

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



**FEUP**

# **Categorização de imagens e pesquisa de base de dados através de exemplos**

**Christophe Rodrigues da Silva**

Mestrado Integrado em Engenharia Electrotécnica e Computadores

Orientador: Aurélio Campilho (Professor Catedrático)

Co-orientador: António Monteiro (Professor Auxiliar)

Julho de 2009



# Resumo

Com o constante crescimento da informação que nos rodeia, cresce também a necessidade de criação de ferramentas de pesquisa. A categorização de imagens através de exemplos e pesquisa de base de dados (*content-based image retrieval*) é uma área importante devido aos desafios que enfrenta.

Nesta dissertação são propostas duas abordagens para a categorização de imagens. Nas duas abordagens, a primeira fase consiste na medição de características e na criação correspondente de um vector de características representativo da imagem. São efectuadas medições da textura, da cor e da forma. Para a caracterização da textura e da cor são calculados previamente *textons* como elementos contendo informação local em vizinhanças de  $2 \times 2$ . A medição da forma é obtida a partir dos histogramas do módulo e da fase dos pontos do contorno da imagem.

Na primeira abordagem mede-se a similitude de uma imagem com as imagens da base de dados, retornando aquelas que são mais semelhantes. A medição de similitude pode ter duas fases, a primeira usando as características globais das imagens a segunda usando características locais, a fim de refinar a pesquisa.

A segunda abordagem consiste em classificar a imagem exemplo a fim de lhe atribuir previamente uma classe. A medição de similitude é posteriormente realizada somente com imagens da mesma classe da imagem de exemplo.

A qualidade da pesquisa foi avaliada, obtendo-se num total de 100 testes uma precisão máxima de 62% e um erro de classificação médio de 5.99%.



# Abstract

With the constant growing of the surrounding information, the need of creating searching tools rises. Content-based image retrieval is an important and challenging research area.

In this dissertation it is proposed two approaches for content-based image retrieval. In both approaches, the first phase consist in feature measurement and in the creation of a representative feature vector of the query image. The measures are made to texture, colour and shape. For the measures of texture and colour first it's computed textons as elements with local information in a neighbourhood of size  $2 \times 2$  . The shape measure is obtained by computing histograms of phase and module of the points in the image contour.

In the first approach the similarity measure is measured with all the images in the database, retrieving the most similar. The similarity measure can have two phases, the first using the global features of images, and the second using local features, with the objective of improve the search.

The second approach consists in classifying the query image so that a class related to it. Next the measure of similarity is made only with the images of the same class as query.

The quality of the retrieval was evaluated, it obtained in a total of 100 tests a maximum precision of 62% and an average classification error of 5.99%.



# Agradecimentos

Gostaria de agradecer ao meu orientador Professor Aurélio Campilho e co-orientador Professor António Monteiro pela ajuda e orientação prestada ao longo deste semestre.

Gostaria também de agradecer ao Instituto Nacional de Engenharia Biomédica pela disponibilização de um espaço de trabalho.

Por fim agradeço a todas as pessoas que estiveram do meu lado durante o desenvolvimento desta dissertação, pelo o apoio e incentivos.

Christophe Silva



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Visão geral do sistema . . . . .	1
1.2	Objectivos e contribuições . . . . .	4
1.3	Estrutura da dissertação . . . . .	4
<b>2</b>	<b>Estado da arte</b>	<b>5</b>
2.1	Introdução . . . . .	5
2.2	Aplicações . . . . .	6
2.3	Sistemas CBIR . . . . .	6
2.4	Análise de imagem . . . . .	7
2.5	Classificação e agrupamento . . . . .	8
2.6	Medidas de distância e semelhança entre imagens . . . . .	9
2.7	Medida de relevância . . . . .	9
2.8	Sumário . . . . .	10
<b>3</b>	<b>Medição de características</b>	<b>13</b>
3.1	Introdução . . . . .	13
3.2	Textura . . . . .	14
3.3	Cor . . . . .	18
3.4	Contorno . . . . .	19
3.5	Sumário . . . . .	21
<b>4</b>	<b>Classificação e Redução de Dimensionalidade</b>	<b>23</b>
4.1	Classificadores Bayesianos . . . . .	23
4.1.1	Classificador Linear . . . . .	24
4.1.2	Classificador Quadrático . . . . .	26
4.2	Classificador K vizinhos mais próximos (KNN) . . . . .	26
4.3	Classificador baseado em Máquinas de Vectors de Suporte (SVM) . . . . .	27
4.4	Combinação de classificadores por votação . . . . .	30
4.5	Redução de Dimensionalidade . . . . .	30
4.5.1	Análise de Componentes Principais . . . . .	30
4.5.2	Análise Discriminante de Fisher . . . . .	32
4.6	Sumário . . . . .	33
<b>5</b>	<b>Metodologia de pesquisa da base de dados</b>	<b>35</b>
5.1	Ordenação: semelhança entre imagens . . . . .	35
5.1.1	Regiões de interesse . . . . .	35
5.1.2	Distâncias . . . . .	38

5.1.3	Ordenação . . . . .	39
5.2	Classificação e ordenação . . . . .	40
5.2.1	Classificação . . . . .	40
5.3	Relevância . . . . .	41
5.4	Sumário . . . . .	42
<b>6</b>	<b>Análise de Resultados</b>	<b>45</b>
6.1	Exemplos . . . . .	45
6.1.1	Relevância . . . . .	48
6.2	Avaliação Quantitativa . . . . .	49
6.3	Sumário . . . . .	50
<b>7</b>	<b>Conclusões e Trabalho Futuro</b>	<b>51</b>
7.1	Trabalho Futuro . . . . .	52
	<b>Referências</b>	<b>53</b>

# Lista de Figuras

1.1	Abordagem de ordenação global. . . . .	2
1.2	Abordagem de ordenação parcial . . . . .	3
2.1	Arquitectura de sistemas CBIR. . . . .	6
2.2	Diagrama de blocos do sistema CBIR de [1] . . . . .	8
3.1	Exemplo de uma imagem médica da base de dados. . . . .	14
3.2	Imagem de gradiente $G_{max}(x,y)$ obtida apartir da imagem da Fig. 3.1 . . . . .	15
3.3	$t_0$ Estrutura de <i>texton</i> . $t_{1-5}$ <i>Textons</i> utilizados . . . . .	15
3.4	Imagem <i>texton</i> de $G_{max}(x,y)$ , Fig. 3.2 . . . . .	17
3.5	Imagem ilustrativa dos <i>offsets</i> utilizados para calcular a matriz de co-ocorrências [2]. . . . .	17
3.6	(a) Módulo do gradiente no contorno da imagem Fig.3.1. (b) Fase do gradiente no contorno da imagem Fig.3.1. (c)(d) Histogramas de módulo e fase das imagens (a) e (b), respectivamente, com diferentes intervalos. . . . .	20
4.1	Exemplo da classificação do LDC para duas classes. . . . .	25
4.2	Exemplo da classificação do QDC para duas classes. . . . .	26
4.3	Exemplo da classificação do KNN para duas classes e $K = 5$ . . . . .	27
4.4	Hiperplano de separação para um conjunto de treino de duas dimensões . . . . .	28
4.5	Exemplo da classificação do SVM para duas classes. . . . .	29
4.6	Diagrama de combinação de classificadores. . . . .	30
5.1	Divisão da imagem em quatro quadrantes. . . . .	36
5.2	As 10 Regiões de interesse da imagem. Os pontos a vermelho são obtidos pelo algoritmo de Harris, os restantes pelo máximo do gradiente da imagem. . . . .	37
5.3	Ilustração do uso de uma métrica como medida de similitude. . . . .	39
5.4	Ilustração do uso de uma métrica como medida de similitude entre RDIs de duas imagens. . . . .	40
5.5	Imagens pertencentes a classe A. (1) Estrutura óssea de um pé. (2) Espinha lombar. (3) Mama. . . . .	41
5.6	Imagens pertencentes a classe B, pulmões. . . . .	41
5.7	Imagens pertencentes a classe C, pulmões posição sagital. . . . .	41
5.8	Exemplo ilustrativo do deslocamento do ponto de consulta. . . . .	42
6.1	Exemplo 1 - obtenção de imagens semelhantes . . . . .	45
6.2	Exemplo 2 - obtenção de imagens semelhantes . . . . .	46
6.3	Exemplo 3 - obtenção de imagens semelhantes . . . . .	46
6.4	Exemplo 4 - obtenção de imagens semelhantes . . . . .	47

6.5	Exemplo 5 - Erro de classificação na obtenção das imagens mais semelhantes . . .	48
6.6	Exemplo 6 - Relevância . . . . .	48
6.7	Gráficos box-plot de precisão do sistemas em função do número de imagens devolvidas pelo sistema. . . . .	49

# Lista de Tabelas

3.1	Tabela de <i>offsets</i> em função do ângulo $\gamma$ e a distância $d$ . . . . .	17
6.1	Tabela de valores médios da precisão do sistema . . . . .	49
6.2	Erros médios dos classificadores . . . . .	50
6.3	Número médio de erros por classe. . . . .	50



# Abreviaturas e Símbolos

CBIR	<i>Content based image retrieval</i>
BING	<i>Brain Image Network Grid</i>
INEB	Instituto de Engenharia Biomédica
HSV	<i>Hue Saturation Value</i>
CIE	<i>Commission Internationale de l'Eclairage</i> - Comissão Internacional da Iluminação
CIELab	Modelo de cor da CIE. $L^*$ luminância, $a^*$ e $b^*$ coordenadas de cor.
RGB	Vermelho (Red), Verde (Green) e Azul (Blue)
LUV	$L^*$ luminância, $u^*$ e $v^*$ coordenadas de cromaticidade
DWF	<i>Discrete Wavelet Frames</i>
LDC	<i>Linear Discriminant Classifier</i>
QDC	<i>Quadratic Discriminant Classifier</i>
KNN	<i>K-nearest Neighbour classifier</i>
SVM	<i>Support Vector Machines</i>
HLM	<i>hypothetical local mapping</i>
PCA	<i>Principal Component Analysis</i>
RDI	Regiões de Interesse

$\sigma$  Desvio padrão

\* Operador de convolução



# Capítulo 1

## Introdução

Com o crescimento das tecnologias de comunicação e de multimédia [3], a pesquisa de imagem em grandes bases de dados é uma área em franco desenvolvimento.

Em várias áreas de actividade como, por exemplo, comerciais, militares, governamentais, académicas e de saúde constroem-se colecções de imagens digitais de grande dimensão, tendo em vista o seu arquivo e proporcionando consultas posteriores. Estas colecções, ou bases de dados, criam oportunidades para o desenvolvimento de sistemas de informação, envolvendo a consulta e a categorização de imagens através de exemplos, sistemas de *content-based image retrieval* (CBIR).

Os sistemas CBIR permitem interrogar uma base de dados com uma imagem, retornando como resultado imagens similares à imagem exemplo. Cada imagem é representada por um vector de características medidas na própria imagem, tais como a cor, a textura e a forma. No final, a semelhança entre vectores é comparada através de uma função de similitude e são retornadas as imagens mais próximas.

O Instituto de Engenharia Biomédica (INEB) em parceria com outras universidades do país, tem em desenvolvimento uma rede para a partilha de imagens médicas e de vários programas de processamento e análise de imagem, BING<sup>1</sup> (*Brain Image Network Grid*). A pesquisa de imagens médicas baseada nas características de imagem é uma ferramenta útil para este projecto, sendo uma das motivações para esta dissertação.

O interesse e fascínio pela área de processamento de imagem e das suas aplicações, serviu de motivação pessoal para a selecção deste tema de trabalho.

### 1.1 Visão geral do sistema

Nesta Secção pretende-se dar uma visão global do sistema desenvolvido no âmbito desta dissertação.

---

<sup>1</sup><http://www.brainimaging.pt/>

Foi criado um sistema automático que permite a obtenção de imagens similares a uma imagem exemplo, um sistema CBIR. Foram adoptadas duas abordagens: a abordagem de ordenação global, envolvendo uma simples ordenação dos resultados por similitude, envolvendo a comparação com todas as imagens da base de dados; a abordagem de ordenação parcial, envolvendo a classificação da imagem exemplo e ordenação dos resultados por similitude, considerando apenas as imagens da classe atribuída à imagem exemplo.

Na abordagem de ordenação global, são medidas na imagem de entrada várias características a fim de formar um vector representativo da imagem, contendo informação quantitativa diversificada, por exemplo sobre a cor, a forma ou a textura.

A base de dados contém  $N$  vectores, cada um representando uma dada imagem. Esses vectores serão comparados com o vector da imagem de entrada, a imagem exemplo, a fim de determinar quais as imagens mais semelhantes. Os vectores são ordenados por semelhança de acordo com uma dada função de similitude. No final, o sistema apresenta ao utilizador um número determinado de imagens que de acordo com a função de similitude se encontravam mais próximas da imagem exemplo.

Esta abordagem está ilustrada na Fig. 1.1.

O bloco de Medição de Características é responsável pela medição das características da imagem de entrada.

O bloco de Ordenação, realiza a medição de similitude entre o vector de características da imagem de entrada e os vectores das imagens da base de dados, após realizadas as medições, as imagens da base de dados são ordenadas por similitude e mostradas ao utilizador.

Na ordenação é possível incorporar informação local, através do uso de vectores de características representativas de regiões de interesse da imagem. Estes vectores são calculados no bloco de medição de características. São escolhidas  $k$  regiões de interesse e calculados os respectivos vectores de características a fim de representar cada uma das regiões.

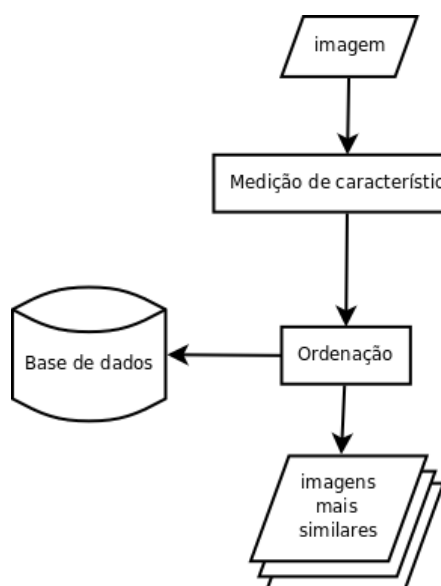


Figura 1.1: Abordagem de ordenação global.

A segunda abordagem, tem uma fase de classificação e outra de relevância. A base de dados está dividida por classes, com imagens previamente etiquetadas. Como na primeira abordagem extraem-se inicialmente um vector de características que é usado na classificação com o objectivo de evitar a medição de similitude com todas as imagens da base de dados, medindo apenas a similitude entre imagens da mesma classe. A relevância permite interactividade com o utilizador, permitindo refinar a pesquisa em busca de melhores resultados.

Esta abordagem, a abordagem de ordenação parcial, está ilustrada na Fig. 1.2, em que o bloco de Medição de Características tem a função já referida.

O bloco de Classificação tem como objectivo atribuir à imagem de entrada uma classe.

O bloco de Ordenação tem como parâmetros de entrada o vector de características da imagem de entrada assim como a classe atribuída ( $e_c$ ). Este bloco mede a similitude entre as imagens da classe  $e_c$  da base de dados com a imagem de entrada. Após as medições, as imagens são ordenadas por similitude e apresentadas ao utilizador.

O bloco de relevância possibilita ao utilizador a interacção com o sistema através da indicação da relevância dos resultados, permitindo uma pesquisa personalizada e refinada.

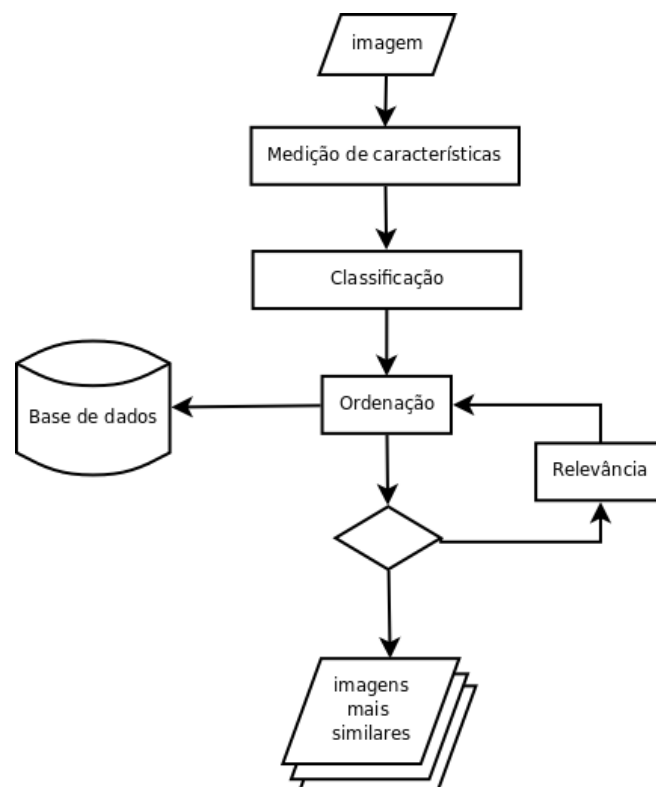


Figura 1.2: Abordagem de ordenação parcial

A abordagem de ordenação global não necessita que as imagens sejam previamente classificadas, dependendo apenas das características medidas nas imagens. Esta abordagem permite a utilização de informação local da imagem para a ordenação das imagens, a fim de refinar os resultados baseando-se em características locais da imagem.

A abordagem de ordenação parcial requer uma base de dados em que as imagens tenham uma classe atribuída *a priori*. É necessário treinar os classificadores, estimando as distribuições e calculando estatísticas por classe. Esta abordagem permite a diminuição do número de imagens a comparar com a imagem de entrada. É uma abordagem de duas fases, a primeira de classificação e a segunda de ordenação.

Nas duas abordagens o bloco de medição de relevância pode ser integrado. No entanto impossibilita a integração da utilização de informação local das imagens na ordenação, pois o método de medição de relevância é realizado considerando apenas características globais das imagens.

## 1.2 Objectivos e contribuições

O desenvolvimento de um sistema de categorização de imagens e pesquisa em base de dados é o objectivo desta dissertação. O sistema tem como base um conjunto de características extraídas de imagens e o cálculo de similitude entre imagens usando uma métrica apropriada e organização dos dados num conjunto de classes. As principais contribuições desta dissertação são:

- Desenvolvimento de métodos de medição de características, em particular de um método de medição de características baseadas no contorno de uma imagem.
- Desenvolvimento de algoritmos de classificação, em particular de um classificador multi-classe baseado em combinação de classificadores.
- Desenvolvimento de métodos de ordenação de imagens por similitude, através de características globais ou locais.
- Desenvolvimento de um novo método de detecção de regiões de interesse de uma imagem médica.

## 1.3 Estrutura da dissertação

Para além desta introdução, no Capítulo 2 são analisados alguns artigos relevantes que se enquadram no tema desta dissertação. É realizado um levantamento das técnicas de medição de características, classificação e agrupamento, medidas de similitude e sistemas de medição de relevância. No Capítulo 3 são descritas as técnicas utilizadas para a medição de características das imagens. As técnicas de classificação utilizadas são apresentadas no capítulo 4. No capítulo 5 é explicado os métodos de pesquisa na base de dados. No Capítulo 6 são discutidos alguns resultados obtidos. Por último, no capítulo 7 são apresentadas conclusões e perspectiva-se o trabalho futuro.

## Capítulo 2

# Estado da arte

Neste capítulo é apresentado o estado da arte descrevendo algumas contribuições apresentadas recentemente em CBIR.

Na Secção 2.1 é feita uma introdução. Na segunda Secção são referidas algumas aplicações desta área. A arquitectura de sistemas CBIR é apresentado na Secção 2.3. Na Secção 2.4, são referidas várias características de imagens usadas em CBIR. Na Secção 2.5 são analisadas técnicas de agrupamento e de classificação frequentemente utilizadas. Medidas de similitude são referidas na Secção 2.6. Por último técnicas de medida de relevância são descritas na Secção 2.7.

### 2.1 Introdução

Sistemas de categorização de imagens e pesquisa de base de dados através de exemplos (sistemas CBIR), foram sendo desenvolvidos ao longo dos últimos anos. No artigo [4] os autores revêem mais de uma centena de artigos, sendo analisada a sua contribuição e apresentados desafios futuros para a pesquisa nesta área. Com este artigo podemos ter uma visão geral do que se está a desenvolver, tal como em [5] onde o autor apresenta os principais avanços em sistemas CBIR.

As técnicas clássicas usam a imagem como exemplo visual, da qual é criada uma assinatura que é composta estritamente por características visuais, como a cor e a textura. Contudo, existem muitas situações em que as imagens são semelhantes do ponto de vista da sua cor e da sua textura mas um observador humano reconhece-as como diferentes. Assim a simples caracterização de uma imagem através de informação primária pode não ser suficiente para avaliar a sua semelhança.

A anotação de imagens, [6], [7] e [8], é muito usada e pode ser uma alternativa aos sistemas CBIR. Este método consiste em atribuir palavras-chave às imagens, de modo que sejam descritas facilmente a nível semântico. A maior vantagem da anotação de imagem em sistemas CBIR é a de facilitar a busca por um utilizador.

## 2.2 Aplicações

As aplicações para sistemas de CBIR são inúmeras. Na pesquisa de imagens do mundo real [7], [9] e [8]. Algumas soluções são mais focadas em determinadas características, por exemplo na textura [2] e [10]. Algumas soluções procuram otimizar o desempenho, as suas aplicações que têm esse requisito, por exemplo, na pesquisa de imagem através da Internet [9]. Existem também aplicações para áreas específicas, como a área médica por exemplo [1], que auxiliam os diagnósticos.

## 2.3 Sistemas CBIR

Esta secção tem como objectivo apresentar a estrutura básica de um sistema CBIR. A organização e os blocos principais de um sistema sistema CBIR estão ilustrados na Fig. 2.1.

Nestes sistemas a medição de características é uma fase essencial. A medição de características das imagens de base de dados é efectuada *offline* (setas a azul). Assim a primeira fase na constituição de um sistema CBIR é a construção de uma base de dados contendo os vectores de características representativos de cada uma das imagens constituintes da base de dados.

A Categorização pode consistir em andares de classificação e/ou de agrupamento, [1], ou basear-se somente numa função de similitude [10], [2]. Como saída pretende-se obter um conjunto de imagens similares à imagem exemplo.

Existem muitos sistemas que incluem a possibilidade de acrescentar um módulo de Relevância. A Relevância permite o refinamento da pesquisa, através da introdução do utilizador no anel de reconhecimento, avaliando cada resultado como a relevância ou não relevância de cada proposta do bloco de categorização.

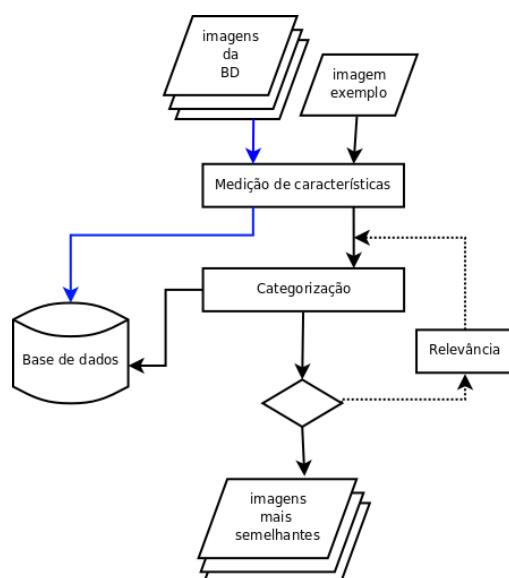


Figura 2.1: Arquitectura de sistemas CBIR.

## 2.4 Análise de imagem

A análise de imagem é uma das fases cruciais do sistema descrito acima, que fundamentalmente envolve medição de características que descrevem a imagem quantitativamente. Há assim nesta fase uma contracção de informação. Nesta secção apresentam-se algumas características frequentemente utilizadas área.

Em [2] a extracção de características, como a cor, a forma e a textura têm por base *textons*. Um *texton* é um conjunto elementar de *pixels* vizinhos. O resultado da fase de medição é uma imagem de *textons*, onde os *pixels* diferentes de zero se ajustam a pelo menos uma das cinco configurações de *textons*. Posteriormente é calculada a matriz de co-ocorrência da imagem de *textons*, que mede o número de ocorrências que um *pixel* de valor  $t$  é adjacente a um *pixel* de valor  $r$ . Cada elemento  $(t, r)$  na matriz de co-ocorrências revela quantas vezes o par  $(t, r)$  ocorre na imagem de *textons*.

Em [1] é usada a cor, textura e o contorno para descrever as características da imagem. As características da cor são extraídas convertendo a imagem para o espaço de cor de HSV (*Hue Saturation Value*), o descritor da cor é representado como um vector com nove elementos entre os quais o desvio médio, o desvio padrão e a *skewness* de cada canal da cor. As características da textura são medidas a partir da matriz de co-ocorrência, em quatro direcções ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ). Destas matrizes são medidas a energia, a probabilidade máxima, a entropia, o contraste e o momento da diferença inversa. Para extrair a característica da forma é construído um histograma da direcção do contorno. Para extrair a informação do contorno presente nas imagens recorre-se ao algoritmo Canny [11]. No final, obtêm um vector de 101 características. A dimensionalidade desse espaço é reduzida através da análise das componente principais (PCA).

Em [10] os autores extraem a informação da cor e da textura usando histogramas de saturação e *wavelets*. A cor é representada no espaço *CIELab*. Apenas a informação da saturação é usada para descrever a cor nas imagens. A distribuição da saturação é representada com histogramas de uma ou duas dimensões. Os autores assumem que a cada coordenada pode ser modelada por uma distribuição Gaussiana ou Laplaciana, tornando o algoritmo mais rápido e diminuindo a necessidade de aumentar o espaço de armazenamento, mas confinando a série de dados de modo a que cada distribuição da saturação da imagem possa ser modelada. Através da transformada discreta de *wavelet* (DWF), os autores extraem as características da textura pela decomposição da imagem em diferentes bandas de frequências. A componente de alta frequência é usada para descrever a textura. Este método de extracção de textura de informação é também usado em [9].

Em [7] os autores combinam meta-informação, como a abertura do obturador, o tempo de exposição, com características naturais (cor, forma e textura), para a anotação automática da imagem. Primeiramente a imagem é convertida do espaço de cor RGB (vermelho, verde e azul) para um valor de imagem indexado. Posteriormente é extraída a característica global usando o histograma da cor da imagem com valores indexados. A extracção de informação espacial, textura e forma, é feita usando correlogramas.

Em [9] a extracção da característica da cor é obtida convertendo a imagem do espaço de cor RGB para LUV ( $L^*$  luminância,  $u^*$  e  $v^*$  coordenadas de cromaticidade). A medida de textura é

obtida através da transformada discreta de *wavelet* (Daubechies-4). Usando o algoritmo *k-means* são calculados grupos a partir da informação da cor e textura, separadamente. O vector de características é construído pelas médias e probabilidades de cada grupo.

## 2.5 Classificação e agrupamento

Técnicas de classificação e agrupamento são muito usados em sistemas CBIR, procuram restringir o espaço de pesquisa ao conjunto de classes ou grupos de imagens. Podemos distinguir dois métodos de Classificação, a classificação supervisionada e não supervisionada.

A classificação não supervisionada é usada quando não existe alguma informação prévia sobre a classe de cada uma das amostras. São métodos que procuram encontrar o agrupamento natural no espaço de características usado, tais como a cor, textura, entre outras. Na classificação supervisionada as imagens são previamente classificadas (ou etiquetadas), de forma a treinar o classificador tendo em conta a classes existentes. Este método é usado frequentemente em algoritmos de anotação.

No artigo [1] são estudadas abordagens de classificação supervisionada não supervisionada. É utilizado *Fuzzy C-Means* para determinar os grupos das imagens. O *Fuzzy C-Means* é uma técnica de classificação não supervisionada, baseada na minimização de uma função objectivo. A classificação supervisionada é explorada usando máquinas de vectores de suporte (SVM) multi-classe, SVMs são máquinas com aprendizagem supervisionada, em que é criado o hiperplano que melhor separa as classes previamente definidas. A decisão final é dada pela combinação das probabilidades dadas pelas duas técnicas (*Fuzzy C-Means* e SVMs), resultando nas probabilidades finais de pertença a cada uma das classes, obtidas a partir de características naturais (*Fuzzy C-Means*) e de amostras previamente classificadas (SVM).

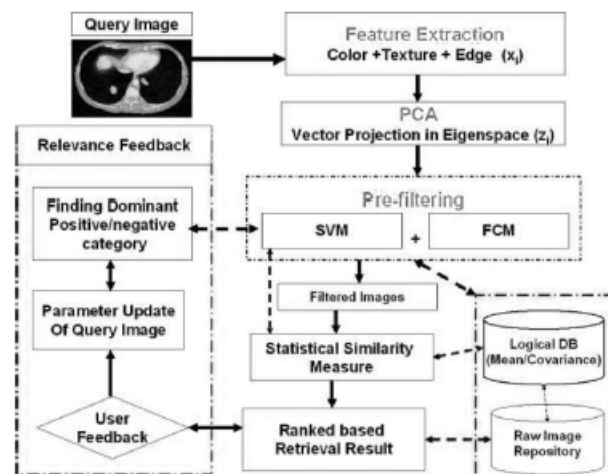


Figura 2.2: Diagrama de blocos do sistema CBIR de [1]

Na Fig. 2.2 está ilustrado o diagrama de blocos deste sistema [1]. Numa primeira fase são extraídas as características, *Feature Extraction*, e a dimensionalidade do espaço de características reduzida

através da análise de componentes principais (*PCA*). É feita uma pré-filtragem da base de dados, que consiste em obter as imagens pertencentes à classe atribuída à imagem exemplo (*Query Image*) pelos classificadores. No final é feita uma ordenação por semelhança recorrendo a uma medida de similitude. O módulo de *Relevance Feedback* permite ao utilizador refinar a pesquisa com vista a melhorar os resultados.

Em [7] os autores usam a informação existente da meta-informação da imagem (abertura, tempo de exposição...) para fazer uma abordagem baseada em restrições a fim de atribuir etiquetas às imagens. As imagens podem ser divididas em cenas, tais como a natureza e a vida selvagem, retrato, paisagem e desportos. Nos desportos, por exemplo, as imagens normalmente são tiradas com a câmara com extrema rapidez de tempo de exposição e lentes telefoto, esta informação é contida na meta-informação de uma imagem, possibilitando a extracção de alguma informação semântica, através deste o tipo das suposições. As regras são induzidas no classificador pelo algoritmo C4.5, recorrendo a um conjunto de imagens de treino previamente anotadas.

## 2.6 Medidas de distância e semelhança entre imagens

Através das medidas de similitude é determinada a probabilidade de similitude entre imagens ou a de pertencer a um grupo. A distância Euclidiana é a medida clássica, é dada pelo comprimento do segmento que junta dois pontos. Esta distância é usada em [2] e [1] como um classificador de distância mínima entre imagens. Em [1] os autores exploram igualmente a distância de Bhattacharyya e Mahalanobis comparando com os resultados dados pela distância Euclideana. A distância de Bhattacharyya mede a similaridade de duas distribuições discretas da probabilidade. Esta é utilizada quando as características da imagem são representadas como distribuições discretas da probabilidade [1], [10].

Outro tipo de medições de similitude é descrita em [12]. Os autores apresentam um novo método de medição de similitude. É baseado em ontologias definidas pelos utilizadores, e tem capacidade para lidar com dados organizados numa estrutura em árvore. A medida de similitude é determinada pela distância entre a correspondência dos nós da estrutura de árvore. Em [9] é apresentado um novo método de medidas, baseado na mistura da aproximação do modelo chamada *hypothetical local mapping* (HLM).

## 2.7 Medida de relevância

Os sistemas de medida de relevância, sistemas RF (do inglês *Relevance Feedback*), acrescentam interactividade ao sistema CBIR, indicando a relevância ou não relevância dos resultados obtidos, permitindo que a pesquisa seja refinada. Pretende-se nesta Secção indicar alguns métodos utilizados para a medição de Relevância.

Em sistemas CBIR uma imagem é representada por um vector de características, num espaço multidimensional. Um ponto que nesse espaço representa a imagem exemplo, é o ponto de consulta. As técnicas de deslocamento do ponto de consulta e de ajuste de pesos das características são as mais populares. Os artigos [13] e [14] fazem uma análise pormenorizada nesta área.

Em [15] os autores propõem uma técnica RF baseada em pontuação. A pontuação de relevância, é calculada através das distâncias da imagem não-relevante e da imagem relevante mais próximas.

No artigo [16] os autores apresentam uma série de técnicas RF, tais como: o deslocamento do ponto de consulta; transformações das funções de distância; selecção de função de distância e atribuição de pesos às características. É considerada também uma abordagem multi-ponto, isto é, vários pontos de consulta em simultâneo.

Em [1] os autores usam uma técnica de deslocamento do ponto de consulta, através do cálculo do vector médio entre o ponto de consulta actual e as imagens classificadas como relevantes pelo utilizador.

Em [17] os autores propõem várias técnicas optimizadas para grandes bases de dados. As técnicas são suportadas por uma estrutura de índice que permite uma grande eficiência em termos computacionais. Os autores também consideram a possibilidade de existir mais que um ponto de consulta.

Um dos problemas típicos dos sistemas RF é considerar que o número de resultados relevantes é igual aos resultados não-relevantes. Na maioria dos casos isso não acontece, já que o número de resultados não-relevantes é francamente superior aos relevantes. Em [18] o autor parte do princípio que os conjuntos não são equilibrados. Usa então *biased minimax probability machine* (BMPM) para fazer a classificação do que é ou não relevante.

Geralmente a análise de relevância é feita somente sobre os resultados positivos (ou relevantes). Em [19], o autor considera que os resultados positivos são agrupados num conjunto, enquanto que os resultados negativos (não-relevantes) estão separados em conjuntos diferentes. Os resultados negativos são agrupados em vários grupos. É construída uma *marginal convex machine* para cada grupo de negativos e também para o grupo dos resultados positivos, dando origem a um conjunto de classificadores. Estes classificadores vão ser incorporados numa *biased marginal convex machine* para construir o classificador do sistema RF.

## 2.8 Sumário

As medidas cruciais num sistema CBIR são as medições de características das imagens e de similitude entre imagens. As medições de características de imagens permitem a construção de vectores representativos das imagens com menor dimensão que a própria imagem. As medidas de similitude permitem a comparação entre imagens em termos de semelhança.

A classificação é um bloco que pode ser integrado a fim de reduzir o espaço de pesquisa, constituído apenas por imagens pertencentes à mesma classe que a imagem exemplo.

A relevância é uma medida que permite ao utilizador interagir com o sistema, refinando a pesquisa de acordo com os seus objectivos.



## Capítulo 3

# Medição de características

Neste capítulo são apresentadas as características utilizadas para representar cada uma das imagens da base de dados. Na secção 3.1 é feita uma breve introdução. Na secção 3.2 é explicado como é retirada informação da textura das imagens e na Secção 3.4 descrevem-se os métodos de caracterização de imagem a partir dos seus contornos.

### 3.1 Introdução

Uma imagem é representada por  $N \times M$  ( $N$  é número de linhas e  $M$  o número de colunas) *pixels* em que a cada *pixel* está associada uma cor. As características de uma imagem correspondem a medições retiradas da própria imagem que a permite descrever num espaço de menor dimensão.

Na medição de características adoptou-se uma abordagem multi-escala, ou seja, da imagem original  $J(x,y)$  é extraído um conjunto de  $n$  imagens, obtidas através da filtragem com  $n$  filtros Gaussianos de diferentes valores de  $\sigma$ . A imagem filtrada  $I(x,y)$  é obtida através da convolução (\*) da imagem de entrada  $J(x,y)$  e o filtro gaussiano  $G_\sigma(x,y)$

$$I(x,y) = G_\sigma(x,y) * J(x,y) \text{ onde } G_\sigma(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2+y^2}{2\sigma^2}\right] \quad (3.1)$$

É adoptada uma solução multi-escala porque permite retirar informação da imagem com diferentes níveis de pormenor.

Os valores de  $\sigma$  escolhidos são [0.8, 1.6, 3.2, 6.4]. A escolha partiu de 0.8, no qual a filtragem na imagem era suave e não se perdia grandes detalhes. Cada escala seria separada por uma oitava terminando em 6.4, escala na qual ainda estava presente informação da imagem após a filtragem.

A seguir são apresentadas as características usadas nesta dissertação para caracterizar as imagens  $I(x,y)$ .

## 3.2 Textura

A textura de uma imagem contém informação sobre a distribuição espacial dos *pixels*. Nesta Secção pretende-se apresentar o método utilizado para a medição da textura nas imagens.

A medição de textura é realizada recorrendo a *textons* e à matriz de co-ocorrências.

O termo *texton* já existe há muitos anos e é um conceito útil em análise de texturas, podendo ser caracterizado por uma estrutura de *pixels*, que ocorre dentro de uma vizinhança. Nesta dissertação foram utilizados *textons* de dimensões  $2 \times 2$ , com *textons* dessa dimensão é possível construir cinco *textons* diferentes (no mínimo com três *pixels*) Fig. 3.3 [2].



Figura 3.1: Exemplo de uma imagem médica da base de dados.

É calculado o gradiente da imagem  $I(x, y)$ . A função gradiente  $G(x, y, \theta)$  [20] é determinada em função de  $g_{xx}$ ,  $g_{yy}$  e  $g_{xy}$ , na direcção  $\theta$  através da expressão

$$G(x, y, \theta) = (0.5 \times [(g_{xx} + g_{yy}) + (g_{xx} - g_{yy}) \cos(2\theta) + 2g_{xy} \sin(2\theta)])^{\frac{1}{2}}. \quad (3.2)$$

Para uma imagem a cores  $g_{xx}$ ,  $g_{yy}$  e  $g_{xy}$  são determinadas em função das três componentes monocromáticas:  $I_R$  para a componente vermelha;  $I_G$  para a componente verde;  $I_B$  para a componente azul, como mostra a Eq. (3.3).

$$\begin{aligned} g_{xx} &= \left| \frac{\partial I_R}{\partial x} \right|^2 + \left| \frac{\partial I_G}{\partial x} \right|^2 + \left| \frac{\partial I_B}{\partial x} \right|^2 \\ g_{yy} &= \left| \frac{\partial I_R}{\partial y} \right|^2 + \left| \frac{\partial I_G}{\partial y} \right|^2 + \left| \frac{\partial I_B}{\partial y} \right|^2 \\ g_{xy} &= \frac{\partial I_R}{\partial x} \frac{\partial I_R}{\partial y} + \frac{\partial I_G}{\partial x} \frac{\partial I_G}{\partial y} + \frac{\partial I_B}{\partial x} \frac{\partial I_B}{\partial y} \end{aligned} \quad (3.3)$$

Para uma imagem monocromática as correspondentes expressões são

$$\begin{aligned} g_{xx} &= \left| \frac{\partial I}{\partial x} \right|^2 \\ g_{yy} &= \left| \frac{\partial I}{\partial y} \right|^2 \\ g_{xy} &= \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \end{aligned} \quad (3.4)$$

O gradiente é calculado em função de uma direcção  $\theta$ , na direcção onde a variação do gradiente é máxima,  $\theta_+$ .

Pretende-se assim encontrar o valor de  $\theta$  que maximize a equação

$$F(\theta) = g_{xx} \cos^2(\theta) + 2g_{xy} \cos(\theta) \sin(\theta) + g_{yy} \sin^2(\theta). \quad (3.5)$$

A solução é dada por

$$\theta_+ = \frac{1}{2} \arctan\left(\frac{2g_{xy}}{g_{xx} - g_{yy}}\right) + k\pi. \quad (3.6)$$

A direcção na qual o gradiente tem variação mínima  $\theta_-$ , é ortogonal a  $\theta_+$ , sendo dada por

$$\theta_- = \theta_+ \pm \frac{\pi}{2}. \quad (3.7)$$

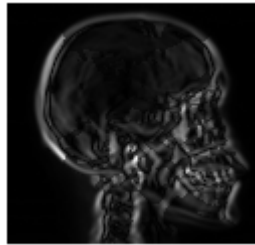


Figura 3.2: Imagem de gradiente  $G_{max}(x,y)$  obtida a partir da imagem da Fig. 3.1

Apartir das duas imagens de gradiente  $G_{max}(x,y)$  e  $G_{min}(x,y)$ , que correspondem ao máximo e mínimo do gradiente respectivamente, são calculadas duas imagem *texton*.

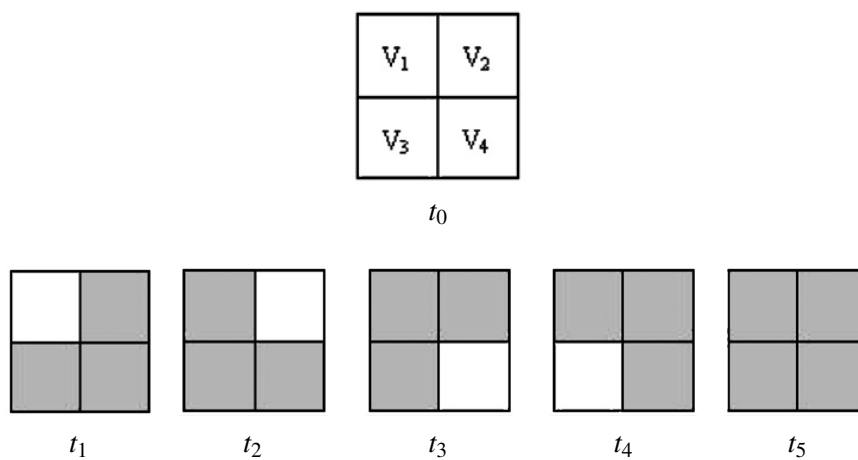


Figura 3.3:  $t_0$  Estrutura de *texton*.  $t_{1-5}$  *Textons* utilizados

Como exemplo, seja  $I(x,y)$  uma imagem de dimensão  $6 \times 6$  dada por

$$I(x,y) = \begin{bmatrix} 1 & 1 & 2 & 2 & 3 & 2 \\ 2 & 1 & 2 & 2 & 1 & 1 \\ 3 & 3 & 0 & 1 & 2 & 2 \\ 3 & 3 & 3 & 1 & 2 & 1 \\ 2 & 2 & 2 & 1 & 3 & 3 \\ 3 & 3 & 3 & 2 & 1 & 1 \end{bmatrix} \quad (3.8)$$

Para a construção da imagem de *textons* as imagens de gradiente são percorridas com cada um dos *textons* da Fig. 3.3. Um *texton* é detectado na imagem sempre que é encontrada uma configuração do *texton* na imagem, ou seja, quando os valores dos *pixels* forem iguais nas posições assinaladas pelos *textons* respectivos (3.3). Esses valores são copiados para imagem de *textons*, caso contrário são nulos. Obtêm-se assim cinco imagens de *textons*,  $T_{1-5}(x,y)$ . As imagens de *textons* de (3.8) são representadas em (3.9)

$$\begin{aligned} T_1(x,y) &= \begin{bmatrix} 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & T_2(x,y) &= \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ T_3(x,y) &= \begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 2 & 2 \\ 3 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & T_4(x,y) &= \begin{bmatrix} 1 & 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ T_5(x,y) &= \begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \quad (3.9)$$

Posteriormente estas imagens são combinadas, resultando numa imagem de *textons* global  $T(x,y)$ , contendo todos os *textons* detectados na imagem. A combinação é feita da seguinte forma: os *pixels* que não são nulos são colocados na imagem *texton* total. Neste exemplo a imagem *texton*

total é

$$T(x,y) = \begin{bmatrix} 1 & 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 & 2 & 2 \\ 3 & 3 & 3 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.10)$$



Figura 3.4: Imagem *texton* de  $G_{max}(x,y)$ , Fig. 3.2

Tabela 3.1: Tabela de *offsets* em função do ângulo  $\gamma$  e a distância  $d$

Ângulos ( $\gamma$ )	<i>Offset</i>
0°	[0 d]
45°	[-d d]
90°	[-d 0]
135°	[-d -d]

A partir desta imagem de *textons*,  $T(x,y)$ , calcula-se a matriz de co-ocorrência  $C_\gamma(i,j,d)$ . A matriz é obtida calculando o número de ocorrências que um *pixel* de valor  $i$  é vizinho de um *pixel* de valor  $j$ , na direcção  $\gamma$  e à distância  $d$ . Cada elemento  $(i,j)$  na matriz de co-ocorrências revela quantas vezes o par  $(i,j)$  ocorre na imagem de *texton*. Na Tabela 3.1 estão representados os vários *offsets* que se podem utilizar para o cálculo desta matriz.

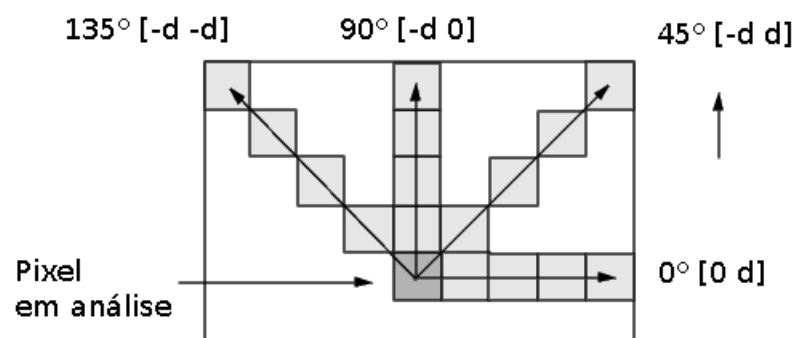


Figura 3.5: Imagem ilustrativa dos *offsets* utilizados para calcular a matriz de co-ocorrências [2].

No exemplo dado a matriz de co-ocorrência de  $T(x,y)$  na direcção  $\gamma = 0^\circ$  e a uma distância  $d = 1$  é dada por

$$C_{0^\circ}(i,j,1) = \begin{bmatrix} O(0,0) & O(1,0) & O(2,0) & O(3,0) \\ O(0,1) & O(1,1) & O(2,1) & O(3,1) \\ O(0,2) & O(1,2) & O(2,2) & O(3,2) \\ O(0,3) & O(1,3) & O(2,3) & O(3,3) \end{bmatrix} = \begin{bmatrix} 13 & 0 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 \\ 1 & 0 & 1 & 3 \end{bmatrix} \quad (3.11)$$

Obtida a matriz de co-ocorrência, são medidas algumas características. Tais como energia (3.12), contraste (3.13), entropia (3.14) e homogeneidade (3.15) [2].

$$En = \sum_x \sum_y C_\gamma^2(x,y,d) \quad (3.12)$$

$$Ct = \sum_{x-y=0} (x-y)^2 \times C_\gamma(x,y,d) \quad (3.13)$$

$$Et = - \sum_x \sum_y C_\gamma(x,y,d) \log(C_\gamma(x,y,d)) \quad (3.14)$$

$$Hm = \sum_x \sum_y \frac{C_\gamma(x,y,d)}{1 + (x-y)^2} \quad (3.15)$$

### 3.3 Cor

A cor é uma característica essencial na representação da imagem. Nesta Secção é apresentado o método de medição de características da cor.

A medição de cor é feita de forma semelhante à extracção da textura. Não se lida com imagens do gradiente da imagem, como na medição da textura, mas com uma imagem monocromática de 256 cores.

Para imagens policromáticas representadas em RGB é feita uma quantização [2], a fim de obter uma imagem de 256 níveis de cinzento. Sejam  $I_R(x,y)$ ,  $I_G(x,y)$  e  $I_B(x,y)$  as três componentes da imagem RGB. A imagem quantizada  $I_Q(x,y)$  é dada pela seguinte expressão

$$I_Q(x,y) = 32 \times I'_R(x,y) + 4 \times I'_G(x,y) + I'_B(x,y) \quad (3.16)$$

onde

$$\begin{cases} I'_R(x,y) = 0, & 0 \leq I_R(x,y) \leq 32 \\ I'_R(x,y) = i, & 32 \times i + 1 \leq I_R(x,y) \leq 32 \times (i + 1) \\ i \in [1, 2, \dots, 7] \end{cases} \quad (3.17)$$

$$\begin{cases} I'_G(x,y) = 0, & 0 \leq I_G(x,y) \leq 32 \\ I'_G(x,y) = i, & 32 \times i + 1 \leq I_G(x,y) \leq 32 \times (i + 1) \\ i \in [1, 2, \dots, 7] \end{cases} \quad (3.18)$$

$$\begin{cases} I'_B(x,y) = 0, & 0 \leq I_B(x,y) \leq 64 \\ I'_B(x,y) = i, & 64 \times i + 1 \leq I_B(x,y) \leq 64 \times (i + 1) \\ i = 1, 2, 3 \end{cases} \quad (3.19)$$

Para imagens monocromáticas, que envolvem só uma componente em nível de cinzento, este passo de quantização não é realizado, pois a imagem  $I(x,y)$  já por si tem de 256 níveis, assim  $I_Q(x,y) = I(x,y)$ .

Após obter-se a imagem  $I_Q$ , é construída a imagem de *textons*, tal como foi descrito na Secção 3.2. A partir da imagem de *textons* é calculada a matriz de co-ocorrências e extraídas medições de energia (3.12), contraste (3.13), entropia (3.14) e homogeneidade (3.15), descritos na Secção 3.2.

### 3.4 Contorno

Para obter informação da forma recorre-se aos contornos da imagem. Nesta Secção é descrito o método de medição das características a partir dos contornos.

Para este efeito, inicia-se pelo cálculo do gradiente da imagem  $I(x,y)$ , calculando-se o módulo e fase pelas expressões

$$|I(x,y)| = \sqrt{\frac{\partial I(x,y)}{\partial x}^2 + \frac{\partial I(x,y)}{\partial y}^2} \quad (3.20)$$

$$\angle I(x,y) = \arctan \frac{\frac{\partial I(x,y)}{\partial y}}{\frac{\partial I(x,y)}{\partial x}} \quad (3.21)$$

As imagens são depois multiplicadas pelo contorno da imagem  $I(x,y)$ , dada pelo algoritmo de Canny [11]. Pretendendo-se assim guardar informação do valor do módulo e fase do gradiente somente no contorno da imagem. Com a imagem resultante é calculado o histograma do módulo e fase.

Depois de obter os histogramas, com  $N_{Bins}$  intervalos, são efectuadas algumas medições a fim de representar os contornos. O vector das medições é constituído pela média (3.22), variância (3.23), *skewness* (3.24) e *kurtosis* (3.25) em que  $x_i$  representa os valores do gradiente e  $f_i$  o número de ocorrências de  $x_i$ ,

$$m = \frac{1}{N_{Bins}} \sum_{i=0}^{N_{Bins}-1} f_i x_i \quad (3.22)$$

$$v = \frac{1}{N_{Bins} - 1} \sum_{i=0}^{N_{Bins}-1} (f_i(x_i) - m)^2 \quad (3.23)$$

$$sk = \frac{\sum_{i=0}^{N_{Bins}-1} (x_i - m)^3}{(N_{Bins} - 1)v^3} \quad (3.24)$$

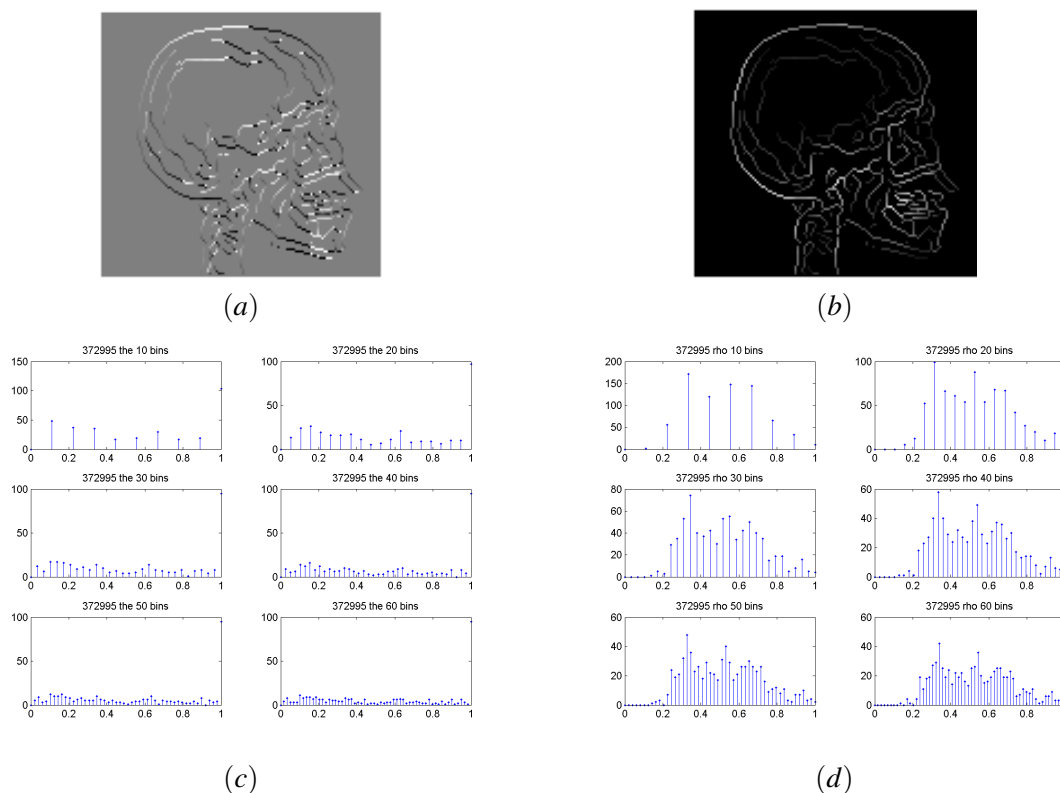


Figura 3.6: (a) Módulo do gradiente no contorno da imagem Fig.3.1. (b) Fase do gradiente no contorno da imagem Fig.3.1. (c)(d) Histogramas de módulo e fase das imagens (a) e (b), respectivamente, com diferentes intervalos.

A *skewness* mede a assimetria do histograma. Um histograma tem uma simetria perfeita se a *skewness* é nula, caso contrário se for positiva a distribuição concentra-se em valores superiores a  $\bar{x}$  caso contrário concentra-se em valores inferiores.

$$ku = \frac{\sum_{i=0}^{N_{Bins}-1} (x_i - m)^4}{(N_{Bins} - 1)v^4} - 3 \quad (3.25)$$

A *kurtosis* mede o grau de achatamento de uma distribuição. Se a *kurtosis* é nula então a distribuição tem o mesmo achatamento que uma distribuição normal, se a *kurtosis* é positiva então a distribuição é mais alta e afunilada que uma distribuição normal, caso contrário a distribuição é mais plana.

São ainda extraídos os 75% e 90% percentis. Percentil é o valor de  $x_k$  que corresponde à frequência acumulativa de  $\frac{N_{Bins}k}{100}$ , ou seja o p-ésimo percentil tem no mínimo p% dos valores de  $f_i$ , inferiores a  $f_k$  e no mínimo  $(100 - p)\%$  dos valores de  $f_i$  superiores a  $f_k$ .

Na figura 3.6 está ilustrado as imagens dos valores do módulo e fase no contorno da imagem e respectivos histogramas.

## 3.5 Sumário

Em suma são retiradas medidas de textura, cor e contorno.

As medidas de textura e cor são obtidas construindo uma imagem de *texton*, através de um conjunto pré-definido de *textons*. Da imagem de *texton* é calculada a matriz de co-ocorrências dos *pixels*. Uma vez calculada a matriz de co-ocorrências são medidas as características: energia, contraste, entropia e homogeneidade.

A medida de contorno, é calculada pela medição de características dos histogramas dos valores de módulo e fase do gradiente da imagem ao longo do contorno. As características são a média, a variância, a *skewness*, a *kurtosis*, e os percentis 75% e 90%.



## Capítulo 4

# Classificação e Redução de Dimensionalidade

Neste capítulo são descritos e exemplificados os classificadores utilizados para construir a estrutura de votação. São também explicados dois métodos de redução de características de forma a diminuir a dimensionalidade espaço de características.

### 4.1 Classificadores Bayesianos

Na Classificação bayesiana [21] [22] intervém a probabilidade de um objecto  $\mathbf{x}$  pertencer à classe  $c_i$ ,  $P(c_i|\mathbf{x})$  dada por

$$P(c_i|\mathbf{x}) = \frac{p(\mathbf{x}|c_i)P(c_i)}{P(\mathbf{x})} \quad (4.1)$$

Para um problema para  $K$  classes

$$P(\mathbf{x}) = \sum_{i=1}^K p(\mathbf{x}|c_i)P(c_i) \quad (4.2)$$

Onde  $P(c_i)$  é a probabilidade *a priori* de um objecto pertencer à classe  $c_i$  e representa o conhecimento prévio sobre a distribuição das classes antes de medir as características do objecto a classificar,  $p(\mathbf{x}|c_i)$  é a probabilidade de uma amostra  $\mathbf{x}$  pertencer à classe  $c_i$ .

Para um problema de duas classes,  $c_1$  e  $c_2$ , a regra de decisão de Bayes é

$$\text{Decidir por } c_1 \text{ se } P(c_1|\mathbf{x}) > P(c_2|\mathbf{x}) \text{ caso contrário decidir por } c_2 \quad (4.3)$$

Assim a probabilidade de erro na classificação do objecto  $\mathbf{x}$ , ou seja, a probabilidade de ser classificado na classe errada é

$$P(\text{erro}|\mathbf{x}) = \min[P(c_1|\mathbf{x}), P(c_2|\mathbf{x})]. \quad (4.4)$$

A probabilidade de erro de um objecto  $\mathbf{v}$  da classe  $c_1$  é igual à probabilidade de  $\mathbf{v}$  ser classificado como sendo da classe  $c_2$ ,  $P(c_2|\mathbf{v})$ .

Este tipo de classificação é supervisionado, pelo que é necessário haver uma fase de treino a fim de construir as funções discriminantes  $h_i(\mathbf{x})$ , que são dependentes das distribuições por classe do objecto  $\mathbf{x}$ .

Um classificador é representado pelas suas  $K$  funções discriminantes, obtidas na fase de treino. Dado um vector de características  $\mathbf{x}$ , o classificador classifica  $\mathbf{x}$  como sendo da classe  $c_i$  se

$$h_i(\mathbf{x}) > h_j(\mathbf{x}) \text{ para todos os } j \neq i \quad (4.5)$$

Pela equação (4.3) pode-se verificar que as equações são semelhantes, então é possível definir a função discriminante da seguinte forma

$$h_i(\mathbf{x}) = P(c_i|\mathbf{x}) = \frac{p(\mathbf{x}|c_i)P(c_i)}{P(\mathbf{x})} \quad (4.6)$$

Simplificando a equação, aplicando o logaritmo à função  $h_i(\mathbf{x})$ , obtém-se

$$g_i(\mathbf{x}) = \ln(h_i(\mathbf{x})) = \ln(p(\mathbf{x}|c_i)) + \ln(P(c_i)) - \ln(P(\mathbf{x})) \quad (4.7)$$

Se a probabilidade  $p(\mathbf{x}|c_i)$  tiver uma distribuição normal multivariável ( $N(\mu_i, \Sigma_i)$ ), isto é

$$p(\mathbf{x}|c_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} (\mathbf{x} - \mu_i)^t \Sigma_i^{-1} (\mathbf{x} - \mu_i) \right] \quad (4.8)$$

onde  $\mathbf{x}$  é um vector de comprimento  $d$ ,  $\mu_i$  é vector da média das amostras da classe  $c_i$ ,  $\Sigma_i$  é a matriz das covariâncias de dimensão  $d \times d$  da classe  $c_i$ . Substituindo na equação 4.7 obtém-se as funções discriminantes pela seguinte equação

$$g_i(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \mu_i)^t \Sigma_i^{-1} (\mathbf{x} - \mu_i) - \frac{d}{2} \ln(2\pi) - \frac{1}{2} \ln(|\Sigma_i|) + \ln(P(c_i)) - \ln(P(\mathbf{x})) \quad (4.9)$$

que correspondem às fronteiras de decisão, de separação das diferentes classes. A equação pode ser simplificada retirando os termos independentes das classes, ou seja os termos independentes de  $i$ , obtendo-se

$$g_i(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \mu_i)^t \Sigma_i^{-1} (\mathbf{x} - \mu_i) - \frac{1}{2} \ln(|\Sigma_i|) + \ln(P(c_i)) \quad (4.10)$$

#### 4.1.1 Classificador Linear

O classificador linear assume que as  $K$  classes têm covariâncias iguais ( $\Sigma_i = \Sigma$ ) e médias diferentes. Como  $\Sigma$  não depende das classes, simplificamos a equação (4.10) retirando o termo

$-\frac{1}{2}\ln(|\Sigma|)$  ficando da seguinte forma

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^t \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \ln(P(c_i)) \quad (4.11)$$

se expandirmos o termo quadrático,  $(\mathbf{x} - \boldsymbol{\mu}_i)^t \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)$  obtém-se

$$g_i(\mathbf{x}) = \mathbf{w}_i^t \mathbf{x} + c_{i0} \quad (4.12)$$

onde

$$\mathbf{w}_i = \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i \quad (4.13)$$

e

$$c_{i0} = -\frac{1}{2} \boldsymbol{\mu}_i^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i + \ln(P(c_i)) \quad (4.14)$$

pois o elemento quadrático, após a expansão  $(\mathbf{x}^t \boldsymbol{\Sigma} \mathbf{x})$ , não depende de  $i$ , daí ser possível retirar da equação.

Como as funções discriminantes dadas pela equação (4.12) são lineares as fronteiras de decisão, neste caso em duas dimensões são rectas. Na Fig 4.1 dá-se um exemplo para um problema com duas classes e com duas dimensões.

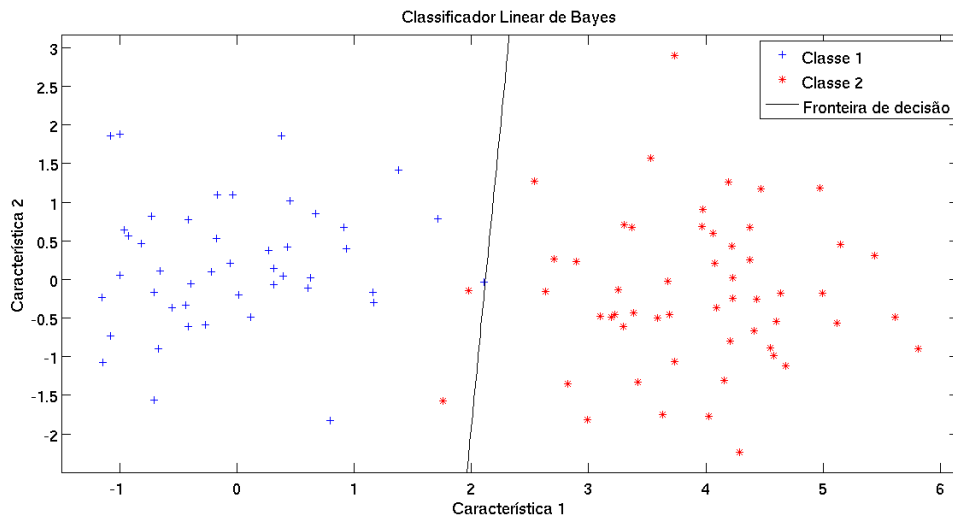


Figura 4.1: Exemplo da classificação do LDC para duas classes.

As distribuições da Fig. 4.1 têm uma distribuição gaussiana com duas dimensões, médias de  $(0,0)$  e  $(4,0)$  para a classe 1 e classe 2 respectivamente e a matriz de covariâncias é a matriz identidade. Estas distribuições são utilizadas para ilustrar os classificadores referidos neste capítulo.

### 4.1.2 Classificador Quadrático

As funções discriminantes são dadas pela equação (4.10). Para evidenciar o termo quadrático, esta equação pode ser reescrita na forma

$$g_i(\mathbf{x}) = \mathbf{x}^t W_i \mathbf{x} + w_i^t \mathbf{x} + c_{i0} \quad (4.15)$$

onde

$$W_i = -\frac{1}{2} \Sigma_i^{-1} \quad (4.16)$$

$$w_i = \Sigma_i^{-1} \mu_i \quad (4.17)$$

e

$$c_{i0} = -\frac{1}{2} \mu_i^t \Sigma_i^{-1} \mu_i - \frac{1}{2} \ln(|\Sigma_i|) + \ln(P(c_i)) \quad (4.18)$$

Obtém-se assim funções de ordem quadrática, então para um problema a duas dimensões as fronteiras de decisão vão ser compostas por curvas. Na Fig. 4.2 está ilustrado este classificador, para duas classes e duas dimensões.

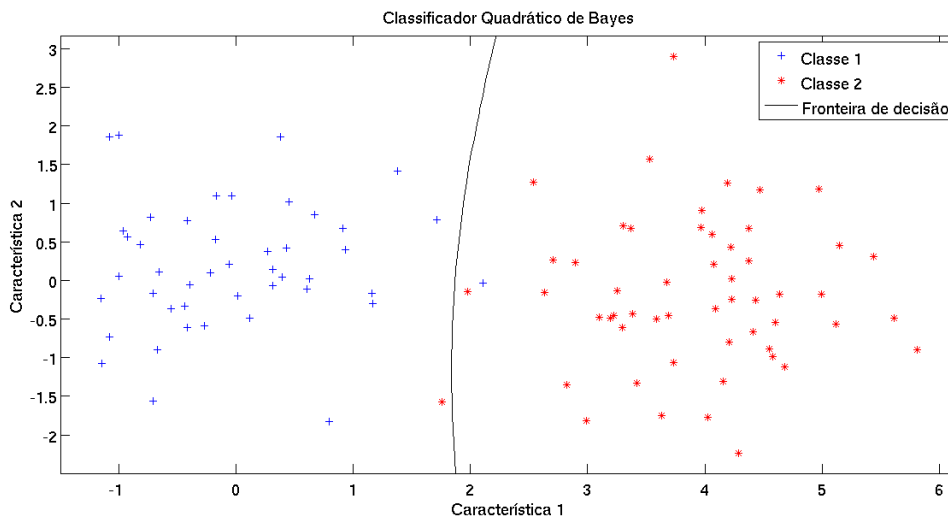


Figura 4.2: Exemplo da classificação do QDC para duas classes.

## 4.2 Classificador K vizinhos mais próximos (KNN)

O classificador KNN é um classificador baseado na vizinhança das amostras. A classificação de uma amostra  $\mathbf{x}$  é realizada analisando as classes dos  $k$  vizinhos mais próximos, sendo atribuída a classe mais comum da sua vizinhança. O classificador necessita de amostras previamente classificadas que vão constituir a vizinhança das amostras a classificar, ou seja é uma classificação supervisionada.

A classificação KNN consiste em determinar a probabilidade *a posteriori*  $P(c_i|\mathbf{x})$ , a partir de um conjunto de  $n$  amostras previamente classificadas, usando as amostras para estimar as densidades envolvidas. Supondo que é colocada uma célula de volume  $V$  em torno de  $\mathbf{x}$  abrangendo também  $k$  amostras,  $k_i$  das quais pertencentes à classe  $c_i$ . A estimativa para a probabilidade conjunta  $p(\mathbf{x}, c_i)$  é

$$p_n(\mathbf{x}, c_i) = \frac{k_i/n}{V} \quad (4.19)$$

então  $P(c_i|x)$  é dado por

$$P_n(c_i|\mathbf{x}) = \frac{p_n(\mathbf{x}, c_i)}{\sum_{j=1}^c p_n(\mathbf{x}, c_j)} = \frac{k_i}{k} \quad (4.20)$$

Isto é, a estimativa da probabilidade *a posteriori*  $P(c_i|x)$  é o quociente entre as amostras da classe  $c_i$  e o número total de amostras dentro da célula.

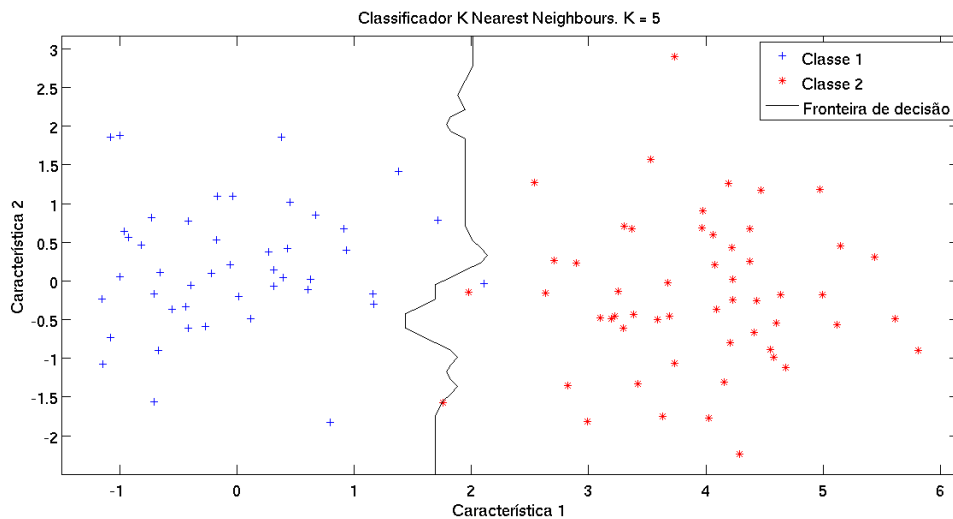


Figura 4.3: Exemplo da classificação do KNN para duas classes e  $K = 5$

Este classificador é um método não paramétrico, ou seja, não existe um conhecimento *a priori* da distribuição de classes ou de fronteiras de decisão. A fase de treino consiste somente em armazenar o conjunto de amostras de treino e as suas etiquetas, toda a computação necessária para este algoritmo é executada na classificação.

Na Fig. 4.3 está ilustrada a classificação para um problema de duas classes com duas dimensões.

### 4.3 Classificador baseado em Máquinas de Vectors de Suporte (SVM)

Máquinas de vectores de suporte são máquinas de aprendizagem supervisionada usadas para classificação. Na fase de treino o classificador tenta construir uma fronteira de separação, no espaço de características, que maximize a separabilidade entre classes.

Considerando um conjunto  $S$  de  $n$  amostras, cada uma caracterizada pelo vector de características  $\mathbf{x}$  e classificada com uma etiqueta  $e$ ,  $S = (\mathbf{x}_1, e_1), \dots, (\mathbf{x}_n, e_n)$ .

Supondo que as classes são linearmente separáveis, a função que separa as classes é dada por

$$f(\mathbf{x}) = \mathbf{w}^t \mathbf{x} + b, \quad (4.21)$$

$\mathbf{w}^t$  é a transposta de  $\mathbf{w}$ . As classes são separáveis pelo hiperplano definido por  $f(\mathbf{x}) = 0$ .

O objectivo do classificador SVM é a de maximizar o valor da margem geométrica, ou seja, a distância do hiperplano  $f(\mathbf{x}) = 0$  às amostras mais próximas, do hiperplano.

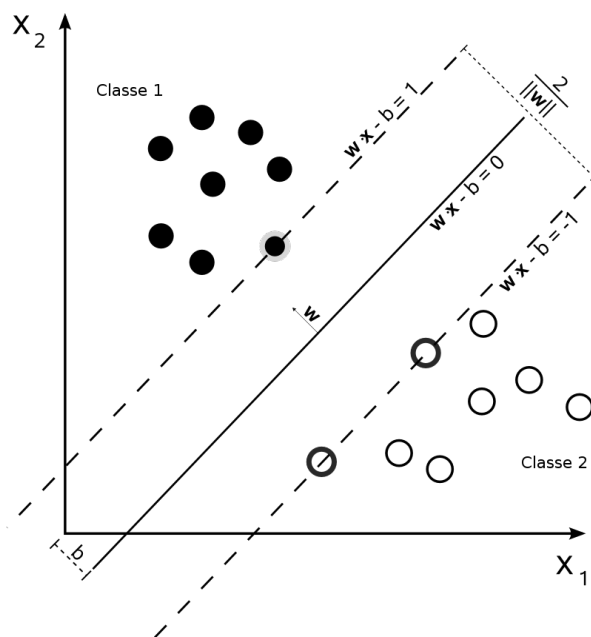


Figura 4.4: Hiperplano de separação para um conjunto de treino de duas dimensões

Na Fig. 4.4 está ilustrado o hiperplano e as margens geométricas, na classificação de um conjunto de treino com um espaço de características de duas dimensões. Os vectores que estão sobre as linhas a tracejado (margens) são chamados vectores de suporte.

Considerando que  $\mathbf{w}^t \mathbf{x} + 1 = 1$  e  $\mathbf{w}^t \mathbf{x} + b = -1$  para os vectores de suporte da classe  $c_1$  ( $\mathbf{v}_{s1}$ ) e da classe  $c_2$  ( $\mathbf{v}_{s2}$ ), a margem geométrica é dada por

$$\zeta = \frac{1}{2} \frac{1}{\|\mathbf{w}\|^2} [(\mathbf{w}^t \mathbf{v}_{s1} + b) - (\mathbf{w}^t \mathbf{v}_{s2} + b)] = \frac{1}{\|\mathbf{w}\|^2} \quad (4.22)$$

onde  $\|\mathbf{w}\|^2 = \mathbf{w}^t \mathbf{w}$ .

Para maximizar a margem  $\zeta$  podemos minimizar  $\|\mathbf{w}\|^2$ . Traduz-se no problema de optimização de encontrar o valor de  $T$ , tal que

$$T = \operatorname{argmin}(\|\mathbf{w}\|^2) \quad (4.23)$$

sujeito à seguinte restrição

$$\begin{aligned} e_i(\mathbf{w}^t \mathbf{x}_i + b) &\geq 1 \\ i &= 1, 2, \dots, n \end{aligned} \quad (4.24)$$

Para determinar a solução do problema acima referido, podemos converter-lo no seguinte problema de optimização quadrático [23]

$$\max(w(\alpha)) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j (x_i^t x_j) \quad (4.25)$$

sujeito à seguinte restrição

$$\begin{aligned} \sum_{i=1}^l \alpha_i y_i &= 0 \\ \alpha_i &\geq 0, 1, \dots, l \end{aligned} \quad (4.26)$$

A solução deste problema de optimização dá origem à seguinte regra de decisão

$$\text{sign}[f(\mathbf{x})] = \text{sign} \left[ \sum_{i=1}^n \alpha_i y_i (\mathbf{x}_i^t \mathbf{x}) + b \right] \quad (4.27)$$

A classificação é decidida pelo sinal da função  $f$ . Por exemplo, na Fig. 4.4 se numa amostra  $u$  o valor de  $f(u)$  é negativo então  $u$  pertence à classe 2 caso contrário pertence à classe 1.

O objectivo do treino de um classificador de máquinas de vectores de suporte, é a determinação do hiperplano tendo em consideração o conjunto de treino previamente classificado. A Fig. 4.5 ilustra este classificador num problema de classificação de duas classes.

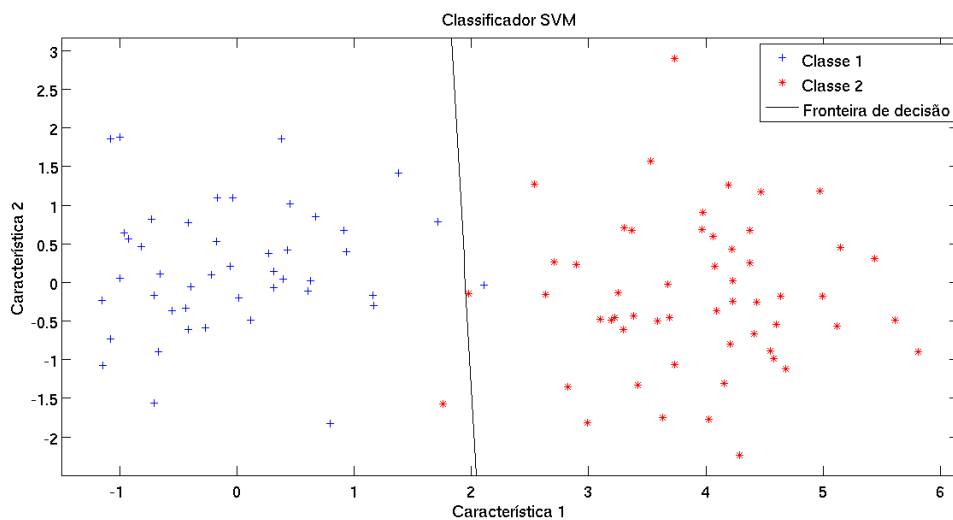


Figura 4.5: Exemplo da classificação do SVM para duas classes.

## 4.4 Combinação de classificadores por votação

A combinação de classificadores permite que não se entregue a classificação a um só classificador mas sim pela combinação de vários classificadores, aumentando assim a precisão e eficiência do sistema. Os classificadores complementam-se obtendo melhores resultados em vez de se tratar apenas de um classificador. [24].

O diagrama da Fig. 4.6 ilustra a forma de como a combinação de classificadores é realizada. Cada um dos classificadores recebe como parâmetro de entrada um vector de características  $x$ , representativo da imagem a classificar. Nesta fase, cada classificador atribui uma classe ao vector  $x$  e, mediante o resultado de uma votação, é atribuída a classe com maior número de votos.

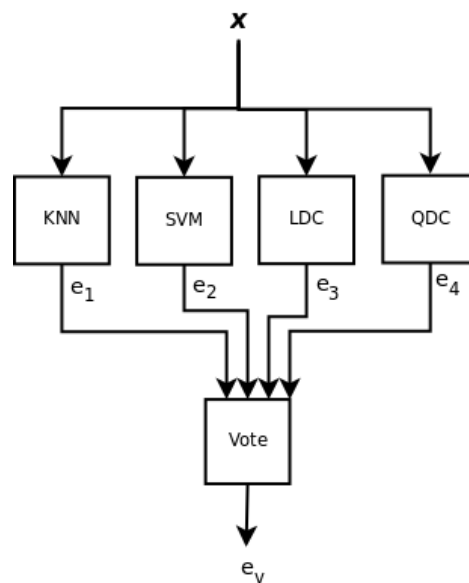


Figura 4.6: Diagrama de combinação de classificadores.

## 4.5 Redução de Dimensionalidade

Nesta Secção serão apresentados apenas os dois métodos de redução de características, a análise de componentes principais (PCA) e análise discriminante de Fischer. As explicações não serão muito detalhadas apenas focando o essencial, podendo-se consultar o livro [22] para uma descrição mais pormenorizada.

### 4.5.1 Análise de Componentes Principais

Na análise de componentes principais é criada uma transformação linear a aplicar no espaço de características, afim de reduzir o espaço de características, obedecendo ao critério de erro quadrado mínimo.

É procurado um vector  $\mathbf{x}'$ , que melhor represente as amostras  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ . O critério de minimização é dado por

$$J_0(\mathbf{x}_0) = \sum_{k=1}^n \|\mathbf{x}_0 - \mathbf{x}_k\|^2 \quad (4.28)$$

a solução para este problema é a média,  $\mathbf{x}' = \bar{\mathbf{x}}$ , sendo esta uma representação de dimensão nula, não representa a variabilidade dos dados.

Para uma representação a uma dimensão, os dados são projectados numa recta que passa pela média. Sendo  $\mathbf{e}$  o vector direcção da recta, fica

$$\mathbf{x}_k = \bar{\mathbf{x}} + a_k \mathbf{e}, \quad (4.29)$$

onde o escalar  $a_k$ , corresponde à distância do ponto de  $\mathbf{x}_k$  à média,  $\bar{\mathbf{x}}$ .

O conjunto valores óptimos de  $a_k$  obtêm-se minimizando o critério de erro quadrado,

$$J_1(a_1, \dots, a_n, \mathbf{e}) = \sum_{k=1}^n \|(\bar{\mathbf{x}} + a_k \mathbf{e}) - \mathbf{x}_k\|^2 \quad (4.30)$$

Sendo  $\mathbf{e}$  um vector de direcção com norma unitária, derivamos a equação (4.30) em função de  $a_k$  e igualamos o resultado a zero, com o objectivo de determinar pontos extremos, obtêm-se o seguinte resultado

$$a_k = \mathbf{e}^t (\mathbf{x}_k - \bar{\mathbf{x}}) \quad (4.31)$$

a solução é obtida com a projecção do vector  $\mathbf{x}_k$  na recta de direcção  $\mathbf{e}$  que passa pela média,  $\bar{\mathbf{x}}$ . Criando-se um novo problema, o de escolher a melhor direcção da recta, ou seja a direcção do vector  $\mathbf{e}$ .

A solução para este problema é dada pela matriz de dispersão

$$\mathbf{S} = \sum_{k=1}^n (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})^t \quad (4.32)$$

obtida substituindo-se na equação (4.30) a expressão de  $a_k$  dada pela equação (4.31), obtendo-se como critério

$$J_1(\mathbf{e}) = -\mathbf{e}^t \mathbf{S} \mathbf{e} + \sum_{k=1}^n \|\mathbf{x}_k - \bar{\mathbf{x}}\|^2 \quad (4.33)$$

ou seja maximizando o termo  $\mathbf{e}^t \mathbf{S} \mathbf{e}$ , minimiza-se a função  $J_1(\mathbf{e})$ .

Recorrendo aos multiplicadores de Lagrange [22], é obtida a solução

$$\mathbf{S} \mathbf{e} = \lambda \mathbf{e} \quad (4.34)$$

ou seja,  $\mathbf{e}$  terá que ser um vector próprio da matriz de dispersão. Então para maximizar o termo  $\mathbf{e}^t \mathbf{S} \mathbf{e}$ , é seleccionado o vector próprio com maior valor próprio da matriz de dispersão  $\mathbf{S}$ . É a melhor projecção uni-dimensional, segundo critério da soma dos erros quadrados.

Este método pode ser estendido a um maior número de dimensões, sendo  $d$  o número de

dimensões após a projecção, a função de critério fica

$$J_d = \sum_{k=1}^n \left\| \left( \bar{\mathbf{x}} + \sum_{i=1}^d a_{ki} \mathbf{e}_i \right) - \mathbf{x}_k \right\|^2, \quad (4.35)$$

a minimização é resolvida quando se obtém  $d$  vectores próprios  $\mathbf{e}_i$ , sendo que este conjunto tem os maiores valores próprios da matriz de dispersão.

Este método é um método não supervisionado e é baseado unicamente nos valores próprios da matriz de dispersão, escolhendo as direcções com mais peso tornando a representação mais eficaz em vez de escolher direcções que melhor discriminem as amostras. As características que são extraídas em nada garantem que sejam as que melhor discriminam as amostras.

#### 4.5.2 Análise Discriminante de Fisher

A análise discriminante de Fisher tem em conta a escolha das melhores direcções para a discriminação das amostras. A análise discriminante de Fisher é uma análise supervisionada, considerando também as etiquetas dadas às amostras para determinar uma direcção que melhor discrimine as amostras ou classes.

A análise discriminante de Fisher envolve a maximização da função

$$J(\mathbf{W}) = \frac{|\tilde{\mathbf{S}}_B|}{|\tilde{\mathbf{S}}_W|}, \quad (4.36)$$

ou seja, pretende-se maximizar a dispersão entre classes, mas minimizar a dispersão das amostras dentro de cada classe.  $\tilde{\mathbf{S}}_B$  é a matriz de dispersão entre classes após a projecção,  $\tilde{\mathbf{S}}_W$  é a matriz de dispersão dentro de cada classe após a projecção.

A projecção de um espaço de características de  $d$  dimensões para  $d' = c - 1$  ( $c$  é o número total de classes) dimensões, obtém-se com  $d'$  funções discriminantes

$$\mathbf{y} = \mathbf{W}^t \mathbf{x}, \quad (4.37)$$

$\mathbf{W}$  é uma matriz  $d \times d'$  em que cada coluna é o vector dos pesos das características para cada classe. Ao calcular as matrizes de dispersão de  $\mathbf{y}$ ,  $\tilde{\mathbf{S}}_W$  e  $\tilde{\mathbf{S}}_B$ , e substituindo  $\mathbf{y}$  pela expressão dada por (4.37), obtemos às seguintes expressões

$$\tilde{\mathbf{S}}_W = \mathbf{W}^t \mathbf{S}_W \mathbf{W} \quad (4.38)$$

e

$$\tilde{\mathbf{S}}_B = \mathbf{W}^t \mathbf{S}_B \mathbf{W}, \quad (4.39)$$

em que  $\mathbf{S}_B$  e  $\mathbf{S}_W$  são as matrizes de dispersão das amostras  $\mathbf{x}$  e são dadas por

$$\mathbf{S}_B = \sum_{i=1}^c N_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}) (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^t \quad (4.40)$$

e

$$\mathbf{S}_W = \sum_{i=1}^c \sum_{\mathbf{x} \in \mathcal{D}_i} (\mathbf{x} - \bar{\mathbf{x}}_i)(\mathbf{x} - \bar{\mathbf{x}}_i)^t \quad (4.41)$$

onde  $N_i$  é o número de elementos da classe  $i$ .

A solução óptima de  $W$  é dada quando as colunas da matriz são os vectores próprios aos quais correspondem os maiores valores próprio de

$$\mathbf{S}_B \mathbf{w}_i = \lambda_i \mathbf{S}_W \mathbf{w}_i. \quad (4.42)$$

Sendo a análise discriminante de Fisher um método supervisionado, ou seja, por considerar as classes existentes nas amostras de treino para a obtenção da projecção, concluí-se que a análise discriminante de Fisher atinge melhor performance que a análise por PCA, as classes são melhor discriminadas. A análise discriminante de Fisher foi a escolha de método para a redução da dimensionalidade do espaço de características do nosso problema.

## 4.6 Sumário

O classificador final, é construído apartir de quatro classificadores. Classificador linear e quadrático de Bayes, o classificador k vizinhos mais próximos e classificador máquina de vector de suporte. São combinados de forma a construírem uma estrutura de votação. O espaço de características é reduzido utilizando a análise discriminante de Fischer.



## Capítulo 5

# Metodologia de pesquisa da base de dados

Neste Capítulo apresenta-se a metodologia adoptada para a pesquisa da base de dados, envolvendo as fases de ordenação, classificação e o cálculo da similaridade entre imagens, através de uma métrica apropriada.

Os algoritmos aqui descritos foram implementados usando Matlab<sup>®</sup> versão 7.5.0.342 (R2007b).

### 5.1 Ordenação: semelhança entre imagens

A ordenação de imagens por similitude é avaliada pela análise de uma função de distância. É calculada a distância entre o vector de características da imagem exemplo e das imagens da base de dados. As características podem ser globais ou locais, características locais são calculadas a partir de regiões de interesse (RDI) da imagem. Primeiramente será explicado o método de escolha das regiões de interesse, posteriormente as métricas utilizadas para a medição de similitude entre imagens e a ordenação das imagens.

#### 5.1.1 Regiões de interesse

O uso das regiões de interesse tem como objectivo melhorar os resultados da pesquisa na fase de ordenação. Usando informação local, a medida semelhança entre duas imagens baseia-se na medida de semelhança entre as suas regiões de interesse. Considera-se que uma imagem  $W$  é semelhante a uma imagem  $V$ , quando as regiões de interesse de  $W$  são semelhantes às regiões de interesse de  $V$ .

A cada imagem determinadas são 10 RDIs, a escolha de regiões está sujeita a algumas condições. A imagem é dividida por quadrantes como ilustra a Fig. 5.1, em cada quadrante serão

permitidos no máximo quatro RDIs, desta forma garante-se que as RDI estão dispersas pela imagem. Uma RDI é representada por um quadrado com 30 pixels de lado. Os pontos centrais das RDI estão a uma distância de mínima 15 *pixels* dos limites da imagem.

Após a obtenção das regiões de interesse, são medidas características de textura, cor e contorno de forma semelhante ao explicado no capítulo 3. Ao calcular os histogramas de módulo e fase das RDI considera-se todos os valores dos pixels e não somente os valores de módulo e fase no contorno, como é feito na medição das características globais da imagem descrita na secção 3.4.

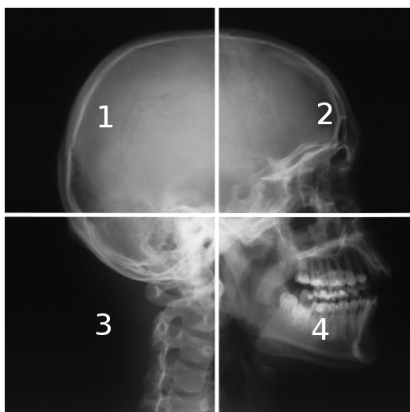


Figura 5.1: Divisão da imagem em quatro quadrantes.

#### 5.1.1.1 Determinação dos pontos centrais das regiões de interesse

A determinação dos pontos centrais das RDIs da imagem tem duas fases.

A primeira consiste em aplicar um algoritmo de detecção de cantos no contorno das imagens [25], algoritmo de Harris.

Após guardar os pontos que satisfazem as condições acima referidas, se nesta fase não se obtém as 10 regiões pretendidas essas são complementadas com os pontos máximos do gradiente da imagem, que também têm que obedecer às condições, acima referidas, a fim de obter pontos dispersos por toda a imagem.

A detecção de cantos é efectuada da seguinte forma: Dado um deslocamento  $(\Delta x, \Delta y)$  e o ponto  $(x, y)$ , a função de auto-correlação é igual a

$$c(x, y) = \sum_W [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \quad (5.1)$$

onde  $I$  é a imagem e  $(x_i, y_i)$  são os pontos da janela gaussiana  $W$  centrada em  $(x, y)$ .

A imagem deslocada  $I(x_i + \Delta x, y_i + \Delta y)$  é aproximada por uma expansão de Taylor obtendo-se

$$I(x_i + \Delta x, y_i + \Delta y) \approx I(x_i, y_i) + [I_x(x_i, y_i) I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (5.2)$$

onde  $I_x$  e  $I_y$  são as derivadas parciais ao longo de  $x$  e  $y$ , respectivamente.

Substituindo Eq. (5.2) em Eq. (5.1) obtém-se

$$c(x,y) = [\Delta x \ \Delta y] C(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (5.3)$$

A matriz  $C(x,y)$  contém informação da intensidade da estrutura da vizinhança. Sejam  $\lambda_1, \lambda_2$  os valores próprios da matriz  $C(x,y)$ . Existem três casos a ter em conta:

1. Se ambos  $\lambda_1, \lambda_2$  são pequenos, tal que a função de auto-correlação local é achatada (pequenas variações de  $c(x,y)$  em qualquer direcção), a região da imagem contida na janela tem aproximadamente intensidade constante.
2. Se um valor próprio é elevado e o outro de baixo valor, então a função de auto-correlação local tem a forma de uma crista, então só deslocações locais numa direcção (ao longo do contorno) causam pequenas variações em  $c(x,y)$  e variações significativas na direcção ortogonal; isto indica a presença de contorno.
3. Se ambos os valores próprios são elevados, então a função de auto-correlação local tem a forma de um pico, e deslocações em qualquer direcção resulta de uma variação significativa de  $c(x,y)$ ; isto indica a presença de um canto.

Quando ambos os valores próprios da matriz  $C(x,y)$  são elevados, estamos perante um candidato a RDI.

A figura 5.2 é um exemplo de uma imagem e as suas RDI.

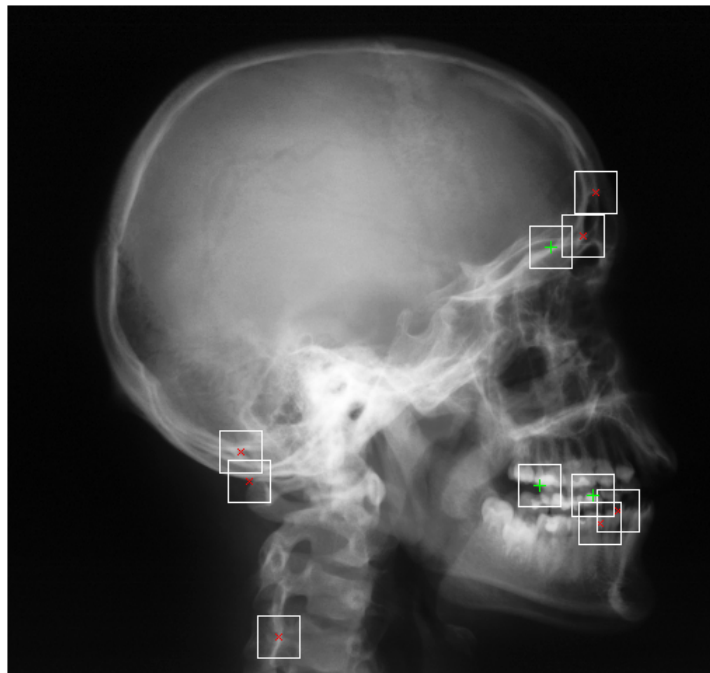


Figura 5.2: As 10 Regiões de interesse da imagem. Os pontos a vermelho são obtidos pelo algoritmo de Harris, os restantes pelo máximo do gradiente da imagem.

Para cada imagem são determinadas 10 RDI, se o número de pontos obtidos pela detecção de cantos não atinge o número de RDI pretendido, as regiões em falta são escolhidas utilizando do módulo do gradiente da imagem (5.4).

$$|I(x, y)| = \sqrt{\frac{\partial I(x, y)^2}{\partial x} + \frac{\partial I(x, y)^2}{\partial y}} \quad (5.4)$$

Os candidatos a RDI fornecidos pelo gradiente da imagem, são os pontos de máximo do gradiente  $(x_k, y_k)$  dados por

$$Max(|I(x, y)|) = \sqrt{\frac{\partial I(x_k, y_k)^2}{\partial x} + \frac{\partial I(x_k, y_k)^2}{\partial y}} \quad (5.5)$$

A escolha dos candidatos está sujeita às restrições já indicadas, se o ponto  $(x_k, y_k)$  não obedecer às restrições o valor do *pixel*, o valor na posição  $(x_k, y_k)$  da imagem do módulo do gradiente, é anulado de forma que não se torne novamente candidato a RDI.

## 5.1.2 Distâncias

As funções de distância são frequentemente utilizadas como medida de similitude entre imagens. As imagens são caracterizadas por pontos no espaço de características. A distância entre os pontos traduzem a semelhança das imagens, ou seja, quanto maior a distância entre dois pontos menor a similaridade.

A distância Euclidiana foi a função de distância utilizada para a avaliação de similitude entre vectores de características. A distância Euclidiana entre dois vectores  $\mathbf{x}$  e  $\mathbf{y}$  é dada por

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=0}^N (x_i - y_i)^2} \quad (5.6)$$

A Fig. 5.3 ilustra esta situação. A imagem exemplo, 5.3(a), é comparada com duas imagens, uma mais semelhante 5.3(b) e outra menos semelhante 5.3(c), usando uma métrica (distância Euclidiana) para medir a similaridade. As imagens são caracterizadas pelos seus vectores de características, as distância das imagens à imagem exemplo são  $d_b = 1.39e7$  e  $d_c = 4.30e8$  respectivamente para as imagens 5.3(b) e 5.3(c), provando-se assim que a menor distância,  $d_b$ , corresponde à imagem mais semelhante e a maior,  $d_c$ , à imagens menos semelhante.

### 5.1.2.1 Distâncias entre regiões de interesse

A medição é executada da seguinte forma: Seja  $r_i$  o vector de características da região  $i$  de uma imagem  $r$ ,  $r'_j$  o vector de características da região  $j$  na imagem  $r'$ ,  $N$  o número de RDIs e  $D(\mathbf{x}, \mathbf{y})$  a função de distância. A distância entre regiões de interesse de duas imagens é dada por

$$d^{rdi} = \sum_{k=1}^N \min_{h \in [1, N]} (D(\mathbf{r}_k, \mathbf{r}'_h)) \quad (5.7)$$

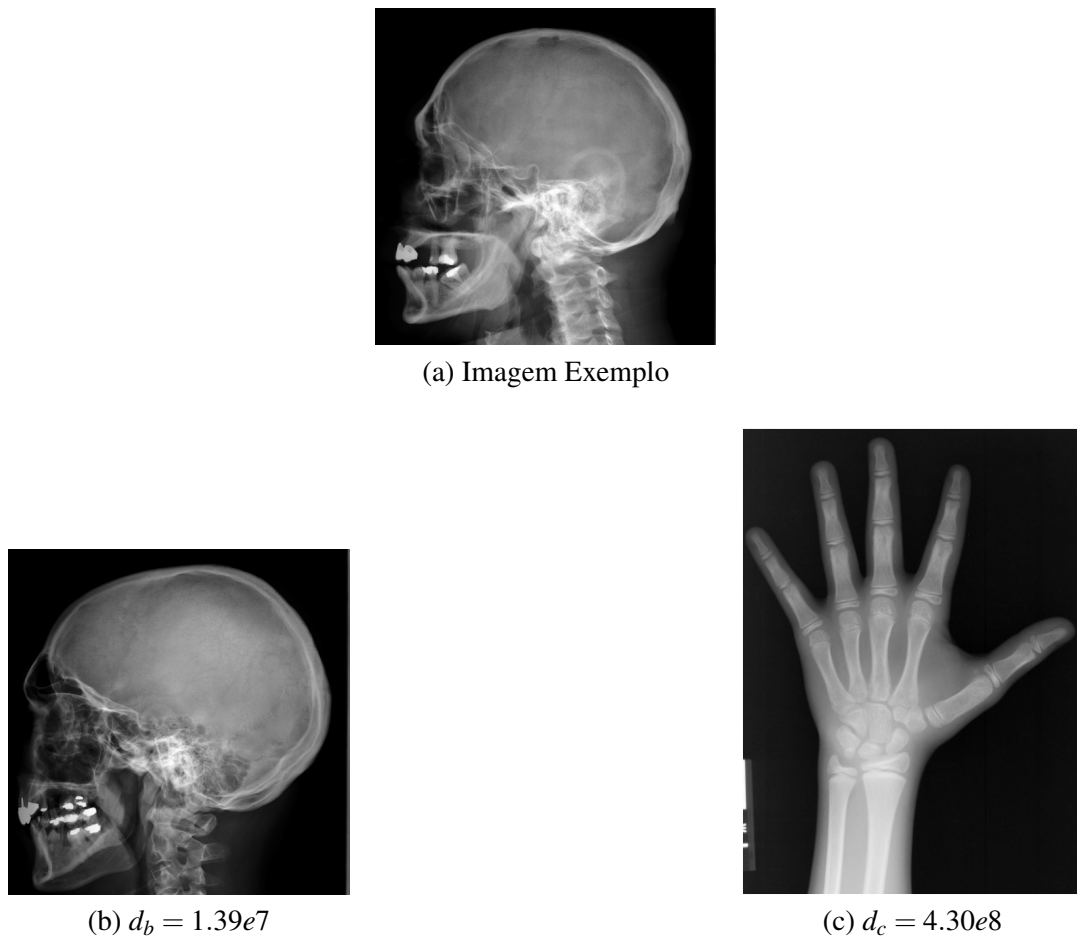


Figura 5.3: Ilustração do uso de uma métrica como medida de similitude.

A função de distância,  $D$ , utilizada na medição de similaridade é a distância Euclidiana dada pela equação (5.6).

Na Fig. 5.4 dá-se o exemplo da medição de similaridade entre duas imagens através das suas RDIs. Cada RDI está caracterizada por um vetor de características. O valor da distância é obtido pela equação (5.7). Verifica-se que imagem mais semelhante 5.4(b) tem a distância menor,  $d_b^{dis} = 7.40e6$ , enquanto a menos semelhante tem uma distância superior,  $d_c^{dis} = 1.39e7$ .

### 5.1.3 Ordenação

A ordenação sem considerar regiões de interesse é feita baseando nos resultados da função de distância, as imagens são ordenadas em ordem crescente, pois o menor valor de distância corresponde à imagem mais semelhante.

A ordenação considerando as RDIs tem duas fases. A primeira fase consiste em avaliar a similitude usando características globais e ordenando por similaridade. A segunda fase avalia a similitude das imagens do conjunto dos primeiros  $S$  resultados com a imagem exemplo usando características das RDIs.

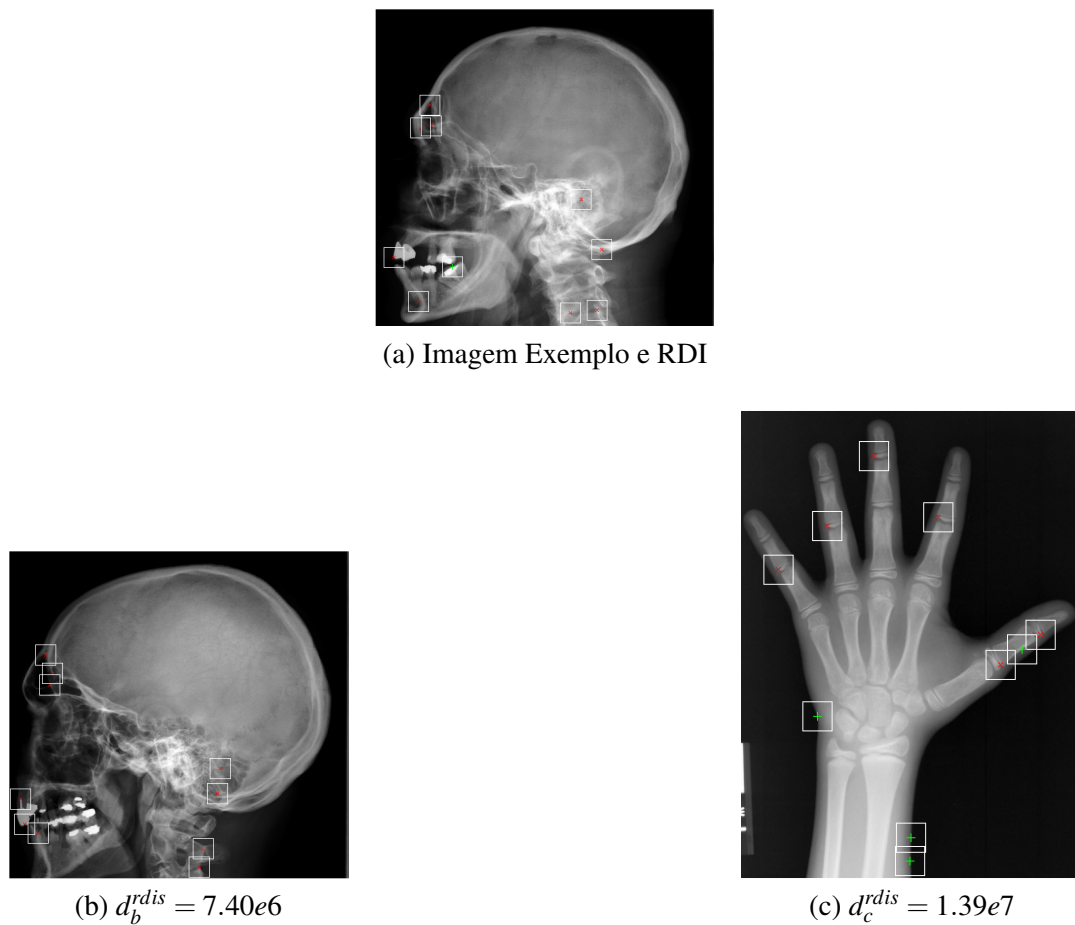


Figura 5.4: Ilustração do uso de uma métrica como medida de similitude entre RDIs de duas imagens.

## 5.2 Classificação e ordenação

A classificação é efectuada na abordagem de ordenação parcial, cujo objectivo é a diminuição do espaço de pesquisa, melhorando os resultados e o desempenho computacional.

### 5.2.1 Classificação

O objectivo da classificação nesta abordagem é atribuição de uma classe à imagem exemplo, permitindo diminuir o espaço de pesquisa, tornando a pesquisa mais rápida (menos comparações) e retornar melhores resultados.

A base de dados contém 1000 imagens e está dividida em três classes, A, B e C. A classe A, Fig. 5.5, é composta por imagens médicas de estruturas ósseas e seios. É uma classe composta por 600 imagens.

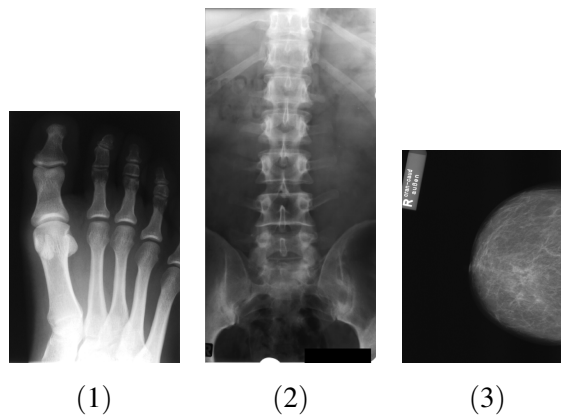


Figura 5.5: Imagens pertencentes a classe A. (1) Estrutura óssea de um pé. (2) Espinha lombar. (3) Mama.

A classe B, Fig. 5.6, é constituída por imagens médicas de pulmões, constituída por 303 imagens.

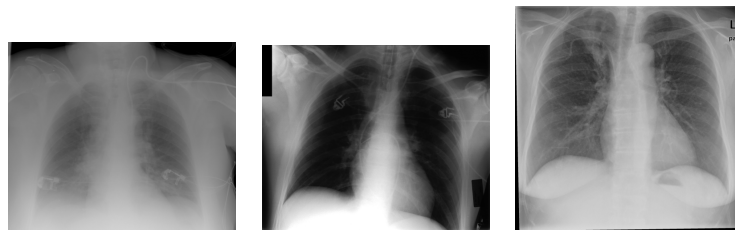


Figura 5.6: Imagens pertencentes a classe B, pulmões.

A classe C, Fig. 5.7, é constituída por imagens médicas de pulmões na posição sagital, em toda a base de dados existem 97 imagens da classe C.

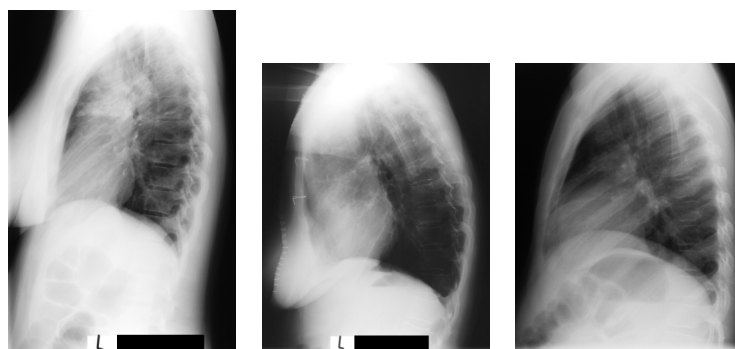


Figura 5.7: Imagens pertencentes a classe C, pulmões posição sagital.

### 5.3 Relevância

O sistema de relevância adotado consiste do deslocamento do ponto de consulta [26][16]. Dada a medida de relevância  $r_{fi}$  são definidos os seguintes conjuntos,  $S_{rel} = a_i | r_{fi} > 0$  como

sendo o conjunto de pontos relevantes e  $S_{non-rel} = a_i | r f_i < 0$  o conjunto de pontos não relevantes. O novo ponto de consulta é dado por

$$p_{new} = \alpha p_{old} + \frac{\beta}{|S_{rel}|} \sum_{a \in S_{rel}} a_j - \frac{\gamma}{|S_{non-rel}|} \sum_{a_j \notin S_{rel}} a_j \quad (5.8)$$

A velocidade à qual o ponto se move em direcção aos pontos relevantes é controlada pelos pesos  $\alpha$ ,  $\beta$  e  $\gamma$  onde  $\alpha + \beta + \gamma = 1$ .

Na Fig. 5.8 está ilustrado o funcionamento do método de relevância, neste sistema.

Foram sintetizadas duas distribuições gaussianas, a cada distribuição foi atribuída uma classe (1 e 2). O ponto de consulta pertence à classe 1, ou seja, o novo ponto de consulta terá que se deslocar para a esquerda do espaço de características.

A abordagem é igual à descrita na Secção 5.1, isto é, são obtidos os 15 pontos mais semelhantes, perante os resultados é avaliada a relevância e calculado o novo ponto de consulta.

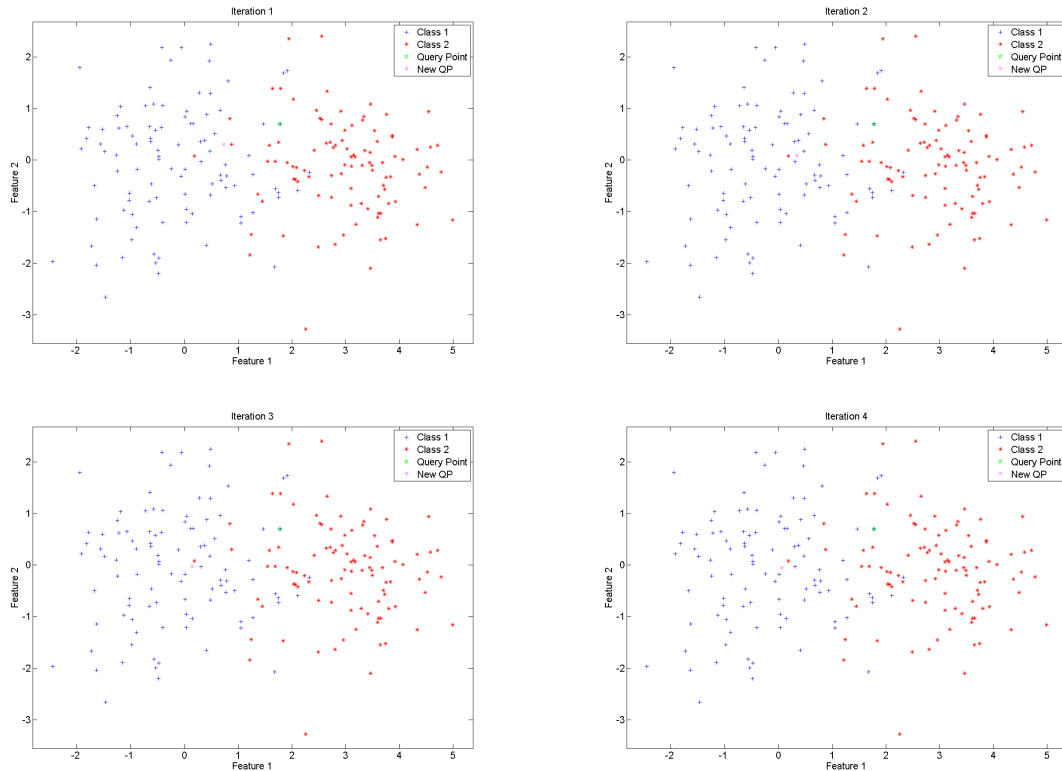


Figura 5.8: Exemplo ilustrativo do deslocamento do ponto de consulta.

## 5.4 Sumário

A ordenação de imagem está em muito ligada às funções de distâncias nestas abordagens. Foram apresentadas as restrições e o método de escolha das regiões de interesse das imagens, as

métricas usadas e a forma como a informação local das imagem é incorporada na ordenação as imagens por similitude. O método de medida de relevância usado foi o método de deslocamento do ponto de consulta.



## Capítulo 6

# Análise de Resultados

Neste capítulo pretende-se analisar resultados da solução apresentada. Identificar os problemas e fazer uma análise quantitativa da solução.

### 6.1 Exemplos

Nesta secção serão analisados e discutidos alguns exemplos ilustrativos.

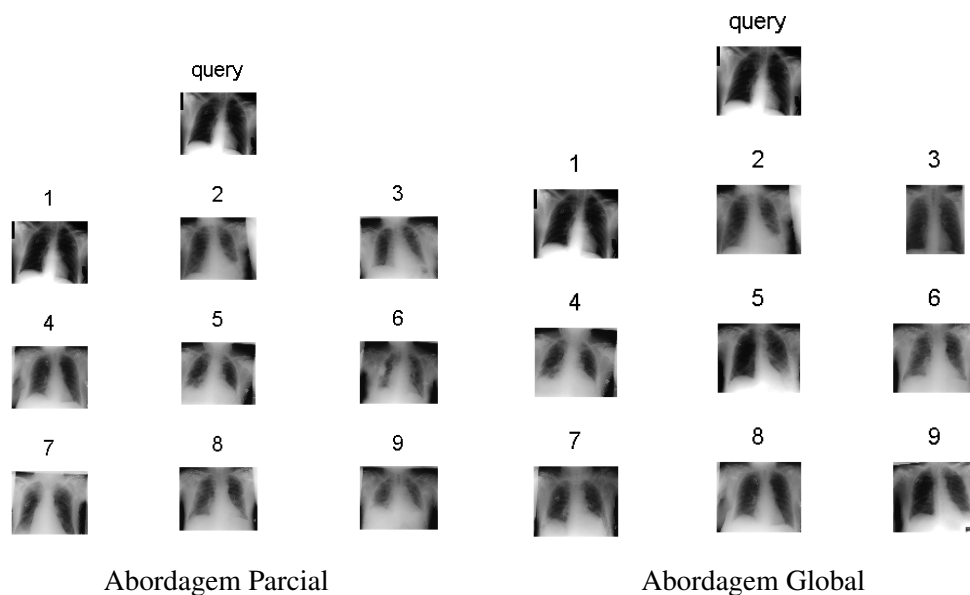


Figura 6.1: Exemplo 1 - obtenção de imagens semelhantes

Analisando exemplo da Fig. 6.1, verifica-se que os resultados das duas abordagens são bastante similares. Na abordagem parcial, a imagem exemplo (*query*) foi classificada correctamente como pertencente à classe B, a pesquisa é realizada no conjunto de imagens pertencentes à classe

B. Na abordagem global a pesquisa é exaustiva, percorrendo toda a base de dados. Verifica-se que o sistema retorna bons resultados nas duas abordagens.

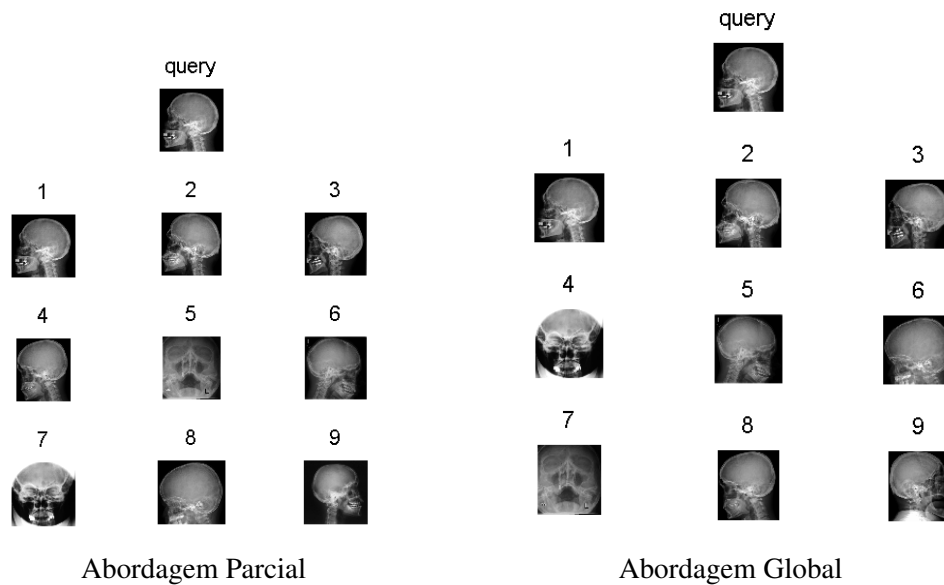


Figura 6.2: Exemplo 2 - obtenção de imagens semelhantes

A Fig. 6.2 ilustra o exemplo de uma imagem pertencente à classe A. Os resultados obtidos são bastante satisfatórios, as duas abordagens retornam as imagens mais semelhantes, porém contêm duas imagens que são consideradas não relevantes. Novamente na abordagem parcial a imagem foi classificada correctamente, logo o espaço de pesquisa foi reduzido para 600 imagens, ao contrário da abordagem global cujo espaço de pesquisa é de 1000 imagens.

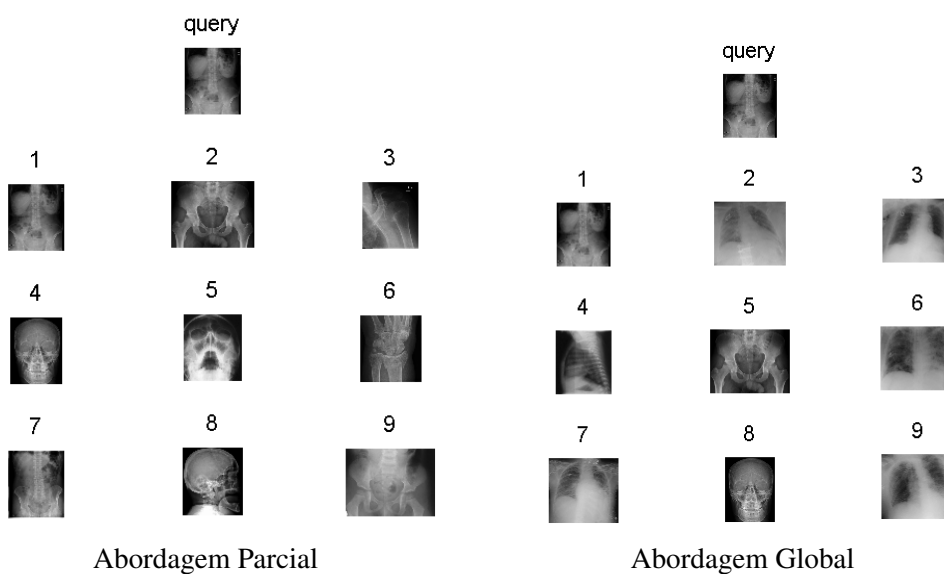
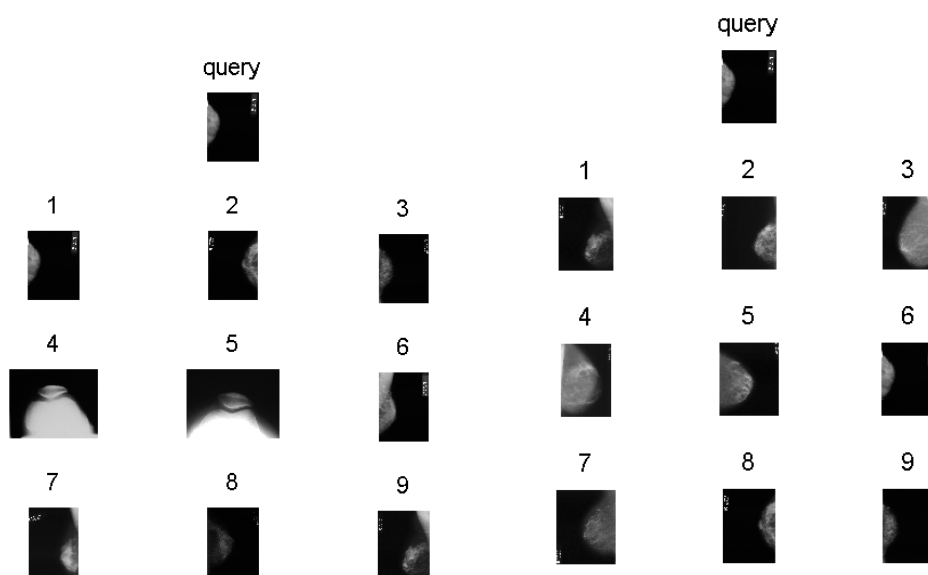


Figura 6.3: Exemplo 3 - obtenção de imagens semelhantes

O exemplo 3, Fig. 6.3, é a ilustração de um resultado menos satisfatório. Apesar de a abordagem parcial retornar 4 imagens semelhantes visualmente. Para imagens de *query* deste tipo o espaço de características não é suficiente para as discriminar. Neste exemplo nota-se a desvantagem da pesquisa exaustiva na base de dados, a abordagem global devolve apenas uma imagem relevante (a própria).

O exemplo da Fig. 6.4 ilustra a vantagem da utilização de RDIs na fase final da ordenação das imagens mais semelhantes. Verifica-se que as duas imagens não semelhantes desaparecem se for utilizada informação local. No entanto esta melhoria só se verifica em alguns tipos de imagens. A utilização de RDIs necessita de ser desenvolvida, no entanto está provado o seu potencial.



Abordagem Global sem informação de RDI    Abordagem Global com informação RDI

Figura 6.4: Exemplo 4 - obtenção de imagens semelhantes

Na Fig. 6.5 está ilustrada outra situação de erro, resultando da utilização do bloco de classificação da abordagem parcial. O classificador não é ideal portanto o erro de classificação não é nulo. Levando portanto a erros de classificação deste tipo: a imagem *query* é da classe A e é classificada como pertencente à classe C.

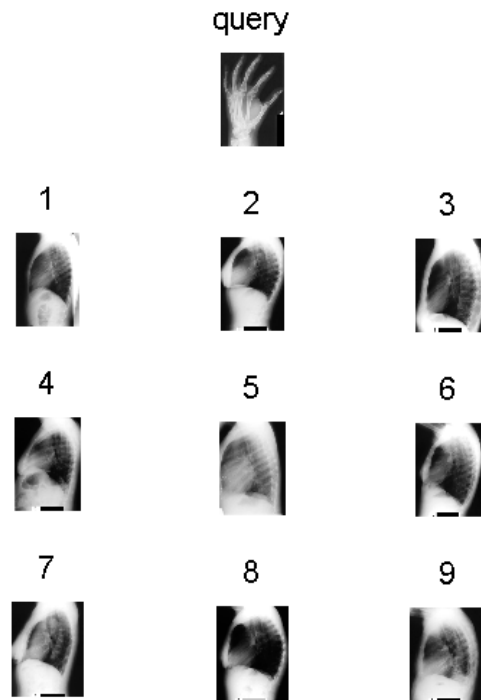


Figura 6.5: Exemplo 5 - Erro de classificação na obtenção das imagens mais semelhantes

### 6.1.1 Relevância

Na Fig. 6.6 está o resultado de uma iteração do método de relevância utilizado. O método funciona, como foi ilustrado no capítulo anterior na Fig. 5.8, mas não no nosso espaço de características. Conclui-se deste resultado que dentro da classe, as características usadas não discriminam as amostras entre si.

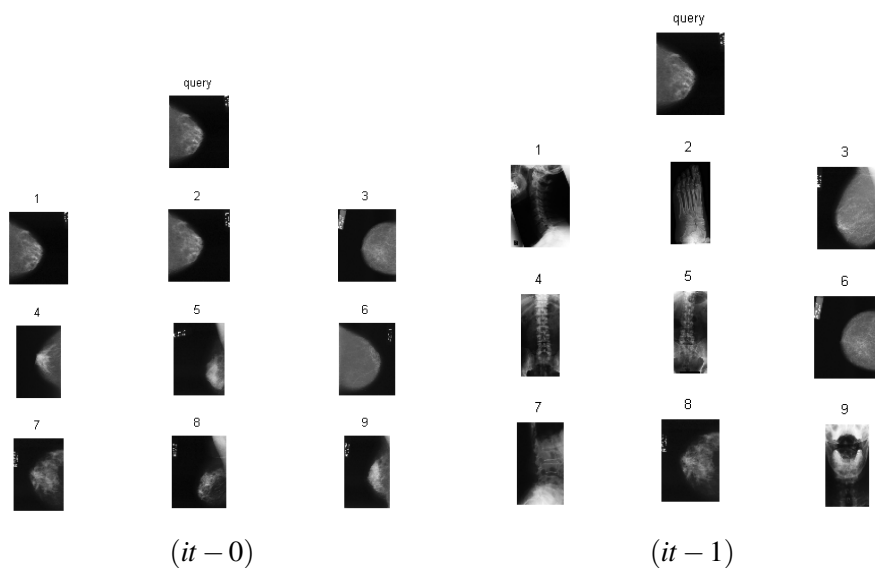


Figura 6.6: Exemplo 6 - Relevância

## 6.2 Avaliação Quantitativa

Nesta secção pretende-se fazer uma análise quantitativa do sistema, apresentando valores que caracterizam a performance do sistema. A análise é feita em termos de precisão e erro de classificação.

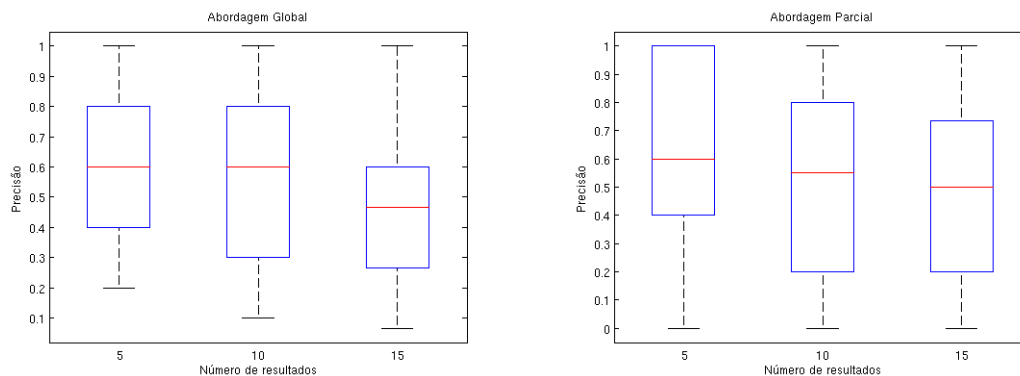


Figura 6.7: Gráficos box-plot de precisão do sistemas em função do número de imagens devolvidas pelo sistema.

Para a avaliação do sistema é calculada precisão a partir da expressão

$$P = \frac{N_{rel}}{N_{rel} + N_{nrel}} \quad (6.1)$$

onde  $N_{rel}$  é número de imagens relevantes do resultado e  $N_{nrel}$  o número de imagens não relevantes, ou seja  $N_{rel} + N_{nrel} = N_{ret}$  é o número de imagens retornadas pelo sistema. Foram realizados 100 testes ao sistema com três valores de  $N_{ret}$  (5, 10 e 15 imagens retornadas), calculada a precisão em cada teste e determinada a média. Os resultados estão ilustrados nos gráficos da Fig. 6.7 e os valores indicados na tabela 6.1.

Tabela 6.1: Tabela de valores médios da precisão do sistema

$N_{ret}$	Abordagem Parcial	Abordagem Global
5	0.620	0.620
10	0.543	0.537
15	0.493	0.480

Para o treino do classificador, a base de dados foi dividida entre treino e teste. A divisão é aleatória, ou seja, os conjuntos de treino e teste não são sempre iguais. Para determinar o erro médio de classificação, o classificador foi treinado e testado 100 vezes e em cada ocorrência foi armazenado o erro de classificação.

O classificador tem uma média de erro de classificação de 5.99% com uma variância de  $1.30e - 04$ .

Nas tabelas 6.2 e 6.3, estão caracterizados os classificadores indicando o erro médio de cada classificador e o número médio de amostras em erro por classe,  $Ec_A$ ,  $Ec_B$  e  $Ec_C$ .

Tabela 6.2: Erros médios dos classificadores

Classificador	Erro médio (%)
LDC	6.09
QDC	6.28
KNN	6.13
SVM	6.37
VOTE	5.99

Tabela 6.3: Número médio de erros por classe.

Classificador	$E_{c_A}$	$E_{c_B}$	$E_{c_C}$
LDC	8.82	3.72	7.70
QDC	9.91	3.63	7.34
KNN	7.35	4.02	9.01
SVM	6.22	3.21	11.7
VOTE	7.37	3.63	8.93

### 6.3 Sumário

Obteve-se um erro de classificação de 5.99% e uma precisão máxima de 62%.

No método de relevância não se obtiveram bons resultados devido ao espaço de características não discriminar bem as amostras. O método foi somente testado com imagens da classe A, ou seja, existe a possibilidade que o método funcione com as outras classes dependendo do poder de discriminação do espaço de características dentro da classe.

## Capítulo 7

# Conclusões e Trabalho Futuro

A medição de características e a medição de similitude entre imagens, são uma parte integrante num sistema CBIR. As medições de características de imagens permitem a construção de vectores representativos das imagens. As medidas de similitude permitem a comparação entre imagens em termos de semelhança.

A classificação pode ser incluída num sistema CBIR tendo como principal objectivo reduzir o espaço de pesquisa, para um espaço constituído apenas por imagens pertencentes à mesma classe que a imagem exemplo.

A relevância possibilita ao utilizador interagir com o sistema, refinando a pesquisa de acordo com os seus objectivos.

No sistema apresentado são retiradas medidas de textura, cor e contorno.

As medidas de textura e cor são obtidas construindo uma imagem de *textons*, através de um conjunto pré-definido de *textons*. Da imagem de *textons* é calculada a matriz de co-ocorrências dos *pixels*. Uma vez calculada a matriz de co-ocorrências são medidas as características: energia, contraste, entropia e homogeneidade.

A medida de contorno, é calculada pela medição de características dos histogramas dos valores de módulo e fase do gradiente da imagem ao longo do contorno. As características são a média, a variância, a *skewness*, a *kurtosis*, e os percentis 75% e 90%.

O classificador por votação, utilizado na abordagem de ordenação parcial, é construído a partir de quatro classificadores: Classificador linear e quadrático de Bayes, o classificador k vizinhos mais próximos e classificador máquina de vector de suporte. São combinados de forma a construir uma estrutura de votação. O espaço de características é reduzido utilizando a análise discriminante de Fischer.

A ordenação das imagens por semelhança está directamente relacionada com as funções de distâncias. A função de distância utilizada é a distância Euclidiana.

Foi criado um método de detecção de regiões de interesse das imagens e de ordenação de imagens semelhantes considerando características locais, regiões de interesse.

O método de medida de relevância usado foi o método de deslocamento do ponto de consulta. Este método provou-se eficaz, contudo não o foi no espaço de características da solução apresentada. Devendo-se ao facto do espaço de características usado não discriminar as amostras dentro de cada classe.

Apresentou-se um novo sistema de CBIR para aplicações em imagens médicas. Foram implementadas e analisadas duas abordagens provando que as duas obtiveram resultados razoáveis. Existem vantagens e desvantagens nas duas abordagens. A abordagem de ordenação global apesar de obter piores resultados não tem associado um erro de classificação ao contrário da abordagem de ordenação parcial, com um erro de classificação de 5.99%. Contudo a pesquisa na abordagem de ordenação global é feita em toda a base de dados ao contrário da abordagem parcial em que a pesquisa é realizada somente dentro da classe atribuída à imagem exemplo.

Obteve-se com o sistema uma precisão máxima de 62% em 100 testes.

## 7.1 Trabalho Futuro

Como trabalho futuro a integração de novas características globais a fim de diminuir o erro de classificação, de discriminar melhor as imagens da mesma classe e de melhorar a eficácia do módulo de Relevância poderá ser uma estratégia.

O desenvolvimento da utilização de RDIs com vista a melhorar o espaço de características, adicionando novas características ou ajustando as restrições, por exemplo, o aumento do tamanho da janela poderá ser uma hipótese a considerar. A utilização de RDIs tem muito potencial, mas não foi desenvolvida exaustivamente.

Procurar uma organização da base de dados a fim de evitar uma pesquisa exaustiva. Uma possível solução seria que cada imagem adquirisse conhecimento da sua vizinhança, ou seja, cada imagem conhece as suas imagens mais semelhantes.

A exploração de novas medidas de similitude também é uma estratégia a considerar, a medida de similitude usada é a medida mais básica havendo a possibilidade de se obter melhores resultados usando outras medidas.

# Referências

- [1] Md.M. Rahman, P. Bhattacharya, and B.C. Desai. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. *Information Technology in Biomedicine, IEEE Transactions on*, 11(1):58–69, Jan. 2007.
- [2] Guang-Hai Liu and Jing-Yu Yang. Image retrieval based on the texton co-occurrence matrix. *Pattern Recogn.*, 41(12):3521–3527, 2008.
- [3] Coffman K. G. and Odlyzko Andrew. The size and growth rate of the internet. 1999.
- [4] Chabane Djeraba Michael S. Lew, Nicu Sebe and Ramesh Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(1):1–19, 2006.
- [5] N. Vasconcelos. From pixels to semantic spaces: Advances in content-based image retrieval. *Computer*, 40(7):20–26, July 2007.
- [6] Halina Kwasnicka and Mariusz Paradowski. Resulted word counts optimization—a new approach for better automatic image annotation. *Pattern Recognition*, 41(12):3562 – 3571, 2008.
- [7] R.C.F. Wong and C.H.C. Leung. Automatic semantic annotation of real-world web images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(11):1933–1944, Nov. 2008.
- [8] Kobus Barnard, Pinar Duygulu, David Forsyth, Nando de Freitas, David M. Blei, and a Michael I. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3(6):p1107 – 1135, Aug. 2003.
- [9] Jia Li and J.Z. Wang. Real-time computerized annotation of pictures. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):985–1002, June 2008.
- [10] S. Liapis and G. Tziritas. Color and texture image retrieval using chromaticity histograms and wavelet frames. *Multimedia, IEEE Transactions on*, 6(5):676–686, Oct. 2004.
- [11] J Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 8(6):679–698, 1986.
- [12] B. Boutsinas and T. Papastergiou. On clustering tree structured data with categorical nature. *Pattern Recognition*, 41(12):3613 – 3623, 2008.
- [13] Xiang S. Zhou and Thomas S. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, 8(6):536–544, April 2003.

- [14] Michel Crucianu, Marin Ferecatu, and Nozha Boujemaa. Relevance feedback for image retrieval: a short survey. In *In State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access and Interaction including Datamodels and Languages (DELOS2 Report)*, 2004.
- [15] Giorgio Giacinto and Fabio Roli. Instance-based relevance feedback for image retrieval. In Lawrence K. Saul, Yair Weiss, and Léon Bottou, editors, *Advances in Neural Information Processing Systems 17*, pages 489–496. MIT Press, Cambridge, MA, 2005.
- [16] Michael Ortega-Binderberger. Relevance feedback techniques in the mars image retrieval system. *Multimedia Systems*, 9:535–547(13), June 2004.
- [17] Danzhou Liu, Kien A. Hua, Khanh Vu, and Ning Yu. Fast query point movement techniques for large cbir systems. *IEEE Trans. on Knowl. and Data Eng.*, 21(5):729–743, 2009.
- [18] Xiang Peng and Irwin King. A biased minimax probability machine-based scheme for relevance feedback in image retrieval. *Neurocomputing*, 72(7-9):2046 – 2051, 2009. Advances in Machine Learning and Computational Intelligence - 16th European Symposium on Artificial Neural Networks 2008, 16th European Symposium on Artificial Neural Networks 2008.
- [19] D. Tao, Xuelong Li, and S.J. Maybank. Negative samples analysis in relevance feedback. *Knowledge and Data Engineering, IEEE Transactions on*, 19(4):568–580, 2007.
- [20] Silvano Di Zeno. A note on the gradient of a multi-image. *Computer Vision Graphics And Image Processing*, 33:116–125, 1986.
- [21] F van der Heijden, R.P.W. Duin, D. de Ridder, and D.M.J. Tax. *Classification, Parameter Estimation and State Estimation*. Wiley, 2004.
- [22] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley, 2004.
- [23] Nello Cristianini and John Shawe-Taylor. *An introduction to support Vector Machines: and other kernel-based learning methods*. Cambridge University Press, New York, NY, USA, 2000.
- [24] Josef Kittler, Mohamad Hatef, Robert P.W. Duin, and Jiri Matas. On combining classifiers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(3):226–239, 1998.
- [25] Chris Harris and Mike Stephens. A combined corner and edge detector. In *The Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [26] J.J. Rocchio. Relevance feedback in information retrieval. *The SMART retrieval system: experiments in automatic document processing*, pages 313–323, 1971.