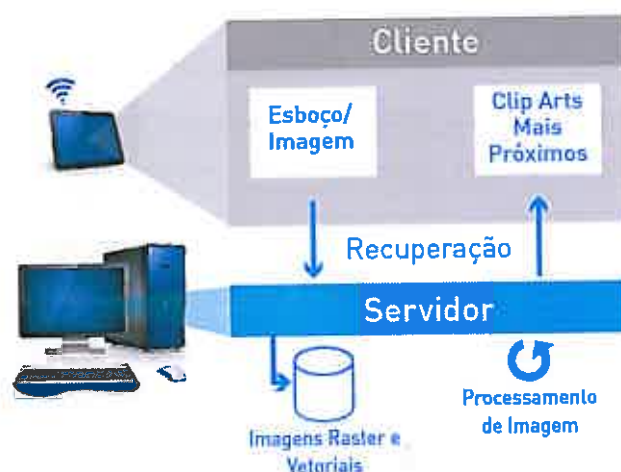


Figura 1.3: Diagrama de blocos, representando a estrutura correspondente à solução apresentada.

### 1.3.1 Recuperação de Imagem baseada no Conteúdo

A recuperação de imagens baseada no seu conteúdo, surgiu de forma a solucionar algumas dificuldades sentidas pelos utilizadores. Quando falamos em extrair informação do conteúdo de imagens, esta extração pode ser feita de forma manual e subjetiva por humanos (análise semântica) ou de forma computacional e objetiva através de extração de informação na imagem (análise sensorial) [13]. Define-se como análise semântica, a falta de correspondência entre a informação que se pode extrair dos dados visuais e da interpretação de um humano em determinado momento. Já a análise sensorial, é a diferença entre o objeto

no mundo real e as informações numa descrição computacional, derivada de uma gravação dessa cena.

A recuperação de imagens baseada no conteúdo, centra-se na comparação entre o vetor de características da interrogação (imagem ou esboço) e representações idênticas das imagens da coleção [4]. Contudo, devido ao problema da falha semântica [14], foram propostas soluções semi-automáticas baseadas na técnica de retroação de relevância (o utilizador atribui uma relevância a cada imagem retornada pelo sistema) [15] e soluções automáticas baseadas em conceitos semânticos [1].

### 1.3.2 Principais Contribuições Previstas

O desenvolvimento desta dissertação consiste em diversas etapas, com o foco na construção de uma aplicação que permite pesquisar *clip arts* através de esboços. Para tal, são efetuados alguns estudos em áreas relativamente distintas, contribuindo cada uma delas para a construção de toda a estrutura necessária para o funcionamento total do sistema. Deste modo, são apresentadas de seguida as principais contribuições previstas:

- Implementar uma interface dedicada à plataforma *tablet* Android. Esta interface deve oferecer as ferramentas necessárias para realizar esboços, de modo a efetuar a pesquisa de *clip arts*. A utilização do *tablet* neste contexto é relevante, dado a adesão cada vez maior a estes dispositivos, bem como o seu bom desempenho. Para além da forma relativamente fácil com que se desenha nestes dispositivos, é ainda importante oferecer portabilidade, cada vez mais procurada pelos utilizadores.
- Extrair características do conteúdo das imagens, de forma a criar vetores de descritores, que representem cada uma das imagens conforme o contexto pretendido (cor e/ ou forma). São conjugadas várias técnicas de processamento de imagens (*raster* e *vetoriais*), para extrair diferentes tipos de informação. Deste modo, é ainda essencial realizar uma pequena análise sobre alguns dos descritores que podem produzir melhores resultados, dentro do âmbito pretendido. O foco centra-se na análise de descritores de cor e forma, bem como a combinação entre ambos.
- Aplicar algumas métricas (Euclidiana e Manhattan), de modo a calcular as distâncias entre os vetores de descritores referentes a cada imagem. Estas métricas permitem calcular distâncias entre os vetores correspondentes às imagens e obter a respetiva

ordem de proximidade. Para além disso, são conjugadas características obtidas com imagens *raster* e vetoriais, sendo interessante analisar os resultados combinando as métricas acima referidas (*rankings* combinados de várias características).

- Desenvolvimento de um sistema de recuperação de informação, que utiliza uma métrica combinada baseada na utilização de descritores diferentes, com imagens de diferentes tipos (*raster* e vetoriais).
- Explorar diversas técnicas, utilizadas no desenvolvimento de aplicações direcionadas para dispositivos móveis. Esta é também uma etapa importante, dada a relevância de desenvolver uma interface centrada no utilizador [11]. Por fim, serão feitos testes de usabilidade, de modo a perceber se existem e quais são as dificuldades e falhas no sistema detetados pelos utilizadores ao utilizar a interface.

Neste contexto, é ainda importante referir que não é foco principal nem objetivo final desta dissertação o estudo de descritores de forma pormenorizada. Será necessário realizar um estudo futuro, de modo a perceber quais as tecnologias mais adequadas para extrair conteúdo das imagens, quando se pretende comparar esboços relativamente simples com imagens mais complexas.

No decorrer do desenvolvimento desta dissertação, resultou ainda a publicação de um artigo intitulado “*Clip art retrieval using a sketch Tablet application*” [42], para a conferência que decorreu no ISEL (Instituto Superior de Engenharia de Lisboa) com o nome *Conference on Electronics, Telecommunications and Computers – CETC 2013*. A versão completa desta publicação, está submetida para a publicação na revista *Elsevier’s PROCEDIA TECHNOLOGY*.

## 1.4 Estrutura da Dissertação

Esta dissertação é composta por sete capítulos principais. São eles:

- **Capítulo 1:** Pretende introduzir o problema abordado nesta dissertação, bem como a solução proposta e as principais motivações e contexto onde esta problemática se



insere. Neste capítulo, são ainda apresentadas as principais contribuições previstas durante o desenvolvimento desta dissertação.

- **Capítulo 2:** Apresenta algumas soluções e tecnologias desenvolvidas no âmbito do problema apresentado, de forma a obter algum *feedback* sobre quais as melhores e piores soluções a ter em conta. São apresentadas também algumas características inerentes às imagens em estudo (*clip arts*), permitindo analisar quais são as melhores tecnologias aplicadas a cada uma delas, tendo em conta as suas vantagens e desvantagens.
- **Capítulo 3:** Descreve o processo de desenvolvimento da interface de pesquisa, direcionada para dispositivos móveis. Neste capítulo, são apresentadas as motivações e decisões tomadas ao longo do desenvolvimento, tendo em conta conceitos e heurísticas associadas ao desenvolvimento de interfaces.
- **Capítulo 4:** Apresenta o sistema de recuperação de imagem desenvolvido, no âmbito de complementar a interface de pesquisa para *tablet*. Este sistema recebe uma imagem *query* (esboço ou *clip art*), realiza todo o processamento relativo às imagens e devolve os resultados mais próximos à interface de pesquisa. Este capítulo apresenta a visão geral de todo o sistema desenvolvido.
- **Capítulo 5:** Neste capítulo, é feita uma descrição relativamente detalhada sobre a extração de características do conteúdo das imagens, bem como as tecnologias utilizadas. São ainda detalhados alguns procedimentos efetuados, deste o pré-processamento que é necessário realizar antes da extração de informação, a outros pormenores necessários para a aplicação de descritores para extrair as características.
- **Capítulo 6:** São apresentados os resultados experimentais obtidos, através da realização de vários testes, de modo a perceber se o sistema apresenta resultados satisfatórios, correspondendo com os objetivos propostos. A realização de testes engloba testes de desempenho, que analisam apenas os resultados obtidos para cada tipo de *query* dependendo de bases de dados maiores ou mais pequenas, e testes de usabilidade, que pretendem analisar o impacto que a aplicação tem no mundo real, bem como perceber se os utilizadores conseguem interagir de forma familiar com a interface de pesquisa desenvolvida.
- **Capítulo 7:** Por fim, neste capítulo são apresentadas as conclusões gerais sobre a realização do trabalho desenvolvido. É feita uma análise geral ao processo de desenvolvimento e ao sucesso ou insucesso dos resultados obtidos. É feita ainda uma

proposta sobre a realização de trabalho futuro, que poderá melhorar os resultados aqui apresentados, bem como criar novas propostas de aplicações baseadas na que aqui é apresentada.

# 2 Trabalho Relacionado

Na elaboração desta dissertação, foram exploradas diversas tecnologias e aplicações que têm vindo a ser desenvolvidas no âmbito do tema aqui apresentado. Neste capítulo, são apresentadas algumas das tecnologias e trabalhos explorados, com o objetivo de contextualizar e enquadrar de forma mais clara as características inerentes a esta dissertação.

## 2.1 Introdução

Nas últimas décadas, têm vindo a ser desenvolvidos estudos e aplicações que incidem na área de recuperação de imagem, seja através de *query by example* ou de *query by sketch*. Com estes estudos, é possível perceber as dificuldades inerentes à utilização de sistemas CBIR, onde são necessárias diversas abordagens para afinar toda a informação que pode ser extraída de uma imagem.

Em 1999, surgiu um sistema para recuperação de desenhos 2D, apresentado por Park e Um [16]. Este sistema baseia-se na interpretação de formas dominantes, onde os objetos são decompostos recursivamente até obter uma forma dominante, componentes auxiliares e as suas relações espaciais. Os elementos visuais presentes nos desenhos são classificados através de geometrias simples (triângulos, círculos, quadrados, entre outros). No entanto, este sistema baseado em formas simples, torna difícil a utilização de bases de dados com grandes dimensões, dado o elevado processamento que é necessário.

Mascio et al. [17], apresentam um sistema para a recuperação de desenhos através de esboços, utilizando técnicas CBIR. É utilizada uma interface gráfica que permite recuperar imagens utilizando *query by sketch* e *query by example* que recupera imagens vetoriais, modelando-as e associando-as com descritores que representam as características visuais da imagem, invariantes a transformações e outras variações. Neste sistema, são ainda utilizadas

diferentes métricas de avaliação de similaridade, de modo a suportar diversos requisitos e domínios de aplicação. Apesar de serem recuperados desenhos vetoriais, estes são convertidos primeiro para *raster*, por forma a aplicar um conjunto de técnicas CBIR. A interface gráfica oferece ferramentas, que ajudam na selecção de critérios e parâmetros necessários para ajustar o sistema para um domínio de aplicação específica.

A recuperação de imagens baseada em conteúdo (CBIR), tem vindo a ser explorada ao longo dos últimos anos, no que diz respeito ao estudo de pesquisa de baixo nível. No entanto, os sistemas CBIR necessitam de imagens como base de consulta, que nem sempre são suficientes para expressar as necessidades de informação de conteúdo dos utilizadores. Deste modo, têm surgido várias abordagens [18], [19], [20], [21], que pretendem solucionar estas e outras questões, proporcionando aos utilizadores a capacidade de desenhar a sua própria imagem de pesquisa (esboço), executando aquela que é conhecida como *query by sketch*. Neste sentido, têm vindo a ser desenvolvidos diversos sistemas, que pretendem melhorar a experiência de usabilidade dos utilizadores, onde para além de ser essencial obter um bom conjunto de resultados, é também importante oferecer ferramentas melhoradas, permitindo que os utilizadores se sintam à vontade na realização de pesquisas através de esboços.

No estudo apresentado por Kreuzer et al. em [22], é explorado o desenvolvimento de uma interface *query by sketch*, através da utilização de papel normal e uma caneta digital de tecnologia interativa. Esta abordagem pretende desenvolver uma interface de pesquisa amigável para o utilizador, uma vez que desenhar com um rato de um computador ou mesmo com um adaptador de desenho (*tablet* de desenho para computador) se torna ineficiente. Para além de o desenho se tornar altamente impreciso utilizando um rato de computador, no caso dos adaptadores de desenho há também a necessidade constante de desviar o olhar da mão, para analisar o resultado do desenho no ecrã do computador. A tecnologia abordada neste estudo pretende utilizar um simples papel e uma caneta com tecnologia *Bluetooth* e outras, realizando a transferência de *streaming* de vetores que representam os traços de caneta desenhados no papel, para um sistema externo que realiza tarefas CBIR.

Ao contrário de outras abordagens, que consideram apenas informação de contorno extraído nos esboços, o estudo apresentado por Giangreco et al. em [23] pretende apresentar um sistema que explora em conjunto informação de contorno e cor, através de uma interface de desenho amigável. Este estudo refere que os descritores de cor geralmente produzem melhores resultados do que descritores baseados em contornos. É apresentada ainda uma

solução que produz uma melhoria significativa, para esboços utilizados como *query*, que consiste na composição entre os descritores de cor e contorno.

De uma forma geral, são muitas as abordagens que se podem ter em conta no que diz respeito à recuperação de imagem baseada no conteúdo, bem como interfaces de pesquisa baseadas em esboço.

## 2.2 Características das Imagens Clip Art

Os *clip arts* já não são, atualmente, desenhos pouco complexos (ver Figura 2.1). Podem ser imagens do tipo *raster* ou vetorial, o que permite utilizar várias abordagens quando se trata de processamento de imagem. Em seguida, são apresentadas algumas características inerentes a cada tipo de imagem, bem como os descritores que melhor se adequam a cada caso.



Figura 2.1: Exemplos de imagens clip art retiradas do site *openclipart.org*.

### 2.2.1 Imagens Raster

As imagens *raster* são constituídas por uma grelha de pixels, onde cada pixel representa um ponto de cor individual, compondo assim uma imagem [8]. Frequentemente, este tipo de imagens surge na representação de fotografias, que são muitas vezes transferidas para computadores através de câmeras digitais, scanners ou mesmo telemóveis, entre outros tipos de dispositivos. Ao contrário das imagens vetoriais, as imagens *raster* não são independentes da resolução, que representa o número de pixels numa imagem por área (e.g., polegadas), normalmente referida como *dpi* (pontos por polegada) ou *ppi* (pixels por polegada). Dado que as imagens *raster* são dependentes da resolução, em algumas condições aumentar ou diminuir a sua dimensão pode danificar o grau de qualidade da imagem (teorema da amostragem) [24].

Este tipo de imagens permite extrair características como cor e textura, onde é possível modelar a variação contínua de uma dada característica. Dado que nesta dissertação se pretende essencialmente utilizar esboços, que na sua maioria são feitos de forma simplificada pelos utilizadores, considerou-se pouco relevante implementar um descritor relativo à textura. Deste modo, para este tipo de imagens é utilizado o descritor de cor Momentos de Cor.

Características das imagens *raster*:

- Pixels em grelha;
- Dependente da resolução (ver Figura 2.2);
- Redimensionar pode reduzir a qualidade.

### 2.2.2 Imagens Vetoriais

As imagens vetoriais podem ser descritas como pontos, linhas ou polígonos, que são representados como coleções de coordenadas  $x$  e  $y$ , com atributos editáveis como cor e contorno. Estes atributos podem ser alterados, sem danificar o objeto principal [8]. Este tipo de imagens são descritas utilizando dados geométricos em vez de pixels fixos, podendo ser redimensionadas, mantendo sempre a mesma qualidade [24]. O modelo vetorial torna-se útil para representar características discretas, apesar de não ser tão útil na descrição de continuidade na variação de uma característica.

As imagens do tipo vetorial permitem extrair características como geometria e topologia, ou seja, representam as formas e as suas relações identificadas no conteúdo das imagens. Nesta dissertação, para descrever e analisar este tipo de conteúdo nas imagens, é utilizado o descritor de topologia e geometria Topogeo.

Características das imagens vetoriais:

- Escalável;
- Independente da resolução (ver Figura 2.2).

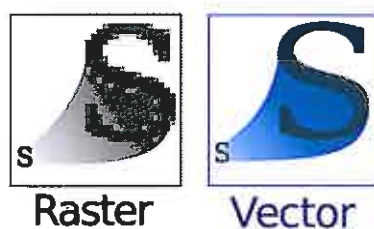


Figura 2.2: Exemplo ilustrativo, apresentando a diferença entre uma imagem do tipo raster e outra do tipo vetorial, quando aumentadas.

## 2.3 Aplicações Relacionadas

Nesta secção, são apresentadas algumas aplicações e tecnologias que, de alguma forma, estão relacionadas com a área de trabalho que aqui se apresenta. Este estudo foi conduzido com o objetivo de perceber quais as principais áreas de trabalho já exploradas e desenvolvidas, no que diz respeito à recuperação e organização de imagens digitais através do seu conteúdo.

### 2.3.1 Pesquisa de Clip Arts Combinando Imagens Raster e Vectoriais

Este trabalho foi desenvolvido no âmbito de uma dissertação inserida no projeto CRUSH, onde se pretendeu desenvolver um sistema de recuperação de *clip arts* baseado nas suas características (CBIR). Neste trabalho, foram implementadas diversas técnicas de simplificação de imagens, convertendo do formato *raster* para vetorial, que permitem posteriormente extrair características de imagens *raster* (cor, textura e pontos de interesse) e de imagens vetoriais (geometria). Neste âmbito, um dos principais objetivos consistiu em recuperar imagens *clip art* combinando técnicas de pesquisa de imagens *raster* com imagens vetoriais.

### 2.3.2 Indagare

O projeto Indagare<sup>1</sup> (ver Figura 2.3), foi desenvolvido com o âmbito de criar um motor de busca, que permite a pesquisa de desenhos vetoriais (*clip arts*), através da utilização de esboços ou imagens como *query*, fornecidas pelos próprios utilizadores. O principal objetivo deste projeto, é efetuar pesquisas com base no conteúdo dos desenhos (geometria e arranjo espacial), deixando de parte a necessidade de gerar toda a meta-informação manualmente. São armazenados numa base de dados uma série de desenhos disponíveis na Web, para posterior recuperação de imagens.



Figura 2.3: Imagem ilustrativa da aplicação Indagare.

### 2.3.3 Retrievr

O Retrievr<sup>2</sup> (ver Figura 2.4) baseia-se num projeto experimental, que permite pesquisar e explorar uma coleção de imagens presentes no Flickr, através de um esboço feito pelo utilizador. Este projeto é baseado numa pesquisa conduzida por Chuck Jacobs, Adam Finkelstein e David Salesin, na Universidade de Washington (*Fast Multiresolution Image Querying* (1995)). Numa abordagem mais recente, é também permitido ao utilizador efetuar a sua pesquisa através do *upload* de uma imagem ou mesmo através de um URL. A ideia é procurar as imagens que sejam mais idênticas ao objeto fornecido, seja através de um esboço ou uma imagem (*query by sketch* e *query by example*).

<sup>1</sup> <http://cgm.dei.ist.utl.pt/propostas/tfcs0607/indagare/index.html> (acedido a 12 de Novembro de 2012)

<sup>2</sup> <http://labs.systemone.at/retrievr> (acedido a 12 de Novembro de 2012)





Figura 2.4: Imagens ilustrativas da aplicação Retrieve.

### 2.3.4 Electronic Cocktail Napkin

O Electronic Cocktail Napkin<sup>3</sup> (ver Figura 2.5), tem como objetivo criar um ambiente baseado em computador para o desenho através de caneta tátil (*pen based*), para o desenho à mão livre de esboços e diagramas. Pretende-se com este trabalho apoiar o desenho computacional, de forma a transformar diagramas em representações mais formais e estruturadas esquematicamente, contando sempre com o apoio do computador. Os utilizadores desenhavam esboços dos edifícios pretendidos, que serão comparados com os diagramas (produzidos manualmente pelos utilizadores) armazenados numa base de dados. Um problema visível deste sistema é o seu desempenho para bases de dados de grandes dimensões, dado que os diagramas representativos dos edifícios para a recuperação têm de ser produzidos manualmente [25]. Se um gráfico é reconhecido, são tidas em conta apenas as restrições de conectividade entre os nós e os arcos, ignorando as relações de disposição espacial.

<sup>3</sup> <http://depts.washington.edu/napkin/> (acedido a 12 de Novembro de 2012)

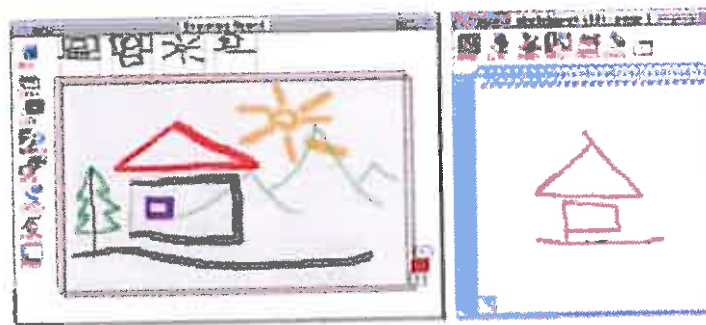


Figura 2.5: Imagem ilustrativa da aplicação Electronic Cocktail Napkin.

### 2.3.5 Google Goggles

Como se pode ver na Figura 2.6, esta é uma ferramenta pertencente à Google e destina-se aos dispositivos móveis (Android e iPhone). Esta aplicação consiste na pesquisa de imagens, através de fotografias tiradas no próprio momento ou imagens já armazenadas no dispositivo, funcionando também para códigos de barras e códigos QR (*Quick Response*). Se o Goggles<sup>4</sup> encontrar a imagem na sua base de dados, deve apresentar informação útil e também imagens idênticas ou iguais à referenciada.



Figura 2.6: Imagens ilustrativas da aplicação Google Goggles.

<sup>4</sup> <http://www.google.com/mobile/goggles/#text> (acedido a 7 de Dezembro de 2012)

### 2.3.6 Google Images

O Google Images<sup>5</sup> (ver Figura 2.7) é uma ferramenta bastante utilizada e conhecida atualmente, devido à sua enorme base de dados (imagens em páginas na Web) e capacidade de resposta. O objetivo é permitir que os utilizadores efetuem a pesquisa de imagens na Web, utilizando imagens exemplo, que podem ser carregadas através de um URL indicando uma imagem na Web, ou através do *upload* de uma imagem armazenada no computador/dispositivo do utilizador.



Figura 2.7: Imagens ilustrativas da aplicação Google Images.

### 2.3.7 Microsoft Word e Microsoft Power Point

Estas ferramentas de edição de documentos e conteúdos (ver Figura 2.8), bem como outras existentes, há muito que permitem aos utilizadores complementar os seus conteúdos através da utilização de *clip arts*. Apesar da grande oferta de imagens que estas ferramentas disponibilizam, nem sempre é fácil para o utilizador encontrar aquilo que procura, como já foi mostrado anteriormente. A pesquisa de *clip arts* pode ser efetuada de diversas formas, nestas ferramentas: palavra-chave, categoria ou outras.

<sup>5</sup> <http://www.google.com/imghp?hl=pt-PT&tab=ii>

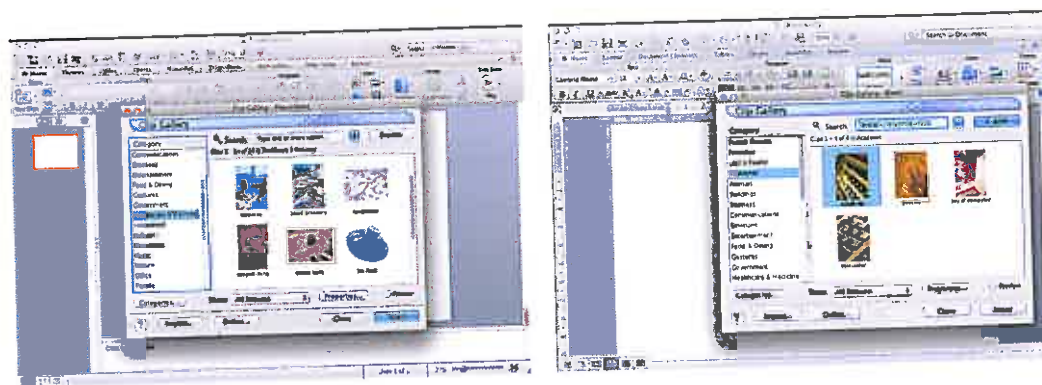


Figura 2.8: Exemplos de pesquisa de imagens clip art nas aplicações Microsoft Word e Microsoft Power Point.

## 2.4 Interfaces de Query

Nesta secção, é feita uma pequena abordagem sobre os vários tipos de *query* que podem ser utilizados atualmente, permitindo assim efetuar pesquisas em diversas áreas. Nesta dissertação, interessa-nos particularmente a pesquisa de imagens *clip art*. Esta análise ajuda-nos a perceber, de entre os métodos existentes, quais apresentam melhores resultados nas diversas situações, tais como pesquisas de documentos, imagens, entre outros. Deste modo, torna-se possível enquadrar de forma mais precisa quais os métodos que se adaptam melhor, nos casos em que o objetivo é a recuperação de imagens. Em seguida, são apresentados os principais métodos de pesquisa utilizados, salientando em cada um deles algumas vantagens e desvantagens.

### 2.4.1 Pesquisa por Texto ou Categoria (Query by Keywords)

Entende-se por *query by keywords*, a pesquisa que se efetua através da utilização de descrições textuais, palavras-chave ou contextos. No estudo realizado em [1], é referido que as consultas baseadas em palavras-chave podem limitar significativamente a expressividade dos utilizadores, prejudicando assim a eficácia da pesquisa. Muitas vezes, enquadrar adequadamente o tema de uma imagem pode tornar-se problemático. Este problema agrava-se ainda com a internacionalização de algumas aplicações, que obrigam os utilizadores a

efetuar pesquisas numa linguagem da qual não são nativos. Como consequência, os utilizadores acabam por desgastar-se e perder demasiado tempo até encontrar um conjunto correto de palavras-chave, atingido muitas vezes através de tentativa e erro [26].

Exemplos da utilização deste tipo de *query* são: Google Images, Microsoft Word.

#### 2.4.2 Pesquisa por Imagem (Query by Example)

Na pesquisa *query by example*, pretende-se que o utilizador forneça uma imagem como exemplo, de modo a que a recuperação seja feita baseada no seu conteúdo, utilizando sistemas CBIR. Nos cenários sem restrições, o tema CBIR enfrenta uma série de desafios, incluindo a falha sensorial (diferença entre a forma como os utilizadores percebem e descrevem um alvo e como o alvo é representado internamente no sistema de pesquisa) e a diferença semântica (dificuldade em extrair informações precisas a partir de imagens) [4]. Tal como é descrito em [5], a *“recuperação de imagens baseada em conteúdo é um problema que envolve muitas questões fundamentais da visão de computador, tais como a forma de representar imagens ou para avaliar distâncias entre as mesmas”*. Dado que este tipo de pesquisa é baseado no conteúdo, é comum existir a necessidade de um pré-processamento, que tem como objetivo decompor cada imagem num ou mais vetores de características. Posteriormente, a recuperação é feita através da comparação entre imagens (através dos vetores de características), realizada através da utilização de uma medida entre vetores, denominada como métrica.

A pesquisa baseada em imagens exemplo tem também algumas falhas, uma vez que para realizar a pesquisa o utilizador tem de possuir previamente uma imagem que sirva de exemplo. Isto significa que, antes de realizar a pesquisa, é necessário proceder a uma “pré-pesquisa” para encontrar imagens exemplo, o que não ajuda a resolver o problema aqui apresentado.

Exemplos da utilização deste tipo de *query* são: Google Goggles, Retrievr.

### 2.4.3 Pesquisa por Esboço (Query by Sketch)

Esta é uma solução mais recente, que permite aos utilizadores pesquisar imagens através da realização de esboços à mão livre. Pretende-se com esta solução facilitar o trabalho dos utilizadores, de modo a que estes consigam expressar mais facilmente a sua intenção sobre o tipo de imagem que procuram. Apesar de esta forma de pesquisa melhorar muitos aspetos, tem também alguns problemas associados, tais como a possibilidade de o utilizador não se sentir à vontade para desenhar, ou não saber bem como posicionar o seu contexto num desenho à mão livre. Como é descrito em [10], testes com utilizadores demonstraram que quando estes pesquisam por *clip arts* através de esboço, há uma tendência para desenharem um pequeno número de formas, especificando apenas a geometria e deixando de parte a topologia.

Na pesquisa através de esboço, é também necessário extrair informação sobre o conteúdo das imagens, que é utilizado na altura da recuperação. A extração de conteúdo visual a partir de imagens é dividido em duas partes: processamento de imagem e construção de *features*. Estas *features* devem ser comparadas quando se pretende comparar diversas imagens.

Exemplos da utilização deste tipo de *query* são: Indagare e Retrieve.

## 2.5 Ferramentas de Desenho em Dispositivos

### Móveis

Os recentes avanços nas novas tecnologias, permitem hoje em dia que um simples *smartphone* com ecrã tátil seja utilizado como uma ferramenta de desenho, que pode ser levada para qualquer lugar. Este conceito aplica-se também e cada vez mais aos dispositivos denominados por *tablets*, que tal como o telemóvel se estão a tornar cada vez mais essenciais no dia a dia de profissionais ou mesmo de utilizadores comuns. De seguida são apresentadas algumas ferramentas, que permitem aos utilizadores realizar esboços e desenhos através dos seus dispositivos móveis.

### 2.5.1 Skitch

O Skitch<sup>6</sup> é uma extensão do Evernote<sup>7</sup>, um aplicativo útil para anotar apontamentos. Quando a aplicação é aberta, o utilizador tem três opções iniciais à sua disposição (ver Figura 2.9): a primeira ativa a câmara do dispositivo móvel e permite capturar uma fotografia na qual o utilizador pode desenhar por cima; a segunda permite ao utilizador procurar uma imagem armazenada no dispositivo e mais uma vez desenhar sobre essa imagem; a terceira permite ao utilizador desenhar livremente sobre um fundo branco.



Figura 2.9: Imagens ilustrativas da aplicação Skitch.

### 2.5.2 Doodledroid

O Doodledroid<sup>8</sup> é uma ferramenta que permite a criação de desenhos livres e com alguma complexidade e pormenor, através da utilização de várias ferramentas disponíveis (caneta, lápis, pastel, entre outros). Permite ainda importar, partilhar, guardar e abrir imagens ou desenhos (ver Figura 2.10).

<sup>6</sup> <http://evernote.com/skitch/> (acedido a 17 de Janeiro de 2013)

<sup>7</sup> <https://evernote.com/intl/pt/> (acedido a 17 de Janeiro de 2013)

<sup>8</sup> <http://www.littlesunsoftware.com> (acedido a 17 de Janeiro de 2013)





Figura 2.10: Imagens ilustrativas da aplicação Doodledroid.

### 2.5.3 Paperless

O Paperless<sup>9</sup> foi desenhado especificamente para ser utilizado em *tablets*, que oferecem maiores dimensões do que os telemóveis e no geral melhores condições no que diz respeito ao desenho. Com esta aplicação instalada no *tablet* o utilizador consegue desenhar com tanta sensibilidade como se estivesse perante um papel normal. É permitido ao utilizador pintar com aguarelas, lápis, pena, pincéis, entre outros, com resposta imediata da interface sem atraso em nenhum dos seus traços (ver Figura 2.11).

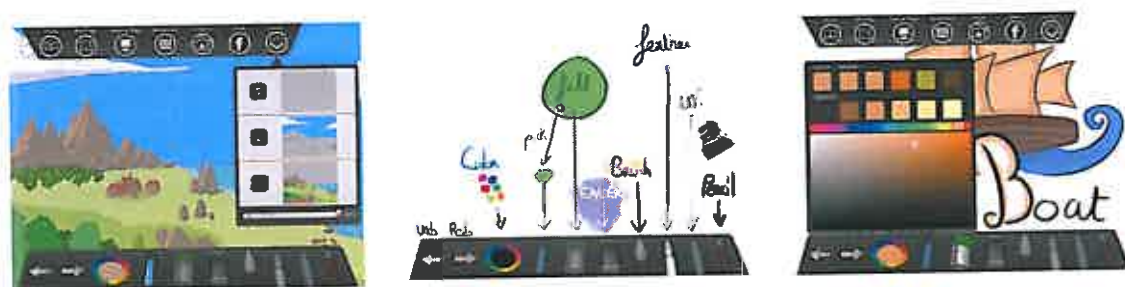


Figura 2.11: Imagens ilustrativas da aplicação Paperless.

### 2.5.4 SketchBook Pro for Tablets

O SketchBook Pro<sup>10</sup> é uma aplicação desenhada também especialmente para ser utilizada em *tablets* Android, com ecrãs de 7 polegadas ou acima. Num âmbito muito mais profissional,

<sup>9</sup> <http://www.draw-app.com/android/> (acedido a 17 de Janeiro de 2013)

<sup>10</sup> <http://www.sketchbook.com> (acedido a 17 de Janeiro de 2013)



esta aplicação disponibiliza um conjunto completo e sofisticado de ferramentas, apesar da sua interface simples e intuitiva (ver Figura 2.12).



**Figura 2.12:** Imagens ilustrativas da aplicação SketchBook Pro for Tablets.

# **3**

## **Interface de Pesquisa para Tablet**

Neste capítulo, são descritas as metodologias e tecnologias que foram utilizadas para o desenvolvimento da aplicação e respetiva interface, direccionadas para a utilização em dispositivos móveis com a plataforma Android.

Em primeiro lugar, é feita uma nota introdutória que descreve os conceitos e heurísticas relevantes que devemos ter em atenção quando desenvolvemos interfaces para dispositivos móveis. Em segundo lugar, são descritas de forma detalhada as decisões que foram tomadas ao longo do desenvolvimento, bem como outras opções que foram analisadas para permitir melhorar a interação e a visualização de resultados. De seguida, são apresentadas as metodologias e tecnologias que foram utilizadas para o desenvolvimento da interface, tais como bibliotecas utilizadas e opções relativas à plataforma móvel utilizada. Por fim, são enumeradas algumas questões de usabilidade, e é explicado a sua relevância na concepção, desenvolvimento e avaliação da aplicação.

### **3.1 Nota Introdutória**

No desenvolvimento de interfaces para dispositivos móveis é importante ter em consideração algumas regras básicas, que permitem proporcionar ao utilizador uma boa experiência na utilização de qualquer aplicação. Como tal, um conceito relevante para desenvolver este tipo de conteúdos, são as heurísticas, que apresentam regras, simplificações e aproximações que permitem afinar e restringir domínios muito vastos e por vezes pouco compreendidos. Para o desenvolvimento da interface proposta nesta dissertação, foram consideradas as heurísticas de

Nielsen [27], que podem ser utilizadas em diversos contextos, bem como outros conceitos utilizados no desenvolvimento de sistemas interativos para dispositivos móveis [11], [28], [29], [30].

Como é descrito no estudo apresentado em [7], é possível identificar o significado semântico que os utilizadores pretendem dar aos seus esboços, mesmo que estes sejam geometricamente distantes da sua forma no mundo real. No entanto, se não for possível identificar a semântica associada, é natural que os motores de recuperação de imagem baseada em esboço tenham problemas. Este estudo propõe a utilização de uma extensão à recuperação baseada em esboço, onde é proposta a utilização de pesquisa tradicional baseada em palavras-chave, que deve complementar e melhorar a pesquisa baseada em esboço.

A solução proposta nesta dissertação, apresenta uma abordagem idêntica à referenciada anteriormente, uma vez que o utilizador deve construir um primeiro esboço para efetuar a sua pesquisa, podendo melhorar a qualidade dos resultados progressivamente. Para potenciar os resultados obtidos, uma vez que o utilizador visualiza o esboço e as imagens retornadas no mesmo ecrã, é possível acrescentar de forma progressiva pequenos elementos ao seu esboço, tentando aproximar-se ao tipo de imagens retornadas. Outra solução, é fazer uma nova *query* clicando sobre uma das imagens retornadas ao esboço. Cada um destes métodos, vai despoletar um novo pedido ao sistema de recuperação de imagem e consequentemente uma nova lista de resultados, eventualmente melhor. Pretende-se com esta abordagem, dar liberdade ao utilizador para interagir com a aplicação de forma intuitiva, construindo uma aprendizagem relativamente ao sistema, onde combinando os métodos de pesquisa utilizando esboço e imagem, naturalmente resultará na obtenção melhores resultados.

### **3.2 Primeira Análise**

Antes de iniciar o processo de desenvolvimento da interface para a aplicação no dispositivo móvel, é necessário perceber quais são os componentes necessários para o funcionamento da mesma. Para além disso, é importante definir a forma como a informação e o conteúdo devem ser estruturados.

### 3.2.1 Menu da Aplicação

Qualquer que seja a aplicação, para dispositivos móveis ou não, é essencial um menu que ofereça diversas funcionalidades e opções ao utilizador. Dado que se trata não só de uma ferramenta de pesquisa, mas também de uma aplicação que pretende a interação com o utilizador através da realização de esboços, é necessário oferecer um conjunto de ferramentas que permita personalizar tanto o quanto possível o desenho.

Deste modo, com base no estudo realizado no capítulo 2, foram criadas as seguintes funcionalidades:

- **Ajuda (Help):** Qualquer aplicação deve conter uma opção que apresente informações relevantes sobre a mesma, bem como algumas instruções básicas de utilização. Ao seleccionar esta opção (ver Figura 3.1), o utilizador visualiza uma mensagem de informação que resume os objetivos e funcionalidades principais da aplicação (heurística de Nielsen – *Help and documentation*).



Figura 3.1: Imagem descritiva do menu Help da interface de pesquisa desenvolvida.

- **Paleta de cores (*Color*):** Permite ao utilizador desenhar esboços mais enquadrados, através da utilização de cor. Uma vez que se trata de pesquisa de imagens baseada no conteúdo das mesmas, que na sua maioria são compostas por cor, é relevante levar o utilizador a usar cores no seu esboço. Quando esta opção é seleccionada (ver Figura 3.2), surge uma paleta com 12 cores, reconhecidas por todas as culturas. Esta restrição à utilização de cor permite uma aproximação às cores das imagens, quando na fase de recuperação é analisada a componente cor.

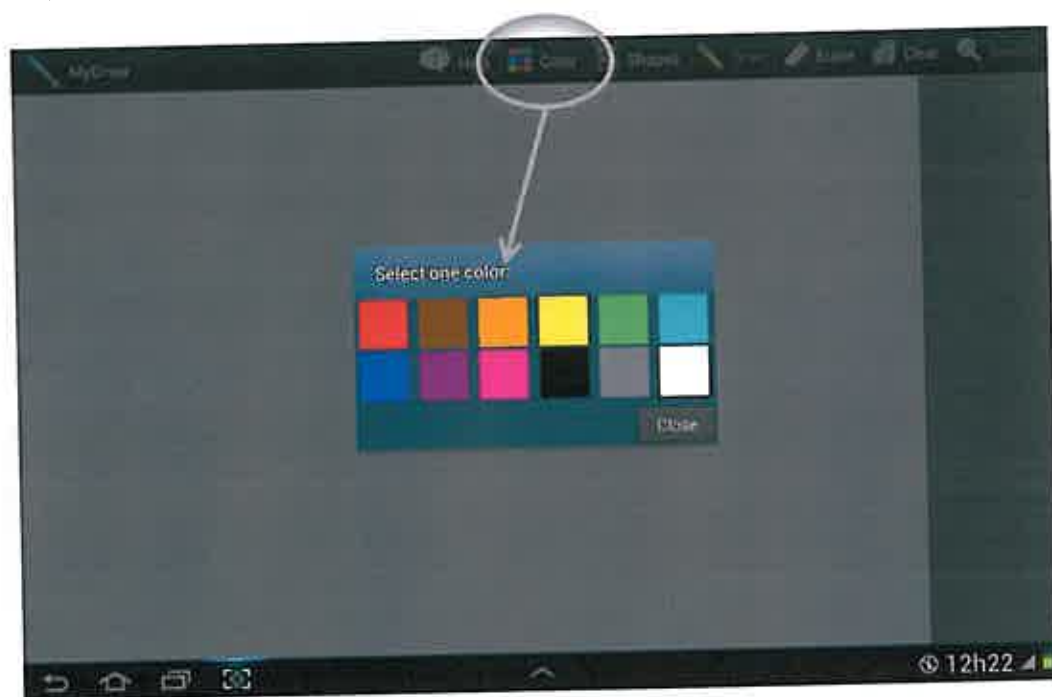


Figura 3.2: Imagem descritiva do menu Color e respetiva paleta de cores disponibilizada na interface de pesquisa desenvolvida.

- **Formas geométricas (*Shapes*):** Para além de poder desenhar à mão livre, o utilizador tem disponíveis algumas formas geométricas básicas (círculo/ oval, quadrado/ retângulo, linha reta). Esta ferramenta não é essencial, mas permite ao utilizador desenhar com mais precisão em alguns casos (ver Figura 3.3).



Figura 3.3: Imagem descritiva do menu Shapes e respectivas formas geométricas disponibilizadas na interface de pesquisa desenvolvida.

- **Desenhar/ Apagar (Draw/ Erase):** Numa aplicação de desenho, é imperativo permitir ao utilizador fazer correções, como se estivesse a utilizar uma borracha. Esta ferramenta permite ao utilizador eliminar pequenos elementos, sem ter de apagar o desenho por completo. Foi criada uma conjugação entre desenho e borracha (ver Figura 3.4), onde a ferramenta em utilização aparece seleccionada no menu, estando uma delas sempre seleccionada (heurística de Nielsen – *User control and freedom*).



Figura 3.4: Imagem descritiva dos menus Draw e Erase e respetiva conjugação entre um e outro, referentes à interface de pesquisa desenvolvida.

- **Limpar a área de desenho (Clear):** Outra funcionalidade importante numa aplicação de desenho. Esta opção permite ao utilizador apagar por completo o desenho e recomeçar novamente. Como forma de prevenção e tendo em atenção alguns conceitos estudados para o desenvolvimento de interfaces, quando esta opção é seleccionada surge uma mensagem de confirmação (ver Figura 3.5), garantindo que o utilizador não seleccionou a opção inadvertidamente (heusística de Nielsen – *Error prevention*).



Figura 3.5: Imagem descritiva do menu Clear da interface de pesquisa desenvolvida.

- **Pesquisar (Search):** Para além das ferramentas de desenho descritas anteriormente, esta pode ser considerada a funcionalidade principal da aplicação. Ao seleccionar este botão (ver Figura 3.6), o utilizador envia o seu esboço para o servidor como uma imagem formato PNG, realizando assim a pesquisa das imagens *clip art*.

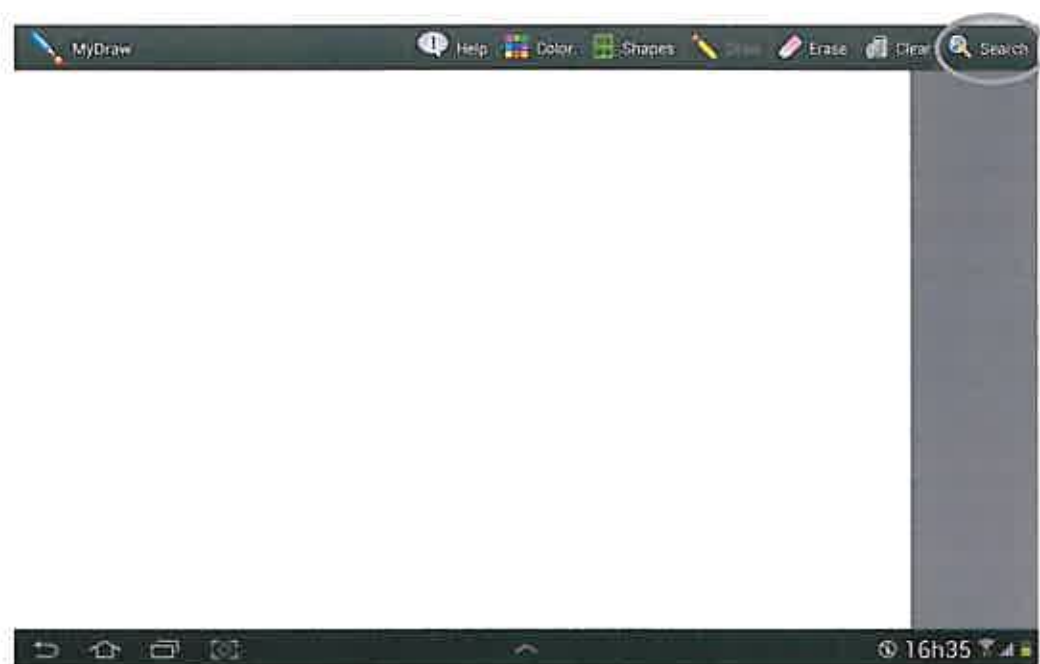


Figura 3.6: Imagem descritiva do menu Search na interface de pesquisa desenvolvida.

Para além destas opções, foram ainda avaliadas outras como permitir a escolha da grossura do lápis ou borracha, mas dado que se trata essencialmente de uma ferramenta para pesquisa de *clip arts* e não de desenho, considerou-se que a utilização dessas opções, nesta fase, era desnecessária.

### 3.2.2 Área de Desenho

Depois de analisar e decidir quais as ferramentas e opções que devem ser disponibilizadas na aplicação, é necessário avaliar a importância da área de desenho. Uma vez que a elaboração do desenho é um ponto-chave para o sucesso da pesquisa, considera-se que é importante oferecer espaço suficiente para o utilizador desenhar livremente sem se sentir demasiado limitado. Assim, foi decidido que esta deveria ser a maior área apresentada no ecrã da aplicação.

A Figura 3.7 mostra um exemplo de utilização da aplicação ao nível do desenho, desde a utilização de cores, formas geométricas, desenho livre e borracha.



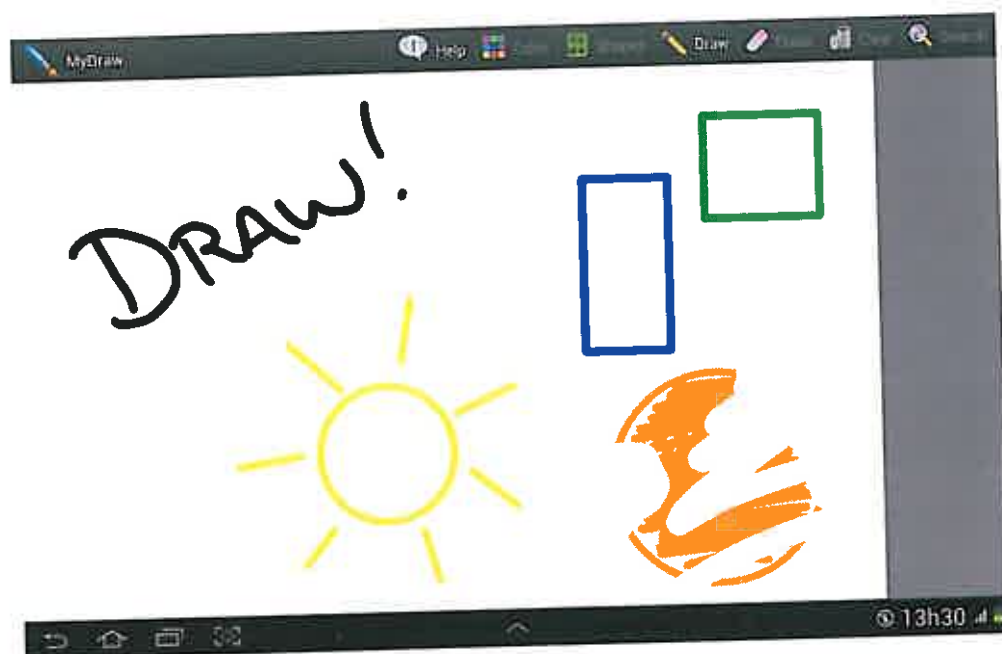


Figura 3.7: Imagem descritiva da utilização da aplicação no *tablet*, com a seleção de cores diferentes, várias formas geométricas, desenho livre e utilização da ferramenta para apagar pequenas áreas

### 3.2.3 Visualização de Resultados

Depois de analisar o menu e a área de desenho, falta ainda outro componente essencial da aplicação: a forma como os resultados (imagens *clip art*) devem ser apresentados ao utilizador.

Numa fase inicial do estudo, a área principal da aplicação continha apenas o espaço para o desenho, sendo os resultados apresentados num novo ecrã, unicamente para este fim. No entanto, com o decorrer do desenvolvimento da aplicação, considerou-se que seria relevante utilizar dois métodos de pesquisa distintos: primeiro seria realizado um esboço e com os resultados obtidos poderia ser feita uma pesquisa mais refinada, utilizando uma das imagens retornadas. A motivação da utilização destes dois métodos utilizados em conjunto, é explicada nos capítulos seguintes de uma forma mais detalhada. Deste modo, foi necessário criar uma nova estratégia e um rearranjo dos componentes necessários para este fim. Para além desta motivação, foi identificada ainda a importância de permitir ao utilizador comparar o seu esboço com o tipo de imagens obtidas, compreendendo de que forma o sistema interpretava os seus esboços e de que forma isso se reflete nos resultados.

Tendo em conta todos os fatores mencionados, foram realizadas diversas experiências para mostrar os resultados no mesmo ecrã em que o esboço é feito. Dado que a aplicação é utilizada com o *tablet* na orientação horizontal (*landscape*), se os resultados fossem apresentados por cima ou por baixo da área de desenho, o utilizador estaria demasiado condicionado a uma área de desenho horizontal e estreita. Assim, considerou-se como melhor opção posicionar os resultados lateralmente à área de desenho, obtendo assim uma interface mais intuitiva. Em segundo lugar, foi necessário considerar de que forma a pesquisa utilizando uma das imagens retornadas deveria ser efetuada. De forma a respeitar uma interface intuitiva e interativa, decidiu-se que a melhor opção seria selecionar uma imagem carregando sobre ela. Ao realizar esta ação, o utilizador envia uma nova *query* ao sistema de recuperação de imagem.

Em relação aos métodos de pesquisa que poderiam ter sido utilizados, foi ainda avaliada a possibilidade de tornar o sistema mais interativo, fazendo várias pesquisas enquanto o utilizador progredia no seu desenho. Esta possibilidade permitiria ao utilizador visualizar resultados à medida que ia desenhando, podendo adaptar-se à medida que desenhava. Esta funcionalidade foi implementada, mas foi desconsiderada ainda durante a implementação, uma vez que o tempo requerido para contactar o servidor e obter as imagens mais próximas provocava um atraso na utilização da aplicação.

Na Figura 3.8 é possível ver a forma como os resultados são apresentados ao utilizador. É importante referir que em ambos os ecrãs, é possível fazer *scroll down* para visualizar todos os resultados. Na pesquisa utilizando esboço são apresentadas as vinte imagens mais próximas e na pesquisa utilizando uma dessas imagens são retornadas as vinte e quatro imagens mais próximas. Todos os resultados são ordenados por ordem de proximidade, ou seja, a primeira imagem é sempre a mais próxima e assim sucessivamente.

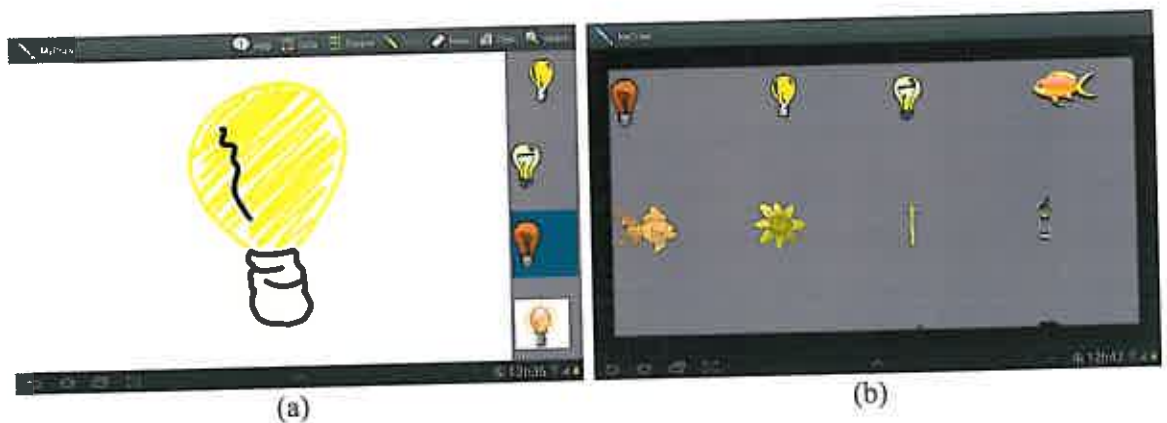


Figura 3.8: A figura (a) mostra um esboço feito pelo utilizador e os resultados obtidos depois da pesquisa. A figura (b) mostra os resultados obtidos, quando o utilizador faz uma nova pesquisa através da seleção de uma das imagens apresentadas na barra lateral cinzenta.

### 3.2.4 Características de Usabilidade

Durante o desenvolvimento da interface da aplicação, foram respeitados os conceitos e regras gerais, algumas já referidas anteriormente. Contextualizando algumas decisões tomadas no âmbito de desenvolvimento, são destacadas em seguida algumas características da aplicação associadas às heurísticas de Nielsen:

- O sistema mantém o utilizador informado sobre qual a ferramenta que está a utilizar (lápiz ou borracha), através da visualização dessa opção seleccionada no menu (*visibility of system status*).
- O utilizador tem total liberdade para efetuar uma pesquisa em qualquer momento, alterar a cor ou forma do seu desenho e apagar parcial ou totalmente o que está no ecrã (*user control and freedom*).
- Cada opção do menu tem uma função específica e é apresentada de forma clara e visual ao utilizador, através da utilização de pequenos ícones em cada uma das opções (*consistency and standards e recognition rather than recall*).
- É apresentada uma mensagem de erro, caso o dispositivo do utilizador não esteja ligado à Internet, uma vez que esta é necessária para efetuar uma pesquisa. Ainda assim, o utilizador pode utilizar a aplicação para desenho, sendo apenas desativada a opção de pesquisa (*error prevention*).
- Existe ainda uma opção que permite descrever de forma simplificada como a plataforma deve ser utilizada (*help and documentation*).

Um dos pontos-chave desta aplicação, é a necessidade de conexão à Internet, uma vez que o processo de recuperação de imagem é feito através de um servidor externo. Desta forma, é necessário prevenir o utilizador para este facto, mostrando-lhe uma mensagem de alerta caso o dispositivo não esteja ligado à rede. Este processo é demonstrado através da Figura 3.9, realçando na Figura 3.9 (b) que depois da mensagem informativa a opção “Search” deixa de estar ativa, uma vez que não é possível contactar o servidor.

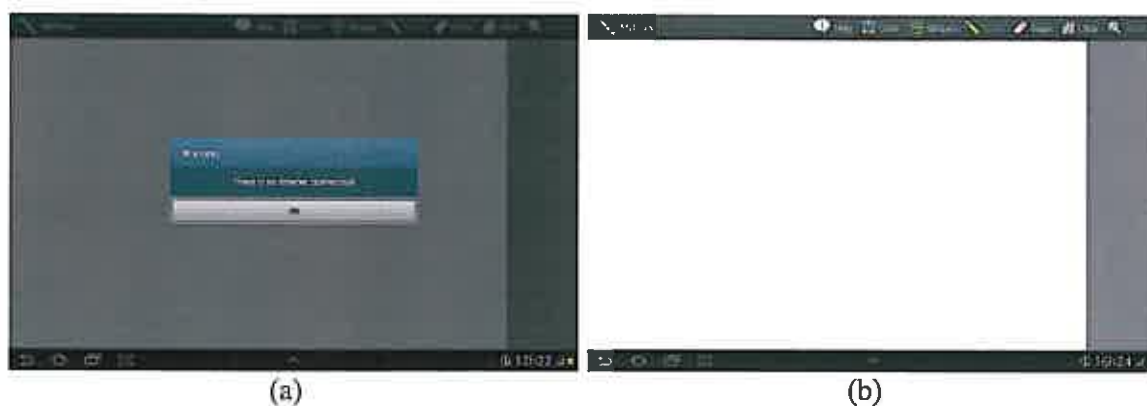


Figura 3.9: A figura (a) apresenta uma mensagem a informar o utilizador que não tem conexão à rede e portanto não pode naquele momento efetuar pesquisas, podendo mesmo assim utilizar a aplicação para desenho. Por fim, a figura (b) mostra o estado da aplicação quando não há ligação à Internet, desabilitando a opção “Search”.

### 3.3 Implementação da Interface

Nas secções anteriores, o desenvolvimento da interface foi abordado de uma forma mais conceptual e visual, descrevendo as opções e decisões que foram tomadas ao nível da interface, apoiada por heurísticas e conceitos de usabilidade. Nesta secção, pretende-se descer ao nível do desenvolvimento detalhando alguns pormenores de implementação bem como decisões ao nível do desempenho.

#### 3.3.1 Especificação da Plataforma

Quando se pretende desenvolver aplicações para dispositivos móveis, é necessário tomar algumas decisões relacionadas com o tipo de dispositivo que pretendemos contemplar. Neste

caso, foi decidido que a aplicação seria desenvolvida para *tablet*, uma vez que estes dispositivos têm maiores dimensões do que os *smartphones*, permitindo uma maior precisão no desenho e melhor visualização de resultados. Depois desta decisão, foi necessário analisar as hipóteses quanto à plataforma que deveria ser considerada. Entre as mais utilizadas, podem destacar-se duas: iOS da Apple e Android da Google. Avaliando alguns aspetos ligados a cada uma das plataformas, o Android pode ser distinguido por ser uma plataforma aberta, reduzindo assim os custos e o esforço de desenvolvimento. Para além disso, a plataforma Android faz parte de um ecossistema amplo, onde é possível encontrar facilmente conteúdos, serviços e aplicações.

Tendo em conta o objetivo de desenvolver a uma aplicação para *tablets*, com a visão e experiência o mais realistas possível, foi adquirido um *tablet* para acompanhar todo o processo de desenvolvimento. Como tal, foi necessário decidir qual o tipo de dispositivo a ser utilizado, avaliando capacidade de processamento, armazenamento e qualidade do *input* para a realização de desenhos à mão livre. Optou-se pela utilização de um *tablet Samsung Galaxy Note* (modelo GT-N8010) de 10.1 polegadas, com Android 4.0 e *pen* integrada. Esta escolha deveu-se à qualidade de ecrã e resolução deste dispositivo, para além de outro factor mais importante ainda, que está relacionado com a adaptação do sistema para interagir com uma caneta táctil, que permite um fácil fluxo de escrita combinado com precisão exata de ponta de caneta. Outras características são: um CPU *quad core* com 1.4 Ghz e 16 GB de memória interna.

### **3.3.2 Desenvolvimento**

Para o desenvolvimento da interface de utilizador proposta, foi utilizada uma plataforma *open source*, o Eclipse<sup>11</sup>, desenhada para o desenvolvimento de *software*. Na sua maioria, a implementação deste projeto é baseada na linguagem Java, dado que a plataforma Android é baseada também em Java. Dado que se trata de uma implementação específica para a plataforma Android, foi necessário utilizar a biblioteca Android SDK<sup>12</sup>, através da utilização de um plugin Android para o Eclipse. Esta biblioteca adiciona ao ambiente de desenvolvimento um conjunto de ferramentas e APIs necessárias para o desenvolvimento de

---

<sup>11</sup> <http://www.eclipse.org>

<sup>12</sup> <https://developer.android.com/sdk>

aplicações para a plataforma Android, bem como um emulador Android que permite testar a aplicação ao longo do processo de desenvolvimento.

O primeiro passo para começar a desenvolver uma aplicação para Android no Eclipse, é criar um novo projeto, mais especificamente um *Android Application Project*. A Figura 3.10 mostra de que forma o novo projeto é criado.

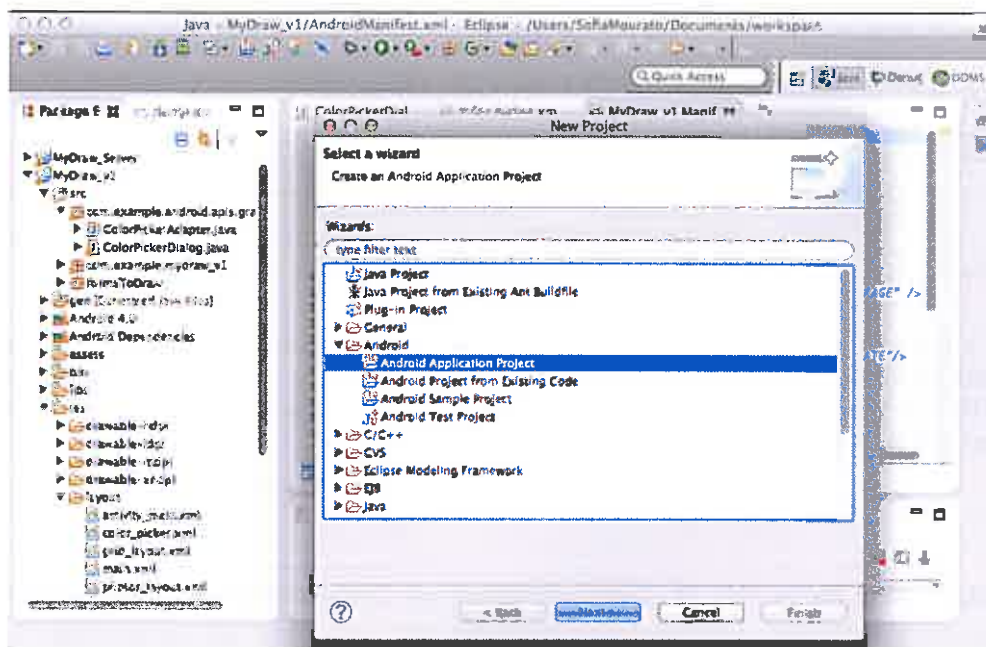


Figura 3.10: Imagem demonstrativa da criação de um novo projeto Android no Eclipse.

Dado que estamos a utilizar um dispositivo móvel real, quando é criada a aplicação é necessário executá-la através do Eclipse, enviando-a para o dispositivo (previamente conectado ao computador). Depois de transferida para o dispositivo, se não houver novas alterações à aplicação, esta passa a poder ser executada diretamente no dispositivo utilizando o ícone da aplicação que é criado no mesmo. As Figura 3.11 e Figura 3.12 mostram a forma como é passada a aplicação para o *tablet* e o ícone que é criado no *Desktop* do dispositivo móvel para executar a aplicação.

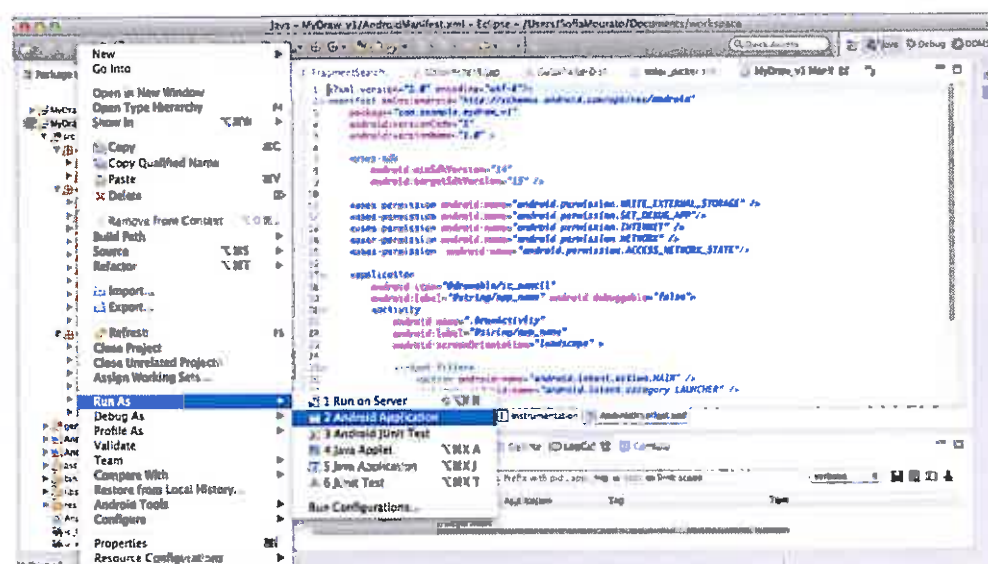


Figura 3.11: Imagem descritiva da forma como deve ser executada uma aplicação Android utilizando o Eclipse. Caso se queira que a aplicação seja executada num dispositivo móvel real, é necessário que este seja previamente ligado ao computador. Caso contrário, a aplicação pode ser executada num emulador no próprio Eclipse.



Figura 3.12: Imagem ilustrativa do ícone que é gerado pela aplicação Android e criado no Desktop do tablet.

A paleta de cores que é oferecida na aplicação, foi desenvolvida a partir de uma solução de *Color Picker*<sup>13</sup> para Android. Depois de analisar o código e os resultados desta pequena aplicação, foram aplicadas as 12 cores que se pretende que estejam disponíveis, bem como um botão “Close” que permite fechar a caixa de cores caso o utilizador não queira escolher outra cor. Para apresentar as cores em forma de grelha é utilizado um documento XML que define a estrutura do *layout* de apresentação de dado componente.

A aplicação correspondente à interface do utilizador, contém ainda uma base de dados com as mesmas imagens que o servidor, em formato PNG. Esta opção foi tomada, de modo a

<sup>13</sup> <http://www.ipcas.com/blog/2011/12/android-color-picker/>



otimizar o processamento na fase de recuperação de *clip arts*. Assim, em vez de serem passadas  $N$  imagens de um lado para o outro, são enviados apenas os nomes de referência das imagens e a aplicação vai buscá-las à sua base de dados para as apresentar no ecrã. As únicas imagens transferidas, são enviadas da aplicação no dispositivo móvel para o sistema de recuperação de imagem. A aplicação final ocupa no *tablet* cerca de 6,79 MB para a base de dados de 100 imagens. O tamanho da aplicação vai depender do tamanho da base de dados em utilização, ou seja, se for utilizada uma base de dados com cerca de 100 MB, esse vai ser o tamanho final da aplicação. Este fator deve se melhorado num trabalho futuro, definindo outras formas de utilizar as imagens, sem ter de as armazenar no dispositivo. A aplicação pode ser executada em qualquer *tablet* com sistema operativo Android, desde que este tenha a versão 4.0 instalada, ou superior.

### **3.4 Avaliação da Interface Utilizador**

Os avanços que se têm verificado no desenvolvimento de aplicações para dispositivos móveis, têm um grande impacto na interação com o utilizador, uma vez que podem ser utilizadas em qualquer altura e em qualquer lugar. Assim, foram realizados alguns testes de usabilidade, de modo a perceber quais as principais dificuldades sentidas pelos utilizadores, na utilização da interface e do sistema de pesquisa. Em seguida, são descritas algumas metodologias que foram utilizadas para a realização destes testes.

Em [31], são analisadas variantes de testes de usabilidade que permitem avaliar a interação do utilizador com a aplicação móvel, como é o caso do conceito mHCI. É necessário avaliar fatores contextuais, tais como: tarefa, ambiente, componentes sociais e tecnológicos e o impacto da aplicação no desempenho do utilizador. É ainda importante avaliar os níveis de desempenho de tarefas alcançados pelos utilizadores [32]. Os testes de usabilidade devem ser feitos por vários utilizadores que devem realizar tarefas típicas. Não se pode dizer que um método é melhor ou pior em todas as situações. Deste modo, é necessário escolher um método apropriado, tendo em conta os objetivos da avaliação e o tempo disponível [33].

Para realizar os testes, foi desenvolvido um questionário que pretende avaliar a usabilidade e o desempenho do sistema. Foram abordados tipos de utilizadores com áreas de trabalho e



conhecimentos distintos, de forma a tentar aproximar o mais possível a ambientes reais. Para testar a aplicação, foi disponibilizado o *tablet* com a aplicação instalada e o sistema de recuperação de imagem a funcionar, pedindo que estes realizassem algumas tarefas e avaliassem não só a interface de desenho mas também a qualidade dos resultados (ver anexos). No capítulo onde são apresentados os testes e resultados experimentais, são também descritos e discutidos os resultados obtidos com os testes de usabilidade.

# 4 Sistema de Recuperação de Imagem

Neste capítulo é descrito o sistema de recuperação de imagem que foi desenvolvido, essencial na utilização da interface de pesquisa apresentada anteriormente. Este sistema deve realizar todo o processamento pesado, na medida em que permite que o dispositivo móvel não seja sobrecarregado. Em seguida, é apresentado de forma detalhada o funcionamento do sistema desenvolvido, no que diz respeito à extração de características do conteúdo das imagens, bem como a sua classificação e recuperação.

## 4.1 Sistema Desenvolvido

Como foi referido no capítulo anterior, o ambiente de desenvolvimento utilizado foi o Eclipse. Para implementar o sistema de recuperação de imagem, foi necessário adicionar algumas bibliotecas externas a esta plataforma, dada a necessidade de utilização de algoritmos associados aos descritores, apresentados de seguida. Como tal, foi utilizado o OpenCV<sup>14</sup>, que contém diversos algoritmos previamente implementados que podem ser diretamente utilizados na fase de implementação. Para utilizar esta biblioteca, foi usada uma adaptação direcionada especificamente para a linguagem Java, o JavaCV<sup>15</sup>, que permite a utilização de funções OpenCV diretamente sobre o código Java.

O sistema de recuperação de imagem é executado num computador externo, separado da interface de utilizador no dispositivo móvel. A sua função é realizar todo o processamento de

---

<sup>14</sup> <http://opencv.org>

<sup>15</sup> <http://code.google.com/p/javacv/>

imagens necessário, de modo a não sobrecarregar o dispositivo móvel com este processamento. Para além disso, foram utilizados componentes que não correm sobre a plataforma Android, tais como o JavaCV e outras ferramentas com extensão JAR para aplicar componentes relacionados aos descritores.

Este sistema realiza tarefas de processamento tais como extração de características, vetorização de imagens *raster*, classificação de resultados e armazenamento das imagens *clip art* numa base de dados. Esta base de dados é composta por duas versões das mesmas imagens *clip art*: imagens vetoriais, com extensão SVG e imagens *raster*, com extensão PNG. A necessidade destas duas versões, está relacionada com a utilização de descritores diferentes, que extraem características de diferentes tipos, ou seja, para um descritor pode ser necessário utilizar imagens PNG e para outro diferente imagens SVG. Numa abordagem inicial, as características foram extraídas e a classificação feita de forma individual, de modo a poder comparar com abordagens diferentes. Posteriormente, numa fase mais avançada do desenvolvimento e análise de resultados, foi feita uma combinação entre as diferentes características, extraídas de ambos os tipos de imagem (SVG e PNG). O objetivo desta combinação é avaliar a possível melhoria de resultados, utilizando uma combinação de características, extraídas utilizando descritores distintos.

#### **4.1.1 Processamento da Base de Dados**

Antes de colocar o servidor disponível para responder a pedidos da interface de pesquisa, é necessário efetuar algum pré-processamento para todas as imagens da base de dados. O processamento consiste na extração de características das imagens da base de dados, utilizando cada versão da imagem (PNG e SVG) para os descritores respetivos. As imagens vetoriais (SVG) permitem extrair informação sobre geometria e topologia e as imagens *raster* (PNG) permitem obter informação sobre cor e textura. Este processamento prévio é executado apenas uma vez, ou quando novas imagens são adicionadas à base de dados, sendo necessário extrair as características correspondentes às novas imagens. De modo a guardar as características de forma permanente no servidor, foram gerados documentos XML com esta informação. Isto permite agilizar o processo de pesquisa, uma vez que o único processamento necessário em tempo real é a extração de características associadas ao esboço ou imagem *query*, comparando-as em seguida com o documento XML, que contém os vetores de

características associados a todas as imagens da base de dados. Estes ficheiros XML são gerados em tempo de execução, ou seja, enquanto a base de dados é percorrida fazendo a extração de características de cada imagem, é adicionado ao ficheiro XML o nome da imagem e o respetivo vetor de características. Para cada descritor, existe um ficheiro XML diferente, permitindo uma melhor distinção entre eles e também permite minimizar o número de linhas que é necessário percorrer para dada pesquisa, se for utilizado apenas um descritor.

Na Figura 4.1, é possível ver a estrutura e o conteúdo dos ficheiros XML descritos anteriormente. Como se pode ver, este ficheiro XML contém os vetores de características extraídos com o descritor Momentos de Cor. As imagens têm nomes genéricos (e.g. img1, img2), uma vez que o nome ou contexto das imagens não é relevante para a recuperação e classificação de imagens. Relativamente ao descritor Topogeo, é gerado um documento XML com a mesma estrutura que é apresentada na figura, mudando apenas o nome do componente base, ou seja, em vez de “<CMomentsDatabase>” é utilizado “<TopoGeoDatabase>”.

```
<?xml version="1.0" encoding="UTF-8"?>
<CMomentsDatabase>
  <image>
    <filename>img1</filename>
    <cmoments_descriptor>0.0 0.542106114362548 0.0 0.0
0.23968723614522325 0.0 0.0 0.7209459839967669 0.0 0.0
0.19000864760357486 0.0 0.0 1.0 0.0 0.0 0.0 0.0 (...)
    </cmoments_descriptor>
  </image>
  <image>
    <filename>img10</filename>
    <cmoments_descriptor>0.038599684592578315 0.9502346450349666
0.11278774638909257 0.011282369693925876
0.027668748361116654 0.08288613856867878 (...)
    </cmoments_descriptor>
  </image>
  .
  .
  .
</CMomentsDatabase>
```

Figura 4.1: Representação ilustrativa de um documento XML contendo os vetores de descritores associados a todas as imagens da base de dados. Neste caso específico, a ilustração representa os descritores extraídos com a utilização do descritor Momentos de Cor.

#### 4.1.2 Visão Geral do Sistema em Execução

Depois de preparar a interface de pesquisa e o sistema de recuperação de imagem, é possível ter a aplicação a funcionar na sua totalidade. Para que estes dois “sistemas” possam comunicar entre si, é necessário ter um protótipo de comunicação, tendo em conta que se

pretende um canal fiável de envio e receção de dados. Este tópico é abordado na secção seguinte.

Em suma, o utilizador deve realizar um esboço, através da interface no dispositivo móvel (cliente), que por sua vez é redirecionado para o sistema de recuperação de imagem (servidor). Caso o utilizador já tenha feito uma pesquisa através de esboço, pode refinar a sua pesquisa utilizando uma das imagens retornadas como *query*. Por sua vez, o sistema de recuperação de imagem extrai o vetor de características do esboço ou da imagem recebida, comparando-a posteriormente com os vetores de descritores armazenados nos documentos XML, consoante o método de extração de características utilizado. Esta comparação é feita através do cálculo de distâncias, de forma a classificar as imagens por ordem de proximidade. No final, o servidor obtém um vetor com o nome das N imagens mais próximas à *query* recebida (da mais próxima à mais distante), retornando-o para a interface do utilizador. A interface interpreta os nomes das imagens recebidos e apresenta as imagens no ecrã.

Na Figura 4.2, é apresentado o modelo que representa o processo descrito anteriormente, destacando as principais operações realizadas no processo de recuperação de *clip arts*.



Figura 4.2: Modelo representativo das principais operações realizadas no processo de recuperação de imagens clip art, para o sistema que foi desenvolvido.

## 4.2 Comunicação entre o Dispositivo Móvel e o Sistema de Recuperação de Imagem

No âmbito do desenvolvimento de uma aplicação para recuperação de *clip arts* através da utilização de dispositivos móveis, houve a necessidade de desenvolver duas componentes distintas: uma interface para realizar esboços e visualizar resultados (cliente) e um sistema de recuperação de imagem que faz a extração do conteúdo nas imagens e a classificação de resultados (servidor). Como tal, cada um destes componentes não funciona de forma individual, sendo necessário interligar os dois para que o processo de pesquisa seja totalmente executado.

Com base nesta problemática, foi utilizada uma arquitetura cliente-servidor, que permite aos dois componentes desenvolvidos comunicar entre si. No que diz respeito à comunicação entre cliente e servidor, optou-se pela utilização de *sockets*, que permitem criar um canal de comunicação entre duas máquinas. Este canal permite realizar as ações necessárias à nossa aplicação, onde o cliente necessita de enviar o seu esboço ou imagem ao servidor (em formato PNG), que por sua vez retorna uma string com o nome das imagens mais próximas. Na Figura 4.3, é possível visualizar o modelo e os principais métodos associados à utilização de *sockets*.

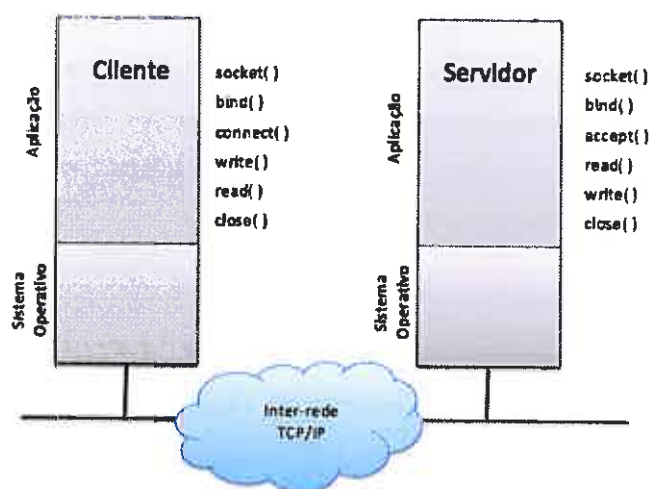


Figura 4.3: Modelo genérico de comunicação entre cliente e servidor utilizando sockets.

# 5

## Extração de Informação em Clip Arts

Este capítulo apresenta as tecnologias utilizadas no âmbito desta dissertação, de forma a extrair informação do conteúdo em imagens *clip art*. Dado que o âmbito desta dissertação assenta principalmente na utilização de esboços como *query*, é importante analisar as componentes que maioritariamente se verificam neste tipo de desenhos, como é o caso da cor e da forma. Deste modo, são utilizados dois descritores distintos, que permitem analisar estas duas componentes individualmente: o descritor Momentos de Cor (ou *Color Moments*) permite analisar a componente cor nas imagens e o descritor Topogeo permite analisar a topologia e geometria das mesmas.

Numa abordagem inicial, cada descritor é utilizado individualmente, no que respeita à recuperação de imagens. Posteriormente, ambos os descritores são utilizados em conjunto, de modo a verificar uma possível melhoria sobre os resultados individuais de cada descritor. Considerou-se que em certas ocasiões, avaliar a cor e a forma num mesmo instante pode ser vantajoso, na medida em que várias características são avaliadas no momento da recuperação.

### 5.1 Pré-processamento

Antes de iniciar o processo de extração de características nas imagens, é necessário aplicar algumas transformações de modo melhorar o desempenho dos descritores. Quando são utilizadas imagens *raster*, o tempo de processamento pode revelar-se um fator importante. Tendo em conta que muitas imagens (*clip arts*) têm fundos de grandes dimensões que não contêm informação útil, é necessário como primeiro passo cortar a imagem pelo seu

conteúdo, ou seja, remover todo o espaço vazio que existe à volta da imagem. O segundo passo consiste em verificar o tamanho da imagem depois do recorte efetuado, que não deve exceder os  $400 \times 400$  pixels. A decisão de utilizar um máximo de 400 pixels para largura e altura, mantendo as proporções da imagem, deve-se ao facto de este valor estar num limiar que permite reduzir o tamanho das imagens, mantendo ainda assim a relação entre velocidade e detalhe, tal como é descrito no estudo apresentado em [34]. Finalmente, é aplicado um filtro de *blur* às imagens, permitindo suavizar os contornos, de modo a remover pequenos elementos que não acrescentam informação útil às imagens.

Dado que é utilizado um descritor que utiliza imagens vetoriais, é ainda necessário na fase de pré-processamento vetorizar as imagens, que estão no formato *raster*. Num capítulo anterior, é referido que o sistema de recuperação de imagem contém uma base de dados com imagens PNG e SVG, mas é importante referir que a versão original das imagens é *raster*, sendo convertidas para SVG através da metodologia descrita neste parágrafo. Para a realização desta tarefa, são utilizadas duas bibliotecas que permitem converter uma imagem PNG (formato *raster*) para SVG (formato vetorial). Primeiro, é usado o ImageMagick<sup>16</sup>, que é um *software* que permite criar, editar, compor ou converter imagens *bitmap* (ou *raster*), convertendo as imagens no formato PNG (*Portable Network Graphics*) para PPM (*Portable Pixmap Format*). Depois, é necessário utilizar o Potrace<sup>17</sup>, uma ferramenta para transformar uma imagem *bitmap* em uma imagem suavizada e escalável para produzir uma imagem SVG (*Scalable Vector Graphics*), utilizando o formato PPM como origem. No fim destas transformações, as imagens estão prontas para a extração de características. Na Figura 4.2 é possível perceber algumas das transformações descritas anteriormente.

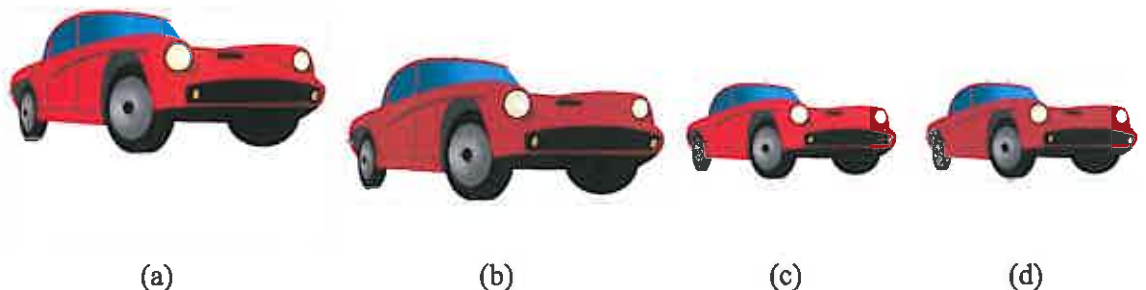


Figura 5.1: Representação da execução do pré-processamento nas imagens. A figura (a) representa a imagem raster original e a figura (b) representa a imagem depois de removido o espaço vazio à sua volta. A figura (c) mostra o redimensionamento aplicado à imagem e a figura (d) apresenta a imagem final a ser utilizada, depois de aplicado o filtro *blur*.

<sup>16</sup> <http://www.imagemagick.org>

<sup>17</sup> <http://potrace.sourceforge.net>



## 5.2 Descritor Momentos de Cor

Aplicado a imagens do tipo *raster*, o descritor Momentos de Cor permite a comparação entre imagens com base nas suas características de cor. A característica “cor” de uma imagem é frequentemente utilizada, dado a sua robustez e independência em relação ao tamanho e orientação das imagens [9]. É ainda importante referir, que o descritor Momentos de Cor apenas obtém informação da cor de cada pixel numa imagem, não envolvendo a relação local entre pixels vizinhos. Em geral, descritores que representam informação de cor apresentam melhores resultados em relação a outros descritores, exceto alguns tipos de imagens especiais onde a cor não é discriminativa. De seguida, são apresentadas as primitivas essenciais utilizadas para implementar o descritor Momentos de Cor.

Uma imagem a cores é normalmente definida por um modelo RGB. Assim sendo, a primeira coisa a fazer é alterar o espaço de cor da imagem para o modelo HSV (ou HSB), uma vez que se adequa melhor à perceção visual humana, principalmente quando se pretende fazer a comparação entre imagens (*matching*).

Para aplicar este descritor, divide-se a imagem em vários blocos (neste caso, numa grelha de 4 x 4), extraíndo de seguida os momentos de cor e agrupando-os em classes (componentes de cor). Por fim, é calculada a cor média e o desvio padrão para cada componente de cor, em cada um dos blocos. Em suma, é construída uma sequência que representa as relações espaciais entre as várias regiões detetadas. “*No entanto, imagens semelhantes podem ter resultados diferentes de segmentação, produzindo sequências 2D muito diferentes, afetando assim os resultados de recuperação*” [9]. O objetivo é obter vetores de descritores e compará-los através da utilização de métricas, neste caso foi calculada a distância Manhattan.

Para aplicar este descritor, podem destacar-se as seguintes etapas:

- Dividir cada imagem num número fixo de blocos (neste caso 16 blocos (4x4)), tal como é mostrado na Figura 5.2.

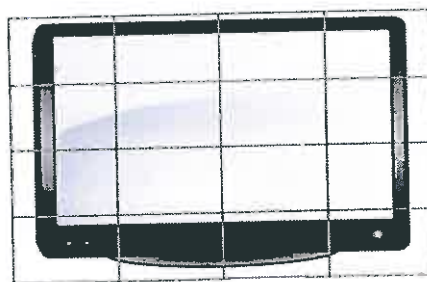


Figura 5.2: Representação ilustrativa da divisão da imagem em blocos, quando é utilizado o descritor Momentos de Cor.

- Calcular a cor média para cada componente de cor de cada bloco. A Equação 5.1 mostra a fórmula utilizada:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad (5.1)$$

- Calcular o desvio padrão para cada componente de cor de cada bloco. A Equação 5.2 mostra a fórmula utilizada:

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2} \quad (5.2)$$

Onde  $f_{ij}$  é o valor  $i$ -th da componente de cor situada no pixel  $j$  e  $N$  é o número total de pixels numa imagem [35].

### 5.3 Descritor Topogeo

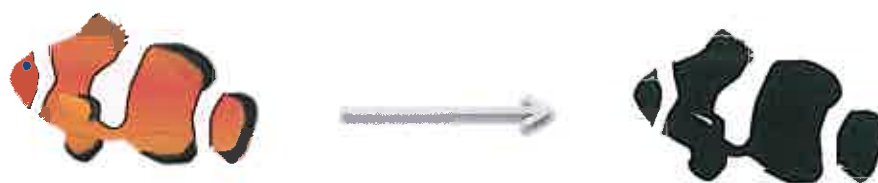
Aplicado a imagens do tipo vetorial, este descritor tem como objetivo analisar a topologia e geometria, ignorando as características de cor presentes nas imagens. Ao contrário do descritor de cor apresentado anteriormente, este é aplicado a imagens vetoriais, o que obriga a uma “vetorização” prévia das imagens *raster*. Neste contexto, pretende-se com o descritor *Topogeo* analisar a geometria (objetos contidos na imagem) e topologia (relação espacial

entre os objetos na imagem). Inicialmente, é feita uma filtragem utilizando a topologia e posteriormente é feita a pesquisa através da geometria [36]. Normalmente, uma característica por si só não é suficiente para distinguir formas, conduzindo a classificações incorretas [37].

Uma abordagem idêntica, que utiliza as diretrizes principais deste descritor, é apresentada em [38]. No que diz respeito a imagens vetoriais, as características relacionadas com a forma dos objetos (geometria) e a sua composição espacial (topologia), descrevem bem o conteúdo do desenho. É necessário aplicar previamente um processo de simplificação, de modo a eliminar alguns detalhes, tentando reter maioritariamente as formas do desenho e as suas relações. Depois da simplificação devem ser identificados os elementos visuais, reconhecidos através de polígonos e linhas, extraindo posteriormente informações geométricas e topológicas. No final do estudo produzido em [38], verificou-se que a análise da topologia produz melhores resultados do que a geometria.

De uma forma geral, foram realizadas as seguintes etapas no que diz respeito ao descritor *Topogeo*:

- Converter cada imagem de *raster* para vetorial (PNG para SVG), tal como é demonstrado na Figura 5.3.



**Figura 5.3:** Representação ilustrativa da transformação de uma imagem no formato raster (PNG) para o formato vetorial (SVG), retirando a componente cor.

- Executar a o ficheiro *CrushFeatureExtraction.jar* (disponibilizada por [36]), que realiza algum processamento sobre as imagens SVG, de modo a gerar os vetores de descritores.

## 5.4 Normalização de Descritores

Existem alguns descritores, que no processo de extração de características não produzem vetores normalizados, sendo necessário normalizar esses valores posteriormente. Esta normalização é importante, de modo a garantir que nenhum elemento avaliado comprometa todo o conjunto de dados, aquando da sua recuperação. Depois da normalização, é possível avaliar as características de diversos vetores proporcionalmente.

A normalização descrita acima, foi aplicada apenas ao descritor de cor (Momentos de Cor), dado que os descritores de topologia são normalizados pelo próprio sistema de extração de características. Este processo é descrito em [10].

O processo de normalização é feito calculando os *z-scores*, através das seguintes etapas:

- Obter uma matriz aplicando o descritor Momentos de Cor a todas as imagens da base de dados. Cada linha contém o vetor correspondente a uma imagem ( $M_i$ ), onde cada coluna contém os valores em dado momento correspondente a cada uma das imagens ( $M_j$ ).

$$M = \begin{matrix} & \begin{bmatrix} Im_{i_1} \\ Im_{i_2} \\ \vdots \\ Im_{i_N} \end{bmatrix} & \begin{matrix} M_i \\ \\ \\ \end{matrix} \\ \begin{matrix} M_j \\ \\ \\ \end{matrix} & \underbrace{\hspace{10em}}_{96 \text{ valores}} & \left. \vphantom{\begin{bmatrix} Im_{i_1} \\ Im_{i_2} \\ \vdots \\ Im_{i_N} \end{bmatrix}} \right\} 100 \text{ Imagens} \end{matrix}$$

- Calcular a média em cada coluna  $M_j$ . Para tal foi utilizada a Equação 5.3:

$$\mu = \frac{\sum_{i=1}^N x_i}{N} \quad (5.3)$$

- Calcular o desvio padrão em cada coluna  $M_j$ , utilizando a equação 5.4:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N - 1}} \quad (5.4)$$

- Normalizar toda a matriz e criar uma nova matriz normalizada, que será guardada num novo documento XML. Para tal, é utilizada a Equação 5.5:

$$Z(M_{i,j}) = \frac{M_{i,j} - \mu_{M_j}}{\sigma_{M_j}} \quad (5.5)$$

## 5.5 Cálculo de Distâncias

Depois de extrair os vetores de características referentes ao conteúdo das imagens, é necessário definir as métricas que devem ser utilizadas. Estas métricas permitem, na fase da recuperação de imagem, calcular as distâncias entre os vetores de descritores das imagens e assim definir uma relação de proximidade entre a *query* e cada uma das imagens da base de dados. Para tal, são utilizadas duas métricas, que permitem definir o grau de semelhança entre cada uma das instâncias: distância de Manhattan [39] e distância Euclideana [40]. A primeira medida é a soma das diferenças entre  $x$  e  $y$  em cada dimensão. A segunda é definida como a soma da raiz quadrada da diferença entre  $x$  e  $y$  nas suas respetivas dimensões. Nas Equações 5.6 e 5.7, são apresentadas as fórmulas utilizadas para o cálculo das distâncias.

- Distância de Manhattan:

$$distMan = |x_1 - x_2| + |y_1 - y_2| \quad (5.6)$$

- Distância Euclideana:

$$distEu = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5.7)$$

Para o descritor Momentos de Cor é utilizada a distância de Manhattan e para o descritor Topogeo é utilizada a distância Euclideana [1]. Uma vez calculadas, as distâncias obtidas necessitam ainda de ser normalizadas, para a posterior combinação entre estes dois os descritores [36]. Para tal, cada valor correspondente à distância da *query* a uma dada imagem é dividido pela máxima distância encontrada entre a imagem *query* e todas as imagens da base de dados. Isto significa que a imagem mais distante à *query*, é utilizada para normalizar todas as distâncias. As Equações 5.8 e 5.9 mostram uma representação dos cálculos efetuados.

- Distância de Manhattan Normalizada:

$$distManNorm = \frac{distMan_i}{distMan_{(maxDistance)}} \quad (5.8)$$

- Distância Euclideana Normalizada:

$$distEuNorm = \frac{distEu_i}{distEu_{(maxDistance)}} \quad (5.9)$$

## 5.6 Descritores Combinados

Para além da utilização individual dos descritores apresentados acima, foi ainda utilizada uma abordagem que combina os dois descritores. Com esta abordagem, pretende-se perceber se estes dois descritores combinados (Momentos de Cor e Topogeo) produzem melhores resultados juntos, do que se forem utilizados individualmente. No capítulo de testes e resultados experimentais, são descritos e analisados os resultados obtidos com esta combinação.

Quando se combinam descritores, existem várias metodologias que podem ser adotadas. Para os testes apresentados, são utilizadas as distâncias combinadas, representadas pelas Equações 5.10 e 5.11. Nas equações,  $n$  representa o total de características combinadas, neste caso  $n=2$

uma vez que são combinados apenas o descritor Momentos de Cor com o Topogeo. Na Equação 5.11,  $i$  representa cada instância presente no vetor de distâncias combinadas, que neste caso pode ser até 100 (base de dados pequena) ou até 12955 (base de dados grande).

$$distComb = \frac{\sum_{i=0}^{i=n} distNorm_i}{n} \quad (5.10)$$

$$distCombNorm = \frac{distComb_i}{maxDistance} \quad (5.11)$$

# 6 Testes e Resultados

## Experimentais

Neste capítulo são apresentados os resultados obtidos, através da realização de alguns testes que permitem avaliar o desempenho dos descritores implementados. Para além disso, é ainda importante analisar o desempenho do sistema de recuperação de imagem e a eficiência da utilização de esboço como *input* de pesquisa utilizando o dispositivo móvel.

Para a realização de testes numa etapa inicial, foi utilizada uma base de dados controlada de 100 imagens, com um grupo de 10 imagens para cada categoria. Posteriormente, foi utilizada uma base de dados de grandes dimensões com 12955 imagens, dividida também em categorias distintas, mas com números de imagens por categoria mais díspares. Esta base de dados de grandes dimensões teve como principal objetivo verificar a qualidade do desempenho do sistema de recuperação de imagem. Em seguida são apresentados os testes realizados e respetivas conclusões.

Todos os testes apresentados, foram realizados utilizando a interface instalada no *tablet Samsung Galaxy Note*, especificado no Capítulo 3, através da comunicação com o sistema de recuperação de imagem, a fim de este realizar todo o processamento necessário e respetiva classificação de resultados.

### 6.1 Dataset

Para a realização dos testes, foi necessário obter um conjunto de imagens para preencher uma base de dados credível e realista. Como tal, foram utilizadas duas bases de dados distintas, com dimensões muito diferentes. Considerámos que inicialmente seria melhor utilizar uma



base de dados de menor dimensão, uma vez que na fase de desenvolvimento é importante conhecer bem a base de dados. Depois de a implementação estar próxima do fim, foi então aplicada a base de dados de grandes dimensões, que permite avaliar o sistema de uma forma mais realista, uma vez que o número de imagens por categoria é muito mais díspar.

No âmbito da realização de testes, foi ainda necessário criar um conjunto de imagens de esboços, para que os resultados dos testes pudessem ser comparados de forma equivalente.

### **6.1.1 Bases de Dados de Clip Arts**

Como foi referido anteriormente, foram utilizadas duas bases de dados distintas para a realização de testes. Na base de dados de 100 imagens, existem 10 categorias diferentes, cada uma delas com 10 imagens *clip art*. As categorias existentes na base de dados são: Bulbs; Car; Fish; Flower; Hammer; Magnifying Glass; Phone; Sword; Tree; TV. A base de dados de 12955 imagens, contém as mesmas categorias que foram mencionadas para a base de dados controlada, acrescentando a categoria Other, que contém todas as imagens que não se enquandram em nenhuma das outras categorias. Assim, temos o seguinte número de imagens para cada categoria: 35 Bulbs; 186 Car; 97 Fish; 365 Flower; 26 Hammer; 13 Magnifying Glass; 47 Phone; 35 Sword; 154 Tree; 65 TV; 11932 Other.

A Figura 6.1 ilustra alguns *clip arts* pertencentes a cada categoria. As imagens *clip art* apresentadas em seguida, foram obtidas através da biblioteca OpenClipart<sup>18</sup>.

---

<sup>18</sup> <http://openclipart.org>



Figura 6.1: Imagem ilustrativa, com alguns exemplos de imagens clip art percentes a cada uma das categorias utilizadas (11 categorias no total).

### 6.1.2 Esboços Utilizados para Teste

Para realizar os testes, foram feitos dois esboços para representar cada categoria existente na base de dados (exceto para a categoria “Other”), que foram utilizados para avaliar o desempenho do sistema de recuperação de imagem e os dois descritores utilizados (Momentos de Cor e Topogeo). As Figura 6.2 e Figura 6.3 ilustram os conjuntos de esboços de forma separada, dado que o primeiro conjunto foi utilizado para realizar os testes definidos como “Esboço Detalhado” e o segundo conjunto para realizar os testes definidos como “Esboço Simples”. Esta separação foi considerada relevante, na medida em que se pode verificar que o primeiro conjunto apresenta esboços mais detalhados e com mais preenchimento de cor, o que não acontece com o segundo conjunto, que apresenta esboços mais simples, representando apenas as linhas gerais que permitem definir o objeto. Esta separação permite, na fase de análise de resultados, definir em que medida a falta de detalhe no esboço é prejudicial para a recuperação de imagens.

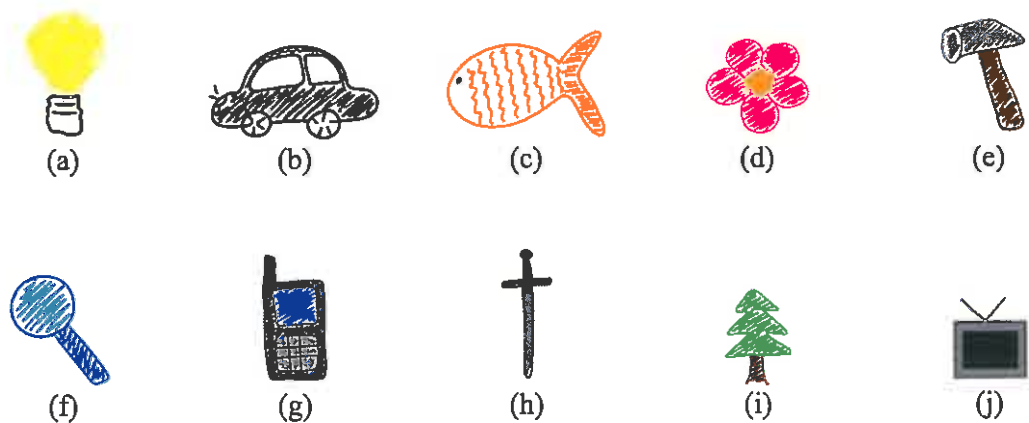


Figura 6.2: Imagem ilustrativa dos esboços que foram utilizados para testes. Estes esboços são referidos nos testes como “Esboço Detalhado”.

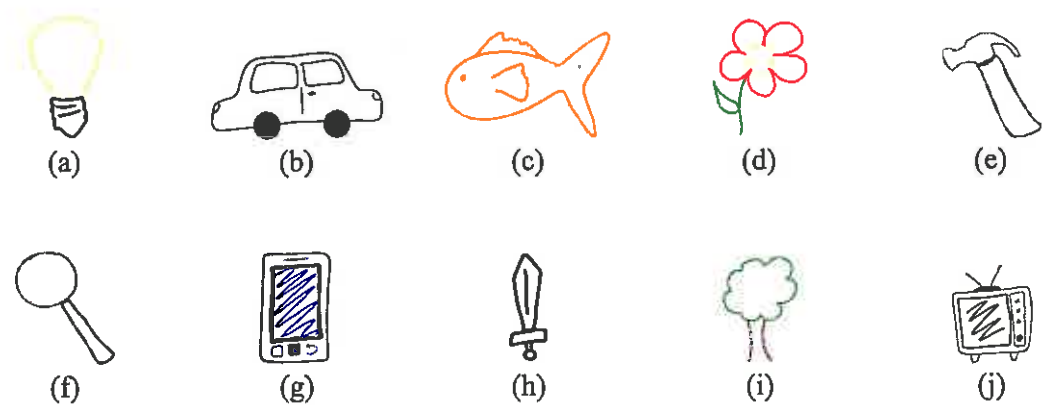


Figura 6.3: Imagem ilustrativa dos esboços que foram utilizados para testes. Estes esboços são referidos nos testes como “Esboço Simples”.

## 6.2 Métodos de Avaliação

O desempenho dos algoritmos desenvolvidos foi avaliado contabilizando a quantidade de imagens relevantes e não relevantes retornadas pelo sistema em cada pesquisa, bem como a ordem pela qual estas são devolvidas. Como tal, foram utilizadas as medidas de *Recall* e *Precision* [41], que permitem avaliar sistemas de recuperação de informação. A medida de *Recall*, permite avaliar a cobertura do sistema, analisando a sua capacidade de encontrar as imagens relevantes na coleção total de imagens. Já a medida *Precision*, permite avaliar a

precisão do sistema, na medida em que consegue retornar ao utilizador as imagens que são relevantes.

Para calcular as medidas de *Recall* e *Precision*, é necessário definir qual o tamanho do resultado que deve ser retornado em cada pesquisa. Em seguida, são apresentadas nas Equações 6.1 e 6.2, as fórmulas utilizadas para calcular as medidas de *Recall* e *Precision*.

$$Recall = \frac{\text{número de imagens relevantes recuperadas}}{\text{número total de imagens relevantes}} \quad (6.1)$$

$$Precision = \frac{\text{número de imagens relevantes recuperadas}}{\text{número total de imagens recuperadas}} \quad (6.2)$$

Nas secções seguintes são apresentados os testes e os resultados obtidos, através da utilização da medida *R-Precision*. Esta medida permite avaliar quantas imagens é necessário retornar, para encontrar um número pré-definido de imagens relevantes (ver Equação 6.3). Uma vez que a interface de pesquisa e de visualização de resultados tem um ecrã relativamente pequeno, foi utilizada esta medida de avaliação de resultados, dado que permite perceber de uma forma geral quantas imagens é necessário retornar para que o utilizador consiga visualizar sempre um pequeno grupo de imagens correspondentes ao que procura. Como várias imagens da mesma categoria podem ter formas e cores completamente distintas, neste âmbito pretende-se apenas visualizar as que se aproximam mais do esboço realizado.

$$R - Precision = \frac{\text{Número pré - definido de imagens relevantes}}{\text{Quantas imagens até obter imagens relevantes}} \quad (6.3)$$

De modo a avaliar o desempenho dos algoritmos e tipos de pesquisa utilizados, as imagens das bases de dados foram previamente classificadas por categorias, manualmente. Esta classificação permite comparar os resultados obtidos por cada descritor e tipo de pesquisa de forma semi-automática. Uma vez que é o sistema de recuperação de imagem que contabiliza o número e as posições das imagens retornadas, para determinada categoria.

### 6.3 Testes para Query by Sketch

Para os testes apresentados, foi utilizado um valor de  $N = 3$  para calcular a medida *R-Precision*, ou seja, foram procuradas as três primeiras imagens correspondentes a cada categoria pesquisada. Apesar de o valor de  $N$  ser baixo, considerou-se que este é um valor razoável, na medida em que são utilizados esboços para realizar a pesquisa. O mais importante, é que numa dada pesquisa seja retornado pelo menos um resultado inserido no âmbito pretendido.

#### 6.3.1 Base de Dados com 100 Imagens

Em primeiro lugar, foram realizados testes com a base de dados controlada de 100 imagens. Dado que esta base de dados tem um número igual de imagens por categoria, é possível comparar de uma forma mais estável os resultados. Em seguida, são apresentados os resultados obtidos utilizando a medida *R-Precision*.

A Tabela 6.1 apresenta os resultados globais obtidos, utilizando o método de pesquisa *query by sketch*. A última linha desta tabela, com o título “Desempenho Geral do Descritor” representa o cálculo da média de todos os valores das *queries* obtidos (“Esboço Detalhado” e “Esboço Simples”) para cada um dos descritores. Estes valores são representados mais abaixo num gráfico que mostra o desempenho de cada descritor utilizado.

	Color Moments		Topogeo		Color Moments + Topogeo	
	Esboço Detalhado	Esboço Simples	Esboço Detalhado	Esboço Simples	Esboço Detalhado	Esboço Simples
<b>Bulbs</b>	1	0,177	0,333	0,143	1	0,167
<b>Car</b>	0,25	0,231	0,158	0,131	0,375	0,6
<b>Fish</b>	0,3	0,15	0,039	0,05	0,15	0,064
<b>Flower</b>	1	0,064	0,143	0,2	1	0,065
<b>Hammer</b>	0,231	0,214	0,111	0,111	0,231	0,231
<b>Magnifying Glass</b>	1	0,158	0,094	0,097	0,75	0,167
<b>Phone</b>	0,333	0,6	0,086	0,097	0,158	0,375
<b>Sword</b>	0,429	0,2	0,111	0,081	0,6	0,2
<b>Tree</b>	0,5	0,054	0,131	0,136	0,5	0,054
<b>TV</b>	0,115	0,115	0,177	0,111	0,115	0,12
<b>Desempenho Geral do Descritor</b>	0,35605		0,127		0,3461	

Tabela 6.1: Resultados obtidos com os testes utilizando Query by Sketch para a base de dados pequena (100 imagens). Os valores apresentados correspondem à medida R-Precision com  $N = 3$ .

Para a realização dos testes, foram utilizados dois conjuntos de esboços diferentes. O primeiro conjunto apresenta um maior nível de detalhe e o segundo apresenta esboços mais simples, sem muito preenchimento de cor.

Como se pode ver nas Figura 6.4 e Figura 6.5, a utilização de esboços detalhados pode conduzir a um maior nível de *matching* entre o esboço produzido e as imagens existentes na base de dados. No caso de esboços mais simples, o *matching* é mais baixo a um nível geral. Um dos principais fatores para a obtenção destes resultados, é o facto de o descritor Momentos de Cor apresentar valores de *matching* mais altos, o que faz com que os esboços mais detalhados e com preenchimento de cor consigam melhores resultados a um nível global. Quando são utilizados esboços mais simples, apenas com as linhas gerais representativas de cada objeto, nota-se uma ligeira melhoria no desempenho do descritor Topogeo, uma vez que este analisa apenas o factor “forma”. Como tal, o conjunto de esboços “Esboço Simples” contém menos linhas (utilizadas para o preenchimento) do que o conjunto “Esboço Detalhado”, o que permite ao Topogeo analisar melhor a forma representada na imagem.

No que diz respeito aos descritores combinados verifica-se que, de uma forma geral os resultados são satisfatórios. Dado que o descritor Momentos de Cor retorna valores muito altos e o descritor Topogeo valores muito baixos, quando os dois são combinados, existe uma ligeira descida no desempenho dos resultados combinados. Isto significa que, uma vez que o

descritor Momentos de Cor retorna valores altos e o Topogeo valores baixos, os resultados referentes ao primeiro descritor são prejudicados em detrimento do segundo, resultando numa combinação que acaba por não ser melhor do que o melhor descritor isolado. Esta combinação permite avaliar dois tipos de informação diferentes na mesma pesquisa, combinando cor e forma, podendo ser melhorada com outros descritores diferentes, ou dando pesos diferentes a cada descritor, de modo a não prejudicar os melhores resultados.

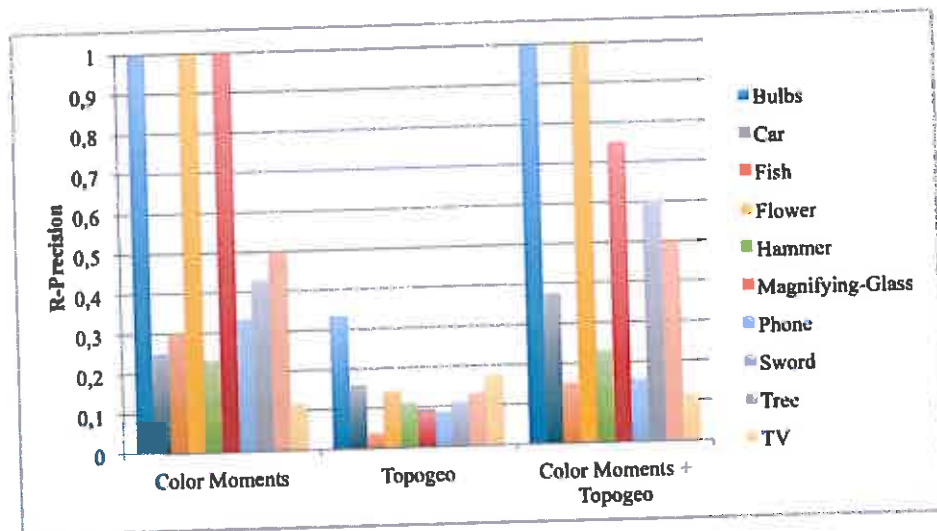


Figura 6.4: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Detalhado" para a base de dados de 100 imagens.

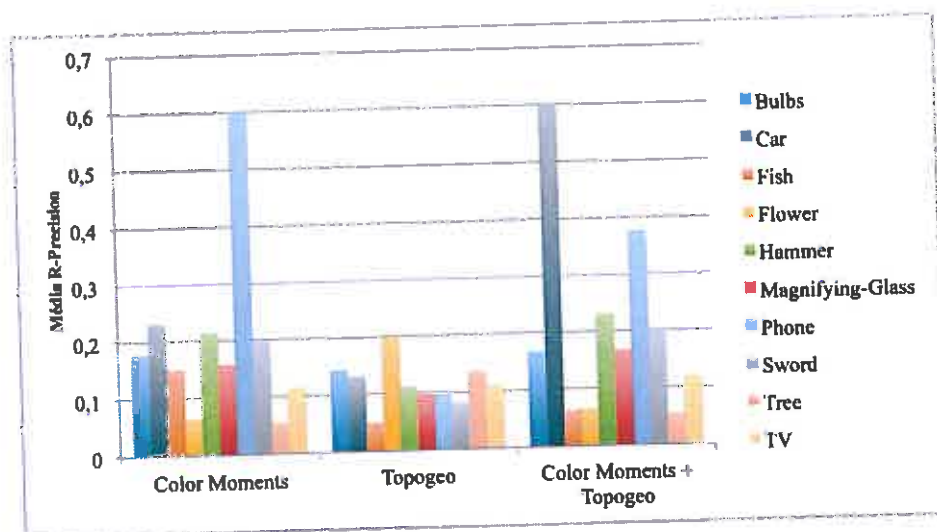


Figura 6.5: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Simples" para a base de dados de 100 imagens.

### 6.3.2 Base de Dados com 12955 Imagens

Em seguida, foram realizados testes com a base de dados grande, com 12955 imagens. É importante, mais uma vez, referir que esta base de dados tem um número bastante díspar de imagens pertencentes a cada categoria. Esta variável, influência consideravelmente os resultados obtidos, como se pode ver em seguida.

Na Tabela 6.2, são apresentados os resultados globais obtidos, utilizando também a medida *R-Precision*. Mais uma vez, a última linha da tabela corresponde ao desempenho geral dos resultados obtidos para cada um dos descritores.

	Color Moments		Topogeo		Color Moments + Topogeo	
	Esboço Detalhado	Esboço Simples	Esboço Detalhado	Esboço Simples	Esboço Detalhado	Esboço Simples
<b>Bulbs</b>	0,167	0,011	0,003	0,002	0,2	0,013
<b>Car</b>	0,021	0,021	0,016	0,05	0,009	0,006
<b>Fish</b>	0,003	0,007	0,008	0,008	0,003	0,017
<b>Flower</b>	0,15	0,022	0,019	0,033	0,115	0,026
<b>Hammer</b>	0,006	0,002	0,005	0,002	0,005	0,004
<b>Magnifying Glass</b>	0,333	0,001	0,001	0,002	0,158	0,001
<b>Phone</b>	0,005	0,002	0,004	0,002	0,002	0,003
<b>Sword</b>	0,022	0,002	0,01	0,002	0,011	0,002
<b>Tree</b>	0,019	0,007	0,012	0,028	0,034	0,008
<b>TV</b>	0,159	0,001	0,009	0,005	0,136	0,002
<b>Desempenho Geral do Descritor</b>	0,04805		0,01105		0,03775	

Tabela 6.2: Resultados obtidos com os testes utilizando Query by Sketch para a base de dados grande (12955 imagens). Os valores apresentados correspondem à medida *R-Precision* com  $N = 3$ .

Como se pode ver nas Figura 6.6 e Figura 6.7, os testes com *query by sketch* para esta base de dados produziram piores resultados que os apresentados para a base de dados pequena. As conclusões destes resultados, são idênticas às apresentadas anteriormente, verificando que o descritor Topogeo funciona melhor com esboços mais simples, e o descritor Color Moments para esboços mais complexos.



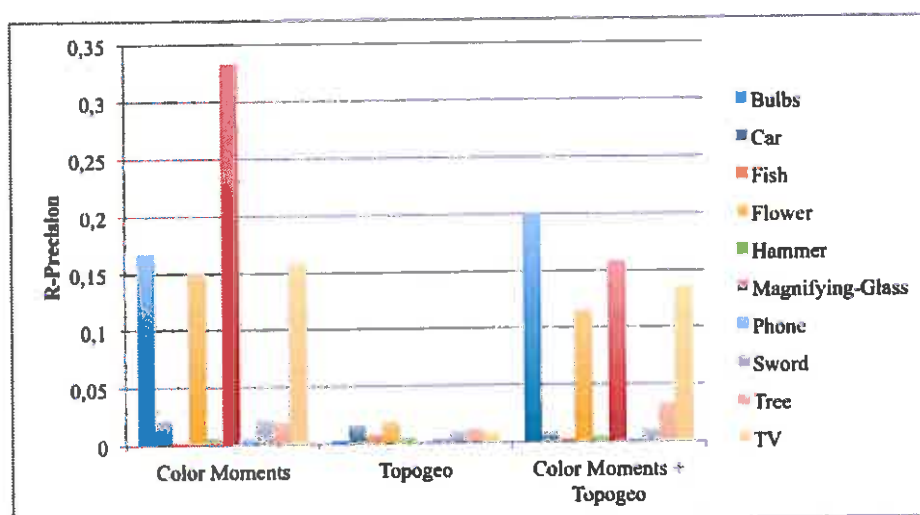


Figura 6.6: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Detalhado" para a base de dados de 12955 imagens.

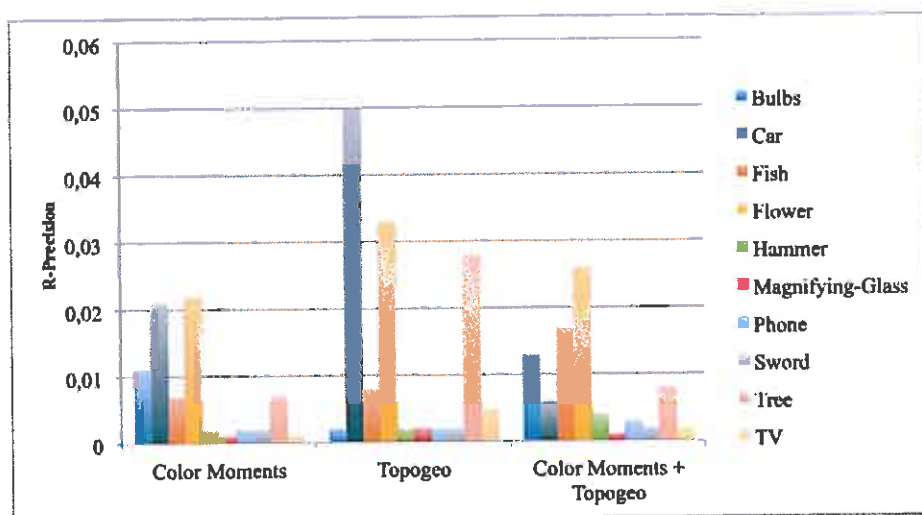


Figura 6.7: Gráfico com os resultados globais dos testes realizados com o conjunto de esboços da categoria "Esboço Simples" para a base de dados de 12955 imagens.

### 6.3.3 Desempenho dos Descritores para Esboço

No final dos testes realizados para avaliar o desempenho da pesquisa utilizando *query by sketch*, foi feita uma análise global sobre quais os descritores e as categorias que

apresentaram maior valor de *matching*. Como se pode ver na Figura 6.8, a categoria “Bulbs” foi a que obteve melhores resultados a um nível geral.

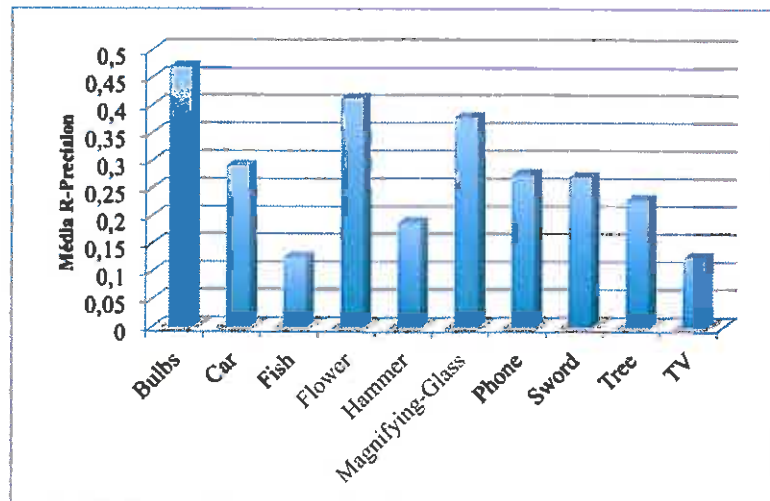


Figura 6.8: Gráfico com a estatística global de desempenho das categorias, para os testes realizados com Query by Sketch.

De uma forma geral (ver Figura 6.9), é possível perceber que o descritor Momentos de Cor teve um melhor desempenho do que o descritor Topogeo, ou mesmo do que estes dois combinados. Este resultado final, deve-se ao facto de o descritor Topogeo oferecer resultados relativamente maus, o que diminui o bom desempenho do descritor Momentos de Cor.

Com estes resultados, podemos concluir que a combinação destes dois descritores não é a melhor opção, uma vez que o descritor Momentos de Cor funciona melhor sozinho, neste caso. Outras combinações, podem trazer melhores resultados em estudos futuros.

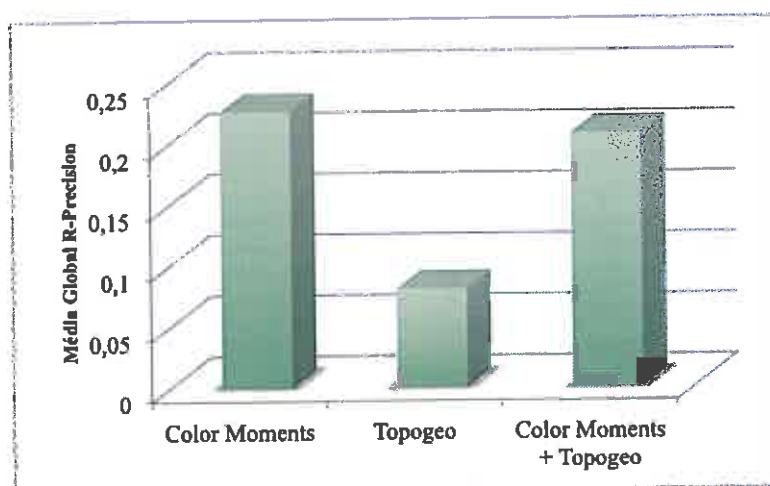


Figura 6.9: Gráfico com a estatística de desempenho global de cada um dos descritores, para os testes realizados com Query by Sketch.

## 6.4 Testes para Query by Image

Nesta secção, são apresentados os resultados obtidos na segunda abordagem de pesquisa referida anteriormente. Esta pesquisa só existe após a realização de um primeiro esboço e obtenção das 20 imagens *clip art* mais próximas ao esboço realizado. Posteriormente, uma dessas imagens pode ser utilizada como *query*, o que permite ao utilizador definir melhor o contexto que procura, permitindo assim que o sistema produza também melhores resultados.

Os resultados apresentados na Tabela 6.3, foram também obtidos através da medida *R-Precision*. Dado que apesar de para estes testes serem utilizadas imagens e não esboços como *input*, considerou-se que seria importante utilizar a mesma medida de classificação, uma vez que se pretende comparar estes resultados com os obtidos através de *query by sketch*. Isto significa, que foi utilizado um  $N = 3$  para calcular o *R-Precision*.

#### 6.4.1 Base de Dados com 100 Imagens

Tal como foi feito para os testes utilizando *query by sketch*, também para os de *query by image*, apresentados em seguida, foi utilizada primeiro a base de dados controlada de 100 imagens. Mais uma vez, a Tabela 6.3 mostra os resultados obtidos através do cálculo da medida *R-Precision*.

	Color Moments	Topogeo	Color Moments + Topogeo
<b>Bulbs</b>	1	0,273	1
<b>Car</b>	0,6	0,12	0,75
<b>Fish</b>	0,231	0,079	0,2
<b>Flower</b>	0,375	0,12	0,429
<b>Hammer</b>	0,75	0,15	0,75
<b>Magnifying Glass</b>	1	0,115	1
<b>Phone</b>	0,429	0,083	0,429
<b>Sword</b>	0,75	0,136	0,75
<b>Tree</b>	1	0,188	1
<b>TV</b>	0,091	0,073	0,091
<b>Desempenho Geral do Descritor</b>	0,6226	0,1337	0,6399

Tabela 6.3: Resultados obtidos com os testes utilizando Query by Image para a base de dados pequena (100 imagens). Os valores apresentados correspondem à medida *R-Precision* com  $N = 3$ .

Como se pode verificar na Figura 6.10, os resultados obtidos com *query by image* apresentam maiores valores de *matching* do que com *query by sketch*, um resultado natural, uma vez que as imagens são muito mais detalhadas do que esboços. Neste caso, a combinação entre os dois descritores utilizados apresenta resultados ligeiramente melhores do que qualquer dos descritores utilizados individualmente. Uma vez utilizando imagens como *input* de pesquisa, o fator cor pode ser mais apropriado. Para *query by sketch*, dado que nem sempre o utilizador consegue aproximar a cor daquilo que é real, ou mesmo pela falta de utilização de cores no esboço para pesquisa, o fator forma pode revelar-se crucial na análise do conteúdo das imagens.

Esta abordagem, de uma segunda pesquisa utilizando as imagens retornadas como uma nova *query*, pode revelar-se muito importante para que o utilizador consiga chegar ao tipo de imagens que procura. Para esta segunda abordagem de pesquisa, foi considerada a utilização da categorização de imagens, ou seja, em vez de realizar processamento sobre o conteúdo da

imagem, procurar pela categoria a que esta pertence (previamente classificada manualmente), retornando apenas imagens pertencentes a essa categoria. Porém, apesar desta metodologia ser mais concreta para o utilizador, uma vez que se é seleccionada a categoria “Car”, por exemplo, surgiriam apenas *clip arts* de carros, seria perdido o foco deste trabalho, uma vez que se pretende recuperar imagens através do seu conteúdo.

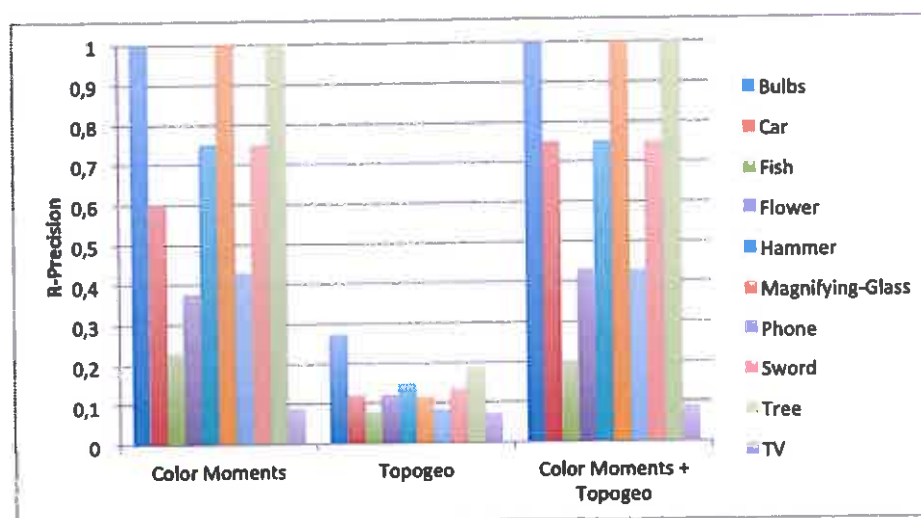


Figura 6.10: Gráfico com os resultados globais dos testes realizados com um conjunto de imagens exemplo para a base de dados de 100 imagens.

#### 6.4.2 Base de Dados com 12955 Imagens

Depois de analisar os resultados obtidos com a base de dados controlada, foi utilizada a base de dados de grandes dimensões, possibilitando assim a comparação entre resultados. A Tabela 6.4 mostra os resultados obtidos com a base de dados de 12955 imagens.

	Color Moments	Topogeo	Color Moments + Topogeo
Bulbs	0,75	0,011	1
Car	1	0,018	1
Fish	0,056	0,231	0,077
Flower	0,273	0,088	0,429
Hammer	0,033	0,002	0,094
Magnifying Glass	1	0,001	1
Phone	0,037	0,004	0,044
Sword	0,021	0,006	0,073
Tree	0,75	0,016	0,75
TV	0,429	0,011	0,5
<b>Desempenho Geral do Descritor</b>	<b>0,4349</b>	<b>0,0388</b>	<b>0,4967</b>

Tabela 6.4: Resultados obtidos com os testes utilizando Query by Image para a base de dados grande (12955 imagens). Os valores apresentados correspondem à medida R-Precision com  $N = 3$ .

Neste caso, dado que se trata de uma base de dados de grandes dimensões (12955 imagens), é notório um nível mais baixo de *matching*, no que diz respeito aos resultados apresentados anteriormente (ver Figura 6.11). Isto acontece, devido à quantidade maioritária de imagens que não pertence a nenhuma das categorias utilizadas para pesquisa (11932 imagens correspondem à categoria "Other").

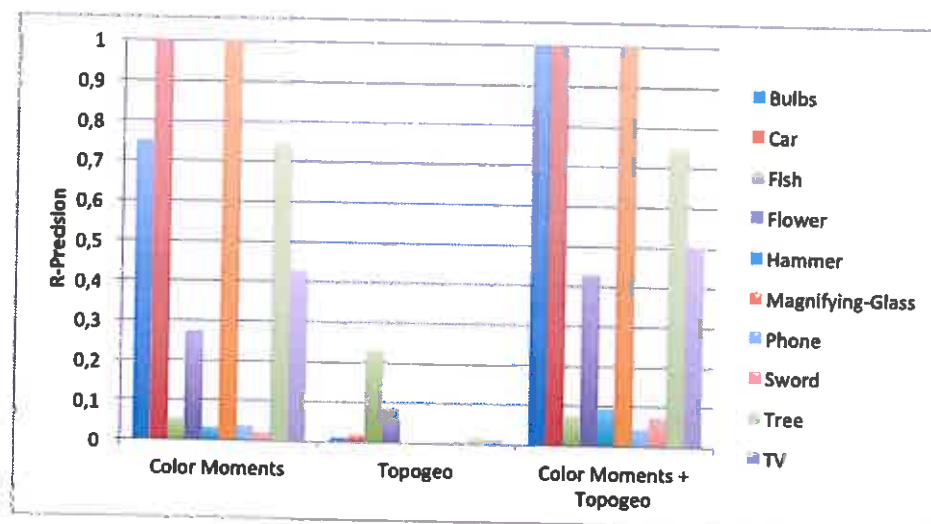


Figura 6.11: Gráfico com os resultados globais dos testes realizados com um conjunto de imagens exemplo para a base de dados de 12955 imagens.

A utilização desta base de dados pretende dar um contexto mais realista ao nível da pesquisa, onde é possível concluir que a qualidade dos resultados depende também da quantidade de imagens para cada categoria. Outro fator importante, referido também na secção anterior, é o facto de a forma nestes casos não ser totalmente discriminativa, ou seja, uma vez que existe uma grande quantidade de imagens, é natural que imagens pertencentes à mesma categoria tenham formas completamente diferentes (ver Figura 6.12), o que prejudica a análise da forma na fase de extração de informação nas imagens.



Figura 6.12: Imagem ilustrativa, apresentando formas muito diferentes em imagens clip art pertencendo à mesma categoria (e.g. "Tree").

Como é possível verificar no gráfico apresentado para as 12955 imagens, as categorias que apresentaram maior nível de *matching* são a "Car" e a "Magnifying Glass". Neste caso, ao contrário do que foi referido anteriormente, verifica-se uma grande disparidade no número de imagens destas categorias: 186 imagens para a categoria "Car" e 13 imagens na categoria "Magnifying Glass". Porém, este valor é justificável, uma vez que enquanto as imagens de carros podem ser totalmente diferentes umas das outras, as imagens referentes às lupas são muito semelhantes (ver Figura 6.13 e Figura 6.14).



Figura 6.13: Imagens ilustrativas de clip arts pertencentes à categoria "Car", mostrando a disparidade entre formas e cores.



Figura 6.14: Imagens ilustrativas de clip arts pertencentes à categoria "Magnifying Glass", mostrando a semelhança entre formas e cores.

### 6.4.3 Desempenho dos Descritores para Imagem

No final dos testes, foi mais uma vez analisado o desempenho geral ao nível das categorias e dos descritores. Como se pode verificar na Figura 6.15, para a *query by image* foi a categoria "Magnifying Glass" que apresentou maior nível de *matching*, logo seguida pela categoria "Bulbs". Este resultado é interessante, uma vez que a categoria com melhor desempenho é a que tem menor número de imagens na base de dados (13 imagens). Este desempenho, pode ser justificado pelo facto de as imagens presentes nesta categoria serem muito idênticas (em relação à cor e à forma), como foi apresentado já na Figura 6.14.

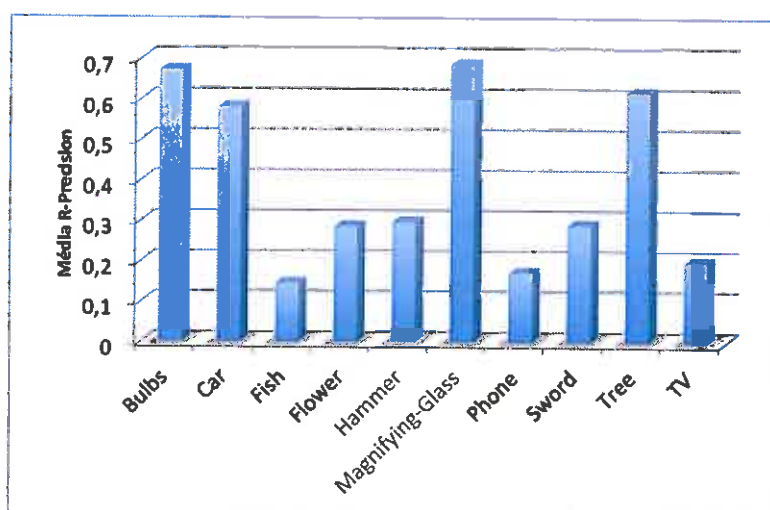


Figura 6.15: Gráfico com a estatística global de desempenho das categorias, para os testes realizados através de Query by Image, para as duas bases de dados.



Depois, foi ainda analisado o desempenho geral dos descritores para a utilização de *query by image* Figura 6.16. Tanto o descritor Momentos de Cor como a combinação deste com o Topogeo, apresentaram melhores resultados do que para o *query by sketch*. O descritor Topogeo manteve um desempenho praticamente igual em relação à pesquisa com esboço, o que significa que não traz muitas vantagens ao ser utilizado para a utilização de *query by image*. Uma abordagem posterior, para uma possível melhoria, não deveria ter em consideração este descritor, mas sim outros mais discriminativos ao nível da análise da forma em imagens.

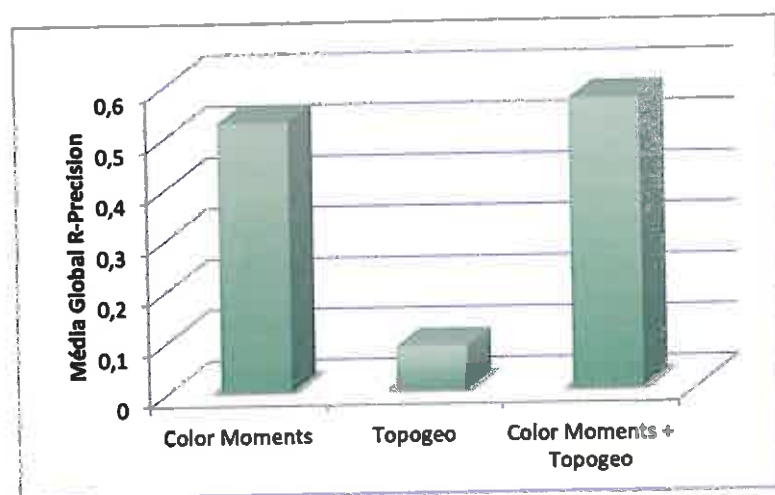


Figura 6.16: Gráfico com a estatística de desempenho global de cada um dos descritores, para os testes realizados com Query by Image.

## 6.5 Testes de Usabilidade

Esta secção apresenta os testes de usabilidade, efetuados com uma versão estável (descrita no capítulo 3) da aplicação de pesquisa no *tablet Samsung Galaxy Note*, referido acima. Estes testes pretendem avaliar a aplicação, tanto a nível de usabilidade e de interface como a nível de desempenho, tendo em conta o que o utilizador considera relevante. Primeiro, é descrita a metodologia utilizada para a construção/ definição dos testes de usabilidade. Em seguida, é feita a caracterização dos perfis dos utilizadores que testaram a aplicação. Por fim, são

apresentados os resultados obtidos e algumas conclusões. Os resultados podem ser vistos na sua totalidade através do anexos.

### **6.5.1 Método**

Os testes foram realizados no campus do ISEL, por colegas (alunos) de várias áreas científicas. Cada teste foi realizado individualmente por cada participante, com a supervisão de um observador/ facilitador. Cada um dos testes teve uma duração média de 15 minutos, que englobam a fase inicial de experimentação da aplicação e a fase de realização das tarefas propostas ao longo do questionário.

Cada teste é constituído por três fases principais:

- Especificação do tipo do utilizador, de modo a perceber se este é experiente ou não na utilização deste tipo de dispositivos móveis. Para além disso, é importante saber se a faixa etária ou área de trabalho/ estudo influencia os resultados obtidos.
- Realização de algumas tarefas pedidas ao utilizador, de modo a perceber se ele consegue interagir de forma intuitiva com a aplicação, percebendo quais as principais dificuldades sentidas e quais os pontos positivos e negativos da aplicação. A realização de tarefas está dividida em três tópicos principais: exploração da interface; avaliação de resultados; avaliação da interface.
- Considerações finais dos utilizadores em relação à interface, respondendo a perguntas de âmbito livre. Estas perguntas permitem obter a opinião geral dos utilizadores sobre a aplicação e a sua utilidade no mundo real.

No questionário, existem questões abertas e outras que visam estabelecer uma pontuação pré-estabelecida (escala tipo Likert), de modo a perceber de que forma os utilizadores avaliaram a interface e o seu desempenho. Nas questões de pontuação, a escala é feita de 1 a 5, onde 1 representa desacordo total com a afirmação ou mau desempenho e 5 representa acordo total ou excelente desempenho.

### 6.5.2 Participantes

Os testes de usabilidade foram realizados por cinco utilizadores, com idades compreendidas entre os 24 e os 51 anos. É importante realçar que nenhum destes utilizadores tinha conhecimento prévio da interface ou ferramentas da aplicação, nem das imagens *clip art* que são utilizadas na base de dados.

De uma forma geral, o primeiro impacto visual da interface no dispositivo móvel foi bastante positivo. Os participantes 1 e 2 são da área de redes e multimédia, o participante 3 é da área de telecomunicações, o 4 é da área de design e multimédia e último não tem estudos em nenhuma área científica. No total, quatro dos participantes possuem *smartphone* e/ ou *tablet* e apenas três utilizam imagens *clip art* nos seus documentos.

### 6.5.3 Resultados

Nesta secção, são detalhados e analisados os resultados obtidos através dos questionários realizados. As respostas e opções descritas pelos diferentes participantes para cada questão são analisadas, calculando para as questões de escala (1 a 5) a média ( $\mu$ ) de cada questão entre todos os participantes, de modo a perceber se existe uma tendência geral para atribuir maior ou menor pontuação. Posteriormente, é ainda calculado o desvio padrão ( $\sigma$ ) da média para avaliar a extensão do consenso em relação a cada questão.

#### 6.5.3.1 Interação

A primeira preocupação durante o desenvolvimento da interface, era criar algo que fosse intuitivo de utilizar e que despertasse o interesse do utilizador. Como tal, um dos aspetos importantes na análise de usabilidade, é perceber de que forma o utilizador interage com a aplicação e se esse método é eficaz. Uma ação comum a todos os participantes, foi tentar arrastar os *clip arts* apresentados na barra lateral (resultado à pesquisa *query by sketch*) para a área de desenho. Não sendo possibilitada essa opção, os utilizadores referiram que esta seria uma opção interessante.

De uma forma geral, os participantes conseguiram interagir facilmente com a aplicação, divertindo-se ainda com a realização de pequenos desenhos. Na avaliação da interface, as

pontuações estiveram sempre entre os valores 4 e 5. À questão “A aplicação é intuitiva e fácil de utilizar.” todos os utilizadores responderam positivamente (Média = 4,4; Desvio Padrão = 0,55). Os participantes responderam também de forma positiva em relação à disponibilização de ferramentas de desenho (Média = 4,4; Desvio Padrão = 0,55) e ao aspeto visual da interface (Média = 4,6; Desvio Padrão = 0,55). Nas Figura 6.17, Figura 6.18 e Figura 6.19 é possível verificar os resultados referidos acima e as questões utilizadas.

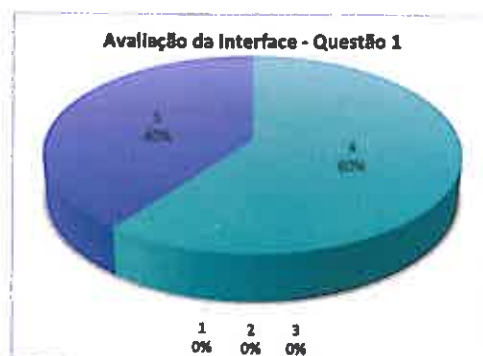


Figura 6.17: Resultados obtidos com a questão “Considero que a interface tem as ferramentas essenciais de desenho”.

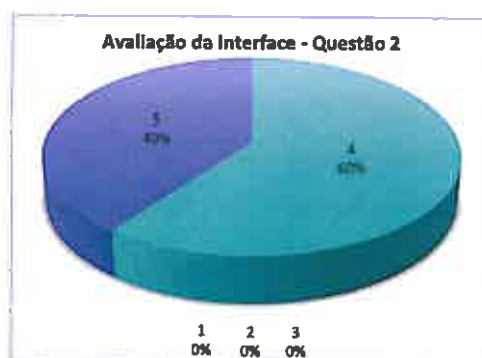


Figura 6.18: Resultados obtidos com a questão “A aplicação é intuitiva e fácil de utilizar”.



Figura 6.19: Resultados obtidos com a questão “O aspeto visual da interface é apelativo”.

#### 6.5.3.2 Desempenho

Para avaliar o desempenho do sistema, foram realizadas algumas tarefas e foi pedido aos participantes que avaliassem a qualidade dos resultados obtidos. Esta avaliação pode ser pouco conclusiva, uma vez que a avaliação do desempenho é dada por cada um dos participantes de forma subjetiva. O que para um podem ser bons resultados, tendo em conta o esboço realizado, para outro podem ser maus resultados, uma vez que surgem poucas imagens da categoria desejada. Como se pode verificar (ver Figura 6.20/Figura 6.21), a avaliação dada ao desempenho de pesquisa para uma lâmpada (Média = 3,6; Desvio Padrão = 0,55) foi superior à avaliação para a pesquisa da categoria carro (Média = 2,4; Desvio Padrão = 0,55). Esta controvérsia pode ser explicada, uma vez que se notou facilidade a nível geral dos participantes para desenharem lâmpadas e mais dificuldades para desenhar um carro. Esta avaliação foi feita propositadamente, dado que no questionário foram pedidos dois tipos de pesquisa, um mais fácil para o utilizador e outro mais difícil, para perceber o impacto que estas dificuldades podem trazer aos resultados.

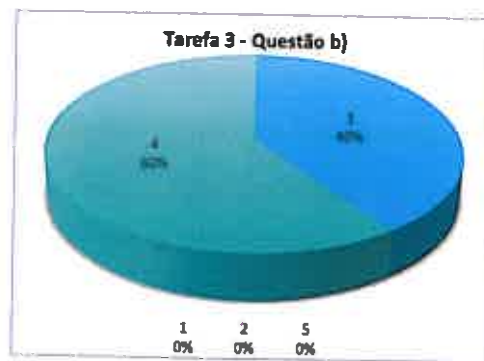


Figura 6.20: Como classifica os resultados obtidos pela query? (realizar o esboço de uma lâmpada)

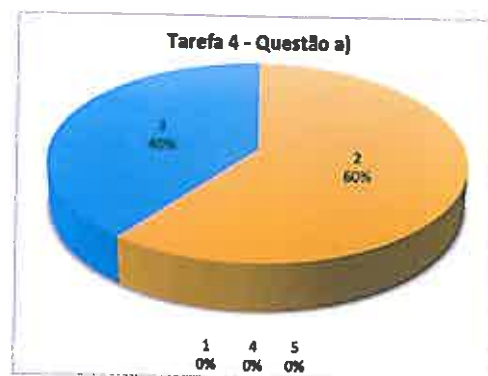


Figura 6.21: Como classifica os resultados obtidos pela query? (realizar o esboço de um carro)

### 6.5.3.3 Utilidade

Na utilização deste tipo de aplicações, é importante ter em conta o conceito de *Human Visual Perception* [10], avaliando até que ponto uma pessoa consegue desenhar aquilo que procura. Este factor pode desempenhar um papel importante no sucesso que este tipo de aplicações pode ter. De uma forma geral, os utilizadores ficaram entusiasmados com a utilização da aplicação, afirmando que se imaginavam a utilizar a aplicação. Tendo em conta os comentários feitos pelos participantes, destacam-se os seguintes:

- “É interessante ter uma interface de desenho, que permite afinar a pesquisa através de imagens exemplo, pois os resultados essenciais nunca são totalmente retornados na pesquisa feita por esboço”.

- “É uma lógica nova e interessante para pesquisar e utilizar imagens. Seria interessante poder redimensionar as imagens e fazer uma composição de várias imagens e esboço, como se estivesse a utilizar um programa de edição de imagem”.

#### 6.5.3.4 Sugestões dos Participantes

No final da realização dos testes de usabilidade, foi pedido aos utilizadores que dessem algumas sugestões relativamente à interface. Das observações feitas, foram indicadas as seguintes alterações/ melhorias:

- Uma vez que durante a utilização da aplicação as opções “Draw” ou “Erase” estão sempre seleccionadas (alternadamente) consoante o utilizador está a desenhar (formas geométricas ou outros) ou a utilizar a função de borracha, este destaque deveria ser mais visível para o utilizador;
- Os *clip arts* que surgem na barra lateral (resultado da pesquisa) deviam poder ser arrastados para o ecrã de desenho;
- Quando é seleccionada uma imagem na barra lateral de resultados, devia surgir um menu que permitisse guardar/ usar esse *clip art* ou fazer uma pesquisa mais refinada;
- Devia poder-se escolher a grossura da linha de desenho;
- Os resultados podiam surgir sempre na barra lateral, sem ser criado um novo ecrã para a pesquisa *query by image*.

Em geral, os utilizadores ficaram satisfeitos com a interface, considerando-a bastante apelativa e dinâmica. Alguns sentiram pequenas dificuldades no impacto inicial, mas depois de alguns minutos de utilização não sentiram grandes dificuldades.

Como trabalho futuro, podem ser aplicadas algumas das sugestões referidas pelos participantes, tais como adicionar mais opções no que respeita ao desenho e dar a possibilidade aos utilizadores para utilizarem as imagens *clip art* encontradas, arrastando-as para a área de desenho.

#### 6.5.4 Síntese

De uma forma geral, os participantes dos testes de usabilidade consideraram que a aplicação é intuitiva e fácil de utilizar. Para além disso, foi referido o interesse na utilização futura

deste tipo de interface, uma vez que é uma abordagem diferente para a pesquisa de *clip arts*, que pode dinamizar e divertir no processo de pesquisa.

Em relação aos resultados obtidos, os participantes ficaram satisfeitos, uma vez que os esboços mais simples (alguns deles sem cor) conseguiram mesmo assim obter imagens *clip art* relacionadas. Por outro lado, os participantes divertiram-se a utilizar a aplicação, alternando entre cores e formas, surpreendendo-se muitas vezes com o facto de os seus esboços serem reconhecidos pelo sistema de recuperação de imagem. Houve ainda algumas sugestões de alteração à interface, que foram descritas ao longo deste capítulo. Algumas alterações foram já feitas, outras mais complexas e de realização mais demorada serão feitas num trabalho futuro, de modo a completar a interface para corresponder à necessidades dos utilizadores.

## 6.6 Conclusões sobre os resultados obtidos

Depois da análise cuidada de todos os resultados obtidos, podemos concluir que de uma forma geral, o utilizador consegue com o seu esboço recuperar um conjunto de imagens que satisfaçam as suas necessidades. Na maioria dos casos, as primeiras 20 imagens retornam sempre pelo menos duas imagens consideradas satisfatórias (inseridas na categoria procurada). Tal como é descrito em [10], quando os utilizadores procuram desenhos, não estão muito preocupados com a ordem pela qual os desenhos são retornados (*precision*), mas sim que os desenhos que procuram apareçam no resultado (*recall*).

Analisando os resultados apresentados relativamente à *query by image*, é possível verificar que os resultados em relação à utilização de esboços como *query*, se alteraram ligeiramente. Quando se utilizam imagens como *query*, a informação de cor presente nas imagens torna-se um fator predominante, exceto alguns casos onde a cor não é discriminativa numa imagem ou num conjunto de imagens (ver Figura 6.22).





**Figura 6.22:** Imagens clip art representando alguns casos onde a cor não é um fator discriminativo na imagem.

Relativamente à combinação de descritores, os resultados também se apresentam satisfatórios, sendo no geral melhores do que cada descritor utilizado de forma individual.

Em primeiro lugar, este tipo de pesquisa “mista”, permite que o utilizador capte progressivamente uma aprendizagem de como os seus desenhos são interpretados, podendo realizar vários esboços no mesmo contexto, percebendo que abordagem retorna melhores resultados (com mais cor, com mais formas, ou outras características inerentes ao desenho). Em segundo lugar, é também importante para o sistema de recuperação de imagem receber um *input* mais detalhado, para poder avaliar melhor o conteúdo das imagens, retornando consequentemente melhores resultados para o utilizador. Em suma, esta mistura oferece a facilidade de desenho à mão livre para o utilizador, ensinando-o a progredir na pesquisa e a afinar os resultados obtidos.

# 7 Conclusão e Trabalho

## Futuro

No âmbito desta dissertação, foi desenvolvida uma interface de pesquisa de *clip arts* para ser utilizada em dispositivos móveis. Esta interface permite pesquisar *clip arts* através de esboços feitos pelo utilizador e imagens exemplo, para refinar os primeiros resultados obtidos. Ao longo deste trabalho, foi possível perceber que estas duas abordagens são importantes, na medida em que apenas um esboço pode não ser representativo o suficiente para satisfazer as necessidades do utilizador.

Para avaliar o sistema desenvolvido, foram realizados vários testes, que apesar de não apresentarem resultados ótimos, mostram que é possível realizar trabalho nesta área de modo a permitir a pesquisa de imagens *clip art* ou outras através do desenho de esboços. De uma forma geral, para todos os esboços realizados, tanto para os testes de desempenho como nos testes de usabilidade realizados por participantes externos ao projeto, foi sempre retornado pelo menos um resultado correspondente à categoria pesquisada (em 20 imagens). No final, foram ainda realizados alguns testes de usabilidade, que permitiram avaliar o desempenho do sistema ao nível da interação com o utilizador. Durante a realização destes testes, foi possível perceber que a reação dos utilizadores a este sistema foi positiva.

Depois de analisar os resultados obtidos, foi possível concluir que a conjugação entre *query by sketch* (15,4 % de sucesso na recuperação) e *query by image* (39,5 %) permite a obtenção melhores resultados finais. Para além disso, através do conjunto de testes “Esboço Detalhado” (21,7 %) e “Esboço Simples” (9,09 %) foi também possível perceber que os resultados através de *query by sketch* vão sempre estar relacionados também com a qualidade do esboço realizado. Relativamente à conjugação de diferentes tipos de informação extraídos nas imagens, foi possível perceber que é importante ter em conta vários tipos de características no momento da recuperação, uma vez que apenas um tipo (cor, forma, etc)

pode não ser suficientemente representativo para algumas imagens. Na utilização de dois descritores que permitem extrair informação de cor e forma, percebeu-se que o fator “cor” obteve bons resultados, principalmente quando é utilizada *query by image*. Por outro lado, o descritor de forma utilizado não apresentou os resultados esperados, mas mesmo assim concluiu-se que este tipo de informação pode ser um fator diferencial e importante quando se utiliza *query by sketch*, dado que os utilizadores muitas vezes não utilizam cores nos seus esboços. Quanto ao *feedback* obtido nos testes de usabilidade, consideramos que este estudo pode levar a uma evolução positiva, fazendo com que os utilizadores usem mais *clip arts* nos seus conteúdos, uma vez que os podem procurar de forma interativa e ao mesmo tempo divertida.

Como trabalho futuro, será importante realizar um estudo mais detalhado sobre quais as tecnologias mais adequadas para extrair informação do conteúdo de esboços, quando se pretende comparar estes com imagens (relativamente mais complexas). Esta é uma problemática importante, na medida em que comparar esboços com imagens não é uma tarefa fácil e pode ser crucial para o sucesso de aplicações deste género. A interface de pesquisa pode ainda ser melhorada, tendo em conta algumas das sugestões feitas pelos participantes dos testes de usabilidade, tais como a utilização e edição das imagens recuperadas, a visualização de resultados à medida que o utilizador vai compondo o seu esboço e a oferta de mais ferramentas de desenho (e.g. espessura do lápis e da borracha). Para além disso, seria ainda interessante realizar as alterações necessárias para permitir a utilização desta aplicação também no sistema iOS da Apple. Outra abordagem interessante para este tipo de aplicações, seria estender a base de dados utilizada para imagens armazenadas na Web, por exemplo, permitindo aos utilizadores escolher bases de dados específicas (e.g. Facebook ou Flickr). Assim sendo, seria interessante para a continuação deste trabalho, a integração deste sistema numa aplicação para criar e imprimir álbuns de fotografias editadas com *clip arts*.

# Referências Bibliográficas

- [1] Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys* , 40 (2), 1-60.
- [2] Murdock, V., Kelly, D., Croft, W. B., Belkin, N. J., & Yuan, X. (2007). Identifying and Improving Retrieval for Procedural Questions. *Information Processing and Management: an International Journal* , 43, 181-203.
- [3] Cavaco, S., Correia, N., Jesus, R., Malheiro, F., & Mateus, J. (2012). Video annotation of tv content using audiovisual information. *The 3rd International Conference on Multimedia Computing and Systems . ICMCS'12*.
- [4] Smeulders , A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 22 (12), 1349-1380 .
- [5] Rasiwasia , N., & Vasconcelos , N. (2008). Image retrieval using query by contextual example. *MIR '08 Proceedings of the 1st ACM international conference on Multimedia information retrieval* (pp. 164-171). New York, NY, USA: ACM Association for Computing Machinery .
- [6] Torres, R. S., Falcão, X. A., Zhang, B., Fan, W., Fox, E. A., Gonçalves, M. A., et al. (2005). A new framework to combine descriptors for content-based image retrieval. *CIKM Conference on Information and Knowledge Management* (pp. 335-336). New York, NY, USA: ACM Association for Computing Machinery .
- [7] Eitz, M., Hays, J., & Alexa, M. (2012). How do humans sketch objects? *ACM Transactions on Graphics (TOG) - SIGGRAPH 2012 Conference Proceedings* , 31 (4), Article No. 44 .

- [8] Chastain, S. (n.d.). *Vector and Bitmap Images, Two Types of 2D Graphics*. Retrieved May 2, 2013, from Graphics Software: <http://graphicssoft.about.com/od/aboutgraphics/a/bitmapvector.htm>
- [9] Shih, J.-L., & Chen, L.-H. (2002). Color Image Retrieval Based on Primitives of Color Moments. In S.-K. Chang, Z. Chen, & S.-Y. Lee (Ed.), *VISUAL '02 Proceedings of the 5th International Conference on Recent Advances in Visual Information Systems* (pp. 88-94). London, UK: Springer-Verlag.
- [10] Sousa, P., & Fonseca, M. J. (2009). Geometric Matching for Clip-Art Drawing Retrieval. *Journal of Visual Communication and Image Representation*, 20 (2), 1-24.
- [11] Nielsen, J., & Budiu, R. (2012). *Mobile Usability*. (T. Croom, Ed.) Berkeley: New Riders.
- [12] Te'eni, D., Carey, J. M., & Zhang, P. (2007). *Human-Computer Interaction: Developing Effective Organizational Information Systems*. New York: Wiley : John Wiley & Sons.
- [13] Jesus, R. M. (2009, September). *Recuperação de informação multimédia em memórias pessoais*. Faculdade de Ciências e Tecnologia, Lisboa.
- [14] Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 2 (1), 1-19.
- [15] Zhou, X. S., & Huang, T. S. (2003). Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, 8 (6), 536-544.
- [16] Park, J. H., & Um, B. S. (1999). A new approach to similarity retrieval of 2-D graphic objects based on dominant shapes. *Pattern Recognition Letters*, 20 (6), 591-616.
- [17] Mascio, T. D., Francesconi, M., Frigioni, D., & Tarantino, L. (2004). Tuning a CBIR system for vector images: the interface support. *AVI '04 Proceedings of the working conference on Advanced visual interfaces* (pp. 425-428). New York, NY, USA: ACM.

- [18] Bimbo, A. D., & Pala, P. (1997). Visual Image Retrieval by Elastic Matching of User Sketches. (R. Kasturi, Ed.) *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 19 (2), 121-132 .
- [19] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., et al. (1995). Query by Image and Video Content: The QBIC System. *Computer* , 28 (9), 23-32.
- [20] Hirata, K., & Kato, T. (1992). Query by Visual Example - Content based Image Retrieval. In A. Pirotte, C. Delobel, & G. Gottlob (Ed.), *EDBT '92 Proceedings of the 3rd International Conference on Extending Database Technology: Advances in Database Technology* (pp. 56-71). London, UK: Springer-Verlag.
- [21] Caduff, D., & Egenhofer, M. J. (2007). Geo-Mobile Query-by-Sketch. *International Journal of Web Engineering and Technology* , 3 (2), 157-175.
- [22] Kreuzer, R., Springmann, M., Kabary, I. A., & Schuldt, H. (2012). An interactive paper and digital pen interface for query-by-sketch image retrieval. In R. Baeza-Yates, A. P. Vries, H. Zaragoza, B. B. Cambazoglu, & V. Murdock (Ed.), *ECIR'12 Proceedings of the 34th European conference on Advances in Information Retrieval* (pp. 317-328). Berlin, Heidelberg: Springer-Verlag.
- [23] Giangreco, I., Springmann, M., Kabary, I. A., & Schuldt, H. (2012). A user interface for query-by-sketch based image retrieval with color sketches. In R. Baeza-Yates , A. P. Vries, H. Zaragoza, B. B. Cambazoglu, & V. Murdock (Ed.), *ECIR'12 Proceedings of the 34th European conference on Advances in Information Retrieval* (pp. 571-572). Berlin, Heidelberg: Springer-Verlag.
- [24] Lawler, B. P. (2002, July 22). Vector that, Charlie! A quick study in file types, history, metaphors and apple pie . San Luis Obispo, California.
- [25] Gross, M. D., & Do, E. Y. (1996). Demonstrating the electronic cocktail napkin: a paper-like interface for early design. In M. J. Tauber (Ed.), *CHI '96 Conference Companion on Human Factors in Computing Systems* (pp. 5-6). New York, NY, USA: ACM.
- [26] Yeh , T., White , B., Pedro , J. S., Katz , B., & Davis , L. S. (2011). A case for query by image and text content: searching computer help using screenshots and keywords.

- WWW '11 Proceedings of the 20th international conference on World wide web* (pp. 775-784). New York, NY, USA: ACM.
- [27] Nielsen, J. (1995, January 1). *Nielsen Norman Group*. Obtido em 7 de Março de 2013, de NN/g: <http://www.nngroup.com/articles/ten-usability-heuristics/>
  - [28] Chen, Q. (2001). *Human computer interaction: issues and challenges*. Hershey: IGI Global.
  - [29] Alshehri, F., & Freeman, M. (2012). Methods of usability evaluations of mobile devices. *23rd Australasian Conference on Information Systems* (pp. 1-10). Geelong: Deakin University.
  - [30] Hammond, T., & Davis, R. (2007). LADDER, a sketching language for user interface developers. *GRAPH International Conference on Computer Graphics and Interactive Techniques* (p. Article No. 35). New York, NY, USA: ACM.
  - [31] Alshehri, F., & Freeman, M. (2012). Methods of usability evaluations of mobile devices. In J. Lamp (Ed.), *Australasian Conference on Information Systems* (pp. 1-10). Geelong, Victoria.
  - [32] Zhang, D., & Adipat, B. (2005). Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction* , 18 (3), 293-308.
  - [33] Ham, D.-H., Heo, J., Fossick, P., Wong, W., Park, S., Song, C., et al. (2008). *Framework and Model of Usability Factors of Mobile Phones*. IGI Global.
  - [34] Pimenta, N. (May de 2008). Photofinder 2: Classificação e pesquisa de fotografias digi- tais.
  - [35] Sergyán, S. (2007, January). Color Content-based Image Classification. *5th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence and Informatics* . Poprad, Slovakia.
  - [36] Martins, P. F. (2012, December). Pesquisa de Clip Arts Combinando Imagens Raster e Vectoriais. Master's thesis, Lisboa, Portugal.
  - [37] Fonseca, M. J., Pimentel, C., & Jorge, J. A. (2002). CALI: An Online Scribble Recognizer for Calligraphic Interfaces. *Sketch Understanding, Papers from the 2002 AAAI Spring Symposium* , 51-58.

- [38] Sousa, P., & Fonseca, M. J. (2010). Sketch-based retrieval of drawings using spatial proximity. *Journal of Visual Languages and Computing* , 21 (2), 69-80.
- [39] Paul E. Black, *Manhattan distance*, in *Dictionary of Algorithms and Data Structures* [online], Vreda Pieterse and Paul E. Black, eds. 31 May 2006. (accessed April 2013) Available from: <http://www.nist.gov/dads/HTML/manhattanDistance.html>.
- [40] Paul E. Black, *Euclidean distance*, in *Dictionary of Algorithms and Data Structures* [online], Vreda Pieterse and Paul E. Black, eds. 17 December 2004. (accessed April 2013) Available from: <http://www.nist.gov/dads/HTML/euclidndstnc.html>.
- [41] Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *ICML '06 Proceedings of the 23rd international conference on Machine learning* (pp. 233-240). New York, NY, USA: ACM.
- [42] Mourato, A. S. (2013). Clip art retrieval using a sketch Tablet application. *Conference on Electronics, Telecommunications and Computers - CETC 2013*. Lisboa: Elsevier's PROCEAIA TECHNOLOGY.



# 8 Anexos

## 8.1 Questionário

### Aplicação MyDraw

A aplicação Draw My Clip Arts destina-se à pesquisa de imagens *clip art*, através da realização de esboços feitos pelo utilizador. Pretende-se que esta interface permita aos utilizadores desenhar esboços e visualizar os resultados à sua *query*. Esta aplicação oferece as ferramentas básicas de desenho, que permitem desenhar esboços com a máxima liberdade possível. É ainda possível navegar nos resultados, de modo a visualizar todos os *clip arts* retornados.

Através deste estudo, pretendemos avaliar a usabilidade e desempenho do sistema. Como tal, pedimos-lhe que realize as tarefas abaixo enumeradas, indicando ainda nas questões de âmbito livre quais foram as suas maiores dificuldades e quais os pontos positivos em relação à utilização da aplicação.

Obrigado pela sua disponibilidade!

#### Dados Pessoais:

Idade: \_\_\_\_\_

Género: \_\_\_\_\_(M/F)

1.1 Possui algum smartphone ou tablet? \_\_\_\_\_

1.2 Utiliza esse dispositivo diariamente? \_\_\_\_\_

1.3 Utiliza o seu dispositivo móvel para trabalhar? \_\_\_\_\_

1.4 Costuma produzir documentos ou conteúdos? \_\_\_\_\_

1.5 Se sim, costuma utilizar imagens ou *clip arts* nos seus conteúdos? \_\_\_\_\_

1.6 Quando pretende pesquisar por um *clip art*, que passos costuma efetuar? \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

1.7 Quando pesquisa *clip arts*, que método utiliza (palavra-chave, categoria, percorrer a lista, outros)? Esse método, é eficiente? \_\_\_\_\_

\_\_\_\_\_

### **Tarefas:**

Por favor, efetue as seguintes tarefas e descreva-nos os resultados obtidos.

#### ***Exploração da interface***

**Tarefa 1:** Explore a aplicação e faça um esboço do que quiser.

a) Que dificuldades encontrou para cumprir a tarefa?

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**Tarefa 2:** Selecione uma cor, desenhe algumas formas geométricas e apague, primeiro uma pequena parte do desenho e depois o desenho por completo.

a) Que dificuldades encontrou para cumprir a tarefa?

---

---

---

### ***Avaliação de Resultados***

**Tarefa 3:** Imagine que quer procurar por *clip arts* de lâmpadas. Realize um esboço e efetue uma pesquisa.

a) Que dificuldades encontrou para cumprir a tarefa?

---

---

---

b) Como classifica os resultados obtidos pela *query*?

**Mau**

**Excelente**

**1**

**2**

**3**

**4**

**5**

Tente afinar a sua pesquisa, utilizando uma das imagens retornadas.

c) Que dificuldades encontrou para cumprir a tarefa?

---

---

---

d) Como classifica os resultados obtidos pela *query*?

**Mau**

**Excelente**

1

2

3

4

5

**Tarefa 4:** Imagine que pretende procurar por clip arts de carros. Realize um esboço e efetue uma pesquisa.

1. Como classifica os resultados obtidos pela *query*?

**Mau**

**Excelente**

1

2

3

4

5

Tente afinar a sua pesquisa, utilizando uma das imagens retornadas.

2. Como classifica os resultados obtidos pela *query*?

**Mau**

**Excelente**

1

2

3

4

5

### ***Avaliação da Interface***

Responda às seguintes questões, desenhando um círculo à volta do número que mais se aproxima da sua opinião.

1. Considero que a interface tem as ferramentas essenciais de desenho.

**Discordo**

**Concordo**

1

2

3

4

5

2. A aplicação é intuitiva e fácil de utilizar.

**Discordo**

**Concordo**

1

2

3

4

5

3. O aspeto visual da interface é apelativo.

**Discordo**

**Concordo**

1

2

3

4

5

4. Eu utilizaria esta aplicação para complementar os meus conteúdos.

**Discordo**

**Concordo**

1

2

3

4

5

**Questões gerais, relacionadas com as tarefas anteriores:**

1. Quais foram as maiores dificuldades sentidas para cumprir as tarefas?

---

---

---

2. Na sua opinião, considera o esboço um bom método para pesquisar *clip arts*?

---

---

---

3. Considera que é interessante utilizar o esboço numa primeira abordagem e posteriormente uma das imagens retornadas como uma nova *query*? Porquê?

---

---

---

4. Na sua opinião, a apresentação de resultados na barra lateral, quando a *query* é o esboço, é adequada?

Inadequado

Adequado

1

2

3

4

5

5. Na sua opinião, é adequada a apresentação de um novo ecrã, quando a pesquisa é feita através de um dos *clip arts* retornados?

Inadequado

Adequado

1

2

3

4

5

6. Observações finais.

---

---

---

## 8.2 Resultados

Utilizador	Idade	Género	Q1.1	Q1.3	Q1.5	Q1.7
1	26	M	Sim	Não	Às vezes	Palavra-chave no Google. Sim.
2	28	M	Sim	Sim, ver e-mails	Não	-
3	24	F	Sim	Não	Às vezes	Categoria. Sim.
4	24	F	Não	-	Sim, para ilustrações	Palavra-chave (em inglês) no Google. Sim.
5	51	F	Sim	Não	Não	Palavra-chave ou percorrer listas. Sim.

Tabela 8.1: Tabela exemplificativa de respostas às perguntas introdutórias do questionário.

Utilizador	Tarefa 1 – a)	Tarefa 2 – a)
1	Depois de utilizar formas geométricas, dificuldade em voltar ao desenho livre.	Depois de fazer Erase, tenho de pôr no Draw para ter a opção Clear clicável.
2	Nenhuma.	Nenhuma.
3	Nenhuma.	Depois de fazer Erase, tenho de seleccionar o Draw para poder apagar o desenho todo.
4	Quando utilizo a opção Erase, tenho de seleccionar Draw primeiro para usar formas geométricas. É um clique a mais.	Perdi-me a perceber como executar a tarefa, não vi logo a opção Clear.
5	Nenhuma.	Nenhuma.

Tabela 8.2: Tabela com exemplos de respostas dos utilizadores, ao realizar as tarefas 1 e 2.

Utilizador	Tarefa 3 – b)	Tarefa 3 – c)	Tarefa 3 – d)
1	4	Perceber que é necessário clicar sobre uma das imagens na barra lateral.	3
2	4	Não é muito intuitivo.	3
3	4	Nenhuma.	3
4	3	Nenhuma.	4
5	3	Perceber que se pode clicar na imagem para fazer uma segunda pesquisa, utilizando a imagem.	3
$\mu$	3,6	-	3,2
$\sigma$	0,548	-	0,447

Tabela 8.3: Tabela exemplificativa de respostas dos utilizadores à tarefa 3, questões b), c) e d).

Utilizador	Tarefa 4 – a)	Tarefa 4 – b)
1	2	4
2	3	3
3	3	3
4	2	4
5	2	3
$\mu$	2,4	3,4
$\sigma$	0,548	0,548

Tabela 8.4: Tabela exemplificativa das respostas dos utilizadores às questões a) e b), pertencentes à tarefa 4.



Utilizador	Q1	Q2	Q3	Q4
1	5	4	5	4
2	4	4	4	3
3	4	5	5	4
4	4	4	5	3
5	5	5	4	4
$\mu$	4,4	4,4	4,6	3,6
$\sigma$	0,548	0,548	0,548	0,548

Tabela 8.5: Tabela exemplificativa das respostas dos utilizadores às questões para avaliação da interface.

Utilizador	Q1	Q4	Q5	Q6
1	Dificuldade em sair do ecrã de imagens mostrado na <i>query by image</i> .	Sim. Para encontrar mais resultados, até chegar ao que procuro.	4	5
2	Ir ao Draw para apagar tudo, ou desenhar formas.	Sim. Porque os resultados essenciais nunca são retornados totalmente na pesquisa por esboço.	4	4
3	Perceber como se efetuava a pesquisa utilizando um dos <i>clip arts</i> .	Sim. Porque dá mais possibilidades.	5	5
4	Os resultados do esboço têm uma lógica vertical, e depois no ecrã inteiro têm uma lógica horizontal (ordem de proximidade das imagens). Cria alguma confusão.	Sim. É uma lógica nova e interessante para comparar os resultados.	5	4
5	Na segunda pesquisa, utilizando uma das imagens retornadas. Não percebi que se podia fazer <i>drop down</i> na lista de resultados.	Sim. Porque torna a aplicação mais interativa.	4	4
$\mu$	-	-	4,4	4,4
$\sigma$	-	-	0,5478	0,548

Tabela 8.6: Tabela exemplificativa das respostas dos utilizadores, às questões finais relacionadas com o teste de usabilidade realizado.

Utilizador	Q8
1	Destacar que o Draw fica selecionado quando o utilizador está a desenhar e que o Erase está selecionado quando o utilizador está a apagar com a borracha.
2	Devia poder-se arrastar os <i>clip arts</i> para o ecrã de desenho. Quado se clica numa imagem da barra lateral, devia surgir um menu de opções: usar <i>clip art</i> e refinar pesquisa. Devia poder escolher a grossura da linha de desenho.
3	Para apresentação dos resultados, agrada-me mais a barra lateral do que o ecrã completo. Podia ser sempre assim.
4	Sinto falta de opções de grossura do lápis. Seria interessante poder desenhar outros polígonos. Gostava de editar as imagens retornadas, compondo um desenho com várias. Era interessante poder digitar texto, na lógica de editar desenhos. Se as imagens pudessem ser usadas, seria interessante poder redimensioná-las.
5	Gostava de poder utilizar uma paleta de cores mais variada.

**Tabela 8.7:** Exemplos de observações finais dadas pelos utilizadores no fim de realizar o teste de usabilidade com a aplicação.

### 8.3 Gráficos Síntese

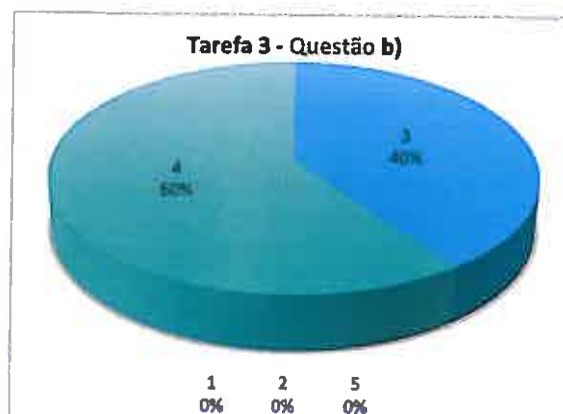


Figura 8.1: Como classifica os resultados obtidos pela query? (realizar o esboço de uma lâmpada)

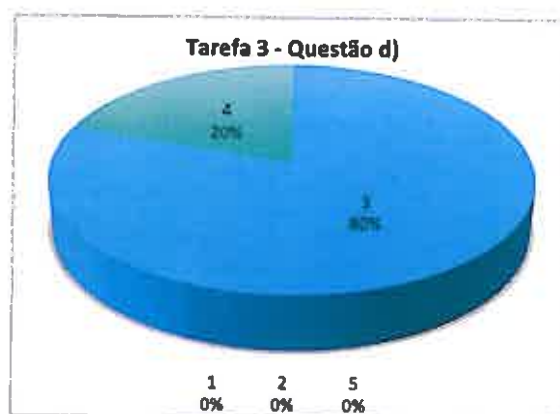


Figura 8.2: Como classifica os resultados obtidos pela query? (refinar a pesquisa com a imagem de uma lâmpada)

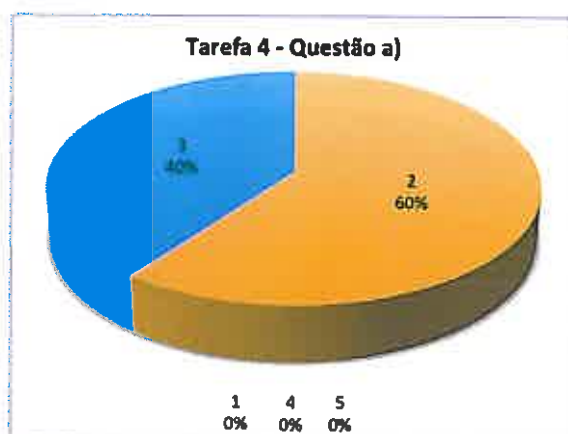


Figura 8.3: Como classifica os resultados obtidos pela query? (realizar o esboço de um carro)

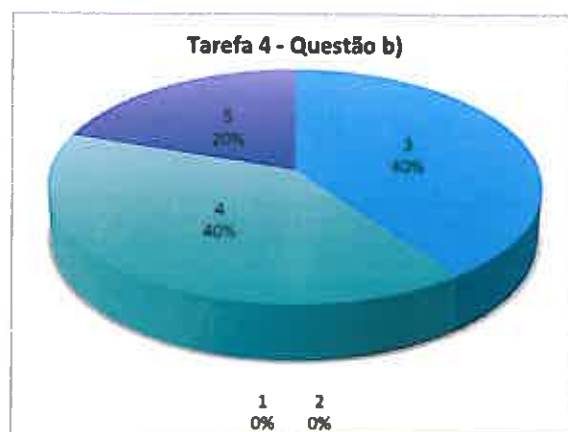


Figura 8.4: Como classifica os resultados obtidos pela query? (refinar a pesquisa com a imagem de uma lâmpada)

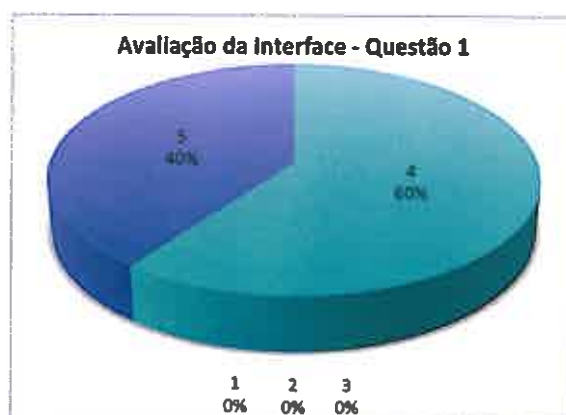
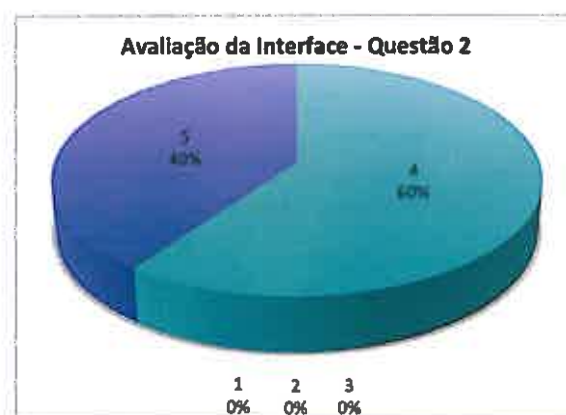
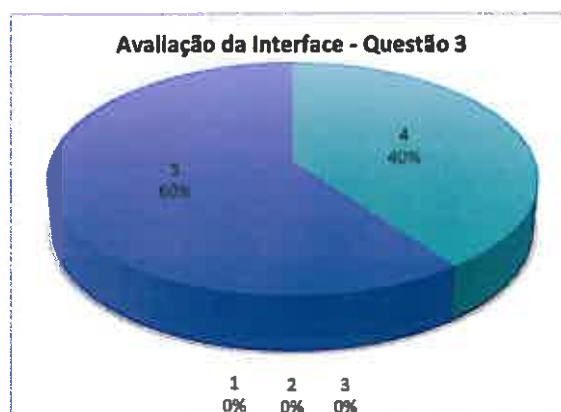


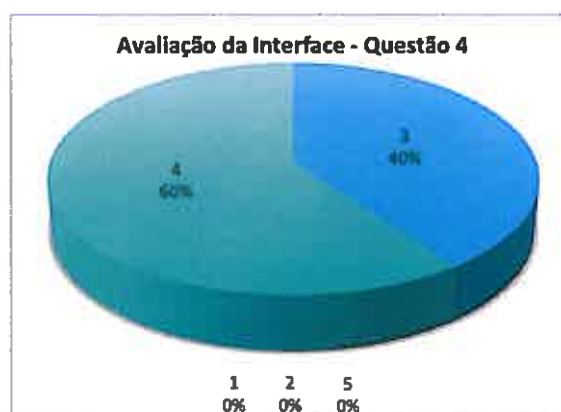
Figura 8.5: Considero que a interface tem as ferramentas essenciais de desenho.



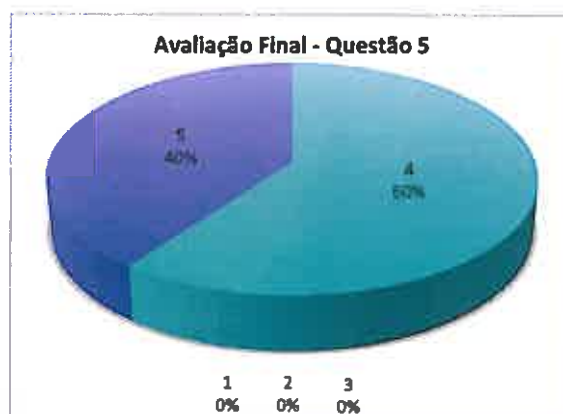
**Figura 8.6:** A aplicação é intuitiva e fácil de utilizar.



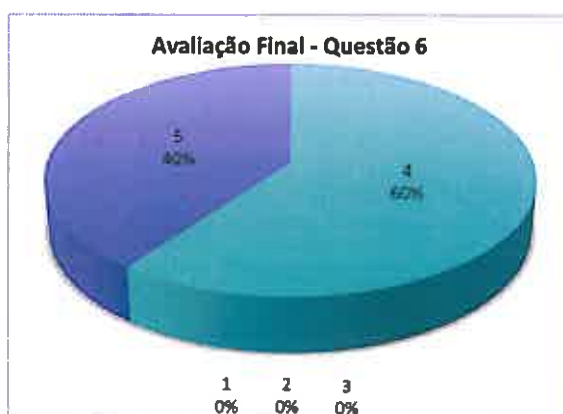
**Figura 8.7:** O aspeto visual da interface é apelativo.



**Figura 8.8:** Eu utilizaria esta aplicação para complementar os meus conteúdos.



**Figura 8.9:** Na sua opinião, a apresentação de resultados na barra lateral, quando a query é o esboço, é adequada?



**Figura 8.10:** Na sua opinião, é adequada a apresentação de um novo ecrã, quando a pesquisa é feita através de um dos clip arts retornados?

Cot.: MOU, 004(043)  
Tit.: Aplicação para pesquisa de clip  
Aut.: MOURATO, Ana Sofia Brás Carrilho  
Cód.: 1026499

