

# UM SISTEMA DE CBIR PARA TRIAGEM VIRTUAL DE IMAGENS DE PATOLOGIA

**Trabalho de Conclusão de Curso**

**Engenharia da Computação**

**Higor Neto Lima**

**Orientador: Prof. Wellington Pinheiro dos Santos**

**Co-Orientador: Mêuser Jorge Silva Valença**



UNIVERSIDADE  
DE PERNAMBUCO

**HIGOR NETO LIMA**

**UM SISTEMA DE CBIR PARA  
TRIAGEM VIRTUAL DE IMAGENS DE  
PATOLOGIA**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia da Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

**Recife 21, dezembro de 2009.**

*A Zélia, Renalíto, Rany e Bruna.*

# Agradecimentos

Agradeço a todos que me apoiaram para o cumprimento desta disciplina principalmente aos meus pais e minha irmã que deram todo o apoio e tranqüilidade para que isso pudesse acontecer e contribuir com a minha formação.

Aos amigos que ajudaram com conversas técnicas ou não-técnicas que ajudaram no trabalho e na forma de encarar os problemas que surgiram ao logo do tempo de implementação e escrita da Monografia.

Ao meu orientador Wellington Pinheiro dos Santos que sempre com sua tranqüilidade contribuiu para que esse trabalho pudesse contribuir para minha formação na universidade. Ao meu Co-orientador Mêuser Jorge Silva Valença que contribuiu com seu pragmatismo na hora de implementar apenas aquilo que atendesse ao meu problema sem perder o foco.

A minha namorada Bruna Katharine que sempre me apoiou nesse período e teve muita paciência com um trabalho que tomou as madrugadas e de certo modo tirou um pouco de meu tempo com ela. Amow Te Amo.

# Resumo

Recuperação de imagens baseada em conteúdo (*Content-Based Image Retrieval* - CBIR) é uma forma de buscar imagens por meio de uma imagem base usando suas características de acordo com o interesse do usuário. Com o grande aumento na geração de imagens, desde a década de 80 vem-se buscando melhorar a busca de imagens. Para se buscar as imagens de interesse são usadas características das imagens tais como cor, textura e forma. Essas características representam imagens na forma de um vetor. Na área médica existe uma necessidade de se obter imagens de uma maior relevância em relação à patologia estudada. Este trabalho usou imagens de imuno-histoquímica de pulmão e placenta. Foram utilizadas imagens coloridas, representadas por vetores de características obtidos pela concatenação dos espectros de padrões de cada banda espectral. Espectros de padrões são descritores morfológicos definidos em Morfologia Matemática. Esses descritores são baseados em operações não lineares dependentes de um determinado elemento básico denominado de elemento estruturante. Assim, os vetores de características possuem informação de forma e de cor. As imagens utilizadas foram segmentadas para se obter apenas as informações necessárias sem o uso de características redundantes para a construção do vetor de características. Esses vetores são combinados e, usando Análise de Componentes Principais, é reduzida a dimensionalidade dos vetores, procurando melhorar a busca das imagens. O sistema proposto obteve resultados que podem indicar uma abordagem capaz de auxiliar nos diagnósticos feitos por patologistas usando informações de imagens semelhantes devidamente agrupadas.

# Abstract

Content-Based Image Retrieval is a set of techniques to perform user-defined searches in image databases using an input image as a key. This field of knowledge has been improved since the 1980's. To perform image seek in image databases it is necessary to model each image as an feature vector composed by certain features like color, texture and shape. In health sciences it is very important to organize images according to their application and relevance to the pathology under analysis. This work employed immuno-histochemical images of placenta and newdeads lungs. Image features were extracted using Pattern Spectra, nonlinear mathematical tools based on Mathematical Morphology. Morphological operators are always based on structure elements, i.e. smaller images used to rebuild the images under analysis according to a specific minimum structure. Therefore Pattern Spectra are composed by shape features in case we analyze monospectral images. In our case of study we used color images. Consequently, the set of pattern spectra we generated also carry color information. The feature vectors are composed by the concatenation of the pattern spectra obtained for each spectral band. Afterwards, these feature vectors are analyzed using Principal Component Analysis, to reduce the dimensionality of the feature vectors and improve image search performance. The proposed approach got results good enough to lead us to conclude the following: CBIR could be an important resource to support health professionals to perform diagnoses in Pathology using well-organized image sets.

# Sumário

## Conteúdo

<b>Resumo</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Sumário</b>	<b>iii</b>
<b>Índice de Figuras</b>	<b>vi</b>
<b>Índice de Tabelas</b>	<b>vii</b>
<b>Tabela de Símbolos e Siglas</b>	<b>viii</b>
<b>Capítulo 1 Introdução</b>	<b>1</b>
1.1 Organização do Trabalho	2
<b>Capítulo 2 Recuperação de Imagens Baseada em Conteúdo</b>	<b>4</b>
2.1 Introdução	4
2.2 Atributos das Imagens	7
2.2.1 Cor	7
2.2.2 Textura	9
2.2.3 Forma	11
2.3 Conclusão	12
<b>Capítulo 3 Morfologia Matemática</b>	<b>13</b>
3.1 Introdução	13
3.2 Operadores Matemáticos	14
3.2.1 Dilatação e Erosão	14
3.2.2 Abertura e Fechamento	17
3.2.3 Propriedades da Abertura e Fechamento	19
3.3 Espectro de Padrões	21

3.3.1	Granulometria	21
3.4	Conclusão	22
<b>Capítulo 4 Análise de Componentes Principais</b>		<b>23</b>
4.1	Introdução	23
4.2	Análise de Componentes Principais	23
4.2.1	Descorrelação de Dados	24
4.2.2	Análise de Componentes Principais – PCA	25
4.3	Conclusão	25
<b>Capítulo 5 Métodos de Classificação</b>		<b>26</b>
5.1	Introdução	26
5.2	K-Médias	27
5.3	Mapas Auto-Organizados - LVQ	29
5.3.1	Quantização Vetorial por Aprendizagem	30
5.4	Perceptron de Múltiplas Camadas	32
5.5	Conclusão	33
<b>Capítulo 6 Triagem Virtual</b>		<b>34</b>
6.1	Introdução	34
6.2	Caracterização do Problema	34
6.3	Proposta de Solução	35
6.3.1	Pré-processamento	37
6.3.2	Segmentação	37
6.3.3	Espectro de Padrões	37
6.3.4	Análise PCA	37
6.3.5	Classificação com Redes Neurais	38
6.4	Conclusão	38



<b>Capítulo 7 Experimentos e Resultados</b>	<b>39</b>
7.1 Introdução	39
7.2 Conjunto de Dados	39
7.3 Vetores de Características	40
7.4 Resultado das Classificações	41
7.5 Discussão e Conclusões	45
<b>Capítulo 8 Conclusão e Trabalhos Futuros</b>	<b>47</b>
<b>Bibliografia</b>	<b>49</b>

# Índice de Figuras

<b>Figura 2.1</b> Imagem original.....	8
<b>Figura 2.2</b> Imagem saturada.....	8
<b>Figura 2.3</b> Histograma de imagem original.....	8
<b>Figura 2.4</b> Histograma de imagem saturada.....	8
<b>Figura 2.5</b> Similaridade baseados em contorno e região [19]. ....	12
<b>Figura 3.1</b> Resultado da dilatação na imagem np001a. ....	16
<b>Figura 3.2</b> Resultado da erosão da imagem np001a.....	17
<b>Figura 3.3</b> (a) Imagem original e (b) Imagem depois de aplicada a abertura [24]. ...	18
<b>Figura 3.4</b> (a) Imagem original e (b) Imagem depois de aplicado o fechamento [24]. .....	19
<b>Figura 4.1</b> Passo 1. ....	28
<b>Figura 4.2</b> (a) Passo 2 e (b) Passo 3.....	28
<b>Figura 4.3</b> Algoritmo converge resultando nos clusters finais.....	29
<b>Figura 4.4</b> Grade do arranjo de uma rede SOM [30]. ....	30
<b>Figura 6.1</b> Diagrama da proposta.....	36
<b>Figura 7.1.</b> Tipos de imagens: (a) np9910a, (b) np9910b, (c) p9918a e (d) p9918a.	40

# Índice de Tabelas

<b>Tabela 7.1</b> Divisão dos dados para treino, validação e teste.....	42
<b>Tabela 7.2</b> Divisão dos dados para treino e teste nos tipos de imagens.....	42
<b>Tabela 7.3</b> Resultados da primeira etapa da classificação.....	43
<b>Tabela 7.4</b> Resultados da primeira etapa da classificação.....	44

# Tabela de Símbolos e Siglas

CBIR – Content-Based Image Retrieval

MRI – Magnetic Resonance Imaging

CT – Computer Tomography

PACS – *Picture Archiving and Communication System*

QBIC – *Query By Image Content*

MRF – *Markov Random Fields*

PDI – Processamento Digital de Imagens

MM – Morfologia Matemática

MLP – *Multilayer Perceptron*

SOM – *Self-Organized Maps*

LVQ – *Learning Vector Quantization*

PCA – *Principal Components Analysis*

# Capítulo 1

## Introdução

A popularização do uso de imagens hoje é uma realidade, devido a uma grande facilidade de geração por meio de câmeras digitais ou aparelhos hospitalares (eletrocardiograma, ressonância magnética etc), que podem gerar imagens de altíssima qualidade. Esse avanço tem contribuído para um crescente uso de imagens em áreas como Medicina, Geologia, Botânica entre outros, assim como o uso pessoal.

Com esse aumento na geração de imagens, grandes quantidades estavam ficando sem utilidade, já que não se conseguia recuperar imagens de maneira satisfatória. Porém, surgiram na década de 80 [1] os primeiros artigos sobre o assunto.

Nos anos de 1990, o uso de banco de dados de imagens de grande escala que utilizavam informações textuais criava anotações referentes a imagens de maneira manual [2]. Com essas anotações surgia o problema da subjetividade das anotações, que são na sua essência, percepções particulares de cada pessoa. As respostas geradas pelo sistema podem ter resultados irrelevantes, adicionando ao sistema inconfiabilidade [1][2].

Recuperação de imagens baseada em conteúdo (*Content-Based Image Retrieval* - CBIR) é uma forma de buscar imagens por meio de uma imagem base usando suas características de acordo com o interesse do usuário. As características mais usadas são cor, textura e forma, mas toda informação que possa representar a imagem pode ser usada. Foi estabelecido como sendo as bases para a recuperação de imagens as seguintes características [3]:

1. Extração de características visuais: dependendo da finalidade do sistema podem usar características citadas anteriormente como cor, textura e forma assim como as anotações inseridas nas bases de dados;
2. Indexação multidimensional: é a representação da imagem na base de dados que normalmente é um vetor numérico e pode ser n-dimensional;

3. Projeto de sistema de recuperação: quando o usuário fornece uma imagem, um vetor é extraído dessa imagem e ele será usado para ser comparado com os vetores das imagens na base de dados [4].

Com as técnicas de CBIR existe um problema para se estabelecer a relação entre as características da imagem e a consulta do usuário. Tal problema está relacionado ao *gap* semântico, que é a diferença entre características de baixo nível, extraídas da imagem, e o alto nível da informação desejada pelo usuário.

Na área médica, com o desenvolvimento de aparelhos de alta definição, temos uma geração de imagens bastante grande, existindo assim, uma grande demanda por bons sistemas de recuperação de imagens. Esse trabalho usará uma base de imagens médicas de imuno-histoquímica da placenta e do pulmão, mas o algoritmo pode ser aplicado a qualquer tipo de base de imagens.

Nesse contexto, o objetivo desse trabalho é desenvolver uma nova abordagem para melhorar a recuperação de imagens. Também, deve fazer a triagem de maneira a obter uma maior exatidão em comparação com os dados reais. Essa abordagem vai testar o uso de técnicas conhecidas e aplicá-la ao problema a fim de obter melhores resultados na recuperação de imagens.

## 1.1 Organização do Trabalho

Este trabalho foi dividido em seis capítulos. Neste primeiro apresenta-se as considerações iniciais e objetivos do trabalho.

O Capítulo 2 introduz conceitos de CBIR e extração de características, expõe os principais atributos utilizados para descrever o conteúdo de uma imagem, com ênfase no atributo forma e cor, nos quais este trabalho está embasado.

O Capítulo 3 apresenta o conceito de Morfologia Matemática, mostrando sua importância no processamento de imagens. São mostradas suas operações básicas e como a combinação gera operadores poderosos, utilizados para os mais variados problemas. Neste capítulo também é mostrado o espectro de padrões, que é capaz de representar imagens de forma única a partir da forma.

No Capítulo 4 vamos apresentar a Análise de Componentes Principais que tem o intuito de reduzir a complexidade dos dados para fazer uma melhora triagem das imagens.

No Capítulo 5 os Métodos de Classificação são apresentados, os algoritmos têm como função, segmentar as imagens e fazer a triagem dos dados.

No Capítulo 6 será discutido o problema da Triagem Virtual, contextualizando o objetivo principal deste trabalho, assim como a proposta de solução.

O Capítulo 7 mostra os experimentos, ressaltando os métodos utilizados e, por fim, os resultados obtidos.

O Capítulo 8, para finalizar, é apresentado conclusões gerais e propostas de trabalhos futuros.

# Capítulo 2

## Recuperação de Imagens Baseada em Conteúdo

### 2.1 Introdução

O aumento do uso de imagens hoje é uma realidade, devido a uma grande facilidade de geração por meio de câmeras digitais, que podem gerar imagens de altíssima qualidade e edição de imagens com *softwares* que podem desde modificar pequenas imperfeições, tais como olhos vermelhos em imagens amadoras até edições que mudam drasticamente as características de uma imagem. Essa facilidade vem contribuindo para as pessoas gerarem grandes quantidades de imagens para fins como Medicina, Engenharia, Geologia, Botânica, para uso pessoal entre outros.

Paralelamente, desde a década de 90 houve um grande crescimento na velocidade dos computadores e capacidade de armazenamento [5]. E devido a essa grande capacidade de armazenamento e ao crescente número de imagens a necessidade de armazenar, representar e prover facilidades na busca de imagens de particular interesse na pesquisa sobre busca de imagens, mas os primeiros artigos sobre o assunto apareceram no início da década de 80 [1].

Com esse aumento na geração de imagens, grandes quantidades estavam ficando sem utilidade, já que não se conseguia recuperar essas imagens de maneira satisfatória. Como as imagens eram armazenadas em banco de dados que utilizavam informações textuais, isso nos anos 1990 com o surgimento de banco de imagens em grande escala, as informações referentes a imagens eram anotadas manualmente [6]. Outro ponto é a subjetividade das anotações que são na sua essência percepções particulares do encarregado de gravar as imagens na base. As respostas geradas por sistemas baseados em anotações podem ser não confiáveis, porque



diferentes informações podem ser descritas por pessoas diferentes sobre a mesma imagem [1][2].

Recuperação de Imagens Baseada em Conteúdo (*Content-Based Image Retrieval* – CBIR) é uma forma de buscar imagens por meio de uma imagem base usando suas características de acordo com o interesse do usuário. As características mais usadas são cor, textura e forma, mas qualquer outra informação que possa ser extraída da imagem base pode ser usada para representa a mesma. Com isso, têm-se três bases fundamentais para a recuperação de imagens baseada em conteúdo [3]:

1. Extração de características visuais;
2. Indexação multidimensional;
3. Projeto de sistema de recuperação.

A extração de características, que reflete as características gerais ou específicas de domínio, é à base dos sistemas CBIR. Sistemas que usam características gerais não dependem da aplicação e podem ser usados em qualquer tipo de sistema. Com o uso de cor e forma os sistemas obtêm características gerais. Características específicas de domínio dependem da aplicação, tais como, por exemplo, reconhecimento de faces humanas e íris para identificação de pessoas. Características podem incluir conteúdo textual ou visual. Textual refere-se às anotações ou chaves de buscas e visual referem-se a cor, textura ou forma. Neste trabalho, extração de características refere-se às características gerais e o enfoque é para a triagem de imagens médicas.

As imagens são representadas nos bancos de dados por um vetor de características multidimensional, que é usado como identificador para as imagens existentes na base de dados. Essa representação numérica caracteriza o objeto e o número de características determina a dimensão do vetor, que pode ser n-dimensional [1][7]. Para recuperar uma imagem o usuário deve fornecer uma imagem base que terá um vetor extraído a com esse vetor será feita a busca baseada na similaridade ou distância entre o vetor dado e os recuperados da base [8].

Sistemas CBIR tem várias implementações e fazem uso de diferentes tipos de consultas tais como a consulta por exemplo visual (*query by visual example* - QBVE) [9] e a recuperação semântica. A consulta por exemplo visual é baseada em uma imagem dada como exemplo para o sistema CBIR, que vai usá-la para a busca de outras imagens. Existe uma série de algoritmos, dependendo da aplicação, mas todos recuperam as imagens que tenham uma similaridade, de acordo com um determinado critério, com a imagem fornecida como exemplo. Esta técnica provê uma facilidade na busca em relação às técnicas que precisam de uma descrição textual para a busca das imagens desejadas.

A consulta por recuperação semântica seria a maneira mais intuitiva de busca pelo usuário. Nessa técnica, o usuário faz sua busca não por palavras chave, mas sim com uma referência às imagens desejadas, por exemplo, “encontre imagens de pulmão”. Essa tarefa não é fácil, pois imagens de pulmão podem parecer muito com outras imagens e dificultar a busca de imagens relevantes. A busca por imagens relevantes tem um peso muito grande nos sistemas CBIR, porque com isso o usuário perde menos tempo analisando imagens que não tem nenhuma relação com a imagem desejada.

Com as técnicas de CBIR existe um problema para se estabelecer a relação entre as características da imagem e a consulta do usuário. Tal problema está relacionado com o *gap* semântico, que é a diferença entre características de baixo nível, extraídas da imagem, e o alto nível da informação desejada pelo usuário.

Na área médica, com o desenvolvimento de aparelhos médicos que geram uma maior quantidade de imagens e de uma melhor qualidade, o uso de imagens digitais tornou-se bastante difundido. Associadas as técnicas existentes como raios-X e mamografia, técnicas como imagens de ressonância magnética (*Magnetic Resonance Imaging* – MRI) e tomografia computadorizada (*Computer Tomography* – CT) podem produzir muitas imagens a cada exame. Logo, a cada ano, um hospital pode produzir uma quantidade imensa de imagens médicas digitais [9].

Com o interesse de integrar todas as informações do paciente (texto, imagem, gráficos temporais entre outros) surgiram recentemente sistemas conhecidos com

PACS (*Picture Archiving and Communication System*) e com eles o uso de técnicas de busca de imagens por conteúdo das próprias imagens.

## 2.2 Atributos das Imagens

Nos sistemas de recuperação de imagens por conteúdo são usualmente usados cor, textura e forma como atributos para geração de vetores de características que descrevem as imagens de interesse do usuário.

### 2.2.1 Cor

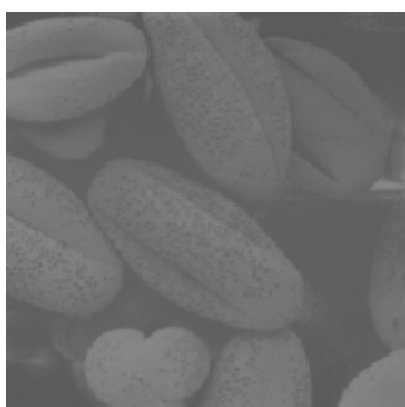
A cor é um atributo importante na análise de imagens realizada pelos seres humanos com ou sem o auxílio de computadores. A identificação de objetos e a interpretação de uma cena podem, em geral, ser simplificada com o uso de cor. Com o uso de experimentos com o sistema visual humano foi descoberto que ele é capaz de discernir milhares de tons e intensidades de cores, comparado com apenas algumas dezenas de tons de cinza. Essa capacidade de discernir milhares de cores é particularmente importante na análise manual de imagens feita por seres humanos.

Existem modelos ou espaços de cores que permitem a especificação de cores em um formato padrão para atender a diferentes dispositivos gráficos ou aplicações que requerem a manipulação de cores. O universo de cores que podem ser reproduzidos por um modelo é chamado espaço ou *gamute* de cores. Por não existirem modelos de cores que descrevam todos os aspectos referentes às cores, modelos diferentes são utilizados para especificar as características das cores. Existem modelos aditivos, em que a cor é gerada pela combinação de vários comprimentos de onda luminosa e modelos subtrativos, em que a cor é gerada pela subtração de vários comprimentos de onda. Existem muitos tipos de modelos como o RGB (Red, Green e Blue), XYZ, HSV, HLS que são modelos aditivos e CMY e CMYK, modelos subtrativos.

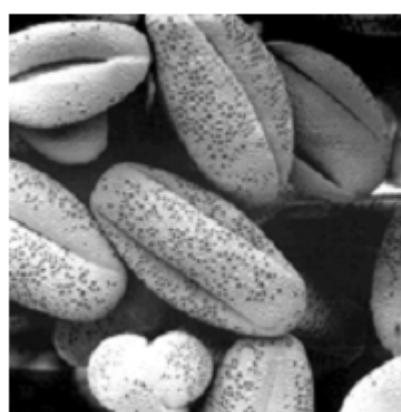
Os trabalhos relacionados com o uso de cor como característica para a busca se concentram no histograma de cor. Outros, mais recentes montam uma estrutura baseada em histogramas de dimensão variável, nomeada histogramas métricos que

são utilizados na recuperação de imagens similares [10]. Os histogramas são invariantes à translação e à rotação, e se normalizados, ficam invariantes à escala.

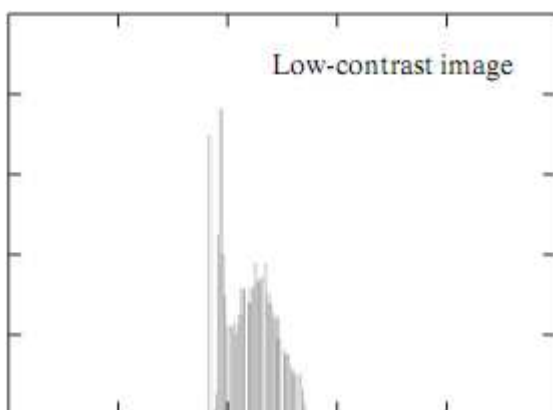
Muitas operações podem ser feitas com histograma de cor, com a simples mudança de parâmetros e limiares, por exemplo, a figura 2.3 que representa o histograma da imagem original com baixo contraste, figura 2.1. Ao se aplicar uma operação de alongamento (*stretch*) sobre a imagem original, o histograma de cor ficaria como na figura 2.4, tendo uma imagem mais bem definida na figura 2.2, de alto contraste, como resultado.



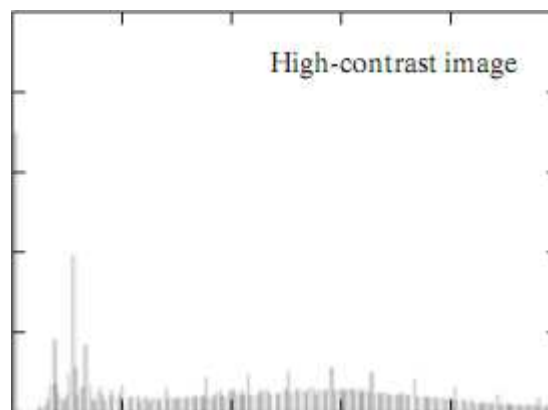
**Figura 2.1** Imagem original.



**Figura 2.2** Imagem saturada.



**Figura 2.3** Histograma de imagem original.



**Figura 2.4** Histograma de imagem saturada.

Sistemas comerciais muito populares usam características de cor com histogramas em sistemas de recuperação de imagens por conteúdo, como QBIC (*Query By Image Content*), *Virage* [11] o *VisualSeek* [12] entre outros.

### 2.2.2 Textura

A textura encontra-se entre as características usadas pelo sistema visual humano para interpretar informações visuais, contendo informações sobre a distribuição espacial e a variação luminosa, além de descrever o arranjo estrutural das superfícies e relações entre regiões.

Na literatura não existem uma definição estrita do conceito de textura de imagens, mas existe um consenso em entendê-la como as mudanças na intensidade da imagem que formam determinados padrões que se repetem [13]. Normalmente esses padrões são conseqüências de características físicas da superfície do objeto como aspereza, ou pode ser resultado da diferença de reflexão da luz sobre o objeto refletindo sua cor. Pode-se definir textura como padrões visuais complexos compostos por subpadrões que possuem características como ser lisa, rugosa, ou ondulada.

Normalmente, as medidas resultantes da aplicação dos métodos de análise de texturas são obtidas por processos de extração de características. Ela é responsável por executar transformações nos dados de entrada, de modo a descrevê-los de forma simplificada, porém representativa.

Processos automáticos para reconhecer uma textura na sua maioria precisam de técnicas computacionais complexas para levar a uma representação satisfatória de uma textura. A extração de características que usa textura como um descritor é muito útil para sistemas de CBIR, pelo fato de quase todas as superfícies naturais exibirem texturas. A análise de texturas difere da análise de cores porque texturas são padrões sobre janelas ou regiões e cores sobre pixels [10].

Conforme [10], a análise de textura tem por objetivo estabelecer o relacionamento de vizinhança dos elementos de textura e seu posicionamento, o número de elementos por unidade espacial, e a sua regularidade.

Embora não haja um consenso na literatura, de acordo com [14], modelos computacionais de características de textura podem ser divididos em abordagens estatística, baseadas em processamento de sinais, geométricas e baseadas em modelos paramétricos.

### 1. Abordagem Estatística

Em métodos estatísticos não se busca compreender explicitamente a estrutura hierárquica da textura, mas por meio de propriedades não determinísticas que definem distribuições e relacionamentos entre os níveis de cinza da imagem, tentam representar a textura indiretamente.

### 2. Abordagem Baseada em Processamento de Sinais

Os métodos de análise de texturas baseados em processamento de sinais têm a característica de extrair descritores a partir da representação obtida depois da aplicação de transformações na imagem de entrada. Um exemplo poderia ser a Transformada de *Fourier* [15].

### 3. Abordagem Geométrica

Nessa abordagem, uma textura é definida como sendo composta por primitivas. Com a identificação das primitivas que a compõem, duas classes de métodos são utilizadas para extração de características. A primeira utiliza medidas extraídas das primitivas para descrever a textura, enquanto a segunda extrai regras para descrever a disposição espacial e o relacionamento existente entre as primitivas.

### 4. Abordagem Baseada em Modelos Paramétricos

Nesta abordagem, a textura é considerada como uma amostra extraída de um processo não determinístico que é definido por um grupo de parâmetros que resume suas características. Como exemplo, temos os modelos paramétricos baseados em campos aleatórios de Markov (*Markov Random Fields* – MRF).

### 2.2.3 Forma

Assim como a cor e a textura, a forma é muito importante para se descrever um objeto. O reconhecimento de um objeto passa também pela análise de sua forma, e assim é possível identificar um objeto apenas pela sua forma.

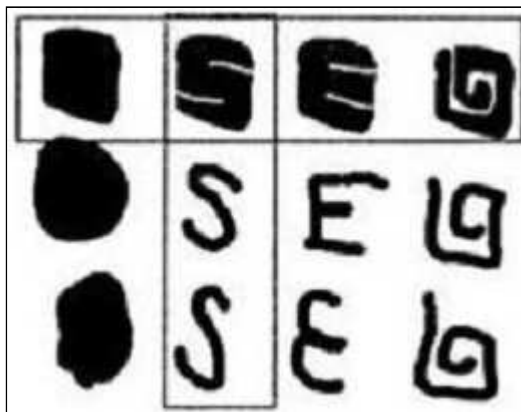
O uso da forma como descritor de características em sistemas de recuperação de imagens é uma das abordagens de maior dificuldade, especialmente pelo fato de se ter que segmentar [16] e conhecer o tamanho dos objetos contidos na imagem. Por exemplo, em aplicações médicas, a forma e tamanho dos tumores são de muita importância, na hora de auxiliar o patologista, em classificá-los como malignos ou benignos [17]. Tumores com bordas irregulares têm uma alta probabilidade de serem malignos e aqueles com bordas regulares geralmente são benignos[18].

Existem ainda, outras dificuldades inerentes as imagens adquiridas. As imagens do mundo real em sua maioria são de três dimensões (3D), porém imagens adquiridas são representadas em duas dimensões (2D), ocasionando uma, possível perda de informações na sua representação.

Vários métodos de representação de imagens, ou descritores de formas, existem na literatura, esses métodos se dividem em duas categorias: baseados em regiões e baseados em contorno. Em técnicas baseadas em regiões, todos os *pixels* dentro de uma forma são levados em consideração para obter uma representação da forma. Comumente métodos baseados em regiões usam descritores de momento para descrever a forma. Representações de regiões em momentos interpretam uma imagem, de tons de cinza com uma função densidade de probabilidade de variáveis aleatórias de duas dimensões (2D) [7]. Os métodos baseados em contorno, em comparação com os métodos baseados em regiões, são mais populares. Representações baseadas em contorno exploram apenas a informação nos limites (bordas) da forma (figura 2.5). Esses métodos de representações podem ser classificados dentro de descritores globais de forma como área, circularidade, excentricidade etc.

Considere-se a imagem da figura 2.5 como exemplo, os objetos na primeira linha possuem uma similar distribuição espacial, sendo semelhantes de acordo com critérios baseados em região. Entretanto, é fácil observar que eles possuem contornos diferentes. Se a similaridade for baseada em contorno, os objetos mostrados em

cada coluna são similares. Nesse caso, se fosse realizada uma consulta a partir de uma imagem exemplo, como a localizada na primeira linha e segunda coluna, as imagens consideradas semelhantes seriam as da primeira linha (baseada em região) ou da segunda coluna (baseada em contorno) [19].



**Figura 2.5** Similaridade baseados em contorno e região [19].

Em um projeto de descritor de imagens, devem ser analisadas características tais como: representar bem a forma; ter baixa complexidade para ser usado em cálculos e comparações; ser conciso; deve ser invariante a mudanças de orientação (rotação e translação), assim como a variações de escala. Descritores de forma devem ter o máximo dessas características para obter um bom resultado na representação de imagens.

## 2.3 Conclusão

Neste capítulo foi visto que com o aumento da quantidade de imagens necessitou-se de sistemas capazes de buscar, da melhor maneira, essas imagens. Os sistemas de buscas de imagens por conteúdo usam, em sua maioria, características como cor, forma e textura. No próximo capítulo veremos como representar essas características usando descritores morfológicos.



# Capítulo 3

## Morfologia Matemática

### 3.1 Introdução

Morfologia matemática consiste em uma metodologia para o processamento não linear de imagens que permite a construção de operadores para a descrição de objetos em imagens.

Originalmente desenvolvida para manipular imagens binárias, a morfologia matemática foi posteriormente estendida para tratar imagens em níveis de cinza. De maneira geral, o objetivo da Morfologia Matemática é descrever quantitativamente as estruturas geométricas e funcionar como uma técnica na concepção de algoritmos na área de Processamento Digital de Imagens – PDI, usando ferramentas básicas como detectores de bordas e filtros morfológicos[20]. Com isso, os operadores morfológicos podem ser utilizados em um grande número de aplicações em processamento e análise de imagens, tais como extração de componentes conexos, busca de padrões na imagem, afinamento de bordas e muitas outras.

Considerada uma técnica construtiva, a Morfologia Matemática com o uso dos operadores de erosão a dilatação pode-se construir os demais operadores, o que torna possível a construção de inúmeros outros operadores dedicados a aplicações específicas tais como filtragem e extrações de formas de interesse[21].

A teoria da Morfologia Matemática (MM) foi formulada na década de 1960, por Georges Matheron, Jean Serra e seus colaboradores da Escola Nacional Superior de Minas de Paris, França. Motivados em experimentar uma nova abordagem para resolver problemas de análise de imagens. Usando a extração de informações de imagens a partir de transformações de forma, com o uso de dois operadores que depois seriam batizados de *dilatação* e *erosão*.

## 3.2 Operadores Matemáticos

As operações fundamentais da Morfologia Matemática são a erosão e a dilatação. Esses dois operadores combinados podem dar origem a outras operações à morfologia.

### 3.2.1 Dilatação e Erosão

A dilatação, também às vezes chama da de dilação, é uma transformação morfológica que combina dois conjuntos usando adição vetorial[22]. Denotada pelo símbolo  $\oplus$  (que representa a soma de Minkowsky, no caso binário) a dilatação de um conjunto A por um conjunto B pode ser definido por:

$$A \oplus B = \{c \mid c = a + b, a \in A, b \in B\} \quad (1)$$

Onde A representa a imagem sendo operada e B é um segundo conjunto que se chama *elemento estrutural*. A composição desse elemento define a natureza específica da dilatação. A imagem original A será modificada de modo a fazer as áreas semelhantes a B se expandirem na imagem (exemplo 3.2.1).

Exemplo 3.2.1 Seja uma imagem A definida como segue[23].

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}_{(4,4)},$$

e o elemento estruturante B definido por:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}_{2,2},$$

o resultado da dilatação de A por B será a imagem  $C_1 = A \oplus B$ , dada por:

$$C_1 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}_{(4,4)} .$$

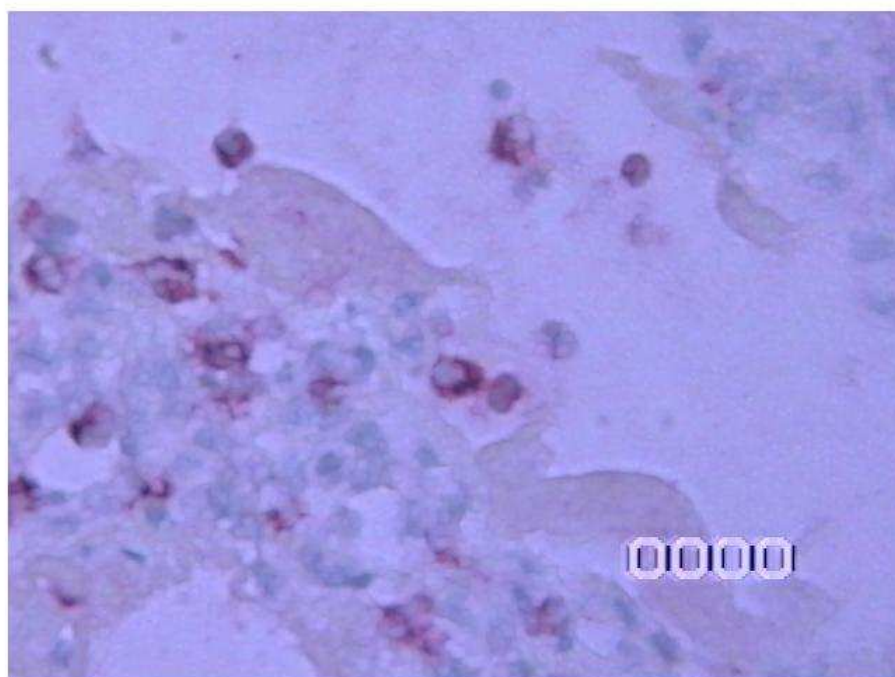
Para imagens de qualquer outro tipo podemos usar a definição de funções:

$$\delta_B(f)(u) = (f \oplus g)(u) = \bigvee_{v \in S} f(v) \wedge g(u - v), \forall u \in S \quad (2)$$

Visualmente a dilatação resulta no crescimento das áreas claras e na eliminação nas áreas escuras menores do que o elemento estruturante (figura 3.1), considerando o fato de que, usualmente, os pixels mais claros são representados por valores próximos de 1, enquanto os mais escuros, por valores mais próximos de 0. Na figura 3.1 e figura 3.2 considere um elemento estruturante quadrada 3x3 definida

por:

$$EE = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}_{3,3}$$



**Figura 3.1** Resultado da dilatação na imagem np001a.

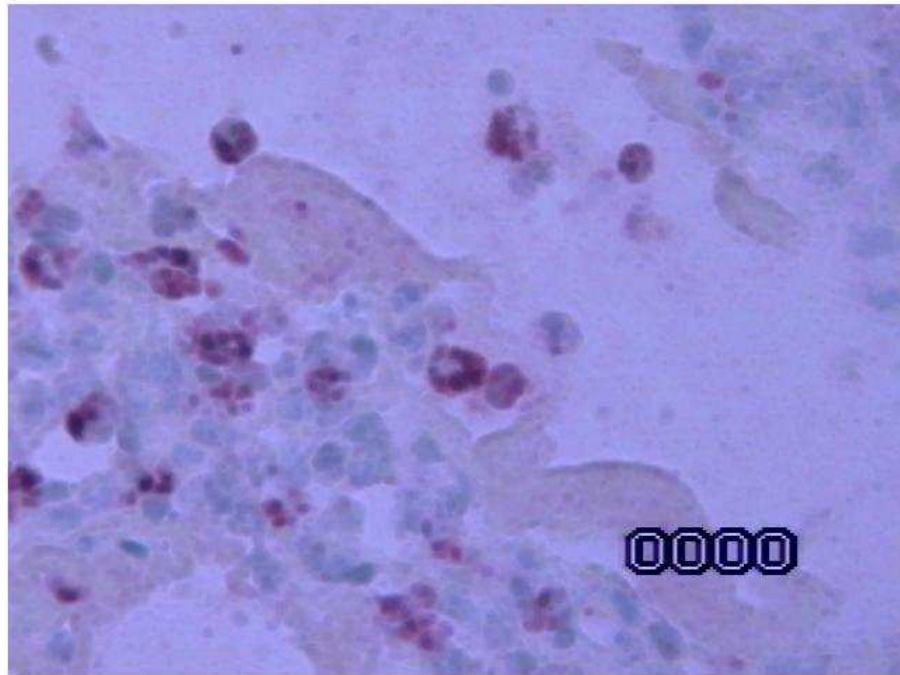
A erosão, ao contrário da dilatação basicamente “encolhe” uma imagem e pode ser vista como uma transformação morfológica que combina dois conjuntos usando vetores de subtração. Ela é a interseção de A e B e é definida por:

$$A \ominus B = \{x \mid x + b \in A, \forall b \in B\} \quad (3)$$

onde B é o elemento estruturante. Como a dilatação, a erosão também pode ser definida para qualquer outro tipo de imagem usando a definição de funções:

$$\varepsilon_g(f)(u) = (f \ominus g)(u) = \bigwedge_{v \in S} f(v) \vee \bar{g}(v - u), \forall u \in S \quad (4)$$

Na prática, a erosão transforma a imagem original  $f$  de modo a fazer as áreas semelhantes “diminuírem”. A erosão resulta no crescimento das áreas escuras e na eliminação nas áreas claras menores do que o elemento estruturante (ver figura 3.2 e exemplo 3.2.2).



**Figura 3.2** Resultado da erosão da imagem np001a.

Exemplo 3.2.2 Considerando a imagem  $f$  e o elemento estruturante  $g$  do exemplo 3.2.1, a erosão de  $f$  por  $g$  será uma imagem  $h_2 = f \ominus g$  dada por [23]:

$$h_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{(4,4)}$$

### 3.2.2 Abertura e Fechamento

Outras importantes operações da Morfologia Matemática são a abertura (*opening*) e o fechamento (*closing*). Abertura e o fechamento são filtros morfológicos, os quais são considerados filtros não lineares baseados em operações morfológicas.

Esses operadores são obtidos pela combinação dos dois operadores vistos na seção 3.2.1, dilatação e erosão.

A abertura é obtida quando se aplica a uma imagem uma erosão seguida imediatamente de uma dilatação, com o mesmo elemento estrutural. A abertura é utilizada para suavizar regiões claras na imagem que não possa estar contidas no elemento estruturante. Assim, a abertura realiza na imagem uma “peneira” nos pontos claros, mas também é usada para suavizar contornos de uma imagem, quebrar estreitos, eliminar proeminências delgadas e remover ruído [23].

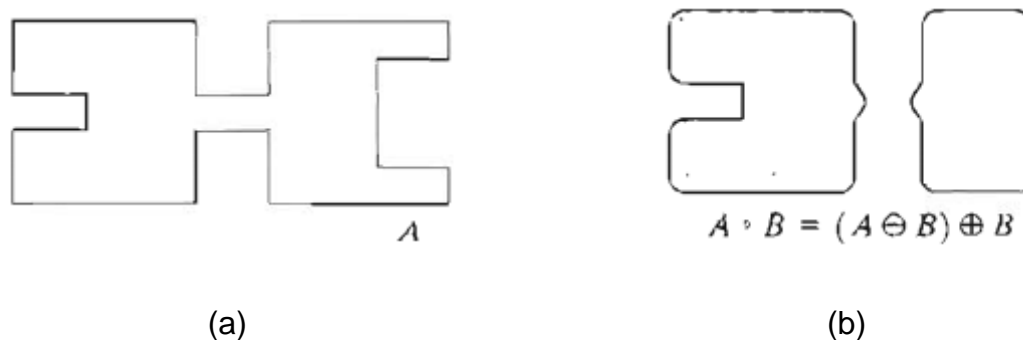
Baseando-se na teoria dos conjuntos, a abertura de um conjunto  $A$  pelo elemento estruturante  $B$  de uma imagem binária temos:

$$A \circ B = (A \ominus B) \oplus B \quad (5)$$

A abertura de uma imagem  $f$  pelo elemento estrutural  $g$  em operações baseado em funções para qualquer tipo de imagens é definida por:

$$\gamma_g(f) = \delta_g \varepsilon_g(f) \quad (6)$$

A figura 3.3 ilustra um exemplo da operação de abertura numa imagem, utilizando um elemento estruturante circular.

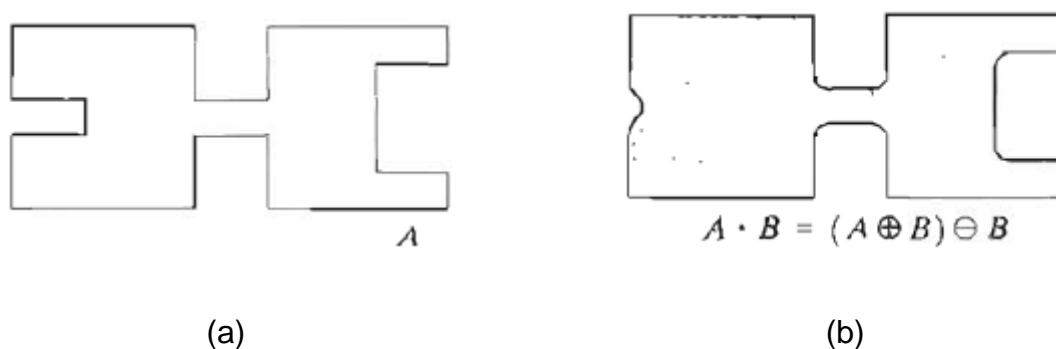


**Figura 3.3** (a) Imagem original e (b) Imagem depois de aplicada a abertura [24].

A operação de fechamento é semelhante a operação de abertura. Esses operadores se diferenciam pela ordem que as operações são realizadas, ou seja, no fechamento a dilatação é feita primeiro e logo em seguida é feita uma erosão. O fe-

chamento une pequenas quebras, alarga golfos estreitos e elimina pequenos orifícios ou vazios. Como resultado, o fechamento realiza uma espécie de eliminação de regiões escuras da imagem que não estejam contidas no elemento estruturante.

A figura 3.4 ilustra um exemplo da operação de fechamento em uma imagem, utilizando um elemento estruturante circular.



**Figura 3.4** (a) Imagem original e (b) Imagem depois de aplicado o fechamento [24].

### 3.2.3 Propriedades da Abertura e Fechamento

A abertura e o fechamento têm três propriedades importantes que devem ser levadas em consideração. As operações são consideradas duais, idempotentes e crescentes.

Entende-se que as operações são duais na condição de que uma abertura possa ser implementada usando um fechamento e vice-versa. Definida:

$$\gamma_B(f) = \overline{\phi_B(\overline{f})} \quad (7)$$

A transformação  $\xi$  é crescente se, dada duas imagens  $f$  e  $g$ , a seguinte condição é satisfeita [21]:

$$f(u) \leq g(u) \Rightarrow \xi(f)(u) \leq \xi(g)(u), \forall u \in S \quad (8)$$

Por sua vez, a operação é idempotente se [21]:

$$\xi\xi(f)(u) = \xi(f)(u), \forall u \in S, \quad (9)$$

ou seja a imagem  $f$  não é alterada por aplicações sucessivas da mesma transformação  $\xi$ , se  $\xi$  é idempotente. São as propriedades de crescimento e idempotência que fazem com que a abertura e o fechamento sejam classificados como filtros morfológicos se possuírem ao mesmo tempo as duas características [21].

Existem também duas propriedades que são específicas para cada operação: a anti-extensividade e a extensividade. Na álgebra, uma operação é dita ser extensiva se o conjunto resultante da operação for superconjunto do conjunto inicial, e é chamada de anti-extensiva no caso em que o conjunto resultante é subconjunto do conjunto inicial [25].

Nesse caso, a abertura é uma transformação anti-extensiva, pois:

$$A_B \subset A \quad (10)$$

Já o fechamento é uma transformação extensiva, ou seja:

$$A^B \supset A \quad (11)$$

A extensão das características de abertura e fechamento no caso binário para a abertura e fechamento em níveis de cinza, é imediata [26]. Dessa forma, as características da anti-extensividade da abertura (equação 12) e a extensividade do fechamento (equação 13) são definidas como:

$$\gamma_g(f)(u) \leq f(u), \forall u \in S \quad (12)$$

$$f(u) \leq \phi_g(f)(u), \forall u \in S \quad (13)$$



## 3.3 Espectro de Padrões

Espectro de padrões é uma função densidade de probabilidade que é usada na Granulometria para a descrição de tamanho e forma em análise de imagens.

### 3.3.1 Granulometria

Dentre as ferramentas da Morfologia Matemática as granulometrias constituem algumas das mais poderosas.

A idéia básica de uma granulometria é realizar diversos “peneiramentos” seguidos de medições quantitativas dos resíduos, tais como medição de área ou perímetro. O procedimento inicia realizando-se um “peneiramento” da imagem original com certa peneira e depois se medem os “grãos” que passaram; em seguida, utiliza-se uma “peneira” com furos menores e novamente se medem os resíduos, e assim por diante. No caso, o tamanho dos furos da peneira é determinado pela dimensão do elemento estruturante utilizado na transformação granulométrica.

Seja  $f$  uma imagem, a função  $V$  é definida com sendo:

$$V(k) = \sum_{u \in S} \gamma_g^k(f)(u), \quad (14)$$

onde  $g$  é uma máscara em cruz ou quadrada  $3 \times 3$ . Assim, a função  $\Xi$  é definida da forma que segue:

$$\Xi[k] = 1 - \frac{V(k)}{V(0)}, k \geq 0, \quad (15)$$

Se considerarmos a imagem  $f$  um conjunto aleatório e observamos que a função  $\Xi$  é uma função monotônica e crescente, pois  $V(k+1) < V(k), \forall k \geq 0$ , podemos afirmar que  $\Xi$  é uma função de distribuição acumulada discreta associada a imagem  $f$  [27]. Assim, podemos definir uma função densidade de probabilidade discreta  $\xi$  por meio da diferença:

$$\xi[k] = \Xi[k+1] - \Xi[k], k \geq 0 \quad (16)$$

Essa função densidade de probabilidade definida na equação (16) é denominada espectro de padrões [27]. O espectro de padrões é uma espécie de histograma de tamanho e forma, sendo único para uma imagem específica  $f$  e um determinado elemento estruturante  $g$ .

### 3.4 Conclusão

Ao longo deste capítulo foram descritos vários conceitos sobre Morfologia Matemática, sendo uma técnica de grande interesse em processamento de imagens. O aprendizado dessa abordagem tem sua base nos operadores básicos de erosão e dilatação, já que com a combinação desses dois é possível construir vários operadores específicos para as mais variadas situações como, por exemplo, a eliminação de ruído de uma imagem através da abertura e fechamento.

Também foi visto conceitos sobre espectro de padrões que fazem parte da granulometria que tem uma importância grande no presente trabalho, pela sua característica de “espalhar” informações de tamanho e forma nas imagens resíduo. Assim, muitas outras representações diferentes do histograma de padrões dos objetos de interesse podem ser geradas. Essas representações por serem únicas se adequam ao problema da triagem de imagem, podendo ser associando a cada imagem um espectro de padrões.

No próximo capítulo vamos verificar a técnica de Análise de Componentes Principais que tem um papel importante o algoritmo proposto. Essa técnica tem como motivação fazer que o sistema possa ter uma menor complexidade, reduzindo a dimensionalidade dos dados que serão usados.

# Capítulo 4

## Análise de Componentes Principais

### 4.1 Introdução

Muitos algoritmos de processamento de sinal recorrem, de uma forma ou de outra, a métodos de descorrelação de dados. As vantagens se baseiam em um princípio simples: ao descorrelacionar os dados, estamos eliminando parte da informação redundante em cada dimensão. Os benefícios são diversos em casos como processamento de imagens, que apresentam uma correlação razoável entre os “pixels” consecutivos, como uma descrição mais concisa dos dados e certas características podem vir a ter uma maior visibilidade depois de transformadas.

### 4.2 Análise de Componentes Principais

A análise de componentes principais (*Principal Components Analysis* – PCA) é um método de descorrelação de dados que tem por objetivo encontrar uma transformação mais representativa e geralmente mais compacta das observações. O método de PCA transforma um vetor aleatório  $x \in \mathbb{R}^m$  em outro vetor  $x \in \mathbb{R}^n$  para  $n < m$  projetando  $x$  nas  $n$  direções ortogonais de maior variância, que são as componentes principais. Geralmente grande parte da variância dos dados é explicada por um número reduzido de componentes, sendo possível eliminar as restantes sem grande perda de informações.

#### 4.2.1 Descorrelação de Dados

Seja  $\Sigma_x$  a matriz de covariância do vetor aleatório, real,  $x$ . Se  $\Sigma_x$  for uma matriz não singular<sup>1</sup>, então  $\Sigma_x$  pode ser decomposto no seguinte produto matricial:

$$\Sigma_x = \Gamma_x \Lambda_x \Gamma_x^T \quad (17)$$

Para

$$\Gamma_x = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & \ddots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nm} \end{bmatrix} = [\gamma_1 \quad \cdots \quad \gamma_m] \Lambda_x = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \quad (18)$$

onde  $\lambda_1, \dots, \lambda_n$  são os autovalores de  $\Sigma_x$  e  $\gamma_1, \dots, \gamma_n$  os seus autovetores correspondentes. Para matrizes simétricas os vetores autovetores são ortogonais:  $\gamma_i^T \gamma_j = 0$  para  $i \neq j$ . Como cada coluna de  $\Gamma_x$  é um autovetor,  $\Gamma_x$  é uma matriz ortogonal e tem a propriedade:  $\Gamma_x \Gamma_x^T = \Gamma_x^T \Gamma_x = I$ .

Com isso, podemos decorrelacionar o vetor  $x$  através de uma transformação linear. Para  $A = \Gamma_x^T$  temos:

$$y = Ax = \Gamma_x^T x \quad (19)$$

obtemos:

$$\Sigma_y = A \Sigma_x A^T = \Gamma_x^T \Sigma_x \Gamma_x, \quad (20)$$

usando a equação 17 temos,

$$\Sigma_x = \Gamma_x^T \Gamma_x \Lambda_x \Gamma_x^T \Gamma_x = \Lambda_x$$

---

<sup>1</sup> Matrizes não singulares são matrizes inversíveis. De notar que matrizes de covariância não singulares também são definidas positivas, que dizer  $A \in \mathbb{R}^{n \times n}$  é uma matriz definida positiva se  $x^T A x > 0$  para qualquer vetor  $x \in \mathbb{R}^n$ , real e não nulo.

Sendo assim  $\Sigma_y$  é uma matriz diagonal, implicando em uma descorrelação entre os elementos do vetor  $y$ . Normalmente os valores da matriz  $\Lambda_x$  são ordenados de maneira decrescente e sendo  $\Gamma_x$  a matriz dos autovetores correspondentes aos autovalores ordenados, a transformação da equação (19) nos dá um vetor com os primeiros valores variando mais que os últimos.

#### 4.2.2 Análise de Componentes Principais – PCA

A transformação da equação (19) projeta  $x$  em direções ortogonais de modo que o vetor transformado  $y$  tenha uma matriz de covariância diagonal. Cada dimensão  $Y_i$  do vetor  $y$  tem uma variância  $\sigma_{Y_i}^2$  igual ao autovalor  $\lambda_n$ . Essa transformação é conhecida como PCA ou *Karhunen e Loève* para o caso de imagens.

As componentes principais são simplesmente os autovetores da matriz de covariância de  $x$ . A primeira componente principal é o autovetor associado ao autovalor mais elevado, a segunda componente principal é o autovetor correspondente ao segundo autovalor mais elevado, e assim por diante.

Trabalhar com imagens, que geralmente tem dados muito correlacionados entre seus *pixels* vizinhos faz com que geralmente exista uma grande diferença entre o maior e o menor autovalor de  $\Sigma_x$ . Isto indica que certas componentes de  $y$  variam muito pouco relativamente a outras, podendo ser descartadas as dimensões de menor variância sem afetar substancialmente a qualidade dos dados.

### 4.3 Conclusão

Neste capítulo foi visto como a análise de componentes principais pode ser importante na descorrelação de dados, e assim eliminando dados irrelevantes ao processamento. A análise PCA é muito usada para compressão de imagens e por isso foi escolhida para este trabalho e com isso obtermos uma melhor relação da imagem original e sua representação.

# Capítulo 5

## Métodos de Classificação

### 5.1 Introdução

Métodos de classificação são usados para dividir ou agrupar os objetos de acordo com sua similaridade. Em sistemas CBIR os usuários que operam o sistema não estão interessados em características de baixo nível como cor, forma e textura, já que o usuário pensa em termos semânticos. O usuário procura recuperar imagens que ele já tenha analisado ou buscar imagens “parecidas” com a imagem sob análise.

Os métodos que serão descritos serão usados para melhorar essa análise inicial. Para dividir as imagens e diminuir a dimensionalidade computacional assim como reduzir o tempo de processamento das imagens. O sistema deve poder fazer o mesmo que o usuário faria. O sistema não tem a pretensão de fazer uma triagem melhor que um especialista treinado, mas o sistema deve poder auxiliar esse especialista tratando essa quantidade crescente de imagens. Fazendo assim aquelas imagens que estariam ficando sem uso poderia ajudar no diagnóstico de possíveis doenças (Capítulo 1).

Nesse capítulo serão apresentados os conceitos de algoritmos de k-médias usado para segmentar as imagens. Logo após, os conceitos sobre as redes de Kohonen conhecidas como Mapas Auto-Organizados serão explicados, assim como sua modificação SOM-LVQ que tem um aprendizado supervisionado. E por fim, as redes neurais Perceptron de Múltiplas Camadas que os resultados serão comparados com as redes SOM-LVQ.

## 5.2 K-Médias

O algoritmo de k-médias é um dos mais simples que utiliza aprendizado não-supervisionado que resolve o problema de segmentação por agrupamento [28]. Tem por objetivo encontrar a melhor divisão  $T$  dados em  $n$  grupos (*clusters*)  $C_i, i = 0 \dots n$  de modo que a distância total entre os dados de um conjunto e seu respectivo centro, somada por todos os conjuntos, seja minimizada.

O método segue passos definidos para cumprir o objetivo de agrupar os objetos em  $n$  grupos. O método consiste em usar os valores dos primeiros  $p$  casos em um arquivo de dados, como estimativas temporárias das medidas dos  $n$  conjuntos. Primeiramente, define-se  $n$  centros, uma para cada conjunto, sendo suas posições iniciais de muita influência no resultado do algoritmo. Em seguida, a região de propriedade de cada é determinada e que agrupa os objetos similares. Com as regiões determinadas, inicia-se a etapa iterativa do algoritmo, no qual o centro do conjunto é recalculado de modo a minimizar uma determinada função objetivo. Esta função é definida com o método dos mínimos quadrados:

$$J = \sum_{j=1}^n \sum_{i=1}^p \left\| x_i^{(j)} - c_j \right\|^2, \quad (21)$$

onde  $\left\| x_i^{(j)} - c_j \right\|^2$  é a medida de distância entre um ponto  $x_i^{(j)}$  até o centro de um conjunto  $c_j$  em questão. Então  $J$  é o indicados dos  $p$  objetos contidos em seus respectivos conjuntos [29]. Após realizar os cálculos, um novo centróide é escolhido pra cada conjunto escolhido, reiniciando o processo. Quando a variação de cada centro do conjunto é mínima, chega-se a uma situação em que todos os pontos já estão nos conjuntos dos seus vetores médios mais próximos.

É visto a seguir os passos para se entender melhor o k-médias:

Passo 1. Aleatoriamente são gerados 3 centróides com  $k = 3$  (figura 4.1);

Passo 2. Atribuir cada elemento ao *cluster* com o centro mais próximo (figura 4.2a);

Passo 3. Recalcular as posições dos centróides (média dos valores de todos os elementos do *cluster*) (figura 4.2b);

Passo 4. Repete-se o segundo e terceiro passo até atingir um critério de convergência<sup>2</sup> (figura 4.3);

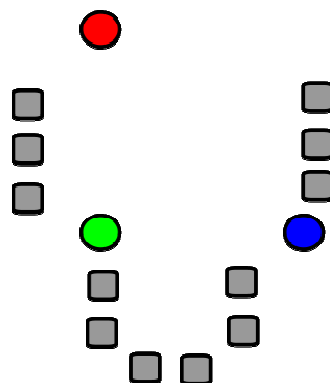


Figura 4.1 Passo 1.

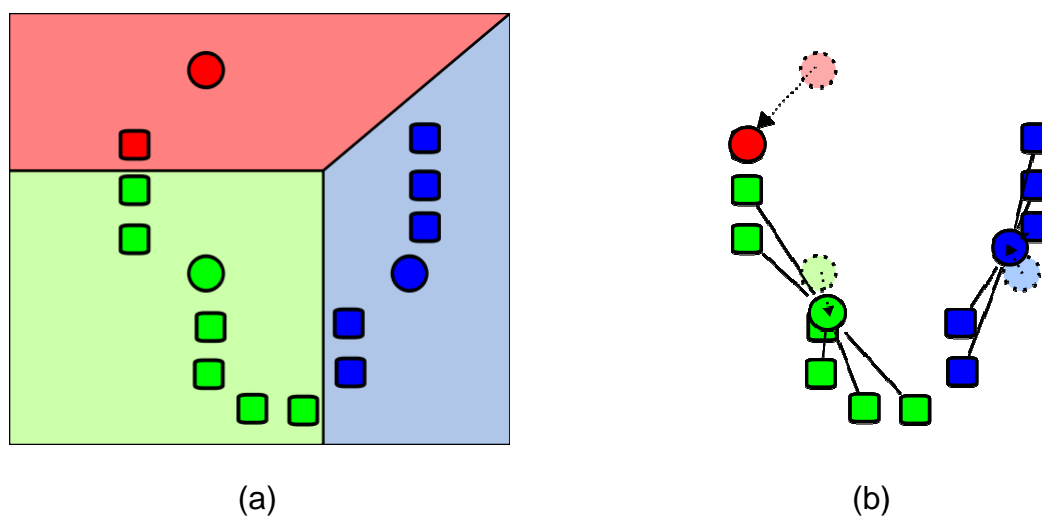
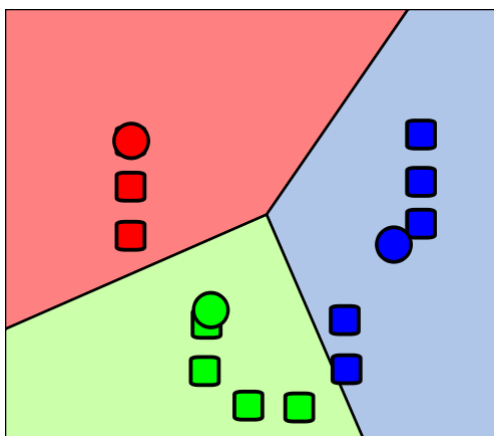


Figura 4.2 (a) Passo 2 e (b) Passo 3.

<sup>2</sup> Como critério de convergência pode ser usado um número máximo de iterações ou executar o algoritmo até que os centros não mudem de local após serem recalculados.





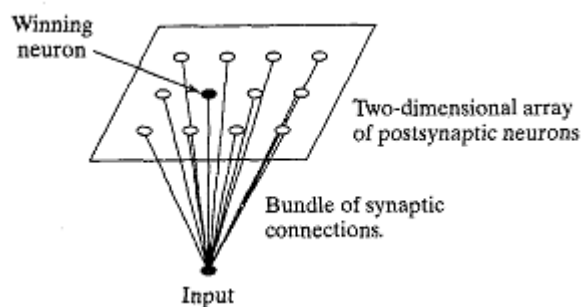
**Figura 4.3** Algoritmo converge resultando nos clusters finais.

Para conjuntos de dados com *clusters* densos, compactos e bem separados k-médias normalmente apresenta bons resultados. Entretanto, a necessidade de indicar o número de clusters  $k$  é visto como uma desvantagem em relação a outros algoritmos de agrupamentos, já que não se tem nenhuma base para definir um valor adequado, podendo assim influenciar no resultado final.

### 5.3 Mapas Auto-Organizados - LVQ

As redes SOM (*Self-Organized Maps*) são redes neurais artificiais que possuem a capacidade de auto-organização. Essa auto-organização é baseada em aprendizagem competitiva; os neurônios de saída da grade competem entre si para serem ativados ou disparados.

Em um mapa auto-organizado, os neurônios estão colocados em nós de uma grade (figura 4.4) que é normalmente uni- ou bidimensional. Os neurônios são seletivamente sincronizados a vários padrões de entrada (estímulos) no decorrer de um processo de aprendizagem. As localizações dos neurônios assim sincronizados (i.e. os neurônios vencedores) se tornam ordenadas entre si de forma que um sistema para diferentes características é criado sobre a grade. Esse mapa auto-organizável é, portanto, caracterizado pela formação de um mapa topográfico de padrões de entrada no qual as localizações dos neurônios são indicativas das características intrínsecas contidas nos padrões de entrada.



**Figura 4.4** Grade do arranjo de uma rede SOM [30].

O algoritmo responsável pela formação do mapa auto-organizado começa primeiramente inicializando os pesos sinápticos de grade. Isto pode ser feito atribuindo-lhes valores pequenos; fazendo dessa forma, nenhuma organização prévia é imposta ao mapa de características. Uma vez a grade tenha sido apropriadamente inicializada, há três processos essenciais envolvidos na formação do mapa auto-organizado:

1. **Competição.** Para cada padrão de entrada, os neurônios da grade calculam seus respectivos valores de uma função discriminante. Esta função fornece a base para a competição e o neurônio que obter o maior valor da função é declarado vencedor da competição;
2. **Cooperação.** O neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo assim a base para a cooperação entre os neurônios vizinhos.
3. **Adaptação Sináptica.** Permite que os neurônios excitados aumentem seus valores individuais da função discriminante em relação ao padrão de entrada através de ajustes em seus pesos sinápticos. Os ajustes têm a função de melhorar a resposta do neurônio se for aplicado um padrão subsequente similar ao anterior.

### 5.3.1 Quantização Vetorial por Aprendizagem

A quantização vetorial é uma técnica que explora a estrutura subjacente dos vetores de entrada para a compressão dos dados. Para o espaço que é dividido em um número de regiões distintas é definido um vetor de reconstrução que é a repre-

sentação do vetor de entrada pertencente a uma região do espaço. Com isso, utilizando uma versão codificada deste vetor de reprodução para armazenamento ou transmissão no lugar do vetor de entrada, pode-se obter uma considerável economia para armazenamento ou transmissão. Porém, existe alguma distorção nessa nova codificação.

Um quantizador vetorial com mínima distorção de codificação é chamado de quantizador de Voronoi ou por vizinho mais próximo, já que as células de Voronoi em torno de um conjunto de pontos em um espaço de entrada correspondem a uma partição daquele espaço de acordo com a regra do vizinho mais próximo.

A quantização vetorial por aprendizagem (*Learning Vector Quantization – LVQ*) é uma técnica de aprendizagem supervisionada que usa a informação sobre as classes para mover os vetores de Voronoi, a fim de melhorar a qualidade das regiões de decisão do classificador. Tomemos como exemplo um vetor de entrada  $x$  aleatoriamente no espaço de entrada. Se os rótulos de classes do vetor de entrada  $x$  e de um vetor de Voronoi  $w$  concordarem, o vetor de Voronoi  $w$  é movido em direção ao vetor de entrada  $x$ . Se, por outro lado, os rótulos de classes do vetor de entrada  $x$  e do vetor de Voronoi  $w$  discordarem, o vetor de Voronoi  $w$  é afastado do vetor de entrada  $x$ .

Considere que  $\{w_j\}_{j=1}^J$  represente o conjunto de vetores de Voronoi e que  $\{x_i\}_{i=1}^N$  represente o conjunto de vetores de entrada. Assumindo que há muito mais vetores de entrada do que vetores de Voronoi, o que é tipicamente o caso prático. O algoritmo de quantização vetorial por aprendizagem opera da seguinte forma:

- i. Suponha que o vetor de Voronoi  $w_c$  seja o mais próximo do vetor de entrada  $x_i$ . Considere que  $\xi_{w_c}$  represente a classe associada com o vetor de Voronoi  $w_c$  e  $\xi_{x_i}$  represente o rótulo de classe do vetor de entrada  $x_i$ . O vetor de Voronoi  $w_c$  é ajustado da seguinte maneira:
  - Se  $\xi_{w_c} = \xi_{x_i}$ , então:

$$w_c(n+1) = w_c(n) + \alpha [x_i - w_c(n)] \quad (22)$$

onde  $0 < \alpha_n < 1$ .

- Se, por outro lado,  $\xi_{w_c} \neq \xi_{x_i}$ , então:

$$w_c(n+1) = w_c(n) - \alpha [x_i - w_c(n)] \quad (23)$$

- ii. Os outros vetores de Voronoi não são modificados.

## 5.4 Perceptron de Múltiplas Camadas

As redes normalmente chamadas de Perceptron de Múltiplas Camadas (*Multi-layer Perceptron* - MLP) são tipicamente redes que consistem um conjunto de unidades sensoriais que constituem a *camada de entrada*, uma ou mais *camadas ocultas* e uma *camada de saída* dos nós computacionais. Os perceptrons de múltiplas camadas têm sido aplicados com sucesso para resolver diversos problemas difíceis, através de seu treinamento de forma supervisionada.

As redes MLP através de seu treinamento supervisionado têm resolvido uma grande quantidade de problemas e com isso, aplicado em diversos problemas difíceis. O algoritmo de treinamento conhecido como *algoritmo de retropropagação do erro* (*error backpropagation*) que é baseado na *regra da aprendizagem por correção de erro* foi bastante difundida e hoje seu nome se confunde, erroneamente, com a rede MLP.

Basicamente a aprendizagem por retropropagação de erro consiste em dois passos:

1. Um passo para frente, a propagação, onde um padrão de atividade é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada produzindo uma resposta as entradas.
2. Um passo para trás, a retropropagação, onde a resposta real e subtraída da saída da rede gerando um sinal de erro. Com a regra de corre-

ção de erro este sinal é usado para ajustar os pesos sinápticos propagando o sinal para trás através da rede, em direção as entradas da rede.

A combinação de algumas características faz que o perceptron de múltiplas camadas derive seu poder computacional. Primeiro, o modelo de cada neurônio inclui uma função de ativação não-linear. A presença de não-linearidades é importante porque, do contrário, a relação de entrada-saída da rede poderia ser reduzida àquela de um perceptron de camada única. A segunda é que a rede contém uma ou mais camadas de neurônios ocultos. Estes neurônios são responsáveis pela rede poder aprender tarefas complexas extraindo progressivamente as características mais significantes dos problemas. E por terceiro, a rede tem um alto grau de conectividade, onde todos os neurônios da rede são potencialmente afetados pela atividade global de todos os outros e conseqüentemente a informação contextual é tratada com naturalidade.

## 5.5 Conclusão

Nesse capítulo foram vistos os conceitos de *clustering* que foram usados de maneira particular no problema em questão. O algoritmo de k-médias usado para segmentação da imagem, as redes SOM-LVQ que usa aprendizado supervisionado adequando-se ao problema e PCA que pode reduzir sensivelmente a complexidade dos vetores de entrada.

No próximo capítulo vamos nos aprofundar no conceito de Triagem Virtual que é um dos propósitos desse trabalho.

# Capítulo 6

## Triagem Virtual

### 6.1 Introdução

Neste capítulo é caracterizado o problema fundamental deste trabalho (seção 6.2) e proposto um algoritmo para melhorar os resultados na triagem de imagens médicas (seção 6.3).

A Triagem Virtual é o passo que vai recuperar as imagens armazenadas em uma base de dados, comparando essas com uma imagem de exemplo e apresentados de forma organizada. Essa comparação pode ser feita de maneira direta, imagem a imagem ou pode ser feita usando uma representação das imagens. Essa representação deve ser fiel as características da imagem e idealmente, diminuir a complexidade do processamento no passo em que uma representação de exemplo é comparada com outra representação extraída da base de dados.

### 6.2 Caracterização do Problema

Nesse trabalho foram usadas imagens de imuno-histoquímica da placenta e do pulmão do tipo CD68, que representa o tipo de antígeno. Essas imagens são utilizadas para se poder auxiliar o patologista no seu diagnóstico fazendo comparação com outros pacientes e seus respectivos diagnósticos. Fazendo a comparação visual ou com o auxílio de algum programa o patologista pode analisar o caso atual de maneira a reduzir a probabilidade de erro.

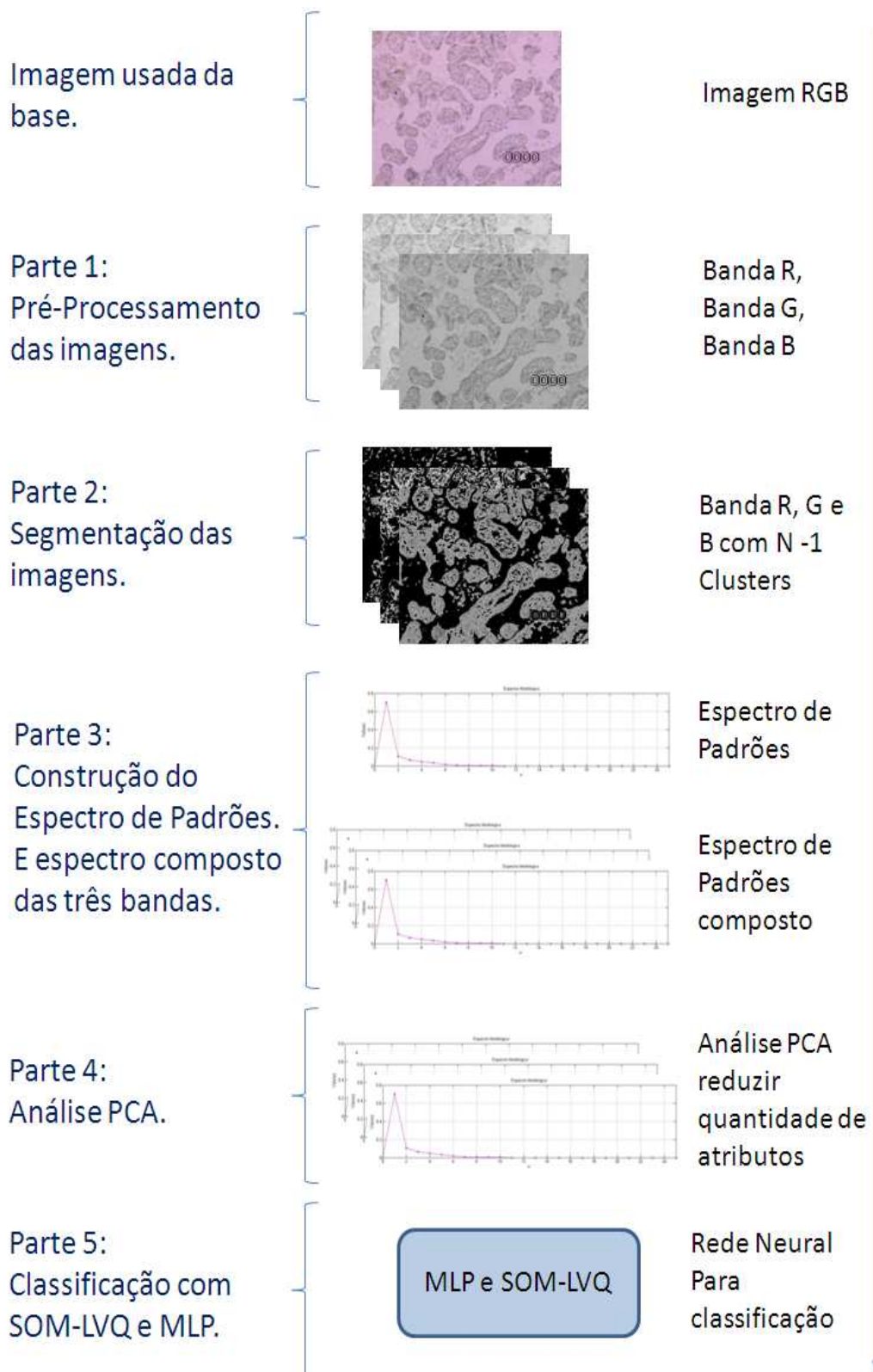
Como uma base de imagens médicas tem uma quantidade gigantesca de imagens essa busca pode ser prejudicada por fatores como:

1. As imagens foram agrupadas de maneira subjetiva: um ou mais patologistas dividiu as imagens quando foi guardá-las na base, deixando assim a subjetividade da análise particular a cada visão do problema;
2. A recuperação das imagens pode ser feita apenas em um número insuficiente: quando o usuário recuperar 10 imagens da base pode ser que essas imagens tenham pouca relevância com o problema;
3. O sistema pode demorar muito para recuperar imagens relevantes: mesmo que o sistema recupere uma quantidade grande de imagens relevantes, a quantidade de tempo gasta com essa recuperação também é importante para o usuário;

O problema atacado nesse trabalho diz respeito a melhorar a representação das imagens usando espectro de padrões, assim como reduzir a dimensionalidade do vetor e com isso melhorar a recuperação dessas imagens trazendo imagens mais relevantes e de maneira mais eficiente.

## **6.3 Proposta de Solução**

A proposta desse trabalho foi dividida em cinco partes: pré-processamento, segmentação das imagens, construção dos espectros de padrões, análise PCA e uso de SOM-LVQ e MLP para classificar as imagens (figura 6.1). Na última fase foi utilizado conjunto de treino e teste para poder avaliar quantitativamente a taxa de acertos na triagem das imagens.



**Figura 6.1** Diagrama da proposta.



### 6.3.1 Pré-processamento

O pré-processamento das imagens foi:

1. Dividir as imagens nas bandas R (*Red*), G (*Green*) e B (*Blue*);

### 6.3.2 Segmentação

A segmentação das imagens (R, G e B) foi feita com uma implementação de k-médias com  $k = 4$ , sendo  $k$  a quantidade de grupos na imagem final.

As imagens utilizadas foram analisadas e na média existiam de 3 a 6 grupos com o maior número de *pixels*. Dentre esses grupos, se usássemos 4 grupos eles seriam suficientes para representar quase totalmente a imagem. Com isso, foi escolhido 3 grupos para representar imagem e o quarto grupo representaria o *background* da imagem. No fim da segmentação o quarto grupo é retirado, ficando assim com três grupos.

### 6.3.3 Espectro de Padrões

Foi construído o espectro de padrões para as imagens (R, G e B) usando a definição descrita na seção 3.3. Foram usadas 25 iterações para compor o espectro, ficando assim, cada banda com um espectro de 25 atributos.

Depois de ser construído o espectro de padrões para cada banda esses espectros foram compostos em um vetor único, ficando no fim com uma matriz 1x75.

Após serem processadas as matrizes das duas classes de dados existentes, de forma independente, os vetores que representam as imagens de cada classe foram unificados em um vetor único. As classes das imagens usadas nesse trabalho serão definidas na seção 7.2.

### 6.3.4 Análise PCA

A análise PCA foi o passo que teve como entrada a matriz unificada gerada, que representava as imagens dos dois tipos. Essa matriz foi processada para obter a menor representação, quer dizer, com a menor quantidade de atributos, mas mantendo uma boa representação dos dados.

### 6.3.5 Classificação com Redes Neurais

Foi usada uma rede SOM-LVQ e uma rede MLP para classificar os dados usando a validação cruzada como condição de parada da rede. As duas redes foram testadas usando a matriz resultante da análise PCA e a matriz sem o uso da análise PCA. A classificação foi dividida em duas etapas como será descrito posteriormente.

Primeiro foi feita uma análise para se descobrir a melhor configuração de rede para o problema. Para isso foi estabelecido uma bateria de testes que fez os treinamentos da rede variando a quantidade de neurônios na camada escondida. Os neurônios foram variados de 10 a 30 neurônios com 20 repetições para cada rede. Em cada passo foi guardado a melhor rede, baseada na taxa de acertos e a média de acertos de todas as redes. O segundo passo foi repetir os testes com a melhor rede e obter a média, desvio padrão e mediana.

A matriz de entrada foi dividida em 19% para teste e 81% para treino e os valores foram embaralhados no início dos treinos.

Com isso, o resultado obtido, que será detalhado no próximo capítulo, foi uma comparação da abordagem usando o PCA, para reduzir a complexidade, mas com o intuito de manter uma boa representação das imagens com uma abordagem de referência usando todos os parâmetros do vetor de entrada.

## 6.4 Conclusão

Neste capítulo foi visto o problema da triagem virtual em sistemas de CBIR. Recuperar imagens irrelevantes, a demora em recuperar essas imagens são alguns problemas que um sistemas de CBR deve se preocupar. Por fim, foi descrito os passos que este trabalho se propõe para melhorar os resultados na busca de imagens por conteúdo.

No próximo capítulo serão apresentados os experimentos e resultados obtidos na bateria de testes

# Capítulo 7

## Experimentos e Resultados

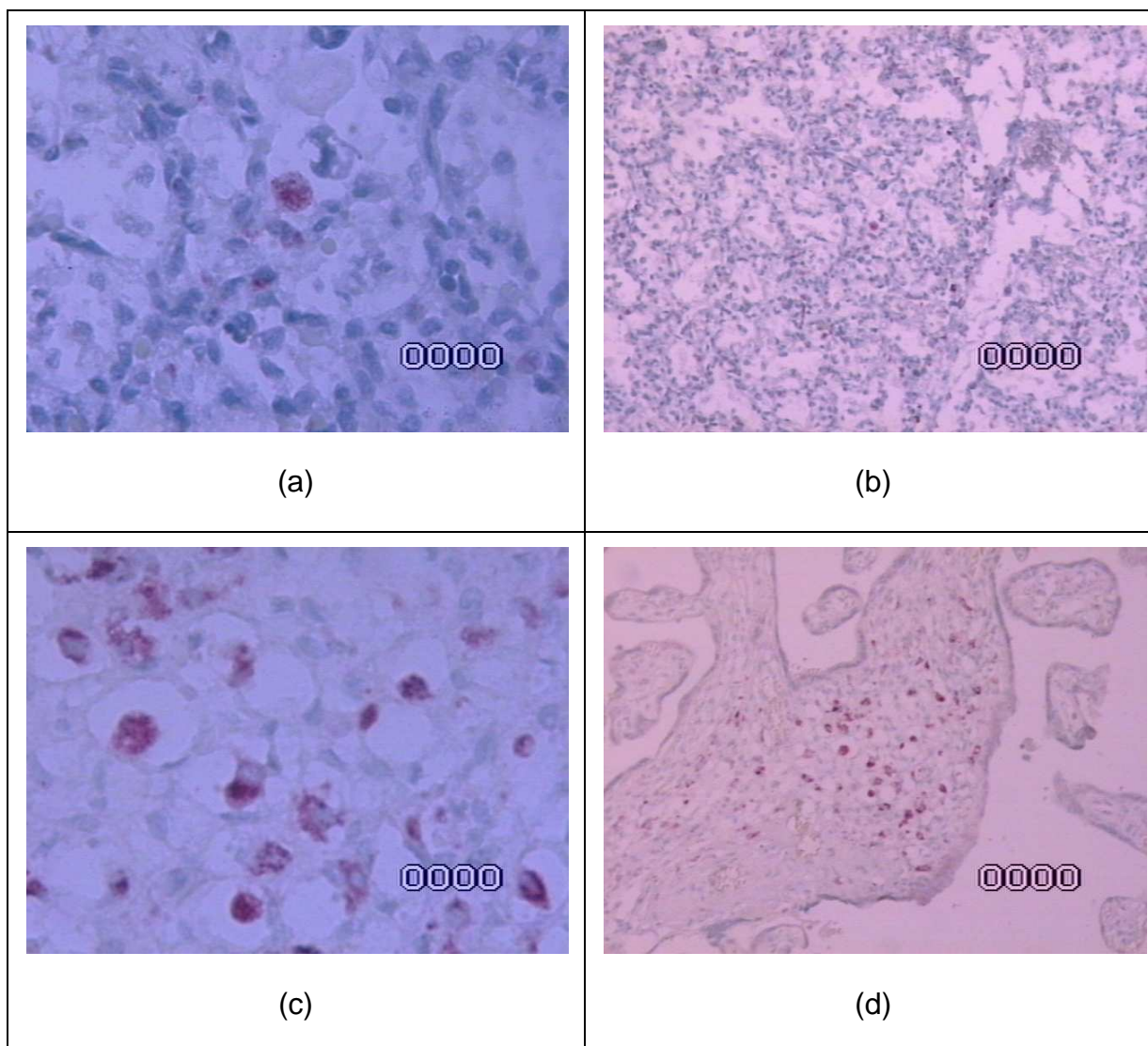
### 7.1 Introdução

Neste capítulo serão apresentados os experimentos desenvolvidos, que usam a extração de características em imagens de imuno-histoquímica usando Morfologia Matemática. De acordo com a proposta de solução anteriormente descrita, também foi reduzindo a complexidade dos vetores de características e utilizou-se redes neurais artificiais para agrupamento.

Na seção 7.2 será mostrado o conjunto de dados utilizados nos experimentos. Na seção 7.3 será visto como os vetores de características foram construídos. Na seção 7.4 os resultados dos experimentos serão mostrados e a seção 6.5 conclui o capítulo.

### 7.2 Conjunto de Dados

Neste trabalho foram usados imagens de imuno-histoquímica da placenta e pulmão, do Tipo CD68, que representa o tipo de antígeno. As imagens foram diferenciadas pela nomenclatura P e NP, onde P significa imagens de pulmão e NP imagens de placenta. As imagens usadas também têm dois tipos de zoom diferentes. Como exemplo, a figura 7.1 ilustra as imagens utilizadas.



**Figura 7.1.** Tipos de imagens: (a) np9910a, (b) np9910b, (c) p9918a e (d) p9918a.

A base de imagens tem um total de 113 imagens, sendo 53 do tipo NP e 60 do tipo P.

## 7.3 Vetores de Características

Os sistemas de CBIR usam informações das imagens para compor um vetor multidimensional que descreve as características visuais das imagens, como dito no Capítulo 2. Para cada imagem foi executado os passos da proposta explicada na seção 5.3.

1. Cada imagem era dividida em suas bandas R, G e B.

2. Cada banda era segmentada resultando uma imagem com quatro clusters e o maior era retirado.
3. Foi construído o espectro de padrões da cada banda resultando em um vetor de  $1 \times 25$ . Depois de serem unificados ficamos com um vetor de  $1 \times 75$ . As classes de imagens NP e P ficaram com matrizes  $53 \times 75$  e  $60 \times 75$  respectivamente. Por fim, foi gerada a matriz de características para representar nossa base de dimensão  $113 \times 75$  com a união dessas duas matrizes.
4. A análise PCA reduziu a complexidade da nossa base para uma matriz de  $113 \times 50$ . Tivemos assim, uma redução de 33,334% na quantidade de atributos da matriz. Ou seja uma redução de cerca de  $1/3$ .

Desta forma, foram obtidos dois vetores com cento e treze padrões. Um vetor com cinquenta atributos e outro com setenta e cinco atributos.

## 7.4 Resultado das Classificações

Os testes de agrupamento das imagens foram feitos em duas etapas, como dito na seção 6.3.5. Para implementar o algoritmo proposto foi usado o ambiente Matlab na versão R2007b como plataforma para construção, teste e treinamento do algoritmo.

Foram utilizados 22 padrões para teste e 91 padrões para treinamento (Tabela 7.1). Desses valores de teste, 11 eram do tipo NP e 11 eram do tipo P. Dos 91 padrões para treinamento 42 eram do tipo NP e 49 eram do tipo P (Tabela 7.2). Todos os treinamentos usaram validação cruzada com parte dos dados de treinamento.

**Tabela 7.1** Divisão dos dados para treino e teste.

	Quantidade	Porcentagem do Total
<b>Teste</b>	22	19%
<b>Treino</b>	91	81%

**Tabela 7.2** Divisão dos dados para treino e teste nos tipos de imagens.

	Tipo	Quantidade	Porcentagem do Total
<b>Teste</b>	NP	11	10%
	P	11	10%
<b>Treino</b>	NP	42	37%
	P	49	43%

Após a primeira etapa de classificação, para seleção da melhor rede obteve os resultados mostradas na tabela 7.3. Onde é mostrada a rede que foi utilizada, a quantidade de atributos que foram usados como entrada na rede, a taxa de acerto da melhor rede, o desvio padrão, a mediana e a quantidade de neurônios na melhor rede.

**Tabela 7.3** Resultados da primeira etapa da classificação.

Rede	Qtd. Atributos	Melhor Rede (%)	Média de Acertos (%)	Desvio Padrão	Mediana	Melhor Rede
MLP	50	95,2381	57,0522	0,4863	0,6818	22
	75	90,4762	53,8322	0,4707	0,5000	19
SOM/	50	66,6667	50,4989	0,2356	0,4091	10
LVQ	75	80,9524	47,1429	0,1565	0,5000	17

Na tabela 7.3 e 7.4 são apresentados os resultados para Perceptron de Múltiplas Camadas e SOM-LVQ considerando quantidades de atributos 50 e 75 com e sem o uso de PCA respectivamente. Para as redes MLP usando padrões obtidos com o uso de PCA foi observada uma média de acertos de 57% e um desvio padrão de 0,49, no entanto foi observada uma mediana distante da media indicando que houve resultados dispares em relação à média. Isso fica evidente quando observamos a taxa de acerto da melhor rede de 95,3%. Neste caso, a rede com 22 neurônios na camada escondida foi escolhida como a melhor rede, baseado na sua alta taxa de acertos. No caso das redes MLP sem usar padrões obtidos com PCA foi observada uma média de acertos de 54% e um desvio padrão de 0,47, porém não foi observada uma mediana tão distante da media indicando que os resultados estavam mais simétricos. Neste caso, a rede com 19 neurônios na camada escondida foi escolhida como a melhor rede, baseado na sua alta taxa de acertos de 90,5%.

Para as redes SOM-LVQ usando padrões com o uso de PCA foi observada uma redução na média de acertos das redes em comparação com as redes MLP. Com uma média de acertos de 50,5% e um desvio padrão de 0,24, como nas redes anteriores, foi observada uma mediana distante da média indicando que houve resultados dispares em relação à média. Contudo, o baixo valor do desvio padrão indica que os valores das taxas de acertos para as redes não oscilaram muito em rela-

ção à média. Assim, a rede com 10 neurônios na camada escondida foi escolhida como a melhor rede com uma taxa de acertos de 66,7%. Para as redes SOM-LVQ sem usar padrões obtidos com PCA, observado também uma redução da efetividade de acertos da rede em comparação com as redes MLP. Com uma média de acertos de 47,2% e um desvio padrão de 0,16, como nas redes anteriores, e com uma mediana distante da média indicando que houve resultados dispares em relação à média. Isso fica evidente quando observamos a taxa de acerto da melhor rede de 81%. Neste caso, a rede com 17 neurônios na camada escondida foi escolhida como a melhor rede, baseado na sua alta taxa de acertos.

Depois de selecionarmos as melhores redes, na segunda etapa foram obtidos os resultados da tabela 7.4.

**Tabela 7.4** Resultados da segunda etapa da classificação.

<b>Rede</b>	<b>Qtd. Atributos</b>	<b>Média de Acertos (%)</b>	<b>Desvio Padrão</b>	<b>Mediana</b>	<b>Melhor Rede (%)</b>
<b>MLP</b>	50	60,9524%	0,4750	0,7045	80,9524
	75	55,7143%	0,4931	0,6818	80,9524
<b>SOM / LVQ</b>	50	52,8571%	0,2185	0,3636	61,9048
	75	49,7619%	0,1581	0,4545	71,4286

Para as redes MLP usando padrões obtidos com PCA foi observada uma média de acertos de 61%, um desvio padrão de 0,5 e uma mediana de 0,7 indicando novamente uma disparidade nas taxas de acertos das redes. Ficando claro quando observado a taxa de acerto da melhor rede de 81%. No caso das redes MLP sem usar padrões obtidos com PCA obteve-se uma média de acertos de 55,7%, desvio padrão de 0,5 e mediana de 0,7. A melhor rede teve uma taxa de acerto de 81%.



Obtiveram assim, um resultado praticamente idêntico as redes com PCA, porém com uma média de acertos inferior.

Para as redes com SOM-LVQ usando padrões com PCA novamente foi observado uma redução na média de acertos das redes em comparação com as redes MLP. Observou-se uma taxa de acertos de 53%, um desvio padrão de 0,22 e uma mediana 0,36. Com um baixo desvio padrão e mediana sugere uma razoável simetria entre os dados, corroborando com uma taxa de acerto da melhor rede, não tão distante de média, em 61,9%. No caso das redes SOM-LVQ sem usar padrões obtidos com PCA obteve-se uma média de acertos de 49,8%, uma média de 0,16 e mediana de 0,46. Porém, mesmo com os valores relativamente baixos de desvio padrão e mediana, obteve uma melhor rede com uma taxa de acerto de 71,4% ficando acima da melhor rede SOM-LVQ com PCA.

## 7.5 Discussão e Conclusões

Uma vez que se busca uma triagem virtual nada se impede que se façam diversas repetições e se obtenha a rede com a melhor taxa de acertos. Resultando que as redes MLP ajudaram a gerar pelo menos uma rede com uma boa taxa de acertos que nos experimentos foram de 81%. Uma vez que estamos interessados em fazer triagem e não em fazer uma análise pura e simples nos dados. Contudo, para podermos concluir que o nosso método de triagem está acima ou abaixo, da taxa de acerto na triagem correspondente ao desempenho de uma triagem humana, é necessário uma validação dos resultados de busca com profissionais da área de saúde ou patologistas.

Os resultados obtidos indicam que o uso de PCA para representar os dados, reduzindo sua complexidade e trouxe benefícios na classificação. Nas redes MLP houve uma melhora de aproximadamente 9% da taxa de acerto com o uso de PCA. Já nas redes SOM-LVQ houve uma melhora de 6% em relação à taxa de acertos com o uso de PCA.

As redes MLP se mostraram mais adequadas ao problema com os resultados mais consistentes e sua alta taxa de acerto na melhora redes de 81%. Porém, as

redes SOM-LVQ obtiveram resultados menos efetivos nas suas classificações. As redes obtiveram uma melhor rede com uma taxa de 71,4% sem usar PCA superior a melhor rede com o uso de PCA. Indicando que nesse caso as redes SOM-LVQ necessita de mais informações para poder ter uma melhor generalização e por conseguinte se adequar ao problema em questão.

## Capítulo 8

# Conclusão e Trabalhos Futuros

Os trabalhos em recuperação de imagens por conteúdo têm evoluído de maneira significativa. Mesmo assim, é uma área que necessita de maiores estudos em virtude dos grandes desafios a serem atingidos e problemas encontrados em outros métodos de recuperação de imagens.

Este trabalho apresentou uma nova proposta para recuperação de imagens baseada em conteúdo, ao utilizar cor e forma para representar as imagens, junto com uma análise para reduzir a complexidade da busca em imagens segmentadas buscando apenas áreas de interesse. A extração de características das imagens se deu através do espectro de padrões, uma técnica de Morfologia Matemática. A redução da complexidade foi obtida utilizando-se Componentes Principais e usando  $k$ -médias se segmentou as imagens. Por fim, a comparação dos resultados da triagem usando Multilayer Perceptron e Mapas Auto-Organizados mostrou que as redes MLP se adequaram melhor ao problema com uma taxa maior e mais consistente nas suas classificações.

Uma das dificuldades encontradas foi a grande quantidade de tempo de processamento no algoritmo para se obter o espectros de padrões das imagens. Já que era calculado o espectro para cada banda de uma imagem e isso era repetido para toda a base. Depois de gerado os vetores o treinamento para se obter as melhores redes exigiam um tempo de processamento também custoso.

Enquanto se construiu essa proposta de triagem virtual, outras idéias foram levantadas. Com o uso de PCA para reduzir a quantidade de atributos, poderiam ser gerado vetores de dimensões maiores, a fim de obter uma melhor precisão mesmo após a redução da quantidade de atributos.

Diferente de como foi aplicado neste trabalho poderia ser testada uma forma diferente de uso de PCA. Em compressão de imagens, a imagem é dividida em blocos de  $m \times m$  pixels, e considerar cada um desses blocos como amostra do vetor

aleatório  $x$ , temos que  $x \in \mathbb{R}^{m \times m}$ . Então, devemos usar essa dimensão, como a dimensão total da imagem e para reduzir a complexidade deveríamos representar a imagem com um fator  $n$ , com  $n$  indicando a quantidade nova de dimensões para ser usada no processamento. Com isso, teríamos a aplicação de PCA ainda na imagem e posteriormente representaríamos essa imagem em um vetor de características com o espectro de padrões.

# Bibliografia

- [1] CASTAÑÓN, C. A. B. Recuperação de imagens por conteúdo através de análise multiresolução por Wavelets, São Paulo, 2003.
- [2] FRANCISCANI, J. F. Redução do Gap Semântico em CBIR utilizando Ontologia de Objeto., Uberlândia – MG, n. Universidade de Uberlândia – UFU.
- [3] YONG, R.; H., T. S.; CHANG, S. F. Image Retrieval: Current Techniques. **The NECI Scientific Literature Digital Library**, 1999.
- [4] FENG, D.; LONG, F.; ZHANG, H. Fundamentals of Content Based Image Retrieval. **Springer-Verlag**, Berlin, 2003.
- [5] GAGAUDAKIS, G.; ROSIN, P. L. Incorporating Shape into Histograms for CBIR, 2002.
- [6] ERPEN, L. R. C. Recuperação de Informação Visual, Porto Alegre.
- [7] EMMANUEL, M. et al. Content-Based Medical Image Retrieval. **IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007)**, Dr. M.G.R. University, Chennai, Tamil Nadu, India., 2007. 712-717.
- [8] FENG, D.; SIU, W. C.; ZHANG, H. J. **Multimedia Information Retrieval and Management Technological Fundamentals and Applications**. Hardcover: Springer, v. XVII, 2003.
- [9] MÜLLER, H. E. A. Evaluation axes for medical image retrieval systems: the image CLEF experience. **ACM Press, The 13th annual ACM international conference on Multimedia**, Cingapura, p. 1014-1022, 2005.
- [10] BUENO, J. M. **Suporte à Recuperação de Imagens Médicas Baseada em Conteúdo Através de Histogramas Métricos**. São Paulo: Tese de doutorado, Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo., 2002.

- [11] AUTONOMY Virage. Disponível em: <<http://www.virage.com>>. Acesso em: 01 nov. 2009.
- [12] VISUALSEEK. **DVMM Lab.** Disponível em: <<http://www.ee.columbia.edu/In/dvmm/researchProjects/MultimediaIndexing/VisualSEEK/VisualSEEK.htm>>. Acesso em: 01 nov. 2009.
- [13] TUCERYAN, M.; JAIN, A. Texture Analysis in Handbook of Pattern Recognition and Computer Vision. **World Scientific Publishing Company**, 1993. 235–276.
- [14] PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais - Princípios, Algoritmos e Aplicações**. São Paulo: Thomson Learning, 2008.
- [15] COGGINS, J.; JAIN, A. A Spatial Filtering Approach to Texture Analysis. **Pattern Recognition Letter**, v. 3, p. 195-203, 1985.
- [16] IAKOVIDIS, D. K. et al. A Pattern Similarity Scheme for Medical Image Retrieval. **IEEE Xplore**, Athens, v. 13, n. 4, p. 442-450, July 2009.
- [17] LIMA, S. M. L. Sistema de Auxílio ao Diagnóstico de Doenças Intracranianas e Câncer de Mama. **Trabalho de Conclusão de Curso**, Recife, 2009.
- [18] THEODORIDIS, S.; KOUTHUMBAS, K. **Pattern Recognition**. 1. ed. [S.l.]: Elsevier Science & Technology Books, 1999.
- [19] BOBER, M. MPEG-7 Visual Shape Descriptors. **IEEE Transactions on Circuits and Systems for Video Technology**, Guildford, v. 11, p. 716-719, jun. 2001.
- [20] STATELLA, T.; SILVA, E. A. Extração de drenagens em imagens digitais através de técnicas de Morfologia Matemática, Cuiabá, 2009. 359-376.
- [21] CANDEIAS, A. L. B. Aplicação da Morfologia Matemática à Análise de Imagens de. São José dos Campos: Departamento de Computação Aplicada do Instituto, 1997.
- [22] WANGENHEIM, A.; DE SOUZA, A. I.; SANTOS, C. A. Morfologia Matemática. **Universidade Federal de Santa Catarina - UFSC**, Santa Catarina. 49-80.

- [23] SANTOS, W. P. Análise de Imagens Digitais em Patologia utilizando Morfologia Matemática e Lógica Nebulosa. **Dissertação de Mestrado em Engenharia Elétrica, Universidade Federal de Pernambuco**, Recife - PE, 2003.
- [24] GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 2nd. ed. New Jersey: Prentice Hall, 2001.
- [25] SERRA, J. **Image Analysis and Mathematical Morphology: Theoretical Advances**. Ecole des Mines de Paris: Academic Press Inc., v. Vol 2, 1988.
- [26] DOUGHERTY, E. R.; LOTUFO, R. A. Hands-on Morphological Image Processingods in Imaging. **SPIE Press**, 2003.
- [27] BRAGA, U. M. N. Reconstrução volumétrica e análise de imagens tridimensionais por morfologia matemática. **Master's thesis, Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas**, Campinas, 1994.
- [28] JAIN, A.; DUBES, R. Algorithms for Clustering Data. **Prentice-Hall**, p. 155-166, 1988.
- [29] MOORE, A. K-means and Hierarchical Clustering - Tutorial Slides. Disponível em: <<http://www-2.cs.cmu.edu/~awm/tutorials/kmeans.html>>. Acesso em: 01 nov. 2009.
- [30] HAYKIN, S. **Redes Neurais**. 2nd. ed. Porto Alegre: Bookman, 2001.