

Universidade do Minho
Escola de Engenharia

Pedro João Soares Rodrigues

**Técnicas de Segmentação e de Classificação
em Imagens. Estudo de um Caso de Aplicação**

Tese de Doutoramento
Área de Informática Industrial/ Ramo de Electrónica Industrial

Trabalho efectuado sob a orientação de
Professor Doutor João Luís Monteiro
Professor Doutor Manuel João Ferreira

Dezembro de 2007

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Agradecimentos

Ao Professor Doutor João Luís Monteiro pela sua orientação, pelo acompanhamento crítico e por todo o apoio dispensado ao longo deste trabalho.

Ao Professor Doutor Manuel João Oliveira Ferreira pela sua orientação, pelas sugestões, pelos esclarecimentos e pela ajuda na preparação de todo o trabalho.

A ambos agradeço também a liberdade científica concedida.

Aos meus colegas do Instituto Politécnico de Bragança que em várias direcções científicas ajudaram no esclarecimento de algumas temáticas abrangidas por este trabalho.

Ao Instituto Politécnico de Bragança e ao programa *PRODEP* que permitiram a dedicação exclusiva a este trabalho.

Ao laboratório de análises clínicas de Maria Montanha e Laura Soares que facultou, desinteressadamente, todos os elementos necessários à obtenção das imagens dos leucócitos.

À minha irmã pelos constantes encorajamentos e pela sua ímpar revisão linguística que executou ao texto desta tese.

Aos meus pais por todo o apoio incessante.

A todos os meus amigos que de múltiplas formas ajudaram a que este trabalho pudesse ser concluído.

**À minha irmã
Aos meus pais**

Resumo

Os processos de segmentação de imagens existentes continuam a estar intensamente relacionados com o problema para o qual foram pensados. Um sistema que produz segmentações de elevado sucesso num determinado problema pode falhar por completo se aplicado, sem modificações, a um outro problema. Mecanismos de segmentação que se aproximem, em capacidade generalista, ao sistema visual humano têm sido procurados pela comunidade científica. Nesta tese são estudadas técnicas e apresentados alguns contributos orientados para a melhoria da capacidade generalista e do desempenho dos sistemas de segmentação.

Os processos de classificação de imagens podem aparecer como parte funcional aos mecanismos de segmentação ou então como utilizadores dos resultados da segmentação. Em qualquer das formas o problema da classificação, no caso do processamento de imagens, relaciona-se, também, com a capacidade generalista do método usado para a classificação. Os classificadores aqui estudados são estabelecidos, de forma automática, a partir de um conjunto de amostras de treino. A eficácia generalista desse treino está subjugada não só à qualidade de treino, mas também intrinsecamente à natureza dos classificadores. Estes tendem a diminuir o seu desempenho à medida que a dimensão estrutural do problema aumenta. Os contributos aqui correntes visam minimizar os problemas do aumento da dimensionalidade e simultaneamente obter um bom grau de generalização na classificação de imagens.

O desempenho temporal destes mecanismos está, normalmente, relacionado de forma exponencial com o aumento do bom comportamento funcional. São consideradas algumas vertentes que visam minimizar este problema.

O classificador estudado e comparado nesta tese é o SVM (Support Vector Machine). É proposto aqui como mecanismo adjuvante da segmentação e da classificação de imagens.

São estudados os momentos de Zernike, descritores numéricos de regiões, que possibilitam a redução estrutural dos problemas de segmentação e de classificação.

Relativamente à segmentação são mostradas as impossibilidades de se conseguir uma parametrização generalista. Deste facto, é descrito um processo que, partindo de uma base de segmentação multi-escala, procura, através de parametrização exaustiva e recorrendo à quântica computacional, atingir as melhores segmentações e por consequência as melhores classificações.

Outros processos propostos nesta tese, que também buscam a generalização na segmentação de imagens e na classificação de imagens, são apresentados por conciliações entre estruturas de redes neuronais não supervisionadas (redes de acoplamento por pulsos –Pulse Coupled Neural Networks–) e estruturas de classificadores supervisionados (Support Vector Machine). Ainda relativamente às redes de acoplamento por pulsos, são apresentadas adições à sua estrutura com o objectivo de estas melhorarem o seu comportamento na codificação de contornos.

Por que todos estes processos, apesar de produzirem algumas melhorias funcionais, demandam uma grande capacidade computacional, são também indicadas duas soluções de aceleração computacional. Estas propostas são de índole quântica. Uma tem vertente quântica menos dirigida e baseia-se nos fenómenos de difracção de partículas em orifícios circulares e na sua respectiva interferência e a outra é de natureza dirigida e integra-se na disciplina da computação quântica. É ainda demonstrada a possibilidade de a difracção de partículas em orifícios ser descrita por equações de funções de onda quântica.

O caso de aplicação estudado que serviu para avaliar, com imagens naturais, a maior parte dos processos, assenta na segmentação e na classificação visual de leucócitos humanos.

Palavras-chave: redes neuronais; segmentação de imagens; classificação de imagens; computação quântica; leucócitos.

Abstract

Image segmentation processes, nowadays, continue to be closely related to the problem for which they had been designed. A system that produces a highly accurate segmentation in one problem can fail if applied without modifications to another problem. Segmentation methods that were able to reach the generalization capacity of the human visual system have been pursued for the scientific community. In this thesis some contributions aimed at the improvement of the generalization capacity in segmentation systems are presented.

Image classification processes can appear as a functional part in the segmentation mechanisms or can use the results of the segmentation to achieve their goals. In both cases, the classification problem in image processing becomes related also with the generalization capacity of the used classification method. The classifying processes studied here are generated, in an automatic form, from a set of training data. The generalist effectiveness of these trainings is a function not only of the quality of trainings, but also of the classifying nature. These tend to diminish their performance in relation to the growth of the structural dimension of the problem. The current contributions aim to minimize the curse of dimensionality problems and simultaneously to get a good generalization level in image classification.

These systems show a temporal exponential growth with the proximity of optimum results. Some approaches towards reducing this problem are suggested.

The classifier studied in this thesis is the SVM (Support Vector Machine). It is proposed as a component in the segmentation and in process of classification of images.

The Zernike moments as numerical region descriptors are studied, since they enable structural dimension reduction in segmentation and classification.

With regard to segmentation, the impracticality of achieving a generalist parameterization is shown. Then, a process is described that, having its origin in

multiscale segmentation, attempts to achieve the better segmentation and consequently the better classification, using an exhaustive parameterization and using quantum computation.

There are other proposed methods in this thesis that attempt to arrive at generalization in the segmentation and in the classification processes. They are formed by the conciliation among unsupervised neural networks (Pulse Coupled Neural Networks) and supervised classifiers (Support Vector Machine). With regard to the Pulse Coupled Neural Networks, components to increase the functional performance of the edge codification are added to its structure.

All these processes bring some functional improvement, but the necessary computational effort to achieve it is huge. Thus, in this thesis two solutions to speed up some image processing methods are presented. One solution is constructed by using the particle diffraction phenomenon in circular apertures and its interferences. The other one uses quantum computation. In the quantum diffraction field, the possibility of describing the particle diffraction in circular apertures using wave function equations is proved.

In this thesis the segmentation and the classification of leukocytes is the case study that was used to verify the behavior of the majority of these methods in natural images.

Key-words: neural networks; image segmentation; image classification; quantum computation; leukocytes.

Índice Geral

Capítulo 1

| | |
|---|----------|
| Introdução | 1 |
| 1.1 A problemática da segmentação e da classificação..... | 1 |
| 1.2 Abrangência da tese..... | 4 |
| 1.3 Trabalho anterior..... | 6 |
| 1.4 Caso de estudo..... | 7 |
| 1.5 Características do suporte físico..... | 9 |
| 1.5.1 Hardware utilizado..... | 10 |
| 1.6 Controlo do microscópio..... | 11 |
| 1.7 Software..... | 11 |
| 1.8 Trabalho de suporte de outros autores..... | 12 |
| 1.9 Contribuições..... | 13 |
| 1.10 Organização da tese..... | 14 |
| 1.11 Notação quântica..... | 15 |

Capítulo 2

| | |
|--|-----------|
| Support Vector Machine | 17 |
| 2.1 Introdução..... | 17 |
| 2.2 Teoria da Aprendizagem Estatística..... | 18 |
| 2.3 Risco..... | 20 |
| 2.4 Dimensão de Vapnik-Chervonenkis (VC)..... | 22 |
| 2.5 Minimização do Risco Estrutural (SRM)..... | 23 |
| 2.6 Classificação por vectores de suporte..... | 24 |
| 2.6.1 Optimização do hiperplano separador..... | 25 |
| 2.6.2 Hiperplano separador óptimo generalizado para a não separabilidade linear..... | 33 |
| 2.6.3 Generalização num espaço de dimensão elevada..... | 36 |
| 2.7 O algoritmo Kernel-Adatron..... | 38 |
| 2.8 Espaço de características..... | 40 |

| | | |
|-------|---|----|
| 2.8.1 | Funções de <i>kernel</i> | 40 |
| 2.8.2 | Escolha de um <i>Kernel</i> | 43 |
| 2.9 | Classificação de múltiplas classes..... | 44 |
| 2.10 | Considerações finais..... | 44 |

Capítulo 3

| | | |
|---|--|----|
| Descritores e Seleccionadores de Características | 47 | |
| 3.1 | Introdução | 47 |
| 3.2 | Descritores de regiões..... | 48 |
| 3.2.1 | Momentos cartesianos não ortogonais | 48 |
| 3.2.2 | Momentos de Hu..... | 49 |
| 3.2.3 | Momentos de Zernike ortogonais complexos..... | 50 |
| 3.3 | Comparação entre momentos de Zernike e de Hu sobre imagens naturais | 53 |

Capítulo 4

| | | |
|--|--|----|
| Comportamento de Métodos de Segmentação não Supervisionados | 57 | |
| 4.1 | Introdução | 57 |
| 4.2 | Métodos de agrupamento..... | 58 |
| 4.2.1 | Algoritmos de agregação clássicos | 58 |
| 4.2.2 | Agrupamento pelo método ISODATA..... | 60 |
| 4.2.3 | Métodos de agregação por histograma..... | 61 |
| 4.2.4 | Crescimento de Região (Region Growing) | 62 |
| 4.3 | Segmentação Multi-escala | 63 |
| 4.4 | Resultados | 69 |
| 4.5 | Discussões finais | 78 |

Capítulo 5

| | | |
|---|---|----|
| Segmentação e Classificação com Redes Neurais de Acoplamento por Pulsos .. | 79 | |
| 5.1 | Introdução | 79 |
| 5.2 | O modelo neuronal PCNN..... | 80 |
| 5.3 | O kernel da PCNN..... | 90 |
| 5.4 | O ruído nas PCNNs | 91 |
| 5.4.1 | Redução do efeito do ruído por oscilação do valor limiar..... | 91 |
| 5.4.2 | Redução do efeito do ruído por propagação rápida do sinal de <i>Linking</i> | 92 |

| | | |
|--------|--|-----|
| 5.4.3 | Redução do efeito do ruído por controlo do decaimento do valor limiar | 93 |
| 5.5 | Parametrização das PCNNs | 94 |
| 5.6 | Séries temporais das PCNNs..... | 95 |
| 5.7 | As PCNNs no processamento de imagens..... | 96 |
| 5.8 | Segmentação produzida pela PCNN em imagens naturais de células sanguíneas..... | 98 |
| 5.9 | Método de segmentação com PCNNs e SVMs..... | 101 |
| 5.9.1 | Resultados obtidos com o segmentador híbrido | 107 |
| 5.9.2 | Um modelo para comparação na segmentação..... | 108 |
| 5.9.3 | Imagens resultado da segmentação híbrida | 109 |
| 5.9.4 | Conclusões sobre o segmentador híbrido..... | 112 |
| 5.10 | Classificação de leucócitos através de um modelo híbrido | 113 |
| 5.11 | Resultados do classificador híbrido..... | 119 |
| 5.11.1 | Um modelo para comparação na classificação | 120 |
| 5.11.2 | Resultados em tabelas de coincidência (confusion tables)..... | 121 |
| 5.12 | Alteração dos campos receptores das PCNN para classificação de contornos..... | 123 |
| 5.12.1 | O problema do <i>aliasing</i> | 126 |
| 5.13 | Resultados da aplicação das PCNN à codificação e classificação de contornos..... | 127 |
| 5.13.1 | Séries temporais resultantes da aplicação das PCNN aos contornos ... | 130 |
| 5.13.2 | Desempenho do método de classificação de contornos do híbrido PCNN-SVM..... | 132 |
| 5.13.3 | Desempenho do método de classificação de contornos do híbrido PCNN-SVM em imagens naturais | 135 |
| 5.14 | Modificação dos Kernels da PCNN com imagens naturais..... | 138 |
| 5.15 | Conclusões | 139 |

Capítulo 6

| | | |
|---|-------------------------------|-----|
| Aceleração Quântica em Processamento de Imagem | 141 | |
| 6.1 | Introdução | 141 |
| 6.2 | Notação de Dirac | 142 |
| 6.3 | Bits quânticos (Qubits) | 144 |

| | | |
|--------|--|-----|
| 6.4 | Esfera de Bloch | 144 |
| 6.5 | Enredo de qubits..... | 146 |
| 6.6 | Operação de observação | 147 |
| 6.7 | Portas quânticas..... | 148 |
| 6.7.1 | Portas quânticas de um único qubit..... | 149 |
| 6.7.2 | A transformação de Walsh-Hadamard | 150 |
| 6.7.3 | Portas de dois e de três qubits..... | 150 |
| 6.8 | Computadores quânticos..... | 151 |
| 6.9 | Paralelismo quântico..... | 153 |
| 6.10 | Algoritmo de Grover | 155 |
| 6.10.1 | Modificação do sinal da amplitude | 158 |
| 6.10.2 | Inversão sobre a média | 158 |
| 6.11 | Adaptação do algoritmo de Grover ao processamento de imagem..... | 159 |
| 6.11.1 | Função de comparação | 162 |
| 6.11.2 | Teste num simulador quântico..... | 164 |
| 6.11.3 | Função quântica para a segmentação multi-escala | 167 |
| 6.12 | Sistema híbrido..... | 168 |
| 6.13 | Considerações finais | 170 |

Capítulo 7

| | | |
|---|---|-----|
| Classificação de Imagens por Interferência Quântica..... | 171 | |
| 7.1 | Introdução | 171 |
| 7.2 | Interferência Quântica | 173 |
| 7.3 | A natureza da probabilidade quântica | 178 |
| 7.4 | A dinâmica de partículas..... | 179 |
| 7.5 | Funções de onda para a difração em orifícios circulares | 180 |
| 7.6 | Interferência quântica aplicada à imagem | 185 |
| 7.6.1 | Pixels numa matriz de orifícios..... | 187 |
| 7.6.2 | Método de interferência..... | 188 |
| 7.6.3 | Comportamento relativo às variações de translação, de rotação e de escala | 195 |
| 7.7 | Resultados..... | 196 |
| 7.7.1 | Aplicação da interferência ao triângulo de Kanizsa..... | 215 |
| 7.8 | Discussões finais | 216 |

| | |
|--|------------|
| Capítulo 8 | |
| Conclusões..... | 219 |
| 8.1 Trabalho futuro..... | 222 |
| Anexos | 223 |
| Anexo A | |
| A Equação de Schrödinger | 225 |
| A.1 A equação de Schrödinger para uma dimensão | 225 |
| A.2 A equação de Schrödinger para duas dimensões | 228 |
| Anexo B | |
| Teoremas de Gödel - Não Computabilidade | 233 |
| Anexo C | |
| Consciência Quântica | 235 |
| Anexo D | |
| Controlador de Microscópio..... | 239 |
| Referências Bibliográficas | 241 |

Índice de Figuras

| | | |
|-------------|--|----|
| Figura 1.1 | Imagens de leucócitos obtidas num microscópio óptico com uma ampliação de 40×10 | 9 |
| Figura 1.2 | Interligação dos principais componentes constituintes do sistema de aquisição de imagens | 10 |
| Figura 2.1 | Erros de modelação perspectivados nos diferentes espaços de soluções..... | 20 |
| Figura 2.2 | Exemplo de uma dimensão VC de valor 3 em \mathfrak{R}^2 . Existem 2^3 combinações de correspondência entre os 3 pontos a classificar e as 2 classes (assinaladas correspondentemente pelas regiões a sombreado e a branco) | 23 |
| Figura 2.3 | Representação do encadeamento de espaços de acordo com as suas dimensões..... | 23 |
| Figura 2.4 | Várias possibilidades de hiperplanos separadores das duas classes, sendo o mais denso o hiperplano que optimiza a separação, ou seja, que eleva a capacidade de generalização | 25 |
| Figura 2.5 | Hiperplanos canónicos..... | 26 |
| Figura 2.6 | Estabelecimento da margem ρ através dos vectores de suporte (SV) no espaço dividido pelo hiperplano (w,b) | 27 |
| Figura 2.7 | Restrição adicional aos hiperplanos canónicos | 29 |
| Figura 2.8 | Hiperplano separador (w,b) num espaço linearmente não separável . | 33 |
| Figura 2.9 | Mapeamento do espaço de entrada num espaço de maiores dimensões..... | 36 |
| Figura 2.10 | Conversão, por mapeamento de uma função f , de um problema não separável linearmente num problema separável linearmente | 36 |
| Figura 3.1 | Sequência de blocos funcionais usados na comparação dos momentos de Hu com os momentos de Zernike | 54 |
| Figura 4.1 | Pilha de imagens a diferentes escalas da imagem original | 67 |
| Figura 4.2 | Hierarquia de ligações entre os vários níveis de escala..... | 68 |
| Figura 4.3 | Área de procura de um pixel pai | 68 |
| Figura 4.4 | Imagem original de quatro células | 69 |

| | | |
|-------------|--|----|
| Figura 4.5 | Componente de intensidade (L) da imagem original de quatro células | 69 |
| Figura 4.6 | Componente de saturação (S) da imagem original de quatro células. | 70 |
| Figura 4.7 | Componente de cor (H) da imagem original de quatro células..... | 70 |
| Figura 4.8 | Processamento sobre a componente de cor (H), por agregação baseada em histograma para doze classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea..... | 73 |
| Figura 4.9 | Processamento sobre a componente de cor (H), por agregação baseada em histograma para seis classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea..... | 73 |
| Figura 4.10 | Processamento combinado dos três componentes (HSL), por agregação baseada em histograma para doze classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea..... | 73 |
| Figura 4.11 | Resultado usando o mesmo processamento feito à imagem anterior, mas com a segunda célula mais à esquerda, dessa imagem, e seus adjacentes, isolados..... | 74 |
| Figura 4.12 | Resultado usando o mesmo processamento feito à imagem da Figura 4.10, mas com a célula mais à direita, dessa imagem, e seus adjacentes, isolados | 74 |
| Figura 4.13 | Processamento combinado dos três componentes (HSL), por agregação baseada em histograma para seis classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea..... | 74 |
| Figura 4.14 | Resultado usando o mesmo processamento feito à imagem anterior, mas com a segunda célula mais à esquerda, dessa imagem, e seus adjacentes, isolados..... | 75 |
| Figura 4.15 | Resultado usando o mesmo processamento feito à imagem da Figura 4.13, mas com a célula mais à direita, dessa imagem, e seus adjacentes, isolados | 75 |
| Figura 4.16 | Segmentação produzida pelo método de crescimento de região, sobre a componente de cor da imagem da Figura 4.4, com o valor | |

| | | |
|-------------|---|----|
| | do limiar de continuidade de crescimento igual a 1 e que compara com o valor da Equação (4.8) | 75 |
| Figura 4.17 | Segmentação produzida pelo método de crescimento de região, sobre a componente de cor, da segunda célula à esquerda, da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a um e que compara com o valor da Equação (4.8)..... | 76 |
| Figura 4.18 | Segmentação produzida pelo método de crescimento de região, sobre a componente de cor da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a 2 e que compara com o valor da Equação (4.8) | 76 |
| Figura 4.19 | Segmentação produzida pelo método de crescimento de região, sobre a componente de cor, da segunda célula à esquerda, da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a 2 e que compara com o valor da Equação (4.8) . | 76 |
| Figura 4.20 | Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor da imagem da Figura 4.4. A parametrização usada empregou o valor 5 para o raio do círculo de procura e o valor 9 para o valor do nível de escala mais elevado | 77 |
| Figura 4.21 | Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor, da segunda célula mais à esquerda, isolada, da imagem da Figura 4.4. A parametrização usada no processo de segmentação foi a mesma que originou o resultado visível na imagem anterior..... | 77 |
| Figura 4.22 | Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor da imagem da Figura 4.4. A parametrização usada empregou o valor 2 para o raio do círculo de procura e o valor 17 para o valor do nível de escala mais elevado | 77 |
| Figura 4.23 | Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor, da segunda célula mais à esquerda, isolada, da imagem da Figura 4.4. A parametrização usada no processo de segmentação foi a mesma que originou o resultado visível na imagem anterior..... | 78 |
| Figura 5.1– | Representação do modelo de um neurónio PCNN | 83 |

| | | |
|-------------|--|-----|
| Figura 5.2 | Estrutura genérica de uma rede PCNN..... | 84 |
| Figura 5.3 | Sinais nos vários constituintes de um neurónio PCNN ao longo das iterações | 85 |
| Figura 5.4 | Evolução de uma <i>auto-onda</i> através de uma imagem. A imagem (a) corresponde à imagem original e as restantes, de (b) a (h), correspondem a imagens obtidas entre a iteração 29 e a iteração 52. <i>i</i> representa o número da iteração | 87 |
| Figura 5.5 | Processamento de uma PCNN sobre uma imagem de características granulares. A imagem original a processar é a (a). A imagem (b) foi obtida numa determinada iteração da PCNN anterior à iteração resultante na imagem (c) | 88 |
| Figura 5.6 | Processamento da PCNN a contornos de descontinuidades geométricas horizontais. A imagem original é a (a). O resultado do processamento à imagem (a) é visível na imagem (b) onde estão assinaladas descontinuidades horizontais da imagem original | 89 |
| Figura 5.7 | A mesma figura representada numa escala e rotação distinta. Relativamente à figura na imagem (a), a figura da imagem (b) representa uma escala inferior e a figura da imagem (c) representa (b) rodada 45 graus | 96 |
| Figura 5.8 | Três séries temporais resultantes da operação de uma rede PCNN nas imagens da Figura 5.7. No eixo vertical são contabilizados os números de neurónios activos por iteração. No eixo horizontal estão assinalados os índices de múltiplas iterações..... | 96 |
| Figura 5.9 | Diagrama de um processo de correlação no espaço de frequências... 98 | |
| Figura 5.10 | Imagem em níveis de cinzento de células a segmentar, o glóbulo branco encontra-se ao centro, circundado por glóbulos vermelhos ... 99 | |
| Figura 5.11 | Alguns resultados das saídas das redes PCNN ao longo de cem iterações do processamento à imagem da Figura 5.10. <i>i</i> representa o número da iteração..... | 99 |
| Figura 5.12 | Alguns resultados das saídas das redes PCNN ao longo de cem iterações do processamento à imagem da Figura 5.10. <i>i</i> representa o número da iteração..... | 100 |
| Figura 5.13 | Composição do método de segmentação por PCNN/SVM | 106 |

| | | |
|-------------|--|-----|
| Figura 5.14 | Modelo de segmentação neuronal exclusivamente supervisionado através de uma rede <i>feed-forward</i> com um algoritmo de aprendizagem backpropagation..... | 109 |
| Figura 5.15 | Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)..... | 111 |
| Figura 5.16 | Comparação entre os resultados de segmentação de um linfócito obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)..... | 111 |
| Figura 5.17 | Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)..... | 111 |
| Figura 5.18 | Comparação entre os resultados de segmentação de um monócito obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)..... | 111 |
| Figura 5.19 | Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)..... | 112 |
| Figura 5.20 | Composição do método de classificação por PCNN/SVM..... | 116 |
| Figura 5.21 | Modelo de classificação neuronal exclusivamente supervisionado através de uma rede <i>feed-forward</i> com um algoritmo de aprendizagem backpropagation..... | 121 |
| Figura 5.22 | Modelo PCNN para codificação de contornos (representação sumária)..... | 124 |
| Figura 5.23 | Triângulo rectângulo a ser processado pela PCNN de contornos | 127 |
| Figura 5.24 | Algumas amostras das saídas da PCNN ao longo de 50 iterações resultantes do processamento à imagem da Figura 5.23. O i representa o número da iteração PCNN..... | 128 |
| Figura 5.25 | Algumas amostras das saídas da PCNN ao longo de 50 iterações resultantes do processamento à imagem da Figura 5.23. O i representa o número da iteração PCNN..... | 129 |
| Figura 5.26 | Séries temporais resultantes da aplicação da PCNN aos contornos de um triângulo, de um quadrado e de um rectângulo, todos à mesma escala..... | 130 |

| | | |
|-------------|--|-----|
| Figura 5.27 | Séries temporais resultantes da aplicação da PCNN aos contornos de dois rectângulos com escalas diferentes..... | 131 |
| Figura 5.28 | Séries temporais resultantes da aplicação da PCNN aos contornos de dois rectângulos com posições angulares diferentes..... | 132 |
| Figura 5.29 | Células com contrastes diferentes | 135 |
| Figura 5.30 | Contornos esquelizados de células com contrastes diferentes | 135 |
| Figura 5.31 | Séries temporais resultantes da aplicação da PCNN aos contornos de três células | 136 |
| Figura 6.1 | Esfera de Bloch | 145 |
| Figura 6.2 | Circuito somador quântico | 152 |
| Figura 6.3 | Porta quântica genérica..... | 153 |
| Figura 6.4 | Operação de conjunção com uma porta de Toffoli | 154 |
| Figura 6.5 | Função quântica para detecção e classificação de figuras em imagens | 160 |
| Figura 6.6 | Circuito quântico para implementação do Oracle quântico, que detecta a posição de segmentos de recta numa dada recta..... | 166 |
| Figura 6.7 | Sistema de processamento quântico híbrido | 169 |
| Figura 7.1 | Experiência das “Duas Ranhuras” com projecteis clássicos..... | 174 |
| Figura 7.2 | Experiência das “Duas Ranhuras” com ondas | 175 |
| Figura 7.3 | Experiência das “Duas Ranhuras” com partículas quânticas..... | 178 |
| Figura 7.4 | Representação do plano de orifícios formantes da figura de interesse. As difracções resultantes das passagens das partículas pelos orifícios são posteriormente detectadas na matriz de pixels da Figura 7.5. Na Figura 7.6 são apresentadas as funções de onda resultantes da passagem das partículas pelos orifícios <i>a</i> , <i>b</i> e <i>c</i> | 192 |
| Figura 7.5 | Representação do plano detector contendo elementos distribuídos uniformemente. Cada um destes elementos observará o produto da interferência entre as várias incidências de funções de onda..... | 193 |
| Figura 7.6 | Representação, em sobreposição, do método de interferência na codificação de um rectângulo. A exemplificação da interferência é feita sobre três dos elementos do rectângulo. Cada <i>pixel</i> do plano detector está sujeito ao valor resultante da interferência entre as múltiplas funções de onda. Neste exemplo, as probabilidades das funções de onda são combinadas de acordo com a expressão: | |

| | | |
|-------------|--|-----|
| | $\frac{1}{N}[\psi(a) + \psi(b) + \psi(c) + \dots + \psi(\cdot)]^2$, em que N representa o número de orifícios abertos | 194 |
| Figura 7.7 | Padrão de interferência (imagem direita) resultante de uma circunferência (imagem esquerda) de níveis binários | 197 |
| Figura 7.8 | Histograma do padrão de interferência da imagem direita da Figura 7.7 | 197 |
| Figura 7.9 | Padrão de interferência resultante de um rectângulo de níveis binários, a uma rotação de 135 graus e a uma escala de índice 10 de escalas entre 1 e 10 | 198 |
| Figura 7.10 | Histograma do padrão de interferência da imagem direita da Figura 7.9 | 198 |
| Figura 7.11 | Padrão de interferência resultante de um rectângulo de níveis binários, a uma rotação de 150 graus e a uma escala de índice 7 de escalas entre 1 e 10 | 198 |
| Figura 7.12 | Histograma do padrão de interferência da imagem direita da Figura 7.11 | 199 |
| Figura 7.13 | Padrão de interferência resultante de um rectângulo de níveis binários, a uma rotação de 70 graus, a uma escala de 7 de escalas entre 1 e 10..... | 199 |
| Figura 7.14 | Histograma do padrão de interferência da imagem direita da Figura 7.13 | 199 |
| Figura 7.15 | Padrão de interferência resultante de um quadrado de níveis binários..... | 200 |
| Figura 7.16 | Histograma do padrão de interferência da imagem direita da Figura 7.15 | 200 |
| Figura 7.17 | Padrão de interferência resultante de um quadrado de cantos arredondados de níveis binários | 200 |
| Figura 7.18 | Histograma do padrão de interferência da imagem direita da Figura 7.17 | 201 |
| Figura 7.19 | Padrão de interferência resultante de uma elipse de níveis binários | 201 |
| Figura 7.20 | Histograma do padrão de interferência da imagem direita da Figura 7.19 | 201 |

| | | |
|-------------|---|-----|
| Figura 7.21 | Padrão de interferência resultante de um triângulo rectângulo de níveis binários | 202 |
| Figura 7.22 | Histograma do padrão de interferência da imagem direita da Figura 7.21 | 202 |
| Figura 7.23 | Padrão de interferência resultante de um triângulo equilátero de níveis binários | 202 |
| Figura 7.24 | Histograma do padrão de interferência da imagem direita da Figura 7.23 | 203 |
| Figura 7.25 | Padrão de interferência resultante dos contornos da letra “T” de níveis binários, e o seu respectivo histograma | 203 |
| Figura 7.26 | Histograma do padrão de interferência da imagem direita da Figura 7.25 | 203 |
| Figura 7.27 | Padrão de interferência resultante dos contornos de uma figura irregular de níveis binários..... | 204 |
| Figura 7.28 | Histograma do padrão de interferência da imagem direita da Figura 7.27 | 204 |
| Figura 7.29 | Padrão de interferência resultante de uma região circular de níveis binários..... | 204 |
| Figura 7.30 | Histograma do padrão de interferência da imagem direita da Figura 7.29 | 205 |
| Figura 7.31 | Padrão de interferência resultante de uma região rectangular de níveis binários | 205 |
| Figura 7.32 | Histograma do padrão de interferência da imagem direita da Figura 7.31 | 205 |
| Figura 7.33 | Padrão de interferência resultante de uma região formada por um triângulo rectângulo | 206 |
| Figura 7.34 | Histograma do padrão de interferência da imagem direita da Figura 7.33 | 206 |
| Figura 7.35 | Padrão de interferência resultante de uma região formada por um triângulo equilátero de níveis binários..... | 206 |
| Figura 7.36 | Histograma do padrão de interferência da imagem direita da Figura 7.35 | 207 |
| Figura 7.37 | Padrão de interferência resultante de uma região que configura a letra “T” de níveis binários | 207 |

| | | |
|-------------|--|-----|
| Figura 7.38 | Histograma do padrão de interferência da imagem direita da Figura 7.37 | 207 |
| Figura 7.39 | Padrão de interferência resultante de uma região de um eosinófilo de níveis de cinzento..... | 208 |
| Figura 7.40 | Histograma do padrão de interferência da imagem direita da Figura 7.39 | 208 |
| Figura 7.41 | Padrão de interferência resultante de uma região de um eosinófilo de níveis de cinzento distintos dos níveis do eosinófilo da Figura 7.39 | 208 |
| Figura 7.42 | Histograma do padrão de interferência da imagem direita da Figura 7.41 | 209 |
| Figura 7.43 | Padrão de interferência resultante de uma região de um neutrófilo de níveis de cinzento..... | 209 |
| Figura 7.44 | Histograma do padrão de interferência da imagem direita da Figura 7.43 | 209 |
| Figura 7.45 | Padrão de interferência resultante de um contorno de um eosinófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central | 210 |
| Figura 7.46 | Histograma do padrão de interferência da imagem direita da Figura 7.45 | 210 |
| Figura 7.47 | Padrão de interferência resultante de um contorno de um eosinófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central | 210 |
| Figura 7.48 | Histograma do padrão de interferência da imagem direita da Figura 7.47 | 211 |
| Figura 7.49 | Padrão de interferência resultante de um contorno de um neutrófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central | 211 |

| | | |
|-------------|--|-----|
| Figura 7.50 | Histograma do padrão de interferência da imagem direita da Figura 7.49 | 211 |
| Figura 7.51 | Blocos sequenciais do processo de análise ao descritor de interferência..... | 213 |
| Figura 7.52 | Padrão de interferência do Triângulo de Kanizsa | 216 |
| Figura D.1 | Janela da aplicação de controlo do microscópio | 239 |
| Figura D.2 | Controlador desenvolvido para a motorização do microscópio..... | 240 |

Índice de Tabelas

| | | |
|---------------|--|-----|
| Tabela 1.I | Percentagens de leucócitos existentes no sangue de uma pessoa sã | 8 |
| Tabela 1.II | Notação quântica | 15 |
| Tabela 3.I | Percentagens de acertos conseguidos para os descritores de Hu e de Zernike na classificação de leucócitos na presença de três níveis de relação sinal-ruído..... | 54 |
| Tabela 5.I | Configuração da PCNN para a segmentação não supervisionada.... | 100 |
| Tabela 5.II | Configuração da PCNN para a segmentação híbrida | 107 |
| Tabela 5.III | Caracterização dos componentes usados na segmentação híbrida... | 108 |
| Tabela 5.IV | Configuração da PCNN para a classificação híbrida..... | 119 |
| Tabela 5.V | Caracterização dos componentes usados na classificação híbrida ... | 120 |
| Tabela 5.VI | Tabela de coincidência de classificação de leucócitos usando um modelo estritamente supervisionado | 122 |
| Tabela 5.VII | Tabela de coincidências de classificação de leucócitos usando um modelo híbrido | 122 |
| Tabela 5.VIII | Comparação de taxas de acertos para cada classe de célula entre o método supervisionado e o método híbrido | 122 |
| Tabela 5.IX | <i>Kernels</i> a aplicar aos estímulos externos para a codificação de contornos por PCNN..... | 125 |
| Tabela 5.X | Configuração da PCNN para a codificação de contornos..... | 127 |
| Tabela 5.XI | Tabela de coincidências de classificação de figuras geométricas usando o <i>código da cadeia</i> modificado | 133 |
| Tabela 5.XII | Tabela de coincidências de classificação de figuras geométricas usando uma PCNN original e três SVM..... | 133 |
| Tabela 5.XIII | Tabela de coincidências de classificação de figuras geométricas usando uma PCNN modificada e três SVM..... | 134 |
| Tabela 5.XIV | Comparação de taxas de acertos para cada classe de figura entre os três métodos testados | 134 |
| Tabela 5.XV | Tabela de coincidências de classificação de células sanguíneas usando o <i>código da cadeia</i> modificado..... | 137 |
| Tabela 5.XVI | Tabela de coincidências de classificação de células sanguíneas usando PCNNs orientadas ao contorno e SVMs | 137 |

| | | |
|---------------|--|-----|
| Tabela 5.XVII | Comparação de taxas de acertos para cada classe de célula entre os dois métodos testados | 137 |
| Tabela 6.I | Transformações quânticas de um <i>qubit</i> | 149 |
| Tabela 6.II | Porta Cnot e porta CCnot..... | 151 |
| Tabela 6.III | Correspondência entre os vários segmentos de estado quântico de $ x\rangle$ e a parametrização da figura a procurar | 160 |
| Tabela 6.IV | Número de <i>qubits</i> e estados para $ i\rangle$, $ y\rangle$ e $ w\rangle$ da função quântica P_f | 161 |
| Tabela 6.V | Tabela verdade entre os bits, x_0, x_1 , de parametrização de segmentos de recta e os bits que estabelecem o correspondente segmento | 165 |
| Tabela 7.I | Parametrização do sistema de interferência..... | 197 |
| Tabela 7.II | Desempenho do descritor de interferência, relativamente a contornos sintéticos, face ao descritor de Zernike e em cooperação com uma SVM | 213 |
| Tabela 7.III | Desempenho do descritor de interferência, relativamente a regiões sintéticas, face ao descritor de Zernike e em cooperação com uma SVM..... | 214 |
| Tabela 7.IV | Desempenho do descritor de interferência, relativamente a regiões de leucócitos, face ao descritor de Zernike e em cooperação com uma SVM..... | 214 |
| Tabela 7.V | Desempenho do descritor de interferência, relativamente a contornos de leucócitos, face ao descritor de Zernike e em cooperação com uma SVM..... | 215 |

Siglas

| | |
|------|-----------------------------------|
| BPNN | BackPropagation Neural Network |
| CCD | Charged-Coupled Device |
| DLL | Dynamic Link Library |
| ERM | Empirical Risk Minimization |
| PC | Personal Computer |
| PCNN | Pulse Coupled Neural Network |
| RBF | Radial Base Function |
| ROC | Receiver Operating Characteristic |
| SRM | Structural Risk Minimization |
| SV | Support Vector |
| SVC | Support Vector Classification |
| SVM | Support Vector Machine |
| SVR | Support Vector Regression |
| VC | Dimensão de Vapnik-Chervonenkis |

Capítulo 1

Introdução

Neste primeiro capítulo são evidenciados alguns temas de contextualização deste trabalho: é discutida a problemática residente nos processos, em foco nesta tese, de segmentação e de classificação de imagens. É apresentado o perfil de abrangência científica deste trabalho. É caracterizado o caso de estudo e são referidos os ambientes de suporte utilizados no procedimento operante do caso de estudo. São realçadas as principais contribuições deste trabalho e é exposta a estrutura geral da organização da tese.

1.1 A problemática da segmentação e da classificação

Em processos de visão, sejam eles biológicos ou artificiais, é possível identificar duas partes fundamentais: a segmentação e a classificação. Normalmente entende-se a segmentação como o assinalar de uma região de interesse e a classificação como a interpretação da região de interesse. Porém, olhando para o funcionamento de um sistema de visão e pretendendo atingir soluções sobre imagens naturais, é sempre difícil dissociar uma da outra. De facto, normalmente, é difusa a fronteira entre o processo de segmentação e o processo de classificação. Em alguns casos pode-se constatar que um não existe sem o outro ou até mesmo pode-se referir, semanticamente, que um é o outro. Concretamente o problema põe-se da seguinte forma: para se segmentar uma região é necessário reconhecer aspectos dessa região, logo, é necessário classificar. Por outro lado, para se classificar uma região é necessário deter informação sobre, pelo menos, algumas partes segmentadas dessa região, logo, é necessário segmentar. Biologicamente

estas ideias têm sido subscritas ao se observarem realimentações de informação nas estruturas em hierarquia neuronal [1;2]. Isto leva a crer que a acção das estruturas menos dirigidas de segmentação, nos primeiros níveis do córtex visual, são por supervisão reajustadas por mecanismos de classificação das camadas mais elevadas do córtex visual, que, por sua vez, interagem com outras zonas do cérebro, como é o caso do hipocampo, onde são estabelecidos alguns tipos de memórias. As camadas mais elevadas do córtex visual estão vocacionadas para um objectivo cuja estrutura poderá ter sido retida numa fase de aprendizagem, pelo menos parcialmente. Isto reforça a ideia de cooperação mútua entre a componente essencialmente de segmentação e a componente essencialmente de classificação.

Em visão por computador, se o problema em causa for de segmentação simples, por exemplo, uma única figura geométrica representada por pixels bem contrastados e totalmente conexos, é exigida uma simples operação de binarização. O passo seguinte, neste caso, é o de classificar a figura mediante as características extraídas do conjunto de pixels que a formam, não existindo uma cooperação forte e mútua entre os dois processos.

Se a segmentação em questão tiver que ser feita sobre imagens naturais complexas, as decisões de como fazer as agregações dos diversos pixels, para obter a segmentação, podem passar por vários tipos e escalas de características. Enquanto algumas características podem ser analisadas como factores típicos de similaridade ou factores típicos de desigualdade para produzir agregação, outras, não sendo típicas, têm que ser consideradas só em alguns contextos. Isto significa que existem características que devem ser tomadas sempre como diferenciadoras para realizar, ou não, a decisão e outras que não podem ser sempre usadas com esse fim e que estão dependentes de um factor efémero contextual. Nesta última situação a decisão fica dependente de uma aprendizagem de um classificador supervisionado. A situação de características típicas pode ser tratada por métodos não supervisionados que intrinsecamente são construídos com base nas assinaturas dessas características. Apesar de tudo, podem existir situações em que as características que eram típicas deixam de o ser. Esse factor contextual pode ser ele próprio difícil de interpretar.

Os problemas relacionados com os classificadores supervisionados estão muitas vezes aliados com a insuficiência, que é natural, dos padrões usados na fase de treino. Por exemplo, como fazer “perceber” a um classificador deste tipo que o facto de um

determinado atributo variar nos padrões de treino significa que esse atributo deve ser sempre e apenas considerado no seu valor relativo aos outros atributos e não no seu valor absoluto? Ou seja, certo atributo pode não funcionar como elemento classificador, mas funcionar como elemento diferenciador, permitindo distinguir um padrão de outro padrão, mas não contribuindo para a determinação de um dado padrão.

Todo este enredo traduz a complexidade deste problema e eleva a capacidade que os mecanismos cerebrais têm ao cooperarem entre si da forma dinâmica como o fazem.

Para agravar o problema, a maior parte das técnicas que estão embutidas nos métodos de decisão supervisionados ou não-supervisionados possuem dificuldades no que refere à obtenção da melhor solução. Uma podem ficar retidas em mínimos locais e outras são demasiadamente exigentes computacionalmente. Estes problemas densificam-se exponencialmente à medida que a dimensão estrutural da solução cresce. As próprias características que se adoptam para decidir sobre a melhor solução podem deixar de ser as melhores dependentemente do contexto. A métrica que é usada para avaliar sobre essas características consiste normalmente em distâncias euclidianas, o que também pode não garantir a melhor solução. Alternativas podem ser usadas, como por exemplo medidas de entropia. Mas mesmo neste caso a melhor utilização de uma ou de outra parece variar de imagem para imagem.

Tudo isto adensa a dificuldade de decidir sobre a estrutura e sobre os métodos a usar na minimização global destes problemas. Uma estrutura de solução que funciona razoavelmente num contexto pode não operar bem em outro contexto não muito distinto. Tentar, de alguma forma, fundir duas estruturas de solução que funcionam bem nos seus respectivos contextos parece por vezes irrealizável, passando a solução composta a comportar-se melhor no geral, mas pior quando comparada com o comportamento que existia sobre os respectivos domínios antes dessa fusão.

Assim, as soluções que minimizam estes problemas são estruturalmente complexas, na medida em que se pretende o aumento do desempenho funcional, e requerem assim suportes físicos de capacidade extrema, por vezes inexistentes.

1.2 Abrangência da tese

A evolução desta tese resulta num formato que adveio essencialmente do próprio percurso que os trabalhos tiveram. A avaliação de técnicas de segmentação e também o conhecimento de técnicas de classificação levou a que se considerassem soluções híbridas formadas por ambas. Uma das deduções a que esta solução híbrida conduziu foi a da carência de suporte físico para o aproveitamento dos métodos propostos. Também o aumento da eficácia funcional destas tarefas visuais, a partir da adição de técnicas mais completas e complexas, resulta na necessidade de se obterem suportes físicos de capacidade ímpar relativamente aos existentes. Neste sentido, são estudados métodos quânticos que proporcionam uma aceleração de magnitude considerável e que em perspectiva portam novas resoluções para o problema.

Convém notar-se que o objectivo permanente foi o de se estudar tanto quanto possível os problemas que limitam a aproximação à visão humana de técnicas artificiais. Não se tentou admitir ou forçar, por processos externos à visão, que as imagens assegurassem uma ou outra característica que simplificasse o processo. De facto, em situações visuais naturais, por uma razão ou por outra, esse facto pode até ser difícil de garantir. A complexidade exigida nas soluções, para se obter algum aumento de desempenho funcional, fica demonstrada nas experiências feitas com os vários métodos construídos. Alguns dos métodos são reformulados ou associados a outros de forma a se procurar garantir o aumento de desempenho. As reformulações algorítmicas aqui apresentadas dirigem a necessidade para um aumento exponencial da capacidade de execução das máquinas que suportam os algoritmos de visão. Presentemente, a tecnologia capaz de vir a garantir aproximação a essa grandeza de aumento computacional é a computação quântica. Sem implementação física, para já, a uma escala desejável, apenas se pode definir teoricamente a sua utilização. Alguns algoritmos apresentados são assim formatados de acordo com as leis da física quântica.

Um dos métodos no qual assentou a procura de soluções para a minimização do problema foi o de segmentação multi-escala. Neste tipo de segmentação são seguidos alguns pressupostos de semelhança para produzir a agregação de pixels. Os atributos de distância entre pixels e de semelhança no que concerne ao valor do pixel são usados para produzir a agregação a múltiplas escalas. No entanto, os pressupostos de agregação utilizados por esse mecanismo não são suficientes para se obter um segmentador

universal. Com o objectivo de minimizar essa insuficiência, é adicionada uma componente de supervisão para a qual, por ser computacionalmente exigente, é sugerida a adaptação ao algoritmo quântico de Grover.

No que toca à utilização de modelos híbridos, menos dirigidos e dirigidos, foram usadas redes neuronais não supervisionadas e redes supervisionadas. As redes não supervisionadas de acoplamentos por pulsos apresentam características singulares e semelhantes às registadas no córtex visual. Funcionando como um mecanismo prévio de segmentação, é apresentada a sua associação a classificadores SVM¹ (Support Vector Machine). Esta associação tem por objectivo ajudar a aumentar a abrangência de actuação do segmentador. As partes não supervisionadas reduzem o espaço de soluções e minimizam o efeito da variabilidade de características que não contribuem para a classificação, mas que operam na diferenciação. O classificador SVM supera, por inerência, o problema, existente em muitos classificadores, de retenção em mínimos locais durante a aprendizagem, contribuindo para o alcance da melhor solução dentro do domínio que se lhe sobpõe. A SVM tem então a tarefa de aprender as características que podem ser singulares e contextuais sobre a pertença de um pixel a uma região, mas já sem o efeito destabilizador de algumas das características, sendo este efeito absorvido pelas PCNN.

No lado da classificação de regiões foi também construído um mecanismo semelhante, que resulta da associação das duas estruturas. Os objectivos para uma rede (PCNN) e para a outra rede (SVM) são exactamente os mesmos usados na segmentação. A diferença reside agora no resultado final, que, em vez de determinar a associação de um pixel a uma região (segmentação), determina a pertença de um conjunto de pixels a uma determinada classe que constituirá a região em análise.

Foi ainda sugerida uma adaptação das redes PCNN no suporte à classificação de contornos, o que permitiu observar o potencial e as limitações da aplicação destas redes a outros domínios que não os originalmente considerados.

A grande exigência computacional que os métodos anteriores apresentam, bem como a natureza do problema, levou a que se estudassem soluções que assentam na física quântica dirigida (computação quântica) e soluções que assentam na física

¹ Normalmente os classificadores SVM são referidos também como redes SVM.

quântica menos dirigida (difracção quântica). Relativamente à primeira, a solução apresentada, exemplificativa, permite a pesquisa numa imagem de um segmento de recta. Uma vez encontrado, a segmentação e a classificação ficam inerentemente concluídas. A extrapolação do processo para imagens naturais é possível, mas, devido à sua dimensão, não foi possível simulá-lo no emulador quântico. O método de descrição de regiões e de formas, baseado em difracção, foi proposto para avaliar o potencial de utilidade de um método físico menos dirigido na sua aceleração e também na capacidade de discriminação e correcta detecção de regiões e de formas.

Todos os diversos domínios explorados na tese, com a excepção, por motivos já referidos, da quântica dirigida, foram confrontados com diversos métodos mais ou menos clássicos e todas as análises feitas atingiram a componente de segmentação e/ou de classificação de imagens naturais do caso de estudo.

1.3 *Trabalho anterior*

Os estudos que foram realizados anteriormente [3] serviram como ponto de partida ao desenvolvimento desta tese. As soluções aí estudadas e apresentadas permitiram ter as primeiras percepções acerca da natureza e das dificuldades neste domínio. O trabalho então desenvolvido focava o problema da segmentação e da classificação de leucócitos em imagens naturais. Nesse trabalho foram usados classificadores neuronais de aprendizagem supervisionada para suportar a acção de segmentação e de classificação. A principal conclusão retida desse estudo foca a problemática de que um sistema baseado estritamente em classificadores supervisionados projecta erros, se as imagens futuras a segmentar e a classificar forem distintas, em algum aspecto, das imagens usadas no treino dos métodos supervisionados. Essas conclusões conduzem à necessidade de se obterem soluções de segmentação e de classificação que minimizem o problema e que permitam deduzir mais profundamente sobre a natureza destes temas.

1.4 *Caso de estudo*

O caso de estudo abordado advém dos trabalhos já elaborados anteriormente. Devido à natureza complexa, no campo da visão artificial, e ao incumprimento funcional que a aplicação continua a deter, esta foi escolhida como cerne de testes para as soluções de segmentação e de classificação estudadas.

A descrição do problema é a seguir exposta:

O sangue humano é uma substância fluida constituída por células e por plasma². A principal população celular existente no sangue divide-se em três classes: leucócitos³, hemácias⁴ e plaquetas⁵. A cardinalidade relativa e absoluta de cada um destes grupos está directamente associada com a existência ou não de anomalias num indivíduo. A classe dos leucócitos, no caso de um indivíduo são, está ainda dividida em cinco subclasses: basófilo, eosinófilo, neutrófilo, linfócito e monócito. Contudo, existem situações anómalas nas quais podem ser visualizados no sangue de um paciente outros tipos de leucócitos. Os leucócitos são produzidos na sua maior parte pela medula óssea e aí permanecem durante um período de amadurecimento até serem difundidos para a circulação sanguínea. Porém, em determinadas afecções estes podem, ainda imaturos, ser lançados na circulação. Alguns destes jovens leucócitos têm o nome de mieloblastos.

Os tipos de análises ao sangue executados nos actuais laboratórios de análises clínicas são centenas, encontrando-se quase todos mais ou menos automatizados. Uma das análises realizadas nestes laboratórios consiste na diferenciação percentual dos diversos tipos de leucócitos existentes numa amostra de sangue. Na Tabela 1.I podem constatar-se os valores percentuais típicos de uma análise deste teor a uma pessoa sã. Desde há já algum tempo que surgiram no mercado máquinas capazes de diferenciar as cinco classes normais de leucócitos. Estas máquinas recorrem à utilização de reagentes químicos que em associação com os vários tipos de células sanguíneas produzem características diferenciadoras num feixe de luz LASER posteriormente observado e analisado electronicamente. No entanto, as restantes, como é o caso dos mieloblastos, continuam a ter que ser discriminadas por acções visuais, executadas por pessoal

² A quase totalidade do plasma sanguíneo é constituído por água, na qual se encontram dissolvidos proteínas, sais, substâncias nutritivas e produtos de degradação.

³ Também denominados por glóbulos brancos. Estas células desempenham funções imunitárias importantes. Ao longo deste trabalho permutar-se-á entre as duas designações (glóbulos brancos / leucócitos).

⁴ Também designadas por glóbulos vermelhos ou eritrócitos, as hemácias têm como tarefa principal a distribuição de oxigénio e de substâncias nutritivas pelo organismo, além da remoção de produtos de degradação.

⁵ Também conhecidas por trombócitos, participam na coagulação sanguínea.

especializado. Estas acções consistem na contagem de cem leucócitos através de um microscópio óptico e, à medida que se vai procedendo à identificação, vai-se registando a quantidade de leucócitos de um e de outro tipo, obtendo-se assim a relação percentual entre os glóbulos brancos encontrados. Todavia, este trabalho é muito moroso e cansativo.

Tabela 1.I Percentagens de leucócitos existentes no sangue de uma pessoa sã

| | Tipos de Leucócitos | | | | | |
|---|---------------------|-----------|-------------|------------|-----------|--------------|
| | Eosinófilos | Basófilos | Neutrófilos | Linfócitos | Monócitos | Mieloblastos |
| % | 2.5 | 0.5 | 52.0 | 40.0 | 5.0 | 0.0 |

De uma forma geral, os diversos passos do processo consistem no seguinte: depois de ser feita a colheita de sangue a um indivíduo, esta é centrifugada⁶ de modo a separar o volume plasmático do volume celular. Este, por sua vez, é adicionado a um composto químico a que se dá o nome de *corante*. Esta aglutinação será espalhada numa lâmina de vidro, dando origem a um esfregaço seco que irá permitir que as células sejam visualizadas num microscópio óptico. No passo seguinte, a lâmina é colocada num microscópio óptico, as imagens são então obtidas por uma câmara de vídeo CCD⁷ a cores e transferidas para um PC onde uma carta digitalizadora as converte e as entrega ao software. As imagens adquiridas apresentam informação visual dos constituintes a nível celular. Glóbulos brancos, glóbulos vermelhos e plaquetas sanguíneas são algumas das células possíveis de encontrar numa imagem deste tipo. Uma vez que as células que neste caso interessam classificar são os glóbulos brancos, o processo de segmentação terá que os isolar na imagem, para que possam ser posteriormente classificados. A natureza visual existente neste género de imagens pode ser observada na Figura 1.1. Repare-se que, tanto numa como noutra imagem, a célula central representa um leucócito circundado por glóbulos vermelhos. A grande variabilidade em termos de cor que este tipo de imagens apresenta adiciona dificuldades no processo de segmentação dos glóbulos brancos. A questão põe-se da seguinte forma: como determinar quais os pixels que constituem o leucócito e quais os que não o constituem? Esta decisão não

⁶ Depois de inserir o sangue num tubo de ensaio, este é colocado numa centrifugadora a alta rotação, para que, por acção da força centrífuga daí resultante, o plasma seja separado das componentes celulares, pois as densidades destes dois componentes são distintas.

⁷ Terminologia vulgarmente usada para designar câmaras de vídeo que utilizam como órgão detector de imagem um CCD **C**harged-**C**oupled **D**evice.

pode ser tomada unicamente através da luminância e crominância do próprio pixel, mas terá também de ser tomada com base na distribuição espacial, a várias escalas, dos pixels vizinhos conexos e não conexos.



Figura 1.1 **Imagens de leucócitos obtidas num microscópio óptico com uma ampliação de 40x10**

Depois de os glóbulos brancos existentes numa imagem terem sido detectados pelo processo de segmentação, é necessário classificá-los. Este trabalho pode ser executado por um classificador, que terá de identificar correctamente as diversas classes de glóbulos brancos. Ao se observarem os dois leucócitos nas imagens da Figura 1.1, conclui-se que a mutabilidade geométrica entre as duas células é elevada, tornando evidente a subtilidade das características visuais que as definem. Portanto, não é simples descrever e transferir para um algoritmo o conhecimento que permite diferenciar as diversas classes de leucócitos. Sem dúvida que, seja qual for o processo usado, o sucesso da classificação depende em grande medida da qualidade da segmentação, estando condenada à partida uma tentativa de classificação de uma região que tenha sido parcial ou equivocadamente segmentada.

Por serem relativamente escassos os casos clínicos que implicam a existência de leucócitos anormais no sangue, não foi possível testar as soluções desenvolvidas com essas classes de células. Daí que se tenham restringido os testes às cinco classes de leucócitos normais.

1.5 Características do suporte físico

O suporte físico utilizado no desenvolvimento deste sistema divide-se em dois grupos. O primeiro reúne os dispositivos necessários à aquisição das imagens, Figura 1.2. O segundo alberga os algoritmos desenvolvidos. Trata-se de um computador pessoal (PC). No primeiro conjunto pode-se encontrar como dispositivo de partida o microscópio óptico, órgão fundamental na qualidade das imagens recolhidas. Acoplado

ao microscópio existe ainda a câmara de vídeo a cores de tecnologia CCD, responsável por converter as imagens ampliadas pelo microscópio em sinais eléctricos do tipo Y/C⁸. Este sinal é desmodulado numa carta de digitalização de imagens que se encontra conectada ao computador através de uma linha USB (Universal Serial Bus).

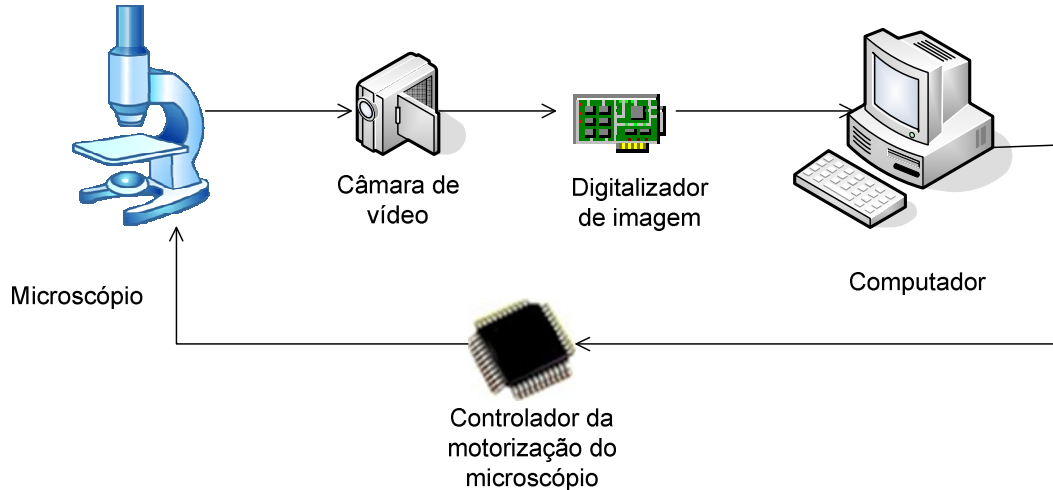


Figura 1.2 Interligação dos principais componentes constituintes do sistema de aquisição de imagens

1.5.1 Hardware utilizado

Microscópio Óptico

- Nikon Labophot;
- Triocular;
- 5 Objectivas;
- Ampliação Max. – 100×10.

Câmara de vídeo CCD

- Sony – SSC-DC50P;
- Resolução Horizontal. (PAL) – 470 Linhas;
- Pixels (CCD) – 752 (H)×582 (V).

Digitalizador de vídeo

- AVerMedia DVD EZMaker Pro USB.

Computador

- Intel – Pentium IV@ 3GHz;
- 2GByte – RAM.

⁸ Y/C Luminância e Crominância separadas. Técnica usada na codificação de imagens de vídeo a cores, que tem como objectivo melhorar a qualidade das imagens em termos de crominância relativamente à vulgar codificação de imagem em vídeo composto.

1.6 Controlo do microscópio

Para tornar mais rápida e autónoma a recolha das imagens de trabalho, foi desenvolvido um sistema de motorização do microscópio, Anexo D. O sistema é instruído pelo computador e sincronizado com a tarefa de aquisição das imagens. O sistema é formado por dois motores de passo, acoplados à mesa *xy* do microscópio, que permitem o posicionamento, sobre o plano da lâmina, da objectiva óptica do microscópio. Esse posicionamento é feito de forma progressiva, ao longo da área de interesse da lâmina e à medida que decorre a aquisição das imagens. Um terceiro motor foi acoplado ao mecanismo de focagem do microscópio para se garantir uma qualidade de focagem semelhante entre as imagens adquiridas da superfície da lâmina. O algoritmo de focagem adoptado procura encontrar a melhor focagem com base na maximização do valor das diferenças entre cada um dos pixels e os seus correspondentes vizinhos mais próximos, até uma ordem de proximidade de dois pixels. Os limites posicionais da mesa *xy* e do sistema de focagem são observados com recurso a sensores de reflexão de luz infravermelha. A comunicação com o computador é executada sobre um barramento série (RS232). O controlador de cada motor de passo recebe os dados de posicionamento, partilhando com os restantes controladores o barramento série que os liga ao computador. Como o barramento é partilhado, foi constituído um sistema de endereçamento simples que identifica cada um dos controladores dos motores. Dispõe também de um regulador de luminosidade do microscópio.

O software construído, que gere todas estas funcionalidades, permite a distribuição alternada das imagens adquiridas por vários computadores ligados em rede. Esta funcionalidade possibilita que o processamento das imagens tenha índole paralela.

1.7 Software

O software usado no desenvolvimento das rotinas dos vários métodos foi escrito na linguagem de programação orientada ao objecto C++. O compilador utilizado foi o Visual Studio .NET 2003 da Microsoft. A interface com o utilizador é do tipo gráfica por janelas.

Com a exceção do classificador SVM e do simulador de computação quântica, todos os métodos descritos na tese foram integralmente ou quase integralmente construídos ao longo dos trabalhos correspondentes. Desde a plataforma de segmentação multi-escala, à rede PCNN, passando pelo processo quântico menos dirigido, todos foram construídos, utilizando C++, em todos os seus níveis de implementação. A SVM foi integrada nos processos desenvolvidos com recurso a uma DLL (Dynamic Link Library) proprietária. O simulador de computação quântica usado é suportado por um grupo de 32 processadores pertencentes ao Instituto Fraunhofer na Alemanha, que também desenvolveu o simulador, e o seu acesso é feito através de interface WEB.

1.8 Trabalho de suporte de outros autores

Relativamente aos métodos de segmentação, procurou-se estabelecer comparações e análises dos métodos desenvolvidos entre alguns daqueles que são considerados potencialmente mais abrangentes e capazes em termos de áreas de aplicação.

A pesquisa produzida conduziu a que se considerassem, em partes deste trabalho, as redes PCNN e a segmentação multi-escala como processos não supervisionados e as SVM como elementos supervisionados.

Relativamente ao caso de estudo, foram analisados alguns trabalhos [4;5] e todos detinham especificidades nas suas soluções que não permitiam um desempenho elevado sobre um conjunto diverso de tipos de imagens. Por exemplo, o trabalho de Ramoser [4] assume que o matiz do núcleo das células tem um valor semelhante ao matiz do citoplasma dessas células. Assume também que o valor da saturação da cor dos núcleos das células é sempre diferente do de qualquer outra região das imagens. O que se verifica é que estes pressupostos não são sempre verdadeiros e estão dependentes de múltiplos factores.

Uma vez que não existe neste domínio uma base de dados de imagens global, partilhada pelos investigadores, não é possível criar comparações directas entre os resultados dos trabalhos existentes.

De realçar que a secção nesta dissertação que descreve o funcionamento das SVM é baseada num trabalho de Steve Gun [6]; a secção que descreve o formalismo matemático dos descritores de Zernike é baseada num trabalho de Shutler [7] e a descrição sobre os aspectos teóricos da computação quântica baseia-se num trabalho de E. Rieffel e de W. Polak [8].

1.9 Contribuições

A contribuição mais geral desta tese pode ser retirada da sequência de capítulos apresentados. Daí surgem compreensões acerca do problema da segmentação e da classificação que, apesar de não serem totalmente novas, reforçam a caracterização e a forma de como estes mecanismos devem ser abordados.

A sugestão de um método de segmentação multi-escala supervisionado por mecanismos quânticos, referida no quarto capítulo e explicada no sexto capítulo, do conhecimento que se tem, é original.

As soluções apresentadas no quinto capítulo procuram a associação entre métodos não-supervisionados e supervisionados. A utilização, em cooperação, de redes de acoplamentos por pulsos, com os descritores de Zernike e com as máquinas de suporte vectorial, pelo menos no caso de aplicação estudado, é original. A reestruturação das redes de acoplamento por pulsos para elevar o seu desempenho na classificação de contornos, de acordo com o formato de solução usada, também é original.

O estudo feito, apresentado no capítulo sexto, relativo à adaptação do algoritmo de Grover de computação quântica ao problema da segmentação e classificação em imagens também é original, na medida em que não foi encontrado um algoritmo análogo no ainda pequeno grupo de algoritmos clássicos já adaptados a algoritmos quânticos.

O descritor quântico apresentado no sétimo capítulo também é original e subscreve a ideia de que é possível, em algumas situações, extrair de forma menos dirigida, em oposição ao que é feito no capítulo anterior, benesses de um suporte quântico. Ainda nesse capítulo é descrito, em apoio ao descritor quântico proposto,

recorrendo, originalmente, a pressupostos quânticos, o comportamento de uma partícula que atravessa uma abertura circular, secção (7.5).

1.10 Organização da tese

Devido à natureza diferencialmente difusa dos dois processos, o de segmentação e o de classificação, não se lhes faz uma distinção ordenada por capítulos, preferindo-se fazer a partição e/ou a agregação apenas e de acordo com os tipos de soluções abordadas e residentes nos respectivos capítulos.

No presente capítulo são feitos os enquadramentos justificativos dos trabalhos desta tese.

O segundo e o terceiro capítulo fazem uma descrição de mecanismos que são comuns a vários capítulos. O segundo capítulo descreve todo o funcionalismo de classificadores supervisionados, Support Vector Machine (SVM). As SVM foram utilizadas em vários processos, aqui apresentados, como entidade de aprendizagem supervisionada. No terceiro capítulo são comparados modelos de descrição numérica de regiões, com foco sobre os descritores de Zernike, que se apresentam úteis na redução dimensional do espaço de representação tanto para problemas de segmentação como de classificação em imagens. Também estes descritores são usados neste trabalho em associação a múltiplos métodos.

No quarto capítulo são estudados alguns métodos de segmentação não supervisionada com o objectivo de se perceberem as debilidades e o potencial desses métodos e de como podem ser aproveitados na adição a um método supervisionado. Neste sentido, é sugerida a interligação de uma segmentação multi-escala a um processo quântico.

Uma vez que os métodos não supervisionados e os métodos supervisionados se complementam mutuamente, no quinto capítulo são associados modelos neuronais de característica impulsional não supervisionados (PCNN) a modelos supervisionados (SVM). São apresentadas soluções de segmentação e soluções de classificação. São também apresentadas, com a mesma associação, soluções de classificações orientadas ao contorno.

No sexto capítulo é sugerida e emulada uma solução de segmentação e de classificação em suporte quântico, recorrendo à computação quântica, com uma perspectiva aceleradora.

No sétimo capítulo é sugerido um mecanismo de descrição de formas que tem como base a interferência e a difracção quântica.

No oitavo capítulo são apresentadas as conclusões gerais obtidas dos trabalhos executados nesta tese.

A maior parte dos capítulos empregam e analisam os resultados sobre imagens relacionadas com o caso de estudo. Além disso, para uma maior compreensão, são também usadas imagens sintéticas com conteúdos de geometria diversa.

Os anexos apresentam, resumidamente, alguns temas cuja natureza serve de suporte à compreensão de alguns dos assuntos abordados na tese.

1.11 Notação quântica

O sumário de algumas notações quânticas com tradução para noções de álgebra linear é mostrado na Tabela 1.II.

Tabela 1.II Notação quântica

| Notação | Descrição |
|--------------------------------------|---|
| z^* | Conjugado complexo de um número complexo z . $(1+i)^* = 1-i$ |
| $ \psi\rangle$ | Representa um vector. Conhecido como <i>ket</i> . |
| $\langle\psi $ | Representa a adjunta do vector $ \psi\rangle$. Conhecido como <i>bra</i> . |
| $\langle\varphi \psi\rangle$ | Representa o produto interno do vector $ \varphi\rangle$ com o vector $ \psi\rangle$. |
| $ \varphi\rangle\otimes \psi\rangle$ | Representa o produto tensorial do vector $ \varphi\rangle$ com o vector $ \psi\rangle$. |
| $ \varphi\rangle \psi\rangle$ | Notação abreviada do produto tensorial entre o vector $ \varphi\rangle$ e o vector $ \psi\rangle$. |
| $ \varphi\rangle\langle\psi $ | Representa o produto tensorial entre o vector $ \varphi\rangle$ e o vector $\langle\psi $. |

| Notação | Descrição |
|--------------------------------------|---|
| A^* | Conjugada complexa da matriz A . |
| A^T | Transposta da matriz A . |
| A^\dagger | Adjunta da matriz A , $A^\dagger = (A^T)^*$. $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}.$ |
| $\langle \varphi A \psi \rangle$ | Produto interno entre $ \varphi\rangle$ e $A \psi\rangle$. Equivalente a obter o produto interno entre $A^\dagger \varphi\rangle$ e $ \psi\rangle$. |

Capítulo 2

Support Vector Machine

O presente capítulo descreve a formulação de base ao funcionamento de um classificador, designado por Support Vector Machine. Perante os domínios em estudo nesta tese houve necessidade de seleccionar, para utilização em várias soluções, um classificador que apresentasse características que se distanciassem, positivamente, das apresentadas por classificadores comuns. A descrição mostrada revela os mecanismos principais de como este classificador pode obter melhores desempenhos em algumas aplicações.

2.1 Introdução

A obtenção de um modelo, a partir de um conjunto de dados, que determine de forma correcta a pertença, ou não, de dados futuramente apreciados a uma determinada classe tem sido um problema de difícil solução. São conhecidos os condicionalismos ostentados pelos próprios dados usados na obtenção do modelo [3]. A qualidade bem como a quantidade desses dados de treino são factores de grande influência que participam na obtenção de um modelo capaz de produzir, ou não, na fase de pós-treino um desempenho elevado. Os dados de treino são amostrados de uma forma não uniforme e o seu número, ao ser finito, apenas permite que seja alcançado um domínio esparso no espaço de entrada. As implicações destas condicionantes projectam-se no âmbito dos problemas *ill posed* [9] em subjugação às teorias de Hadamard [10]. As redes neuronais, como as *feed-forward/back-propagation*, tipicamente apresentam

dificuldades desta índole, com capacidades pouco generalizadoras, ao produzirem modelos que tendem a representar os dados com demasiado detalhe e prendimento (*over-fitting*). Estes problemas são consequência, também, dos métodos de optimização usados na obtenção dos parâmetros que representam os modelos e das formas estatísticas que se empregam nas medidas de desempenho do modelo a ser encontrado. As bases das *Support Vector Machine* foram propostas por Vapnik em 1995 [11]. A sua popularidade foi aumentando devido a possuírem características que favorecem o seu comportamento relativamente às redes neuronais tradicionais. Esse facto resulta de se servirem do conceito de minimização estrutural do risco (Structural Risk Minimization – SRM), que foi mostrado ser superior [12] ao conceito de minimização do risco empírico (Empirical Risk Minimization – ERM) que é usado pelas redes neuronais mais convencionais. O SRM minimiza um limite superior do risco esperado, ao contrário das redes convencionais que minimizam o erro sobre os padrões de treino. Esta diferença possibilita que as SVM generalizem melhor, o que é um factor fundamental para um bom desempenho em aprendizagem estatística, feita a partir de dados. Inicialmente propostas para a resolução de problemas de classificação, foram mais tarde introduzidas, também, nos problemas de regressão [13]. Genericamente usa-se a mesma terminologia para referir o tratamento dos dois casos pelas *Support Vector Machine*. Mas em situações de necessidades de diferenciação é comum chamar aos métodos, em caso de classificação, Support Vector Classification (SVC) e, no caso da regressão, Support Vector Regression (SVR).

2.2 Teoria da Aprendizagem Estatística

Na teoria da aprendizagem estatística [11], é considerado um espaço de soluções, designado por espaço T ou por *espaço conceptual*, Figura 2.1, no qual se considera existir uma função $f_0(x)$ que modeliza o problema de aprendizagem de forma óptima para os padrões x . Este espaço é demasiadamente complexo de tratar pois suporta funções com uma infinidade de parâmetros. Além disso, a partir de um conjunto de padrões de treino, em número limitado, não é possível conhecer e caracterizar o espaço T.

Para tornar possível a aprendizagem através de padrões, em número limitado, é importante confinar as funções de modelização a um espaço menos complexo e assim tratável. Este espaço é designado por H_n onde n representa o índice de complexidade deste espaço. Neste espaço existe uma função $f_n(x)$ que melhor se aproxima da função óptima $f_0(x)$. A diferença produzida por estas duas funções para um mesmo padrão x resulta no Erro de Aproximação. Este erro não está directamente relacionado com os padrões de treino, mas sim, com as diferenças estruturais entre os dois espaços, o T e o H_n .

O estimar de uma função pertencente ao espaço de hipóteses é feito com um processo de aprendizagem. O objectivo é o de conseguir obter uma função $\hat{f}_{n,l}(x)$ que iguala $f_n(x)$ com base na informação dada pelos padrões de treino. A diferença entre estas duas funções resulta no Erro de Estimativa. Com o aumentar da complexidade n do espaço de hipóteses tende a aumentar o Erro de Estimativa para um mesmo número l de padrões de treino. Este erro pode diminuir aumentando o número l de padrões de treino.

A soma do Erro de Aproximação com o Erro de Estimativa perfaz o Erro de Generalização. É este erro que num problema prático interessa minimizar pois reflecte o comportamento do sistema na presença de padrões que não existiam na fase de treino.

Realce-se que o Erro de Estimativa não existirá se x representar apenas os padrões de treino. Nesta situação, apenas existe o Erro de Aproximação e o sistema apresentará *over-fitting*. Quando a aprendizagem é realizada unicamente com base em dados de treino sem se conhecer a distribuição probabilística destes, diz-se que a aprendizagem é empírica. Por estes motivos o Erro de Aproximação é também conhecido por Erro Empírico.

O desafio está em conseguir um sistema de aprendizagem que, de acordo com o número l de padrões de treino existente, escolha a complexidade n do espaço de Hipóteses de forma a minimizar o Erro de Generalização.

Segundo a teoria da aprendizagem estatística o processo de aprendizagem das redes neuronais relativamente às SVM pode ser diferenciado da seguinte forma:

Nas redes neuronais *feed-forward*, é escolhido o número de neurónios da camada escondida, mantida uma relação, conhecida por *Intervalo de Confiança*, entre o

número de padrões de treino e a complexidade do respectivo espaço H_n e é minimizado o erro de treino (erro empírico).

Nas SVMs, é estabelecido um determinado valor para o erro empírico e é minimizado o *intervalo de confiança*.

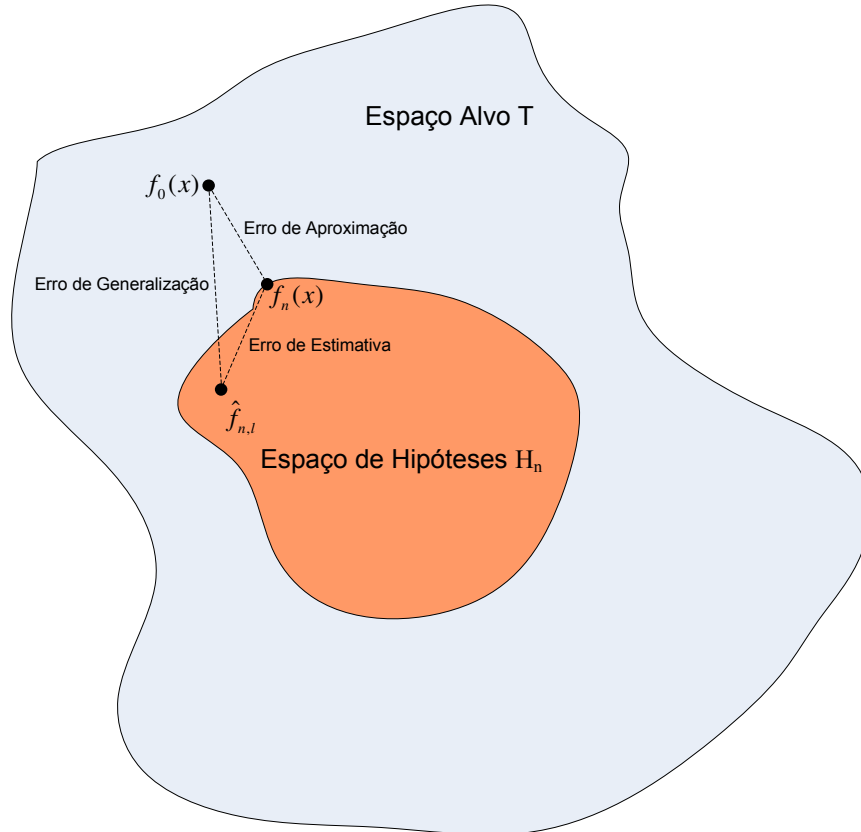


Figura 2.1 Erros de modelação perspectivados nos diferentes espaços de soluções

2.3 Risco

O conceito de risco está associado à probabilidade de um par de padrões (x,y) acontecer e a um custo sobre esse mesmo par de padrões para o contexto em causa. O risco integra todo o domínio de par de padrões existentes no contexto e possibilita quantificar o desempenho de um classificador f . O par de padrões é constituído pelo padrão a analisar x e pelo padrão y que deverá indicar a classe de pertença de x .

Para aumentar a capacidade de generalização deve garantir-se que o hiperplano particionador do espaço de classificação não apenas separe os elementos de uma classe dos da outra classe, em elementos de treino, mas também que o faça para os elementos

que apareceram em fase após treino. Para isto ser realizável, o hiperplano separador das classes deverá estar situado o mais afastadamente possível dos elementos de ambas as classes que lhe estão mais próximos. Se isto acontecer, reduz-se o risco de que apareçam, futuramente, elementos a classificar que acabem por cruzar erradamente o hiperplano separador.

Assim, pretende-se encontrar uma função f que minimize esse risco:

$$R[f] = \int_{X \times Y} L(y, f(x)) P(x, y) dx dy. \quad (2.1)$$

Em que L é chamada função de *perda* que pretende avaliar a diferença entre o valor a obter e o valor obtido sobre um determinado par de padrões (x, y) , usando um modelo de classificação f no qual x é argumento. Se $f(x)$ tiver um valor igual a y , o valor representado por L será o zero. Mas se y e $f(x)$ forem diferentes e se $y \in \{-1, 1\}$ e $f(x) \in \{-1, 1\}$, então L poderá sugerir a perda com um valor superior a zero. $P(x, y)$ corresponde à distribuição probabilística do par (x, y) e, sendo estas variáveis independentes entre si, podem ser representadas pelo produto das distribuições de cada uma, $P(x)P(y)$.

Porém, a distribuição $P(x, y)$ é desconhecida. Mas é possível obter uma aproximação do risco seguindo o conceito do risco empírico (R_{emp}), em que l representa o número de pares de padrões (x, y) :

$$R_{emp}[f] = \frac{1}{l} \sum_{i=1}^l L(y^i, f(x^i)). \quad (2.2)$$

Então, o modelo f que minimiza o risco, num espaço H_n , será dado por:

$$\hat{f}_{n,l}(x) = \arg \min_{f \in H_n} R_{emp}[f]. \quad (2.3)$$

A determinação do modelo a partir do risco empírico apenas pode ser considerada sabendo que:

$$\lim_{l \rightarrow \infty} R_{emp}[f] = R[f]. \quad (2.4)$$

O qual poderá ser válido quando l tomar valores relativamente elevados. Mas é também obrigatório que seja cumprida a seguinte relação:

$$\lim_{l \rightarrow \infty} \min_{f \in H_n} R_{emp}[f] = \min_{f \in H_n} R[f]. \quad (2.5)$$

A qual é apenas válida quando o espaço de hipóteses dos modelos, H_n , é suficientemente pequeno quando comparado com o número, l , de pares (x, y) .

Assim, é necessária a seguinte inequação, que permite estabelecer uma fronteira superior para o risco, com uma probabilidade de $1-\delta$, e na qual h representa a dimensão Vapnik-Chervonenkis (ver secção 2.4) do espaço de hipóteses:

$$R[f] \leq R_{emp}[f] + \sqrt{\frac{h \ln\left(\frac{2l}{h} + 1\right) - \ln\left(\frac{\delta}{4}\right)}{l}}. \quad (2.6)$$

Note-se que a Equação (2.6) não é dependente da função de densidade de probabilidade P .

2.4 Dimensão de Vapnik-Chervonenkis (VC)

A teoria de aprendizagem estatística [11], ou teoria VC, mostra que é imperativo restringir o conjunto de funções, das quais f é escolhido, a um conjunto adequado à dimensão do conjunto de padrões de treino. A redução desse conjunto está dependente do risco empírico e da capacidade da função de classificação f . Essa noção de capacidade está ligada à chamada dimensão de VC. Cada função classificadora separa os padrões de entrada de uma certa forma, levando à classificação de cada um através da correspondência, sobre esses padrões, de valores que tipicamente são dois valores, o $\{\pm 1\}$. No máximo, existem 2^m formatos de classificação para m padrões. Uma função eficaz consegue, potencialmente, combinar todos os arranjos possíveis de classificação num número de 2^m . A dimensão de VC é definida como o maior valor que m pode adquirir, para um conjunto de m combinações nas quais a função classificadora se pode moldar (Figura 2.2). Se esse limite m não existir, então a dimensão é infinita. O valor de

m pode ser traduzido como o valor de capacidade de aprendizagem de um sistema classificador. Existem outros métodos, mais precisos, usados no estabelecimento da capacidade do modelo, mas são mais difíceis de tratar, como é o caso da *annealed VC entropy* [14].

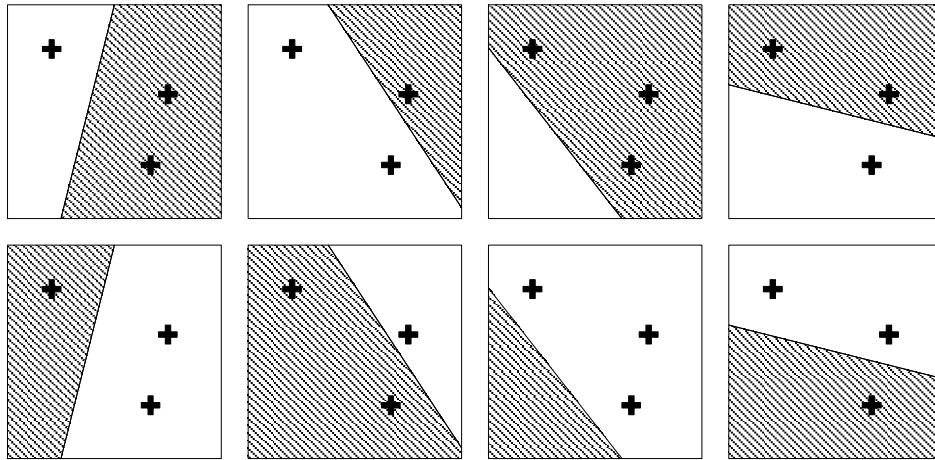


Figura 2.2 Exemplo de uma dimensão VC de valor 3 em \mathcal{R}^2 . Existem 2^3 combinações de correspondência entre os 3 pontos a classificar e as 2 classes (assinaladas correspondentemente pelas regiões a sombreado e a branco)

2.5 Minimização do Risco Estrutural (SRM)

É considerada uma estrutura, Figura 2.3, onde S_h é um espaço de hipóteses de dimensão VC h , tal que:

$$S_1 \subset S_2 \subset S_h \cdots \subset S_\infty. \quad (2.7)$$

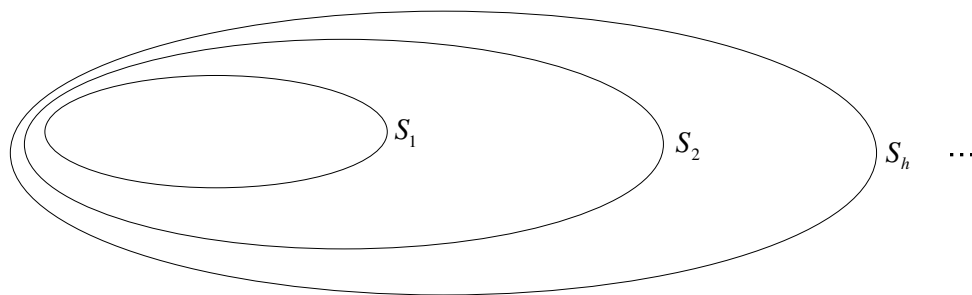


Figura 2.3 Representação do encadeamento de espaços de acordo com as suas dimensões

O SRM consiste na resolução do seguinte problema:

$$\min_{S_h} R_{emp} [f] + \sqrt{\frac{h \ln\left(\frac{2l}{h} + 1\right) - \ln\left(\frac{\delta}{4}\right)}{l}}, \quad (2.8)$$

de onde se deduz que a minimização do risco estrutural consiste na procura de um espaço de hipóteses que minimiza a combinação do valor de risco empírico com o valor resultante da capacidade, h , da máquina face ao número de dados de treino l , para um determinado valor probabilístico de $1 - \delta$.

2.6 Classificação por vectores de suporte

Considere-se o problema de classificação de um conjunto de padrões em uma de duas classes possíveis. A situação exige que se encontre um classificador que seja capaz de separar cada um dos padrões de acordo com a classe a que estes pertencem. O classificador é uma função que é modelada indutivamente através da informação contida em cada uns dos padrões. O obstáculo está em conseguir obter-se uma função classificadora que, não só classifique correctamente os padrões de treino, mas também os padrões que nunca apareceram na fase de treino. Ou seja, a função deverá ter capacidades de generalização. Considerando um espaço de classificação dividido por um de vários possíveis hiperplanos, a questão coloca-se na obtenção do hiperplano que maximiza a generalização. Esse plano será o que se distanciar mais dos pontos que lhe estão mais próximos pertencentes a cada uma das classes (Figura 2.4).

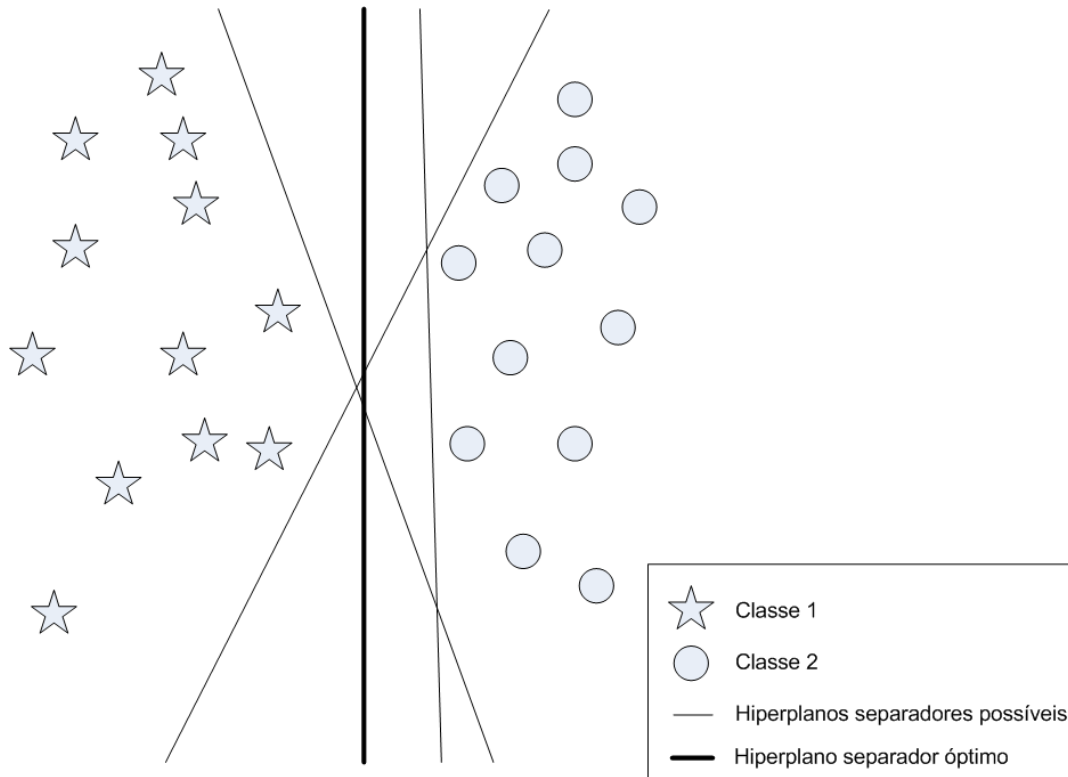


Figura 2.4 Várias possibilidades de hiperplanos separadores das duas classes, sendo o mais denso o hiperplano que otimiza a separação, ou seja, que eleva a capacidade de generalização

2.6.1 Optimização do hiperplano separador

Defina-se o problema da separabilidade de dois subconjuntos do conjunto de vectores de treino partindo da seguinte representação:

$$D = \{(x^1, y^1), \dots, (x^l, y^l)\}, \quad x \in \mathfrak{R}^n, y \in \{-1, 1\}. \quad (2.9)$$

O hiperplano separador pode ser tratado, através de um produto interno, como:

$$\langle w, x \rangle + b = 0, \quad (2.10)$$

no qual w é o vector perpendicular ao hiperplano, b é a menor distância do hiperplano à origem e x é um qualquer ponto do plano.

O conjunto de vectores de treino dizem-se separados por um hiperplano óptimo quando não existem vectores atribuídos erradamente a classes e quando a distância ao hiperplano, dos vectores que lhe estão mais próximos, é máxima.

De forma a tornar menos redundante a Equação (2.10), é apropriado apresentar uma equação canónica (com margem funcional igual a 1) do hiperplano [11] que estreita o domínio possível dos valores de w e de b :

$$\min_i |\langle w, x^i \rangle + b| = 1. \quad (2.11)$$

O significado empírico que se pode ler na Equação (2.11) é que a norma do vector w será igual ao inverso da distância do ponto ao hiperplano, do conjunto de dados, que está mais perto do hiperplano. A figura seguinte (Figura2.5) ilustra esta abordagem, na qual qualquer hiperplano se relaciona com o ponto que lhe está mais próximo.

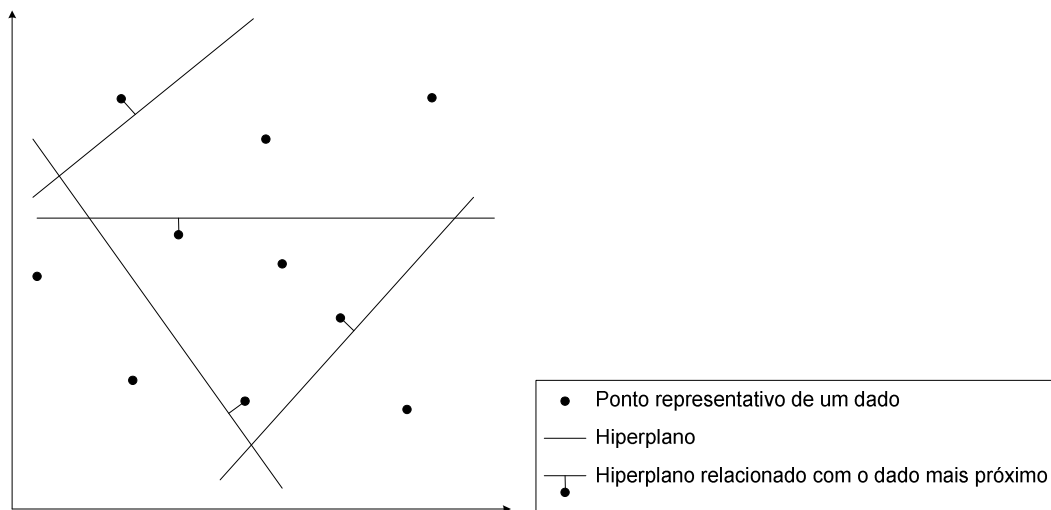


Figura2.5 Hiperplanos canónicos

Um hiperplano separador na forma canónica deve obedecer à restrição imposta pela Equação (2.12):

$$y^i [\langle w, x^i \rangle + b] \geq 1, \quad i = 1, \dots, l. \quad (2.12)$$

A distância $d(w,b;x)$ do ponto x ao hiperplano (w,b) é dada por:

$$d(w,b;x^i) = \frac{|\langle w, x^i \rangle + b|}{\|w\|}. \quad (2.13)$$

O hiperplano óptimo é conseguido maximizando a margem, ρ , ficando sujeita às restrições da Equação (2.12). No exemplo gráfico da Figura 2.6 estão assinalados, com a cor vermelha, os limites que definem a margem ρ . Estes são determinados, como se concluirá posteriormente, pelos denominados vectores de suporte (SV – dados que definem os limites da margem) durante o processo de aprendizagem. O hiperplano separador situa-se a meia distância entre os limites da margem ρ .

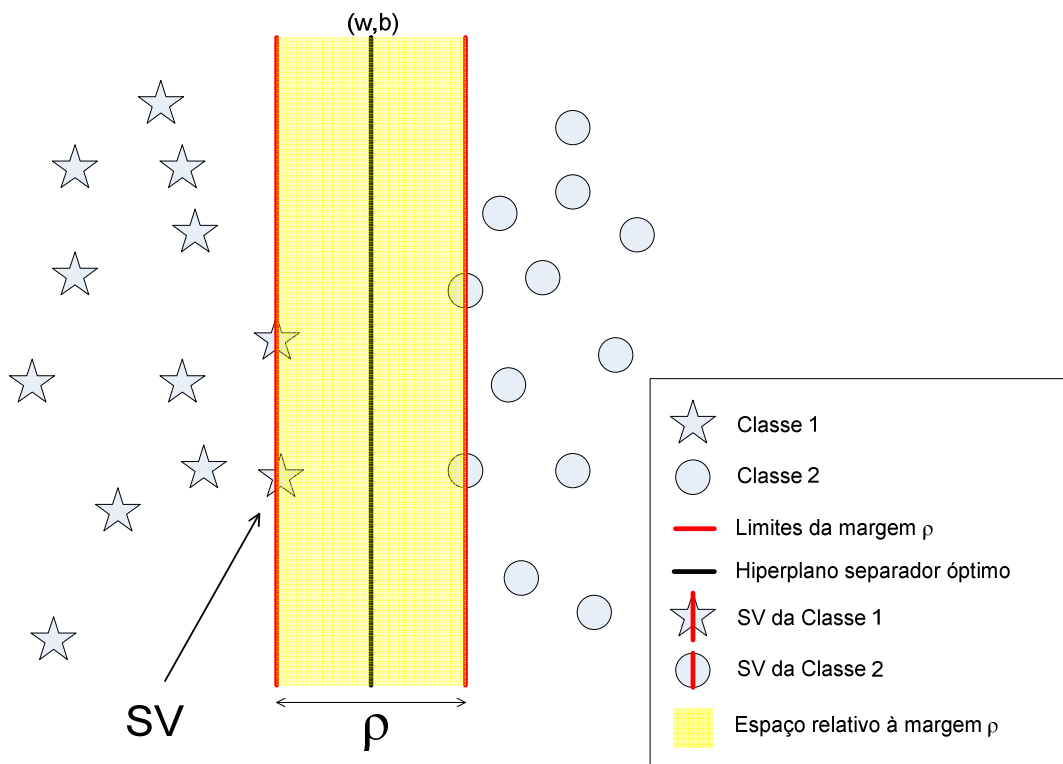


Figura 2.6 Estabelecimento da margem ρ através dos vectores de suporte (SV) no espaço dividido pelo hiperplano (w,b)

A margem ρ é então dada por:

$$\begin{aligned}
 \rho(w, b) &= \min_{x^i: y^i = -1} d(w, b; x^i) + \min_{x^i: y^i = 1} d(w, b; x^i) \\
 &= \min_{x^i: y^i = -1} \frac{|\langle w, x^i \rangle + b|}{\|w\|} + \min_{x^i: y^i = 1} \frac{|\langle w, x^i \rangle + b|}{\|w\|} \\
 &= \frac{1}{\|w\|} \left(\min_{x^i: y^i = -1} |\langle w, x^i \rangle + b| + \min_{x^i: y^i = 1} |\langle w, x^i \rangle + b| \right) \\
 &= \frac{2}{\|w\|}
 \end{aligned} \tag{2.14}$$

Devido às características matemáticas da função $\|w\|$, e por simplicidade, maximizar a margem ρ , $\frac{2}{\|w\|}$, é equivalente a minimizar $\frac{1}{2}\|w\|^2$. Então o hiperplano que separa de forma óptima os dados de entrada é um que minimize,

$$\Phi(w) = \frac{1}{2}\|w\|^2. \tag{2.15}$$

Repare-se que não existe dependência do factor b porque variar b significa variar a posição do hiperplano ao longo do seu vector perpendicular e a restrição imposta pela Equação (2.11) é utilizada na simplificação da Equação (2.14). Isto implica que o hiperplano separador óptimo se situe a igual distância do dado mais próximo de uma classe e do dado mais próximo da outra classe. Se a formulação estabelecida pela Equação (2.14) não fosse cumprida, o hiperplano poderia não ser o óptimo, pois poderia estar mais próximo de uma classe do que da outra.

Ao se minimizar a Equação (2.15) implementa-se o princípio de minimização do risco estrutural (SRM – Secção 2.5), supondo a seguinte desigualdade e considerando A um escalar no domínio de \mathfrak{R} ,

$$\|w\| < A. \tag{2.16}$$

Então, a partir da Equação (2.13) pode escrever-se,

$$d(w, b; x) \geq \frac{1}{A}. \tag{2.17}$$

Assim, os hiperplanos não podem estar a uma distância inferior a $\frac{1}{A}$ de qualquer dos pontos de dados, levando a que o número de hiperplanos possível seja reduzido, tal como se pode concluir da figura seguinte (Figura 2.7):

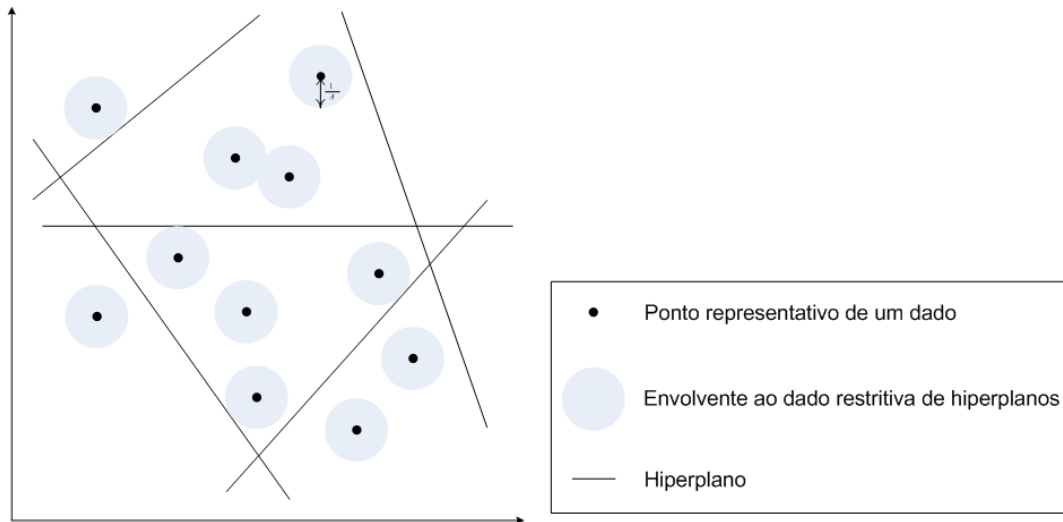


Figura 2.7 Restrição adicional aos hiperplanos canónicos

Desta forma, minimizar $\|w\|$ permite minimizar A , o que se traduz, como pretendido, numa minimização da capacidade da SVM.

A dimensão VC, h , no conjunto de hiperplanos canónicos num espaço de dimensão n é limitada por,

$$h \leq \min [R^2 A^2, n] + 1, \quad (2.18)$$

na qual R corresponde ao raio da hiperesfera que envolve todos os elementos de dados. Assim, minimizar a Equação (2.15) é equivalente a minimizar um limite superior da dimensão VC. A optimização da Equação (2.15) submetida à restrição formulada em (2.12) é dada pelo ponto estacionário de inflexão do funcional de Lagrange [15], Equação (2.19),

$$\Phi(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha^i (y^i [\langle w, x^i \rangle + b] - 1), \quad (2.19)$$

onde α corresponde aos multiplicadores de Lagrange. O Lagrangiano tem que ser minimizado relativamente a w, b e maximizado relativamente a $\alpha \geq 0$. Note-se que cada ponto de dados corresponderá a uma restrição, que por sua vez terá um coeficiente de Lagrange associado, α^i . A dualidade Langrangiana clássica estabelece a solução para o problema. Transformando a Equação (2.19) nessa dualidade obtém-se:

$$\max_{\alpha} W(\alpha) = \max_{\alpha} \left(\min_{w, b} \Phi(w, b, \alpha) \right). \quad (2.20)$$

O mínimo em função de w e de b do Langrangiano, Φ , é dado por,

$$\begin{aligned} \frac{\partial \Phi}{\partial b} = 0 &\Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \\ \frac{\partial \Phi}{\partial w} = 0 &\Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \end{aligned} \quad (2.21)$$

Combinando as Equações (2.19), (2.20) e a Equação (2.21), a forma de dualidade do problema aparece como,

$$\max_{\alpha} W(\alpha) = \max_{\alpha} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle + \sum_{k=1}^l \alpha_k. \quad (2.22)$$

Podendo a solução de α óptimo (α^*) ser dada por,

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle - \sum_{k=1}^l \alpha_k, \quad (2.23)$$

sujeita às restrições de,

$$\begin{aligned} \alpha_i &\geq 0 \quad i=1, \dots, l \\ \sum_{j=1}^l \alpha_j y_j &= 0 \end{aligned} \quad (2.24)$$

Resolvendo a Equação (2.23), sujeita as restrições de (2.24), determinam-se os coeficientes de Lagrange; então o hiperplano separador óptimo (w^*, b^*) será dado por:

$$\begin{aligned} w^* &= \sum_{i=1}^l \alpha_i y_i x_i \\ b^* &= -\frac{1}{2} \langle w^*, x_r + x_s \rangle \end{aligned}, \quad (2.25)$$

onde x_r e x_s são quaisquer vectores de suporte de cada uma das classes de acordo com,

$$\alpha_r, \alpha_s > 0, \quad y_r = -1, y_s = 1. \quad (2.26)$$

Pode-se, agora, formatar o classificador de acordo com,

$$f(x) = \text{sgn}(\langle w^*, x \rangle + b^*). \quad (2.27)$$

Ou então, pode-se usar um classificador, $h(z)$, que interpola linearmente a margem:

$$f(x) = h(\langle w^*, x \rangle + b^*), \quad h(z) = \begin{cases} -1 & : z < -1 \\ & : -1 \leq z \leq 1. \\ +1 & : z > 1 \end{cases} \quad (2.28)$$

Esta última forma poderá ser mais apropriada do que a sua antecedente, porque devolve um valor real situado no intervalo $[-1, 1]$, quando a resposta do classificador a um determinado dado se situar dentro dos valores limites da *margem*, onde não existem dados.

De acordo com a condição de Kuhn-Tucker, Equação (2.29),

$$\alpha_i \left(y^i \left[\langle w, x^i \rangle + b \right] - 1 \right) = 0, \quad i = 1, \dots, l \quad , \quad (2.29)$$

apenas os pontos que satisfazem,

$$y^i [\langle w, x^i \rangle + b] = 1, \quad (2.30)$$

terão multiplicadores de Lagrange diferentes de zero. A estes chamam-se vectores de suporte (Support Vectors – SV). Se os dados forem separáveis linearmente, todos os SV residirão na margem e o número de SVs poderá ser então muito pequeno. Isto faz que o hiperplano seja determinado por um subconjunto pequeno do conjunto de dados de treino. Ou seja, se os restantes pontos de dados fossem eliminados do conjunto de padrões de treino, o hiperplano assim obtido seria o mesmo que o obtido com a totalidade de pontos de padrões de treino. Desta forma, as SVM podem ser usadas para reduzir o volume de dados do conjunto de padrões de treino através dos elementos SV. Se os dados forem linearmente independentes, a seguinte igualdade é estabelecida,

$$\|w\|^2 = \sum_{i=1}^l \alpha_i = \sum_{i \in SVs} \alpha_i = \sum_{i \in SVs} \sum_{j \in SVs} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle. \quad (2.31)$$

A partir da Equação (2.18) deduz-se que a dimensão VC do classificador fica limitada por,

$$h \leq \min \left[R^2 \sum_{i \in SVs} \alpha_i, n \right] + 1. \quad (2.32)$$

Se os dados de treino, x , forem normalizados de forma a serem confinados a uma hiperesfera unitária, então a equação anterior escreve-se,

$$h \leq \min \left[\sum_{i \in SVs} \alpha_i, n \right] + 1. \quad (2.33)$$

Geralmente, na prática, a dimensão do espaço de entrada n é maior do que o número de vectores de suporte, o que faz com que a capacidade do classificador seja independente dessa dimensão e que seja apenas função do número de vectores de suporte o que se adequa à redução estrutural do risco (SRM).

2.6.2 Hiperplano separador óptimo generalizado para a não separabilidade linear

O desenvolvimento apresentado na secção anterior é apenas funcional para o caso de os dados serem linearmente separáveis. Todavia, normalmente, não é essa a situação que surge, Figura 2.8. Existem abordagens que permitem a generalização, mas que dependem de alguma parametrização inicial sobre o problema e um conhecimento do ruído existente no conjunto de dados.

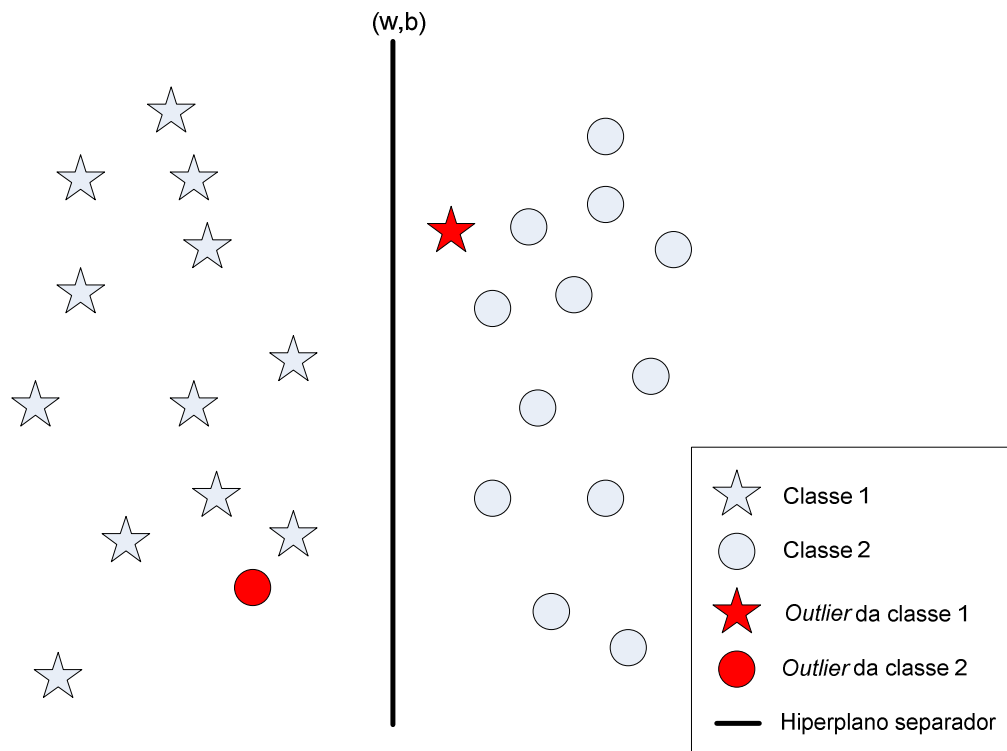


Figura 2.8 Hiperplano separador (w,b) num espaço linearmente não separável

Para que o método da separabilidade de hiperplanos possa ser generalizado, Cortes e Vapnik [16] introduziram um conjunto de variáveis, $\xi_i \geq 0$, e uma função de custo,

$$F_{\sigma}(\xi) = \sum_i \xi_i^{\sigma} \quad \sigma > 0, \quad (2.34)$$

onde as variáveis ξ_i são os erros relativos à classificação falhada dos dados de treino de índice i . O tipo de função está relacionado com o índice σ . O problema de otimização

é agora apresentado em duas vertentes: a minimização do erro de classificação e a minimização dos limites na dimensão VC do classificador. Daí que as restrições que eram apresentadas na forma da Equação (2.12) sejam alteradas para,

$$y^i [\langle w, x^i \rangle + b] \geq 1 - \xi_i, \quad i = 1, \dots, l \quad . \quad (2.35)$$

Em que $\xi_i \geq 0$. A generalização do hiperplano separador óptimo é obtida através do vector w , que minimiza o funcional,

$$\Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i. \quad (2.36)$$

C é um valor sujeito às restrições induzidas pela Equação (2.35). Este parâmetro pode ser entendido como um parâmetro de regularização; [17]. A solução para o problema de optimização associado à Equação (2.36), que está também sujeita à Equação (2.35), é dada pelo ponto estacionário de inflexão do Lagrangiano [15],

$$\Phi(w, b, \alpha, \xi, \beta) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_{i=1}^l \alpha^i (y^i [\langle w, x^i \rangle + b] - 1 + \xi_i) - \sum_{j=1}^l \beta_j \xi_j, \quad (2.37)$$

onde α e β são multiplicadores de Lagrange. O Lagrangiano terá que ser minimizado relativamente a w , b e ξ e terá que ser maximizado relativamente a α e a β . Tal como se fez na Equação (2.20), a Equação (2.37) terá que ser construída num formato de optimização dualizado,

$$\max_{\alpha, \beta} W(\alpha, \beta) = \max_{\alpha, \beta} \left(\min_{w, b, \xi} \Phi(w, b, \alpha, \xi, \beta) \right). \quad (2.38)$$

O mínimo do Lagrangiano, Φ , em função de w , b , ξ é obtido pelas derivadas parciais,

$$\begin{aligned}
 \frac{\partial \Phi}{\partial b} = 0 &\Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \\
 \frac{\partial \Phi}{\partial w} = 0 &\Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i . \\
 \frac{\partial \Phi}{\partial \xi} = 0 &\Rightarrow \alpha_i + \beta_i = C
 \end{aligned} \tag{2.39}$$

Das equações (2.37), (2.38) e da (2.39) a dualidade de optimização fica escrita como,

$$\max_{\alpha} W(\alpha) = \max_{\alpha} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle + \sum_{k=1}^l \alpha_k . \tag{2.40}$$

Então a equação de optimização é dada por,

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle - \sum_{k=1}^l \alpha_k . \tag{2.41}$$

Sujeita às condições,

$$\begin{aligned}
 0 \leq \alpha_i \leq C \quad i=1, \dots, l \\
 \sum_{j=1}^l \alpha_j y_j = 0
 \end{aligned} . \tag{2.42}$$

A solução apresentada para o caso da não separabilidade dos dados de forma linear é semelhante à apresentada para quando os dados são garantidamente separáveis linearmente. A única excepção reside na modificação dos limites dos coeficientes de *Lagrange*. Porém, a solução de Cortes [16] exige que o parâmetro C tenha que ser determinado, normalmente, empiricamente. Blanz et al. [18] usaram um valor, para C , de 5. A escolha do valor de C para ser feita correctamente deve reflectir a existência de ruído nos dados.

2.6.3 Generalização num espaço de dimensão elevada

Nas situações em que a separação do espaço de entrada, por um hiperplano, não é adequada, devido à não separabilidade linear dos dados, é possível mapear os elementos do vector de entrada num espaço de características de dimensão mais elevada. Desta forma, a SVM estabelece um hiperplano separador óptimo, mas no espaço de características, Figura 2.9. Este espaço pode permitir tratar o problema, originalmente não linearmente separável, como um problema linearmente separável, Figura 2.10.

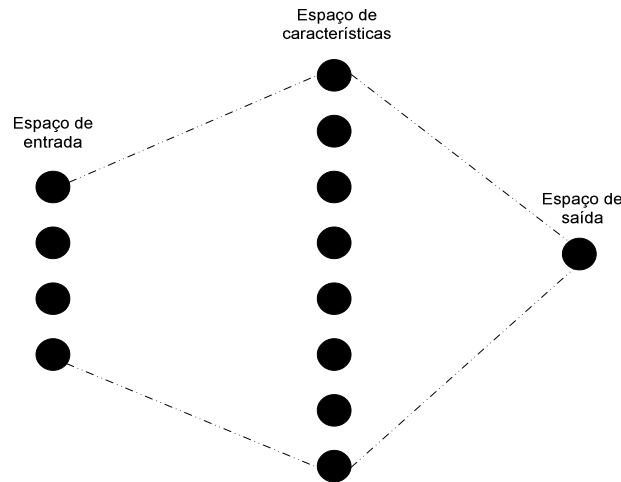


Figura 2.9 Mapeamento do espaço de entrada num espaço de maiores dimensões

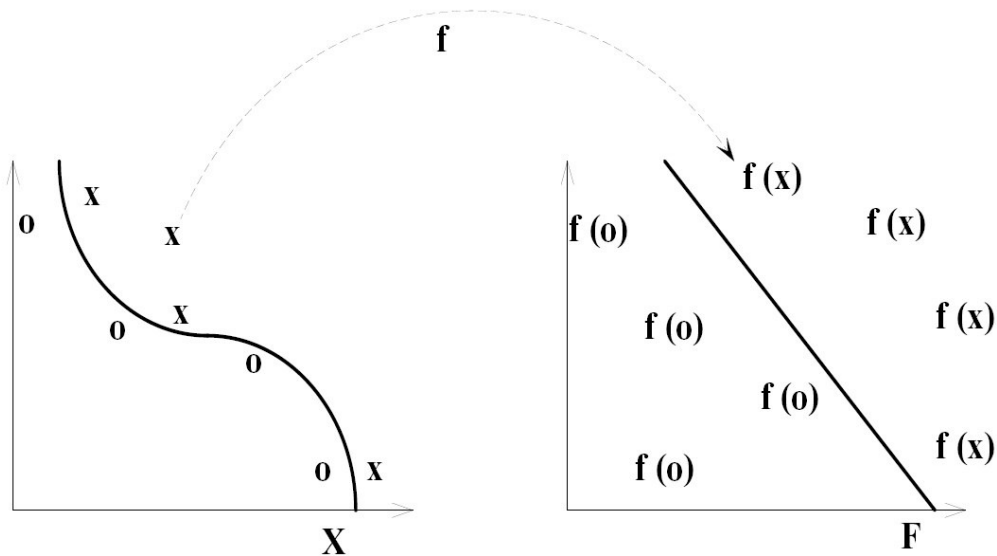


Figura 2.10 Conversão, por mapeamento de uma função f , de um problema não separável linearmente num problema separável linearmente

Existem algumas restrições para se obter um mapeamento não linear através de funções. Porém, existe um grande número de funções que podem ser utilizadas nesse

processo. Funções como as polinomiais, funções de base radial e funções sigmóide permitem constituir mapeamentos do espaço de entrada para o espaço de características. O problema de optimização descrito pela Equação (2.41) é alterado como se apresenta na equação seguinte:

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{k=1}^l \alpha_k, \quad (2.43)$$

onde $K(x_i, x_j)$ representa uma função de *kernel*, a qual executa o mapeamento não linear no espaço de características, enquanto que as restrições se mantêm inalteradas:

$$\begin{aligned} 0 \leq \alpha_i \leq C \quad i = 1, \dots, l \\ \sum_{j=1}^l \alpha_j y_j = 0 \end{aligned} \quad (2.44)$$

Resolvendo a Equação (2.43), sujeita às condições da Equação (2.44), determinam-se os coeficientes de Lagrange. Pode então obter-se um classificador que funciona a partir da separação, por um hiperplano, do espaço de características e que é escrito da seguinte forma:

$$f(x) = \text{sgn} \left(\sum_{i \in SV_s} \alpha_i y_i K(x_i, x) + b \right). \quad (2.45)$$

Em que,

$$\langle w^*, x \rangle = \sum_{i=1}^l \alpha_i y_i K(x_i, x), \quad (2.46)$$

$$b^* = -\frac{1}{2} \sum_{i=1}^l \alpha_i y_i [K(x_i, x_r) + K(x_i, x_s)]. \quad (2.47)$$

Aqui, b , que representa o deslocamento do plano à origem, é calculado através de dois vectores de suporte. No entanto, seria possível fazê-lo usando todos os vectores

de suporte [19]. Se a função de *kernel* detiver um termo de desvio (*bias*), então b poderá ser integrado pelo *kernel* e assim o classificador pode ser reduzido a:

$$f(x) = \text{sgn} \left(\sum_{i \in SV_s} \alpha_i y_i K(x_i, x) \right). \quad (2.48)$$

Muitos *kernels* têm um termo de desvio e qualquer *kernel* finito pode ser estruturado para possuir um [20]. Isto simplifica o problema de otimização ao ser possível retirar a condição que era dada pela igualdade da Equação (2.44). Porém, continua a ser necessário cumprir alguns requisitos que tornam adequada a utilização de um determinado *kernel* no mapeamento (ver Secção 2.8.1).

2.7 O algoritmo Kernel-Adatron

O método de otimização referido nas secções anteriores é bastante robusto no que concerne aos problemas de *over-fitting*. Porém, a sua implementação computacional não é elementar e os custos computacionais que lhe estão agregados podem ser elevados. Adicionalmente, existem problemas de índole numérica que em certas circunstâncias apresentam condicionalismos difíceis de solucionar (*ill-Conditioned* da matrix *hessiana* do problema de otimização⁹). A aplicação do algoritmo de aprendizagem *Adatron* a métodos que têm como base de funcionamento *kernels* vem diminuir os mencionados problemas que aparecem associados às otimizações de programação quadrática. O *Adatron* é um algoritmo com funcionamento paulatino (on-line) e que é usado, na aprendizagem, com estruturas de *perceptrons*. A sua adaptação às SVMs resulta no seguinte algoritmo:

1. Inicializar $\alpha_i = 1$;

⁹ Genericamente um problema de otimização quadrática pode ser representado por,

$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \alpha^T H \alpha + c^T \alpha$, onde H representa a matriz hessiana e que da análise da Equação

(2.23) corresponde a $y_i y_j \langle x_i, x_j \rangle$.

2. Partindo do par de padrões de índice $i = 1$, sobre os pares x_i, y_i , calcular

$$z_i = \sum_{j=1}^l \alpha_j y_j K(x_i, x_j);$$

3. Para cada um dos pares de padrões i calcular $\gamma_i = y_i z_i$ e executar os passos 4, 5 e 6;

4. Calcular a possível modificação dos multiplicadores α_i de acordo com:

$$\delta\alpha^i = \eta(1 - \gamma^i), \text{ sendo } \eta \text{ um factor de crescimento};$$

5. Se $(\alpha^i + \delta\alpha^i) \leq 0$, então a actualização proposta neste passo resultará num valor inferior a zero. Para evitar este problema faz-se $\alpha^i = 0$;

6. Se $(\alpha^i + \delta\alpha^i) > 0$ então os multiplicadores são actualizados segundo a atribuição $\alpha^i \leftarrow \alpha^i + \delta\alpha^i$;

7. Calcular o termo de desvio b ,

$$b = \frac{1}{2}(\min(z_i^+) + \max(z_i^-)),$$

em que z_i^+ são os padrões da classe indicada por +1 e em que z_i^- são os padrões da classe indicada por -1;

8. Se o número máximo de apresentações do conjunto de padrões for ultrapassado, então parar; se não, voltar ao ponto 2.

No final o classificador será obtido pela seguinte função:

$$f(x) = \text{sgn}\left(\sum_j \alpha_j y_j K(x_j, x) + b\right). \quad (2.49)$$

O *kernel* $K(x_i, x_j)$ pode ser um qualquer que satisfaça as condições de Mercer (ver Secção 2.8.1). Por exemplo, podem ser usados *kernels* polinomiais ou funções de base radial (RBF).

2.8 Espaço de características

O objectivo de um *kernel* é permitir que as operações de classificação sejam executadas sobre o espaço de entrada em vez de o serem no espaço de elevada dimensionalidade, o espaço de características. Dessa forma o produto interno não necessita de ser calculado no espaço de características. Isto evita que o processo fique associado aos problemas do aumento de dimensionalidade. Apesar disso, o método possui algumas dependências funcionais críticas que se relacionam com o número de padrões de treino, já que, de forma a garantir uma favorável distribuição dos dados num problema dimensionalmente grande é necessário, normalmente, um conjunto extenso de padrões de treino. Por outro lado, embora as SVMs implementem o princípio do SRM, o que lhes confere uma boa capacidade de generalização, deve existir algum cuidado na escolha do *kernel*. Um *kernel* que execute mapeamentos para espaços de dimensões relativamente grandes tenderá a produzir SVMs com dificuldades de generalização.

2.8.1 Funções de *kernel*

Tendo como base as teorias (RKHS) – *Reproducing Kernel Hilbert Spaces* de [20-23], um produto interno, de funções de mapeamento ϕ , no espaço de características, tem um *kernel* equivalente no espaço de entrada (Equação (2.50)),

$$K(x, x') = \langle \phi(x), \phi(x') \rangle, \quad (2.50)$$

desde que sejam cumpridas algumas condições que permitem manter a convexidade do problema de optimização quadrática. Se K é uma função definida positiva e simétrica, então satisfaz as condições de Mercer, tal que,

$$\iint K(x, x')g(x)g(x')dxdx' > 0, \quad \int g^2(x_i)dx_i < \infty, \quad (2.51)$$

onde g representa uma qualquer função do espaço L_2 . Se K cumpre a Equação (2.51), então existe uma base ortogonal $\{\phi_i\}_i$, em $L_2[a, b]$, formada por funções próprias tal que a correspondente sequência de valores próprios $\{\lambda_i\}_i$ é não negativa. As funções

próprias com valores próprios diferentes de zero são contínuas em $[a,b]$ e K tem a seguinte representação (teorema de Mercer):

$$K(x, x') = \sum_m^{\infty} \lambda_m \phi_m(x) \phi_m(x'), \quad (2.52)$$

sendo a convergência absoluta e uniforme.

Desta forma é garantido que o *kernel* representa um produto interno no espaço de características. Algumas funções que satisfazem as condições de Mercer são discutidas nas secções a seguir (2.8.1.2 a 2.8.1.8).

2.8.1.1 Normalização dos dados

Alguns *kernels*, por restrição de domínio, obrigam a que sejam normalizados os dados de entrada, mas podem surgir também vantagens mesmo nos casos em que o domínio dos *kernels* é suficientemente amplo face aos dados. Por exemplo, a normalização dos dados reduz alguns dos problemas que derivam dos condicionalismos numéricos da matriz *hessiana*.

2.8.1.2 Polinomial

Um mapeador polinomial é definido pelas Equações (2.53) ou (2.54) e é comumente usado em modelos não lineares.

$$K(x, x') = \langle x, x' \rangle^d, \quad (2.53)$$

$$K(x, x') = (\langle x, x' \rangle + 1)^d. \quad (2.54)$$

Para se evitar que a *hessiana*, que corresponde a $y_i y_j \langle x_i, x_j \rangle$ na Equação (2.23), se anule, o *kernel* apresentado na Equação (2.54) é normalmente preferido ao *kernel* (2.53). O valor de d relaciona-se com a dimensão do espaço de características.

2.8.1.3 Função Gaussiana de base radial (RBF)

A função Gaussiana de base radial é dada pela seguinte equação:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right). \quad (2.55)$$

Métodos clássicos que utilizam funções de base radial servem-se de alguma técnica para determinar os centros dos subconjuntos. Normalmente é usado um método de *clustering* para esse fim. Um das vantagens da utilização nas SVMs é que essa tarefa é implícita, uma vez que cada vector de suporte contribui para uma função Gaussiana local em que o centro é o ponto definido pelo padrão de treino. No caso do algoritmo *Kernel-Adatron* é vulgar usar-se esta função de *kernel*.

2.8.1.4 Função exponencial de base radial

Uma função radial com a forma,

$$K(x, x') = \exp\left(-\frac{\|x - x'\|}{2\sigma^2}\right), \quad (2.56)$$

produz uma solução descontínua, o que pode ser vantajoso se para um dado problema as discontinuidades forem aceitáveis.

2.8.1.5 Perceptron multicamada

Uma função deste tipo, Equação (2.57), com uma única camada escondida, para alguns valores de ρ e de ζ , também é uma representação válida para um *kernel*,

$$K(x, x') = \tanh(\rho \langle x, x' \rangle + \zeta). \quad (2.57)$$

Os vectores de suporte correspondem aos pesos da camada escondida e os multiplicadores de Lagrange correspondem aos *pesos* da camada de saída.

2.8.1.6 Séries de Fourier

Uma série de Fourier pode ser considerada uma expansão dimensional de $2N+1$ no espaço de características. O *kernel* é definido para o intervalo $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$,

$$K(x, x') = \frac{\sin\left(N + \frac{1}{2}\right)(x - x')}{\sin\left(\frac{1}{2}(x - x')\right)}. \quad (2.58)$$

A capacidade de regularização para este *kernel* não é elevada, não sendo, geralmente, um kernel com um bom desempenho [17].

2.8.1.7 Kernel aditivo

Podem-se construir *kernels* a partir de um somatório de *kernels*, Equação (2.59). Esta possibilidade é verdadeira já que a soma de duas funções definidas positivas é definida positiva.

$$K(x, x') = \sum_i K_i(x, x'). \quad (2.59)$$

2.8.1.8 Produto tensorial

Kernels multidimensionais podem ser obtidos a partir de produtos tensoriais de *kernels* [23]:

$$K(x, x') = \prod_i K_i(x_i, x'_i). \quad (2.60)$$

2.8.2 Escolha de um Kernel

A selecção do *kernel* mais adequado a um determinado problema, geralmente, não é determinística. Embora métodos determinísticos possam ser integrados para obter esta selecção, estes obrigam a que seja estimado o raio da hipersfera que envolve os dados no espaço de características não linear. Assim, a melhor forma de o fazer é incluir

vários *kernels* na plataforma de simulação de SVMs e compará-los em termos dos seus desempenhos para um determinado problema. Esses testes devem ser suportados por abordagens como a da validação cruzada, avaliando-se assim a capacidade de generalização de cada *kernel* para o problema em causa.

2.9 Classificação de múltiplas classes

Uma SVM é um classificador de duas classes apenas. Porém, no domínio prático os problemas que surgem geralmente obrigam a uma classificação de mais do que duas classes. As soluções existentes abordam a associação de várias SVMs para a solução deste problema. Num sistema de K classes, são necessárias K SVMs. A k -ésima SVM com saída $y_k(x)$ é treinada usando os dados da classe C_k como exemplos de padrões positivos e os exemplos restantes das $K-1$ classes são usados como exemplos de padrões negativos. A obtenção da resposta global é dada por:

$$y(x) = \max_k y_k(x). \quad (2.61)$$

Contudo, cada uma das SVMs é treinada num domínio diferente e isso impossibilita a correcta comparação entre os seus valores já que a escala de uma SVM poderá ser diferente da escala das outras SVMs. Outro problema pode resultar do não balanceamento dos grupos de padrões positivos e negativos entre as SVMs, o que fragiliza o processo de comparação de valores. Estes problemas contribuem para a descida do desempenho funcional das SVMs em aplicações de classificação com mais do que duas classes.

2.10 Considerações finais

Os aspectos teóricos em que assentam as SVMs vêm minimizar vários problemas que outras índoles de classificadores detêm. Isso, geralmente, é constatado na produção de melhores resultados sobre as aplicações em que são inseridas. Porém, em

algumas situações os condicionalismos de implementação computacional podem ser problemáticos.

As características principais que as SVMs apresentam são:

- a retenção em mínimos locais não acontece (o mínimo é global);
- tentam suster alguns dos problemas derivados do aumento da dimensionalidade (*curse of dimensionality*);
- a ordem de execução computacional é directamente dependente do número de padrões de treino;
- a aprendizagem tende a ser independente da irregularidade do espaço do problema;
- a informação de decisão está contida nos vectores de suporte (SV);
- está implícita a minimização do risco estrutural (SRM);
- podem surgir problemas aquando da utilização em mais do que duas classes.

Capítulo 3

Descritores e Seleccionadores de Características

O descritor e/ou seleccionador de características constitui um importante componente nos sistemas de reconhecimento de padrões. Da correcta escolha e do desempenho funcional destes mecanismos depende a simplificação e o bom desempenho dos órgãos posteriores como, por exemplo, os classificadores. Neste capítulo são estudados e testados dois descritores numéricos de dados, especializados na descrição de regiões. São eles: os descritores de Hu e os descritores de Zernike. Os testes feitos visam o caso de estudo em questão nesta tese, o reconhecimento de leucócitos.

3.1 Introdução

Num sistema de reconhecimento de padrões, normalmente, existem duas grandes entidades funcionais. A primeira, que pode ser genericamente entendida como um descritor ou um seleccionador, tem como função a de tratar os dados de entrada para que a segunda, o classificador, mantenha a sua tarefa o mais simples possível e assim proporcionar uma eficácia de resultados elevada.

A manipulação dos dados de entrada, por acção de um descritor, leva a que os dados sejam apresentados ao classificador, presumivelmente, da forma mais adequada e a que as representações que esses dados sugerem possam ser reduzidas a espaços que à partida são suficientemente discriminadores, e não redundantes, dos dados de entrada

segundo a partição de classificação a obter. Estes processos assentam o seu funcionamento, de redução do espaço de entrada, análise e combinação das variáveis de entrada, em algum tipo de conhecimento prévio e na relação de características produzidas pelas variáveis. Por outro lado, os seleccionadores também pretendem diminuir a dimensão do espaço de entrada, mas eliminando determinados dados de entrada que não são discriminantes nem agregadores, de uma forma coerente sobre o universo dos dados de entrada em relação às partições de classificação. No limite e no que diz respeito ao resultado que produzem, ambos os processos, de descritores e de seleccionadores, podem não ser dissociáveis e as técnicas tratadas neste capítulo, apesar de serem designadas de descritores, apresentam inerentemente uma componente de seleccionadoras.

3.2 Descritores de regiões

Uma aplicação dos momentos clássicos a imagens bidimensionais foi apresentada por Hu [24] nos anos 60. Hu testou a sua validade sobre um problema de reconhecimento de caracteres. Os testes que executou não abrangiam imagens com ruído, mas outros estudos [25] demonstraram que a consideração de imagens com ruído diminuía intensamente o desempenho desses descritores. Já os momentos de Zernike têm-se revelado mais robustos na representação de objectos em imagens. Estes testes foram realizados em [26].

3.2.1 Momentos cartesianos não ortogonais

Este tipo de descritores quando aplicado a imagens descreve o conteúdo destas relativamente aos seus eixos. Aplicando as transformações adequadas às funções de base que definem os momentos, é possível obter descritores que são invariáveis à translação, rotação e escala. A aplicação de momentos estatísticos à análise de imagens digitais usando a base cartesiana [24], sendo P a imagem, resulta em,

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q P_{xy}, \quad p, q = 0, 1, 2, \dots \quad (3.1)$$

onde p e q representam a ordem do momento, x e y representam a posição dos pixels da imagem e onde M e N representam as dimensões da imagem.

A modificação que conduz à invariância à translação, produzindo momentos centralizados, é dada por:

$$\begin{aligned}\mu_{pq} &= \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q P_{xy}, \quad p, q = 0, 1, 2, \dots \\ \bar{x} &= m_{10} / m_{00}, \\ \bar{y} &= m_{01} / m_{00},\end{aligned}\tag{3.2}$$

onde M e N são as dimensões da imagem. \bar{x} e \bar{y} definem o centro de massa provocado por todos os *pixels* em x e todos os pixels em y respectivamente. P_{xy} corresponde à intensidade de cada *pixel* xy .

A normalização seguinte permite obter momentos centralizados que são invariantes à escala:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \quad \gamma = \frac{(p+q+d)}{d}.\tag{3.3}$$

Para imagens binárias a variável d terá o valor 2 e para imagens em níveis de cinzento a variável d terá o valor 3.

A independência à rotação destes momentos pode ser obtida a partir da Equação(3.3), através de os momentos de Hu.

3.2.2 Momentos de Hu

Baseado nos momentos centralizados normalizados, Equação (3.3), Hu introduziu sete funções não lineares que permitem a independência à translação, à escala e à rotação. São elas:

$$\phi_1 = \eta_{20} + \eta_{02}\tag{3.4}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2\tag{3.5}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3.6)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \quad (3.7)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + \\ & (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned} \quad (3.8)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11} (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (3.9)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2 \right] + \\ & (3\eta_{21} - \eta_{30})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned} \quad (3.10)$$

3.2.3 Momentos de Zernike ortogonais complexos

Os momentos ortogonais de Zernike [27] são construídos através de um conjunto de polinómios complexos, os quais constituem uma base ortogonal completa definida no disco unitário, $x^2 + y^2 \leq 1$. A seguir é apresentada a expressão que os define:

$$A_{mn} = \frac{m+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x,y) [V_{mn}(x,y)]^* dx dy, \quad (3.11)$$

onde $m = 0, 1, 2, \dots, \infty$ define a ordem do momento de Zernike. $f(x,y)$ é a função de que se pretende obter uma descrição. O símbolo * representa um valor complexo conjugado e n um valor inteiro, positivo ou negativo, que representa a dependência angular, sob as seguintes restrições:

$$m - |n| = \text{valor par}, \quad |n| \leq m. \quad (3.12)$$

Os polinómios de Zernike [28] apresentados em coordenadas polares são dados por,

$$V_{mn}(r, \theta) = R_{mn}(r) \exp(jn\theta), \quad (3.13)$$

onde (r, θ) está definido sobre o disco unitário e $R_{mn}(r)$ é o polinómio radial ortogonal definido como,

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s F(m, n, s, r), \quad (3.14)$$

onde,

$$F(m, n, s, r) = \frac{(m-s)!}{s! \left(\frac{m+|n|}{2} - s\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s}. \quad (3.15)$$

Os primeiros seis polinómios radiais ortogonais são a seguir apresentados:

$$\begin{aligned} R_{00}(r) &= 1 \\ R_{11}(r) &= r \\ R_{20}(r) &= 2r^2 - 1 \\ R_{22}(r) &= r^2 \\ R_{31}(r) &= 3r^3 - 2r \\ R_{33}(r) &= r^3 \end{aligned} \quad (3.16)$$

Para o cálculo dos momentos de Zernike em imagens representadas por $P(x, y)$, a Equação (3.11) adquire a seguinte forma,

$$A_{mn} = \frac{m+1}{\pi} \sum_x \sum_y P(x, y) [V_{mn}(x, y)]^*, \quad x^2 + y^2 \leq 1. \quad (3.17)$$

No cálculo dos momentos de Zernike, as imagens ou as regiões de interesse são mapeadas, em primeiro lugar, para o disco unitário através de coordenadas polares. Os

pixels que se situarem fora do disco unitário não são considerados nos cálculos. As coordenadas polares são descritas por r e por θ . O r corresponde ao comprimento do vector estabelecido pelo ponto a considerar e pela origem. O θ corresponde ao ângulo que o mesmo vector faz com o eixo dos xx . Por convenção, o ângulo é obtido a partir do eixo positivo x na direcção contrária aos ponteiros do relógio. A transformação de coordenadas polares para coordenadas cartesianas é determinada por:

$$x = r \cos \theta \quad y = r \sin \theta. \quad (3.18)$$

Em que,

$$r = \sqrt{x^2 + y^2} \quad \theta = \tan^{-1} \left(\frac{y}{x} \right). \quad (3.19)$$

A independência à translação e à escala pode ser conseguida pela utilização de momentos cartesianos previamente ao cálculo dos momentos de Zernike. A invariância à translação é obtida deslocando a origem para o centro de massa da imagem usando centróide de momento e fazendo com que $m_{01} = m_{10} = 0$. Com abordagem semelhante é obtida a invariância à escala. O conteúdo da imagem é alterado para que a sua área corresponda a uma constante, $m_{00} = \beta$. Esta constante é definida previamente, de acordo com um valor de área de referência. A transformação que caracteriza estas duas normalizações pode ser conseguida usando:

$$g(x, y) = f \left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y} \right), \quad \text{onde } a = \sqrt{\frac{\beta}{m_{00}}}, \quad (3.20)$$

onde $g(x, y)$ é a nova função representativa da imagem original, agora normalizada à posição e à escala. O erro que está associado à discretização da imagem pode ser reduzido pela utilização de técnicas de interpolação.

Por exemplo, se as novas coordenadas não coincidirem de forma exacta com as coordenadas da grelha de discretização actual, o valor do pixel em questão pode ser interpolado pelos quatro pixels adjacentes.

Como resultado da normalização, os momentos de Zernike $|A_{00}|$ e $|A_{11}|$ adquirem valores bem determinados. Devido à translação do centróide da figura para o centro do sistema de coordenadas, $|A_{11}|$ adquire o valor zero. No entanto, os efeitos da discretização fazem com que esse valor apenas possa ser assumido para conteúdos de imagem de maior escala, já que os erros de discretização são menores à medida que a imagem do objecto em representação aumenta. A dependência de $|A_{00}|$ sobre m_{00} implica que esse seja função de β :

$$|A_{00}| = \frac{\beta}{\pi}. \quad (3.21)$$

O *módulo* de um momento de Zernike é invariante à rotação. Este facto é um reflexo do mapeamento da imagem no disco unitário. A rotação do objecto sobre o disco unitário é representada por uma alteração do valor de fase. Sendo ϕ o ângulo de rotação, sendo A_{mn}^R um momento de Zernike de uma imagem rodada e A_{mn} um momento de Zernike da imagem original, então:

$$A_{mn}^R = A_{mn} \exp(-jn\phi). \quad (3.22)$$

3.3 Comparação entre momentos de Zernike e de Hu sobre imagens naturais

Para fazer a comparação funcional entre os dois descritores foram usadas imagens naturais de leucócitos. O processo consistiu em entregar a cada tipo de descritor as mesmas células, que já tinham sido previamente segmentadas na imagem capturada por um microscópio. O classificador usado foi uma SVM e a sequência de blocos para um e para o outro caso ficou assim organizado, Figura 3.1:

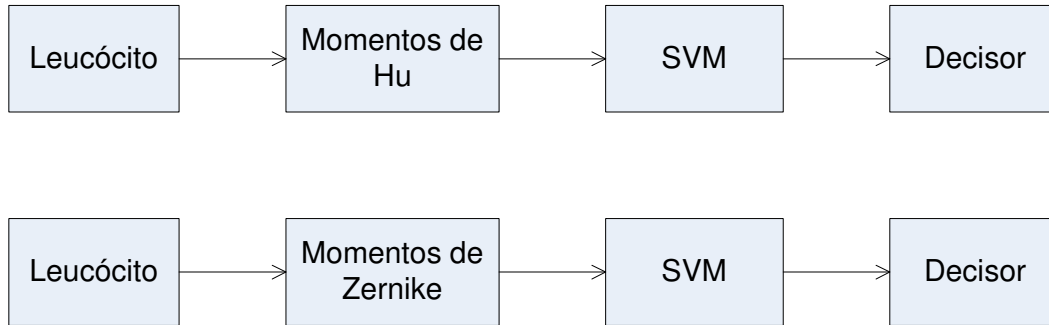


Figura 3.1 Sequência de blocos funcionais usados na comparação dos momentos de Hu com os momentos de Zernike

Foram usados os seis primeiros momentos de Hu e os seis primeiros momentos de Zernike como entradas do classificador SVM.

Os resultados foram obtidos adicionando ruído gaussiano às imagens que tinham sido usadas na fase treino. Na Tabela 3.I são indicadas as percentagens de classes de leucócitos correctamente decididas para três níveis de relação sinal-ruído:

Tabela 3.I Percentagens de acertos conseguidos para os descritores de Hu e de Zernike na classificação de leucócitos na presença de três níveis de relação sinal-ruído

| SNR | Hu | Zernike |
|------|-----|---------|
| 30dB | 72% | 75% |
| 25dB | 68% | 73% |
| 17dB | 51% | 58% |

Da observação feita à Tabela 3.I conclui-se sobre o pior comportamento dos descritores de Hu face aos descritores de Zernike. Tal facto é manifestamente agravado com a descida da relação sinal ruído. Esta observação leva a que a maior robustez dos descritores de Zernike seja também assumida em imagens naturais. Embora estes resultados possam apenas ser válidos para este tipo de imagens naturais, as imagens de leucócitos, parecem ser um indicador da robustez destes descritores. No entanto, é possível que em outro tipo de imagens naturais, e até na utilização de um outro classificador, a relação de comportamento entre os dois descritores seja favorável aos descritores de Hu.

Ao longo do trabalho tornou-se também claro que nas duas técnicas as ordens mais elevadas dos descritores são mais representativas das maiores frequências espaciais existentes nas imagens. Deste facto resulta uma maior sensibilidade ao ruído por parte dos descritores de ordens mais elevadas.

Na vertente de desempenho, embora não se tenha quantificado, os descritores de Zernike, pela análise das equações que os implementam, são claramente mais exigentes, computacionalmente, do que os descritores de Hu.

Capítulo 4

Comportamento de Métodos de Segmentação não Supervisionados

Neste capítulo são estudados alguns métodos de segmentação não supervisionada com o objectivo de se perceberem as debilidades e o potencial desses métodos e de como podem ser aproveitados na adição a um método supervisionado. É mostrada a dificuldade de parametrizar correctamente o mecanismo não supervisionado para obter resultados positivos e em simultâneo é realçada a indispensabilidade de a parametrização ser, de alguma forma, variável em função das mutabilidades, em alguns casos pequenas, do contexto visual. É sugerida a interligação de um processo de segmentação multi-escala a um processo quântico para atenuar as dificuldades de execução temporal de um sistema de parametrização exaustiva.

4.1 Introdução

A expressão *segmentação de imagens* está associada à partição de uma imagem num conjunto de regiões Ω_i , onde Ω representa o domínio da imagem, não sobrepostas e homogéneas, de acordo com algum critério semântico existente sobre esses fragmentos de forma relacional ou não.

$$\bigcup_{i=1}^n \Omega_i = \Omega, \quad \Omega_i \cap \Omega_j = \emptyset, \quad \text{se } i \neq j \quad (4.1)$$

Num sentido mais abrangente, as regiões podem ser consideradas como um aglomerado de *pixels* que definem uma linha, como é o caso de um segmento de recta, ou então, mais vulgarmente, como um conjunto de *pixels* que definem uma forma com o seu respectivo contorno, como é o caso de um círculo. Um dos objectivos da segmentação é o de decompor a imagem em partes, para que apenas as regiões de interesse sejam processadas na posterior fase de análise. Um outro objectivo é o de modificar convenientemente a representação adjacente à agregação original dos *pixels*, para que o processo de análise seja menos complexo e mais eficaz em todas as suas vertentes. O grande desafio está em encontrar um método de segmentação que seja suficientemente global, isto é, que não se construa apenas à base de conhecimento prévio de um determinado problema; mas que, sem ter o conhecimento prévio específico, possa tratar, com desempenhos elevados, quaisquer outras situações de segmentação que sejam apresentadas. A obtenção deste equilíbrio está associada a um custo. Esse custo aumenta à medida que o equilíbrio aumenta, pois relaciona-se com a complexidade do sistema. A complexidade existe principalmente na forma estrutural com que a solução ao problema é abordada. Aqui surgem essencialmente dois aspectos, um ligado às limitações impostas pelos dispositivos de suporte físico, que actualmente se têm à disposição, e o outro ligado às dificuldades de perceber ou encontrar mecanismos funcionais que atinjam o mencionado equilíbrio.

4.2 Métodos de agrupamento

Os métodos de agrupamento produzem subgrupos de dados (*pixels*) que são agregados segundo critérios de semelhança.

4.2.1 Algoritmos de agregação clássicos

Em processamento de imagens a avaliação de similaridade de grupos pode ser feita sobre o valor das intensidades (luminância) dos *pixels*, sobre os canais de cor (RGB), sobre propriedades geométricas que os *pixels* constroem, sobre a textura que os *pixels* representam, entre outras.

No método clássico de agregação existem K grupos C_1, C_2, \dots, C_k com médias m_1, m_2, \dots, m_k . A função de medida conhecida por *erro dos mínimos quadrados* pode ser definida como,

$$D = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - m_k\|^2, \quad (4.2)$$

e mede o quanto os dados (vectores) estão próximos dos actuais grupos. A construção dos grupos através deste método pode ser executada considerando todos os arranjos possíveis de agregações dos dados aos grupos, de forma a minimizar D . Porém, o método, por não ser determinístico, é computacionalmente impróprio. Os métodos usados na determinação dos grupos são métodos de aproximação à melhor solução, não garantem a minimização de D , mas são válidos no que respeita à sua ordem de execução temporal. A exigência de conhecimento à partida, ou não, do número de grupos é função do tipo de algoritmo usado.

Um dos métodos mais utilizado é o agrupamento iterativo por *K-médias* cujo algoritmo é o seguinte:

1. Iniciar o contador j (contador de iterações) com o valor 1;
2. Escolher aleatoriamente os elementos que formam os K conjuntos com médias $m_1(1), m_2(1), \dots, m_k(1)$;
3. Para cada vector x_i calcular $D(x_i, m_k(j))$ para cada $k = 1, \dots, K$ e atribuir x_i ao grupo, C_i , com a média que lhe está mais próxima;
4. Incrementar o contador j uma unidade e calcular os novos valores das médias $m_1(j), m_2(j), \dots, m_k(j)$;
5. Repetir os passos 3 e 4 até que $C_k(j) = C_k(j+1)$ para todos os k .

4.2.2 Agrupamento pelo método ISODATA¹⁰

O algoritmo das K-médias é intensamente dependente dos parâmetros de inicialização já que o número K é escolhido no início e mantido até ao fim do algoritmo. O algoritmo ISODATA [29] permite tornar variável o número de grupos reagrupando grupos redundantes e desagrupando grupos inadequados.

Assume-se inicialmente, também neste método, K grupos C_1, C_2, \dots, C_k com médias m_1, m_2, \dots, m_k e estabelece-se uma matriz de co-variância Σ_k para cada grupo k . Sabendo que os dados x_i são vectores na forma,

$$x_i = [v_1, v_2, \dots, v_n], \quad (4.3)$$

então cada média m_k é um vector,

$$m_k = [m_{1k}, m_{2k}, \dots, m_{nk}], \quad (4.4)$$

e Σ_k ,

$$\Sigma_k = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_{nn} \end{bmatrix}, \quad (4.5)$$

onde $\sigma_{ii} = \sigma_i^2$ é a variância do i -ésimo elemento v_i dos vectores do grupo k e onde $\sigma_{ij} = \rho_{ij} \sigma_i \sigma_j$ é a co-variância entre o i -ésimo elemento e o j -ésimo elemento dos vectores. O ρ_{ij} corresponde ao coeficiente de correlação entre o i -ésimo elemento e o j -ésimo elemento, σ_i representa o desvio padrão do i -ésimo elemento e σ_j corresponde ao desvio padrão do j -ésimo elemento.

¹⁰ O termo ISODATA deriva de Iterative Self-Organizing Data Analysis Techniques.

O Algoritmo de agregação por ISODATA é a seguir apresentado:

1. Atribuir x_i ao grupo l que minimiza,

$$\left| [x_i - m_l]^T \Sigma_l^{-1} [x_i - m_l] \right|$$

2. Agregar os grupos i e j se,

$$|m_i - m_j| < \tau_v$$

Onde τ_v corresponde a um valor limiar de variância.

3. Dividir o grupo k se o máximo valor próprio de Σ_k é superior a τ_v .
4. Terminar quando,

$$|m_i(t) - m_i(t+1)| < \varepsilon,$$

para todos os grupos i ou terminar quando um número máximo de iterações for atingido. O valor ε define o limite superior de não convergência sobre o centro dos grupos (médias).

4.2.3 Métodos de agregação por histograma

Os métodos referidos nas secções anteriores, quando aplicados ao processamento de imagens, podem ser de um desempenho temporal reduzido. As formas iterativas, recombinadoras dos grupos, acarretam um número elevado de operações sobre o conjunto de dados. Os métodos que têm como base o histograma recorrem a uma única passagem operacional pelos dados. Estes métodos funcionam se as imagens detiverem uma estatística de distribuição de pixels que esteja de acordo com a importância e a semântica dos objectos a segmentar. Caso contrário, os resultados podem variar intensamente.

Este processo assume que os segmentos de interesse existentes na imagem manifestam-se eles próprios como grupos no espaço de agregação, ou seja, no histograma. A segmentação resulta do mapeamento reverso, do histograma para o domínio da imagem onde as componentes máximas (modas) do histograma assinalam os segmentos. A determinação das modas pode ser feita pela análise da variação do histograma, por exemplo, detectando os vales do histograma e supondo que os segmentos se encontram na zona de união desses vales. Um pixel que se encontre no i -

ésimo intervalo ganhará a marca do índice i . Todos os pixels situados dentro desse intervalo serão assinalados com o mesmo índice, o que corresponderá a um grupo de pixels de intensidades semelhantes. A determinação das modas de um histograma multimodal pode ser obtida pelo processo de Otsu [30], o qual procura encontrar os vários valores limiares de separação modal pela minimização da variância interna de cada uma das classes (modas). Estas são definidas, no histograma, pelos respectivos intervalos de valores limiares.

4.2.4 Crescimento de Região (Region Growing)

Ao contrário dos métodos anteriores que seguem um mecanismo de divisão, este método tem início num determinado *pixel* da imagem e procura aumentar, por agregação, essa região até que os *pixels* que estão a ser comparados, para posterior agregação, deixem de ser semelhantes à região em causa. A comparação pode ser executada através de um teste estatístico. Um método de crescimento de região foi apresentado por Haralick [31] e assume que uma região é formada por um conjunto de *pixels* adjacentes que configuram uma determinada média dentro de uma determinada variância.

Assuma-se R como uma região formada por N *pixels* constituídos na matriz I e indexados por r e por c . Considere-se também y como a intensidade de um *pixel* adjacente a essa região. O método é formulado através da média, \bar{x} , das intensidades dos *pixels* na região e pela dispersão, S^2 , desses mesmos *pixels* relativamente à média:

$$\bar{X} = \frac{1}{N} \sum_{[r,c]} I[r,c], \quad (4.6)$$

$$S^2 = \sum_{[r,c]} (I[r,c] - \bar{X})^2. \quad (4.7)$$

Então, Haralick assume que todos os *pixels* da região e o *pixel* y a testar são independentes e que seguem uma distribuição Gaussiana para deduzir a seguinte equação estatística:

$$T = \left[\frac{(N-1)N}{N+1} (y - \bar{X})^2 / S^2 \right]^{\frac{1}{2}}. \quad (4.8)$$

Se T for suficientemente pequeno, y é adicionado à região R , a média e a dispersão são actualizadas usando o valor de y :

$$\bar{X}_{actual} \leftarrow \frac{(N\bar{X}_{anterior} + y)}{N+1}, \quad (4.9)$$

$$S_{actual}^2 \leftarrow S_{anterior}^2 + (y - \bar{X}_{actual})^2 + N(\bar{X}_{actual} - \bar{X}_{anterior})^2. \quad (4.10)$$

Se T é demasiadamente alto, o valor de y não deverá pertencer à população de pixels da região R . Se y é diferente de todos os seus vizinhos, então ele mesmo irá iniciar uma nova região.

4.3 Segmentação Multi-escala

Os vários métodos de agregação vistos atrás têm por base de agregação uma estatística que pode ser mais ou menos local ou mais ou menos global. Porém, na maior parte das imagens naturais, a correcta agregação de sub-elementos segue uma hierarquia semântica que parte das características mais básicas das regiões e evolui até às mais elevadas. Os métodos atrás descritos, na sua constituição básica, não contemplam explicitamente esta relação hierárquica. Esta lacuna faz com que surja, de caso para caso, uma diversidade de erros de agregação.

O sistema visual humano contém mecanismos indicadores de realizar um processamento estruturado em hierarquia [32-33]. A estrutura observada possui características próximas de uma representação multi-escala [32].

Uma representação multi-escala pode ser conseguida através de um espaço de escalas [34]. Nessa representação é considerada uma pilha de elementos que vão progredindo da imagem original até às sucessivas versões de escalas cada vez maiores relativamente à imagem original. À medida que a escala aumenta, diminui a informação localizada da imagem original e, por oposição, prevalece a informação global. O

processo relaciona os elementos de mais alta resolução de forma a obter os elementos com menor resolução, mas com informação mais global. O operador de escala a aplicar pode ser de natureza diversa [35], mas existem dois principais grupos: os operadores de espaços de escala lineares e os operadores de espaços de escala não lineares. Os lineares derivam da equação da transferência de calor devido às características de multi-resolução desta [33]. A principal propriedade dos operadores lineares é que executam a difusão de escala para escala sem considerarem particularidades do sinal. Os não lineares, apesar de manterem o fundamento base da difusão do sinal de escala para escala, obedecem também a algumas restrições que se relacionam com o sinal em si. Por exemplo, podem propagar a difusão entre escalas mediante a existência ou não de um determinado gradiente no sinal [36].

Como os métodos não lineares possuem restrições que são escolhidas mediante o tipo de imagens a processar, nas implementações deste método aqui testadas, preferiu-se manter o processo generalizante, optando-se pelo método linear. Como será descrito, pode ser associada uma fase de supervisão para se executar a tarefa de especialização. A implementação do método aqui usado, de índole linear, baseia-se em [37].

A partir da imagem a processar $I(\vec{x})$, $\vec{x} = (x, y)$, é obtida uma pilha de imagens $I(\vec{x}, t)$, cada uma à sua escala t , que obedece à seguinte equação [33]:

$$\frac{\partial I(\vec{x}, t)}{\partial t} = \Delta I(\vec{x}, t), \quad (4.11)$$

em que a condição inicial é: $I(\vec{x}, t = 0) = I(\vec{x})$, sendo o parâmetro de escala t monótono crescente. Pode-se obter $I(\vec{x}, t)$ através da operação de convolução. O único *kernel* que satisfaz a Equação (4.11) é a função Gaussiana [33-34], [38]. Daí resulta:

$$I(\vec{x}, t) = \int_D G(\vec{x} - \vec{x}', t) \cdot I(\vec{x}', 0) d\vec{x}', \quad (4.12)$$

sendo,

$$G(\vec{x}, t) = \frac{1}{4\pi t} e^{-\left(\frac{\|\vec{x}\|^2}{4t}\right)}. \quad (4.13)$$

De facto, a obtenção das diferentes escalas a partir da imagem original é conseguida por um conjunto de operações de média.

O processo de segmentação multi-escala é obtido através da análise da estrutura de pilha cujos respectivos elementos se estabelecem a escalas diferentes, Figura 4.1.

O princípio de relacionamento entre os elementos da pilha baseia-se no seguimento de pixels com intensidades próximas entre escalas adjacentes [39-41]. O algoritmo constrói uma estrutura em árvore que descreve a referida relação entre os elementos de escalas adjacentes, Figura 4.2.

Um elemento de escala n apenas tem pixels conectados aos pixels do nível $n+1$ se, e só se, esses pixels de n tiverem ligações a pixels do nível $n-1$. A determinação da ligação dos pixels filhos aos pixels pais é feita mediante a utilização de uma região de procura, situada sobre os pixels pais, Figura 4.3. A região de procura é centrada sobre o pixel filho e abrange um conjunto de pixels pais, num raio que é proporcional ao nível de escala t [39]. Todos os pixels sitiados nessa área são candidatos a pixel pai. A determinação do pixel pai é feita avaliando as diferenças de valor de pixel filho e de potenciais pixels pais. Isto é executado entre os dois níveis de escala adjacentes e em causa. O pixel pai efectivo é aquele que tiver uma menor diferença para o pixel filho dentro dessa sub-região de procura. Com o aumento da escala existirá uma maior convergência de ligações entre a camada inferior e a camada imediatamente seguinte. Isto é um efeito directo da operação de suavização, usada na obtenção das diferentes escalas. A segmentação dada por um determinado nível resulta do conjunto de ligações que se estabelecem, no sentido descendente, a partir de um dado nó desse nível. Esta determinação é executada no sentido inverso ao da construção da árvore de ligações, dos nós pais para os nós filhos. A escolha dos nós de uma determinada escala estabelece os segmentos agregados finais, resultantes de todo o processo de segmentação multi-escala.

O problema que surge nesta fase é o de escolher a escala correcta onde se encontram esses nós finais. O tamanho da região de pesquisa também contribui para o resultado final. Uma possibilidade que pode tornar automática a escolha da escala máxima e do raio de pesquisa é a de considerar uma gama de escalas e uma gama de valores de raio, escolhidas empiricamente, de forma a que o resultado da combinação desses parâmetros seja avaliado por um sistema supervisor. Entre as várias segmentações geradas, a partir de parametrização diversa, é escolhida a segmentação

mais promissora de entre as avaliadas. A avaliação pode passar por submeter cada um dos segmentos obtidos por cada uma dessas combinações de parâmetros a um processo de determinação do seu valor de coincidência com um *kernel*, ou vários *kernels*, que formulam a região tipo que se espera segmentar. O nível de escala que obtiver maior magnitude de coincidência é o escolhido para retornar os segmentos de interesse. Em vez de *kernels*, o processo poderia ser melhorado pela utilização de um bloco de descritores de Zernike seguido de uma SVM. Assim, a SVM ou o sistema clássico de *kernels* classificariam os segmentos obtidos, e por consequência, reflectiriam nessa classificação a qualidade, a avaliação, da segmentação. O problema seria agora perspectivado em quem, se existir, avalia o resultado da avaliação feita pelos mecanismos supervisionados (*kernel* ou SVM). Construindo uma hierarquia de avaliadores, o sistema pode melhorar, mas os suportes actuais não permitem a execução em tempo útil a um sistema com uma estrutura dessa magnitude. No capítulo desta tese sobre aceleração quântica, na secção 6.11.3, é sugerida uma técnica quântica a aplicar a este processo, a qual minimiza o problema de execução temporal.

O algoritmo apresentado a seguir permite obter e avaliar um segmento de uma determinada região de suposto interesse. O algoritmo considera o sistema de *kernels* clássico (*matching*) para fazer a avaliação dos segmentos. Do passo 1 ao passo 4 o algoritmo foca o seu processamento na obtenção de segmentações sobre as diferentes escalas. Do passo 5 ao passo 9 é feita a análise das segmentações obtidas para se produzirem as correspondentes classificações do que são essas segmentações. Se as classificações tiverem um nível de coincidência baixo, são então rejeitadas, não produzindo qualquer segmentação e classificação finais.

Realce-se a cooperação bidireccional da segmentação-classificação para o resultado final. A classificação é executada sobre um conjunto de segmentações resultantes das várias escalas analisadas. Só as classificações dos segmentos que atingem maior magnitude de coincidência, relativamente ao procurado, é que são consideradas. Isto permite rejeitar as segmentações presentes em escalas que não têm sentido geométrico para a semântica do problema em causa.

Algoritmo:

1. Escolher uma combinação de parametrização a usar (escala máxima; raio da região de pesquisa).

2. Para se obter um novo nível de escala, suavizar a imagem (nível actual) com a operação de convolução subjacente às equações (4.12) (4.13).
3. Imagem original: Para cada pixel da imagem original procurar o pixel pai de intensidade mais próxima dentro da área de pesquisa do nível de escala superior. Níveis de escala superior: Para cada pixel de um determinado nível n , com ligação a pixels filhos, procurar o pixel pai de intensidade mais próxima dentro da área de pesquisa do nível de escala superior.
4. Repetir 2 e 3 até ao nível em que apenas é estabelecido um único nó ou até a um nível preestabelecido de escala máxima.
5. Percorrer a árvore das raízes (nós superiores) para as folhas, de acordo com cada nível de escala, de forma a obter a segmentação das imagens para cada valor de escala distinto. Fazê-lo apenas para as escalas com potencial de correcta segmentação. (Usar um sistema de *trial-error* prévio para essa determinação).
6. Para cada segmento obtido fazer a convolução com um ou mais *kernels* de segmentos de interesse para obter o respectivo grau de coincidência.
7. Relacionar cada *kernel* com o nível de escala que maior coincidência devolveu.
8. Reter apenas os segmentos pertencentes às escalas que implicaram maior grau de coincidência.
9. Classificar os segmentos de acordo com o grau de coincidência e com o tipo de *kernel*.

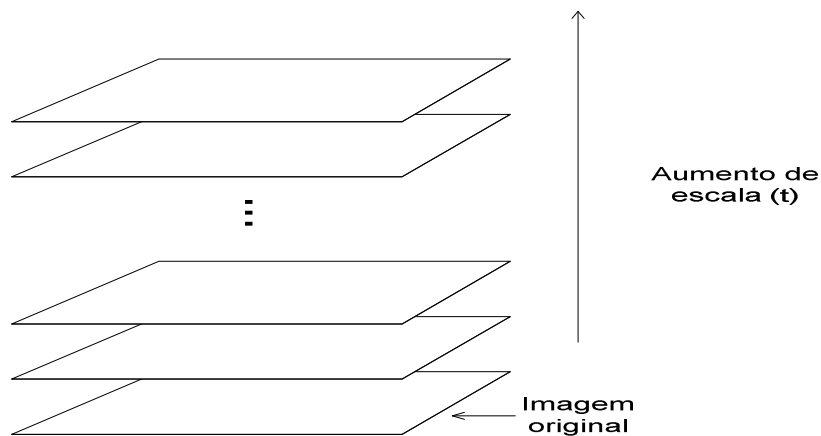


Figura 4.1 Pilha de imagens a diferentes escalas da imagem original

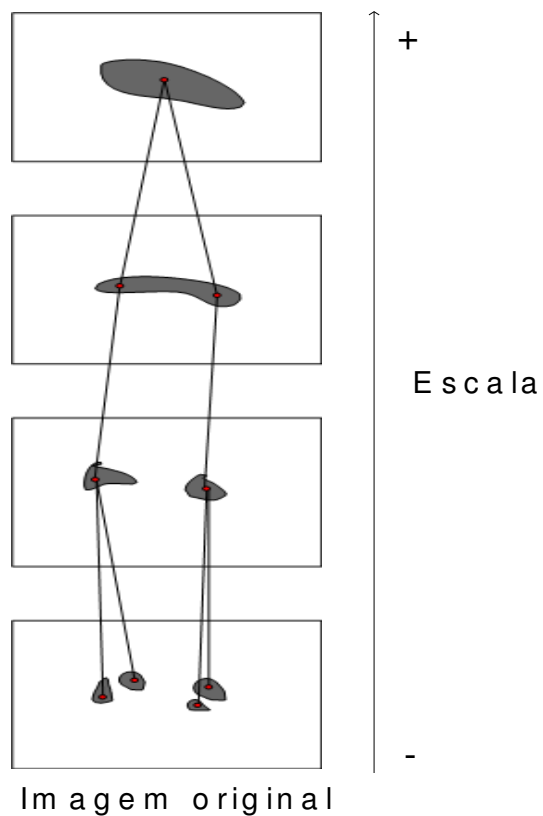


Figura 4.2 Hierarquia de ligações entre os vários níveis de escala

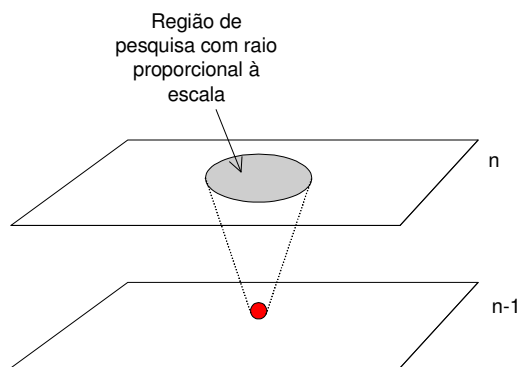


Figura 4.3 Área de procura de um pixel pai

4.4 Resultados

Algumas das imagens utilizadas para a obtenção dos comportamentos de três dos métodos de segmentação estudados são mostradas no formato agregado de componentes na Figura 4.4 e são mostradas nas respectivas componentes parciais (HSL) da Figura 4.5 à Figura 4.7.

O primeiro grupo de imagens, Figura 4.8 à Figura 4.15, pretende exibir o comportamento, evidenciando as dificuldades que os processos de segmentação baseados em agregação por histograma apresentam. Com o mesmo objectivo são apresentados resultados no segundo grupo de imagens, da Figura 4.16 à Figura 4.19, para o método de agregação por crescimento de região.

O terceiro grupo de imagens, Figura 4.20 à Figura 4.23, mostra os resultados que a técnica de segmentação multi-escala atinge sobre o mesmo tipo de imagens.

Com estes grupos de imagens pretende-se mostrar o comportamento evidenciando as dificuldades que os processos de segmentação não dirigida detêm.

São referidos exemplos que realçam a dificuldade de escolha da parametrização através da comparação dos resultados para as diversas parametrizações.

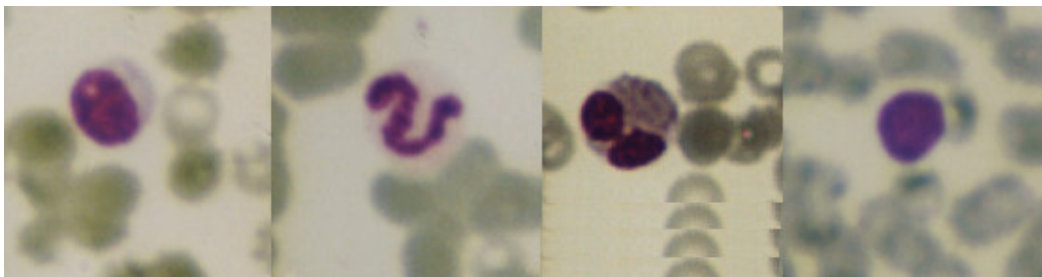


Figura 4.4 Imagem original de quatro células

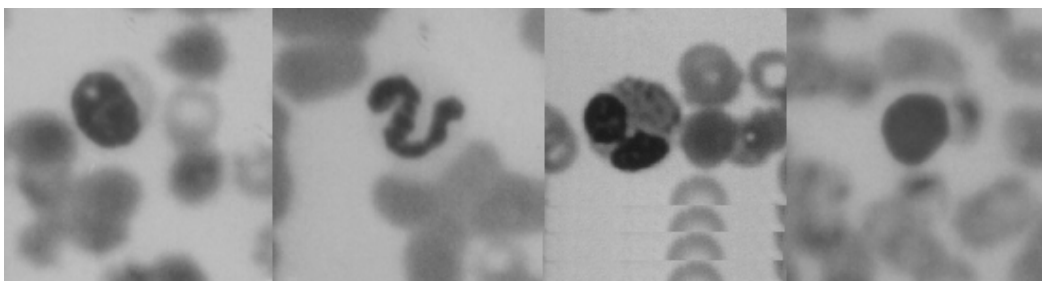


Figura 4.5 Componente de intensidade (L) da imagem original de quatro células

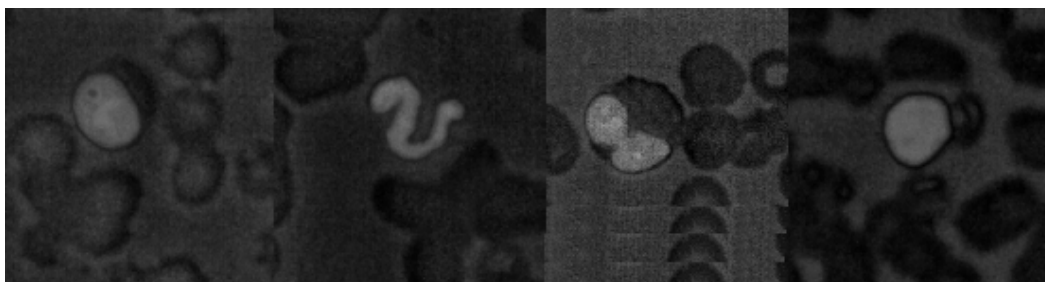


Figura 4.6 Componente de saturação (S) da imagem original de quatro células

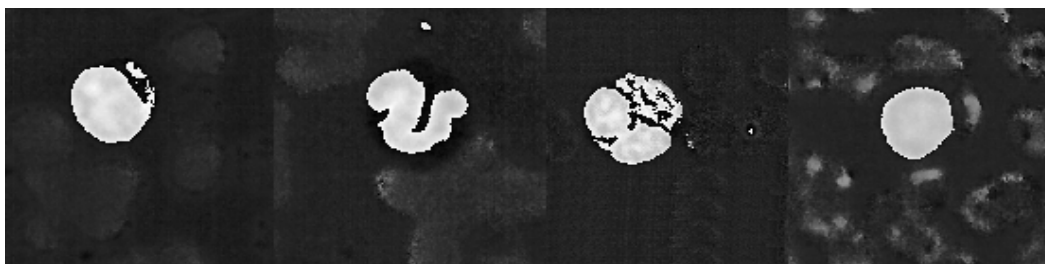


Figura 4.7 Componente de cor (H) da imagem original de quatro células

Da extração dos três componentes HSL é possível, neste caso de aplicação, observar que há uma tendência para que os leucócitos sejam assinalados pela componente de matiz (H). Para uma maior percepção da transformação resultante do processamento, segundo as várias técnicas consideradas, as imagens em processamento são confinadas a esta componente (com a excepção dos métodos de histograma), já que esta componente, por si mesma e generalizando, gera resultados próximos à detecção correcta dos leucócitos. Porém, qualquer um dos planos poderia ser utilizado de forma independente para a análise dessas técnicas. Numa perspectiva de obtenção de um sistema com um desempenho maior, os três planos necessitariam de ser combinados. Essa combinação pode ser feita antes do processamento, adaptando as técnicas ao tratamento cooperativo das três componentes, ou pela fusão dos resultados individuais do processamento. Esta última abordagem poderá ser mais difícil de coadunar deterministicamente com o objectivo a atingir, enquanto que a primeira abordagem é de mais natural adaptação pela reformulação das técnicas em causa em direcção à nova dimensão vectorial.

Relativamente aos resultados do processo de agregação por histograma, são apresentadas segmentações que utilizaram apenas a componente de matiz H e também são mostradas segmentações que utilizaram as três componentes HSL. A agregação das

três componentes foi feita distribuindo os bits mais significativos de cada uma por uma única *word* de 16 bits e dando mais significância ao H, depois ao S e depois ao L.

Nas imagens da Figura 4.8 e da Figura 4.9 é evidente que o método de agregação por histograma depende, no que concerne aos resultados de segmentação, da estatística intrínseca de pixels da imagem a processar. A segunda célula à esquerda foi diferenciada relativamente ao seu núcleo e ao seu citoplasma, em oposição à célula situada mais à direita que acaba por apresentar um erro na sua segmentação. Esta situação piora se o número de classes de agregação for escolhido como sendo seis. Como se deduz, vários parâmetros iniciais do processo têm que ser escolhidos em função do tipo de conteúdo a processar. O número de classes de agregação e os elementos do histograma, que funcionam em associação às classes, determinam fundamentalmente o sucesso final do processo. Neste processamento foram considerados, sempre, os n valores do histograma com maior ocorrência para as n classes de agregação correspondentes. A pretensão de se obter um método totalmente adaptativo parece ser inatingível. A adaptação parte sempre de uma base, de uma série de considerações que parecem nunca estar completas.

De forma semelhante, os resultados obtidos com processo de crescimento de região projectam-se na mesma natureza de problemas. Comparando as imagens da Figura 4.16 e da Figura 4.17 com as imagens da Figura 4.18 e da Figura 4.19, são visíveis as variações de resultados com as variações de parâmetros. De modo diferente, a comparação entre a imagem da Figura 4.16 e a imagem da Figura 4.17 e ainda a comparação entre a imagem da Figura 4.18 com a imagem da Figura 4.19 revelam as variações de resultados relativamente à alteração do conteúdo (outro fundo; outras células vizinhas) adjacente da imagem. Qualquer das situações, também neste método, mostra a necessidade de proceder a reajustes de parametrização mediante novas imagens e mediante as alterações de estatística intrínseca às novas imagens. Não significa isto que para se obterem resultados universais seja suficiente a correcta adaptação dos parâmetros. Estas técnicas são dirigidas à agregação orientada ao pixel. Mas para se conseguirem resultados mais universais, são necessárias, dependendo também do problema, outras abordagens que diversifiquem a base de agregação ou de crescimento. Por exemplo, para decidir sobre o crescimento da região pode ter-se como base informação, também, de textura ou de outros aspectos de semântica geométrica mais elevada. Num formato de teste à agregação poder-se-ia averiguar: o pixel a agregar

à região em crescimento é circundado por outros pixels que formam uma determinada textura ou padrão? A escala verificada dessa textura é a esperada? Na verdade, critérios em número aparentemente indeterminável poderiam ser considerados na determinação desse crescimento ou agregação.

O método da segmentação multi-escala, verificando-se os resultados através da Figura 4.20 até à Figura 4.23, é o mais consistente a produzir segmentos. Porém, continua a ser dependente da parametrização usada relativamente às imagens a processar. O método da segmentação multi-escala é dos três aqui testados o mais exigente computacionalmente. Isto parece criar uma relação clara entre o aumento de desempenho funcional e a complexidade dos algoritmos, apesar de parecer que essa relação não é linear mas sim exponencial.

De facto, de uma análise empírica, parece haver uma necessidade exponencial da ordem de execução computacional para produzir um aumento de desempenho linear de um processo de visão deste tipo. Mais ainda, empiricamente parece que a curva de aproximação ao desempenho visual humano é de natureza logarítmica. A dificuldade expande-se em duas grandes áreas, que se inter-relacionam: a necessidade de suportes computacionais poderosos e a necessidade de algoritmos mais completos e complexos.

A extracção da informação segmentada para ser usada num sistema classificador pode ter vários formatos. Relativamente ao caso de estudo, uma solução é a de classificar as várias classes de sub-regiões resultantes da agregação, assinaladas pelo sistema de segmentação. Nessa classificação, os segmentos reconhecidos como pertencentes a uma mesma classe de célula são reavaliados pelo seu posicionamento relativo. Se a proximidade se verifica, por exemplo, entre um suposto núcleo e um suposto citoplasma, a célula é então considerada. O problema do método é que é vastamente dependente dos segmentos que foram utilizados durante a fase de constituição do classificador, neste caso de estudo, de células. Isto reduz a eficácia funcional do sistema.

Imagens dos resultados por agregação com histograma:

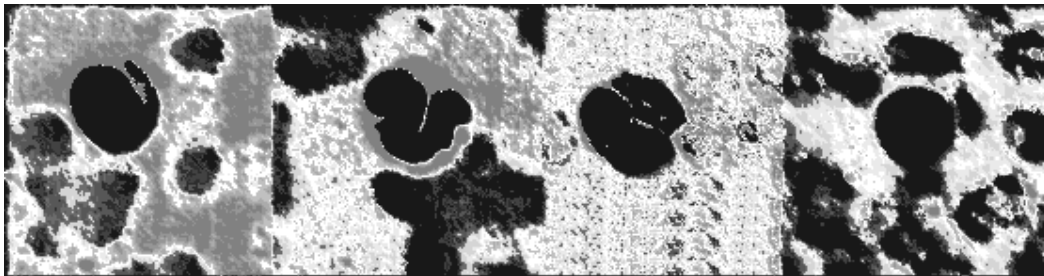


Figura 4.8 Processamento sobre a componente de cor (H), por agregação baseada em histograma para doze classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea

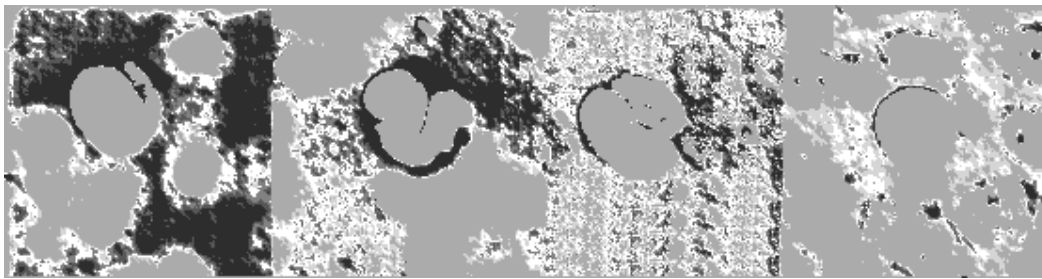


Figura 4.9 Processamento sobre a componente de cor (H), por agregação baseada em histograma para seis classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea

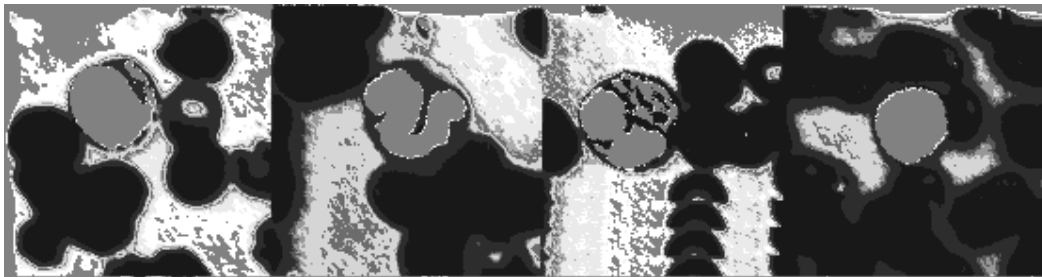


Figura 4.10 Processamento combinado dos três componentes (HSL), por agregação baseada em histograma para doze classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea

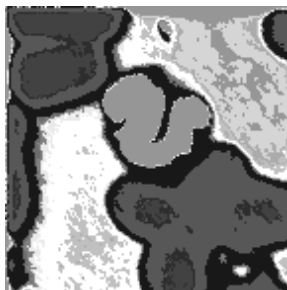


Figura 4.11 Resultado usando o mesmo processamento feito à imagem anterior, mas com a segunda célula mais à esquerda, dessa imagem, e seus adjacentes, isolados

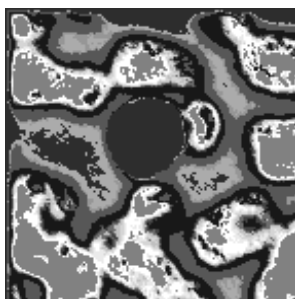


Figura 4.12 Resultado usando o mesmo processamento feito à imagem da Figura 4.10, mas com a célula mais à direita, dessa imagem, e seus adjacentes, isolados

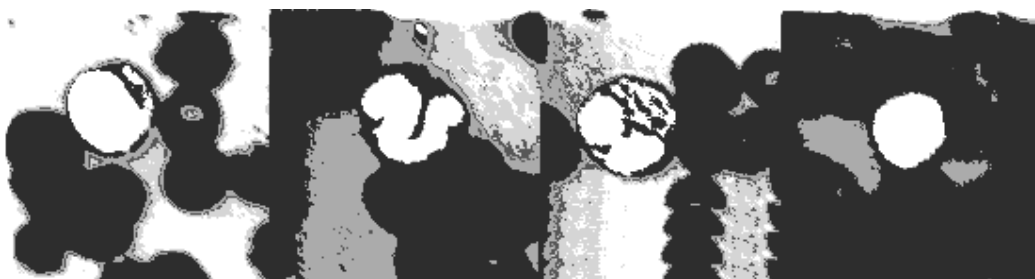


Figura 4.13 Processamento combinado dos três componentes (HSL), por agregação baseada em histograma para seis classes, a quatro imagens de células todas concatenadas na mesma imagem e todas adquiridas da mesma amostra sanguínea

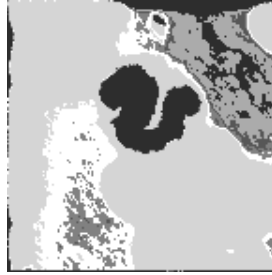


Figura 4.14 Resultado usando o mesmo processamento feito à imagem anterior, mas com a segunda célula mais à esquerda, dessa imagem, e seus adjacentes, isolados

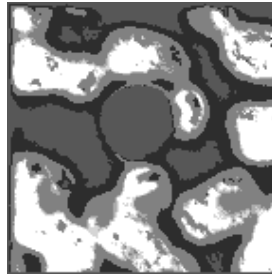


Figura 4.15 Resultado usando o mesmo processamento feito à imagem da Figura 4.13, mas com a célula mais à direita, dessa imagem, e seus adjacentes, isolados

Imagens dos resultados por agregação em região de crescimento:

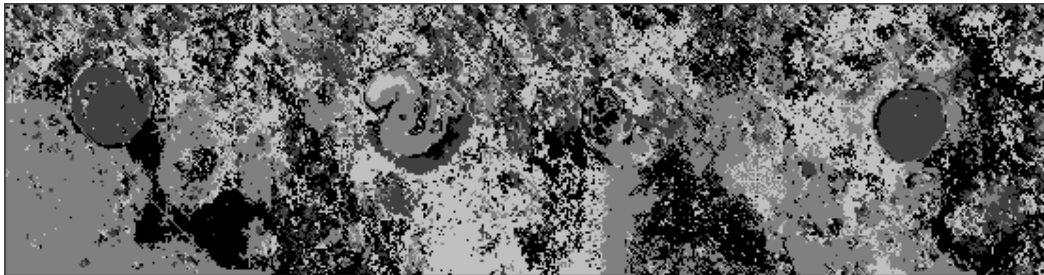


Figura 4.16 Segmentação produzida pelo método de crescimento de região, sobre a componente de cor da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a 1 e que compara com o valor da Equação (4.8)

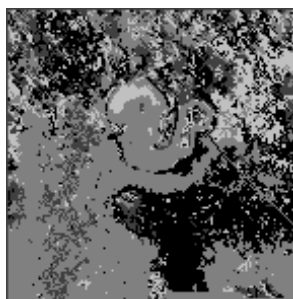


Figura 4.17 Segmentação produzida pelo método de crescimento de região, sobre a componente de cor, da segunda célula à esquerda, da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a um e que compara com o valor da Equação (4.8)

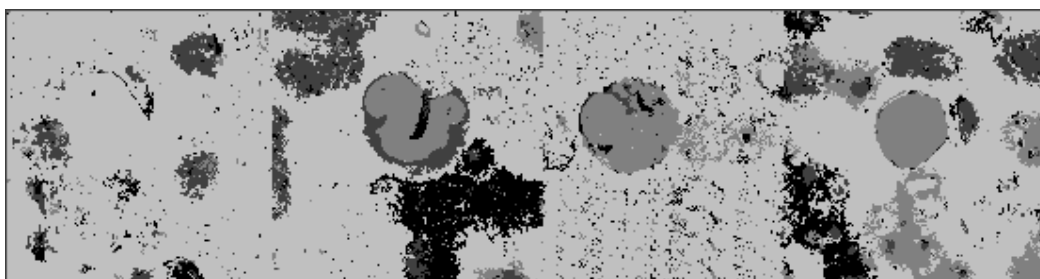


Figura 4.18 Segmentação produzida pelo método de crescimento de região, sobre a componente de cor da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a 2 e que compara com o valor da Equação (4.8)

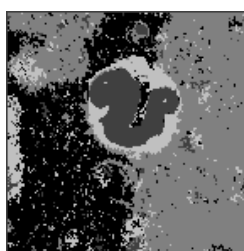


Figura 4.19 Segmentação produzida pelo método de crescimento de região, sobre a componente de cor, da segunda célula à esquerda, da imagem da Figura 4.4, com o valor do limiar de continuidade de crescimento igual a 2 e que compara com o valor da Equação (4.8)

Imagens dos resultados por segmentação multi-escala:

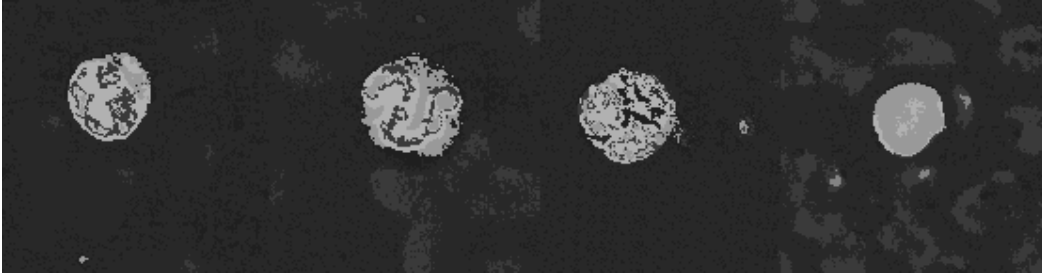


Figura 4.20 Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor da imagem da Figura 4.4. A parametrização usada empregou o valor 5 para o raio do círculo de procura e o valor 9 para o valor do nível de escala mais elevado

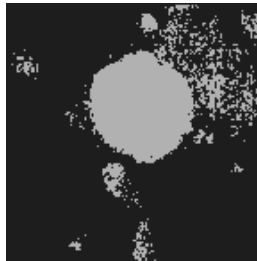


Figura 4.21 Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor, da segunda célula mais à esquerda, isolada, da imagem da Figura 4.4. A parametrização usada no processo de segmentação foi a mesma que originou o resultado visível na imagem anterior

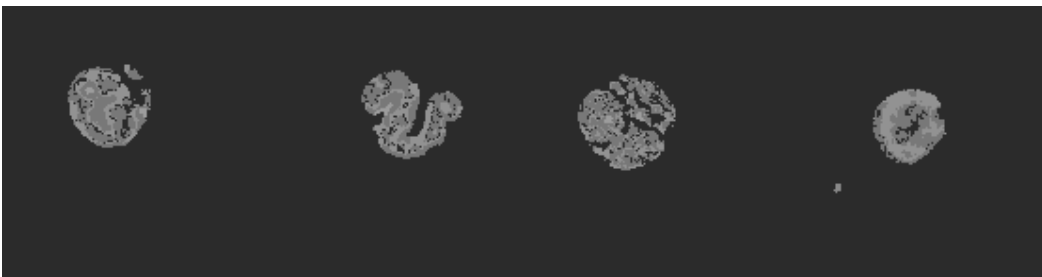


Figura 4.22 Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor da imagem da Figura 4.4. A parametrização usada empregou o valor 2 para o raio do círculo de procura e o valor 17 para o valor do nível de escala mais elevado

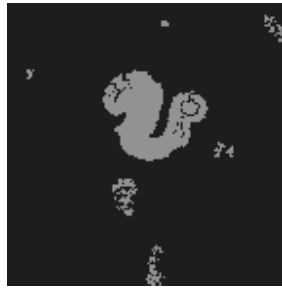


Figura 4.23 Segmentação produzida pelo método de segmentação multi-escala, sobre a componente de cor, da segunda célula mais à esquerda, isolada, da imagem da Figura 4.4. A parametrização usada no processo de segmentação foi a mesma que originou o resultado visível na imagem anterior

4.5 Discussões finais

Dos resultados apresentados atrás é essencialmente deduzida a necessidade de um mecanismo supervisor que detecte quais as segmentações que são de interesse para a resolução de um determinado problema de visão. A utilidade da segmentação obtida depende intensamente da parametrização empregada. A variabilidade das imagens naturais, bem como os seus múltiplos aspectos geométricos conduz à impossibilidade de se poder ter uma parametrização constante e universal. Mecanismos de supervisionamento, como o que foi descrito, podem ajudar a minimizar o problema. Mas, como foi referido, a necessidade de parametrizar os próprios mecanismos de supervisionamento leva a que não se encontre um fim na hierarquia de supervisão. De qualquer forma, o aumento da densidade dessa supervisão tende a produzir melhores soluções. A utilização de tecnologias promissoras, como a computação quântica, pode converter soluções deste tipo, computacionalmente muito exigentes, em soluções praticáveis. Podem mesmo permitir que a consistência da procura da solução seja de tal forma aumentada que a melhor solução seja muitas vezes alcançada. Porém, mesmo um suporte com essas capacidades não resolve o problema que está associado, quase sempre, a situações de imagens naturais; a grande dificuldade continua a ser como saber que aquele resultado é efectivamente o melhor? A multiplicidade de métricas adoptadas para cobrir melhor esse fim pode não ser suficiente, mesmo em suportes quânticos.

Capítulo 5

Segmentação e Classificação com Redes Neurais de Acoplamento por Pulsos

As redes neuronais de acoplamento por pulsos constituem um tipo de rede menos comum, mas que, por se basearem no funcionamento pulsante dos neurónios do córtex visual, possuem algumas características adequadas ao processamento de imagem. Neste capítulo, estas redes são utilizadas em conjunto com SVMs e descritores de Zernike no processo de segmentação e de classificação de leucócitos. São produzidas comparações com métodos que usam redes neuronais clássicas. São também sugeridas modificações a estas redes para funcionarem em domínios, não de regiões, mas de contornos.

5.1 Introdução

A utilização de redes neuronais clássicas em problemas de visão artificial tem sido multiplamente documentada e demonstrada através de muitos trabalhos científicos que têm sido desenvolvidos ao longo das últimas décadas. Essas redes neuronais podem apresentar certas dificuldades no seu uso, essencialmente, na vertente de treino. É, de facto, notada uma grande dependência do cumprimento do objectivo da rede ao universo de padrões de treino que lhe são fornecidos. Variabilidade de todo o tipo de características ópticas, nas imagens, que não se coadune com as imagens de treino, proporcionará um decréscimo do êxito de processamento dessas redes. As redes neuronais de acoplamento por pulsos não surgem como uma alternativa às redes

clássicas, mas sim como um complemento, em determinados planos, a estas. Os modelos originantes destas redes têm como base mecanismos neuronais do córtex visual de alguns animais. Esses mecanismos surgem já funcionais nos primeiros momentos de vida desses animais e não carecem de treino. No entanto, mostram-se importantes nas acções de processamento visual. Da mesma forma, as redes neuronais de acoplamento por pulsos, pelo menos as que se vão aqui tratar, não necessitam de ser treinadas antes de executarem as suas tarefas e conduzem a processamentos simplificados que ajudam a eliminar algumas das dificuldades patentes nas redes clássicas, principalmente no que concerne à variabilidade e dissemelhanças frente às imagens entregues ao processo de treino. Algumas das características de invariabilidade que podem ser conseguidas com este tipo de rede são a translação, a escala e a rotação.

5.2 O modelo neuronal PCNN

O primeiro modelo de uma rede neuronal de acoplamento por pulsos (PCNN) foi apresentado por Eckhorn [42]. Este modelo apresenta um neurónio dividido em duas partes principais, o *Feeding* e o *Linking* (Figura 5.1). Ambos recebem sinais de neurónios vizinhos, mas apenas o *Feeding* recebe sinais (S) de estímulos externos à rede e que no caso do presente estudo representam a imagem a processar. Os sinais provenientes dos neurónios vizinhos são condicionados por *matrizes sinápticas*, a M para o *Feeding* e a W para o *Linking*. Ambos, o *Feeding* e o *Linking*, retêm o estado das iterações anteriores com base numa memória de decaimento exponencial. Essa retenção é conseguida, no domínio discreto, por um filtro digital de 1.^a ordem conhecido na literatura por integrador *leaky*.

O valor do *Feeding* é dado pela Equação (5.1) e o do *Linking* pela Equação (5.2)

:

$$F_{ij}[n] = e^{\alpha_f} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1], \quad (5.1)$$

$$L_{ij}[n] = e^{\alpha_L} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1]. \quad (5.2)$$

Na Equação (5.1) $F_{ij}[n]$ corresponderá ao valor do *Feeding* do neurónio ij , estando este integrado numa matriz bidimensional de neurónios, à iteração n e em que S_{ij} corresponde ao valor do estímulo externo, à rede neuronal, proporcionado pelo *pixel* ij da imagem a processar. O $Y_{kl}[n-1]$ representa o valor das saídas, na iteração anterior, dos neurónios adjacentes de índice kl . Este valor será modificado de acordo com o elemento $ijkl$ da matriz M , sendo este uma representação da ligação sináptica entre o neurónio ij a computar e a saída do neurónio vizinho kl . Note-se que a parcela $\sum_{kl} M_{ijkl} Y_{kl}[n-1]$ executada para todos os elementos ij equivale a uma operação de convolução entre a matriz M e a matriz Y . Uma vez que a alteração dos valores de M modifica o comportamento do campo receptor dos sinais provenientes dos neurónios vizinhos, foi adicionado um factor normalizador, o V_F . O valor da constante V_F permite compatibilizar a magnitude do sinal resultante dos neurónios vizinhos com a dos sinais S_{ij} e $e^{\alpha_F} F_{ij}[n-1]$. O valor de α_F controla a magnitude do decaimento exponencial.

De forma semelhante, $L_{ij}[n]$ diz respeito ao valor da componente de *Linking* do neurónio ij na iteração n . A componente de memória, com decaimento exponencial, é estabelecida pela parcela $e^{\alpha_L} L_{ij}[n-1]$. O valor de α_L controla a magnitude do decaimento exponencial.

Por sua vez, o campo receptor dos sinais dos neurónios vizinhos é formalizado pela matriz W .

Os dois componentes, o *Feeding* e o *Linking*, são então combinados no mecanismo seguinte do neurónio, criando o estado interno, U . A equação seguinte determina essa combinação:

$$U_{ij}[n] = F_{ij}[n] \{1 + \beta L_{ij}[n]\}. \quad (5.3)$$

A proporção das duas componentes, F e L , na combinação é controlada pelo valor de β .

Na etapa seguinte, no neurónio, é comparado o valor do seu estado interno, U , com o valor de um *threshold* dinâmico, Θ . Este processo vai determinar o estado da saída, Y , do neurónio, sendo esse estado estabelecido pelo valor 1 ou o valor 0 , tal como a equação seguinte o define:

$$Y_{ij}[n] = \begin{cases} 1 & \text{se } U_{ij}[n] > \Theta_{ij}[n-1] \\ 0 & \text{outro valor} \end{cases}. \quad (5.4)$$

O Θ é dinâmico da forma seguinte. Na iteração em que o neurónio dispara, na qual a saída adquire o valor 1, o *threshold* (Θ) recupera o seu valor máximo. Na iteração seguinte a saída do neurónio volta para 0 e a partir daí o valor de Θ começa a desvanecer, com uma magnitude controlada por α_Θ , de forma exponencial através do integrador de *leaky*. Este decaimento cessa quando $U_{ij}[n] > \Theta_{ij}[n-1]$; nessa situação o valor da saída do neurónio é comutado novamente para 1, levando, uma vez mais, o valor de Θ ao seu máximo. A Equação (5.5) descreve a actualização do valor de Θ explicada neste parágrafo.

$$\Theta_{ij}[n] = e^{\alpha_\Theta} \Theta_{ij}[n-1] + V_\Theta Y_{ij}[n]. \quad (5.5)$$

O valor da constante V_Θ deve ser superior ao valor médio de U .

A Figura 5.1 esquematiza os mecanismos de processamento de um neurónio PCNN. Repare-se que cada dendrite possui um integrador *leaky*. Os quais, por todos terem a mesma constante temporal T , aparecem, na Equação (5.1) e na Equação (5.2), em parcela externa ao somatório. Esta simplificação reduz a magnitude do esforço computacional sem prejudicar os resultados finais.

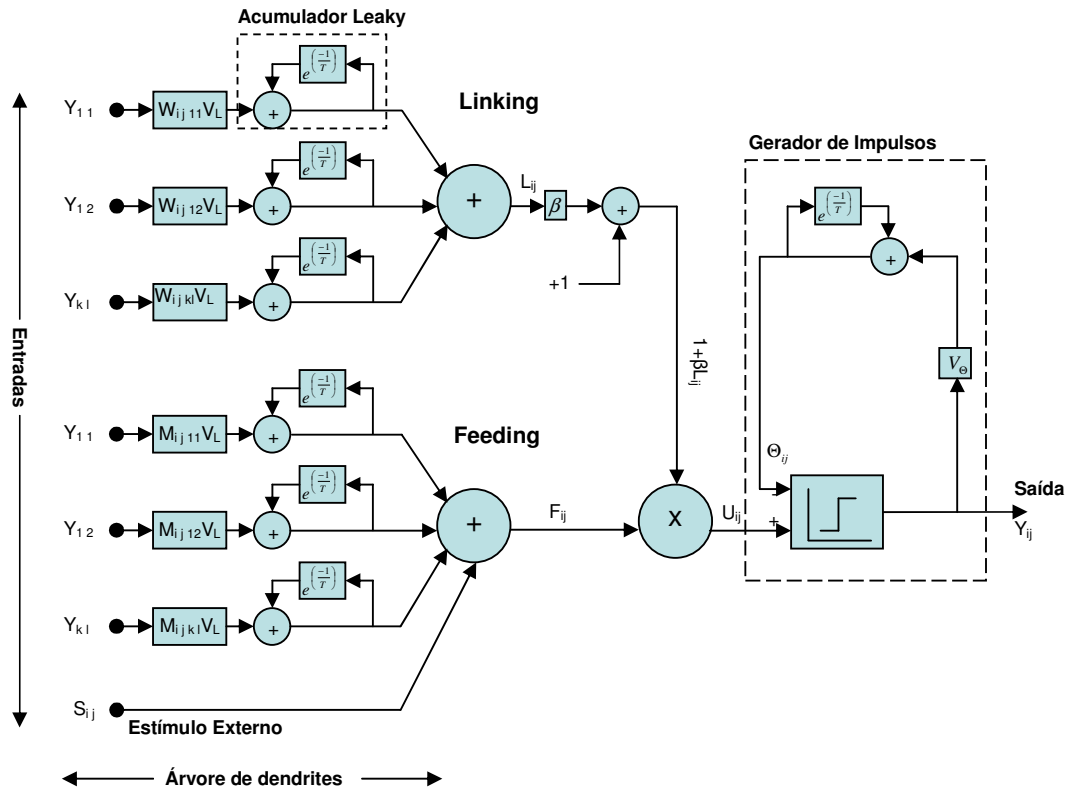


Figura 5.1– Representação do modelo de um neurónio PCNN

A estrutura de uma rede neuronal composta por neurónios deste tipo é esquematizada na Figura 5.2. Cada neurónio é circundado por um conjunto de neurónios homólogos e que se lhe ligam, através dos *Linkings* e dos *Feedings*, em raios correspondentemente bem estabelecidos. No modelo de neurónio, esquematizado na Figura 5.1, os raios destas ligações encontram-se associados às dimensões das matrizes M e W . O estímulo externo procede directamente do valor do *pixel* que cada neurónio PCNN tem associado e que integra a imagem a processar.

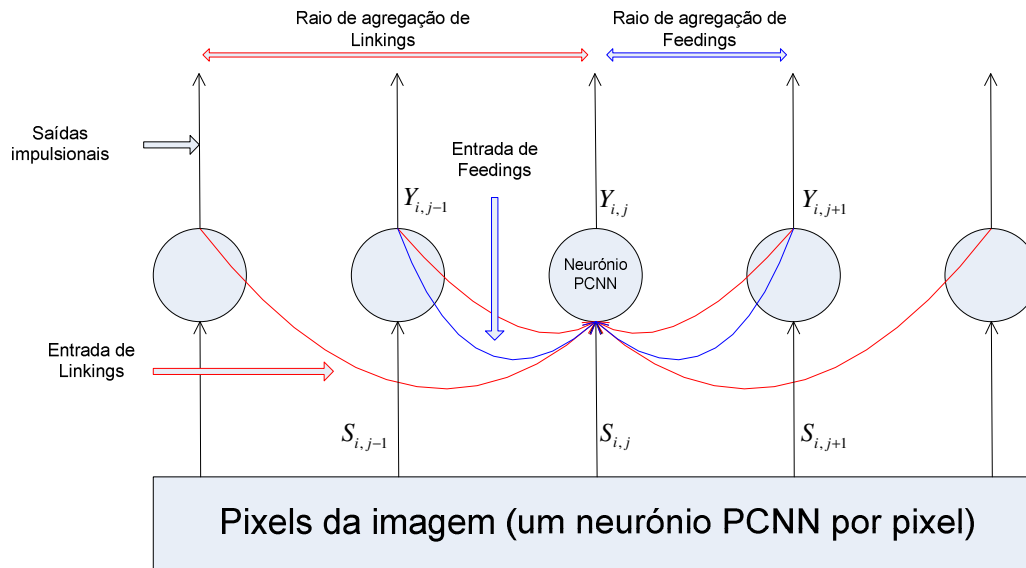


Figura 5.2 Estrutura genérica de uma rede PCNN

Antes da primeira iteração, as matrizes, F , L , U e Y são iniciadas a zero.

A computação das equações (5.1) até à (5.5), de forma repetida, constitui o percurso algorítmico que constrói a solução para a qual este modelo neuronal foi concebido. Neste modelo neuronal a evolução do processamento não está sujeito a nenhum critério de paragem. No entanto, consoante a sua utilização, esses podem ser adicionalmente definidos e constituídos.

O resumo funcional do modelo é a seguir apresentado. Os impulsos que chegam dos neurónios vizinhos, através das partes de *Feeding* e de *Linking*, contribuem, bem como os estímulos externos (imagem), para o valor, U , do estado interno do neurónio. À medida que as iterações avançam, o estado interno do neurónio aproxima-se do valor de Θ . Uma sobre-elevação desse valor, relativamente a Θ , faz que o neurónio dispare, ou seja, faz que a saída do neurónio mude de estado momentaneamente. Após isto, Θ acompanha esse disparo iniciando um desvanecimento do seu valor até que o valor do estado interno do neurónio volte a atingir o valor do Θ . Este mecanismo proporciona a natureza pulsante deste modelo neuronal, o que é visível na Figura 5.3.

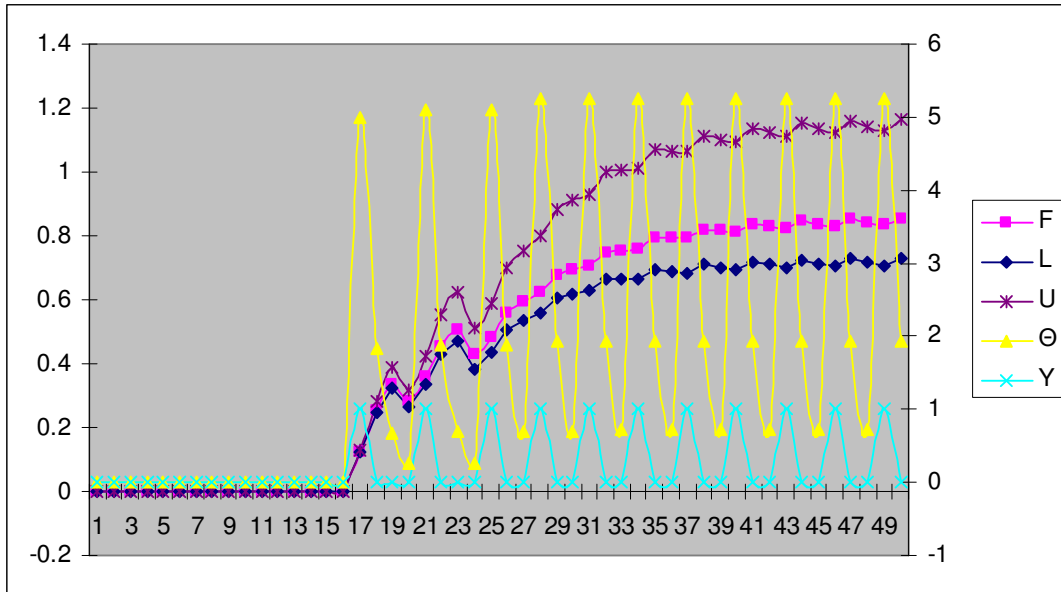


Figura 5.3 Sinais nos vários constituintes de um neurônio PCNN ao longo das iterações

Todo este processo traz consigo características que se revelam importantes para o processamento de imagens. Uma das características está ligada ao sincronismo entre os neurónios em consonância com múltiplos aspectos da imagem. Por exemplo, um determinado conjunto de neurónios poderá apresentar disparos síncronos, numa determinada iteração, de acordo com uma dada textura que se apresenta em correspondência na imagem a esses mesmos neurónios. Este comportamento deve-se, pelo menos em parte, à adjacência dos neurónios, que faz que um disparo de um neurónio tenda a provocar um disparo do neurónio vizinho, mas condicionado aos disparos dos restantes e aos estímulos externos de imagem.

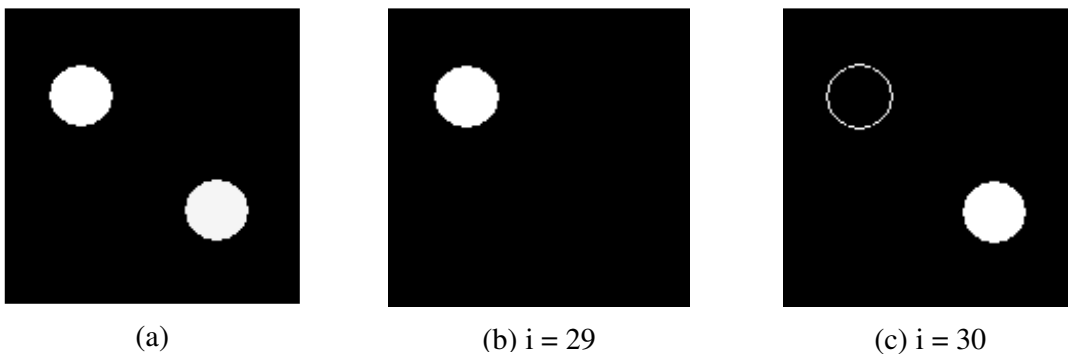
Considerem-se três neurónios, o **A**, o **B** e o **C**, colocados espacialmente de uma forma sequencial. O **A** está ligado ao **B** e o **B** está ligado ao **C**, em que apenas o **A** está sob a acção de um estímulo externo que vai provocar o seu disparo. Por adjacência, o neurónio **B** irá disparar a seguir. No entanto, este disparo apenas induzirá o disparo, desta vez, do neurónio **C**. Apesar de o sinal se propagar também a **A** e por localidade também a **B**, estes possuem agora, ao contrário de **C** que ainda não disparou, um valor de *threshold* ainda elevado, o que os inibe de disparar. Esta propagação, aqui exemplificada, com os neurónios **A**, **B** e **C** é normalmente referida como o princípio das *auto-ondas* [43] que são observáveis em múltiplos fenómenos físicos e que são, pelo seu intrínseco funcionamento, uma propriedade, também, das redes neuronais de acoplamento por pulsos (PCNN). Fisicamente uma *auto-onda* é caracterizada por não

sofrer reflexões nem refrações. Ou seja, quando duas ondas deste tipo colidem, nenhuma *passa* através da outra. Mas também é necessário referir que as redes PCNN não estão estritamente limitadas ao uso de mecanismos que respeitem as propriedades de uma *auto-onda* natural. Podem ser usadas algumas soluções nas PCNN que adulteram as propriedades de uma *auto-onda* de forma a se atingirem certas soluções para o processamento das imagens.

As imagens da Figura 5.4 são exemplos que mostram algumas das características descritas atrás do processamento de imagens digitais por uma PCNN.

A imagem origem, (a), que serve de estímulo à PCNN, é composta por um fundo preto e dois círculos, um branco, à esquerda, e o outro, com uma tonalidade próxima do branco, à direita. Na imagem (b), à iteração 29.^a, a saída neuronal resultante mostra o círculo da esquerda com toda a sua área, em correspondência ao círculo da esquerda da imagem original. A magnitude superior da intensidade dos *pixels* desta região fez que as saídas dos neurónios, que são estimulados por este círculo, atingissem o estado de disparo mais cedo que os neurónios associados à área circular da direita. Estes acabam por disparar na iteração seguinte, como é visível na imagem (c). Repare-se também no aparecimento, nessa iteração, do contorno exterior da área esquerda.

Os neurónios que dispararam e que constituem este contorno foram estimulados pela área circunscrita pelo contorno. As saídas dos neurónios que na iteração anterior assinalaram a área ficam agora inibidas de o fazer, por acção da elevação do valor do *threshold* após o disparo. Por outro lado, a não existência de disparos nos neurónios adjacentes e, neste caso, externos à região circular em causa, potencia o aparecimento de valores baixos de *threshold* que possibilitam, em complemento à estimulação referida, o disparo neuronal.



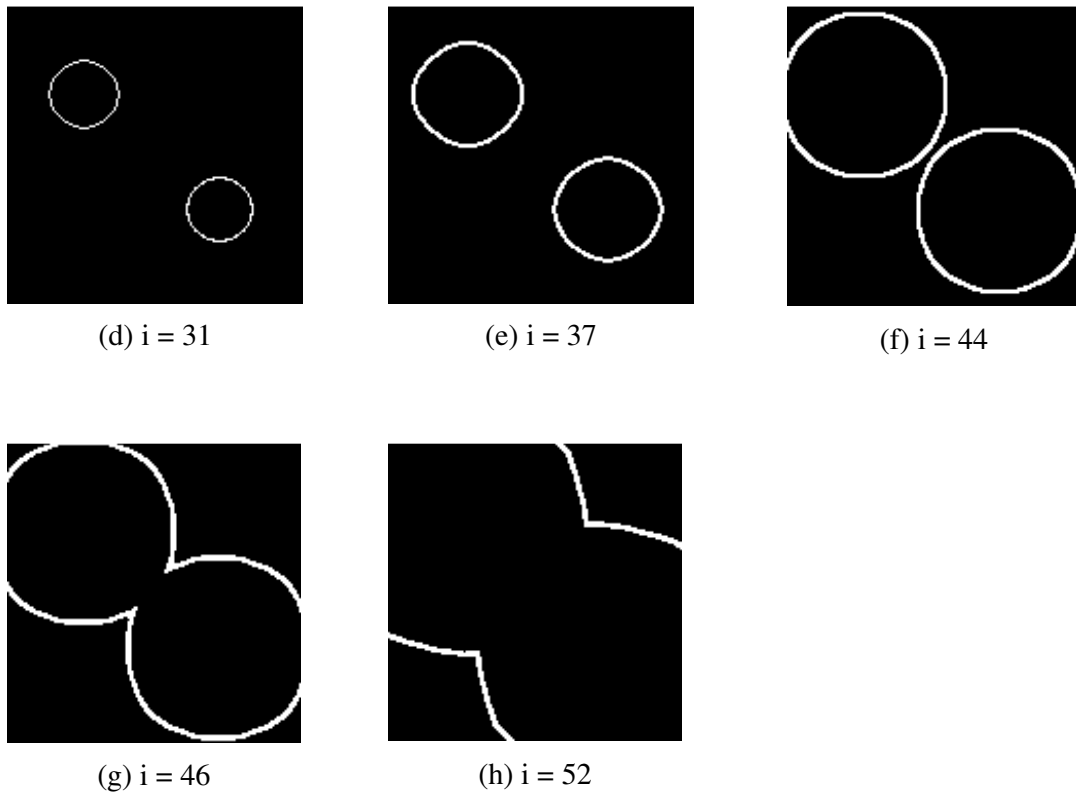


Figura 5.4 Evolução de uma *auto-onda* através de uma imagem. A imagem (a) corresponde à imagem original e as restantes, de (b) a (h), correspondem a imagens obtidas entre a iteração 29 e a iteração 52. i representa o número da iteração

Na iteração seguinte (31.^a), o aparecimento do contorno externo acaba também por acontecer na área do lado direito. Por sua vez, o contorno do círculo superior começa agora a propagar-se em conformidade com a característica de *auto-onda*, como se vai analisar a seguir.

A imagem (d) foi obtida na saída da PCNN à 31.^a iteração da rede neuronal. Nesta iteração, a imagem resultante evidencia os contornos externos dos dois círculos originais. As próximas iterações darão um resultado que é função, como se pode deduzir do exemplo apresentado atrás, em parte, dos estados anteriores das saídas de cada neurónio e da imagem estimuladora (imagem original). Daí que se torne visualmente observável a evolução de *auto-ondas* nas imagens. Esta evolução da *auto-onda* traduz-se, sinteticamente, por uma propagação, num plano, com direcções opostas aos estímulos que lhe deram origem. A imagem (e) mostra a saída neuronal obtida na 37.^a iteração, onde é notória a dilatação circunferencial dos contornos obtidos em (d), resultante da propagação da *auto-onda*. Na imagem (f) observam-se as duas *auto-ondas*,

duas iterações antes de se interceptarem. É observável na imagem (g) e na imagem (h) a cessação da propagação das *auto-ondas* no pontos onde se interceptam.

Analisou-se, com este exemplo, o mecanismo de *auto-onda* que está subjacente às redes neuronais de acoplamento por pulsos. Este fenómeno pode apresentar-se como útil no processamento digital de imagens. O estado que a *auto-onda* exhibe numa determinada iteração e posição pode ser essencialmente função de várias características locais, como a textura e a geometria de uma sub-região da imagem, para na iteração seguinte, por colisão das ondas, passar a ter uma representação de contributos de imagem mais globais. Veja-se o exemplo seguinte:

Na Figura 5.5 a imagem (a) mostra um conjunto de aglomerados de pontos que formam texturas granulares.

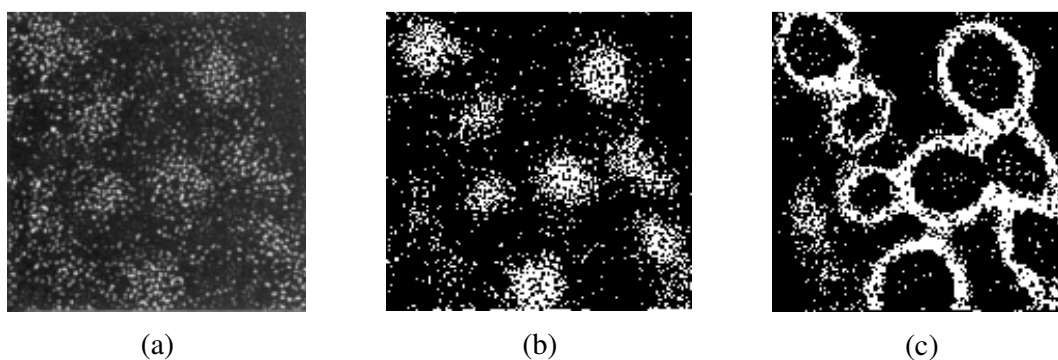


Figura 5.5 Processamento de uma PCNN sobre uma imagem de características granulares. A imagem original a processar é a (a). A imagem (b) foi obtida numa determinada iteração da PCNN anterior à iteração resultante na imagem (c)

À 12.^a iteração obtém-se, na imagem (b), uma segmentação das regiões que, predominantemente, na imagem original pertenciam aos aglomerados de maior densidade e de maior intensidade. Se se esperar pelas iterações seguintes, obtém-se à 39.^a iteração contornos delimitativos, próximos dos contornos definidores dos principais aglomerados em causa. De referir que estes resultados são obtidos fazendo alguns ajustes aos parâmetros da rede neuronal de forma a adaptá-la ao tipo de imagem com que se está a trabalhar e ao tipo de resultado pretendido. Por exemplo, os resultados da Figura 5.5 foram obtidos com uma configuração sensivelmente diferente da estabelecida para obter os resultados da Figura 5.6. Alguns parâmetros da rede neuronal foram modificados de um caso para o outro. É exemplo disso a dimensão do *kernel M*.



Figura 5.6 Processamento da PCNN a contornos de discontinuidades geométricas horizontais. A imagem original é a (a). O resultado do processamento à imagem (a) é visível na imagem (b) onde estão assinaladas discontinuidades horizontais da imagem original

O resultado que foi obtido à 42.^a iteração é visível na imagem (b) da Figura 5.6, onde estão assinaladas discontinuidades geométricas horizontais da imagem (a) e que podem representar um contorno constituído peculiarmente por este tipo de discontinuidades.

Esta série de exemplos demonstra o potencial de ajuda que estas redes podem trazer ao processamento de imagens digitais. Embora também deixe, indubitavelmente, questões acerca do seu aproveitamento. Como saber a que iteração se vai obter um resultado de interesse; ou ainda mais importante, se se vai, ou não, obter de forma aproveitável e correcta para a situação de processamento que venha a estar em causa? O modo de transpor estas questões passa por conduzir esse processamento à medida que as iterações avançam, interligando esse processamento com uma base de conhecimento de interacção mútua. Outra abordagem, passa por garantir que os processos posteriores a este estão preparados para reconhecerem essas situações e agirem para que o resultado não seja inesperado. As soluções que vão ser apresentadas seguem esta última estratégia numa figura implícita ao terem mecanismos de aprendizagem supervisionada. No treino desses mecanismos deve ser suportada, inerentemente, a capacidade de tratar correctamente as situações referidas.

5.3 O kernel da PCNN

O *kernel* utilizado no modelo de Eckhorn contém elementos que estão organizados de acordo com uma distribuição Gaussiana. O esforço computacional exigido por este *kernel* é dependente da sua dimensão. É verificável que para uma determinada escala de imagens a expansão espacial das ligações entre os neurónios PCNN não produz alterações significativas nos resultados. Em imagens naturais a substituição dos elementos do *kernel* também não produz variações elevadas nos resultados, como se pôde constatar nos ensaios que foram feitos nesse sentido. No entanto, em aplicações específicas, a substituição dos elementos do *kernel* ou até mesmo do próprio mecanismo de convolução pode trazer aumentos do desempenho funcional da rede PCNN.

O *kernel* mais utilizado computacionalmente, que faz variar a magnitude dos seus elementos de forma inversa à distância que os separa do elemento central do *kernel*, é referido por *kernel 1/r*:

$$K_{ij} = \begin{cases} 1 & \text{se } i = m \text{ e } j = m \\ 1/r & \text{outro valor} \end{cases}, \quad (5.6)$$

onde r corresponde à distância do centro do *kernel* ao elemento ij e onde m é metade da dimensão linear do *kernel*.

Um outro *kernel* passível de ser usado é o *On-Centre/Off-Surround*:

$$K_{ij} = \begin{cases} K_{ij} & \text{se } i = m \text{ e } j = m \\ -K_{ij} & \text{outro valor} \end{cases}. \quad (5.7)$$

Este tipo de configuração é detectável entre as células da retina do olho humano. As *auto-ondas* resultantes são espacialmente descontínuas. Tal facto, deve-se à concorrência entre neurónios adjacentes, pois neste caso, devido à existência de valores de *kernel* negativos e positivos, os neurónios têm tendência a alternar a sua actividade.

5.4 O ruído nas PCNNs

O ruído aleatório representa um grande problema para as PCNNs. O ruído destrói o sincronismo entre neurónios que a determinada iteração deveriam pulsar para assinalar um qualquer segmento. O ruído pode aparecer na rede por três formas distintas. Primeiro, pode advir dos estímulos externos S , ou seja, da própria imagem a processar. Segundo, pode formar-se em U , na conjugação do sinal dos vários compartimentos do neurónio. Terceiro, pode formar-se pela iniciação aleatória do valor limiar de Θ . Todas estas situações contribuem para a perda de capacidade de segmentação das PCNNs.

5.4.1 Redução do efeito do ruído por oscilação do valor limiar

Uma das formas que ajudam a insensibilizar as redes PCNN ao ruído consiste em introduzir um gerador de sinal na obtenção do valor limiar, Equação (5.8), que determina o disparo do neurónio. O gerador adicionará pequenas oscilações de frequência f ao valor limiar de disparo. Desta forma, periodicamente, o valor limiar crescerá, atrasando nesses instantes o disparo de certos neurónios que de outra forma iriam disparar, por ruído, intemporalmente. Nos instantes em que o valor do sinal oscilante decresce, os neurónios voltam a ter maior possibilidade de disparar, mas agora de forma sincronizada entre eles.

$$Y_{ij}[n] = \begin{cases} 1 & \text{se } U_{ij}[n] > \Theta_{ij}[n-1] + \left(\cos\left(f * \frac{n}{2\pi}\right) + 1.0 \right) \\ 0 & \text{outro valor} \end{cases} \quad (5.8)$$

Este método funciona correctamente se o ruído apresentar uma distribuição aleatória de iteração para iteração. Nas imagens digitais, após aquisição, o ruído permanece estático. Isto faz que o compartimento de *Feeding* acumule os valores de ruído de iteração para iteração. Em resultado, nas iterações posteriores, a eliminação do ruído pelas oscilações do valor limiar deixa de ser efectiva. Se o ruído em causa tiver proveniência na *figura de ruído* intrínseca aos órgãos electrónicos de aquisição de imagem, então torna-se conveniente fazer corresponder uma nova instância da imagem adquirida a cada iteração da rede PCNN. O valor de f é funcionalmente comparável à frequência dos sinais cerebrais globais, observáveis nos electroencefalogramas (EEG),

que parecem estar relacionados com a sincronização das funções cerebrais distribuídas [44].

5.4.2 Redução do efeito do ruído por propagação rápida do sinal de *Linking*

Este método consiste em fazer propagar a *auto-onda*, provinda apenas do compartimento de *Linking*, por toda a imagem, por cada iteração da rede PCNN. Ao contrário do método original em que a *auto-onda* deriva paralelamente do compartimento de *Linking* e de *Feeding*, aqui existe uma *auto-onda* intermédia que se propaga, não só pela região abrangida pelo *kernel*, mas também por toda a imagem. A propagação da *auto-onda* de *Linking* tende a colocar os neurónios num estado final de grupo, minimizando os efeitos de ruído que surgem pela existência de estados de disparo próximos, mas não simultâneos, entre neurónios. Este método foi originalmente desenvolvido [45] para minimizar os efeitos da discretização temporal das PCNN relativamente aos modelos contínuos. O que se denota ao adoptar esta técnica é uma efectiva redução dos efeitos do ruído nas segmentações finais. As equações seguintes, da (5.9) à (5.16), representam a modificação ao modelo de funcionamento das redes PCNN de acordo com este método. De realçar as equações (5.13) à (5.15), que formulam a propagação da *auto-onda*, derivada do *Linking*, por toda a imagem, repetidas vezes, até que as saídas dos neurónios deixem de variar. O método reduz o efeito do ruído, mas adiciona esforço computacional ao processo.

$$F_{ij}[n] = e^{\alpha_F} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1], \quad (5.9)$$

$$L_{ij}[n] = e^{\alpha_L} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1], \quad (5.10)$$

$$U_{ij}[n] = F_{ij}[n] \{1 + \beta L_{ij}[n]\}, \quad (5.11)$$

$$Y_{ij}[n] = \begin{cases} 1 & \text{se } U_{ij}[n] > \Theta_{ij}[n-1] \\ 0 & \text{outro valor} \end{cases}, \quad (5.12)$$

REPETIR

$$L_{ij}[n] = V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1], \quad (5.13)$$

$$U_{ij}[n] = F_{ij}[n] \{1 + \beta L_{ij}[n]\}, \quad (5.14)$$

$$Y_{ij}[n] = \begin{cases} 1 & \text{se } U_{ij}[n] > \Theta_{ij}[n-1] \\ 0 & \text{outro valor} \end{cases}, \quad (5.15)$$

ATÉ QUE Y DEIXE DE VARIAR

$$\Theta_{ij}[n] = e^{\alpha_\Theta} \Theta_{ij}[n-1] + V_\Theta Y_{ij}[n]. \quad (5.16)$$

5.4.3 Redução do efeito do ruído por controlo do decaimento do valor limiar

A técnica anterior eleva o número de operações que são necessárias para a PCNN funcionar, o que agrava o custo computacional da utilização das redes PCNN. O tipo de imagens que foram processadas no decorrer deste trabalho pelas redes PCNN fez com que fosse necessário deduzir uma outra forma, com menos operações, para a redução de ruído. Essa forma consiste em escolher o valor, α_Θ , relativo à Equação (5.5), para que provoque um decaimento lento no valor de comparação. É necessário aumentar o valor de V_Θ com o objectivo de que a contribuição do sinal de saída para o valor de comparação, após o disparo do neurónio, seja elevada. Estas acções fazem que o valor limiar diminua a sensibilidade de um neurónio imediatamente após o seu disparo, contribuindo para a sincronização entre neurónios. Isto conduz à sincronização de grupos de neurónios em consonância com as regiões a segmentar, mesmo na presença de algum ruído. Porém, é verificável que o aumento em demasia dessa dessensibilização impossibilita a segmentação de certas sub-regiões, como as que se encontram diferenciadas por *pixels* de valores próximos. Daí que este processo acabe por ser dependente do tipo de imagens que se deseja processar. Para as imagens do caso que se analisou, estes valores foram empiricamente obtidos, num processo de *trial-error*, de forma a garantir o equilíbrio entre o ponto do efeito da redução do ruído e o ponto de não perda de sub-regiões de interesse.

5.5 *Parametrização das PCNNs*

A parametrização das redes PCNN é conseguida através de uma abordagem de *trial-error*. É sem dúvida o maior aspecto negativo destas redes neuronais. Os parâmetros são ajustados empiricamente com base na funcionalidade conhecida de cada componente do modelo do neurónio PCNN e com base na resposta impulsional que é observada à saída da rede ao longo das várias iterações. É uma tarefa entregue à interpretação humana. O ajuste é executado mediante a percepção que se vai tendo das imagens de saída obtidas.

Quanto melhor for a partição global, julgada pelo sistema visual do humano que realiza o ajuste, maior será, em princípio, o desempenho segmentador/codificador da PCNN.

É uma tarefa complexa, uma vez que está em causa a configuração de mais de uma dezena de aspectos. Entre os quais, a distribuição dos elementos dos *kernels*, a *dimensão* dos *kernels* e os coeficientes de cada componente do neurónio PCNN. Dentro dos coeficientes encontram-se os de magnitude e os temporais, estando ambos aplicados aos campos receptores do neurónio. Têm-se também os coeficientes que regulam o limiar de disparo do neurónio. Alguns relacionam-se intensamente entre eles e outros são altamente não lineares. O próprio ajuste está dependente da natureza das imagens. Apesar de estas redes serem não supervisionadas, estão desta forma dependentes dos parâmetros escolhidos para um determinado tipo de problema. Assim, é necessário algum treino humano para se poder ajustar convenientemente uma rede deste tipo às aplicações de segmentação/codificação em causa.

O comportamento final das PCNNs pode ser ajustado através dos vários parâmetros que constituem um neurónio PCNN. A modificação do parâmetro β revela diferentes tipos de segmentação. Alterar este valor significa obter sub-segmentações orientadas a uma determinada escala de textura ou a uma determinada escala de sub-regiões homogéneas mas com tons distintos entre si. Enquanto que este parâmetro, em conjunto com as matrizes sinápticas M e W , permite escalonar as entradas de *Feeding* e de *Linking*, os três parâmetros, V , escalonam os sinais internos dos neurónios PCNN. Os valores associados à constante de tempo do limiar de disparo são usados para adaptar o comportamento impulsional às magnitudes dos sinais do corpo neuronal. O aumento da dimensão do *kernel* permite a comunicação entre neurónios mais afastados e também

uma propagação mais distante da *auto-onda* relativamente a um neurónio. Este factor permite controlar a magnitude e até o tipo de interferência entre neurónios. Uma vez mais, também o ajuste deste tamanho poderá ser função do tipo de segmentações a obter e por isso poderá estar sujeito a um processo de *trial-error*.

5.6 Séries temporais das PCNNs

Uma observação importante, que pode ser útil em algumas aplicações e que foi demonstrada [46] com redes deste tipo, está ligada à vectorização temporal dos valores das saídas dos neurónios. Foi verificado que a série resultante das múltiplas interacções, ao longo do tempo, produz informação invariante à translação, escala e rotação e até às intensidades das regiões, sobre o conteúdo das imagens. A série temporal é obtida pela Equação (5.17).

$$G[n] = \sum_{ij} Y_{ij}[n]. \quad (5.17)$$

Nas imagens seguintes, Figura 5.7, podem observar-se uma cruz (a), uma cruz numa escala reduzida (b) relativamente à primeira e uma cruz com a escala reduzida e também rodada em 45 graus (c) relativamente, também, à primeira.

Quando estas figuras são apresentadas a uma rede PCNN e as saídas activas da rede são contabilizadas, iteração a iteração, através da Equação (5.17), obtêm-se três séries muito semelhantes para cada figura. Como pode analisar-se na Figura 5.8, a divergência, entre as séries, apenas começa a acontecer nas últimas iterações. Assim, as séries podem ser usadas para caracterizar com invariância à escala e invariância à rotação arbitrarias figuras. Este facto também é válido para o processamento de regiões. Tal pode elevar o potencial de utilidade destas redes neuronais. No entanto, os experimentos que foram feitos usando imagens naturais mostraram que a aleatoriedade geométrica e o ruído contribuem de forma caótica para a dessincronização entre os neurónios PCNN, tornando, no caso das imagens naturais, estas características de invariabilidade pouco robustas e confiáveis. Este problema foi também observado por outros autores [47] e acaba por se tornar um factor de precariedade nos resultados que

foram obtidos sobre o caso de aplicação implementado. Isto, apesar de terem sido adoptados mecanismos coadjuvantes com o sentido de minimizar esse problema.

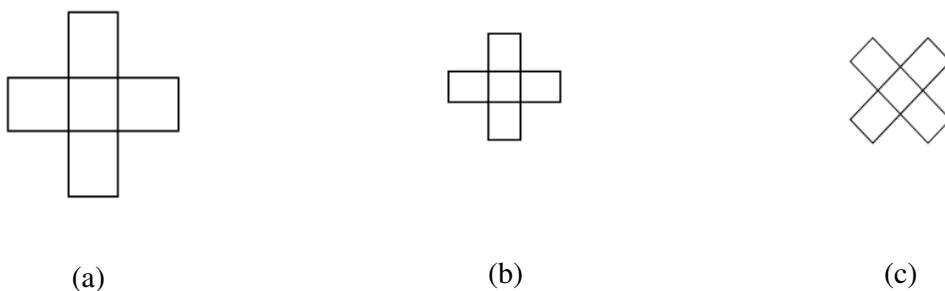


Figura 5.7 A mesma figura representada numa escala e rotação distinta. Relativamente à figura na imagem (a), a figura da imagem (b) representa uma escala inferior e a figura da imagem (c) representa (b) rodada 45 graus

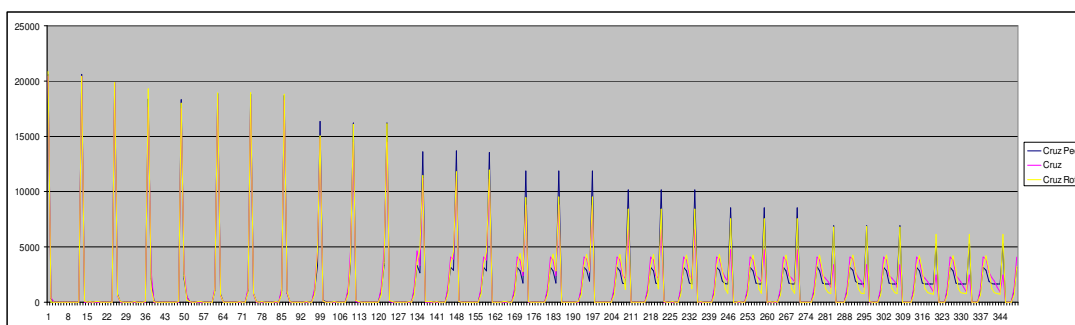


Figura 5.8 Três séries temporais resultantes da operação de uma rede PCNN nas imagens da Figura 5.7. No eixo vertical são contabilizados os números de neurónios activos por iteração. No eixo horizontal estão assinalados os índices de múltiplas iterações

5.7 As PCNNs no processamento de imagens

A utilização de redes neuronais em processos de análise de imagens digitais é algo que surgiu com as primeiras aplicações de redes neuronais. Quer o motivo estivesse relacionado com a proximidade destas à visão humana, quer por mero interesse exploratório, a verdade é que as funcionalidades apresentadas pelas redes neuronais nesta área detêm razoáveis benesses. De facto, as redes neuronais adequam-se à resolução de problemas em processos de visão que, pelos métodos mais tradicionais, por vezes, se mostra difícil de conseguir. Um dos processos classicamente usados na ajuda à segmentação de imagens e no reconhecimento de imagens é a correlação. A

técnica consiste em criar um *kernel* com a informação a procurar na imagem. Este é correlacionado com a imagem a processar. São avaliados, então, os pontos de correlação. Se a informação a procurar estiver na imagem, aparecerão pontos de correlação elevados. Geralmente, por uma questão de desempenho e de simplicidade, a correlação é feita no domínio das frequências e recorrendo à FFT (Fast Fourier Transform) ou a qualquer outro processo da mesma natureza, Figura 5.9.

Porém, este método possui uma série de problemas. Se a região a detectar não for exactamente a mesma que a região estabelecida pelo filtro, os níveis de correlação poderão ser diminutos. Com este método, a insensibilidade à mudança de rotação, escala, brilho e à existência de ruído é pouco elevada. Técnicas que trabalham no domínio das frequências a escalas e rotações variáveis são também comuns e podem aumentar a eficácia, como é o caso da transformada de *Gabor* e das *Wavelets*. Esses problemas são também comuns a outras técnicas. Redes neuronais clássicas, técnicas de morfologia ou processamento estatístico e probabilístico estão abrangidos por essas mesmas dificuldades. Qualquer técnica que use *comparação*, sem particionar as regiões em primitivas mais pequenas e que não analise a semântica que está associada à composição entre esses elementos incorre na elevada possibilidade de falha.

Alguns autores [48] defendem que é neste aspecto que as redes PCNN podem trazer um contributo importante. A capacidade que revelam, pelo menos em algumas situações, de, de uma forma inerente, extraírem por partes a informação fundamental de caracterização de imagens será a justificação dessa alegação. Eles afirmam que a base biológica destas redes, ao estarem no córtex visual de certos mamíferos, é justificação para o adequado funcionamento que apresentam. É verdade que o não carecerem de treino potencia estas redes relativamente aos métodos referidos anteriormente. Mas, como já foi referido, outra natureza de problemas está associado a estas redes. E do estudo feito e das experimentações realizadas não se pode dizer que estas redes sejam totalmente imunes aos problemas apresentados pelas outras técnicas. As vantagens que as PCNNs trazem estão ligadas aos fenómenos de activação simultânea de neurónios ao longo das regiões adjacentes e de energia similar. Isto faz que regiões de brilho semelhantes, texturas semelhantes ou contornos semelhantes tendam a ser assinaladas pelo disparo em grupo dos respectivos neurónios. Poder-se-ia aqui fazer alguma analogia com técnicas não supervisionadas de agrupamento e que são também estudadas nesta tese.

Relativamente à extracção de contornos em imagens é classicamente utilizado o filtro de *Sobel* [49] que consiste num pequeno *kernel* com os elementos centrais positivos e os elementos adjacentes negativos. No resultado da convolução deste *kernel* com uma imagem, são assinalados os pixels que são constituintes de contornos delimitadores de regiões. Porém, os contornos já existentes na imagem original, ao serem convolvidos, passam a ser representados por um duplo contorno. A extracção de contornos usando PCNN realiza-se sem este inconveniente [48].

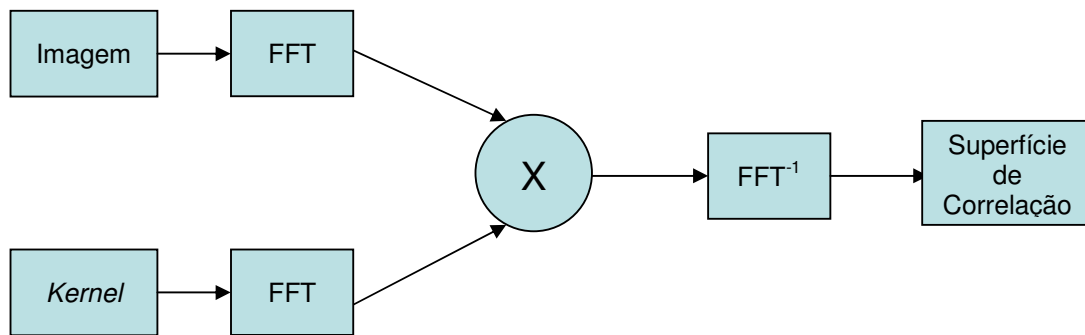


Figura 5.9 Diagrama de um processo de correlação no espaço de frequências

5.8 Segmentação produzida pela PCNN em imagens naturais de células sanguíneas

Embora pudesse ser usada a informação de cor, utilizando, em vez de um único componente de níveis de cinzento, três componentes (RGB), pareceu conveniente não o fazer. O primeiro motivo está relacionado com o desempenho temporal e o segundo emerge da capacidade humana de conseguir segmentar as células sem recorrer à informação de cor. A cor pode mesmo trazer perda à capacidade de generalização num sistema computadorizado [3]. Assim, pareceu à partida esclarecedor poder-se comparar os resultados deste método, que não usa cor, com um método que usa cor.

As imagens da Figura 5.11 e da Figura 5.12 representam um subconjunto das imagens de activação das saídas da rede PCNN, ao longo de cem iterações, resultantes do processamento à imagem da Figura 5.10.

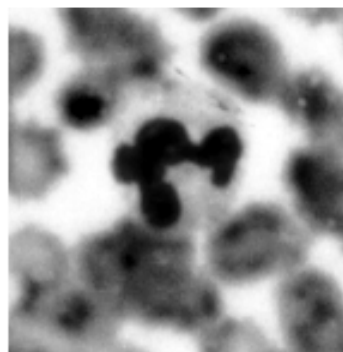


Figura 5.10 Imagem em níveis de cinzento de células a segmentar, o glóbulo branco encontra-se ao centro, circundado por glóbulos vermelhos

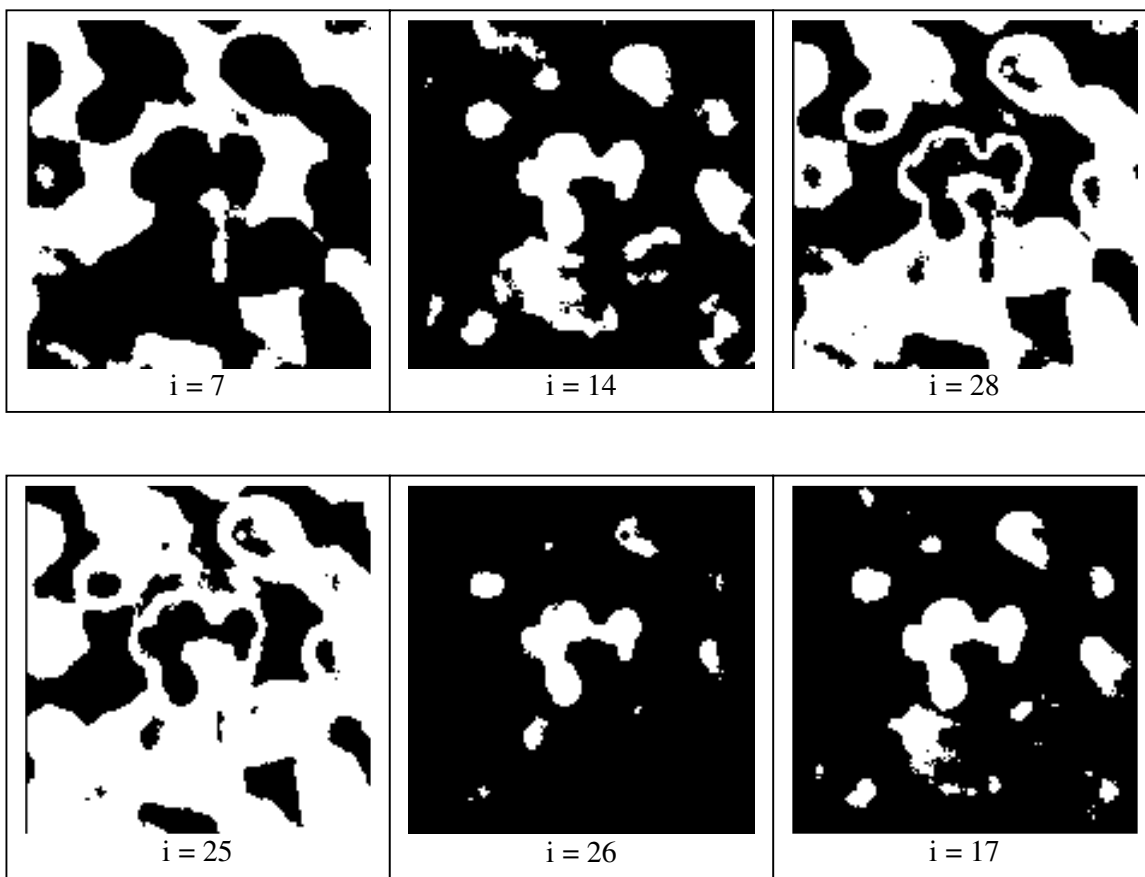


Figura 5.11 Alguns resultados das saídas das redes PCNN ao longo de cem iterações do processamento à imagem da Figura 5.10. i representa o número da iteração

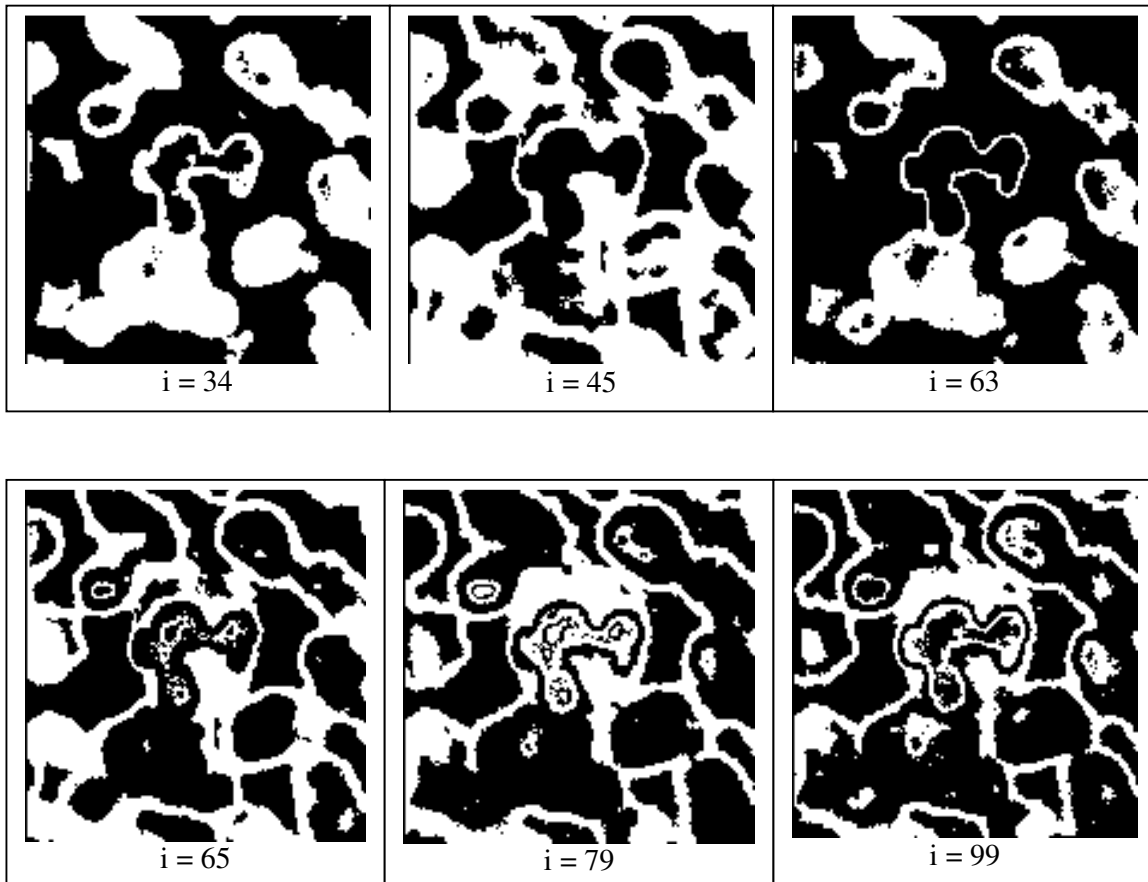


Figura 5.12 Alguns resultados das saídas das redes PCNN ao longo de cem iterações do processamento à imagem da Figura 5.10. i representa o número da iteração

A parametrização usada para a PCNN é mostrada na Tabela 5.I.

Tabela 5.I Configuração da PCNN para a segmentação não supervisionada

| Componentes da PCNN | Configuração |
|---|-----------------|
| kernel M | 5x5; tipo l/r |
| kernel W | 5x5; tipo l/r |
| Vf | 0.1 |
| Vl | 0.1 |
| $V\theta$ | 33.0 |
| β | 0.1 |
| Af ($\alpha_f = -1/Af$) | 10.0 |
| Al ($\alpha_l = -1/Al$) | 6.0 |
| A θ ($\alpha_\theta = -1/A\theta$) | 18.0 |

Estas imagens evidenciam a capacidade que as redes PCNN têm de extrair em determinadas iterações segmentos de específico interesse. No entanto, também é visível que nem sempre os segmentos pertencentes a uma sub-região em particular são totalmente isolados. Na iteração 7, grande parte do fundo da imagem é extraído, mas também o é uma porção do citoplasma da célula. Na iteração 17 o núcleo da célula é completamente isolado. O citoplasma da célula é inteiramente assinalado na iteração 25, mas não é totalmente isolado do resto das sub-regiões. Os glóbulos vermelhos são segmentados na iteração 34, ainda que parte do núcleo do glóbulo branco apareça agregado a um aglomerado de glóbulos vermelhos. O contorno do núcleo do glóbulo branco é detectado na iteração 63, mas não é completamente isolado. Nas iterações 79 e 99 o citoplasma do glóbulo branco é assinalado, de forma ainda mais correcta, nomeadamente a sua parte superior, embora não seja totalmente isolado.

Estas imagens reforçam, ainda que com algumas condicionantes, a possível vantagem de se utilizarem as redes PCNN para segmentar regiões de interesse em imagens. Todavia, devido a essas condicionantes, não devem ser utilizadas sem um mecanismo de supervisão. A secção seguinte aborda uma forma de aproveitar o potencial segmentador das redes PCNN através de um mecanismo híbrido, composto por redes PCNN, inerentemente não supervisionadas, e composto por classificadores SVM, estes últimos supervisionados.

5.9 Método de segmentação com PCNNs e SVMs

A forma original de segmentação usando PCNNs não possui um processo de supervisão. O objectivo que se procurou atingir foi o de obter um método que associa as redes neuronais de acoplamento por pulsos a métodos de aprendizagem supervisionados. A ideia é a de entregar os resultados obtidos na segmentação não supervisionada das redes PCNNs a redes de aprendizagem supervisionada, como as SVM ou as *back-propagation*.

Da análise que se fez e que se apresentou atrás verifica-se que o resultado obtido com a PCNN é dependente de iteração para iteração. Qual a iteração de segmentação mais favorável? Existe alguma iteração favorável? Poderiam ter sido escolhidos métodos de procura da iteração mais favorável. O método que se adoptou parece à partida menos arriscado e mais generalista do que tentar procurar a melhor segmentação, pois utiliza todas as segmentações produzidas, iteração a iteração, pela

PCNN. Contudo, a elevada dimensionalidade que esta solução implica é a vários níveis não suportável. Gerar uma rede neuronal com um número de elementos de entrada que agregue todos os pixels resultantes de todas as segmentações produzidas pela rede neuronal é, para a capacidade dos suportes físicos que se possuem hoje em dia, totalmente impraticável. Isto é verdade ao nível de capacidade de processamento e ao nível da capacidade de armazenamento. Mais do que isso, estas duas problemáticas estão intensivamente relacionadas.

Para diminuir este problema decidiu-se introduzir entre a rede PCNN e a rede supervisionada um processo de redução de dimensionalidade, no que refere ao número de pixels resultantes da segmentação iteração a iteração da PCNN. Como as imagens resultantes à saída da PCNN são binárias, pareceu adequada a utilização de descritores de regiões de Hu [24], tendo sido posteriormente usados descritores de Zernike [7;26;28]. O superior desempenho, relativamente aos descritores de Hu, mostrado pelos descritores de Zernike, fez que se abandonasse a utilização dos descritores de Hu. Assim, as saídas da PCNN, descritas numericamente pelos descritores de Zernike, são entregues à entrada do classificador supervisionado. Desta forma, reduz-se consideravelmente a dimensão do espaço de entrada desse classificador. Com certeza que assim também se juntam outros riscos, nomeadamente riscos relativos ao possível mau desempenho destes descritores. É possível que certos aspectos geométricos, importantes, deixem de ser evidenciados pelos valores devolvidos no descritor. Se esses aspectos forem diferenciadores ou caracterizadores da região em causa, a sua omissão diminuirá a capacidade funcional do sistema. Esse risco pode ser atenuado aumentando a ordem dos descritores de Zernike, mas também aumentando a dimensionalidade do espaço de entrada da rede supervisionada. A escolha de uma ordem mais elevada fica associada, uma vez mais, ao equilíbrio entre o desempenho computacional e o desempenho funcional.

Esta técnica permite considerar o relacionamento semântico resultante das várias fases de segmentação obtidas na PCNN. Porém, algumas fases poderão ser menos coerentes. Cabe então ao classificador supervisionado o discernimento correcto das fases importantes. Isto durante a aprendizagem. De qualquer forma, a ideia fundamental está em ganhar alguma independência às características redundantes e não sinalizadoras, na caracterização das regiões de interesse. Por outro lado, não deverão as características unicamente diferenciadoras e não caracterizadoras ser perdidas. Esta tarefa é proporcionada, mais directamente, pela PCNN. A fase posterior é considerada pelo

classificador supervisionado. Este, sim, deverá aprender aquilo que realmente participa no caracterizar da região a classificar. Como a experiência tem mostrado, a tentativa de agregar todas estas tarefas em métodos de aprendizagem supervisionada tem fracassado. A justificação é simples: não é possível treinar um classificador supervisionado com todos os padrões possíveis de aparecerem num determinado problema. Principalmente por duas causas: a dificuldade na obtenção dos padrões em número elevado e o crescimento dimensional que isso implica. Um crescimento dimensional da estrutura do classificador provoca múltiplos problemas na convergência óptima do classificador. São comuns os problemas de bloqueio por mínimos locais, os problemas por escassez de recursos computacionais e os problemas de aumento no número de operações de processamento. Nos últimos anos várias soluções têm sido apresentadas no que respeita ao problema da retenção em mínimos locais. São exemplo disso métodos de base genética [50] e as próprias SVMs. Mas, tradicionalmente, essas soluções produzem um agravar do número de operações a cumprir, para a finalização da tarefa. O método aqui descrito incorre na mesma incapacidade, a de solucionar o problema sem um aumento muito considerável do número de operações; embora, com este método, se procure deslocar um pouco esse problema para o bloco funcional dos descritores e da PCNN ao invés de o produzir directamente no classificador. Isto tende a confinar o aumento das operações às dimensões da imagem a analisar. Na situação oposta, o aumento do número das operações fica normalmente relacionado com a dimensão da estrutura interna do classificador que por sua vez se relaciona com o número de padrões a aprender, o que normalmente implica um maior custo do que a situação mais favorável anteriormente referida.

Descrição das fases constituintes do treino:

1. Percorrer a imagem, pixel a pixel, numa progressão bem definida. Poderá ser da esquerda para a direita e de cima para baixo.
2. Para cada pixel da imagem, considerar uma vizinhança de pixels que pode ser quadrada e com uma determinada dimensão.
3. As subáreas resultantes dos passos anteriores são processadas pela PCNN em *i-iterações*.

4. As subáreas binárias resultantes nas saídas das PCNNs, ao longo de *i-iterações*, são descritas, numericamente, por um conjunto de descritores de regiões, os descritores de Zernike.
5. Para cada pixel, as séries que descrevem as regiões têm um tamanho de *i-elementos* vezes o número de descritores usados. Cada uma destas séries corresponde a um padrão de treino.
6. Normalização dos padrões de treino para um desvio padrão de uma unidade e para uma média de zero.
7. Cada padrão de treino é apresentado a um classificador de aprendizagem supervisionada, SVM. Cada padrão está associado a um padrão de saída do classificador. Os pixels pertencentes às regiões a segmentar têm um valor na saída da SVM situado, dentro do domínio de valores de saída da SVM, num dos limites desse domínio; enquanto que os pixels não inclusos nas regiões de interesse têm um valor de saída de SVM no limite oposto.
8. O processo descrito no ponto 7 será repetido um número de vezes de acordo com as regras de convergência do classificador de aprendizagem supervisionada em causa.

Descrição das fases constituintes da segmentação (pós-treino):

1. Percorrer a imagem a segmentar, pixel a pixel, numa progressão bem definida. Poderá ser da esquerda para a direita e de cima para baixo.
2. Para cada pixel da imagem, considerar uma vizinhança de pixels que pode ser quadrada e com uma determinada dimensão. A morfologia da janela de abrangência de pixels e a sua dimensão terão que ser a mesma usada na fase de treino.
3. As subáreas resultantes dos passos anteriores são processadas pela PCNN em *i-iterações*. A parametrização da PCNN e o número de iterações terá que corresponder ao que foi usado na fase de treino.

4. As subáreas binárias resultantes nas saídas das PCNN, ao longo de *i-iterações*, são descritas, numericamente, por um conjunto de descritores de regiões, descritores de Zernike, os mesmos empregues na fase de treino.
5. Para cada pixel, as séries que descrevem as regiões têm um tamanho de *i-elementos* vezes o número de descritores usados. Cada uma destas séries corresponde a um padrão a analisar.
6. Normalização dos padrões a analisar para um desvio padrão de uma unidade e para uma média de zero.
7. Cada padrão é apresentado a um classificador de aprendizagem supervisionada, SVM. Os pixels pertencentes às regiões a segmentar são assinalados pelo classificador, com um valor situado, dentro do domínio de valores de saída do classificador, num dos limites desse domínio; enquanto que os pixels não inclusos nas regiões de interesse têm um valor de saída de classificador no limite oposto.

A organização dos constituintes de todo o processo de segmentação com o híbrido PCNN/SVM é observável na Figura 5.13.

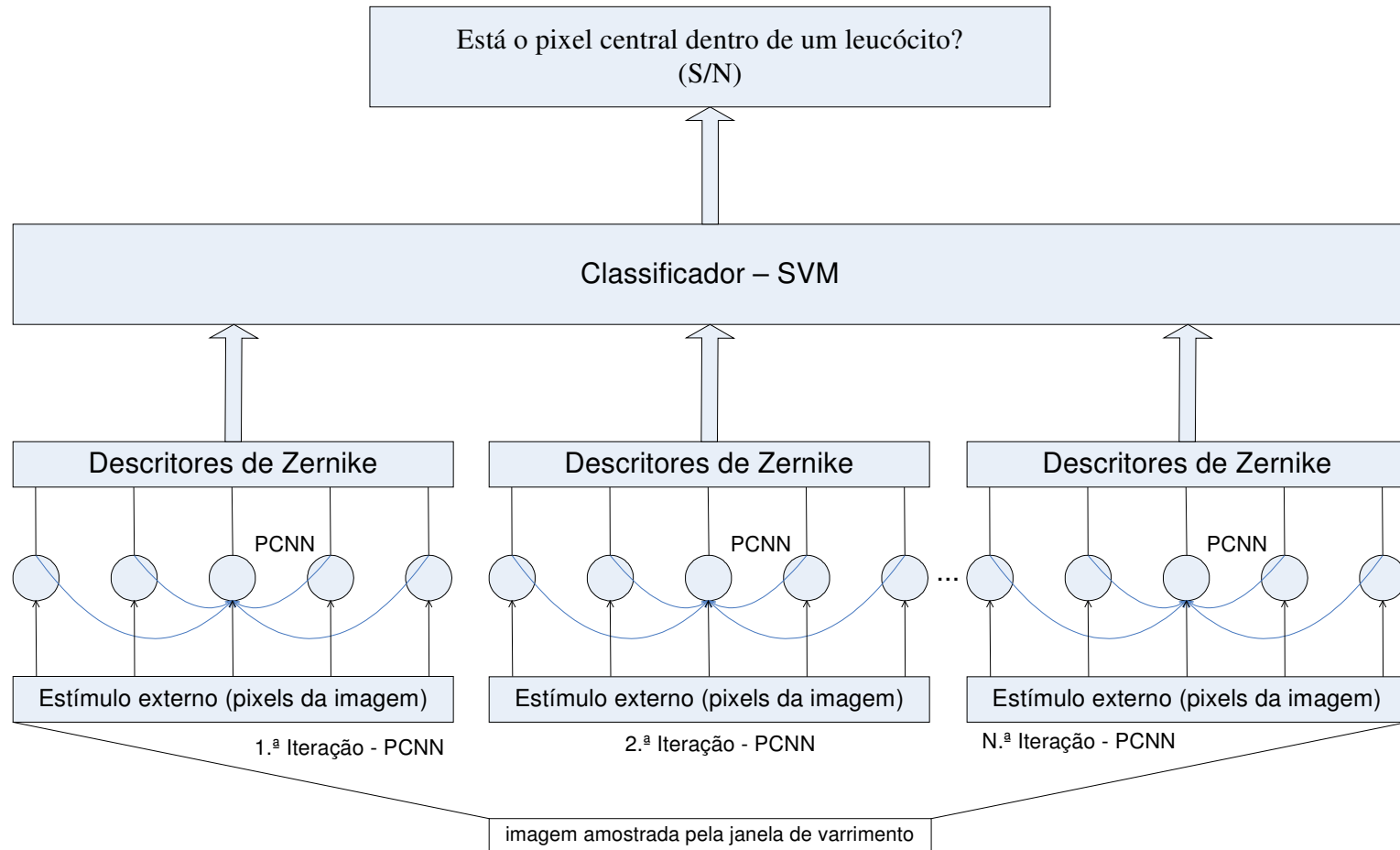


Figura 5.13 Composição do método de segmentação por PCNN/SVM

5.9.1 Resultados obtidos com o segmentador híbrido

Relativamente à estrutura apresentada na Figura 5.13, foram feitos testes substituindo os descritores de Zernike pelos descritores de Hu, assim como no bloco de aprendizagem supervisionada onde também foram feitos testes nos quais se tentou usar a rede *back-propagation* (BPNN) em oposição à SVM. O par Zernike/SVM mostrou, como era esperado, uma eficiência relativa superior nos resultados. As imagens de resultados apresentadas sobre o segmentador híbrido foram alcançadas recorrendo a esse último par. Nos testes feitos com as redes *back-propagation* e os testes feitos com os descritores de Hu não se obtiveram comportamentos finais excepcionalmente distantes dos obtidos com os descritores de Zernike e a SVM, apesar de estes últimos terem sido, em média, superiores. Foram até registadas algumas dificuldades associadas ao treino do classificador SVM. Devido ao número relativamente elevado de entradas da SVM, a dimensão de trabalho da sua estrutura causou problemas de operação à máquina que a implementava. As SVMs detêm duas vantagens, relativamente às redes *feed-forward/back-propagation*, que são evidentes: a separação de classes é feita linearmente sem os problemas de retenção em mínimos locais e a generalização é melhorada, evitando-se o *over-fitting*, pelo mecanismo de regularização de margens.

A parametrização determinada para a PCNN é dada na Tabela 5.II. Os componentes usados e que estruturam todo o modelo de segmentação híbrida são referidos na Tabela 5.III.

Tabela 5.II Configuração da PCNN para a segmentação híbrida

| Componentes da PCNN | Configuração |
|--|-----------------|
| $Kernel M$ | 5x5; tipo $1/r$ |
| $Kernel W$ | 5x5; tipo $1/r$ |
| V_f | 0.1 |
| V_l | 0.1 |
| V_Θ | 33.0 |
| β | 0.1 |
| A_f ($\alpha_f = -1/A_f$) | 10.0 |
| A_l ($\alpha_l = -1/A_l$) | 6.0 |
| A_Θ ($\alpha_\Theta = -1/A_\Theta$) | 18.0 |

Tabela 5.III Caracterização dos componentes usados na segmentação híbrida

| Componentes gerais | Configuração |
|---------------------------------|---|
| Tipo de pixel | Cinzento – 8 bits |
| Janela circundante a cada pixel | Quadrada – 51x51 |
| Comprimento da série PCNN | 20 Iterações |
| Descritores | Os 6 primeiros descritores de Zernike. – (0,0); (1,1); (2,0); (2,2); (3,1); (3,3) |
| Classificador supervisionado | SVM |
| Dimensão de entrada da rede | 240 – 20 iterações por 6 descritores de valores complexos |
| <i>Kernel</i> usado na SVM | Adatron – Radial Base Function |
| Número de SVMs | 1 |
| Número de padrões de treino | 95354 |

5.9.2 Um modelo para comparação na segmentação

Para que se pudesse ter uma referência sobre o desempenho do modelo híbrido, foi feita uma comparação com um modelo não-híbrido que utiliza apenas um classificador supervisionado para obter a segmentação [3]. Este modelo segue uma estratégia semelhante no que refere à estrutura de resolução do problema, ou seja, cada pixel é classificado de acordo com a pertença ou não à região interna do leucócito, Figura 5.14. A rede neuronal, de característica supervisionada, usada é de propagação única, das entradas para as saídas, mas utiliza como algoritmo de aprendizagem o método de propagação para trás (Backpropagation). Não são utilizados descritores auxiliares.

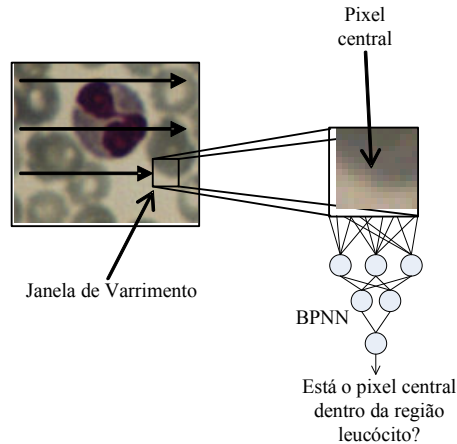


Figura 5.14 Modelo de segmentação neuronal exclusivamente supervisionado através de uma rede *feed-forward* com um algoritmo de aprendizagem *backpropagation*

5.9.3 Imagens resultado da segmentação híbrida

Nas imagens seguintes é possível observar e comparar as segmentações produzidas por um modelo exclusivamente supervisionado e pelo modelo híbrido que integra uma rede PCNN e um classificador SVM. Da Figura 5.15 até à Figura 5.19 a primeira imagem corresponde à imagem original. Ainda da Figura 5.15 à Figura 5.19, a segunda imagem mostra a segmentação produzida pelo mecanismo da primeira abordagem (BPNN). A última imagem das mesmas figuras apresenta a segmentação resultante da aplicação do mecanismo híbrido (PCNN-SVM). A imagem original é codificada em níveis de cor para o caso do método supervisionado e codificada em níveis de cinzento para o caso do modelo híbrido. No primeiro caso a cor funciona como uma característica discriminadora excelente e provoca eficazes segmentações se as situações de cor das imagens forem semelhantes às existentes no conjunto de padrões aquando do treino neuronal da BPNN. Se as imagens das células tiverem uma tonalidade diferente, tanto a nível da distribuição de croma como da distribuição de brilho, das do conjunto de treino, então os resultados poderão ser negativamente inesperados. Em termos práticos é impossível garantir que as imagens de treino abrangam qualquer situação de imagem futura. Isto deve-se principalmente a três motivos. Primeiro, são irrealizáveis a obtenção e a previsão de imagens de treino para todas as arbitrarias possibilidades futuras. Segundo, a tarefa de criar os padrões de treino de saída pelo assinalar da região leucócito por um especialista tem evidentemente

limites temporais e de recursos humanos. Terceiro, o acréscimo do conjunto de padrões de treino provoca dificuldades no próprio treino da rede neuronal, em vertentes como a otimização da aprendizagem e a velocidade da aprendizagem.

A tentativa de se usar imagens em níveis de cinzento na primeira abordagem demonstrou elevada precariedade na discriminação da região leucócito. Só um aumento computacionalmente insuportável do tamanho da janela de amostragem consegue levar à produção de medianos resultados.

Relativamente à abordagem mista (PCNN-SVM), por se tratar de uma abordagem, de princípio, computacionalmente mais intensa, optou-se por tentar usar apenas imagens em níveis de cinzento. Note-se que, ainda assim, esta abordagem exige um esforço computacional superior, relativamente à primeira abordagem, que emprega cor. Isto é verdade se, nos dois casos, forem utilizadas as mesmas dimensões de janelas de amostragem e o mesmo número de padrões de treino. Para se obterem os mesmos resultados generalistas com a primeira abordagem, teriam que se aumentar o tamanho da janela de amostragem e também o número de padrões de treino e conseqüentemente as dimensões da rede neuronal. Daí que se possa concluir que a partir de um certo nível de resultados, a segunda abordagem acaba por ser temporalmente mais eficaz que a primeira.

Nos exemplos apresentados a seguir, os contornos produzidos pela segmentação estão assinalados a azul. As regiões assinaladas a verde representam o citoplasma e as regiões assinaladas a vermelho representam o núcleo da célula.

Em todos os exemplos, com excepção do da Figura 5.18, o citoplasma constituinte dos leucócitos é sub-limitado pelo modelo que usa a rede *back-propagation*. Em oposição, a associação do modelo não supervisionado com o modelo supervisionado (PCNN-SVM) consegue produzir segmentações mais perfeitas ao assinalar a quase totalidade do citoplasma em todas as células com a excepção das células relativas aos exemplos da Figura 5.18 e da figura Figura 5.19. É necessário destacar que estas imagens não eram constituintes do conjunto de padrões de treino de ambos os mecanismos.

A região conjunta núcleo/citoplasma é assinalada integralmente por um dos mecanismos descritos de segmentação. A diferenciação núcleo/citoplasma é executada por um processo que tem como base a análise do histograma dos valores dos pixels da região segmentada [3].

Apesar de a capacidade generalista da segunda abordagem ser superior à da primeira, a capacidade especificista, que refere a íntegra segmentação das imagens existentes nos padrões de treino, não diminui acentuadamente. No entanto, em algumas imagens, foram detectadas ligeiras imperfeições aquando da comparação desta vertente de resultados com a primeira abordagem. É verdadeiramente difícil garantir a capacidade de generalização sem perder a capacidade de especificação e vice-versa.

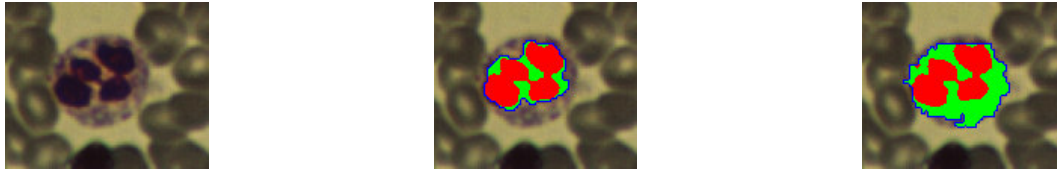


Figura 5.15 Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)

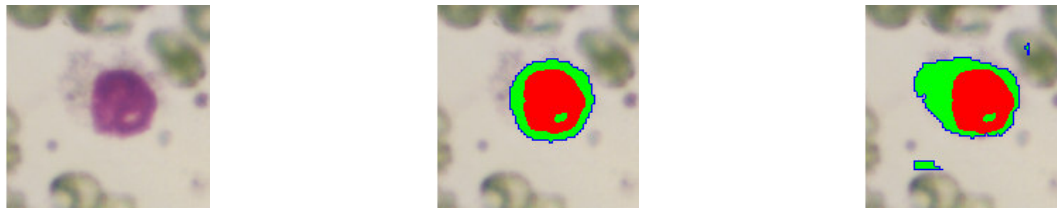


Figura 5.16 Comparação entre os resultados de segmentação de um linfócito obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)

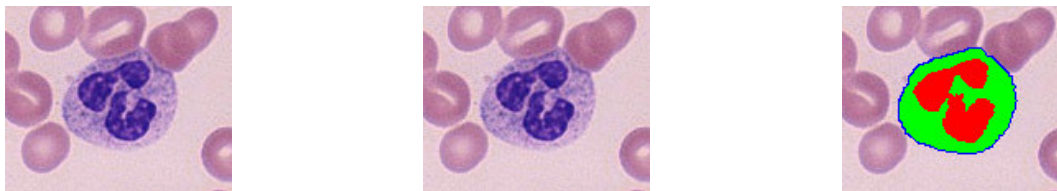


Figura 5.17 Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)

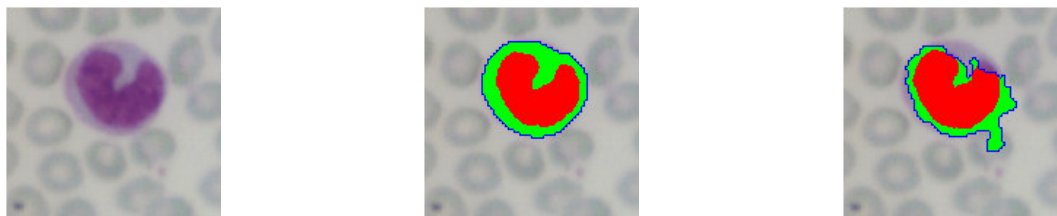


Figura 5.18 Comparação entre os resultados de segmentação de um monócito obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)



Figura 5.19 Comparação entre os resultados de segmentação de um neutrófilo obtidos pela abordagem não híbrida (BPNN) e pela abordagem híbrida (PCNN-SVM)

5.9.4 Conclusões sobre o segmentador híbrido

As imagens apresentadas mostram fundamentalmente que o método híbrido é mais generalista do que o método estritamente supervisionado baseado na rede *back-propagation*. Esse método consiste em entregar à entrada da rede neuronal supervisionada os valores dos pixels circundantes ao pixel a segmentar. O sucesso dessa segmentação fica intensamente condicionado à existência de padrões de treino representativos dos padrões apresentados após treino.

As imagens usadas para testar o método constituem um grupo que não foi usado na fase de aprendizagem e algumas apresentam tons e brilho muito díspares das usadas em treino. Este facto não é tão tolerado pelo método estritamente supervisionado. No método híbrido a segmentação de imagens ausentes e diferentes do grupo de treino é satisfatória, uma vez que na maior parte das imagens as regiões constituintes dos glóbulos brancos são assinaladas. É no entanto de referir que, relativamente ao método unicamente supervisionado, as próprias imagens de treino são segmentadas de forma mais precária. É também de realçar que este método falha a segmentação em algumas imagens, embora a eficácia geral de segmentação generalista seja superior à do método unicamente supervisionado. Porém, nem sempre o desempenho positivo de âmbito generalista foi confirmado ao utilizar-se este modelo. A Figura 5.18 e a Figura 5.19 mostram uma situação em que o método exclusivamente supervisionado obteve melhores resultados. O desempenho temporal do método híbrido é, sem se ter procurado optimizá-lo, bastante inferior ao método unicamente supervisionado. Por exemplo, uma imagem de 138x114 pixels é processada em 0.4 segundos pelo método unicamente supervisionado e em 190 segundos pelo método híbrido. No caso de uma imagem de dimensão igual a 188x157 o método de base *back-propagation* demora 0.65 segundos,

enquanto que o método híbrido demora 570 segundos. Importa referir aqui que o método de base *back-propagation* foi sujeito a técnicas de optimização a vários níveis, enquanto que o híbrido não obteve qualquer laboração nesse sentido. Isto contribui para que estes números de desempenho computacional sejam tão díspares.

Existe pelo menos uma justificação para os problemas de desempenho funcional denotados nesta abordagem híbrida. A série produzida pelas saídas da PCNN e que no seu agregado constitui, a cada iteração, uma nova segmentação da imagem em processamento não é abrangentemente independente de características elementares dessa imagem. Uma variação, por exemplo, no brilho da imagem altera a iteração em que uma determinada segmentação é produzida. Quer isto significar que pelo menos a fase da série sofrerá um deslocamento, o que implicará uma interpretação diferente por parte do bloco de aprendizagem seguinte. Este facto reduz significativamente o desempenho esperado para este método. São exemplos deste facto os resultados obtidos na Figura 5.18 e na Figura 5.19. Seria assim necessário integrar no método um mecanismo que tornasse a aprendizagem da série PCNN independente à sua variação de fase. Uma outra possível melhoria levaria a que se alterasse o modelo PCNN de forma a ser possível induzir e direccionar as várias segmentações, a serem produzidas, de acordo com uma referência de segmentação. Estruturas neurobiológicas que operam nesse sentido são observáveis no cérebro humano [1;2]. Este mecanismo possibilitaria ter um controlo e uma previsão mais adequada da segmentação a ser produzida pela PCNN. A utilização da cor no método híbrido representaria à partida uma melhoria na capacidade discriminadora do processo; mas devido aos problemas de complexidade de estrutura referentes à PCNN, já mencionados, a utilização da cor poderia piorar o desempenho computacional deste processo.

5.10 Classificação de leucócitos através de um modelo híbrido

O modelo de segmentação híbrida pode ser facilmente adaptado à classificação de regiões. No modelo segmentador existe um classificador que determina a pertença, ou não, de um determinado *pixel* à região de interesse, decidindo-se sobre uma das duas classes possíveis. Esta decisão é tomada a partir da relação que os pixels adjacentes ao pixel a classificar estabelecem. Na classificação de uma região, como no caso de um leucócito, os pixels já agrupados pelo método de segmentação constituem o padrão de

entrada para o classificador. O conceito funcional do modelo é em todo semelhante ao modelo usado na classificação. A principal diferença relaciona-se com a possibilidade de o número de classes a classificar ser agora maior do que dois. Este facto conduz a que se tenham de ter classificadores supervisionados treinados cada um deles, especificamente, para cada uma das classes de regiões consideradas. A razão de aplicar um modelo híbrido também neste problema é a mesma que levou à sua aplicação no problema da segmentação. A micro-segmentação executada na secção não supervisionada, no modelo híbrido, produz diversos aspectos padronizantes das regiões de interesse que permitem evitar um pouco o problema da métrica estática que suporta a decisão do que é importante e do que não é importante considerar numa região. Sem a percepção e o posterior entendimento de novas situações, o sistema pode falhar mais frequentemente. No problema da classificação de regiões a variabilidade em múltiplas direcções pode naturalmente surgir, como no caso do problema da segmentação, conduzindo a decisões erradas. A componente supervisionada tem assim uma tarefa mais simples, com resultados menos incertos, ao ter a cooperação da componente não-supervisionada.

Também neste caso a rede PCNN é responsável por obter um particionamento geométrico caracterizador da região a classificar. A assinatura obtida pelo decorrer das múltiplas iterações neuronais é numericamente codificada por descritores de Zernike. Estes, por sua vez, vão fornecer à rede supervisionada uma descrição mais independente das características altamente variantes das regiões em causa. A rede supervisionada deverá, nesta situação, ter apenas de aprender os traços fundamentais seleccionados e extraídos pela etapa não supervisionada que a antecede de cada classe em causa. A organização de blocos do modelo apresentado é mostrada na Figura 5.20. Como se pode visualizar, a decisão final passa por analisar cada uma das saídas dos classificadores SVM no bloco de *comparação*. A saída que tiver um valor mais elevado indicará a classe a que pertence a região em processamento. São consideradas ainda duas situações de não classificação. O comparador não indicará (a partir de um formato binário) qualquer classe, se existirem classes separadas por um valor inferior a κ ou se todas as classes tiverem um valor inferior a ρ . No primeiro caso, o comparador codifica a resposta num correspondente valor binário à interpretação de “Não tenho a certeza” e, no segundo caso, um outro valor binário corresponde à resposta de “Não sei”. O domínio numérico das saídas dos classificadores SVM situa-se no intervalo $[-1,1]$. Assim,

dentro deste intervalo, por uma análise da curva ROC, (Receiver Operating Characteristic), procurou-se encontrar os valores de κ e de ρ que minimizam o número de células classificadas erradamente. O valor final encontrado para κ foi o de 0.1 e o valor final encontrado para ρ foi o de -0.4. O procedimento foi feito avaliando, de forma independente, cada uma das curvas ROC de cada classe e conseqüentemente obtendo o valores de κ e de ρ para cada classe. Os valores finais de κ e de ρ foram obtidos pela média ponderada (relativamente ao número de instâncias de células existentes em cada classe) dos valores parciais de κ e de ρ de cada classe. No entanto, a forma mais correcta de se obterem esses valores seria pela análise conjunta das curvas ROC através de um método de optimização.

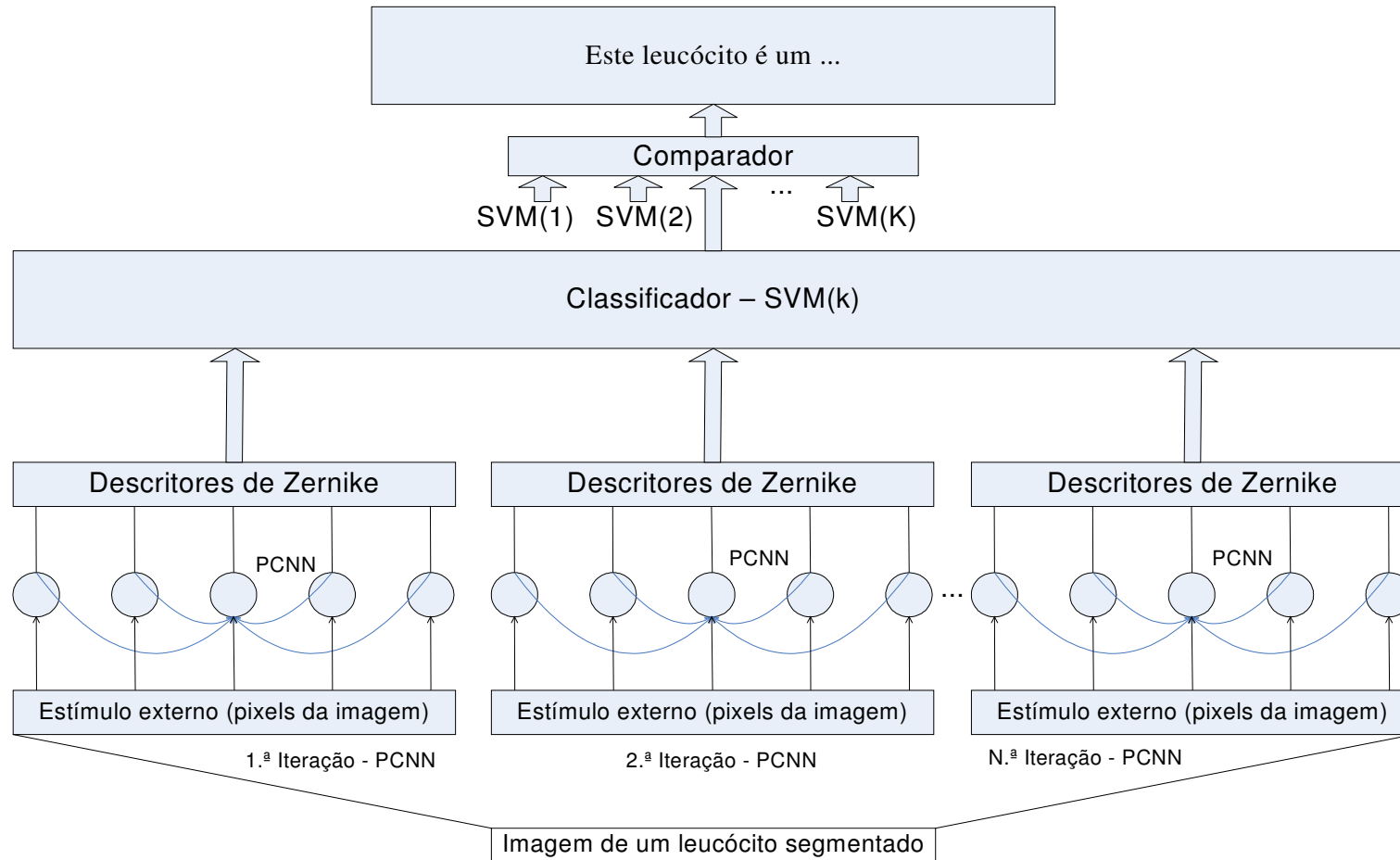


Figura 5.20 Composição do método de classificação por PCNN/SVM

Descrição das fases constituintes do treino para a classificação das regiões leucócito:

1. Percorrer a imagem de segmentações, pixel a pixel, numa progressão bem definida. Poderá ser da esquerda para a direita e de cima para baixo.
2. Após encontrar um primeiro pixel, assinalado pelo processo de segmentação e pertencente a uma célula, assinalar o contorno da célula na imagem de segmentação. O método usado para assinalar o contorno tem por base o processo de representação de contornos, *chain code*.
3. Assinalar/reassinalar todos os pixels internos ao contorno por um processo de preenchimento de regiões.
4. Os valores dos pixels são extraídos da região assinalada e são recolocados no plano bidimensional de entrada da PCNN. Cada região leucócito resultante é processada, de forma independente, pela PCNN em *i-iterações*.
5. As regiões binárias resultantes nas saídas da PCNN, ao longo de *i-iterações*, são descritas, numericamente, por um conjunto de momentos do descritor de regiões, o descritor de Zernike.
6. Para cada célula, a série que descreve a região tem um tamanho de *i-elementos* vezes o número de momentos do descritor usado. Cada uma destas séries corresponde a um padrão de treino para a SVM.
7. Normalização dos padrões de treino para um desvio padrão de uma unidade e para uma média de zero.
8. Cada padrão de treino é apresentado a um classificador de aprendizagem supervisionada, a SVM. Cada padrão está associado a um padrão de saída da SVM. Existe um número de SVMs igual ao número de classes de células a classificar. Cada SVM é treinada para que o padrão de treino na sua entrada tenha uma correspondência positiva ou negativa, na sua saída, de acordo com a classe de leucócitos associada ao respectivo classificador SVM.

9. O processo descrito no ponto 8 será repetido um número de vezes de acordo com as regras de convergência do classificador, de aprendizagem supervisionada, em causa.
10. Pela análise das curvas de ROC, da saída de cada classificador SVM, são determinados valores limiares internos, κ e ρ , do comparador, que permitem otimizar o número de identificações correctas.

Descrição das fases constituintes do pós-treino na classificação das regiões leucócito:

1. Percorrer a imagem de segmentações, pixel a pixel, numa progressão bem definida. Poderá ser da esquerda para a direita e de cima para baixo.
2. Após encontrar um primeiro pixel, assinalado pelo processo de segmentação e pertencente a uma célula, assinalar o contorno da célula na imagem de segmentação.
3. Assinalar/reassinalar todos os pixels internos ao contorno por um processo de preenchimento de regiões.
4. Os valores dos pixels são extraídos da região assinalada. Cada região leucócito resultante é processada, de forma independente, pela PCNN em *i-iterações*.
5. As regiões binárias resultantes nas saídas das PCNN, ao longo de *i-iterações*, são descritas, numericamente, por um conjunto de momentos do descritor de regiões, o descritor de Zernike.
6. Para cada célula, a série que descreve a região tem um tamanho de *i-elementos* vezes o número de momentos do descritor usado. Cada uma destas séries corresponde a um padrão a identificar.
7. Normalização do padrão a identificar para um desvio padrão de uma unidade e para uma média de zero.
8. Cada padrão a identificar é apresentado a um classificador de aprendizagem supervisionada, SVM, já treinado. Cada SVM

codificará uma resposta que corresponderá positivamente ou negativamente à classe de leucócitos associada ao respectivo classificador SVM.

9. Cada valor de saída de cada SVM será comparado com os valores de saída restantes das SVMs restantes.
10. O valor maior, entre todos, será indicador da classe de célula identificada. Se o maior valor de saída cair dentro dos intervalos de rejeição de identificação do comparador, este devolve uma codificação binária que indica essa situação.

5.11 Resultados do classificador híbrido

Tal como se fez no processo de segmentação, os resultados a seguir apresentados, para um maior esclarecimento da sua validade, serão confrontados com os resultados extraídos da utilização de um modelo unicamente supervisionado.

A parametrização determinada para a PCNN é dada na Tabela 5.IV. Os componentes usados e que estruturam todo o modelo de classificação híbrida são referidos na Tabela 5.V.

Tabela 5.IV Configuração da PCNN para a classificação híbrida

| Componentes da PCNN | Configuração |
|--|---------------|
| $Kernel M$ | 7x7; tipo 1/r |
| $Kernel W$ | 7x7; tipo 1/r |
| V_f | 0.5 |
| V_l | 0.8 |
| V_θ | 20.0 |
| β | 0.2 |
| $A_f (\alpha_f = -1/A_f)$ | 10.0 |
| $A_l (\alpha_l = -1/A_l)$ | 1.0 |
| $A_\theta (\alpha_\theta = -1/A_\theta)$ | 5.0 |

Tabela 5.V Caracterização dos componentes usados na classificação híbrida

| Componentes gerais | Configuração |
|--------------------------------|--|
| Tipo de pixel | Cinzento – 8 bits |
| Pixels considerados na entrada | Todos os pixels pertencentes a uma célula segmentada |
| Comprimento da série PCNN | 40 iterações |
| Descritores | Os 6 primeiros momentos de Zernike. – (0,0); (1,1); (2,0); (2,2); (3,1); (3,3) |
| Classificador supervisionado | SVM |
| Dimensão de entrada da rede | 480 – 40 iterações por 6 descritores de valores complexos |
| <i>Kernel</i> usado na SVM | Adatron – Radial Base Function |
| Número de SVMs | 5 (o número de classes de células que foram consideradas) |
| Número de padrões de treino | 276 |
| Valores limiares de rejeição | $\kappa = 0.10$; $\rho = -0.4$ |

5.11.1 Um modelo para comparação na classificação

Os resultados conseguidos na classificação usando o modelo híbrido são comparados com os resultados conseguidos com um modelo estritamente supervisionado, Figura 5.21. Neste modelo os elementos da imagem constituintes de uma região célula, previamente segmentada, são codificados por momentos de Hu. Esses valores constituem os padrões de treino e o padrão a classificar nas respectivas fases, a de treino e a de pós-treino. A rede neuronal supervisionada é treinada para que seja aprendida a relação entre os padrões descritores da região célula e uma codificação binária que está associada à classe de leucócito. A aprendizagem faz-se minimizando o erro entre os valores de correspondência anunciados pela rede neuronal e os valores de correspondência verdadeiros estipulados por um especialista. A operação de minimização está em incumbência de um método de *back-propagation* que tem como base a minimização pela direcção de descida do gradiente. Este tipo de abordagem, comparativamente à do modelo híbrido, é mais sensível à existência de novos padrões uma vez que a aprendizagem da rede faz-se no pressuposto de que insignificantes

diferenças aparecerão nos padrões futuros externos ao domínio de treino. Nos casos em que isso não acontece, a rede não tem uma capacidade, por falta de generalização, de *discernir* um novo entendimento sobre os novos tipos de imagem e os resultados são em sua maior parte um colapso na capacidade de acerto. Também neste modelo, o não híbrido, existe um mecanismo de comparação que, distinguindo cada saída da rede neuronal, decide qual a saída com a melhor hipótese de correcta indicação da classe de célula.

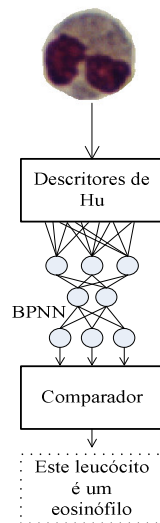


Figura 5.21 Modelo de classificação neuronal exclusivamente supervisionado através de uma rede *feed-forward* com um algoritmo de aprendizagem *backpropagation*

5.11.2 Resultados em tabelas de coincidência (confusion tables)

A Tabela 5.VI apresenta o número de correspondências entre a classe efectiva de leucócito (1.^a coluna) e a classe anunciada pelo sistema estritamente supervisionado (1.^a linha). A mesma estrutura é seguida na Tabela 5.VII, mas para o sistema de classificação híbrido. As imagens usadas foram, na íntegra, imagens de teste. Note-se que a dispersão de valores relativamente à diagonal principal é maior na Tabela 5.VI relativamente à Tabela 5.VII, indicando um melhor desempenho funcional do modelo híbrido.

Tabela 5.VI Tabela de coincidência de classificação de leucócitos usando um modelo estritamente supervisionado

| | Basófilo | Eosinófilo | Linfócito | Monócito | Neutrófilo | Incerto | Desconhecido |
|--------------|----------|------------|-----------|----------|------------|---------|--------------|
| Basófilo | 4 | | | 1 | | | |
| Eosinófilo | 1 | 38 | | 4 | 4 | 2 | 2 |
| Linfócito | | | 55 | 6 | 1 | 3 | |
| Monócito | 1 | 1 | 3 | 19 | 2 | 2 | 3 |
| Neutrófilo | 2 | 10 | | 4 | 61 | 8 | 4 |
| Incerto | 1 | | | 3 | | 12 | 2 |
| Desconhecido | | 4 | 1 | 3 | 2 | 4 | 22 |

Tabela 5.VII Tabela de coincidências de classificação de leucócitos usando um modelo híbrido

| | Basófilo | Eosinófilo | Linfócito | Monócito | Neutrófilo | Incerto | Desconhecido |
|--------------|----------|------------|-----------|----------|------------|---------|--------------|
| Basófilo | 4 | | | 1 | | | |
| Eosinófilo | 1 | 41 | | 2 | 4 | 1 | 2 |
| Linfócito | | | 57 | 5 | | 3 | |
| Monócito | 1 | | 1 | 22 | 2 | 2 | 3 |
| Neutrófilo | 1 | 7 | | 4 | 70 | 6 | 1 |
| Incerto | 1 | | | 2 | | 11 | 2 |
| Desconhecido | | 2 | 1 | 3 | 2 | 3 | 20 |

Na Tabela 5.VIII são mostradas as porcentagens de acertos sobre cada classe de leucócitos para o modelo estritamente supervisionado e para o modelo híbrido. Deve ser notado que alguns dos resultados, nomeadamente o relativo aos basófilos, são também função do escasso número de células disponível para teste.

Tabela 5.VIII Comparação de taxas de acertos para cada classe de célula entre o método supervisionado e o método híbrido

| | BPNN | PCNN-SVM | Número de imagens de teste |
|------------|--------|----------|----------------------------|
| Basófilo | 80.0 % | 80.0 % | 5 |
| Eosinófilo | 80.8 % | 85.4 % | 51 |
| Linfócito | 88.7 % | 91.9 % | 65 |
| Monócito | 73.0 % | 84.6 % | 31 |
| Neutrófilo | 79.2 % | 85.3 % | 89 |

Para quase todos os tipos de células houve um aumento de acertos no uso do modelo híbrido. Já que foram apenas usadas imagens não existentes no conjunto de padrões de treino, os resultados indicam uma maior capacidade de generalização do

método híbrido, que, é de relembrar, usa apenas imagens codificadas em níveis de cinzento.

5.12 Alteração dos campos receptores das PCNN para classificação de contornos

A Figura 5.7 apresenta imagens que são constituídas não por um contorno único, mas por um conjunto de contornos que, ao serem submetidos à acção de uma PCNN, despoletaram, cada um, uma *auto-onda* que foi interferir com as restantes. Esta interferência entre *auto-ondas* pode ser suficiente para caracterizar relações entre contornos complexos.

Se for necessário codificar, através de séries temporais, obtidas com PCNNs, utilizando a Equação (5.17), um contorno simples, tal como um triângulo ou uma circunferência, formado por pixels de iguais valores, a utilização de um *kernel* clássico ($1/r$), nas conexões de *Feeding* e de *Linking*, não permite diferenciar adequadamente zonas de contorno mediante a sua variação angular, uma vez que os elementos do *kernel* não variam angularmente. Isto significa que cada neurónio é insensível ao gradiente de direcção local do contorno e como os pixels têm todos o mesmo valor, todos os neurónios dispararão, ou inibirão, a sua actividade em simultâneo. Mesmo que os contornos sejam representados por níveis distintos, a série resultante não codificará, convenientemente, a informação relativa à geometria dos contornos; mas sim codificará a informação relativa à distribuição desses níveis de cinzento pelo contorno.

Para resolver o problema foi escolhida a manipulação do estímulo externo em oposição à escolha dos *kernels* dos *Feedings* ou dos *Linkings*, porque só aqui é que se pode ter a certeza da geometria de contorno a processar. Nas restantes partes, por efeito das *auto-ondas*, essa certeza não existe.

Como se pretende obter uma codificação que seja independente da posição rotacional, torna-se necessário adicionar ao neurónio PCNN vários *kernels* diferentes sobre os estímulos externos, mediante a morfologia de contornos locais que se pretende codificar. A combinação linear do resultado da convolução destes *kernels* com as imagens de contornos intervém no valor final de *Feeding* de cada neurónio. Além disso, cada *kernel* tem associado um coeficiente que garante uma diferenciação numérica, após convolução, entre as várias morfologias associadas a cada *kernel*. A adição deste

componente ao modelo do neurónio PCNN é mostrada na Figura 5.22. Com este modelo, é agravado o esforço computacional que um neurónio PCNN representa. Porém, se os pixels de fundo da imagem tiverem um valor bem conhecido, como no caso de imagens binárias, é apenas necessário fazer corresponder neurónios PCNN aos pixels activos que representam os contornos. Reduz-se assim, para a grande maioria de imagens de contornos, o esforço computacional final. Esse esforço é claramente menor do que o que seria necessário em caso de se usar o modelo não alterado do neurónio PCNN sobre todos os pixels da imagem.

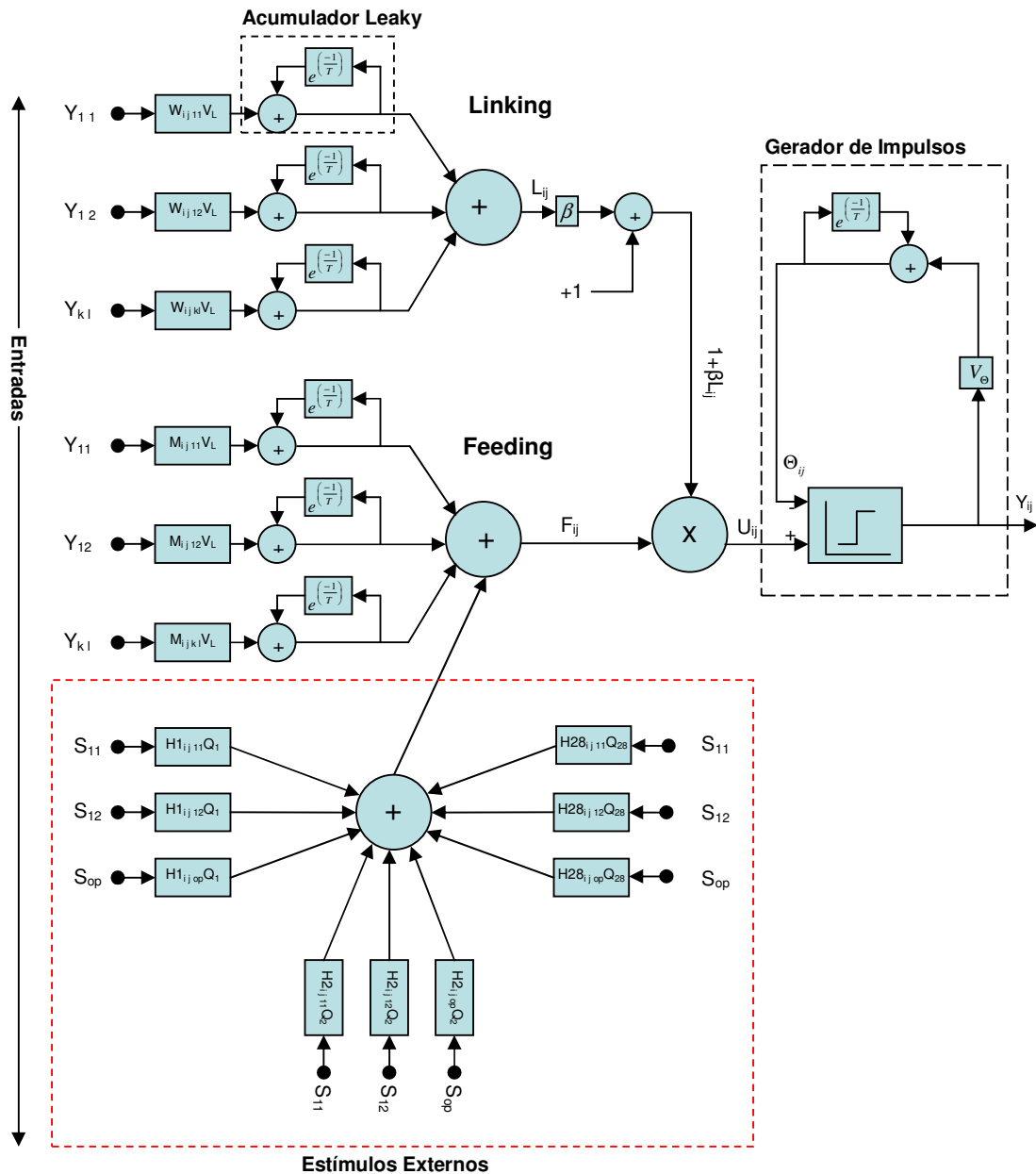


Figura 5.22 Modelo PCNN para codificação de contornos (representação sumária)

Os *kernels* (campos receptores) utilizados possuem 3x3 elementos e representam todas as configurações possíveis de descontinuidade angular de contornos, num total de 28 combinações. As configurações que detêm ângulos mais abruptos irão produzir valores de maior magnitude do que as que possuem ângulos mais suaves, de acordo com o coeficiente Q a que são sujeitas, Tabela 5.IX.

A Equação (5.1), que traduz o cálculo do valor de *Feeding*, de acordo com a alteração descrita nesta secção, é agora reescrita da seguinte forma:

$$F_{ij}[n] = e^{\alpha_F} F_{ij}[n-1] + \sum_{c=1}^{28} Q_c \sum_{op} H_{ijop}^c S_{op} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1], \quad (5.18)$$

onde H^c representa um dos 28 *kernels* da Tabela 5.IX.

Tabela 5.IX *Kernels* a aplicar aos estímulos externos para a codificação de contornos por PCNN

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|
| $Q = 0.125$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr></table> | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table> | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr></table> | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td></tr></table> | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $Q = 0.250$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td></tr></table> | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr></table> | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table> | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td></tr></table> | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $Q = 0.250$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table> | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>1</td><td>0</td><td>0</td></tr></table> | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr></table> | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | <table border="1" style="border-collapse: collapse; text-align: center;"><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr></table> | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Q = 0.750

| | | |
|---|---|---|
| 0 | 1 | 0 |
| 0 | 1 | 1 |
| 0 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 1 | 0 |
| 0 | 1 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 0 | 1 | 0 |

| | | |
|---|---|---|
| 0 | 1 | 1 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |

Q = 0.750

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |

| | | |
|---|---|---|
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 0 |
| 0 | 1 | 1 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |

Q = 1.000

| | | |
|---|---|---|
| 1 | 1 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 1 | 0 |
| 1 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 0 | 0 | 1 |

| | | |
|---|---|---|
| 0 | 1 | 1 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |

Q = 1.000

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 1 | 0 |

| | | |
|---|---|---|
| 1 | 0 | 0 |
| 1 | 1 | 0 |
| 0 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 0 | 0 | 0 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 1 | 1 |

Estes *kernels* foram pensados para processar contornos com uma espessura igual a um pixel. Se os contornos tiverem uma largura superior, para caracterizá-los, poderá ser suficiente a interferência entre *auto-ondas* que desse facto emergem. Porém, a robustez da codificação poderá ser incerta. Outra abordagem possível é a de se usar um processo de *thinning* [51] para reduzir a espessura do contorno a um pixel.

5.12.1 O problema do *aliasing*

Como se trata de contornos com a espessura de um pixel, devido a efeitos de *aliasing*, a representação destes não consegue ser totalmente independente da rotação. Este facto, à partida, pode ser minimizado utilizando técnicas de *anti-aliasing* sobre o contorno. O que se verificou ao usar-se estas técnicas foi que a codificação era menos robusta, ao ser menos discriminante. Isto porque a detecção angular produzida nos

kernels está prevista para contornos de espessura de 1 pixel e os filtros de *anti-aliasing* [52] usados alteram a espessura do contorno.

5.13 Resultados da aplicação das PCNN à codificação e classificação de contornos

As imagens da Figura 5.24 e da Figura 5.25 representam algumas amostras das saídas da PCNN ao longo de 50 iterações sobre o processamento à imagem da Figura 5.23.

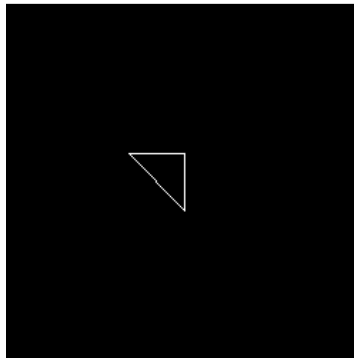


Figura 5.23 Triângulo rectângulo a ser processado pela PCNN de contornos

A parametrização determinada para a PCNN foi a seguinte (Tabela 5.X):

Tabela 5.X Configuração da PCNN para a codificação de contornos

| Componente | Configuração |
|--|----------------------|
| <i>Kernel M</i> | 3x3; tipo <i>l/r</i> |
| <i>Kernel W</i> | 3x3; tipo <i>l/r</i> |
| <i>V_f</i> | 0.65 |
| <i>V_l</i> | 0.9 |
| <i>V_θ</i> | 18.0 |
| <i>β</i> | 0.15 |
| <i>A_f</i> ($\alpha_f = -1/A_f$) | 15.0 |
| <i>A_l</i> ($\alpha_l = -1/A_l$) | 2.0 |
| <i>A_θ</i> ($\alpha_\theta = -1/A_\theta$) | 3.0 |

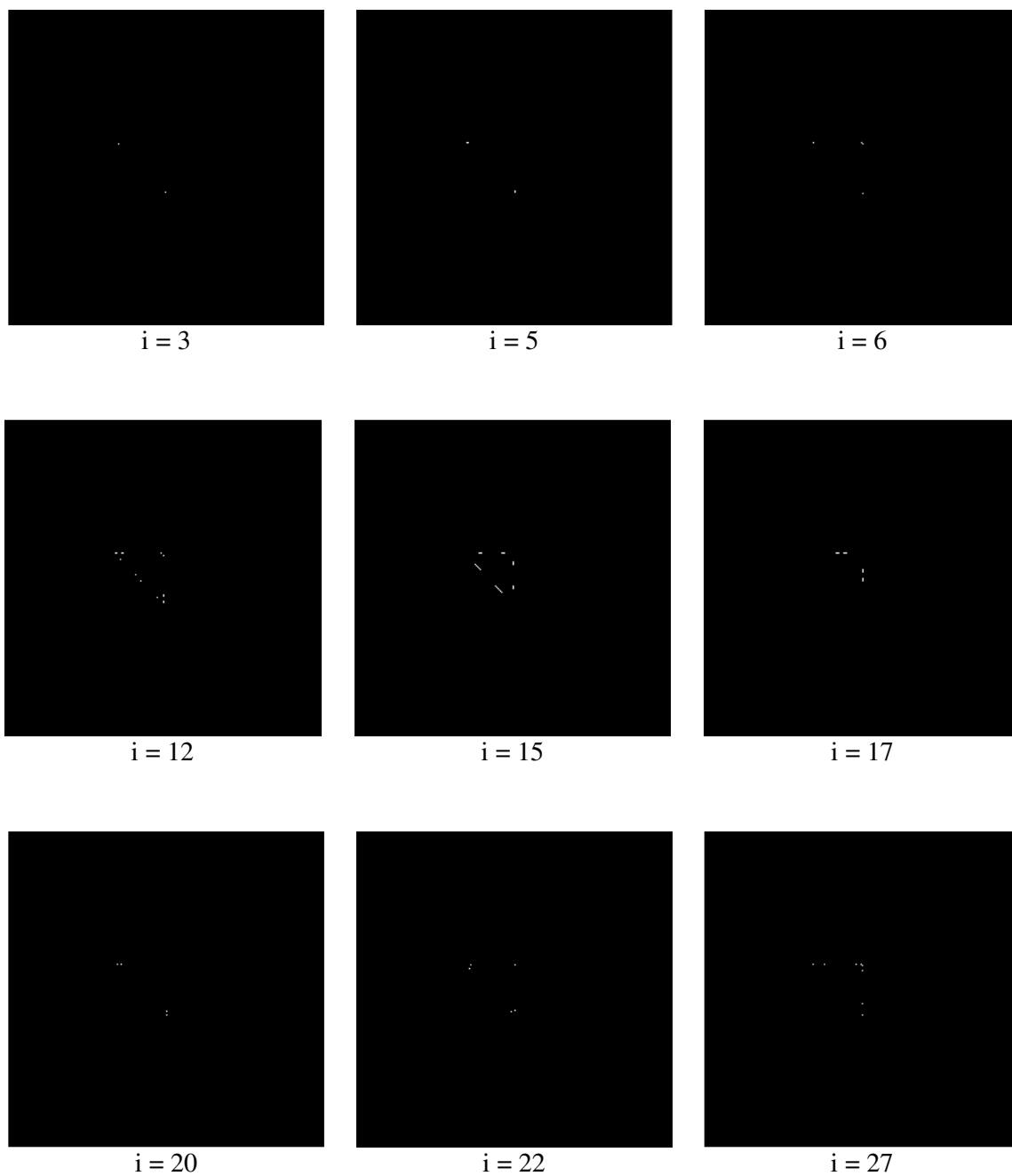


Figura 5.24 Algumas amostras das saídas da PCNN ao longo de 50 iterações resultantes do processamento à imagem da Figura 5.23. O i representa o número da iteração PCNN

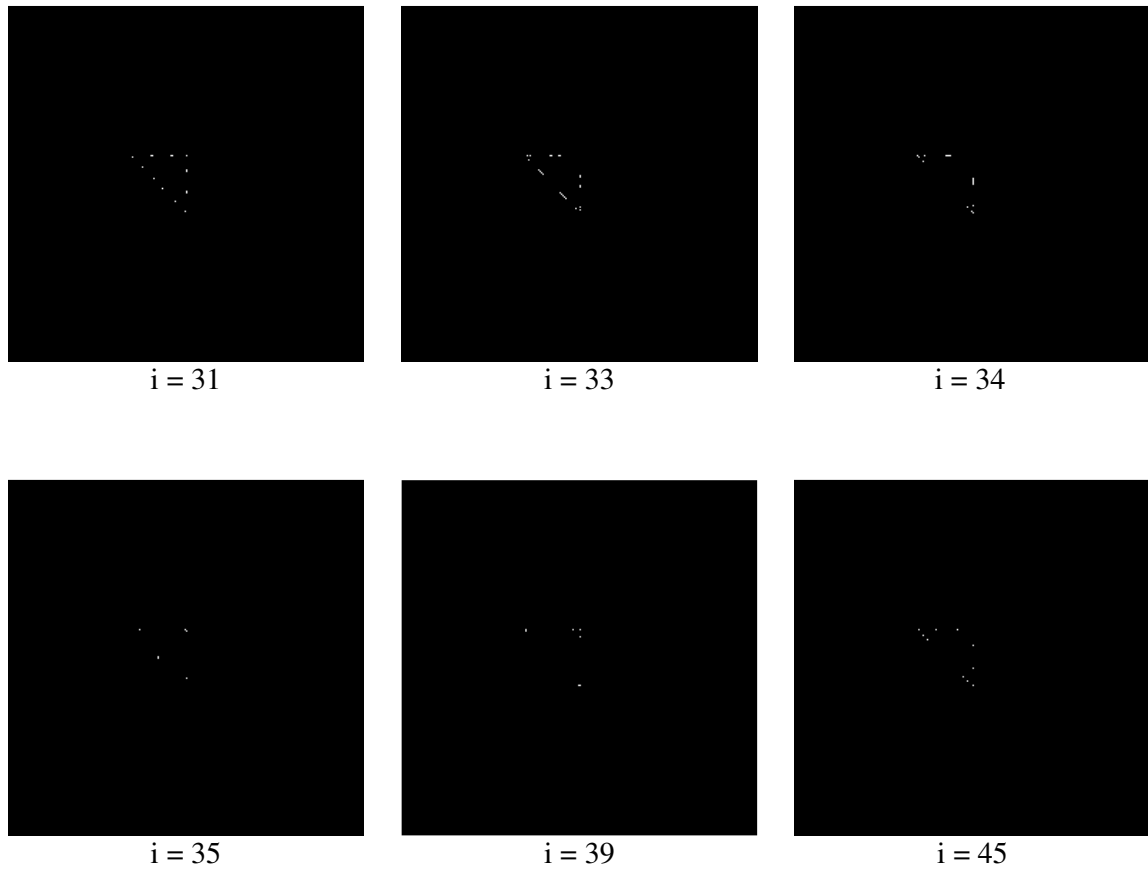


Figura 5.25 Algumas amostras das saídas da PCNN ao longo de 50 iterações resultantes do processamento à imagem da Figura 5.23. O i representa o número da iteração PCNN

É observável nas imagens da Figura 5.24 e da Figura 5.25 a forma como a partição do contorno de um triângulo é apresentada pela PCNN de iteração para iteração. Repare-se que as zonas de descontinuidade angular tendem a prevalecer num maior número de iterações do que as zonas de contornos lineares. O importante é conseguir diferenciar os diversos aspectos do contorno. Os *kernels* escolhidos, para aplicar aos estímulos externos, fazem-no com maior magnitude nas zonas de maior descontinuidade. Daí que os neurónios dessas zonas fiquem activos com maior frequência.

5.13.1 Séries temporais resultantes da aplicação das PCNN aos contornos

Utilizando a Equação (5.17) foram obtidas séries que representam, para cada iteração da PCNN, o número de neurónios activos. O comprimento da série foi escolhido empiricamente através da observação do *período* aproximado da série. Estas imagens apresentaram séries que detêm uma assinatura sensivelmente repetida a partir da quinquagésima iteração. Este facto levou a que fosse escolhido um comprimento de 50 iterações.

As três séries visíveis no gráfico da Figura 5.26 revelam as diferenças e as igualdades entre as assinaturas de um contorno que é um triângulo, de um contorno que é um quadrado e de um contorno que é um rectângulo. A assinatura do triângulo é amplamente não coincidente com as duas restantes. A assinatura do quadrado e do rectângulo apresentam variações semelhantes, em alguns segmentos de iterações, embora a fase e a amplitude sejam ligeiramente distintas. Isto é coerente com as coincidências e com as diferenças geométricas existentes entre estas figuras.

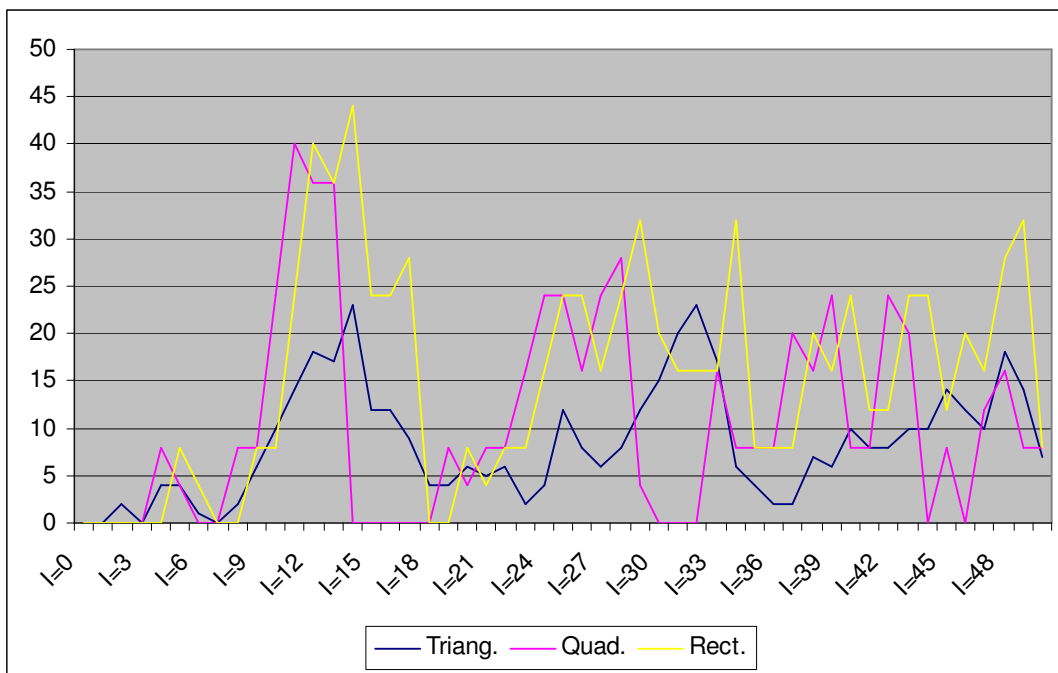


Figura 5.26 Séries temporais resultantes da aplicação da PCNN aos contornos de um triângulo, de um quadrado e de um rectângulo, todos à mesma escala

Na Figura 5.27 o gráfico mostra duas séries temporais correspondentes a dois rectângulos de diferentes escalas. Embora existam segmentos das séries com coincidência perfeita, outros não a possuem. Este facto pode ser indicador que o processo usado não evidencia uma invariabilidade à mudança de escala. Uma justificação para este caso poderá relacionar-se com a discretização da imagem, que no caso de escalas mais pequenas deforma a essência geométrica da figura. De facto, as experiências que foram feitas entre figuras com escalas diferentes, mas com ambas as escalas relativamente elevadas, mostraram séries de maior coincidência.

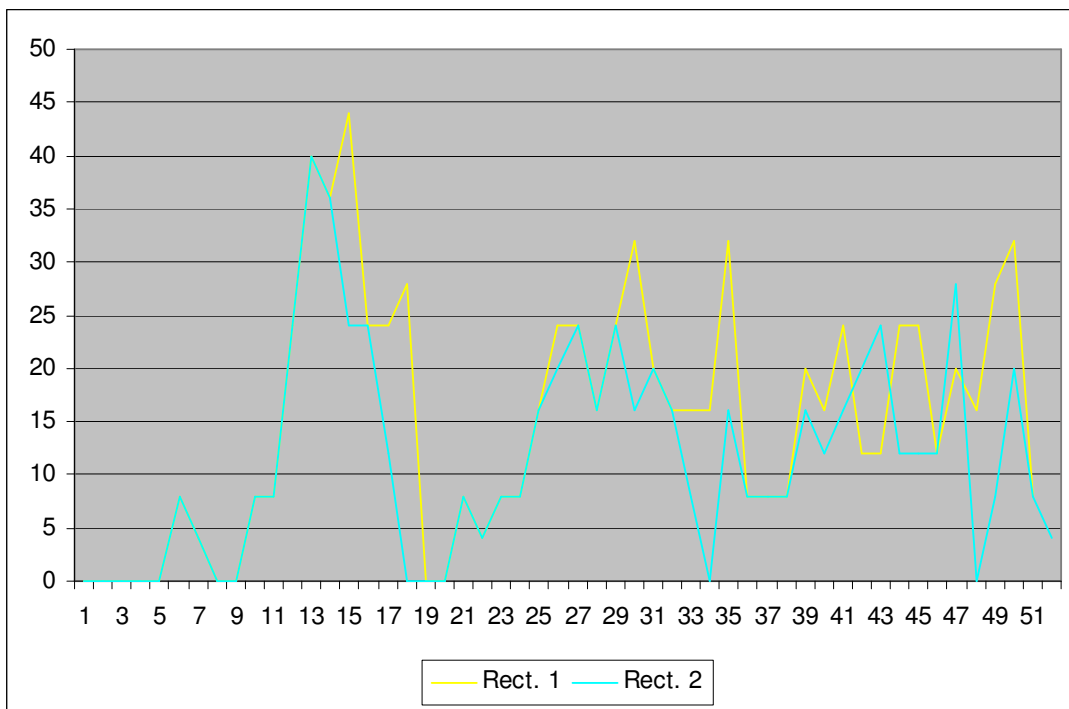


Figura 5.27 Séries temporais resultantes da aplicação da PCNN aos contornos de dois rectângulos com escalas diferentes

A discrepância entre séries verificada na Figura 5.28 tem especialmente um motivo. O eixo principal do rectângulo 1 encontra-se, relativamente aos eixos cartesianos da imagem, numa posição de 15 graus. As arestas que o constituem, devido ao fenómeno de *aliasing*, são formadas por descontinuidades angulares constantes. Estas descontinuidades provocam disparos mais frequentes nos neurónios PCNN.

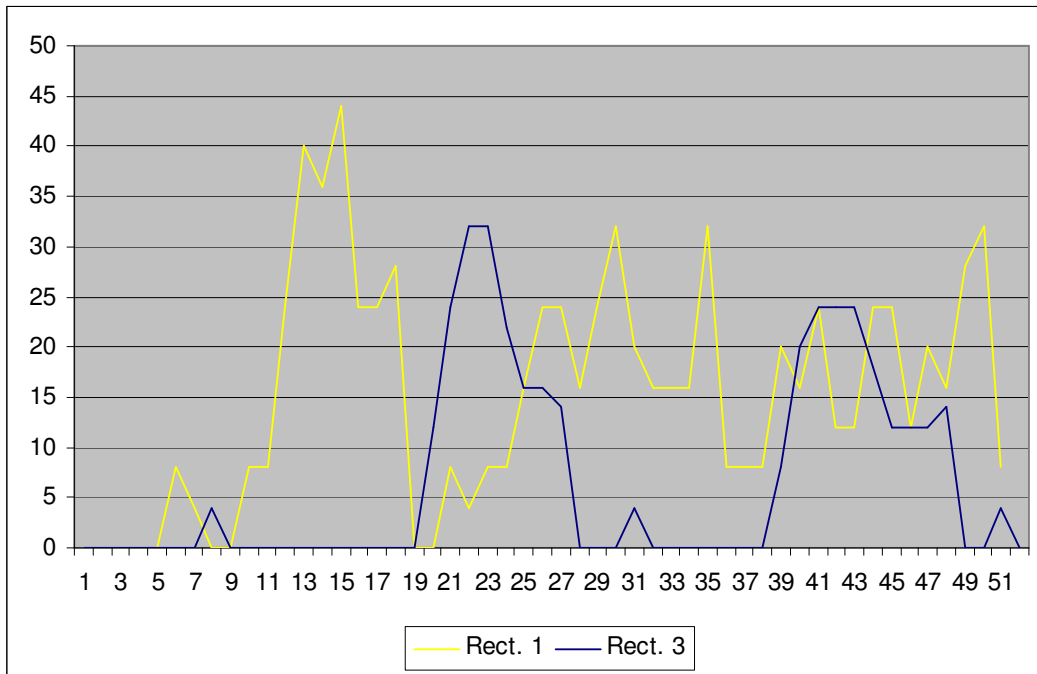


Figura 5.28 Séries temporais resultantes da aplicação da PCNN aos contornos de dois rectângulos com posições angulares diferentes

5.13.2 Desempenho do método de classificação de contornos do híbrido PCNN-SVM

As séries resultantes das saídas das PCNNs ao longo de várias iterações foram apresentadas às entradas de um classificador SVM. A estrutura funcional deste mecanismo é no essencial semelhante à que foi apresentada para a classificação das regiões na secção 5.10. A maior diferença situa-se na substituição do modelo original da rede PCNN pelo modelo PCNN orientado aos contornos.

Foi associado um classificador SVM a cada uma das três figuras que se determinou classificar. Uma SVM detecta um triângulo, outra SVM detecta um quadrado e outra detecta um rectângulo.

O valor de saída de cada uma das SVMs é comparado com os restantes. Aquele que tiver uma maior magnitude sugerirá a classe a que pertence a figura em processamento. As excepções a esta situação são geridas pelas margens de rejeição que já foram explicadas e usadas no processo de classificação de regiões e que se relacionam com a classificação de “Incerto” e de “Desconhecido”.

Para comparação usou-se um método de classificação de contornos por um código de cadeia relativo modificado¹¹ [53]. Também se usou o modelo original das PCNN (sem modificação) para esse mesmo objectivo. Assim, o esquema do sistema classificador de contornos, relativo à segunda técnica de comparação, é idêntico ao acima explicado. Cada imagem de treino e de teste integra uma representação de uma das três figuras consideradas. Essa representação detém variação relativamente à posição, à rotação e à escala dessas figuras.

5.13.2.1 Resultados em tabelas de coincidência (confusion tables)

Os resultados obtidos, e que podem ser constatados da Tabela 5.XI à Tabela 5.XIV, indicam que para este tipo de figuras a PCNN modificada consegue produzir taxas de acerto mais elevadas, superior a 95%, do que os dois outros métodos opositores.

Tabela 5.XI Tabela de coincidências de classificação de figuras geométricas usando o código da cadeia modificado

| | Triângulo | Quadrado | Rectângulo | Incerto | Desconhecido |
|--------------|-----------|----------|------------|---------|--------------|
| Triângulo | 36 | 1 | 2 | 1 | 0 |
| Quadrado | 0 | 35 | 3 | 2 | 0 |
| Rectângulo | 2 | 2 | 34 | 2 | 0 |
| Incerto | 0 | 0 | 0 | 0 | 0 |
| Desconhecido | 0 | 0 | 0 | 0 | 0 |

Tabela 5.XII Tabela de coincidências de classificação de figuras geométricas usando uma PCNN original e três SVM

| | Triângulo | Quadrado | Rectângulo | Incerto | Desconhecido |
|--------------|-----------|----------|------------|---------|--------------|
| Triângulo | 20 | 6 | 10 | 3 | 1 |
| Quadrado | 8 | 15 | 12 | 3 | 2 |
| Rectângulo | 6 | 17 | 15 | 2 | 0 |
| Incerto | 0 | 0 | 0 | 0 | 0 |
| Desconhecido | 0 | 0 | 0 | 0 | 0 |

¹¹ Este processo é invariante à rotação e à escala e possui um classificador que se baseia no erro quadrático médio para encontrar o contorno com a rotação de maior coincidência sobre a figura em análise. A independência à escala é conseguida pela normalização das curvas *ângulo-perímetro*. Estas curvas retratam a variação angular do contorno em função da variação posicional de cada elemento do contorno.

Tabela 5.XIII Tabela de coincidências de classificação de figuras geométricas usando uma PCNN modificada e três SVM

| | Triângulo | Quadrado | Rectângulo | Incerto | Desconhecido |
|--------------|-----------|----------|------------|---------|--------------|
| Triângulo | 39 | 0 | 1 | 0 | 0 |
| Quadrado | 0 | 38 | 2 | 0 | 0 |
| Rectângulo | 0 | 1 | 39 | 0 | 0 |
| Incerto | 0 | 0 | 0 | 0 | 0 |
| Desconhecido | 0 | 0 | 0 | 0 | 0 |

Tabela 5.XIV Comparação de taxas de acertos para cada classe de figura entre os três métodos testados

| | Código da cadeia modificado | PCNN-SVM | PCNN mod.- SVM | N.º de imagens de teste |
|------------|-----------------------------|----------|----------------|-------------------------|
| Triângulo | 92.3 % | 55.5 % | 97.5 % | 40 |
| Quadrado | 92.1 % | 42.8 % | 95.0 % | 40 |
| Rectângulo | 89.5 % | 39.5 % | 97.5 % | 40 |

A codificação de contornos a partir do código da cadeia modificado mostrou também bons níveis de acerto, com uma taxa média ligeiramente acima dos 90%. O método que faz uso da PCNN original consolida com estes resultados a inadequação prevista pela análise do funcionalismo das PCNNs relativamente à codificação de contornos. Por outro lado, já era esperado um bom desempenho associado ao método do código da cadeia relativo, pois este codifica de forma determinística e coerente os contornos que lhe são sujeitos. Os *kernels* atribuídos aos *estímulos* da PCNN modificada seguem o mesmo conceito codificador que o código da cadeia modificado. No entanto, o superior nível de acerto por parte da PCNN modificada pode ser atribuído principalmente a um factor. A característica de *auto-onda* da PCNN possibilita gerar interferências entre os diferentes níveis de energia, devolvidos pelos *kernels*, sobre os diferentes elementos do contorno. Isto permite a combinação de informação a várias escalas, aumentando o grau de diferenciação codificante entre figuras. Este facto acaba também por atenuar a problemática do *aliasing*, que no caso do método do código da cadeia relativo conduz à perda de capacidade diferenciadora na sua codificação.

5.13.3 Desempenho do método de classificação de contornos do híbrido PCNN-SVM em imagens naturais

Foi também analisado o desempenho deste método usando imagens naturais de células sanguíneas humanas.

A análise foi feita submetendo a processamento as imagens de contornos das células. Neste caso, cada imagem contém uma única célula e através de operadores de *Sobel* são extraídos, binariamente, os contornos da célula e de seguida são reduzidos à espessura de um pixel por um método de *Thinning* [54]. As imagens com estes contornos são sujeitas aos métodos de descrição por PCNN modificada e são sujeitas ao método de descrição pelo código da cadeia modificado. Um dos factores de que depende o método de determinação dos contornos é o contraste da imagem. As células apresentam uma textura granular distribuída pela sua superfície. Dependendo do contraste da imagem, os contornos associados a essa granularidade aparecerão, ou não, adulterando os contornos característicos das células. Por exemplo, na imagem esquerda da Figura 5.29, os contornos característicos da célula são minimamente bem definidos, Figura 5.30, em oposição à precária caracterização dos contornos da imagem da direita.



Figura 5.29 Células com contrastes diferentes



Figura 5.30 Contornos esqueletizados de células com contrastes diferentes

Este problema condiciona o desempenho final de todo o método aqui apresentado. Isto conduziria a que fosse proposta uma solução de diminuição desta dificuldade. Optou-se, apenas, por avaliar o desempenho do método sobre imagens de contornos seleccionadas adequadamente de forma a minimizar o citado problema.

Na Figura 5.31 são mostradas as séries temporais que resultaram do processamento das PCNNs aos contornos de três células: dois eosinófilos e um neutrófilo.

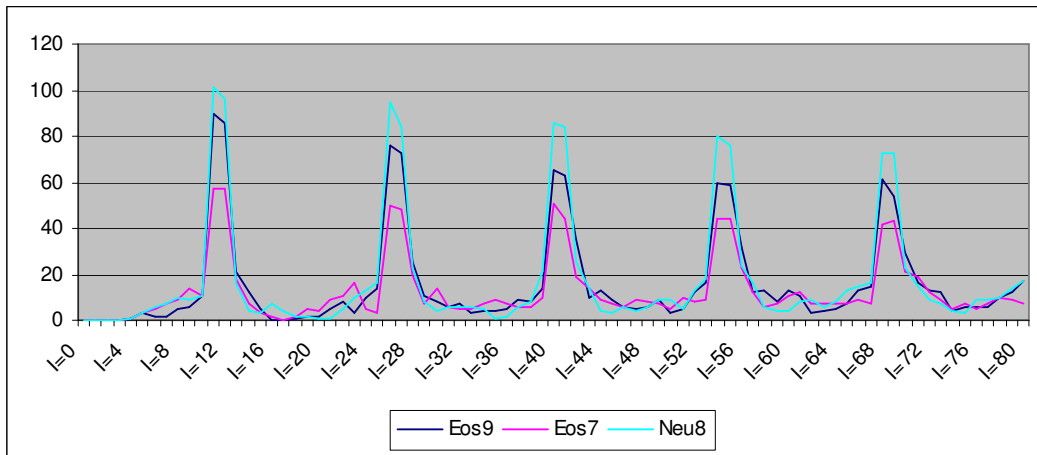


Figura 5.31 Séries temporais resultantes da aplicação da PCNN aos contornos de três células

As assinaturas são muito idênticas, variando essencialmente as amplitudes entre séries. Ao contrário das figuras geométricas, a composição destes contornos, de célula para célula, é maioritariamente variante em ângulo, provocando o disparo dos neurónios em instantes semelhantes de célula para célula, pois a energia nos estímulos é quase a mesma. Estas dificuldades, embora reflectindo-se de maneira diferente, são também reveladas por outros métodos, como o método de comparação aqui usado, o código da cadeia modificado. Este facto coloca-os no mesmo nível de partida, permitindo considerar as diferenças de desempenho final entre os dois métodos.

Para se diminuir esse problema, inclusive o problema do *aliasing*, seria necessário criar campos receptores (*kernels*) com variadíssimas configurações geométricas em escalas diferentes. Essa abordagem seria semelhante à abordagem existente nas diferentes camadas do córtex visual humano [2]. Computacionalmente essa solução é para já incomportável.

5.13.3.1 Resultados em tabelas de coincidência (confusion tables)

Os resultados obtidos podem ser constatados da Tabela 5.XV à Tabela 5.XVII.

Tabela 5.XV Tabela de coincidências de classificação de células sanguíneas usando o código da cadeia modificado

| | Basófilos | Eosinófilos | Linfócitos | Monócitos | Neutrófilos | Incerto | Desc. |
|-------------|-----------|-------------|------------|-----------|-------------|---------|-------|
| Basófilos | 2 | | 2 | 1 | | | |
| Eosinófilos | 2 | 29 | 1 | 2 | 9 | 2 | 3 |
| Linfócitos | 3 | | 43 | 10 | 2 | 3 | 2 |
| Monócitos | 5 | 1 | 3 | 10 | 2 | 4 | 5 |
| Neutrófilos | 1 | 12 | 3 | 6 | 49 | 10 | 4 |
| Incerto | 1 | 1 | | 2 | | 9 | 2 |
| Desconhec. | 2 | 2 | 4 | 4 | 3 | 4 | 12 |

Tabela 5.XVI Tabela de coincidências de classificação de células sanguíneas usando PCNNs orientadas ao contorno e SVMs

| | Basófilos | Eosinófilos | Linfócitos | Monócitos | Neutrófilos | Incerto | Desc. |
|-------------|-----------|-------------|------------|-----------|-------------|---------|-------|
| Basófilos | 3 | | 1 | 1 | | | |
| Eosinófilos | 2 | 35 | 1 | 2 | 4 | 2 | 2 |
| Linfócitos | 2 | | 48 | 6 | 2 | 3 | 2 |
| Monócitos | 2 | 0 | 2 | 17 | 2 | 4 | 3 |
| Neutrófilos | 0 | 9 | 2 | 4 | 61 | 6 | 3 |
| Incerto | 1 | 1 | | 2 | | 9 | 2 |
| Desconhec. | 1 | 3 | 3 | 5 | 4 | 5 | 8 |

Tabela 5.XVII Comparação de taxas de acertos para cada classe de célula entre os dois métodos testados

| | Código da cadeia modificado | PCNN mod-SVM | Número de imagens de teste |
|------------|-----------------------------|--------------|----------------------------|
| Basófilo | 40.0 % | 60.0 % | 5 |
| Eosinófilo | 60.4 % | 72.9 % | 48 |
| Linfócito | 68.2 % | 76.2 % | 63 |
| Monócito | 33.3 % | 56.6 % | 30 |
| Neutrófilo | 57.6 % | 71.7 % | 85 |

Foi usado o mesmo grupo de imagens que serviu para avaliar as estruturas classificadoras orientadas à região. No entanto, nem todas as imagens desse grupo puderam ser usadas nos processos orientados aos contornos, devido aos problemas já referidos no que refere à extração de contornos. Mesmo assim, é visível uma diminuição geral do desempenho, se a comparação for feita sobre os métodos orientados às regiões. O método que codifica os contornos com base no código da cadeia relativo

tem um desempenho consideravelmente menor do que o método que utiliza a PCNN modificada. Este facto pode uma vez mais ser atribuído ao efeito produzido pela *auto-onda* das PCNNs, que permite relacionar os vários aspectos dos contornos codificando-os ao longo das múltiplas iterações neuronais. O desempenho conseguido pela PCNN modificada não é tão elevado como o que foi conseguido pelos processos orientados às regiões. A razão para que isto aconteça reside na menor quantidade de informação que está normalmente associada aos contornos de entidades existentes em imagens naturais em oposição à maior informação que pode ser extraída se essas entidades estiverem representadas por regiões. Por outro lado, em imagens naturais, a variação angular dos constituintes dos contornos, à sua escala mais baixa, é muito irregular, independentemente do tipo de contorno. Isto leva a que as máscaras usadas na PCNN modificada retornem constantemente energias elevadas, o que diminui a capacidade de produzir padrões diferenciáveis entre os contornos. A incapacidade, já mencionada, de se obter uma extracção correcta e coerente de contornos em imagens naturais usando mecanismos primários de extracção de contornos é outra fundamentação.

5.14 Modificação dos Kernels da PCNN com imagens naturais

Com base nos problemas de arbitrariedade segmentadora da PCNN, tentou-se alterar a distribuição de valores pelos elementos das matrizes M e W (*kernels*) com o objectivo de induzir algum direccionamento ao processamento da rede PCNN. Foram experimentados vários níveis de composição semântica visual. Os níveis mais baixos possuíam *kernels* com detalhes geométricos mais decompostos. Nos níveis mais elevados os próprios *kernels* eram formados por elementos trazidos de imagens de células em níveis de cinzento. No global, não se verificaram significativas melhorias relativamente à utilização do *kernel* $1/r$. Foi mesmo observado, em algumas situações, um diminuir na capacidade de segmentação da rede PCNN. No entanto, como seria esperado, no caso das imagens que eram constituídas com padrões, em geometria ou em luminância global, semelhantes ao padrão do *kernel*, o desempenho era ligeiramente melhor.

A justificação para este facto relaciona-se com o factor dimensional do *kernel* de convolução, com as características do próprio valor resultante da convolução e com a forma de disparo do neurónio PCNN. Um *kernel* formado por elementos

correspondentes a uma região de interesse é normalmente um *kernel* de dimensões relativamente elevadas. O valor que resulta da convolução acaba por ser um valor representativo médio de toda a área abrangida pelo *kernel*. O instante de disparo do neurónio provém da comparação dos valores da convolução com um valor limiar que varia no tempo. Quer isto dizer que o neurónio dispara independentemente de o valor de convolução ultrapassar o valor limiar por uma grande margem ou uma pequena margem. Estes factores conjugados levam ao pouco desempenho apresentado por estes *kernels*.

5.15 Conclusões

As técnicas analisadas e algumas propostas partilham todo um mesmo princípio para atingirem o objectivo de funcionamento com maior generalização. Esse princípio defende a existência de uma componente não-supervisionada que coopera com uma componente supervisionada. A componente não supervisionada evita que certas características, não essenciais na classificação, mas essenciais na diferenciação, sejam utilizadas pelo bloco supervisionado, o que eleva a capacidade de generalização. A utilização de redes PCNN para preencher o bloco não supervisionado mostrava-se à partida muito promissora, dado o seu modo autónomo, fragmentador de regiões, ser direccionado a características geométricas fundamentais das imagens. No domínio do bloco supervisionado, a escolha de classificadores SVM deveu-se ao seu mecanismo de aprendizagem ultrapassar alguns dos problemas, como por exemplo a retenção em mínimos locais, patentes em outros métodos supervisionados. Foram propostos arranjos para retirar cooperativamente as melhores funcionalidades dos dois blocos. Apresentou-se uma estratégia de natureza segmentadora e várias de natureza classificadora. No caso desta última foi constituída uma abordagem orientada às regiões e outra orientada aos contornos.

O desempenho apresentado pelo segmentador híbrido mostrou-se ligeiramente superior, em termos de generalização, relativamente aos modelos unicamente supervisionados. Todavia, além de este nível de desempenho não ter sido relativamente muito elevado, outros factores, como o desempenho computacional, revelaram a dificuldade de se usar um modelo híbrido de base PCNN. O automatismo segmentador da PCNN revelou também precariedades, ao não ser possível, no seu modelo original, direccionar de forma precisa o tipo de características a ter como base na segmentação.

Complementaridades ao modelo seriam necessárias para garantir esse cumprimento. Factores de realimentação informativa nessa vertente deveriam ser considerados. Embora a tarefa do classificador seja inerentemente mais simples do que a do segmentador, o classificador híbrido de base PCNN sofreu também das mesmas dificuldades trazidas pela PCNN. Mostrou sensível melhoria no desempenho generalizador, mas acabou também por ser imprevisível em situações que não estavam consideradas no conjunto de padrões de treino. A abordagem classificadora orientada aos contornos mostrou um comportamento favorável na codificação de figuras geométricas, mas não na codificação de contornos derivados de imagens naturais. Tudo isto mostra que, para executar qualquer aproximação generalista, aos níveis segmentadores e classificadores, idêntica aos conseguidos pelo córtex visual humano, é necessário aumentar exponencialmente a complexidade do sistema. No entanto, a este domínio clássico, não quântico, os indicadores retirados parecem subscrever a vantagem de se fazer a junção entre bloco não supervisionado e bloco supervisionado. Foi observada ainda a necessidade de conjugar estes mecanismos quase de uma forma difusa, não só numa direcção de sinais, dos mecanismos não supervisionados para os supervisionados, mas também no sentido contrário, para que os últimos sejam directores das actividades dos primeiros.

Capítulo 6

Aceleração Quântica em Processamento de Imagem

A necessidade de aumentar a complexidade dos métodos computacionais para produzir melhorias de desempenho funcional, nomeadamente em processos de visão artificial, conduz a que se procure encontrar suportes físicos que adequam o aumento da complexidade ao aumento da execução temporal do processo. O capítulo actual descreve a teoria elementar que sustenta a computação quântica. Este tipo de computação permite atingir, para certos problemas, desempenhos temporais inigualáveis em computação clássica. Recorrendo ao algoritmo quântico de Grover é demonstrada a implementabilidade de técnicas de processamento de imagem em suportes quânticos.

6.1 Introdução

Os conceitos de computação quântica foram primordialmente sugeridos, de forma independente, por Paul Benioff [55-58] e por Richard Feynman [59-60] no início da década de 80. Estes investigadores, ao observarem que para simular o comportamento das partículas quânticas, em computadores clássicos, era necessário um poder computacional que crescia exponencialmente com o número de partículas, sugeriram que o aproveitamento das características funcionais do universo quântico pudesse ser usado para resolver problemas de ordem temporal e espacial complexa. De facto, para simular um processo quântico de 100 *qubits* num computador clássico é necessário armazenar 2^{100} números complexos [61]. Esses mesmos 100 *qubits*,

processados por um conjunto de portas quânticas adequadas, resultam numa equivalência de 2^{100} operações quando realizadas num computador clássico [61]. Apesar das características invulgares do comportamento quântico, convém realçar que um computador quântico não tem a capacidade de solucionar problemas que não têm solução num computador clássico. O elemento básico da computação quântica é o *qubit*. Os *qubits* possuem uma concreta similaridade com os *bits* clássicos, uma vez observados: ou valem 0 ou valem 1. Os *qubits* encontram-se na natureza em múltiplas formas. Podem representar o estado de diversas partículas quânticas, como o fóton, em que o *qubit* pode ser estabelecido pela sua polarização, vertical ou horizontal, ou como o núcleo dos átomos, em que o *qubit* pode ser estabelecido pelo seu *spin* magnético, *up* ou *down*. Além destas, outras propriedades quânticas e outras partículas quânticas podem representar um *qubit*. A computação quântica consiste em aplicar determinadas transformações unitárias a conjuntos de *qubits*, ou individualmente, de forma a serem obtidos cálculos de específico interesse. As transformações adquirem forma física por intermédio de portas quânticas de natureza diversa; podem ser ópticas, magnéticas, entre outras. Por exemplo, no caso dos *qubits* representados pelos momentos magnéticos dos núcleos atómicos, podem ser manipulados e observados através de sistemas de ressonância magnética nuclear [62].

Neste capítulo será apresentada uma abordagem para acelerar técnicas de processamento de imagem através de computação quântica.

6.2 Notação de Dirac

Os estados quânticos e as transformações que actuam sobre eles podem ser descritos através de vectores e matrizes ou então, de forma conveniente e condensada, através da chamada notação *bra/ket* proposta por Dirac [63].

Kets, $|x\rangle$, representam vectores coluna e são usados para descrever estados quânticos. O seu complemento *bra*, $\langle x|$, representa a transposta da conjugada (adjunta) de $|x\rangle$. A base ortogonal $\{|0\rangle, |1\rangle\}$ pode ser expressa como $\{(1,0)^T, (0,1)^T\}$. Qualquer combinação linear complexa de $|0\rangle$ e de $|1\rangle$, $a|0\rangle + b|1\rangle$, pode ser apresentada como $(a,b)^T$. A escolha da ordem da base vectorial é arbitrária. Representar $|0\rangle$ por $(0,1)^T$ e

representar $|1\rangle$ por $(1,0)^T$ é também válido, desde que se mantenha esta representação em todas as operações envolvidas.

A combinação de $\langle x|$ e de $|y\rangle$ como $\langle x||y\rangle$, também escrita como $\langle x|y\rangle$, refere o produto interno de dois vectores. Assim, como $|0\rangle$ é um vector unitário, $\langle 0|0\rangle = 1$, e como o $|0\rangle$ e o $|1\rangle$ são ortogonais, $\langle 0|1\rangle = 0$.

A notação $|x\rangle\langle y|$ corresponde ao produto externo (outer-product) entre $|x\rangle$ e $\langle y|$. Daí que $|0\rangle\langle 1|$ corresponde à transformação que mapeia $|1\rangle$ para $|0\rangle$ e $|0\rangle$ para $(0,0)^T$, pois:

$$|0\rangle\langle 1||1\rangle = |0\rangle\langle 1|1\rangle = |0\rangle \quad (6.1)$$

$$|0\rangle\langle 1||0\rangle = |0\rangle\langle 1|0\rangle = 0|0\rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (6.2)$$

De forma equivalente, $|0\rangle\langle 1|$ é escrito na forma matricial como:

$$|0\rangle\langle 1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (0,1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (6.3)$$

já que, $|0\rangle = (1,0)^T$, $\langle 0| = (1,0)$, $|1\rangle = (0,1)^T$ e o $\langle 1| = (0,1)$.

Esta notação é uma forma conveniente de representar transformações de estados quânticos. Por exemplo, a transformação que intermodifica os estados de $|0\rangle$ e de $|1\rangle$ é dada por:

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|. \quad (6.4)$$

Uma forma mais intuitiva de escrever a mesma transformação é:

$$X : \begin{array}{l} |0\rangle \rightarrow |1\rangle \\ |1\rangle \rightarrow |0\rangle \end{array}. \quad (6.5)$$

6.3 Bits quânticos (Qubits)

Um *bit* quântico é um vector unitário num espaço vectorial complexo de duas dimensões referido a uma base previamente estipulada. A base ortogonal $\{|0\rangle, |1\rangle\}$ pode corresponder, no caso dos fotões, às polarizações $\{|\uparrow\rangle, |\rightarrow\rangle\}$ ou às polarizações $\{|\nearrow\rangle, |\nwarrow\rangle\}$. No caso dos electrões corresponderá, respectivamente, ao *spin-up* e ao *spin-down* [8]. Nas situações de omissão assume-se que a base usada é a $\{|0\rangle, |1\rangle\}$.

Em analogia aos bits clássicos, usam-se os bits quânticos para a obtenção de dois estados, o 0 e o 1. De facto, num *qubit*, apenas é possível trazer para o mundo clássico estes dois estados. Mas previamente à operação de observação desses bits, antes do colapso da sua função de onda, os bits quânticos podem estar em sobreposição de acordo com $a|0\rangle + b|1\rangle$. Onde a e b são números complexos, designados por amplitude, e que respeitam a seguinte igualdade: $|a|^2 + |b|^2 = 1$. Na situação de sobreposição, a probabilidade de se observar o estado $|0\rangle$ é igual a $|a|^2$ e a probabilidade de se observar o estado $|1\rangle$ é igual a $|b|^2$.

Quando um estado é medido ou observado, a própria medida modifica o estado para uma das duas hipóteses dos estados da base de medida. Como a própria medida pode modificar o estado quântico, não é possível usar duas bases para realizar duas medidas em bases diferentes sobre o estado original. Se for usada uma base que não corresponde à base original, surge a probabilidade de se medir um estado errado e, após este facto, perde-se a informação do estado original. Isto impossibilita a clonagem de um *qubit* sem erros e capacita aplicações como a criptografia quântica. Se a medida for executada conhecendo a base que deu origem ao estado a ser medido, torna-se possível a clonagem do estado do *qubit* e torna-se possível teleportar partículas quânticas.

6.4 Esfera de Bloch

A esfera de Bloch constitui uma forma de visualizar o estado quântico de um *qubit*. Como $|\psi\rangle = a|0\rangle + b|1\rangle$ e $|a|^2 + |b|^2 = 1$, é possível escrever [62]:

$$|\psi\rangle = e^{i\gamma} \left(\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right)|1\rangle \right), \quad (6.6)$$

na qual os valores γ , θ e φ são números reais.

Por não produzir efeitos físicos observáveis é possível remover, de (6.6), o factor $e^{i\gamma}$, no que resulta:

$$|\psi\rangle = \left(\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right)|1\rangle \right). \quad (6.7)$$

Então, geometricamente, $|\psi\rangle$ pode ser representado numa esfera tridimensional onde θ e φ determinam um ponto na superfície da esfera unitária (Figura 6.1). Esse ponto representa um estado quântico de um *qubit*. As transformações quânticas modificam o valor de θ e de φ reposicionando o ponto, que define o estado quântico, na superfície da esfera. Após uma operação de observação, esse estado é alterado, projectando-se, sempre na respectiva base, em $|0\rangle$ ou em $|1\rangle$.

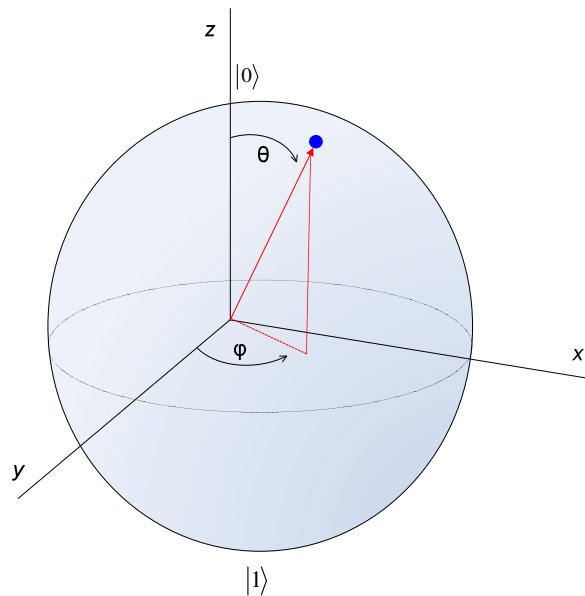


Figura 6.1 Esfera de Bloch

6.5 Enredo de qubits

Quando são considerados conjuntos de *qubits* quânticos, existe a possibilidade de que não se possa descrever o estado dessa agregação a partir do estado de cada um dos *qubits*. No mundo clássico é sempre possível descrever um sistema através do estado de cada componente que o constitui. Essa singularidade quântica, em conjunto com a sobreposição de estados, é apontada como a responsável pelo poder computacional verificado nos processos quânticos.

O estado de um *qubit* pode ser representado por um vector no espaço complexo bidimensional. Em física clássica, um sistema de n partículas, no qual os estados individuais de cada partícula são descritos por um vector num espaço bidimensional, forma um espaço vectorial de $2n$ dimensões. Porém, o espaço vectorial de estados de uma representação de n *qubits* atinge 2^n dimensões. Este crescimento exponencial do espaço de estados também sugere um aumento exponencial da capacidade computacional de um sistema quântico relativamente aos sistemas computacionais clássicos.

Enquanto que num sistema clássico a combinação dos espaços de estados individuais de partículas é produzida pelo produto cartesiano, num sistema quântico essa combinação é produzida por um produto tensorial. Por exemplo, considerando V e W dois vectores bidimensionais complexos de bases $\{v_1, v_2\}$ e $\{w_1, w_2\}$, respectivamente, então o produto cartesiano resultará numa união dos componentes espaciais de cada uma das bases, $\{v_1, v_2, w_1, w_2\}$, enquanto que o produto tensorial, no caso quântico, resultará em $\{v_1 \otimes w_1, v_1 \otimes w_2, v_2 \otimes w_1, v_2 \otimes w_2\}$. A dimensão da primeira combinação é dada por $\dim(X \times Y) = \dim(X) + \dim(Y)$; enquanto que a segunda combinação é obtida por $\dim(X \otimes Y) = \dim(X) \times \dim(Y)$. A base de espaço de estados de dois *qubits* com uma base individual, $\{|0\rangle, |1\rangle\}$, passa a ser $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$, que numa forma menos densa pode ser escrita como $\{|00\rangle |01\rangle |10\rangle |11\rangle\}$. É comum escrever-se $|x\rangle$ para representar $|b_n b_{n-1} \dots b_0\rangle$, onde b_i são os dígitos binários do número x .

A representação da base associada a 3 *qubits* será:

$$\{|000\rangle |001\rangle |010\rangle |011\rangle |100\rangle |101\rangle |110\rangle |111\rangle\}. \quad (6.8)$$

Um exemplo de um estado que não pode ser descrito, separadamente, pelos seus componentes é o: $|00\rangle + |11\rangle$. Não é possível calcular a_1 , a_2 , b_1 e b_2 tal que $(a_1|0\rangle + b_1|1\rangle) \otimes (a_2|0\rangle + b_2|1\rangle) = |00\rangle + |11\rangle$, já que

$$(a_1|0\rangle + b_1|1\rangle) \otimes (a_2|0\rangle + b_2|1\rangle) = a_1a_2|00\rangle + a_1b_2|01\rangle + a_2b_1|10\rangle + b_1b_2|11\rangle \quad (6.9)$$

e $a_1b_2 = 0$ implica que $a_1a_2 = 0$ ou que $b_1b_2 = 0$.

Os estados que não podem ser decompostos individualmente são chamados “estados de enredo quântico”.

6.6 Operação de observação

A observação de um *qubit* projecta o estado quântico para um dos estados da base do dispositivo de observação. O resultado tem natureza probabilística e o próprio processo de observação modifica o estado do *qubit*. A observação, nestes casos, pode conduzir ao conhecimento, no mundo clássico, de um dos estados de base sobre o *qubit* em medição relativamente ao dispositivo de medição.

Para o caso de se tratar de 2 *qubits*, a expressão que representa o seu estado é: $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, em que a , b , c e d são números complexos sujeitos à seguinte igualdade: $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. Pode reescrever-se a expressão de estado da seguinte forma:

$$\begin{aligned} & a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \\ &= |0\rangle \otimes (a|0\rangle + b|1\rangle) + |1\rangle \otimes (c|0\rangle + d|1\rangle) \quad , \\ &= u|0\rangle \otimes \left(\frac{a}{u}|0\rangle + \frac{b}{u}|1\rangle \right) + v|1\rangle \otimes \left(\frac{c}{v}|0\rangle + \frac{d}{v}|1\rangle \right) \end{aligned} \quad (6.10)$$

onde u é igual a $\sqrt{|a|^2 + |b|^2}$ e onde v é igual a $\sqrt{|c|^2 + |d|^2}$.

Assim, da expressão (6.10) podem determinar-se as seguintes probabilidades: a probabilidade de que o primeiro *qubit* lido seja um $|0\rangle$ é $|a|^2 + |b|^2$ e a probabilidade de o primeiro *qubit* ser um $|1\rangle$ é $|c|^2 + |d|^2$. A probabilidade de o segundo *qubit* ser um $|0\rangle$,

condicionada ao facto de se ter lido o primeiro como $|0\rangle$, é igual a $\frac{|a|^2}{|a|^2 + |b|^2}$. A

probabilidade de o segundo *qubit* ser um $|1\rangle$, condicionada ao facto de se ter lido o primeiro como $|0\rangle$, é igual a $\frac{|b|^2}{|a|^2 + |b|^2}$. Se o primeiro *qubit* fosse um $|1\rangle$, então as probabilidades de se encontrar o estado $|1\rangle$ ou o estado $|0\rangle$ no segundo *qubit* são calculadas de forma equivalente, mas tendo em conta c e d .

Assim se demonstra que *qubits* em enredo quântico são interdependentes. Ou seja, o facto de se observar um dos *qubits* num determinado estado, pode determinar a probabilidade de se observar um determinado estado nos restantes *qubits*.

A partir da observação de *qubits* é possível deduzir sobre a existência de estados em enredo quântico. Considerando o estado, $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, pode dizer-se que os *qubits* se encontram em enredo quântico, porque, se for executada uma observação sobre o primeiro *qubit*, fica automaticamente determinado o estado do segundo. A probabilidade de se observar o primeiro *qubit* a $|0\rangle$ ou a $|1\rangle$ é igual a 0.5, desde que o segundo ainda não tenha sido observado. Depois de se observar o primeiro *qubit*, a probabilidade de se encontrar o segundo *qubit* a $|0\rangle$ será de 1.0, se o primeiro foi lido como $|0\rangle$, e será de 0.0, se o primeiro foi lido como $|1\rangle$. Na situação oposta estará o estado $\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$, que pode ser escrito como $|0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Qualquer medida feita em primeiro lugar a um ou a outro *qubit* não alterará a probabilidade de ser medido um qualquer estado em nenhum deles. Isto porque o primeiro *qubit* tem sempre o estado $|0\rangle$; o segundo irá ter uma projecção na sua base, que pode ser $|0\rangle$ ou $|1\rangle$, ambas com uma probabilidade de 0.5.

6.7 Portas quânticas

A dinâmica de um sistema quântico, antes da operação de observação, é regida pela equação de Schrödinger. Essa dinâmica produz-se de estados para estados preservando a ortogonalidade. Num espaço de vectores complexos a ortogonalidade é mantida por transformações lineares unitárias. Qualquer transformação deste tipo e neste espaço pode ser descrita matricialmente. Uma matriz M é unitária se $MM^* = I$,

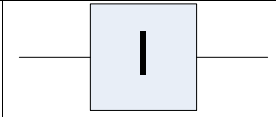
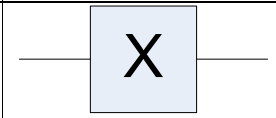
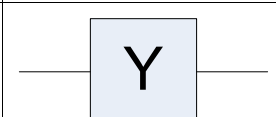
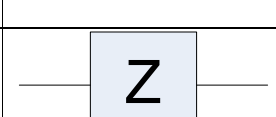
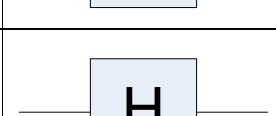
em que M^* corresponde à transposta da conjugada da matriz M . Qualquer transformação unitária do espaço de estados quântico é uma transformação válida. As transformações unitárias podem ser vistas como rotações no espaço de vectores complexos.

Uma sequência directa do facto de as transformações serem unitárias é a de serem reversíveis. Este facto é importante frente a questões que relacionam a quantidade de energia com a informação gerada/manipulada [64]. Alguns investigadores, como Bennett, Fredkin e Toffoli mostraram que todas as operações clássicas podem ser executadas reversamente.

6.7.1 Portas quânticas de um único qubit

As transformações a seguir apresentadas, Tabela 6.I, correspondem a portas quânticas de um único *qubit*. As quatro primeiras são conhecidas por transformações de Pauli.

Tabela 6.I Transformações quânticas de um *qubit*

| Operação | Forma de transição de estados | Forma matricial | Representação gráfica |
|--|--|--|---|
| Identidade | $I : 0\rangle \rightarrow 0\rangle$ $ 1\rangle \rightarrow 1\rangle$ | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ |  |
| Negação | $X : 0\rangle \rightarrow 1\rangle$ $ 1\rangle \rightarrow 0\rangle$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ |  |
| Negação seguida de deslocamento de fase (Y = ZX) | $Y : 0\rangle \rightarrow - 1\rangle$ $ 1\rangle \rightarrow 0\rangle$ | $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ |  |
| Deslocamento de fase | $Z : 0\rangle \rightarrow 0\rangle$ $ 1\rangle \rightarrow - 1\rangle$ | $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ |  |
| Sobreposição de estados | $H : 0\rangle \rightarrow \frac{1}{\sqrt{2}}(0\rangle + 1\rangle)$ $ 1\rangle \rightarrow \frac{1}{\sqrt{2}}(0\rangle - 1\rangle)$ | $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ |  |

Como se pode verificar, todas as transformações são unitárias, garantindo a reversibilidade necessária ao cumprimento dos fenómenos físicos quânticos. Por exemplo:

$$YY^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = I. \quad (6.11)$$

6.7.2 A transformação de Walsh-Hadamard

A transformação que na Tabela 6.I está designada por H é muito importante na maior parte das aplicações de aceleração de computação quântica. Quando aplicada a $|0\rangle$ ou a $|1\rangle$ gera uma sobreposição de estados de $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ou de $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectivamente. Aplicada a n qubits individualmente, a transformação gera uma sobreposição de todos os possíveis 2^n estados. Estes podem ser vistos como a representação binária dos números compreendidos entre 0 e $2^n - 1$. A transformação sobre n qubits pode ser calculada através do produto tensorial dos estados sobrepostos de cada qubit individual,

$$\begin{aligned} & (H \otimes H \otimes \dots \otimes H) |00\dots 0\rangle \\ &= \frac{1}{\sqrt{2^n}} ((|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes \dots \otimes (|0\rangle + |1\rangle)). \quad (6.12) \\ &= \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \end{aligned}$$

6.7.3 Portas de dois e de três qubits

A chamada porta *NOT controlada* (controlled-NOT), também conhecida por porta de Feynman, usa dois qubits. Modifica o valor do segundo qubit, se o valor do primeiro for $|1\rangle$, e não altera o valor do segundo qubit se o valor do primeiro for $|0\rangle$. A transformação $Cnot$ é unitária, já que $Cnot^* = Cnot$ e que $Cnot^* Cnot = I$. A transformação $Cnot$ não pode ser decomposta num produto tensorial de transformações de um único bit. Na Tabela 6.II é apresentada a representação esquemática da porta $Cnot$. O círculo aberto indica o qubit de controlo e o \times indica o qubit sujeito ao qubit de controlo. Na mesma tabela é apresentada a porta *NOT duplamente controlada* (CCNot). Neste caso, o último qubit apenas é negado se os dois primeiros tiverem o seu estado a $|1\rangle$.

Tabela 6.II Porta Cnot e porta CCnot

| Operação | Forma de transição de estados | Forma matricial | Representação gráfica |
|---------------------------|---|--|-----------------------|
| NOT controlado | Cnot: $ 00\rangle \rightarrow 00\rangle$ $ 01\rangle \rightarrow 01\rangle$ $ 10\rangle \rightarrow 11\rangle$ $ 11\rangle \rightarrow 10\rangle$ | $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ | |
| NOT duplamente controlado | CCnot: $ 000\rangle \rightarrow 000\rangle$ $ 001\rangle \rightarrow 001\rangle$ $ 010\rangle \rightarrow 010\rangle$ $ 011\rangle \rightarrow 011\rangle$ $ 100\rangle \rightarrow 100\rangle$ $ 101\rangle \rightarrow 101\rangle$ $ 110\rangle \rightarrow 111\rangle$ $ 111\rangle \rightarrow 110\rangle$ | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ | |

6.8 Computadores quânticos

Para que seja possível implementar um computador quântico, as portas que o constituem têm que ser reversíveis. A única porta que em termos clássicos é reversível é a porta NOT. As portas clássicas AND, OR e NAND não o são. No entanto, a porta CCnot, também conhecida por porta de Toffoli [65], representada por T , permite realizar qualquer computação clássica num computador quântico. A porta pode ser usada para construir um qualquer circuito combinacional, já que alguns estados afectos a esta porta se encontram no paralelo domínio de estados das portas clássicas NOT e AND. As expressões de (6.13) representam as situações em que uma porta de Toffoli estabelece uma porta NOT e uma porta AND, respectivamente.

$$\begin{aligned} T|1,1,x\rangle &= |1,1,-x\rangle \\ T|x,y,0\rangle &= |x,y,x \wedge y\rangle \end{aligned} \tag{6.13}$$

O exemplo seguinte, Figura 6.2, mostra um circuito de portas quânticas de Toffoli e de Feynman para obter a soma entre dois *qubits*,

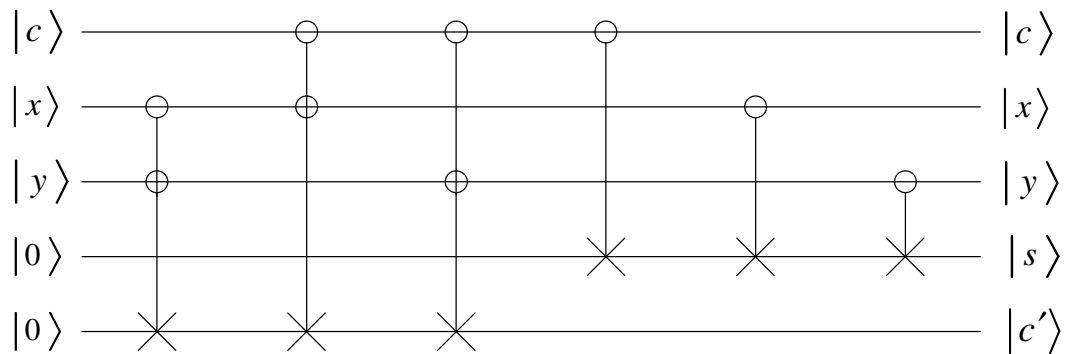


Figura 6.2 Circuito somador quântico

onde, x e y são os *bits* a somar, s é o resultado, c é o *bit* a transportar e c' é o novo *bit* de transporte.

Deutsch [66] demonstrou que é possível construir portas quânticas reversíveis para qualquer função matemática de resolução clássica. Bernstein e Vazirani demonstraram que é possível implementar uma máquina de Turing num suporte quântico [67].

Supondo-se uma qualquer função clássica f com m *bits* de entrada e k *bits* de saída, é possível a sua implementação quântica através de uma matriz de portas quânticas (*quantum gatearray*) U_f . U_f é uma transformação com $m + k$ *qubits* na forma,

$$U_f : |x, y\rangle \rightarrow |x, y \oplus f(x)\rangle, \quad (6.14)$$

onde \oplus simboliza a operação *bit a bit* de *ou-exclusivo*. Matrizes quânticas definidas desta forma são unitárias para qualquer função f . É necessário que o número de *qubits* da saída seja igual ao da entrada. Assim, para computar uma função $f(x)$ submete-se $|x\rangle$ a U_f , fazendo o produto tensorial dos elementos de x com k estados de $|0\rangle$.

A transformação quântica de uma função genérica é graficamente definida na Figura 6.3.

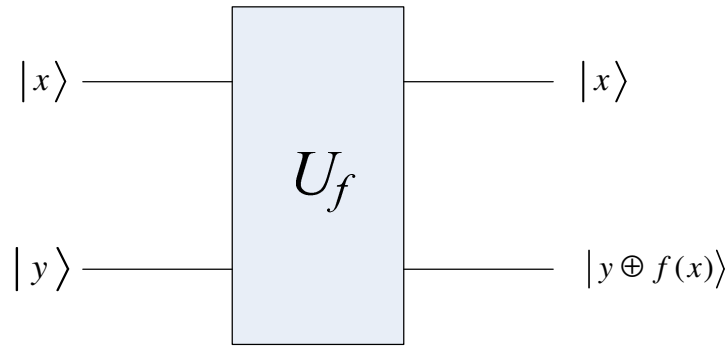


Figura 6.3 Porta quântica genérica

As portas quânticas de Toffoli permitem criar um equivalente quântico para qualquer circuito combinacional clássico, mas não abrangem todas as operações quânticas possíveis. É por isso necessário incluir operações explícitas de rotação de *qubits*.

Barenco et. al. [68] mostraram que a porta de Feynman (Cnot) em agregação com todas as portas de 1 *qubit* constituem um conjunto completo de portas quânticas. Para isso é suficiente incluir, em conjunto com a porta Cnot, as seguintes transformações de 1 *qubit*,

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}, \quad (6.15)$$

nas quais α pode variar entre 0 e 2π . Tais transformações são essenciais para se extraírem as capacidades de computação ímpares dos sistemas quânticos.

6.9 Paralelismo quântico

Se estados quânticos em sobreposição forem submetidos às entradas de uma função quântica U_f , todos esses estados serão processados pela função simultaneamente. Como a função é uma transformação linear, a sua saída será uma sobreposição de resultados. Assim, de uma só vez é possível criar todos os resultados que U_f devolve para cada uma das combinações de entrada. A isto chama-se “paralelismo quântico”.

A maior parte dos algoritmos quânticos começa por processar uma função de interesse em que as suas entradas são colocadas em sobreposição. n qubits no estado $|00\dots 0\rangle$ são sujeitos à transformação de Walsh-Hadamard para se obter a sobreposição,

$$\frac{1}{\sqrt{2^n}}(|00\dots 0\rangle + |00\dots 1\rangle + \dots + |11\dots 1\rangle) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle, \quad (6.16)$$

a qual é vista como a sobreposição de todos os inteiros, $0 \leq x < 2^n$. Adicionando, às entradas, um registo de k qubits a $|0\rangle$, então, por linearidade, pode escrever-se

$$U_f \left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, 0\rangle \right) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} U_f |x, 0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, f(x)\rangle, \quad (6.17)$$

em que $f(x)$ é a função de interesse.

Considere-se a transformação de Toffoli (CCnot), Figura 6.4, que processa a conjunção de 2 qubits:

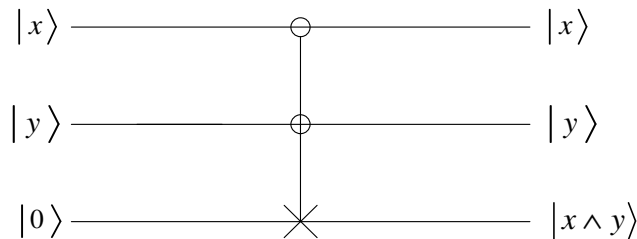


Figura 6.4 Operação de conjunção com uma porta de Toffoli

No primeiro passo colocam-se em sobreposição as entradas,

$$\begin{aligned} H|0\rangle \otimes H|0\rangle \otimes |0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle \\ &= \frac{1}{2}(|000\rangle + |010\rangle + |100\rangle + |110\rangle) \end{aligned} \quad (6.18)$$

Como se pode verificar, resultam desta operação quatro estados possíveis e em sobreposição. Ao sujeitar estes estados à porta de Toffoli originará em,

$$T(H|0\rangle \otimes H|0\rangle \otimes |0\rangle) = \frac{1}{2}(|000\rangle + |010\rangle + |100\rangle + |111\rangle). \quad (6.19)$$

A sobreposição de estados resultantes pode ser vista como a tabela verdade da função de conjunção. Uma operação de observação sobre este registo de saída terá como resultado uma das 4 linhas da tabela verdade, cada uma, com uma probabilidade de ser medida de 0.25, independentemente da ordem com que são observados cada um dos *qubits*. Operar observações neste patamar do processo não produz qualquer vantagem nos sistemas quânticos relativamente aos sistemas clássicos. Nem sequer permite escolher qual dos resultados a observar. Os algoritmos quânticos, para poderem extrair vantagens do paralelismo quântico, têm que ser capazes de, de alguma forma, aumentar a probabilidade de um resultado de interesse ser observado.

Existem actualmente duas técnicas que permitem consegui-lo:

- Amplificando os valores de amplitude do(s) estado(s) de interesse. A técnica consiste em aplicar transformações quânticas a estados de interesse, de forma a que a probabilidade de se observarem esses estados seja aumentada e a probabilidade de se observarem os restantes seja diminuída.
- Encontrando propriedades comuns dos valores de $f(x)$. A ideia foi proposta por Peter Shor em 1994 [69]. Através da transformada de Fourier quântica, Shor obtém o período da função $f(x)$, conseguindo factorizar em tempo polinomial, m^2 , um qualquer valor de m bits.

6.10 Algoritmo de Grover

Este algoritmo [70] consiste num método de procura de um elemento i , em listas desordenadas, de tamanho N . A ordem de resolução temporal para este algoritmo é $O(\sqrt{N})$. Em termos clássicos a ordem média é de $O\left(\frac{N}{2}\right)$. Considere-se um valor n tal que $2^n \leq N$. Considere-se ainda uma função quântica U_p que suporta uma função clássica $p(x,i)$ que testa se um determinado elemento i é igual a um valor que se encontra indexado por x , devolvendo o valor 1 caso essa condição se verifique:

$$U_p : |x, i, 0\rangle \rightarrow |x, i, p(x, i)\rangle. \quad (6.20)$$

O primeiro passo consiste em colocar todos os *qubits* de entrada $|x\rangle$ em sobreposição, juntamente com o *qubit* $|0\rangle$. Desta forma, todos os valores da lista são indexados e comparados simultaneamente. O estado de $|i\rangle$ define, sem sobreposição, o valor do elemento a procurar. A saída da função quântica U_p ficará estabelecida como:

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, i, p(x, i)\rangle. \quad (6.21)$$

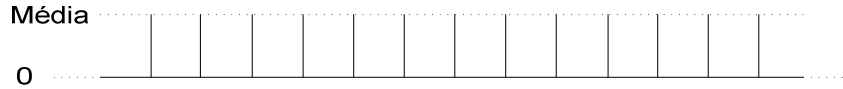
A dificuldade está em obter um resultado de interesse entre todos os estados em sobreposição. Para um determinado par x_j, i_l tal que $p(x_j, i_l) = 1$, resultará um estado, $|x_j, i_l, 1\rangle$, como parte dos estados em sobreposição. Já que a amplitude de tal estado é $\frac{1}{\sqrt{2^n}}$, a probabilidade de uma observação devolver o estado produzido por x_j, i_l é apenas de 2^{-n} . Grover propôs transformações, a aplicar à Equação (6.21), que aumentam a probabilidade de se observarem os estado de $|x, i, 1\rangle$, para os quais p é verdadeiro, e que diminuem a probabilidade de se observarem os estados de $|x, i, 0\rangle$, para os quais p é falso.

Assim, é suficiente começar por observar o último *qubit*. Se este estiver a $|1\rangle$, os restantes *qubits* corresponderão a informação relativa de um elemento que se procurava na lista. Caso esteja a $|0\rangle$, é necessário reiniciar todo o processo, pois existe a probabilidade de ser observado um estado desses, que pode indicar que não existe nenhum elemento na lista, mas também pode surgir, embora com baixa probabilidade, por mera incerteza na projecção resultante da observação.

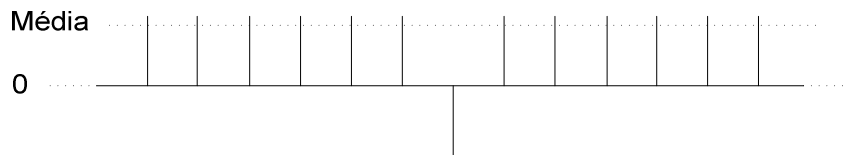
Fases do algoritmo de Grover:

1. Preparar o registo de entrada $|x\rangle$ de forma a obter a sobreposição de todos os estados possíveis;

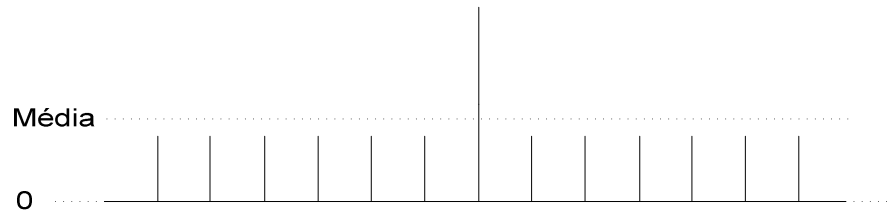
2. Definir o valor a procurar $|i\rangle$;
3. Computar U_p sobre o registo de entrada $|x, i, 0\rangle$. As amplitudes de cada estado no registo de saída de U_p são mostradas a seguir:



4. Modificar a amplitude de a_j para $-a_j$, do estado $|x_j, i, 1\rangle$ que cumpre a condição $p(x_j, i) = 1$. O diagrama de amplitudes é alterado de acordo com a seguinte figura:



5. Aplicar a inversão sobre a média para aumentar a amplitude do estado $|x_j, i, 1\rangle$. A inversão da amplitude sobre a média faz com que o estado que anteriormente tinha amplitude negativa adquira agora um valor superior à média:



6. Repetir os passos, do 3 ao 5, $\frac{\pi}{4}\sqrt{2^n}$ vezes. Agora o registo de entrada corresponde ao estado obtido na última iteração do passo 5. Com esse valor de repetição a probabilidade de ler um estado errado cai para 2^{-n} . Devido à natureza circular do mapeamento do espaço de estados quânticos, a partir deste valor a probabilidade de se observar o estado pretendido volta a decrescer [71];
7. Observar o resultado sobre o registo de saída.

6.10.1 Modificação do sinal da amplitude

É necessário garantir que os estados respeitantes à condição $p(x_j, i_l) = 1$, identificados pelo último *qubit*, sofram uma alteração no sinal do seu valor de amplitude. Esta alteração pretende estabelecer uma diferenciação entre os estados de interesse e os estados restantes. No entanto, no mundo clássico, após observação, esta alteração do sinal de amplitude não provoca qualquer divergência probabilística. Para que isso aconteça, o passo 4 tem que ser executado como complemento ao passo 5.

O método que produz a negação do sinal do valor de amplitude aos estados que apresentam o seu último *qubit* a $|1\rangle$ é muito simples. Consiste em estabelecer o estado do último *qubit*, do registo de entrada, da função quântica que implementa $p(x, i)$, como

$$\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

6.10.2 Inversão sobre a média

Com esta operação pretende-se aumentar a amplitude dos estados quânticos relacionados com a condição $p(x_j, i_l) = 1$. Usando-se a transformação (6.22), pode entender-se que o valor da nova amplitude se reduz para os estados que já tinham um valor positivo e ligeiramente acima da média (após a operação de modificação do sinal de amplitude), representada por A , e que os estados com amplitudes negativas passam a ter um valor igual ao dobro da média mais o valor absoluto da sua amplitude actual. y representa o estado de $p(x_j, i_l)$.

$$\sum_{i=0}^{2^{n-1}-1} a_i |x_i, i_l, y\rangle \rightarrow \sum_{i=0}^{2^{n-1}-1} (2A - a_i) |x_i, i_l, y\rangle. \quad (6.22)$$

Esta transformação pode ser representada por uma matriz $N \times N$:

$$D = \begin{pmatrix} \frac{2}{N} - 1 & \frac{2}{N} & \dots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} - 1 & & \frac{2}{N} \\ & & & \\ \frac{2}{N} & \frac{2}{N} & & \frac{2}{N} - 1 \end{pmatrix}. \quad (6.23)$$

Grover mostrou que esta matriz pode ser construída através de portas quânticas elementares em número igual a $\log(N)$. Mostrou também que a matriz D é igual a WRW , em que W corresponde à transformação de Walsh-Hadamard e em que R é definido como:

$$R = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots \\ 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & -1 \end{pmatrix}. \quad (6.24)$$

6.11 Adaptação do algoritmo de Grover ao processamento de imagem

O algoritmo quântico que a seguir se descreve tem como base fundamental o algoritmo de Grover. São abordados dois tipos de problemas. Um refere-se à detecção e à classificação de contornos em imagens e outro refere-se à segmentação de regiões.

Uma forma de detectar e de identificar contornos numa imagem binária é fazê-lo exaustivamente. Ou seja, dentro de uma determinada classe de figuras, gerar todas as translações, todas as rotações e todas as escalas de contornos possíveis. Para cada elemento gerado é feita uma comparação com a imagem em análise, pixel a pixel, e assim é detectada qual a figura, a sua localização, a sua posição angular e a sua escala. Esta técnica, apesar de funcional, é impraticável de implementar nos computadores clássicos. Por exemplo, para uma imagem de 1200 pixels por 1200 pixels, em que fosse necessário comparar uma classe de figuras com 300 escalas diferentes e com uma resolução angular de 1 grau, seriam necessárias 155000 milhões de comparações. Num computador quântico, e segundo o algoritmo de Grover, essa exigência decairia para 309000 operações de comparação. Supondo que uma operação de comparação demora, em média, 500us para ser realizada, em ambos os circuitos, clássico e quântico, então o resultado final demoraria a ser obtido, no processador clássico, 896 dias, enquanto no quântico demoraria 150 segundos. A diferença temporal entre os dois tipos de processos amplifica-se geometricamente à medida que a dimensão da imagem a processar aumenta e à medida que o número de classes de figuras a procurar aumenta.

Suponha-se a implementação numa matriz de portas quânticas, P_f , (Figura 6.5), uma função clássica, $f(x,i,w)$, que compara os pixels constituintes de uma figura, codificada em x , com os pixels de uma imagem, i , que poderá ou não conter a figura. Na função f , w representa os *qubits* internos da função P_f . A função f devolverá o valor 1 no(s) caso(s) em que a figura se encontre na imagem. Cada figura de comparação pode ocupar um qualquer local da imagem, com uma qualquer posição angular e com uma qualquer escala. Se aparecer na imagem uma figura que coincida exactamente com alguma das formas de comparação, o estado associado a essa figura terá a sua amplitude quântica aumentada relativamente aos outros estados. Como os *qubits* a observar, $|x\rangle$, são, além do *bit* relacionado com o valor devolvido pela função f , especificadores da combinação parametrizadora da figura, sabe-se de imediato qual o tipo de figura encontrada, a sua escala, a sua rotação e a sua posição na imagem. A função $f(x,i,w)$ é implementada de forma quântica como a seguir se descreve a partir da Tabela 6.III e Tabela 6.IV.

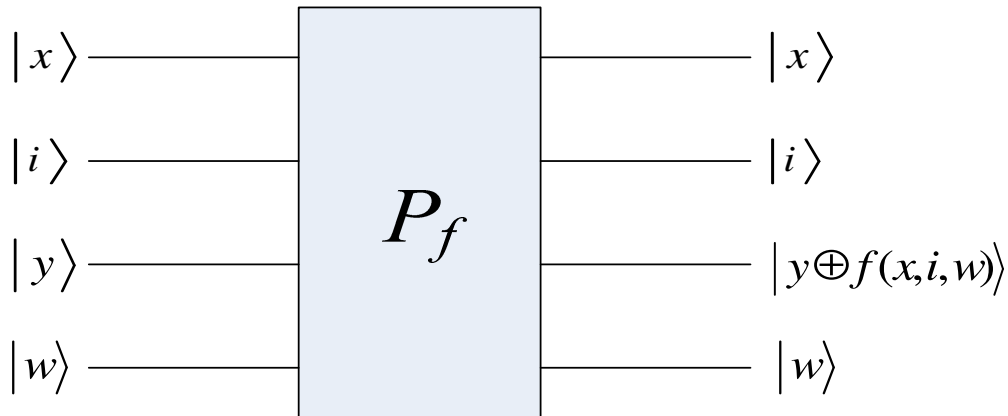


Figura 6.5 Função quântica para detecção e classificação de figuras em imagens

Tabela 6.III Correspondência entre os vários segmentos de estado quântico de $|x\rangle$ e a parametrização da figura a procurar

| $ x\rangle$ | | | | |
|---|---|---|---|---|
| Tipo de figura | Posição x | Posição y | Escala | Rotação |
| $\lceil \log_2(c \text{ figuras}) \rceil$ | $\lceil \log_2(\dim(X)) \rceil$ | $\lceil \log_2(\dim(Y)) \rceil$ | $\lceil \log_2(\min(\dim(X), \dim(Y))) \rceil$ | $\left\lceil \log_2\left(\frac{2\pi}{\text{resolução angular}}\right) \right\rceil$ |
| $(H \otimes H \otimes \dots \otimes H) 00\dots 0\rangle$ | $(H \otimes H \otimes \dots \otimes H) 00\dots 0\rangle$ | $(H \otimes H \otimes \dots \otimes H) 00\dots 0\rangle$ | $(H \otimes H \otimes \dots \otimes H) 00\dots 0\rangle$ | $(H \otimes H \otimes \dots \otimes H) 00\dots 0\rangle$ |

O significado de cada uma das linhas da Tabela 6.III é o seguinte:

1. A primeira linha refere o segmento do registo quântico de entrada;
2. A segunda linha refere cada um dos parâmetros constituintes do segmento referido em 1;
3. A terceira linha mostra o comprimento de cada parâmetro em *qubits*. X e Y representam o conjunto de valores possíveis que podem tomar as coordenadas de pixels na imagem a analisar.
4. A última linha mostra os estados iniciais a submeter a um produto tensorial entre todos os subregistos quânticos de $|x\rangle$.

Tabela 6.IV Número de *qubits* e estados para $|i\rangle$, $|y\rangle$ e $|w\rangle$ da função quântica P_f

| $ i\rangle$ | $ y\rangle$ | $ w\rangle$ |
|---|---|--|
| $\lceil \log_2(\dim(X) \cdot \dim(Y)) \rceil$ | 1 | Número de qubits dependente da organização interna e da complexidade da função P_f |
| Ex.: $ 001010011111000\dots111001\rangle$ (<i>qubits</i> que representam o valor de cada pixel na imagem a processar) | $\frac{1}{\sqrt{2}}(0\rangle - 1\rangle)$ | Normalmente a $ 000\dots00\rangle$ |

Na Tabela 6.IV são mostrados os registos quânticos $|i\rangle$, $|y\rangle$ e $|w\rangle$ no que concerne a cada um dos seus comprimentos (segunda linha) e no que concerne ao seu domínio de estados, (terceira linha). O registo $|i\rangle$ codifica a imagem a analisar, cada *qubit* corresponderá a um pixel de valor binário, ou, então, subgrupos de *qubits* de $|i\rangle$ podem representar pixels em níveis de cinzento. O $|y\rangle$ é o *qubit* que irá identificar um estado de interesse, por exemplo, um estado de coincidência entre a imagem $|i\rangle$ e a figura a procurar $|x\rangle$. $|w\rangle$ representa os qubits internos da função P_f necessários para a sua construção.

Algoritmo quântico de procura de figuras:

1. Iniciar todos os *qubits* do registo $|x\rangle$ a $|0\rangle$;

2. Submeter cada *qubit* do registo $|x\rangle$ à transformação de Walsh-Hadamard;
3. Estabelecer o estado inicial do registo $|y\rangle$ a $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$;
4. Fornecer a cada *qubit* do registo $|i\rangle$ o correspondente valor de cada pixel da imagem a processar;
5. Iniciar todos os *qubits* do registo $|w\rangle$ a $|0\rangle$;
6. Repetir os passos, do 7 ao 9, $\frac{\pi}{4}\sqrt{2^n}$ vezes ($n =$ número de *qubits* de $|x\rangle$);
7. Computar $P_f(|x, i, y, w\rangle)$ sobre o último estado obtido no registo $|x, i, y, w\rangle$;
8. Modificar o sinal da amplitude do estado de solução;
9. Aumentar a amplitude do estado de solução;
10. Observar os estados dos *qubits*. Se $|y\rangle = 0$, regressar a 1 ou finalizar; se não, reter o resultado.
11. Identificar no registo de saída $|x\rangle$ a figura encontrada.

6.11.1 Função de comparação

Suponha-se que se está interessado em procurar um único tipo de figura e que é uma circunferência. Quer descobrir-se se esta existe numa determinada imagem binária, em que local e qual a sua escala. Na verdade têm que se considerar duas funções integradas numa só: a que produz os pixels a partir da parametrização de escala, rotação e posição e a que compara esses pixels com os pixels da imagem a analisar.

Na versão discretizada, o conjunto de pares de posição de pixels que formam o contorno é dado pelas equações paramétricas:

$$C \times R = \begin{cases} x = r \cdot \cos(\alpha) \\ y = r \cdot \sin(\alpha) \end{cases} \quad 0 \leq \alpha < 2\pi, \quad (6.25)$$

em que r representa o raio da circunferência e α representa o ângulo que o vector que parte da origem até ao pixel, de coordenadas (x, y) , faz com o eixo de referência xx .

O conjunto finito $C \times R$ define os pares de coordenadas (x, y) em \mathbb{Z}^2 , constituintes do contorno da figura a estabelecer.

Os problemas de discretização obrigam a que, pelo menos, se escolha com cuidado a resolução para o α , pois desta depende o aparecimento de menores ou maiores discontinuidades no mapeamento do contorno sobre a escala a considerar. Quanto maior a escala, maior deverá ser a resolução de α .

Sendo o conjunto de pixels que formam a imagem em análise dado por I , então, para que se possa considerar que uma figura existe nessa imagem, tem que,

$$C \times R \xrightarrow{g} I = \{1, 1, 1 \dots 1, 1\}, \quad (6.26)$$

ou seja, qualquer par (x, y) , com $x \in C$ e $y \in R$, pertencente à figura a testar, que posiciona um pixel (x, y) na imagem a analisar I , deve encontrar esse pixel a 1 na imagem a analisar (imagens binárias).

O *bit* indicador de comparação verdadeira seria o resultado de uma operação de conjunção sobre todos os valores que resultam da aplicação $C \times R \xrightarrow{g} I$.

Neste caso está a considerar-se que as figuras são primitivas e que não existem figuras compostas. Caso isto acontecesse, as condições poderiam ter de ser mais restritas consoante a análise que se pretendesse.

Se fosse necessário comparar mais tipos de figuras, então bastaria considerar mais *bits* que estabeleceriam a condição de selecção para um conjunto de equações paramétricas de definição de figuras. As equações activas estariam em correspondência com a figura em causa. No caso quântico, todas as figuras seriam comparadas em simultâneo, mas apenas seria activado o estado da figura que fosse encontrada na imagem. Se estiverem várias figuras na imagem, o processo quântico terá que ser repetido, pois apenas uma figura é identificada por cada observação. Quer isto significar que o número de repetições do processo quântico estará directamente relacionado com a própria estatística da imagem. Este facto, inerente ao algoritmo de Grover, pode ser negativo. Mas continua a ser vantajoso, em termos temporais e em termos de eficácia funcional, usar um algoritmo quântico.

O mapeamento feito pela Equação (6.25) está sujeito a problemas de arredondamento de posições, o que poderá causar comparações *falsas*, sem o serem na realidade. O algoritmo de Bresenham [72] é uma forma adequada de se obter o mapeamento das posições contínuas para as discretas. Para resolver esse problema pode

ser usado um mecanismo de comparação sobre um valor limiar. Assim, não é necessário obter uma coincidência de 100% de pixels para considerar que determinada figura existe na imagem em análise. Escolhe-se um valor limiar, por exemplo 90%. Desta forma, todas as comparações que devolverem um número de pixels coincidentes superior a 90% do total serão consideradas verdadeiras. Por consequência, irão produzir estados quânticos de interesse.

Outra forma de estabelecer contornos a partir de parâmetros é pelos descritores de Fourier [73]. Estes trazem uma possível ajuda ao serem implementados quanticamente.

Segundo o algoritmo de Peter Shor é possível, de forma inerente, implementar a transformada de Fourier quântica e a inversa da mesma transformada, em vez de a mapear de uma função clássica para uma função quântica. Isto minimiza o número de portas quânticas necessárias à sua implementação. De qualquer forma teria que existir um mecanismo que associasse os coeficientes de Fourier aos diferentes tipos de figuras que aí codificariam.

A implementação quântica destas funções pode ser feita recorrendo a um pressuposto, o de que todas as funções clássicas podem ser implementadas por funções quânticas. Também se sabe que qualquer função pode ser, após discretizada, implementada por funções de conjunção negadas (NAND). Quer isto dizer que a obtenção dos circuitos quânticos pode ser feita a partir da representação destas funções por portas *booleanas* clássicas e consequente mapeamento de NANDs clássicas para NANDs quânticas.

6.11.2 Teste num simulador quântico

Para que se pudesse minimamente verificar, experimentalmente, a validade do algoritmo proposto, recorreu-se a um simulador quântico [74]. Os simuladores quânticos emulam todos os comportamentos das partículas quânticas a partir dos modelos matemáticos matriciais que representam as respectivas transformações quânticas. Claramente, e porque algumas das transformações são de natureza tensorial, o desempenho verificado nas emulações é de custo exponencial. Por este motivo e para simplificação da visualização do problema, decidiu-se tratar o algoritmo sobre um problema de dimensão menor. Ou seja, em vez de se considerar a procura de um padrão numa imagem bidimensional, procurou-se sobre uma recta.

Nesse caso, vai considerar-se que um segmento de recta contido na recta, representada por 5 pixels binários, é o padrão a procurar. Assim, como exemplo, os valores *00110* definem uma recta na qual o 3.º e o 4.º pixel constituem o segmento de recta a encontrar. A função terá assim que codificar os 4 padrões possíveis a partir de 2 bits. Estes bits vão permitir identificar qual a posição do segmento de recta situado sobre a recta submetida a análise. A tabela verdade seguinte, Tabela 6.V, mostra a relação entre cada bit de parametrização *x* e as diferentes posições de cada segmento de recta.

Tabela 6.V Tabela verdade entre os bits, x_0, x_1 , de parametrização de segmentos de recta e os bits que estabelecem o correspondente segmento

| x_1 | x_0 | w_1 | w_2 | w_3 | w_4 | w_5 |
|-------|-------|-------|-------|-------|-------|-------|
| 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 1 | 1 |

As equações *booleanas* obtidas pela aplicação de mapas de *Karnaugh* são as seguintes:

$$\begin{aligned}
 w_1 &= \bar{x}_1 \cdot \bar{x}_0 \\
 w_2 &= \bar{x}_1 \\
 w_3 &= x_1 \oplus x_0 \\
 w_4 &= x_1 \\
 w_5 &= x_1 \cdot x_0
 \end{aligned}
 \tag{6.27}$$

A cada par $w_j \times i_j$, em que *i* representa a imagem a analisar (unidimensional neste caso) e *j* é o índice de um elemento dessa imagem, é aplicada uma operação XOR para avaliar a sua igualdade. Por fim, se todos os pares forem iguais, uma operação de AND que abrange todas as saídas das operações do XOR, referido atrás, terá o seu valor de saída a 1. Repare-se que, para a função quântica em causa, os registos *w* são registos intermédios que são usados como suporte auxiliar às operações de comparação.

A conversão entre o circuito lógico clássico e o circuito lógico quântico é relativamente directa. A única excepção está em ter que se garantir a relação unitária das entradas para as saídas, de forma a cumprir a regra da reversibilidade quântica. Isto,

por vezes, provoca a necessidade de usar registos de *trabalho* adicionais na implementação quântica.

Imaginando que os *qubits* usados no circuito da Figura 6.6 são estabelecidos por fótons, então o percurso dos fótons executa-se da esquerda para a direita. À medida que os fótons vão interagindo com os dispositivos de transformação de estado quântico, o seu estado modifica-se até encontrarem novo dispositivo. Neste sentido, devido à velocidade finita de propagação dos fótons, o tempo evolui também da esquerda para a direita.

O último *qubit*, o y , é o responsável por assinalar o estado de interesse.

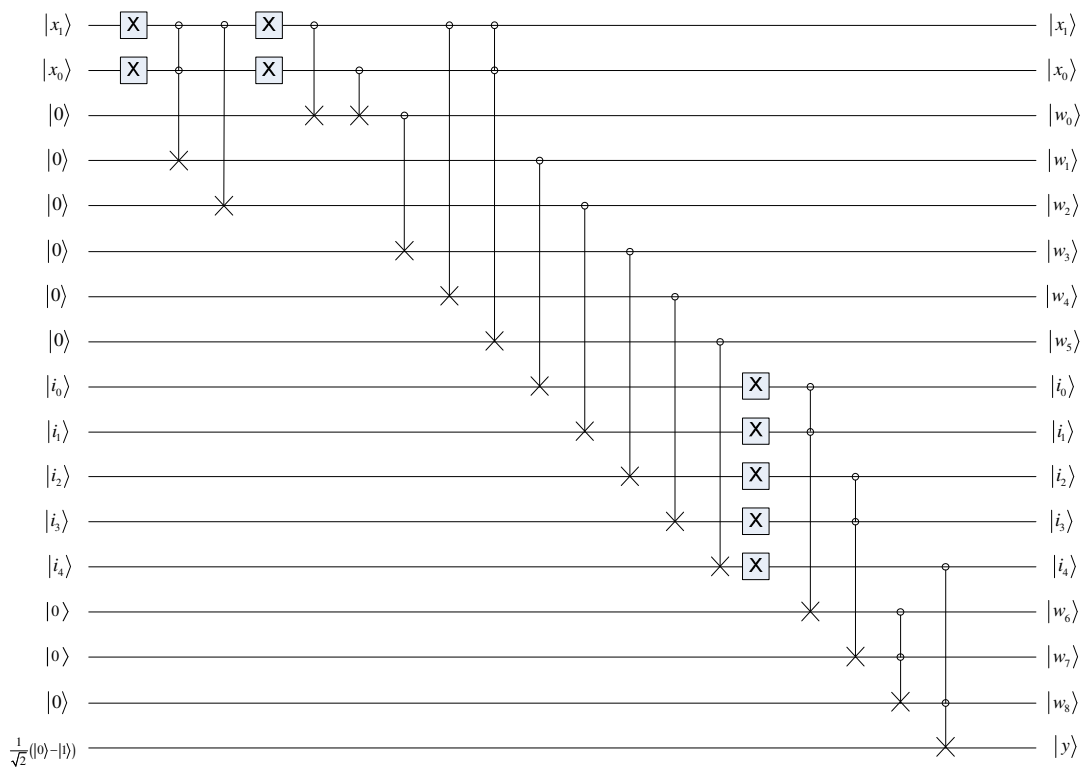


Figura 6.6 Circuito quântico para implementação do Oracle quântico, que detecta a posição de segmentos de recta numa dada recta

O circuito foi testado individualmente no simulador quântico, combinando os valores de x e avaliando a saída y para um determinado segmento de recta i , como se de um circuito clássico se tratasse. Quer isto dizer que, nesta fase de teste, as entradas quânticas não tinham estados sobrepostos.

Depois, agregou-se este Oracle¹² aos módulos de implementação do algoritmo de Grover existentes no simulador. As entradas x foram colocadas em sobreposição pela transformada de Walsh-Hadamard. O resultado devolvido pelo simulador foi o estado que anunciava a posição do segmento de recta na linha (imagem) considerada. Comprovou-se desta forma o esperado funcionamento do algoritmo.

6.11.3 Função quântica para a segmentação multi-escala

A utilização do simulador [74] é disponibilizada via Internet pelo Instituto Fraunhofer e está limitada a algoritmos de 31 *qubits*. A implementação, no simulador quântico, da função quântica de segmentação multi-escala ultrapassaria facilmente o número de *qubits* disponíveis nesse simulador. Apesar de se tratar de um simulador que é suportado por 64 processadores, facilmente é sobrecarregado se os algoritmos quânticos se aproximarem dos 31 *qubits*. Por este motivo, não é possível testar uma versão concreta da função quântica de segmentação multi-escala, a própria representação gráfica em esquema de portas quânticas seria densa. Assim, são apenas apresentados, descritivamente, os aspectos fundamentais desta função.

No Capítulo 4, onde é tratada a segmentação multi-escala, é referida uma técnica exaustiva que permite determinar o melhor nível de escala para obter a melhor segmentação. A técnica baseia-se no melhor grau de coincidência obtido por operações de convolução. Por cada escala, é necessário obter o nível de coincidência entre cada segmento resultante e cada *kernel*, de interesse, considerado. O objectivo é encontrar a escala que melhor define os resultados de segmentação estabelecidos nos *kernels*. Se o conteúdo da imagem a analisar variar na escala e variar na rotação e se o tipo de objecto variar também, o número de operações de convolução terá que crescer intensamente de acordo com essas variabilidades. Neste aspecto, é um problema semelhante ao que foi referido nas secções anteriores relativamente à detecção e à classificação de figuras. A aplicação do algoritmo de Grover a este problema é adequada, pois assim todas as operações de convolução podem ser executadas de uma só vez. A função a implementar quanticamente é descrita como se segue.

A função quântica tem que colocar em sobreposição os resultados do grau de coincidência resultante das convoluções a todas as escalas possíveis. Isto permitirá

¹² A função que determina um estado de interesse, em computação quântica, é designada por Oracle, em associação às entidades da Grécia antiga que representavam um conhecimento elevado.

comparar os diversos graus de coincidência, entres os vários níveis de escala. Posteriormente é seleccionada a escala que tiver tido o maior grau de coincidência. Seria assim necessário construir uma função quântica que directamente determinasse o estado de maior valor [75].

Neste processo revela-se uma outra vantagem da computação quântica: a possibilidade de melhorar a eficácia funcional e não só a temporal, embora isso seja conseguido pela permissão de execução intensa. No entanto, é necessário, na função quântica, um processo de construção do *kernel*, para que este varie, não só em escala e rotação, mas também na sua composição relativa. Por exemplo, no caso dos glóbulos brancos, a utilização de um *kernel* de uma determinada célula a várias escalas e posições não é suficiente, devido à diversidade de núcleos e relativos citoplasmas. Também aqui o problema do universo de padrões incompletos se põe, mas sem as exponenciais imposições temporais. Também aqui a possibilidade de resolver este problema pode passar pela descrição prévia das regiões a procurar e só em fase posterior, com o resultado desses descritores, é que se executa a operação de convolução. Uma possibilidade para parametrizar o *kernel* com factores além da posição, rotação e escala seria a de criar primitivas de núcleo e de citoplasma que permitissem ser controladas por parâmetros. Esses parâmetros seriam: o seu posicionamento relativamente ao citoplasma, o seu fraccionamento, a posição relativa dos seus próprios fragmentos, a sua elasticidade e rigidez, as suas dimensões, a sua textura e a sua cor.

A construção quântica da pilha multi-escala não foi considerada, mas poderia tê-lo sido através de abordagens adicionais.

6.12 Sistema híbrido

Embora os computadores quânticos ainda só existam em laboratório, é possível que, mais tarde ou mais cedo, estes venham a ser disponibilizados em maior escala. Contudo, é demonstrável que, pela sua natureza, um computador quântico não é uma máquina universal de computação [76]. Isto inibe a possibilidade de se ter uma máquina computacionalmente útil com uma organização estritamente quântica. É assim necessário que os sistemas quânticos sejam um híbrido entre o próprio processo

quântico e um computador clássico (Figura 6.7). O processador clássico terá a função de executar o algoritmo e, à medida que vão aparecendo operações de execução quântica, fará pedidos ao equipamento físico que implementa as operações quânticas. Após este passo, são executadas observações, pelas interfaces do sistema quântico, de forma a projectar os estados quânticos em estados clássicos. Actualmente, estes processos têm sido ensaiados a nível óptico, trabalhando-se com fotões, e a nível de ressonância magnética, trabalhando-se com *spins* atômicos. Mas outros métodos, como os pontos quânticos (*quantum dot*), encontram-se também em desenvolvimento.

No caso do processamento de imagens, a aquisição, num sistema híbrido, terá lugar de forma convencional; ou seja, através da integração de unidades de aquisição digital de imagens no computador clássico. Poderá especular-se acerca da possibilidade de introduzir directamente os fotões, provenientes de uma imagem, num processador quântico. Fisicamente, desde que se conheça a base de polarização, não parece haver impedimento algum. No entanto, isolar fotões impõe um conjunto de dispositivos de alta precisão [62] e devido à natureza computacionalmente não universal das máquinas quânticas continuará a ser necessária a interligação à componente de computação clássica.

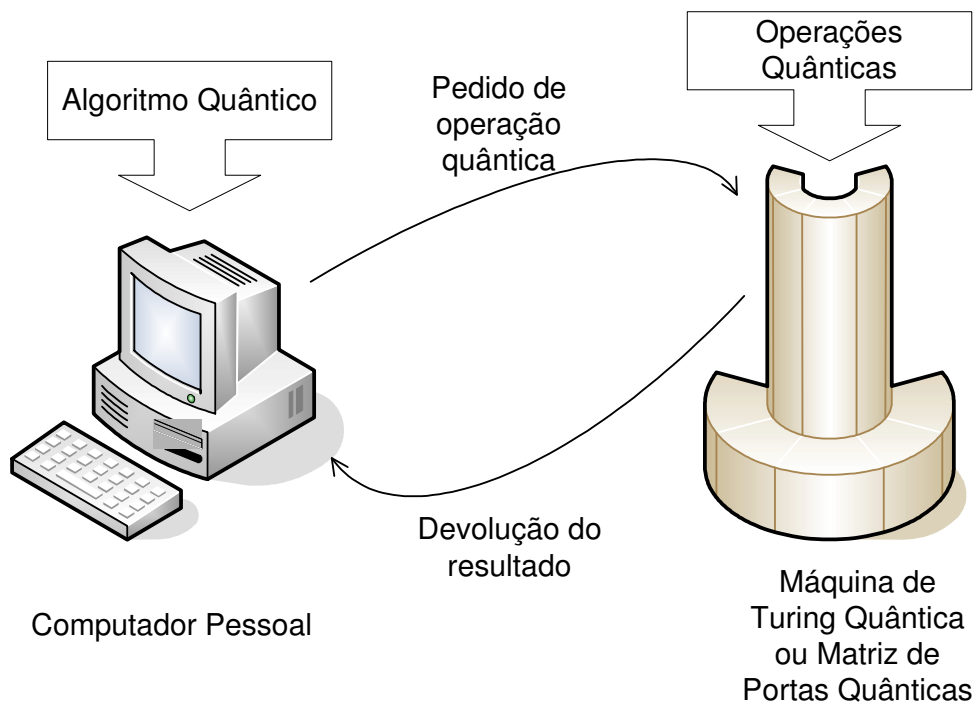


Figura 6.7 Sistema de processamento quântico híbrido

6.13 Considerações finais

O algoritmo de Grover não permite atingir ordens de resolução logarítmica e portanto não permite resolver eficazmente qualquer problema do tipo **NP** (non-deterministic polynomial time) ou **NP completo**. Por outro lado, permite acelerações de ordem quadrática, o que, para certos problemas, como é o caso da visão artificial, é bastante útil. Para que a execução do algoritmo de Grover seja possível numa ordem quadrática, é necessário que a ordem de execução do próprio Oracle não seja superior à quadrática. Esta exigência foi cuidada sobre o Oracle apresentado.

A abordagem aqui exposta pretende essencialmente mostrar uma possível configuração para a integração de soluções de processamento de imagem em algoritmos quânticos.

Como se pode deduzir, esta temática tem um sem fim de aplicações. São precisamente as aplicações como as de visão artificial que hoje em dia fazem denotar a precariedade do poder computacional clássico para resolver muitas situações. Aplicações de resolução temporalmente exponencial, como o conhecido problema do caixeiro-viajante, têm resolução menos exponencial, não polinomial, nas técnicas de processamento quântico. O problema do caixeiro-viajante, tal como parte de problemas sobre processamento de imagens naturais, encontra actualmente solução nos computadores clássicos através de heurísticas ou de aproximações que no caso da visão acabam por destruir parte da sua utilidade em problemas reais. As duas frentes que se opõem normalmente ao sucesso destes problemas caracterizam-se quase sempre na vertente temporal e na vertente da eficácia do bom funcionalismo e é confirmável que ambas se relacionam intensamente. Embora a computação quântica não pareça ser a solução para todos os problemas da visão artificial, pode ser um bom complemento aos algoritmos de visão, que são cada vez mais completos.

Capítulo 7

Classificação de Imagens por Interferência Quântica

O capítulo anterior mostra como podem ser resolvidos, em suportes de computação quântica, alguns problemas de processamento de imagens. Estes suportes são estabelecidos organizacionalmente por primitivas de portas quânticas dirigidas e estruturadas a um determinado problema. Porém, as estruturas obtidas podem ser complexas. Este facto pode tornar ainda mais difícil a sua implementação em máquinas quânticas. O presente capítulo procura avaliar a possibilidade de se retirarem ajudas de mecanismos quânticos, mas sem se recorrer a uma organização de componentes quânticos tão complexa, aproveitando fenómenos físicos, também existentes no domínio quântico, como a difracção e a interferência. Para se obter consistência e prova na representação dos fenómenos de difracção quântica que possibilitam a interferência, foram apresentadas funções de onda de difracção para orifícios circulares a partir de equações de onda quântica, ao invés de se partir de equações de onda electromagnética.

7.1 Introdução

As relativas incapacidades computacionais que os dispositivos actuais continuam a apresentar quando se pretende obter processamentos a sinais de dimensões elevadas conduz a que se procurem soluções de suporte físico inerentemente mais rápidas. As nanotecnologias que têm surgido nos últimos tempos são normalmente propostas como soluções genéricas capazes de dar progressão à Lei de Moore sem, no

entanto, poderem acelerar os processos de forma exponencial, contrapondo a necessidade exponencial de processamento exigido normalmente em problemas se inseridos em aplicações de complexidade **NP**. O trabalho apresentado neste capítulo pretende explorar a possibilidade de se aproveitarem certas características da mecânica quântica para se conseguirem classificadores de figuras bidimensionais, nos quais o resultado do processamento à imagem é atingido de forma directa e não num processamento sequencial. Este tipo de aproveitamento poderia chamar-se aproveitamento quântico não dirigido ou menos dirigido, em oposição ao aproveitamento dirigido associado aos algoritmos quânticos.

No século XVII, Isaac Newton apresentou um conjunto de ideias que indicavam que a luz deveria ser entendida como um fluxo de partículas. Até ao século XIX poucos foram os que se opuseram à teoria de Newton. Mas, nesse século, Thomas Young e outros apresentaram experiências que demonstravam que a teoria da luz de Newton poderia estar errada. Eles propunham que a luz fosse percebida através do movimento ondulatório. O comportamento de interferência que é possível observar quando duas ondas, provenientes de fontes diferentes, se encontram, por exemplo, numa superfície de água, foi observado por Thomas Young também na luz. A experiência ficou conhecida por “Double-slit”, na qual duas ranhuras colocadas a uma determinada distância permitiam a obtenção das duas fontes de luz. O resultado permitiu detectar um padrão de interferência, tal como acontece na água. Porém, já no final do século XIX novas experiências deixaram de poder ser explicadas a partir da teoria ondulatória da luz. Quando uma porção de metal se encontra carregada negativamente e é colocada sob uma luz ultravioleta, a carga é perdida. Mas o mesmo já não acontece se esse metal for atingido por luz, no espectro visível. Este efeito, conhecido por fotoelétrico, foi explicado por Albert Einstein. A explicação aprovava novamente a ideia de Newton de que a luz poderia ser entendida como um fluxo de partículas. Einstein propunha que essas partículas de luz fossem vistas como porções de energia. As partículas de luz ultravioleta tinham energia suficiente para provocar a descarga do metal ao invés da luz visível, que, por mais intensidade que tivesse, não provocava qualquer descarga. Só em 1920, com as teorias de mecânica quântica de físicos como Heisenberg, Schrödinger e Dirac, é que foi possível chegar a um consenso sobre o paradoxal comportamento da luz. A mecânica quântica estabelece meios que permitem prever a dinâmica de partículas, tal como se consegue prever a dinâmica de objectos com a mecânica clássica.

A teoria quântica permite também o esclarecimento do paradoxo onda/partícula. Não existe analogia perfeita entre o mundo quântico e o mundo clássico, porque, nas suas essências, são de natureza diferente. Mas é possível, dentro de certas condições, analisar elementos quânticos como ondas e em outras condições analisá-los como partículas.

7.2 Interferência Quântica

De forma a se entender melhor o mecanismo de interferência quântica, vai proceder-se à descrição de duas experiências recorrendo a utensílios do mundo clássico. Na primeira experiência, Figura 7.1, um dispositivo automático de disparo de projecteis em série vai actuar como se de uma fonte de partículas quânticas se tratasse. Posteriormente à fonte de projecteis, é colocado um plano com duas ranhuras posicionadas de forma simétrica relativamente ao centro do plano. Na seguinte posição, e já depois do plano, é colocado um detector que conta o número de projecteis que chegam a cada uma das sub-regiões do detector. Os disparos são feitos de forma aleatória e em todas as direcções. Apenas um projectil atravessa uma das ranhuras de cada vez. Desta forma é possível determinar a probabilidade de um projectil atingir uma dada sub-região, x , do plano detector. É então determinada essa probabilidade no caso da primeira ranhura estar aberta e a segunda estar fechada, $P_1(x)$. É também determinada essa probabilidade estando fechada a primeira ranhura e aberta a segunda, $P_2(x)$. Por último, a probabilidade é medida com as duas ranhuras abertas, $P_{12}(x)$.

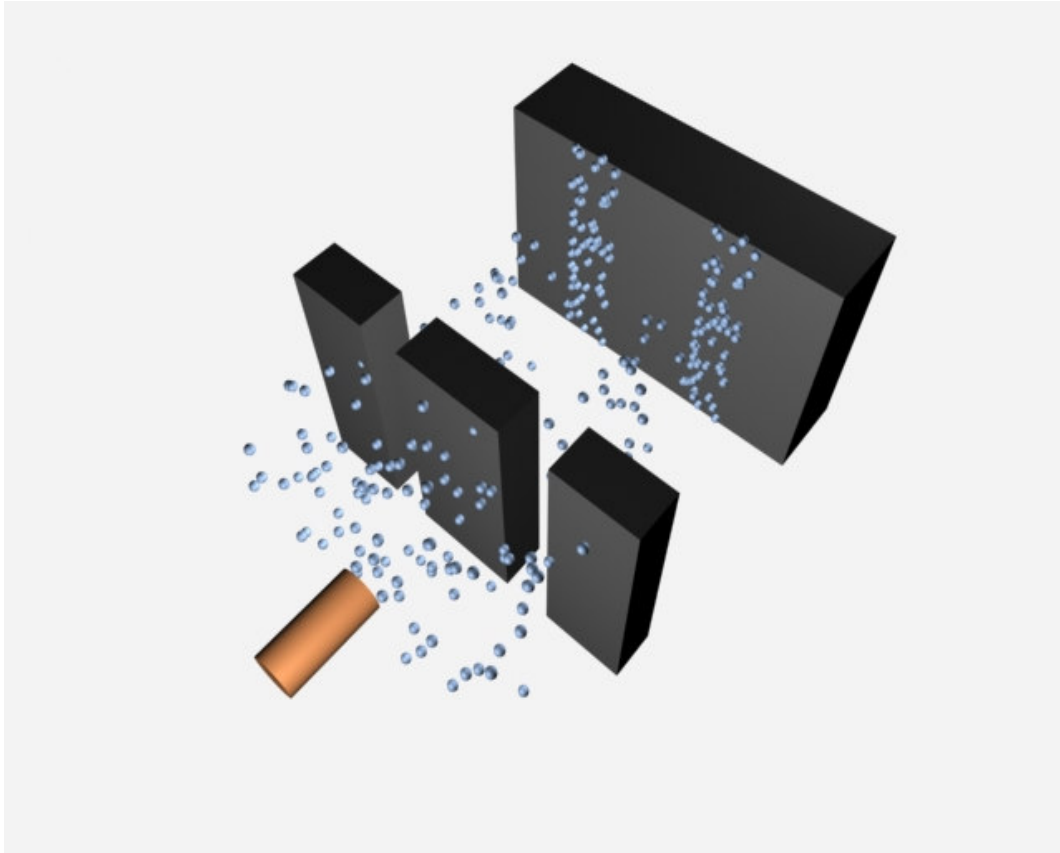


Figura 7.1 Experiência das “Duas Ranhuras” com projéteis clássicos

Pode então deduzir-se que a probabilidade de um projétil atingir um determinado ponto do plano detector, quando as duas ranhuras estão abertas, é dada pela expressão (7.1):

$$P_{12}(x) = \frac{1}{2}(P_1(x) + P_2(x)). \quad (7.1)$$

A segunda experiência, Figura 7.2, possui uma configuração semelhante; mas a fonte de projéteis é substituída por um dispositivo que produz um movimento ondulatório e que o transmite a um meio como a água, originando ondas. Essas ondas propagam-se até ao plano que contém as duas ranhuras. Parte dessas ondas continua a sua propagação através das ranhuras até que atinge o plano de detectores que aqui têm a particularidade de medir a intensidade (energia) com que a onda os alcança, de acordo com a Equação (7.2):

$$I(x) = |h(x)|^2. \quad (7.2)$$

Em que $|h(x)|$ corresponde à amplitude de onda e é dada através da Equação (7.3):

$$h(x) = e^{i\Phi(x)}. \quad (7.3)$$

Obtendo a intensidade $I_1(x)$ com a primeira ranhura aberta e a segunda fechada; obtendo a intensidade $I_2(x)$ com a primeira ranhura fechada e a segunda aberta; e, por fim, obtendo os valores das intensidades $I_{12}(x)$ para os diferentes valores de x , com as duas ranhuras abertas. Pode verificar-se que este último resultado está de acordo com a interferência da teoria das ondas, onde em alguns valores de x as energias das ondas se anulam e onde em outros pontos de x as energias das ondas se reforçam. Esta construção de interferência é dada pelo quadrado da amplitude da soma das equações de onda para um dado valor de x , Equação (7.4).

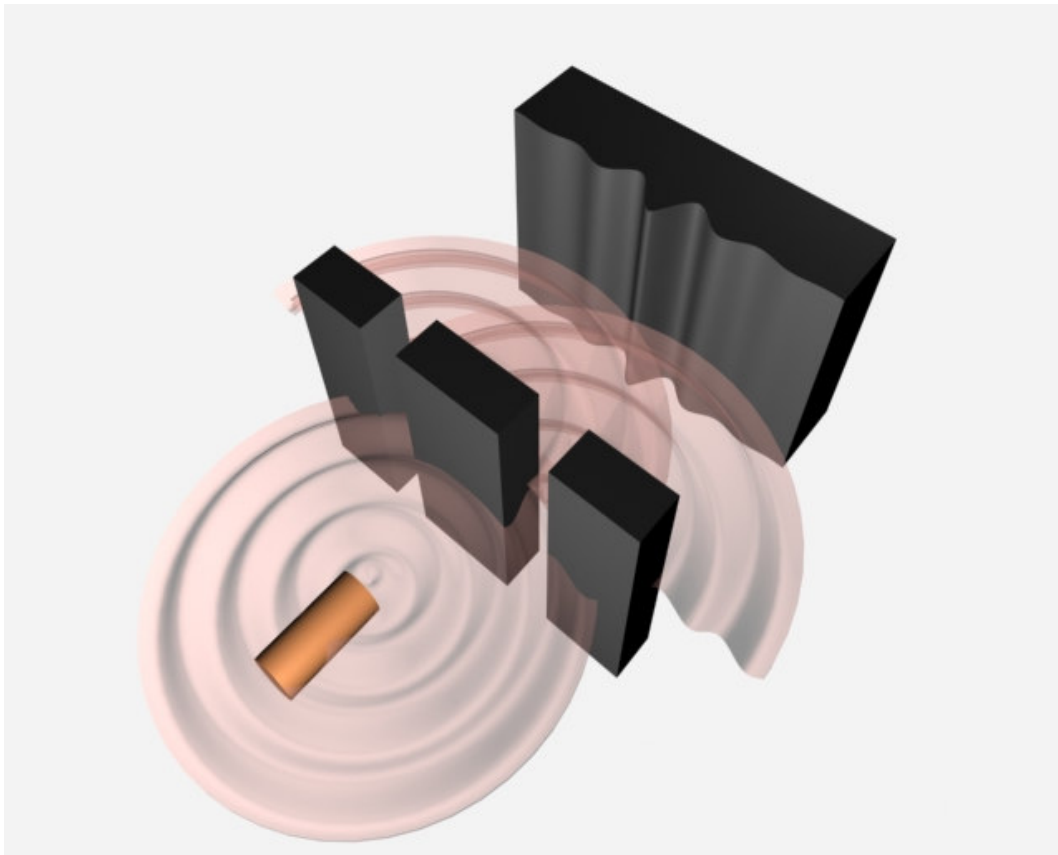


Figura 7.2 Experiência das “Duas Ranhuras” com ondas

$$I_{12}(x) = |h_1(x) + h_2(x)|^2. \quad (7.4)$$

A terceira experiência, Figura 7.3, é semelhante à primeira; mas em vez de se usarem projecteis usam-se electrões.

O plano que está colocado a seguir à fonte de electrões tem igualmente duas pequenas ranhuras distanciadas a uma escala quântica e com uma secção, também, de dimensões quânticas. O plano detector permite a contagem do número de electrões que acabam por atingir cada elemento da superfície desse plano. Também aqui as distribuições parciais podem ser verificadas abrindo e fechando cada um dos furos de forma complementar, $P_1(x)$ e $P_2(x)$. Essas distribuições estão de acordo com aquilo que se esperava intuitivamente e com o observado na primeira experiência feita com os projecteis. Mas a distribuição, $P_{12}(x)$, resultante da abertura dos dois furos em simultâneo não coincide nem com o esperado, nem com o obtido na primeira experiência, sendo semelhante à distribuição de interferência conseguida na segunda experiência. Contrariando o senso comum, existem posições no plano detector que recebem menos electrões quando as duas ranhuras estão abertas do que quando uma delas está fechada. Daqui se deduz que o electrão, mesmo sendo uma partícula, apresentou um comportamento semelhante a uma onda. Se se tentar perceber, em termos de partículas, o fenómeno e tendo em consideração que é produzido um electrão de cada vez, então cada electrão aparenta conhecer o trajecto do seu precedente. Ou então, se se comparar com o que acontece com as ondas, a partícula electrão comporta-se como que atravessasse ambas as ranhuras em simultâneo, de forma a produzir a interferência. Não é possível deterministicamente saber o trajecto para um electrão, mas é possível conhecer alguns elementos da curva probabilística descrita no detector. Com as duas ranhuras abertas, existem posições do detector que registam um número de zero electrões, enquanto que outras posições registam um aumento de electrões relativamente à situação de fecho de uma das ranhuras. As posições de mínimos e de máximos locais estão distanciadas segundo a seguinte relação:

$$\frac{l\lambda}{L}. \quad (7.5)$$

Na qual, l corresponde à distância entre os dois planos, λ ao comprimento de onda dos electrões e L à distância entre as duas ranhuras.

Pretendendo determinar-se a curva de probabilidades de um electrão atingir uma posição no detector, tal como se fez para o caso da onda, é necessário recorrer a uma representação por fasores. Assim, $P_1(x)$, referente à condição em que a primeira ranhura está aberta e a segunda fechada, é dado por:

$$P_1(x) = |\psi_1(x)|^2. \quad (7.6)$$

Na situação complementar, referente às ranhuras, $P_2(x)$ é obtido em (7.7):

$$P_2(x) = |\psi_2(x)|^2. \quad (7.7)$$

A equação de probabilidades resultante da interferência ao abrir as duas ranhuras é mostrada em (7.8):

$$P_{12}(x) = |\psi_1(x) + \psi_2(x)|^2. \quad (7.8)$$

Decompondo (7.8) resulta em (7.9):

$$\begin{aligned} P_{12}(x) &= |\psi_1(x) + \psi_2(x)|^2 = |\psi_1(x)|^2 + |\psi_2(x)|^2 \\ &\quad + 2 \operatorname{Re}(\psi_1(x)) \operatorname{Re}(\psi_2(x)) + 2 \operatorname{Im}(\psi_1(x)) \operatorname{Im}(\psi_2(x)), \quad (7.9) \\ &= P_1(x) + P_2(x) + 2 \operatorname{Re}(\psi_1(x)) \operatorname{Re}(\psi_2(x)) + 2 \operatorname{Im}(\psi_1(x)) \operatorname{Im}(\psi_2(x)) \end{aligned}$$

onde $\operatorname{Re}(y)$ e onde $\operatorname{Im}(y)$ correspondem à parte real e à parte imaginária do número imaginário y . O valor da amplitude de onda ao estar ligado a um valor de fase permite obter um desenvolvimento matemático que explica a existência de interferência ao serem usados electrões em vez de projecteis. Estando as oscilações obtidas na curva de probabilidade P_{12} inerentemente associadas ao termo:

$$2 \operatorname{Re}(\psi_1(x)) \operatorname{Re}(\psi_2(x)) + 2 \operatorname{Im}(\psi_1(x)) \operatorname{Im}(\psi_2(x)). \quad (7.10)$$

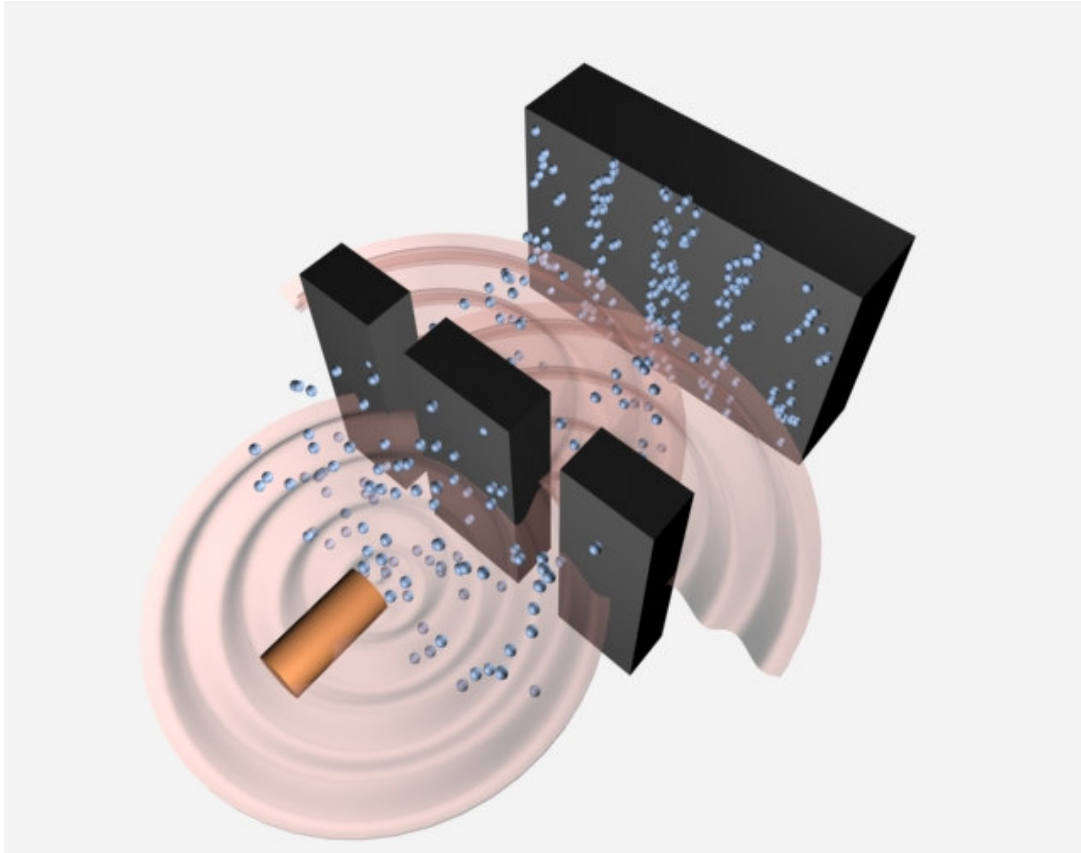


Figura 7.3 Experiência das “Duas Ranhuras” com partículas quânticas

Mesmo que os electrões deixem a sua origem com um intervalo de anos, em vez de com um intervalo ínfimo de tempo, a curva de probabilidades obtida será exactamente a mesma.

Estas experiências permitem demonstrar a particular essência de toda a mecânica quântica.

7.3 A natureza da probabilidade quântica

A natureza aleatória dos disparos feitos na experiência dos projecteis pode ser atribuída à falta de rigor com que o dispositivo lançador foi construído ou então a uma aleatoriedade provocada com o propósito de se conseguir este comportamento. Quer isto dizer que, se o lançador for construído com um rigor elevado, o fenómeno probabilístico que lhe está associado deixará de existir. Poderia pensar-se que o mesmo se passaria com a fonte de elementos quânticos; que uma construção cuidada da fonte de partículas

quânticas eliminaria os trajectos aleatórios que resultam na escolha de uma ranhura ou da outra. Mas a natureza aleatória, implícita nos mecanismos quânticos, não é do mesmo tipo que o da natureza associada à experiência dos projecteis, o qual se prende com imperfeições na construção da máquina. A natureza da aleatoriedade quântica é parte da física e não é possível eliminá-la de forma alguma. Esta lei da física quântica está associada ao princípio de incerteza de Heisenberg, do qual é extraída a seguinte desigualdade $\Delta x \Delta p \geq \frac{1}{2} \hbar$, em que o produto da incerteza da posição x de uma partícula com a incerteza do *momento* p dessa partícula nunca é inferior a um meio da constante de Planck. Todas as tentativas de se observar a partícula de forma a ter-se informação determinística sobre o trajecto da partícula terminam com perda do padrão de interferência e com uma distribuição no alvo detector idêntica à distribuição verificada na experiência dos projecteis. Não existe nada que possa refinar a experiência, sem a perda do comportamento interferente, mantendo-se sempre o princípio da incerteza. Todas as teorias que propõem a existência de variáveis escondidas, naturais, que definiriam estados determinísticos das partículas, têm sido sistematicamente eliminadas com recurso a experiências. Essas experiências são dirigidas a provar que, se existissem estados bem definidos antes da observação da partícula, então o comportamento detectado nessas experiências seria outro. A conhecida desigualdade de Bell [77] valida essa situação, mesmo no caso experimental, conduzindo à não existência real da partícula antes da observação. Assim, a natureza da probabilidade quântica está enraizada nas leis da mecânica quântica.

7.4 A dinâmica de partículas

A dinâmica dos objectos à nossa escala pode ser descrita através de equações determinísticas que permitem conhecer e prever com precisão a posição que um objecto adquire no espaço depois de ser sujeito a uma força de determinada direcção e de determinada magnitude. A natureza quântica não permite conhecer com determinismo todas as “variáveis” em questão. O que se consegue saber, através dos dados quânticos iniciais, usando equações da dinâmica quântica, são resultados probabilísticos. É através da constante de Planck que se pode perceber uma das razões para esta desigualdade de comportamentos do mundo clássico para o mundo quântico. A conhecida relação de

Broglie possibilita o cálculo do comprimento de onda para um determinado *momento* de uma porção de matéria e é dada por:

$$\lambda = \frac{h}{p}, \quad (7.11)$$

onde h é a constante de Planck e p é o *momento* do objecto, dado por:

$$p = \nu m, \quad (7.12)$$

sendo ν a velocidade e m a massa do objecto em repouso.

O que se verifica na experiência das duas ranhuras é que, para haver interferência quântica, as duas ranhuras têm que estar separadas por uma distância da ordem do comprimento de onda das partículas que as atravessam. Então, se for calculado o comprimento de onda de um objecto do mundo clássico, como a constante de Planck é muito pequena e a massa relativamente elevada, obtém-se um comprimento de onda ainda mais pequeno do que as dimensões atómicas. Parece inibir-se assim a natureza quântica no mundo clássico. Outras situações podem também ser atribuídas à não existência de comportamento quântico no mundo clássico. Por exemplo, a função de onda de cada partícula constituinte de um objecto clássico comporta o seu colapso por acção de observações intrínsecas ao próprio objecto. Os fenómenos de propagação de variação térmica podem constituir uma forma dessas observações e o colapso da função de onda destrói o comportamento quântico.

7.5 Funções de onda para a difracção em orifícios circulares

A experiência das duas ranhuras encerra nela própria a natureza do comportamento quântico, principalmente na vertente onda-partícula, e é classicamente usada em demonstrações desse comportamento. A maior parte dos autores enfatiza a impossibilidade de se descrever o comportamento da experiência das duas ranhuras usando partículas. Defendem que a interferência, que é inerente a esta experiência, não é uma propriedade apresentada por partículas e é uma propriedade apresentada por ondas, deduzindo, então, que as partículas usadas na experiência das duas ranhuras, ao

exibirem interferência, estão a comportar-se estritamente como uma onda. Assim, a análise que é feita para se estudar e representar a experiência das duas ranhuras, normalmente, recorre à teoria de ondas em detrimento da mecânica quântica. Desta forma, as deduções apresentadas para a experiência das duas ranhuras provêm das bases de difracção da óptica clássica. Isto significa que, usando esse método, não se adquirem os resultados na sua verdadeira natureza quântica, que é a probabilidade de amplitude. Em oposição, alguns autores apresentaram possíveis soluções para a descrição desta experiência usando a mecânica quântica [78-82]. Em [78-80] são discutidas as possibilidades e as dificuldades de se obter uma função de onda conseguida directamente da resolução da equação de Schrödinger para as condições da experiência das duas ranhuras. É problemático estabelecer o conjunto de condições fronteira para as ranhuras, assim como a resolução matemática da equação de Schrödinger nessas condições. Porém, recorrendo ao integral de caminho de Feynman, Barut e Basri [81] apresentam uma solução exacta com base no formalismo quântico. Uma solução mais simples é dada por Marcella [82] partindo do postulado de Born. O postulado de Born refere que, se um sistema está inicialmente no estado $|\psi_{inicial}\rangle$, então a probabilidade P , resultante de uma observação, de encontrar esse sistema num estado final $|\psi_{final}\rangle$ é dada por:

$$P = \left| \langle \psi_{final} | \psi_{inicial} \rangle \right|^2. \quad (7.13)$$

O trabalho de Marcella consiste em tratar o problema através da combinação de subestados coadunados ao sistema em causa. Mas esse trabalho restringe-se ao sistema das ranhuras. Neste capítulo, como será percebido, o interesse situa-se num sistema de difracção não de ranhuras mas de orifícios. A descrição que é aqui desenvolvida tem como base o trabalho de Marcella, mas é dirigido ao sistema de orifícios e é apresentada a seguir.

O objectivo é calcular a probabilidade que uma partícula detém em sair de um orifício com um determinado ângulo θ . Na prática esta determinação pode ser feita com um detector cujo princípio de funcionamento assenta na sensibilidade sobre uma determinada gama de momento de uma partícula.

O próprio orifício pode ser visto como um sistema observador [83] que projecta a partícula num estado de posição $|\psi\rangle$. Então,

$$|\psi_{inicial}\rangle = |\psi\rangle. \quad (7.14)$$

Assim, o orifício causará a posição (x, y) da partícula no próprio orifício. Como posição e momento, p , não são atributos comutativos, a partícula que atravessa o orifício apresentará um momento no qual as suas componentes x e y são incertas. O outro elemento observador será um detector que permitirá conhecer os ângulos de difracção. A relação que existe entre as componentes do momento e os ângulos de difracção é dada por:

$$\begin{aligned} p' &= p \sin(\theta) \\ p_x &= p' \cos(\phi), \\ p_y &= p' \sin(\phi) \end{aligned} \quad (7.15)$$

onde $-\pi/2 \leq \theta \leq \pi/2$ e onde $-\pi \leq \phi \leq \pi$. θ refere o valor do ângulo que o vector do momento da partícula faz com o eixo que atravessa o orifício. ϕ refere o valor do ângulo que a projecção, obtida por $\sin(\theta)$, do vector do momento da partícula adquire. Como se verá, o ângulo ϕ , quando modificado, devido à total simetria circular que o fenómeno da difracção apresenta num orifício circular, não altera o valor probabilístico de difracção da partícula. O mesmo já não acontece para o ângulo θ .

Conhecer estes dois ângulos consiste em observar as componentes do momento p_x e p_y . Assim, uma partícula difractada com um determinado ângulo θ e um determinado ângulo ϕ é caracterizada num estado próprio de momento $|p_{xy}\rangle$:

$$|\psi_{final}\rangle = |p_{xy}\rangle. \quad (7.16)$$

A probabilidade de que uma partícula seja difractada pelo orifício, com momento $p_{xy} = (p_x, p_y)$, é obtida em:

$$P(p_{xy}) = \left| \langle p_{xy} | \psi \rangle \right|^2. \quad (7.17)$$

Na representação para o atributo de posição (x, y) , a função própria de momento para uma partícula livre num espaço bidimensional é:

$$\langle xy | p_{xy} \rangle = e^{i(p_x/\hbar)x} e^{i(p_y/\hbar)y} A, \quad (7.18)$$

onde A representa um factor normalizador [84]. A amplitude de probabilidade final terá que ser considerada para todas as posições observáveis pelo orifício, então:

$$\langle p_{xy} | \psi \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle p_{xy} | xy \rangle \langle xy | \psi \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A e^{-i(p_x/\hbar)x} e^{-i(p_y/\hbar)y} \psi(x, y) dx dy, \quad (7.19)$$

em que $\langle xy | \psi \rangle$ é dado pela função de onda $\psi(x, y)$ que caracteriza a amplitude de onda posicional de percurso da partícula pelo orifício e a amplitude de onda posicional do percurso da partícula pela região externa ao orifício. A probabilidade de a partícula atingir uma determinada posição do orifício é inversamente proporcional à área desse e é zero para qualquer ponto fora do orifício:

$$\psi(x, y) = \begin{cases} 1/\sqrt{\pi a^2} & \sqrt{x^2 + y^2} \leq a \\ 0 & \text{outras posições} \end{cases}, \quad (7.20)$$

em que a representa o raio do orifício.

Através da substituição de x e de y :

$$\begin{aligned} x &= \rho \cos \alpha \\ y &= \rho \sin \alpha \end{aligned} \quad (7.21)$$

é possível realizar a integração da Equação (7.19) usando coordenadas polares:

$$\langle p_{xy} | \psi \rangle = \int_0^a \int_0^{2\pi} A e^{-i(p_x/\hbar)\rho \cos \alpha} e^{-i(p_y/\hbar)\rho \sin \alpha} \frac{1}{\sqrt{\pi a}} \rho d\rho d\alpha. \quad (7.22)$$

Recorrendo-se a transformações trigonométricas a Equação (7.22) pode ser rescrita como:

$$\langle p_{xy} | \psi \rangle = \int_0^a \int_0^{2\pi} A e^{-i \left(\frac{\rho}{\hbar} \sqrt{p_x^2 + p_y^2} \cos \left(\alpha - \arctg \left(\frac{p_y}{p_x} \right) \right) \right)} \frac{1}{\sqrt{\pi a}} \rho d\rho d\alpha. \quad (7.23)$$

Nesta expressão é reconhecível o integral que representa uma função de Bessel e que é escrito como:

$$\frac{i^{-n}}{2\pi} \int_0^{2\pi} e^{ix \cos \beta} e^{in\beta} d\beta = J_n(x). \quad (7.24)$$

Fazendo as correspondentes substituições, a Equação (7.23) fica reduzida a:

$$\langle p_{xy} | \psi \rangle = A \frac{2\pi}{\sqrt{\pi a}} \int_0^a J_0 \left(\frac{\rho}{\hbar} \sqrt{p_x^2 + p_y^2} \right) \rho d\rho. \quad (7.25)$$

A resolução deste integral é conseguida utilizando a seguinte relação recursiva das funções de Bessel:

$$\frac{d}{dx} \{ x^{n+1} J_{n+1}(x) \} = x^{n+1} J_n(x). \quad (7.26)$$

Sendo $n = 0$, a integração deduzida da Equação (7.26) resulta em:

$$\int_0^x x' J_0(x') dx' = x J_1(x). \quad (7.27)$$

Realizando-se as adequadas transformações na Equação (7.25), respeitando a Equação (7.27), resulta na seguinte solução:

$$\langle p_{xy} | \psi \rangle = A \sqrt{\pi a} \frac{2J_1 \left(\frac{a}{\hbar} \sqrt{p_x^2 + p_y^2} \right)}{\frac{a}{\hbar} \sqrt{p_x^2 + p_y^2}}. \quad (7.28)$$

Substituindo, em (7.28), $\sqrt{p_x^2 + p_y^2}$ pelo seu equivalente $p \sin(\theta)$, a apresentação de $P(p_{xy}) = |\langle p_{xy} | \psi \rangle|^2$ origina:

$$P(p_{xy}) = A^2 \pi a^2 \left[\frac{2J_1\left(a \frac{p}{\hbar} \sin(\theta)\right)}{a \frac{p}{\hbar} \sin(\theta)} \right]^2, \quad (7.29)$$

que facilmente se reconhece como a expressão da difracção de Fraunhofer no âmbito da teoria óptica.

Esta equação suportará as simulações computacionais que vão ser descritas nas posteriores secções.

7.6 Interferência quântica aplicada à imagem

Nesta secção apresentar-se-á uma forma de aproveitar os mecanismos de interferência quântica para estabelecer um classificador que poderá descrever uma região ou um contorno de uma imagem. Deve notar-se que não se pretende aqui introduzir o detalhe físico que possibilitaria a implementação do classificador. Uma abordagem de implementação levaria a que se considerassem particularidades de índole física que *apenas* condicionam a possibilidade e a facilidade de implementação de um mecanismo deste tipo e que pouco condicionam a característica essencial do processo, que é a interferência. Além disso, actualmente, os meios para o conseguir não são de existência comum, embora esse passo pudesse ser exposto numa perspectiva unicamente teórica. No entanto, o objectivo foi dirigido à validação do fenómeno de interferência. A utilidade deste mecanismo, sobre implementação física, prende-se com a necessidade de se possuírem cada vez mais soluções de processamento que inerentemente paralelizam o processo, em vez de se recorrer a soluções baseadas em sistemas que serializam o processo. Soluções serializantes, mesmo com as velocidades de relógio dos processadores actualmente existentes, não são suficientes para os requisitos da maior parte das tarefas de visão que se pretende que estejam ao nível da capacidade humana. Ao referir-se a capacidade humana, estão a abranger-se essencialmente três vertentes: a da velocidade com que as imagens são processadas, a complexidade funcional do processo e a dimensão espacial do processo. A dimensão do processo está normalmente associada à própria complexidade, que tende a conduzir ao bom desempenho apresentado pelos humanos. É verdade que, sem se entenderem os

mecanismos complexos que fazem emergir um elevado desempenho do sistema de visão humano, não será muito útil possuírem-se soluções que apenas oferecem rapidez. Mas o que se tem verificado recentemente é que muitas das hipóteses que são oferecidas pela comunidade científica, explicadoras dos processos cerebrais, não conseguem ser depois muito proveitosas devido à sua dimensão, que é funcionalmente executada, maioritariamente, de forma sequencial. O modelo híbrido das PCNNs apresentado num capítulo anterior é também exemplo dessa dificuldade. Mais do que isso, por vezes é mesmo difícil, se não impossível, a sua experimentação e simulação por limitação física dos suportes, seja em termos de memória ou em termos de velocidade. Isto é válido, embora não tão vincadamente, mesmo se se pensar em computação paralela, que pode recorrer a alguns milhares de processadores e a centenas de *Terabytes* de memória. Estas condicionantes levam a que novos métodos sejam caracterizados e testados de forma insuficiente, sem que se possa levá-los a uma dimensão verdadeiramente próxima da dimensão a que estes se constroem no córtex visual humano. Com a análise aqui apresentada, ao processo de interferência, como órgão classificador, pretende-se sugerir que, sendo um processo inerentemente paralelizado, de interferência quântica, permite obter respostas temporalmente menos dependentes da dimensão do problema.

De qualquer forma, da pesquisa que se fez, podem apontar-se algumas estruturas físicas quânticas que permitem construir padrões de interferência e que poderiam ser adaptadas a uma imagem de *pixels*. É o caso de uma extensão à experiência das duas ranhuras, na qual é usada uma matriz de orifícios e na qual a fonte quântica, por exemplo, produz fotões. A combinação das funções de onda destes fotões estabelece o padrão de interferência. Outra estrutura quântica que permite observar interferência é organizada a partir de mecanismos de efeito de túnel. Átomos de ferro são colocados numa superfície de cobre e arrastados, por efeito túnel, para as localizações pretendidas [85]. Neste caso os átomos de ferro representariam os *pixels*. É possível depois observar, com um microscópio de efeito de túnel, as amplitudes de probabilidade da função de onda interferente, resultante do confinamento provocado pelos átomos de ferro nos electrões da superfície de cobre. A simulação computacional deste mecanismo obriga à resolução da equação de Schrödinger para confinamentos arbitrários, o que na maior parte dos casos não é elementar (ver anexo A). A solução mais clara consiste em combinar uma série de funções de Bessel em que cada uma se relaciona directamente com a posição dos elementos confinadores. No entanto, a estrutura quântica que parece mais viável e simples de utilizar no classificador apresentado é uma estrutura que tem

como base os fótons, ou mesmo electrões, e a “experiência das duas ranhuras”. Actualmente, uma forma vulgarizada de controlo da passagem de um fóton gerado por excitação electrónica é conseguida pela acção de pequenos campos eléctricos sobre os chamados *cristais líquidos*. Estes permitem a passagem dos fótons em função da polarização que o fóton possui e em função da polarização dos próprios cristais que é modificada pela acção do campo eléctrico a que estes estão submetidos. Assim, as ranhuras (orifícios) podem ser dinâmicas se cada uma for constituída por elementos dessa natureza.

Uma vantagem prática de se usar interferência ao nível quântico, partícula a partícula, é a de minimizar-se o custo energético na obtenção da operação, bem como o espaço ocupado pelos seus órgãos. Em qualquer caso, uma vez que a descrição da interferência quântica pode assentar em equações de onda, seria possível apresentar mecanismos de base não quântica directa, mas que permitissem, de igual forma, a interferência. Por exemplo, fazer passar um feixe de luz através de ranhuras de dimensão e inter-espacamento adequado permite o surgimento de padrões de interferência que podem ser adquiridos por uma vulgar câmara de vídeo.

7.6.1 Pixels numa matriz de orifícios

Supondo-se que a região ou o contorno a classificar, a partir da interferência, se encontram representados binariamente, nessa imagem, os valores dos *pixels* concordam com os valores definidos no conjunto $\{0,1\}$. Nesta situação, a utilização de orifícios colocados ao longo de uma superfície bidimensional segue a mesma estrutura matricial que numa imagem de *pixels*. Pontos da matriz com *pixels* a 1 corresponderão a pontos na superfície da barreira que estão num estado de permissão de passagem dos fótons. Mas, se se quiser representar valores de *pixels* numa variação mais ampla que o conjunto $\{0,1\}$, por exemplo, num conjunto $\{0,1,2 \dots 255\}$, pode-se fazê-lo recorrendo a diferentes valores de comprimento de onda dos fótons ou até mesmo de electrões, modificando o valor do *momento* destes. No entanto, esta solução levaria a que tivesse de existir uma fonte de partículas quânticas, com uma determinada abertura, para cada orifício da barreira. Outra solução será a de estabelecer na barreira de orifícios tamanhos distintos de forma a obter diferentes ângulos de difracção em associação aos diferentes valores de *pixels*. Esta solução exigiria mecanismos mais complexos de construção na obtenção de aberturas variáveis. Uma forma possível seria a de se usarem dispositivos

LCD (Liquid Crystal Display) com luz polarizada, na qual a probabilidade de um fóton ultrapassar um orifício seria proporcional ao nível de similaridade de polarização entre o fóton e os cristais polarizados pelo controlador do LCD. Esta variação nas probabilidades de passagem de cada orifício codificaria de forma correspondente o padrão de interferência, em acordo com as intensidades de pixels pretendidas. Mais uma vez, o processo seria válido em termos ópticos utilizando agregados de fótons, em vez de um único fóton. A simulação por *software* do processo de interferência exige um esforço computacional que aumenta intensamente com o número de *pixels* a considerar. Mas é precisamente o produto que está em adjacência a este esforço que se espera que represente informação mais completa da relação entre os *pixels* e que se espera que permita uma classificação mais acertada das figuras e regiões codificados pela interferência. Substituindo, no futuro, essa simulação, por um dispositivo físico que construa intrinsecamente a interferência, o processo teria uma ordem de execução muito menos elevada.

7.6.2 Método de interferência

Geralmente a caracterização de uma região ou de um contorno é obtida, seja por métodos clássicos, menos clássicos, ou biológicos, pela conjugação relativa das existências espaciais dos *pixels* em associação aos seus correspondentes valores. Os processamentos e as análises feitas a esses conjugados *pixels* permitem obter níveis semânticos diferentes de acordo com um objectivo particular. Porém, normalmente, cada problema exige um processo singular que apenas abrange aquela situação específica, o que conduz a que na prática apareçam quase sempre situações novas para as quais os métodos de processamento e de análise não estavam dedicados. Ao escolher-se um mecanismo de interferência espera-se que a “existência” de um *pixel* de um contorno ou de um *pixel* de uma região seja de alguma forma distribuída aos *pixels* mais adjacentes e menos adjacentes. Mais do que isso, espera-se que a influência que um determinado *pixel* recebe dos outros se traduza, de alguma maneira, pelo caracterizar da região ou contorno que o afecta semanticamente de alguma forma.

A proposta aqui apresentada tanto pode assentar na equação de Schrödinger que descreve a dinâmica de partículas quânticas que constituem o sistema como nas equações de ondas electromagnéticas que descrevem a dinâmica de uma onda dessa natureza. Quer isto significar que o mesmo padrão de interferência poderia ser obtido

exclusivamente por dispositivos ópticos sem a preocupação de se trabalhar com partículas quânticas individualmente.

É sabido que um sistema quântico como o da “experiência das duas ranhuras” apresenta características de onda estacionária. Este facto implica que a equação de Schrödinger a aplicar aqui seja independente do tempo.

A função de onda geral nesta situação é conseguida pela combinação linear das funções de onda parciais de cada orifício obtidas quando um orifício é aberto e os restantes são fechados:

$$\Psi(x, y) = \frac{1}{\sqrt{(N+1)(M+1)}} [\Psi_{00}(x, y) + \Psi_{01}(x, y) + \dots + \Psi_{NM}(x, y)]. \quad (7.30)$$

Cada função de onda parcial terá um domínio espacial de actuação que depende, entre outros, do orifício que lhe corresponde e da grandeza desse mesmo orifício. Quanto maior o orifício, menor são os ângulos de difracção e por consequência menor é a sua área de incidência de partículas sobre o plano de detectores. Cada função parcial Ψ_{nm} corresponde ao orifício de coordenadas n, m , na matriz de dimensões NM , e representa a amplitude de onda para um determinado ponto (x, y) no plano de detectores com os restantes orifícios bloqueados.

É importante, por uma questão de equilíbrio probabilístico, que a probabilidade de uma partícula quântica atingir um qualquer orifício seja exactamente a mesma. Para isso, a fonte de partículas deve estar colocada a uma distância tal que exija uma abertura angular da fonte de partículas muito pequena. Assim, os limites posicionais dos orifícios da barreira de partículas devem possuir uma distância vertical e uma distância horizontal muito mais pequena do que a distância da barreira à fonte de partículas. Os fenómenos que estipulam a direcção das partículas à saída da fonte são verdadeiramente aleatórios e encarregar-se-ão de cumprir esse requisito de aleatoriedade em complemento à relativa pequena abertura angular da fonte.

A função de onda geral, $\Psi(x, y)$, ao ser o resultado das combinações das funções de onda parciais, estabelece um padrão de onda interferente. A probabilidade de se observarem partículas quânticas, no plano detector, numa posição arbitrária (x, y) , para um único orifício situado na posição (n, m) é dada por:

$$|\Psi_{nm}(x, y)|^2, \quad (7.31)$$

em que a função de onda Ψ corresponde à Equação (7.32) já deduzida em (7.29)

:

$$\Psi_{nm}(x, y) = A^2 \pi a^2 \left(\frac{2J_1(ka \sin(\theta))}{ka \sin(\theta)} \right), \quad (7.32)$$

onde J_1 representa a função de Bessel do tipo 1 e de ordem 1, a representa o raio da abertura do orifício e A representa um factor normalizador da amplitude máxima de onda. O k simboliza o *número de onda* e é dado por:

$$k = \frac{2\pi}{\lambda}, \quad (7.33)$$

com λ a representar o comprimento de onda da partícula.

O θ representa o ângulo formado pelo eixo da abertura com a semi-recta que parte do centro da abertura e termina no ponto de projecção de coordenadas (x, y) .

Assim, θ é dado por:

$$\theta = \tan^{-1} \left(\frac{\sqrt{(x-n)^2 + (y-m)^2}}{l} \right), \quad (7.34)$$

onde l representa a distância do plano de orifícios ao plano de projecção (plano detector). Segundo as leis da óptica, para que estas equações sejam válidas, é necessário que se cumpra o critério que permite a obtenção da difracção de Fraunhofer, que corresponde aos padrões de difracção em causa, ditado pelo número F de Fresnel:

$$F = \frac{a^2}{l\lambda} \ll 1. \quad (7.35)$$

A partir da Figura 7.4, da Figura 7.5 e da Figura 7.6, descreve-se a simulação em computacional do método. Suponha-se uma imagem com um conteúdo de vários pixels

“acesos” a formarem um contorno que constitui um retângulo, Figura 7.4. Cada um destes pixels contribui com uma função de onda ψ . Neste exemplo, são referidos três pixels. O pixel a , o pixel b e o pixel c . Cada um estabelece a sua correspondente função de onda, $\psi(a)$, $\psi(b)$, $\psi(c)$, em acordo com a Equação (7.32). Na Figura 7.6 é observável que as funções de onda resultantes expandem-se, a partir de cada um dos *pixels* da Figura 7.4, por todos os *pixels* do plano detector, Figura 7.5. Assim, a interferência de funções de onda acontece em dependência às posições relativas dos *pixels* considerados no plano formulador da figura. Em vez de uma figura, pode ser considerada uma região. A interferência entre a função de onda $\psi(a)$, a função de onda $\psi(b)$ e a função de onda $\psi(c)$ resulta num padrão de franjas, representado, de forma aproximada, na Figura 7.6. Cada elemento do plano detector aglutinará, assim, a probabilidade final de uma partícula atingir esse mesmo elemento. Genericamente a contribuição de amplitude de cada elemento da superfície de interferência é dada pela Equação (7.36),

$$|\Psi_{nm}(x, y)|, \quad (7.36)$$

onde (n, m) representa as coordenadas dos pixels (orifícios) da figura a codificar e que se situam no plano da Figura 7.4. O par (x, y) representa as coordenadas de cada elemento detector do plano da Figura 7.5.

Sobre a imagem resultante, que descreve um padrão de interferência, podem ser aplicados operadores simples de descrição numérica desse resultado.

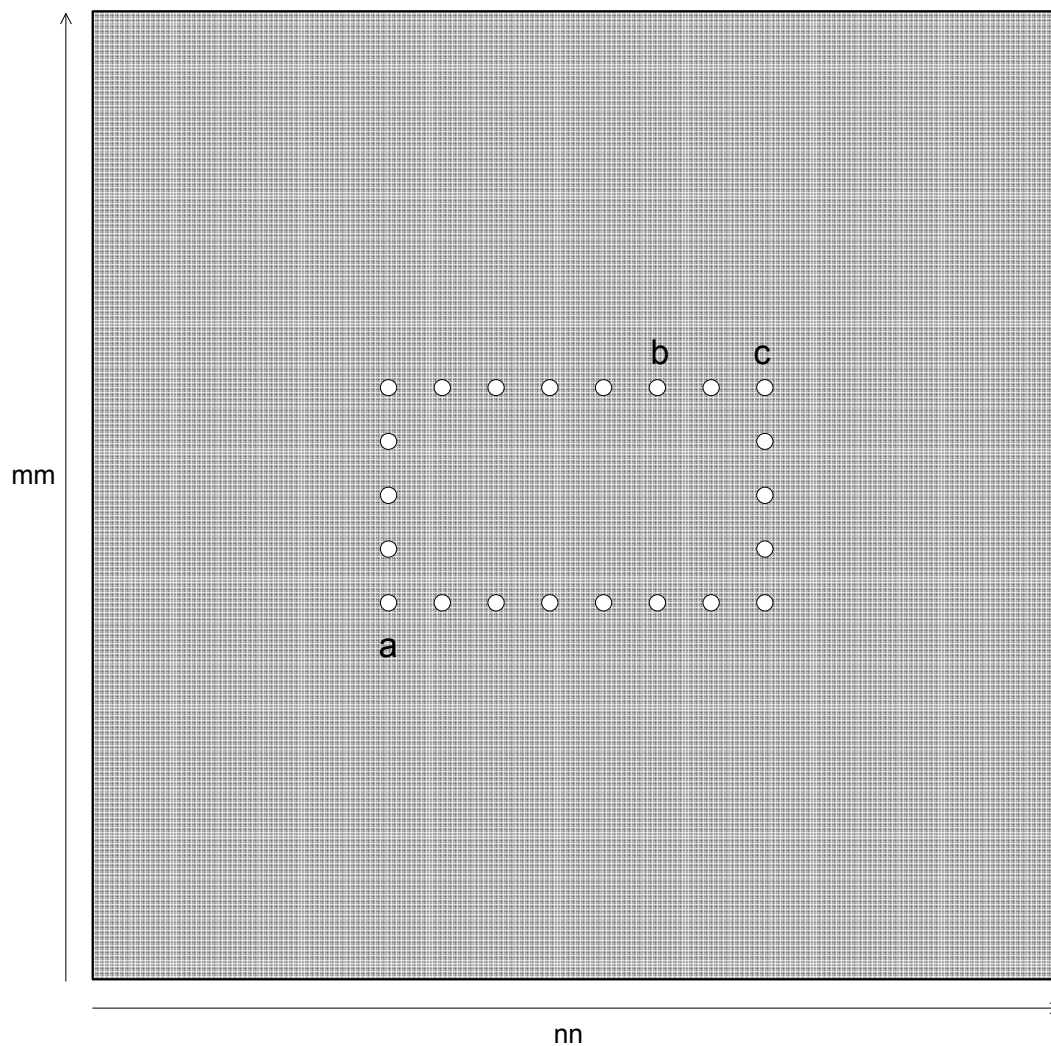


Figura 7.4 Representação do plano de orifícios formantes da figura de interesse. As difracções resultantes das passagens das partículas pelos orifícios são posteriormente detectadas na matriz de pixels da Figura 7.5. Na Figura 7.6 são apresentadas as funções de onda resultantes da passagem das partículas pelos orifícios *a*, *b* e *c*

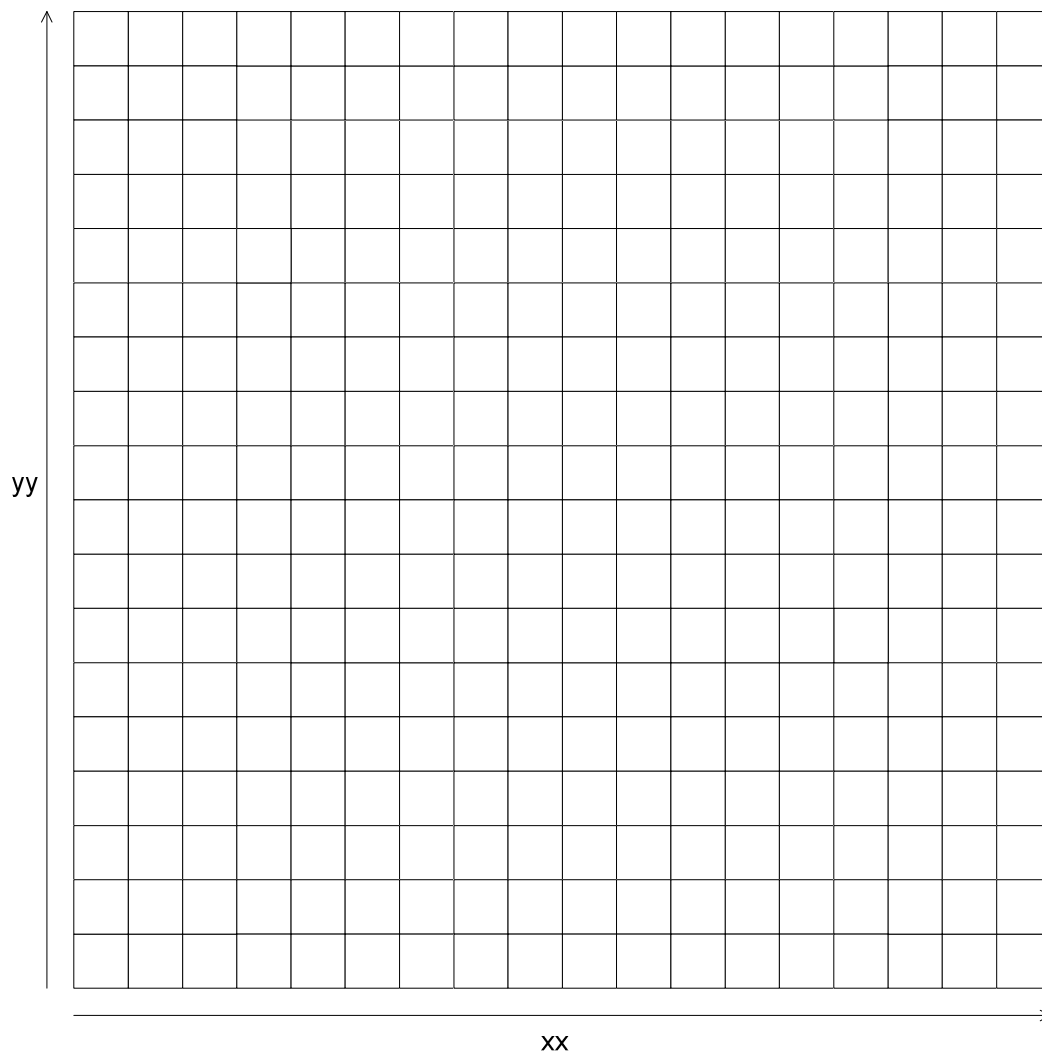


Figura 7.5 Representação do plano detector contendo elementos distribuídos uniformemente. Cada um destes elementos observará o produto da interferência entre as várias incidências de funções de onda

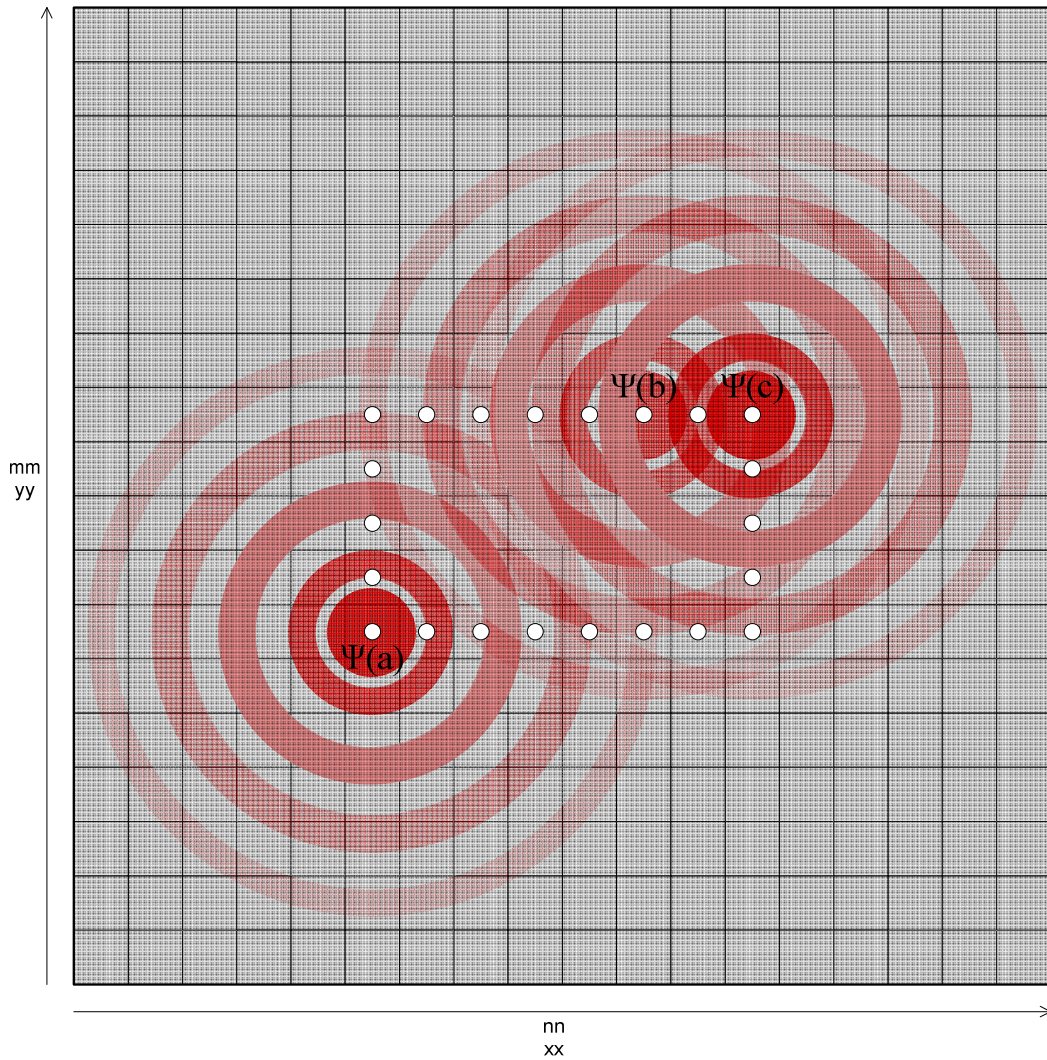


Figura 7.6 Representação, em sobreposição, do método de interferência na codificação de um rectângulo. A exemplificação da interferência é feita sobre três dos elementos do rectângulo. Cada *pixel* do plano detector está sujeito ao valor resultante da interferência entre as múltiplas funções de onda. Neste exemplo, as probabilidades das funções de onda são combinadas de acordo com a expressão: $\frac{1}{N} [\psi(a) + \psi(b) + \psi(c) + \dots + \psi(\cdot)]^2$, em que N representa o número de orifícios abertos

7.6.3 Comportamento relativo às variações de translação, de rotação e de escala

Pela análise matemática às equações de interferência, onde o método assenta, é imediato concluir que os padrões de interferência gerados são independentes da translação e da rotação. De acordo com a Equação (7.29) o valor posicional, absoluto, de um pixel não contribui para uma formulação distinta da função de onda. A dependência rotacional também não existe, já que o ângulo de difracção ϕ não participa no valor final da Equação (7.29). Assim, as novas posições de pixels, resultantes de operações de translação ou de rotação, não têm influência no padrão de interferência obtido. Apenas as posições relativas entre os pixels evidenciam essa contribuição. Outra perspectiva que justifica este comportamento advém do facto de que a posição relativa entre pixels, após uma operação de translação ou de rotação, não se altera. Isto, na prática, é verdade, assumindo a não modificação das posições relativas por mapeamento, interpolador, na grelha discretizada de pixels. Para determinadas relações entre a resolução da grelha de pixels e a escala dos conteúdos geométricos a mapear, essa assunção é na prática verdadeira. No entanto, em casos em que se pretende ter variação à rotação, é possível obter padrões de interferência dependentes da rotação. Para isso é necessário escolher aberturas quadradas em vez de aberturas circulares. Relativamente às variações de escala, tudo o que foi mencionado sobre as duas primeiras operações geométricas não é cumprido após uma operação de escala. Fundamentalmente porque a posição relativa dos pixels é alterada e, mais ainda, por interpolação e por discretização da grelha de projecção, pixels que não estavam activos podem passar a estar após uma mudança de escala. Como o padrão de interferência é dependente da posição relativa entre pixels e da existência ou não de pixels, torna-se claro que a independência à operação de escala não existe.

A função de onda, que está associada a cada pixel do plano de orifícios, está dependente do comprimento de onda da partícula e isto afecta o padrão de interferência de acordo com a escala espacial a que os restantes elementos, inclusos na imagem, aparecem relativamente às suas funções de onda. Por isto, é possível minimizar a dependência do padrão de interferência à operação de escala através da utilização de energias adaptativas (comprimentos de onda) em função da escala com que se opera. Embora isso obrigue a que, de alguma forma, a adaptação de escala seja quantificada previamente, por exemplo, através da computação do valor da área ou do valor do perímetro do conteúdo a codificar. Esta opção exige um esforço computacional

relativamente reduzido, quando comparado com outros métodos de compensação à escala. Por exemplo, a normalização à escala feita através de um mapeamento com interpolação bilinear, apesar de ser mais eficaz, é também mais exigente na perspectiva computacional. Na Equação (7.37) o comprimento de onda é adaptado através de N , o qual representa o número de pixels do conteúdo da imagem (pixels com valores diferentes de 0). S é obtido através do valor médio do número de pixels segundo um conjunto de figuras em escalas diversas e também pode ser obtido segundo um conjunto de regiões.

$$\lambda_{adaptado} = \lambda \frac{N}{S}. \quad (7.37)$$

Com este processo a dependência à escala é mais reduzida.

7.7 Resultados

Nas imagens da Figura 7.7 à Figura 7.28, pode observar-se o padrão de interferência resultante da aplicação do método a contornos sintéticos. As imagens da Figura 7.29 à Figura 7.38 mostram o resultado da aplicação do método a regiões sintéticas. O comportamento sobre imagens naturais de leucócitos é mostrado da Figura 7.39 à Figura 7.44. Da Figura 7.45 à Figura 7.50 o processo é executado sobre contornos extraídos de imagens reais de leucócitos.

Os resultados são apresentados directamente pelos padrões de interferência obtidos e indirectamente pelo histograma desses mesmos padrões.

A parametrização usada para a simulação do sistema de interferência foi escolhida utilizando uma abordagem *trial-error* com o objectivo de os padrões resultantes serem o mais coerentes possível entre figuras da mesma classe e o mais distintos possível entre figuras de classes diferentes. Foram escolhidos factores de adaptação de escala de acordo com o tipo de conteúdo, ou seja, conteúdos orientados ao contorno ou conteúdos orientados à região. A Tabela 7.I apresenta os vários parâmetros usados no simulador do sistema de interferência. As imagens resultado foram processadas para que a sua gama dinâmica fosse máxima e assim se obter uma melhor

visualização das mesmas. Estes parâmetros podem ser escalonados de forma a adaptar, num ou noutro aspecto, o sistema a um dispositivo físico.

Tabela 7.I Parametrização do sistema de interferência

| Parâmetros | Configuração |
|---|------------------|
| Abertura do orifício (a) | $4\mu m$ |
| Comprimento de onda (λ), base, da partícula | 1000 nm |
| Distância entre o plano de orifícios e o plano detector (l) | $0.001m$ |
| Distância entre <i>pixels</i> | $100\mu m$ |
| Factor de normalização (S), relativo à escala, para figuras | 114 |
| Factor de normalização (S), relativo à escala, para regiões | 1376 |



Figura 7.7 Padrão de interferência (imagem direita) resultante de uma circunferência (imagem esquerda) de níveis binários

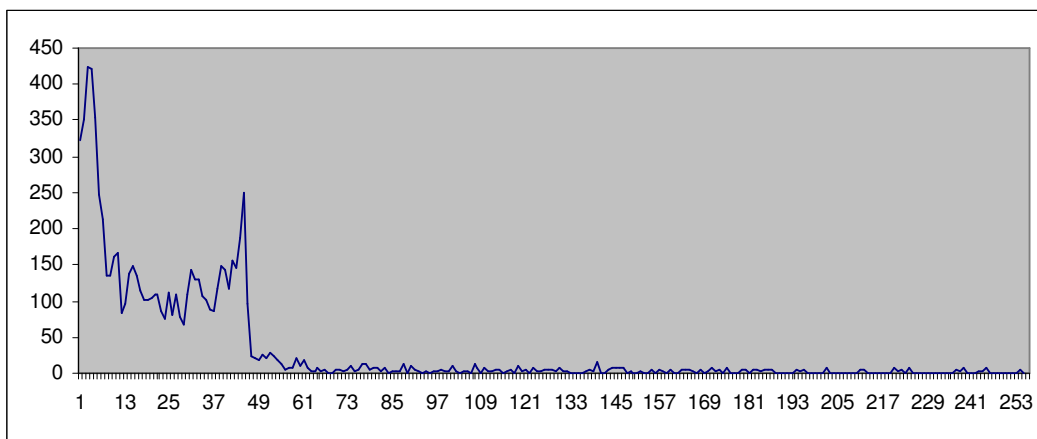


Figura 7.8 Histograma do padrão de interferência da imagem direita da Figura 7.7



Figura 7.9 Padrão de interferência resultante de um retângulo de níveis binários, a uma rotação de 135 graus e a uma escala de índice 10 de escalas entre 1 e 10

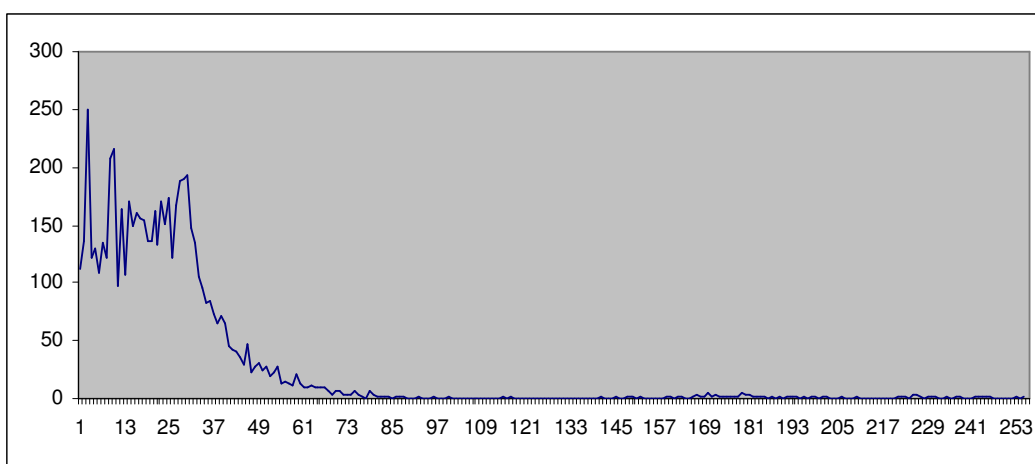


Figura 7.10 Histograma do padrão de interferência da imagem direita da Figura 7.9



Figura 7.11 Padrão de interferência resultante de um retângulo de níveis binários, a uma rotação de 150 graus e a uma escala de índice 7 de escalas entre 1 e 10

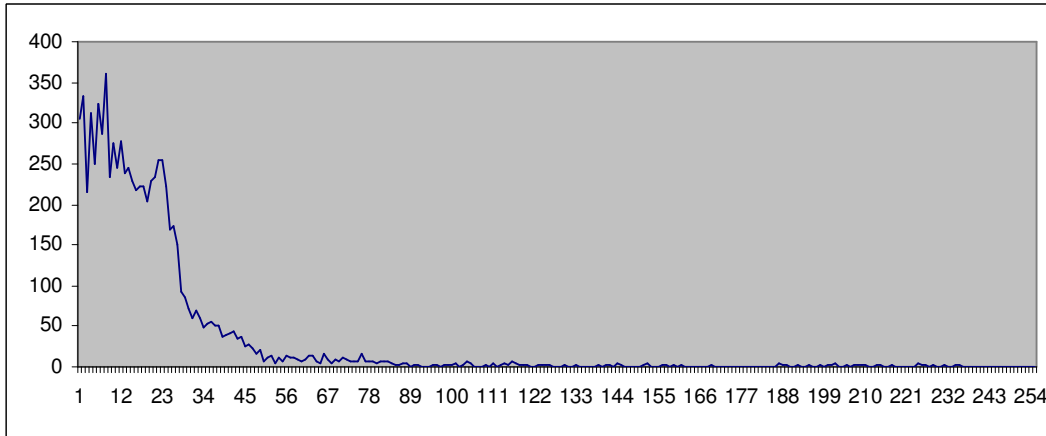


Figura 7.12 Histograma do padrão de interferência da imagem direita da Figura 7.11



Figura 7.13 Padrão de interferência resultante de um retângulo de níveis binários, a uma rotação de 70 graus, a uma escala de 7 de escalas entre 1 e 10

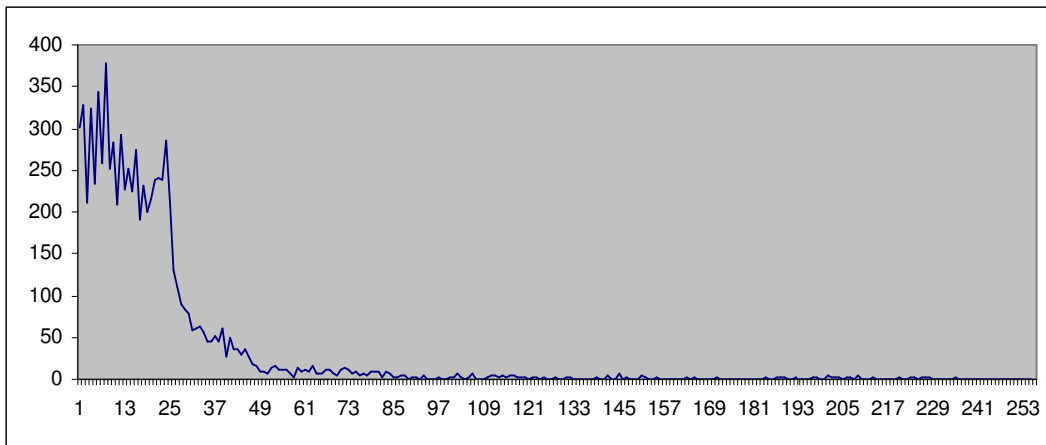


Figura 7.14 Histograma do padrão de interferência da imagem direita da Figura 7.13



Figura 7.15 Padrão de interferência resultante de um quadrado de níveis binários

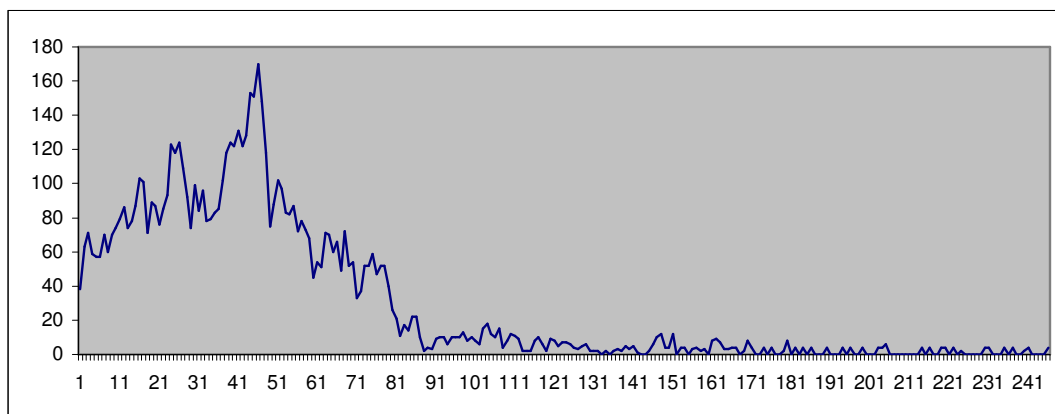


Figura 7.16 Histograma do padrão de interferência da imagem direita da Figura 7.15



Figura 7.17 Padrão de interferência resultante de um quadrado de cantos arredondados de níveis binários

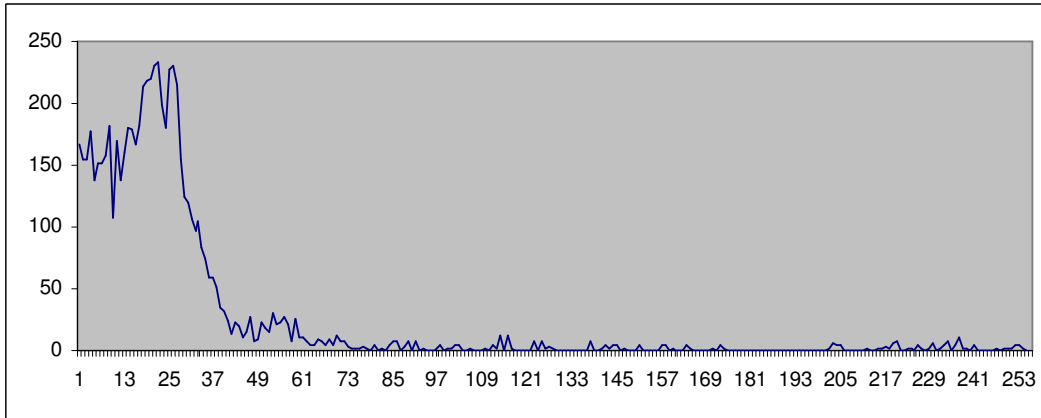


Figura 7.18 Histograma do padrão de interferência da imagem direita da Figura 7.17



Figura 7.19 Padrão de interferência resultante de uma elipse de níveis binários

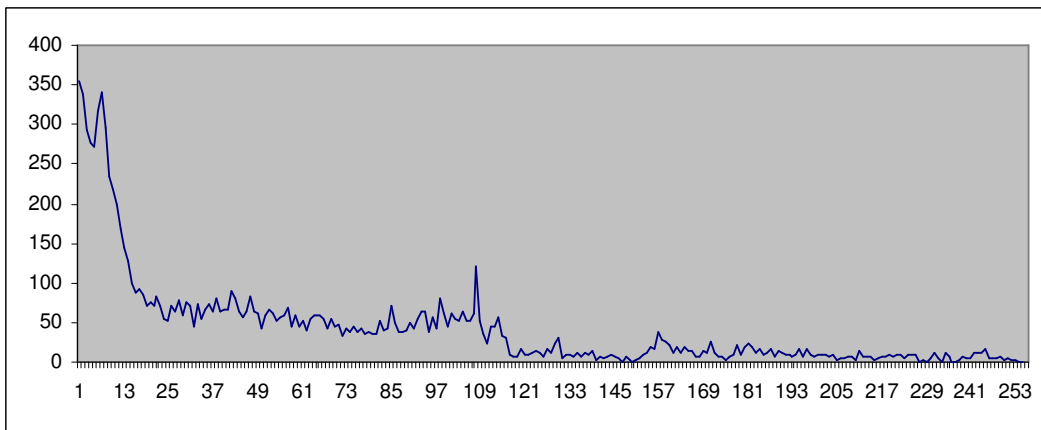


Figura 7.20 Histograma do padrão de interferência da imagem direita da Figura 7.19



Figura 7.21 Padrão de interferência resultante de um triângulo rectângulo de níveis binários

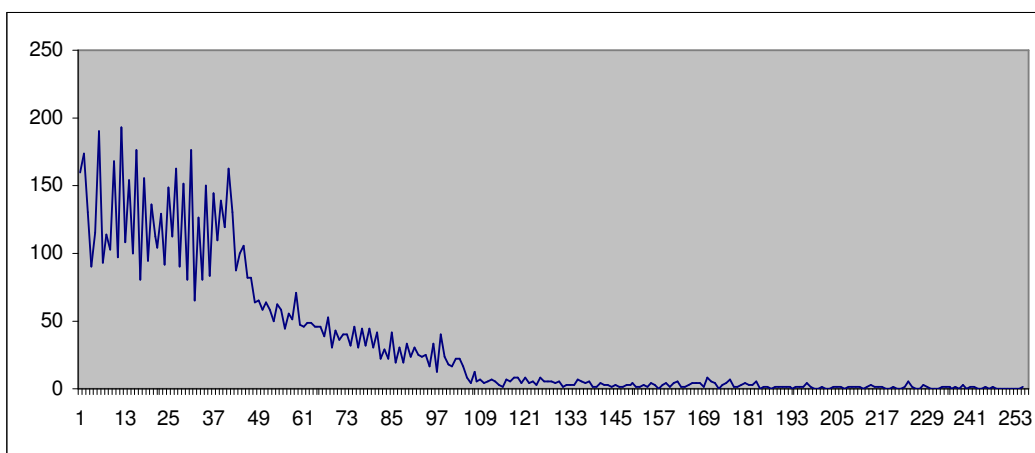


Figura 7.22 Histograma do padrão de interferência da imagem direita da Figura 7.21



Figura 7.23 Padrão de interferência resultante de um triângulo equilátero de níveis binários

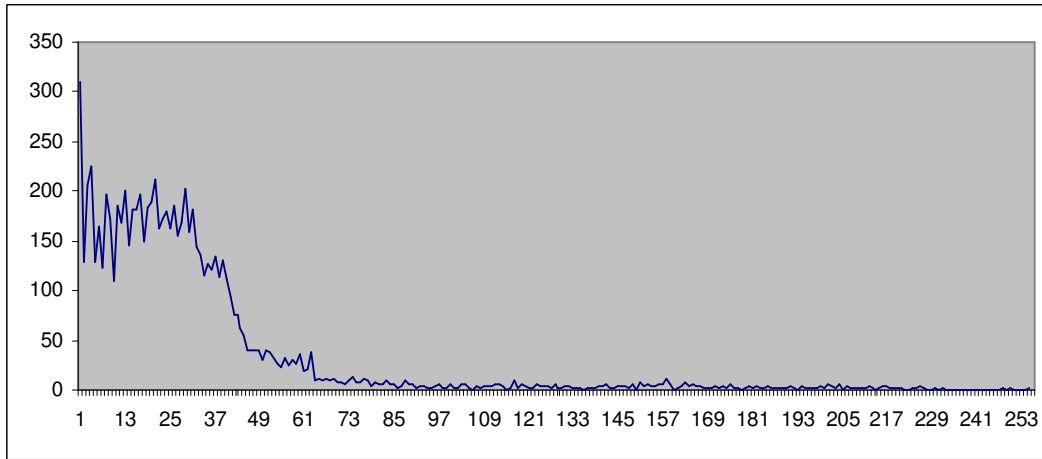


Figura 7.24 Histograma do padrão de interferência da imagem direita da Figura 7.23



Figura 7.25 Padrão de interferência resultante dos contornos da letra “T” de níveis binários, e o seu respectivo histograma

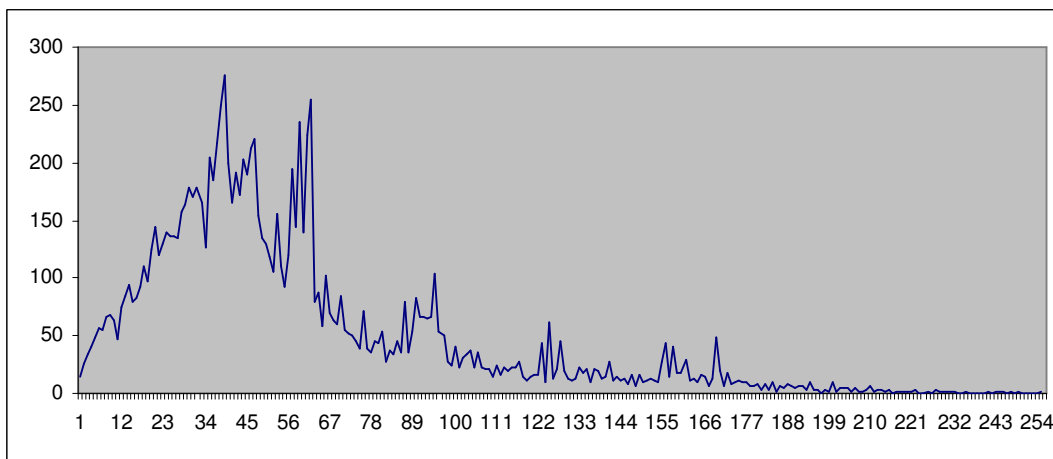


Figura 7.26 Histograma do padrão de interferência da imagem direita da Figura 7.25

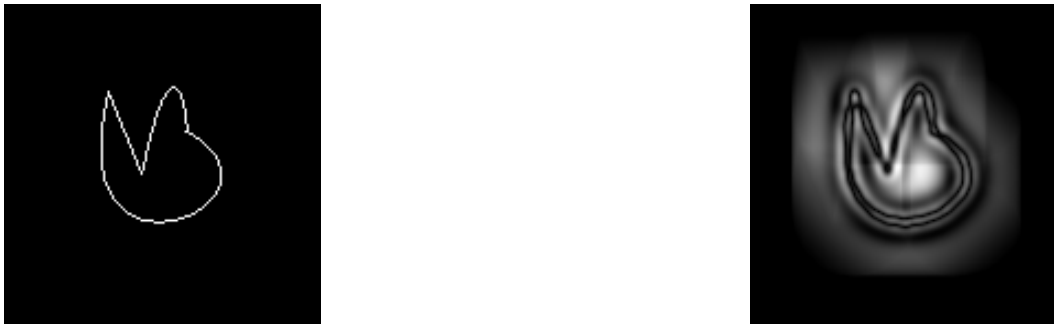


Figura 7.27 Padrão de interferência resultante dos contornos de uma figura irregular de níveis binários

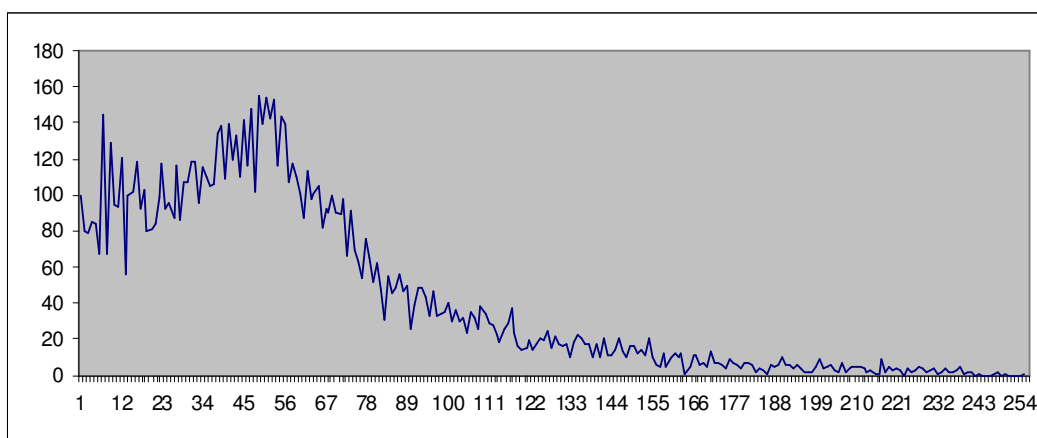


Figura 7.28 Histograma do padrão de interferência da imagem direita da Figura 7.27



Figura 7.29 Padrão de interferência resultante de uma região circular de níveis binários

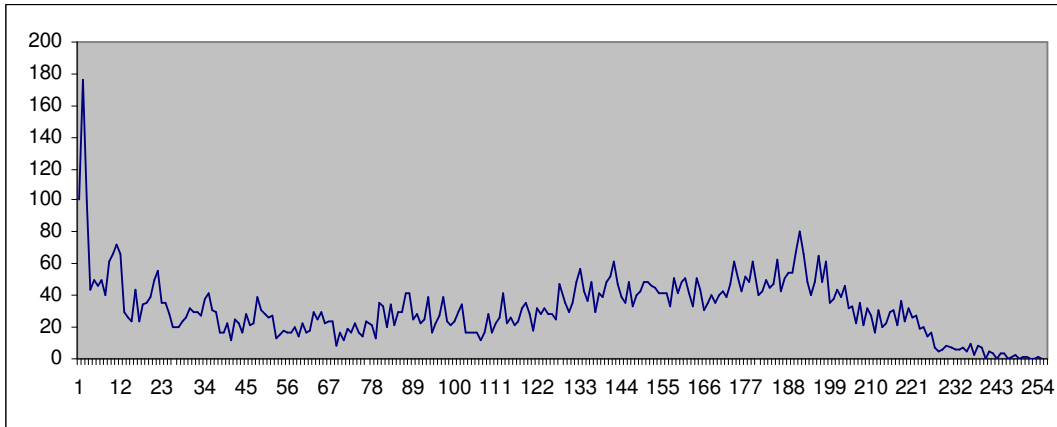


Figura 7.30 Histograma do padrão de interferência da imagem direita da Figura 7.29



Figura 7.31 Padrão de interferência resultante de uma região rectangular de níveis binários

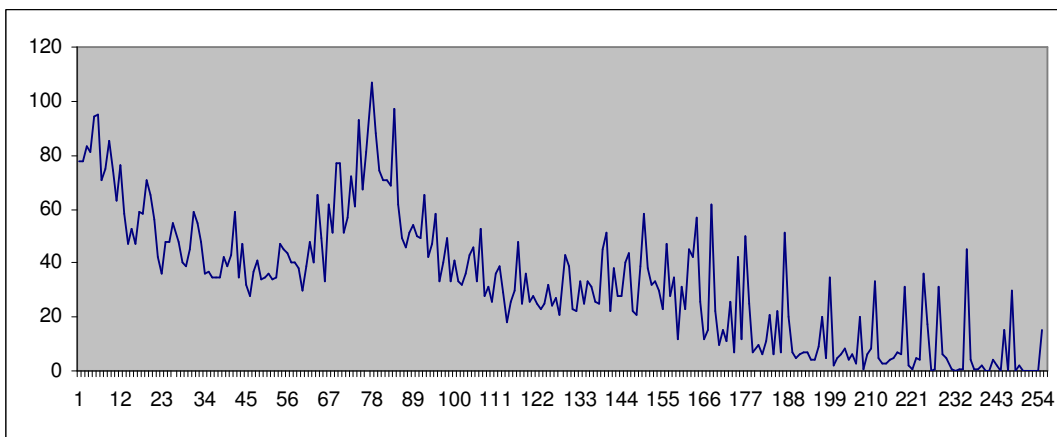


Figura 7.32 Histograma do padrão de interferência da imagem direita da Figura 7.31



Figura 7.33 Padrão de interferência resultante de uma região formada por um triângulo rectângulo

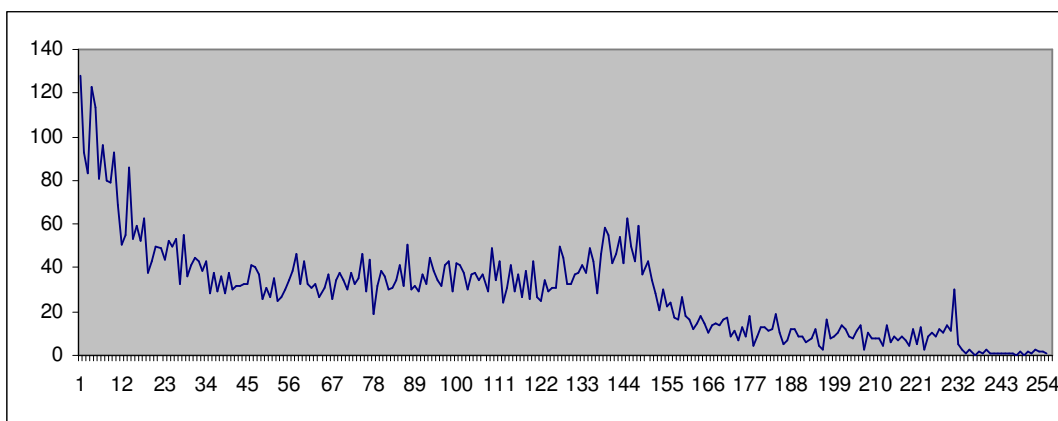


Figura 7.34 Histograma do padrão de interferência da imagem direita da Figura 7.33



Figura 7.35 Padrão de interferência resultante de uma região formada por um triângulo equilátero de níveis binários

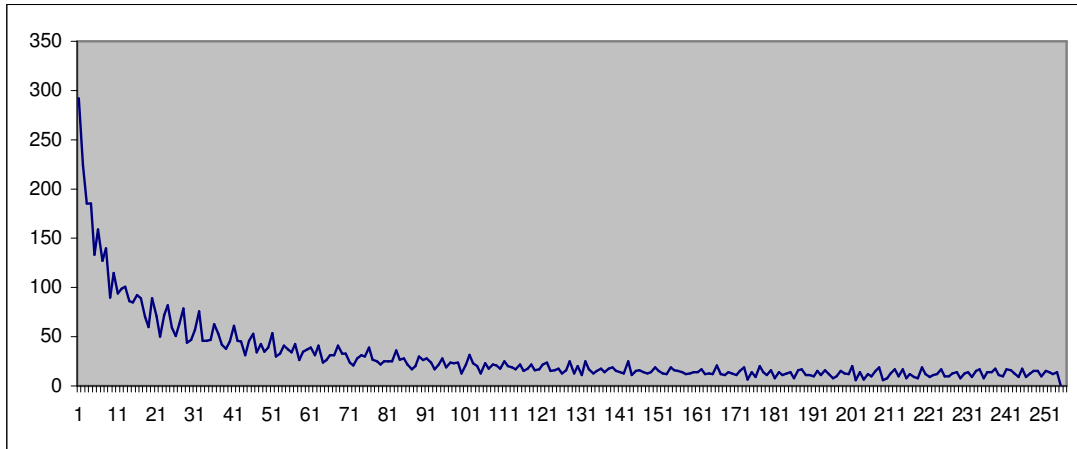


Figura 7.36 Histograma do padrão de interferência da imagem direita da Figura 7.35



Figura 7.37 Padrão de interferência resultante de uma região que configura a letra "T" de níveis binários

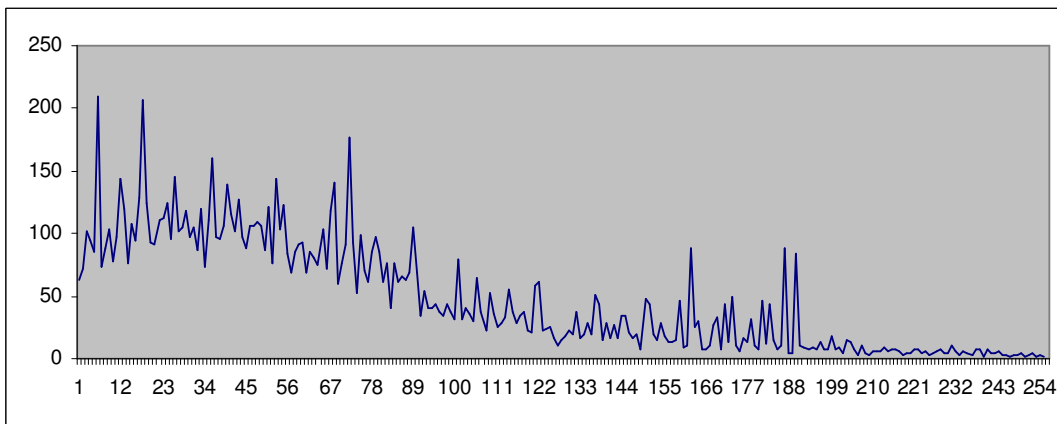


Figura 7.38 Histograma do padrão de interferência da imagem direita da Figura 7.37



Figura 7.39 Padrão de interferência resultante de uma região de um eosinófilo de níveis de cinzento

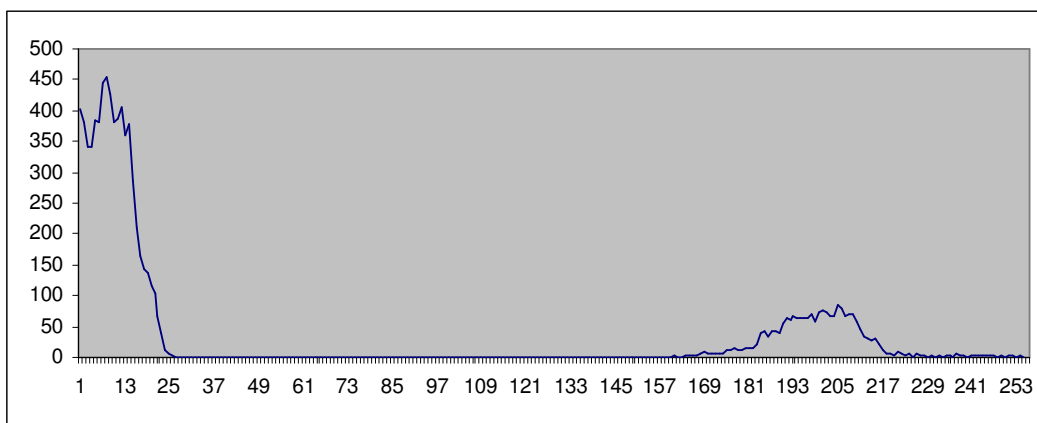


Figura 7.40 Histograma do padrão de interferência da imagem direita da Figura 7.39



Figura 7.41 Padrão de interferência resultante de uma região de um eosinófilo de níveis de cinzento distintos dos níveis do eosinófilo da Figura 7.39

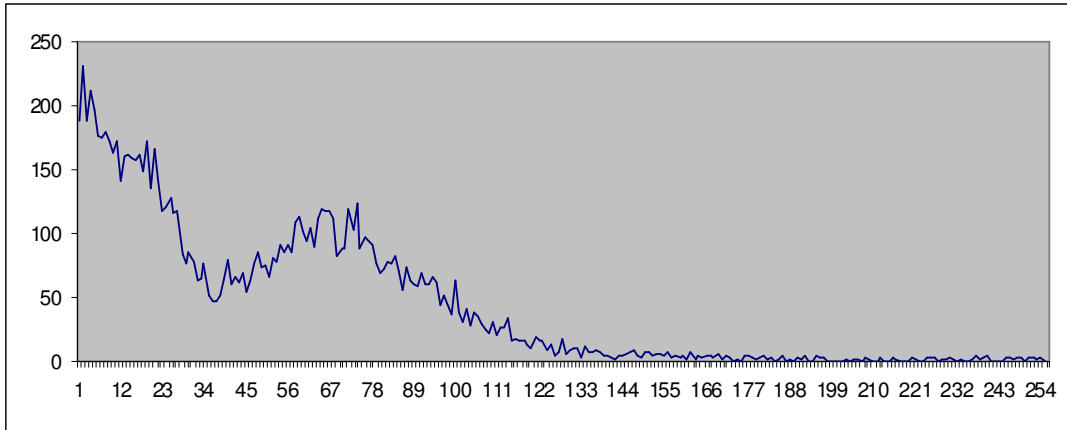


Figura 7.42 Histograma do padrão de interferência da imagem direita da Figura 7.41



Figura 7.43 Padrão de interferência resultante de uma região de um neutrófilo de níveis de cinzento

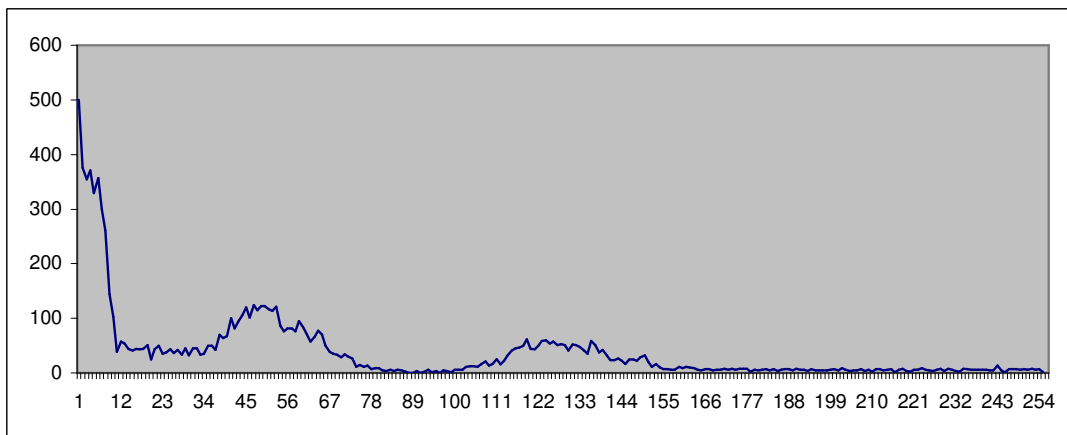


Figura 7.44 Histograma do padrão de interferência da imagem direita da Figura 7.43

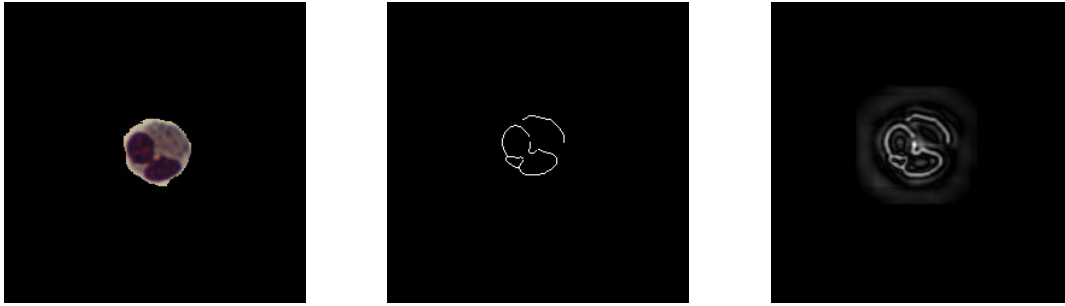


Figura 7.45 Padrão de interferência resultante de um contorno de um eosinófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central

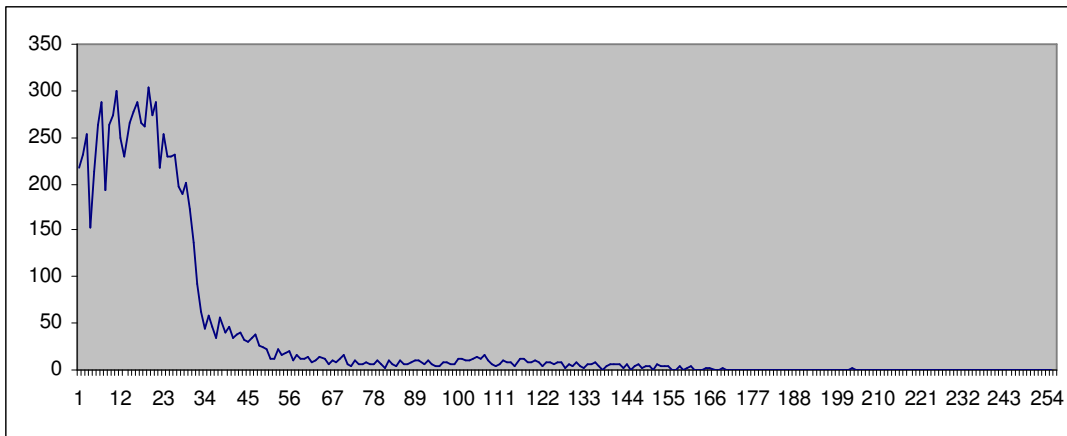


Figura 7.46 Histograma do padrão de interferência da imagem direita da Figura 7.45

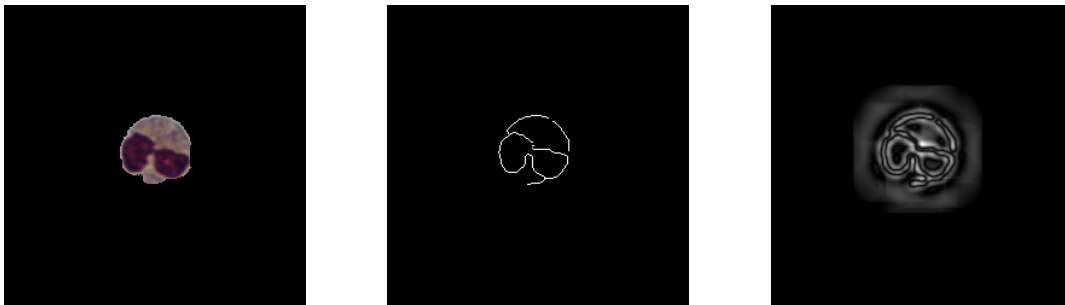


Figura 7.47 Padrão de interferência resultante de um contorno de um eosinófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central

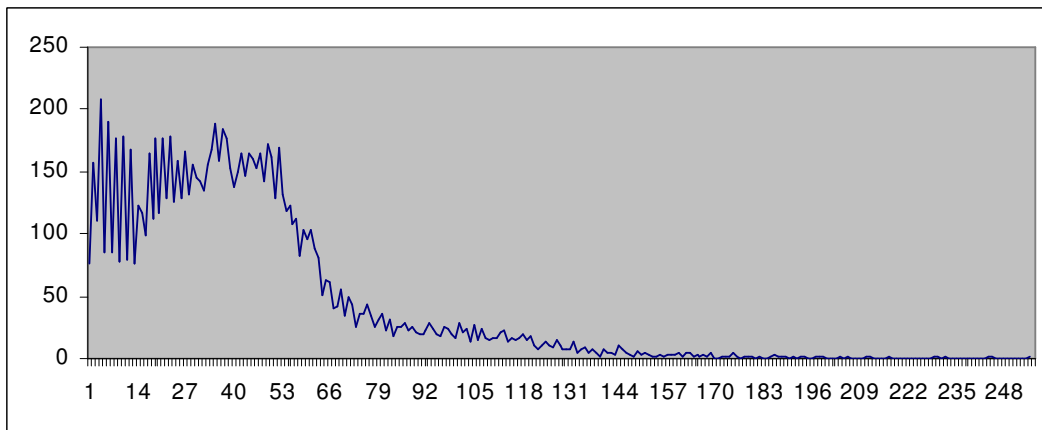


Figura 7.48 Histograma do padrão de interferência da imagem direita da Figura 7.47

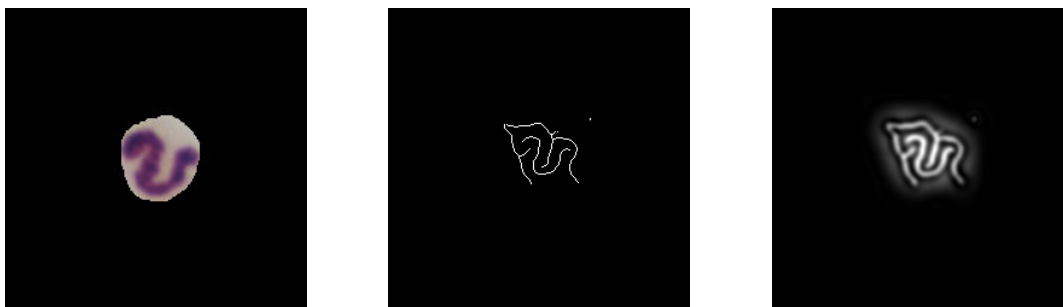


Figura 7.49 Padrão de interferência resultante de um contorno de um neutrófilo. A imagem original é a da esquerda, a seguinte resulta da aplicação de operadores de Sobel à primeira, obtendo-se os respectivos contornos. A imagem da direita apresenta o padrão de interferência da imagem central

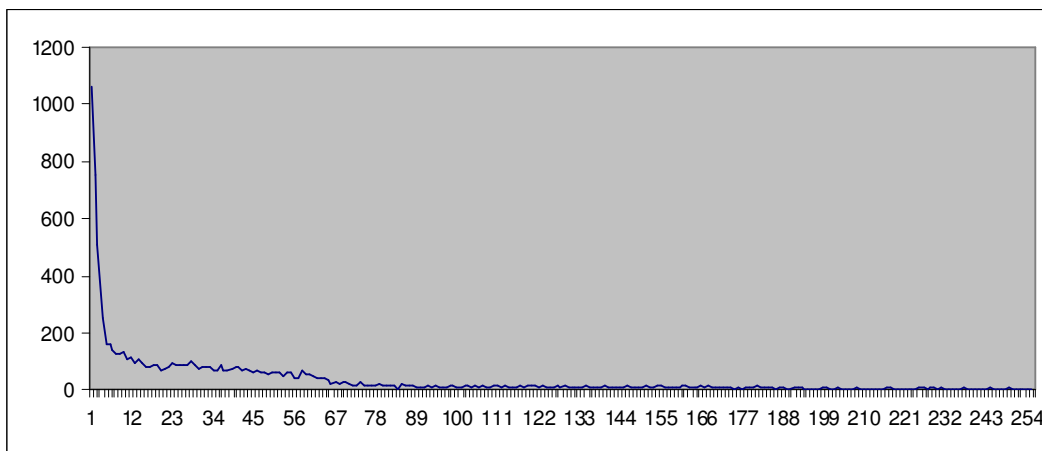


Figura 7.50 Histograma do padrão de interferência da imagem direita da Figura 7.49

O desempenho foi analisado comparando os resultados cedidos por estes descritores com os resultados dados pelos descritores de Zernike. Os valores dos descritores foram introduzidos nas entradas de um classificador SVM. No caso dos descritores de interferência, a descrição numérica final é executada pela obtenção do histograma do padrão de interferência. A utilização do histograma não carece de um esforço computacional exponencial na sua obtenção e a invariância à posição, rotação e escala não sofre prejuízo pela utilização do histograma. Os vários blocos constituintes deste sistema são mostrados na Figura 7.51.

Durante a etapa de treino, a SVM aprende a mapear os dados do histograma dos descritores, para as representações numéricas estabelecidas nas suas saídas e que codificam cada classe de figura apresentada ao descritor. O número de entradas da SVM é assim igual ao número de níveis distintos de cinzento do histograma, que neste caso é 256. O número de saídas, SVM, é igual ao número de classes a classificar numa etapa de treino. Foram analisados quatro grupos de imagens. Os dois primeiros grupos são compostos por imagens sintéticas de figuras geométricas. Um desses grupos ostenta as figuras a partir das suas regiões e o outro grupo utiliza as mesmas figuras, mas num formato orientado ao contorno. Os restantes dois grupos são constituídos por imagens naturais de leucócitos. Um destes grupos estabelece os leucócitos através da sua região codificada em níveis de cinzento, enquanto que o outro o faz a partir dos contornos que foram extraídos pela utilização de detectores de contornos de Sobel.

Com esta configuração pretende observar-se o comportamento do descritor de interferência na perspectiva da representação de região e de contorno. Pretende também denotar-se o comportamento sobre imagens naturais, mais complexas, e imagens sintetizadas, menos complexas. Estas observações são comparadas com os resultados obtidos substituindo o descritor de interferência por um descritor de Zernike e retirando o bloco do histograma. Dentro de cada grupo foram criadas imagens com diferentes posições, diferentes rotações e diferentes escalas. Isto permite perceber o comportamento do descritor de interferência relativamente a estes atributos em cada um dos grupos. Com o mesmo objectivo foram, nesses grupos, adicionadas imagens que sofreram ligeiras modificações à sua geometria original. O comportamento avaliado daí conduz à percepção do quanto estes descritores podem ser utilizados na descrição de imagens que contêm ruído e cujo objecto em representação, apesar de não coincidir em absoluto com o objecto de treino, pertence a uma mesma classe, que lhe corresponde, do objecto de treino.

Da Tabela 7.II à Tabela 7.IV são exibidos os resultados de acordo com as configurações supra mencionadas.

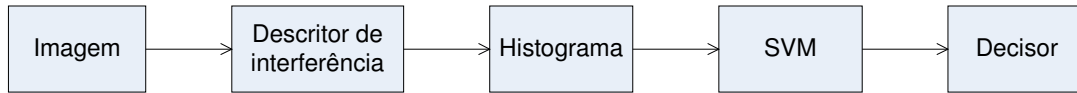


Figura 7.51 Blocos sequenciais do processo de análise ao descritor de interferência

Tabela 7.II Desempenho do descritor de interferência, relativamente a contornos sintéticos, face ao descritor de Zernike e em cooperação com uma SVM

| | Zernike+SVM | Interferência Quântica+SVM |
|--|-------------|----------------------------|
| Figuras de contornos com translações e rotações diferentes das figuras utilizadas na fase de treino | 96% | 100% |
| Figuras de contornos com escalas diferentes das figuras utilizadas na fase de treino | 85% | 76% |
| Figuras de contornos com morfologias ligeiramente alteradas relativamente às morfologias originais de treino e que pertencem às correspondentes classes de figuras | 79% | 72% |

Tabela 7.III Desempenho do descritor de interferência, relativamente a regiões sintéticas, face ao descritor de Zernike e em cooperação com uma SVM

| | Zernike+SVM | Interferência Quântica+SVM |
|---|-------------|----------------------------|
| Regiões com translações e rotações diferentes das figuras utilizadas na fase de treino | 98% | 94% |
| Regiões com escalas diferentes das utilizadas na fase de treino | 83% | 71% |
| Regiões com morfologias ligeiramente alteradas relativamente às morfologias originais de treino e que pertencem às correspondentes classes de regiões | 75% | 65% |

Tabela 7.IV Desempenho do descritor de interferência, relativamente a regiões de leucócitos, face ao descritor de Zernike e em cooperação com uma SVM

| | Zernike+SVM | Interferência Quântica+SVM |
|--|-------------|----------------------------|
| Regiões de leucócitos com translações e rotações diferentes das utilizadas na fase de treino | 87% | 88% |
| Regiões de leucócitos com escalas diferentes das utilizadas na fase de treino | 85% | 64% |
| Regiões de leucócitos com morfologias diferentes relativamente às morfologias originais de treino e que pertencem às correspondentes classes de leucócitos | 80% | 61% |

Tabela 7.V Desempenho do descritor de interferência, relativamente a contornos de leucócitos, face ao descritor de Zernike e em cooperação com uma SVM

| | Zernike+SVM | Interferência Quântica+SVM |
|--|-------------|----------------------------|
| Contornos de leucócitos com translações e rotações diferentes das utilizadas na fase de treino | 84% | 91% |
| Contornos de leucócitos com escalas diferentes das utilizadas na fase de treino | 81% | 70% |
| Contornos de leucócitos com morfologias diferentes relativamente às morfologias originais de treino e que pertencem às correspondentes classes de leucócitos | 75% | 68% |

7.7.1 Aplicação da interferência ao triângulo de Kanizsa

O triângulo de Kanizsa foi apresentado como um teste da capacidade que os mecanismos mentais têm de fazer emergir na consciência visual representações de componentes visuais sem que estes se encontrem totalmente definidas na imagem observada. O triângulo de Kanizsa é definido a partir de regiões que contribuem de forma parcial para a definição dos contornos de um triângulo que acaba por emergir na mente de quem o observa. Ao longo dos anos foram propostos vários métodos explicativos que os circuitos cerebrais implementarão para induzirem à consciência do triângulo. Alguns autores propõem mecanismos de interferência associados, não directamente aos circuitos neuronais, mas à interferência e ao enredo quântico entre moléculas constituintes dos corpos dos neurónios.

Quis-se, de uma forma meramente experimental e superficial, verificar o resultado da aplicação do método de interferência ao triângulo de Kanizsa. O padrão resultante aponta para um prolongar efectivo dos contornos que eram omissos na imagem original. Associado a outras técnicas poderia conduzir à extracção efectiva do triângulo de Kanizsa. No entanto, técnicas baseadas em operadores de dilatação poderiam produzir idênticos resultados.

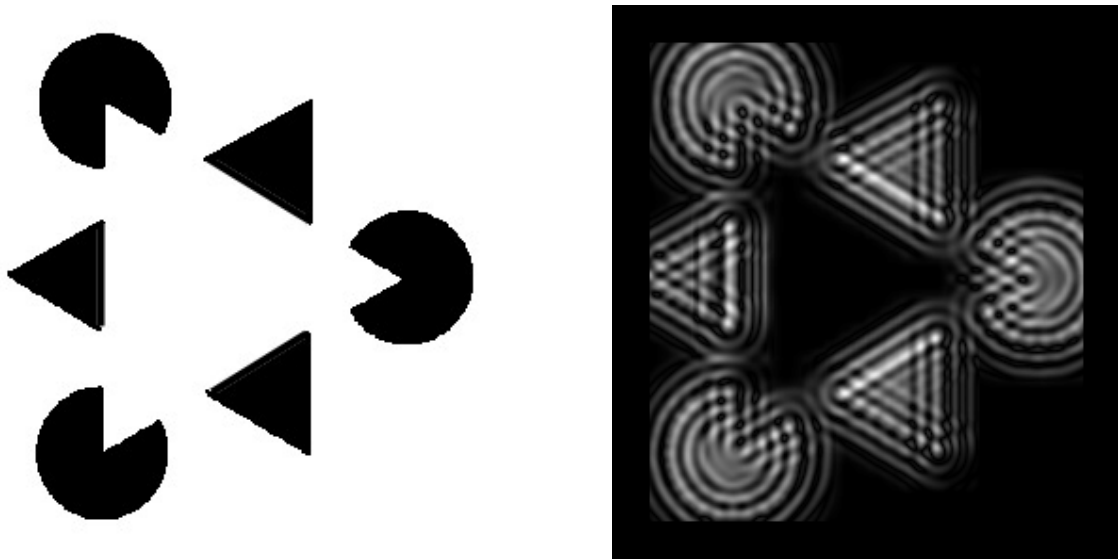


Figura 7.52 Padrão de interferência do Triângulo de Kanizsa

7.8 Discussões finais

Uma vez que o descritor por interferência apresenta, intrinsecamente, invariabilidade à posição e à rotação, os melhores resultados são dados nessas vertentes, quer se trate de imagens geométricas de regiões ou de imagens geométricas de contornos. O mesmo pode afirmar-se para os casos de imagens naturais, sejam elas provenientes de regiões ou de contornos. Os resultados nestas situações são sempre superiores aos denotados aquando da utilização do descritor de Zernike. Este facto encontra explicação na maior dimensão dos vectores de resultado relativamente ao descritor de interferência. Os momentos de Zernike possuem uma dimensão consideravelmente menor impossibilitando melhores desempenhos, pois assim parte da informação é aglutinada, o que pode conduzir à perda de informação diferenciadora.

Os resultados conseguidos com imagens de escalas diferentes, relativamente às imagens usadas na fase de treino, são inferiores, na globalidade, aos conseguidos pelo método de Zernike. A sensibilidade que o método de interferência apresenta à alteração da escala das imagens foi atenuada utilizando um mecanismo simples, fazendo depender

o valor do momento da partícula na proporção inversa ao número de pixels da figura a processar. Este método não tem em atenção a direcção em que a variação de escala acontece. No caso dos descritores de Zernike o processo é mais optimizado, complexo e computacionalmente exigente, pois tem como base a normalização prévia da imagem por métodos interpoladores. Isto permite a obtenção de melhores resultados com os descritores de Zernike, relativamente à escala, do que com o método de interferência. À custa do aumento de esforço computacional seria aplicável o mesmo processo de normalização à escala ao descritor de interferência. Pelos superiores resultados obtidos pelo método de interferência na invariabilidade à posição e à rotação será de prever, pelo menos, um desempenho aproximado dos conseguidos pelos descritores de Zernike, no caso de ser utilizado o mesmo método de normalização à escala.

O desempenho dos descritores de interferência é também inferior na situação de classificação de imagens com alterações nos seus objectos relativamente aos objectos contidos nas imagens durante a fase de treino. Neste caso o resultado pode também estar relacionado com o deficiente processo de normalização à escala, mas parece ser também indicador de que não existe uma relação coerente entre o tipo de objecto em representação e o padrão de interferência representador. Esta componente exigiria um estudo mais aprofundado da relação entre esses dois domínios para se poder sugerir uma forma solucionadora dessa dificuldade.

Esta inferioridade acentua-se mais nos casos de imagens de regiões modificadas, sejam elas sintéticas ou naturais, embora com uma tendência ligeiramente mais negativa nestas últimas.

Enquanto que a utilização de imagens geométricas permite um certo grau de sincronismo entre as funções de onda, a operação em imagens naturais tende a anular a coerência de relação de fase entre as funções de onda, o que conduz a um menor desempenho, já que o padrão de interferência, assim formado, tende a ser indiferenciador. Uma parte do problema reside nos vários níveis de energia resultantes, que são função dos diferentes níveis de pixels. Isto gera funções de onda com comprimentos de onda muito variáveis, o que contribui para a destruição dos padrões geométricos de onda. A outra parte está associada à própria distribuição dos pixels na imagem, que é mais densa, o que também contribui para a entropia do padrão de interferência final. O método pode, mesmo neste caso, ser usado com maior sucesso se a aplicação deste for feita de forma mais dirigida, codificando numericamente padrões de

interferência dirigidos isoladamente a certos aspectos locais, em vez de se usar um único padrão global.

Como no caso das imagens naturais existe uma relação entre os diversos níveis de *pixels* e a energia da função de onda, o padrão resultante depende não só da distribuição de *pixels*, mas também da tonalidade global da imagem. Isto torna o método algo dependente dessa tonalidade.

É interessante realçar que o desempenho é sensivelmente superior para a classificação de leucócitos, se forem usados, como imagem inicial, os contornos das células, em vez de se usarem as regiões células no formato original. A leitura que se pode fazer deste facto é que a informação que é perdida na extracção de contornos não é maior do que aquela que é perdida na construção dos padrões de interferência que partem das imagens de regiões. Esta observação subscreve o débil desempenho que o descritor de interferência apresenta na descrição de regiões, principalmente sobre as regiões de imagens naturais.

De qualquer forma, em situações de não existência de qualquer descritor é mais difícil garantir um comportamento de invariância à posição, à rotação e à escala. De facto, os resultados obtidos com a aplicação dos contornos directamente num classificador SVM, sem a camada intermédia de um qualquer descritor, foram inferiores aos obtidos com a aplicação de descritores, à SVM, dos contornos das células. Só um aumento dos padrões de treino a contemplarem uma multiplicidade de posições, rotações e de escalas é que pode atenuar essa diferença.

O processo de interferência pode ser usado em imagens em cores actuando o processo de interferência de forma individual em cada um dos planos de cor. A agregação final dos 3 resultados pode ser executada pelo sistema classificador, por exemplo, pelo classificador SVM.

Por último, referir que a simulação do processo de interferência exige, para imagens com esta dimensão, dez segundos de processamento. Este tempo sobe intensamente com a dimensão da imagem. A utilização do método num suporte de interferência físico desceria a ordem temporal de forma significativa e a ordem seria pouco dependente do tamanho da imagem.

Capítulo 8

Conclusões

O estudo desenvolvido ao longo dos vários capítulos assentou sempre na procura da aproximação de métodos de segmentação e de classificação ao desempenho da visão humana. Essa procura começou por ser baseada na associação de métodos não supervisionados a métodos supervisionados e terminou na adaptação de técnicas de segmentação e de classificação a métodos quânticos.

A associação de métodos não supervisionados a métodos supervisionados é necessária para que os problemas inerentes a cada um dos métodos sejam compensados pelo outro. Todavia, essa associação só por si não é suficiente para resolver o problema na totalidade. Muitos outros problemas continuam a ser detidos por cada um dos métodos sem que sejam atenuados pela cooperação do método adjacente.

Os métodos não supervisionados, pela independência preambular a alguns aspectos, deveriam à partida apresentar um melhor comportamento aos conteúdos de imagens menos particulares. A tentativa de minimizar esses problemas passou por se obterem várias soluções distintas fazendo variar a parametrização do método não supervisionado em causa. O método de segmentação multi-escala permite seguir esta abordagem fundamentalmente pela obtenção de múltiplas soluções (segmentações) que derivam da escala considerada. Numa linha semelhante, o conjunto de segmentações sequencialmente obtidas pelas redes PCNN minimizam a dependência dos processos posteriores ao conteúdo das imagens. Em qualquer dos casos, essa pluralidade de soluções pode não ser caracterizante da imagem em causa e pode mesmo não ser coerente entre imagens com conteúdos semelhantes mas com uma ou outra característica sensivelmente diferente. Ou seja, pode não garantir a generalização

adequada. Os pressupostos que por construção fazem parte desses métodos baseiam-se em métricas que, pela sua natureza, não são as adequadas para todos os casos. Modificá-las ou adicioná-las pode aumentar ligeiramente o grau de bom funcionamento do processo, mas surge sempre um grupo de conteúdos de imagens que elimina a possibilidade da universalidade do método. O aumento de desempenho funcional que estas técnicas geram representa uma pequena fracção daquilo que estas mesmas técnicas exigem computacionalmente.

A adaptação do algoritmo quântico de Grover permite que se pesquise numa imagem um determinado conteúdo e que uma vez encontrado se saiba imediatamente quais os pixels que o constituem ou qual a parametrização que o define. Realizam-se assim a classificação e a segmentação em simultâneo. Contudo, os testes feitos, nesse domínio, tiveram que ser confinados a uma figura geométrica, a de uma recta. Em imagens de conteúdos naturais, o problema de determinar qual dos conteúdos, nas suas variadíssimas morfologias, da imagem é que representa a classe em pesquisa situar-se-á ao mesmo nível do que se fosse usado um método clássico de execução. Nesta perspectiva, a vantagem de um método quântico está na aceleração, a qual permite que o grau de complexidade e de completitude dos métodos aumente consideravelmente, aumentando com menos intensidade o tempo de execução.

A classificação executada sobre os *pixels* extraídos directamente do resultado da segmentação fica dependente do correcto cumprimento da segmentação. Os processos que executam a classificação de forma estritamente supervisionada não permitem obter um bom desempenho sobre imagens futuras.

A classificação de contornos através das auto-ondas, produzidas pelas PCNNs, mostrou ser exequível e obtém um maior desempenho funcional quando comparada com um método clássico.

A utilização de classificadores SVM na classificação mostrou-se, pelo menos nas comparações sobre as imagens usadas, superior à utilização de modelos neuronais que têm por base de aprendizagem o algoritmo da direcção de descida do gradiente (Backpropagation). Mas o custo computacional adicionado pelas SVM é também evidenciado. De qualquer forma, do estudo realizado acerca da computação quântica,

pode dizer-se que a adaptação a esse domínio de métodos de aprendizagem neuronal supervisionada e não supervisionada seria possível, o que produziria acelerações de execução importantes.

A utilização de dispositivos quânticos computacionalmente menos dirigidos, que permitem a descrição de formas ou de regiões pela análise do padrão de interferência de sinais difractados, revelou ser possível, principalmente sobre conteúdos de imagens sintéticos. O resultado prático desta funcionalidade é também a aceleração obtida, que por consequência pode permitir a existência de estruturas de soluções mais completas. Comparativamente à quântica computacionalmente dirigida, a quântica menos dirigida é mais simples de implementar e numa perspectiva temporal poderia dizer-se que a sua implementação seria imediata, já que, como demonstrado, o mesmo processo tem validade no domínio óptico. Sobre conteúdos de imagens naturais, ficou a ideia de que seria necessário um método mais estruturado que fizesse uso da difracção a um nível de composições locais e em associação com outros métodos particionadores prévios. Posteriormente poderia ser feita a fusão dos resultados parciais utilizando também a interferência. No entanto, todos estes intentos estariam condicionados pela sua capacidade de gerarem franjas de interferência pouco entrópicas.

Na forma de uma conclusão geral registre-se que os métodos actuais, bem como os métodos estudados nesta tese, têm um perfil que se aproxima à filosófica ideia do *homunculus*. Esta filosofia, quando aplicada à visão humana, menciona a existência de uma espécie de pequeno homem que integra o cérebro e que neste recebe os sinais visuais e que então lhes dá interpretação consciente. Sendo assim, o *homunculus*, teria ele próprio de possuir no seu cérebro um outro *homunculus* e recorrentemente o mesmo aconteceria sobre a restante hierarquia de *homunculi* até ao infinito. Por este facto, a filosofia do *homunculus* é cientificamente rejeitada. É apenas usada para mostrar que mecanismos, ou explicações sobre o cérebro, que assentam, de alguma forma, numa recorrência semelhante não podem estar correctos ou completos. No caso da segmentação e da classificação o problema põe-se de forma análoga. Como garantir que um determinado segmentador numa situação ímpar “sabe” que produziu um mau ou um bom resultado? A resposta, tal como foi feito, passa por adicionar um supervisor que fará essa avaliação. A recorrência continua. Quem vai garantir que esse supervisor fez uma avaliação correcta? Parece haver necessidade, tal como no *homunculus*, de uma hierarquia sem fim.

Parece sempre faltar uma regra, ou então a nova regra parece colidir com as já existentes. Ao mesmo tempo, a aproximação ao desempenho humano, ao aumentar-se o esforço funcional, parece suceder, mas sem nunca se atingir uma solução o suficientemente abrangente para poder ser encarada como universal. Isto induz, também, o problema sob as teorias da não computacionalidade de Gödel (ver anexo B).

Crê-se que, para se evitar a recorrência do *homunculus*, os mecanismos físicos teriam que estar inerentemente dotados de uma propriedade física chamada *qualia*. Mas apesar de existirem alguns indicadores que suportam a natureza física dessa propriedade, o tema ainda se mantém num campo especulativo (ver anexo C).

8.1 Trabalho futuro

A adaptação de algoritmos clássicos à computação quântica é importante para se tentar ultrapassar os limites temporais impostos pelos meios existentes. Assim, a conversão de técnicas neuronais ou de técnicas clássicas, tanto de segmentação como de classificação, seria importante ser considerada em termos futuros.

Ultimamente as tecnologias quânticas que têm permitido mais facilmente a implementação computacional, ainda que em ambientes muito difíceis de criar, situa-se numa variante à computação quântica clássica. Essa variante é denominada de computação quântica adiabática [86-88]. Ao partir de abordagens ligeiramente diferentes das que foram explicadas no capítulo 6, esta tecnologia exige novas adaptações algorítmicas. Técnicas de processamento de imagens, como as que foram estudadas, seriam candidatas a essa associação. Esta tecnologia aproveita os mecanismos quânticos de efeito de túnel para, em problemas de optimização, contornar os mínimos locais. Máquinas de suporte vectorial com *kernels* que não respeitam as condições de Mercer seriam também candidatas a serem adaptadas segundo as regras da computação quântica adiabática.

Seria importante procurar tornar os algoritmos mais completos ao atacar a problemática da variação de conteúdos tanto na componente da segmentação como na componente de classificação e considerar métricas e escalas diferentes a aspectos como os de textura, de região e outros.

Seria ainda de interesse estudar estruturas que se avaliem mutuamente em ciclos iterativos que conduzam à convergência de forma a poder-se escapar à estrutura recorrente habitual.

Anexos

Anexo A

A Equação de Schrödinger

A.1 A equação de Schrödinger para uma dimensão

A equação de Schrödinger permite, deterministicamente, no mundo quântico, estabelecer a evolução de uma propriedade quântica de uma partícula. A partir desta equação é possível calcular, por exemplo, a probabilidade de se observar uma partícula numa arbitrária posição x .

Abordar-se-á o problema de encontrar as amplitudes quânticas posicionais de uma partícula numa única dimensão. Para isso resolve-se a equação de Schrödinger (A.1) na sua versão não dependente do tempo, já que, neste caso, a distribuição de energia não varia com o tempo, para uma partícula que está confinada a uma região unidimensional limitada por um potencial eléctrico elevado e que possui um comprimento L . \hbar é o valor da constante de Planck dividida por 2π . m é a massa da partícula. O E corresponde à energia da partícula para o comprimento L .

$$E\psi(x) = -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) \quad (\text{A.1})$$

Considere-se que nos limites da região unidimensional se tem um valor de potencial eléctrico tão elevado que elimina a possibilidade de a partícula electrão se posicionar nesses pontos e para além desses. Admita-se, também, que o potencial eléctrico V , ao longo de toda a região, é 0. Então a equação ficará reduzida à seguinte forma:

$$E\psi(x) = -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2}. \quad (\text{A.2})$$

Por conveniência faz-se a seguir uma substituição na Equação (A.2), usando a próxima expressão:

$$k^2 = \frac{2mE}{\hbar^2}. \quad (\text{A.3})$$

A equação de Schrödinger fica, em termos de k , assim escrita:

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x). \quad (\text{A.4})$$

Verifica-se, segundo a Equação (A.4), que a dupla diferenciação resulta, novamente, na mesma função multiplicada por $-k^2$. Isto sugere que a equação possa ser representada por um conjunto de funções seno e de funções co-seno. Isto resulta numa expressão geral da função que é dada por uma soma de uma função seno e de uma função co-seno:

$$\psi(x) = A \sin(kx) + B \cos(kx). \quad (\text{A.5})$$

Para o caso da região unidimensional, sabe-se que, nos seus limites, os potenciais eléctricos não permitem o aparecimento da partícula nessas posições. Então define-se como uma das condições fronteira dessa região a seguinte igualdade:

$$\psi(x=0) = A \sin(0) + B \cos(0) = 0. \quad (\text{A.6})$$

Para que a igualdade (A.6) seja possível, o valor de B tem que ser 0. A outra condição de fronteira é estabelecida para a extremidade oposta da região unidimensional. Nessa extremidade, na qual x assume o valor do comprimento da região, L , a amplitude quântica também será 0, pois também aqui o potencial elevado impede o aparecimento do electrão.

$$\psi(x=L) = A \sin(kL) = 0. \quad (\text{A.7})$$

Como se sabe que a função não pode ser sempre nula, tem que se garantir que se respeita a condição de fronteira para $x=L$ sem quebrar esse pressuposto. Então A não pode aqui ser igual a zero. Mas o constituinte k do argumento da função *seno* tem que ser tal que garanta a igualdade da condição fronteira. Sabe-se, da função seno, que se anula sempre que o seu argumento assume um múltiplo inteiro de π radianos. Seguindo este funcionalismo k terá que valer:

$$k = \frac{n\pi}{L}, \quad n \in [1, 2, 3 \dots]. \quad (\text{A.8})$$

Usando a Equação (A.3) e fazendo-se a substituição de k , pode escrever-se:

$$E = n^2 \frac{\hbar^2 \pi^2}{2mL^2}, \quad n \in [1, 2, 3 \dots]. \quad (\text{A.9})$$

É observável, na Equação (A.9), a quantização energética assegurada pelo valor inteiro de n .

A determinação do valor de A é conseguida partindo de um pressuposto de normalização. Segundo Max Born, o quadrado da amplitude quântica, dada pela função de onda $\psi(x)$, fornece a probabilidade de encontrar a partícula numa determinada posição x . O valor acumulativo de probabilidades deve ser igual à unidade. Escrevendo-se a condição de normalização das seguinte forma:

$$\int_0^L \psi(x)^2 dx = 1. \quad (\text{A.10})$$

Sendo a função de onda dada por:

$$\psi(x) = A \sin(kx). \quad (\text{A.11})$$

Para simplificar a integração pode recorrer-se às equivalências trigonométricas, em que:

$$\cos(2kx) = 1 - 2\sin^2(kx). \quad (\text{A.12})$$

Substituindo a variável k pela Equação (A.8), a integração resultará em:

$$\frac{A^2 L}{2} = 1. \quad (\text{A.13})$$

Assim, a constante A será dada por:

$$A = \sqrt{\frac{2}{L}}. \quad (\text{A.14})$$

A expressão final da equação de onda quântica de uma partícula confinada a uma região unidimensional de comprimento L toma a seguinte forma:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \quad n \in [1, 2, 3, \dots]. \quad (\text{A.15})$$

A.2 A equação de Schrödinger para duas dimensões

Considere-se agora a situação em que a partícula está confinada a uma região bidimensional, representando L_x o comprimento em x e L_y o comprimento em y . Também aqui vai considerar-se que o potencial eléctrico na região é sempre zero e que os potenciais confinantes são tendencialmente infinitos. De forma semelhante à Equação (A.2), mas agora para duas dimensões, escreve-se:

$$E\psi(x, y) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi(x)}{\partial x^2} + \frac{\partial^2 \psi(y)}{\partial y^2} \right). \quad (\text{A.16})$$

Para resolver a Equação (A.16) usa-se o método da separação de variáveis. Para isso é necessário assumir que ψ pode ser representado como um produto de duas funções independentes:

$$\psi(x, y) = X(x)Y(y). \quad (\text{A.17})$$

Fazendo a substituição da Equação (A.17) na Equação (A.16) e resolvendo as derivadas parciais, origina:

$$EX(x)Y(y) = -\frac{\hbar^2}{2m} \left(Y \frac{\partial^2 X(x)}{\partial x^2} + X \frac{\partial^2 Y(y)}{\partial y^2} \right). \quad (\text{A.18})$$

Se todos os termos forem divididos por $X(x)Y(y)$, $\frac{\partial^2 X(x)}{\partial x^2}$ escrito como $X''(x)$ e $\frac{\partial^2 Y(y)}{\partial y^2}$ escrito como $Y''(y)$, obtém-se:

$$E = -\frac{\hbar^2}{2m} \left(\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} \right). \quad (\text{A.19})$$

Sabe-se que $\frac{X''(x)}{X(x)}$ não depende de y e que uma variação de y poderia apenas provocar uma variação em $\frac{Y''(y)}{Y(y)}$. Mas, analisando a Equação (A.19), conclui-se que variando unicamente $\frac{Y''(y)}{Y(y)}$, sem variar $\frac{X''(x)}{X(x)}$, irá variar também o valor de E . Mas pelo princípio de conservação de energia sabe-se que o valor de E não pode variar. Assim, o termo $\frac{Y''(y)}{Y(y)}$ também não pode variar, independentemente de y . Pelas razões idênticas $\frac{X''(x)}{X(x)}$ é independente de x . Então, se $\frac{X''(x)}{X(x)}$ e $\frac{Y''(y)}{Y(y)}$ são constantes e $E = E_x + E_y$, é permitido escrever:

$$E_x = -\frac{\hbar^2}{2m} \frac{X''(x)}{X(x)}. \quad (\text{A.20})$$

$$E_y = -\frac{\hbar^2}{2m} \frac{Y''(y)}{Y(y)}. \quad (\text{A.21})$$

Substituindo $X''(x)$ e $Y''(y)$ pelos respectivos diferenciais, multiplicando todos os membros por $X(x)$, na Equação (A.20), e multiplicando todos os membros por $Y(y)$, na Equação (A.21), resulta em:

$$E_x X(x) = -\frac{\hbar^2}{2m} \frac{\partial^2 X(x)}{\partial x^2}, \quad (\text{A.22})$$

$$E_y Y(y) = -\frac{\hbar^2}{2m} \frac{\partial^2 Y(y)}{\partial y^2}. \quad (\text{A.23})$$

Assim, a Equação (A.22) e a Equação (A.23) têm a mesma forma da Equação (A.2) que tinha sido analisada para o caso unidimensional. Através da equação de onda apresentada em (A.15), pode adaptar-se a Equação (A.22) e a Equação (A.23) para a seguinte forma:

$$X_n(x) = \sqrt{\frac{2}{L_x}} \sin\left(\frac{n_x \pi x}{L_x}\right), \quad n_x \in [1, 2, 3, \dots], \quad (\text{A.24})$$

$$Y_n(y) = \sqrt{\frac{2}{L_y}} \sin\left(\frac{n_y \pi y}{L_y}\right), \quad n_y \in [1, 2, 3, \dots]. \quad (\text{A.25})$$

Resta agora escrever a equação final, sabendo que $\psi(x, y) = X(x)Y(y)$ e que $E = E_x + E_y$:

$$\psi_{n_x n_y}(x, y) = \sqrt{\frac{4}{L_x L_y}} \sin\left(\frac{n_x \pi x}{L_x}\right) \sin\left(\frac{n_y \pi y}{L_y}\right), \quad n_x \in [1, 2, 3, \dots], \quad n_y \in [1, 2, 3, \dots]. \quad (\text{A.26})$$

Tal como se fez para o caso unidimensional, pode obter-se o valor da energia através da seguinte equação:

$$E_{n_x n_y} = \frac{\hbar^2}{8m} \left[\left(\frac{n_x}{L_x} \right)^2 + \left(\frac{n_y}{L_y} \right)^2 \right]. \quad (\text{A.27})$$

Anexo B

Teoremas de Gödel - Não Computabilidade

Teorema 1: Um sistema baseado em axiomas, formal e consistente, não permite provar ou refutar a validade de todos as suas proposições consequentes dentro desse mesmo sistema.

Teorema 2: A consistência de um sistema não pode ser garantida dentro do próprio sistema. Ou seja, se for possível provar a consistência e a completitude de um sistema axiomático dentro do próprio sistema, então o sistema é inconsistente.

Os teoremas matemáticos de Gödel têm sido usados para argumentar sobre a impossibilidade de um dia se poderem obter máquinas verdadeiramente inteligentes.

Já que o funcionalismo computacional é estabelecido através do conceito de uma máquina de Turing, em acordo com a variante *Church-Turing thesis*, o sistema de axiomas é formal e consistente, logo, existem proposições que o próprio sistema não consegue decidir sobre a sua validade. Isto é verdade mesmo que o conjunto axiomático aumente.

Nos Humanos parece não ser observável a consistência e por consequência são capazes de descobrir verdades “inesperadas”. Por exemplo, para um humano aprender o conceito de números naturais não são necessários axiomas.

A seguinte sequência é frequentemente usada para demonstrar a impossibilidade que um sistema axiomático consistente pode deter na validação real de uma proposição:

- Supondo uma Máquina Universal da Verdade (MUV).
- Considerando a seguinte afirmação G:
- “MUV nunca dirá que G é verdadeiro”.
- Pergunta-se a MUV se G é verdadeiro ou falso.
- Se MUV responde “G é verdadeiro”, então “MUV nunca dirá que G é verdadeiro” é falso.
- Se “MUV nunca dirá que G é verdadeiro” é falso, então G é falso, pois (G= “MUV nunca dirá que G é verdadeiro”).
- Assim, se MUV diz que G é verdadeiro, G é de facto falso. Mas MUV não erra!, então “MUV nunca dirá que G é verdadeiro” é de facto verdadeiro.
- Mas, G=“MUV nunca dirá que G é verdadeiro”, então G é verdadeiro. Mas a máquina não o poderá dizer...
- Conclusão de Gödel: A máquina universal da verdade não existe!

Anexo C

Consciência Quântica

Se há assunto enigmático dentro e fora da ciência, é a consciência¹³. O que é a consciência? Como é formulada? É propriedade apenas dos humanos? É possível a sua interpretação funcional ou a sua descrição queda-se confinada ao círculo do divino? As máquinas poderão vir a ter consciência? Poderá a consciência ser descrita por modelos matemáticos que permitem a sua emulação proporcionando a construção de máquinas conscientes? Que produto traria a sua utilização no mundo das máquinas? São tudo questões que ao longo dos tempos atraíram todas as classes de pensadores. Desde os remotos e conhecidos filósofos providos de um senso grandioso, aos cientistas actuais que além do senso dispõem de tecnologia adjuvante.

O filósofo Sócrates tentou perceber a consciência propondo que esta seria o resultado da actividade informativa que se passa no cérebro. Já Demócrito entendia que a consciência seria uma propriedade base e natural das coisas.

Alguns trabalhos científicos têm vindo a suportar a conjugação das duas correntes.

Segundo esses trabalhos, a consciência apenas poderá existir havendo uma propriedade natural que a formule. Mas ao mesmo tempo, a consciência apenas atingirá níveis elevados se mecanismos informativos complexos alimentarem e se alimentarem da propriedade base de consciência.

A propriedade de consciência supostamente existente naturalmente nas coisas poderá estar abrangida por mecanismos de origem quântica. As observações lideradas

¹³ O conceito de consciência referido ao longo do texto foca a unidade mais básica de consciência, geralmente designada por *Qualia*.

por Hameroff assim o fazem crer [89-93]. Hameroff descobriu que num estado de não consciência, como no exemplo de uma anestesia geral, em que a actividade impulsional do cérebro não se reduz completamente, existem estruturas moleculares, chamadas *microtubules*, integrantes dos neurónios, que cessam a sua actividade. Essa actividade é de natureza quântica, indicando que a referida propriedade de consciência nasce no mundo quântico, presente em qualquer elemento.

Se assim é, a consciência não poderá ser obtida a partir de uma simulação informativa, resultante de modelos matemáticos.

Num computador pode simular-se a existência de uma cadeira, e daí pode observar-se a imagem como a de uma cadeira real, pode até ouvir-se sons, como os emitidos pelo movimento de uma cadeira real, mas não se pode utilizá-la como objecto de apoio. Com a emulação meramente informativa da consciência passa-se algo similar: pode nela conseguir-se obter algumas características espelhantes da consciência, mas o núcleo, a base objectiva dessa consciência, não existirá. Logo, tal como não se pode sentar-se na cadeira simulada pelo computador também não se pode tirar partido do principal objectivo da consciência.

Mas se um dia for possível efectuar uma emulação que não se isole no campo informativo, mas que consiga associar-se às propriedades físicas da consciência, então será que se torna exequível a obtenção de máquinas conscientes?

As hipóteses actualmente apresentadas sobre o suporte à consciência assentam em quatro direcções [94]:

A Todo o pensamento é computação; em particular, as sensações de pensamento consciente que são produzidas pela mera execução de cálculos apropriados.

B A consciência é uma característica da acção física do cérebro. Apesar de qualquer acção física poder ser simulada computacionalmente, essa simulação computacionalmente não pode, por algum motivo, fazer emergir a consciência.

C A acção física apropriada do cérebro produz consciência, mas esta acção é impossível de ser simulada computacionalmente.

D A consciência não pode ser simulada em termos físicos, computacionais ou por um outro qualquer método científico – Hipótese do divino.

A hipótese **C** é defendida pelo matemático e físico Roger Penrose. A sua posição tem por base os trabalhos da incompletude de Gödel e também certos aspectos da incompletude geométrica.

Segundo este autor existem duas classes de **C**:

C fraca:

A física actualmente conhecida é suficiente para tratar **C**. É só uma questão de se procurarem em detalhe os fenómenos físicos que participam na consciência e que são computacionalmente impróprios.

C forte:

A física actualmente conhecida é insuficiente para tratar **C**. Segundo Penrose a física contém partes que não se integram na compreensão actual da física e da matemática; são partes com uma natureza completamente distinta da percebida actualmente. Se houver capacidade para, no futuro, criar essa matemática e essa física, então será provavelmente possível perceber, mas não “simular” a consciência.

De facto, a não compreensão da natureza da redução da função de onda quântica parece sugerir a falta de uma matemática e de uma física novas. Nesse sentido uma teoria consistente acerca da unificação dos campos poderá conduzir a novos entendimentos do problema.

Anexo D

Controlador de Microscópio

A imagem da Figura D.1 apresenta a janela da aplicação de controlo do microscópio. É possível fazer o varrimento automático à lâmina e adquirir nesse varrimento as sequências de imagens correspondentes. A focagem do microscópio é reajustada automaticamente de aquisição em aquisição. A intensidade luminosa do microscópio é também definida no controlador.

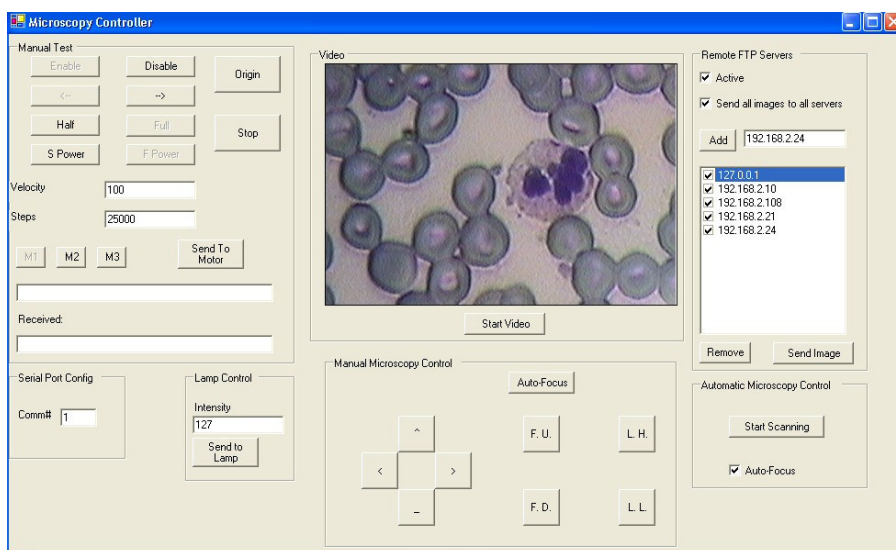


Figura D.1 Janela da aplicação de controlo do microscópio

Existe a possibilidade de indicar servidores de FTP (File Transfer Protocol) para distribuir automaticamente as imagens a processar.

A aplicação envia e recebe dados entre o controlador físico da motorização acoplada ao microscópio, Figura D.2.

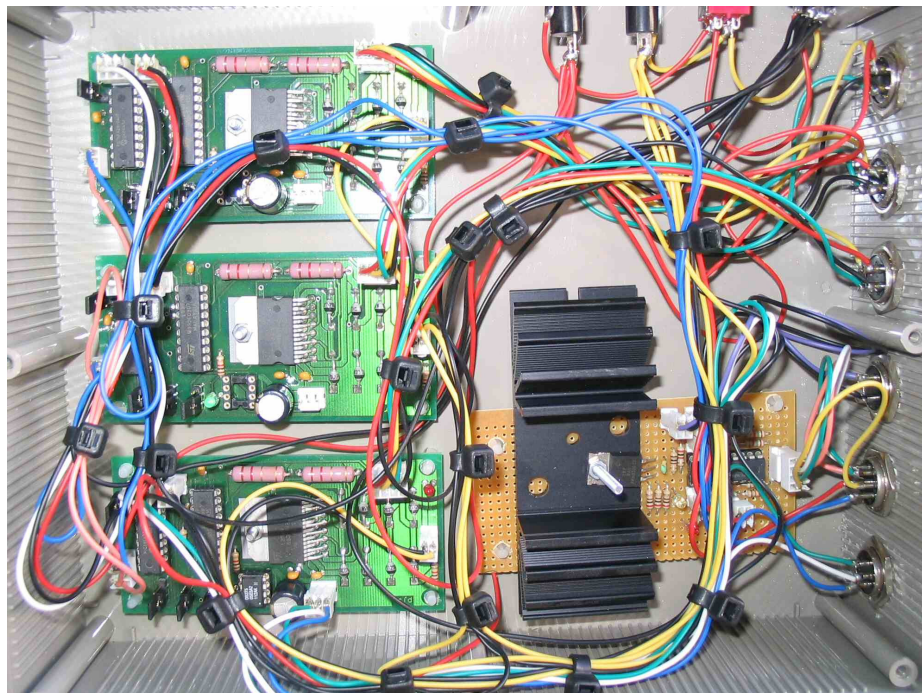


Figura D.2 Controlador desenvolvido para a motorização do microscópio

Referências Bibliográficas

- [1] G. Deco and J. Zihl, "Top-down selective visual attention: A neurodynamical approach," *Visual Cognition*, vol. 8, no. 1, pp. 119-140, Feb.2001.
- [2] E. T. Rolls and G. Deco, *Computational neuroscience of vision*. Oxford England: Oxford University Press, 2002.
- [3] P. J. Rodrigues, "Redes neuronais aplicadas à segmentação e classificação de leucócitos em imagens." Dissertação de Mestrado apresentada à Universidade de Aveiro, 2000.
- [4] H. Ramoser, V. Laurain, H. Bischof, and R. Ecker, "Leukocyte segmentation and classification in blood-smear images ," in *EMBS 2005 IEEE*, 2005, pp. 3371-3374.
- [5] T. Hirono, K. Yabusaki, and Y. Yamada, "The development of leukocyte counter using fluorescence imaging analysis," *Optical and Quantum Electronics*, vol. 37, no. 8, pp. 735-753, June2005.
- [6] S. Gunn, "Support vector machines for classification and regression," University of Southampton , Technical report, 1998.
- [7] J. D. Shutler and M. S. Nixon, "Zernike velocity moments for sequence-based description of moving features," *Image and Vision Computing*, vol. 24, no. 4, pp. 343-356, Apr.2006.
- [8] E. Rieffel and W. Polak, "An introduction to quantum computing for non-physicists," *Acm Computing Surveys*, vol. 32, no. 3, pp. 300-335, Sept.2000.
- [9] T. Poggio, V. Torre, and C. Koch, "Computational vision and regularization theory," *Nature*, vol. 317, no. 6035, pp. 314-319, 1985.
- [10] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*. New York: Yale University Press, 1923.
- [11] V. N. Vapnik, *The nature of statistical learning theory*. New York: Springer, 1995.
- [12] S. R. Gunn, M. Brown, and K. M. Bossley, "Network performance assessment for neurofuzzy data modelling" *Advances in Intelligent Data Analysis*, vol. 1280, pp. 313-323, 1997.

- [13] B. Scholkopf, K. K. Sung, C. J. C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing support vector machines with Gaussian kernels to radial basis function classifiers," *Ieee Transactions on Signal Processing*, vol. 45, no. 11, pp. 2758-2765, Nov.1997.
- [14] B. Schölkopf and A. J. Smola, *Learning with kernels support vector machines, regularization, optimization, and beyond*. Cambridge, Mass: MIT Press, 2002.
- [15] M. Minoux and M. Minoux, *Mathematical programming theory and algorithms*. Chichester W. Sussex: Wiley, 1986.
- [16] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, Sept.1995.
- [17] A. J. Smola and B. Scholkopf, "On a kernel-based method for pattern recognition, regression, approximation, and operator inversion," *Algorithmica*, vol. 22, no. 1-2, pp. 211-231, Sept.1998.
- [18] V. Blanz, H. Schollkopf, H. Bulthoff, C. Burges, V. Vapnik, and T. Vetter, "Comparison of view - based object recognition algorithms using realistic 3D models.," in *Springer Lecture Notes in Computer Science, Vol. 1112* Berlin: 1996, pp. 251-256.
- [19] V. Vapnik, S. Golowich, and A. J. Smola, "Support vector method for function approximation, regression estimation, and signal processing.," in *Advances in Neural Information Processing Systems 9* Cambridge,MA: MIT Press, 1997, pp. 281-287.
- [20] F. Girosi, "An equivalence between sparse approximation and support vector machines," *Neural Computation*, vol. 10, no. 6, pp. 1455-1480, Aug.1998.
- [21] N. Heckman, "The theory and application of penalized least squares methods or reproducing kernel hilbert spaces made easy," 1997.
- [22] G. Wahba, *Spline models for observational data*. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics, 1990.
- [23] N. Aronszajn, *Introduction to the theory of Hilbert spaces*. Stillwater, Okla: Reasearch sic Foundation, 1950.
- [24] M. Hu, "Visual-pattern recognition by moment invariants," *Ire Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, 1962.
- [25] C. H. Teh and R. T. Chin, "On image-analysis by the methods of moments," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496-513, July1988.
- [26] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489-497, May1990.

- [27] M. R. Teague, "Image analysis via the general-theory of moments," *Journal of the Optical Society of America*, vol. 69, no. 10, p. 1468, 1979.
- [28] F. Zernike, "Beugungstheorie de Schneidverfahrens und seiner verbesserten Form, der Phasenkontrastmethode," *Physica*, pp. 689-704, 1934.
- [29] J. T. Tou and R. C. González, *Pattern recognition principles*. Reading, Mass: Addison-Wesley Pub. Co, 1974.
- [30] N. Otsu, "Threshold selection method from gray-level histograms," *Ieee Transactions on Systems Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [31] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, vol. 548, pp. 2-9, 1985.
- [32] D. Marr, *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman, 1982.
- [33] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, no. 5, pp. 363-370, 1984.
- [34] A. Witkin, "Scale-space filtering," in *Proc. 8th Int. Joint Conf. Artificial Intelligence*, Karlsruhe, Germany, 1983, pp. 1019-1022.
- [35] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel, "Axioms and fundamental equations of image-processing," *Archive for Rational Mechanics and Analysis*, vol. 123, no. 3, pp. 199-257, 1993.
- [36] P. Perona and J. Malik, "Scale-space and edge-detection using anisotropic diffusion," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, July 1990.
- [37] A. Petrovic, O. D. Escoda, and P. Vandergheynst, "Multiresolution segmentation of natural images: From linear to nonlinear scale-space representations" *Ieee Transactions on Image Processing*, vol. 13, no. 8, pp. 1104-1114, Aug. 2004.
- [38] T. Lindeberg, *Scale-space theory in computer vision*, SECS 256 ed. Boston: Kluwer Academic, 1994.
- [39] K. Vincken, "Probabilistic multi-scale image segmentation by the hyperstack." Thesis, Utrecht Univ., Utrecht, The Netherlands, 1995.
- [40] K. L. Vincken, A. S. E. Koster, and M. A. Viergever, "Probabilistic multiscale image segmentation," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 109-120, Feb. 1997.
- [41] K. Vincken, W. J. Niessen, and M. A. Viergever, "Blurring strategies for image segmentation using a multiscale linking model," in *Computer Society Conf. Computer Vision Pattern Recognition*, 18-20 ed IEEE, 1996, pp. 21-26.

- [42] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex," *Neural Comp.* 2, pp. 293-307, 1990.
- [43] Y. I. Balkarey, M. G. Evtikhov, and M. I. Elinson, "Autowave media and neural networks," *SPIE* 1621, 1991, pp. 238-249.
- [44] W. Maass and C. Bishop, *Pulsed neural networks* MIT Press, 1998.
- [45] J. L. Johnson, M. L. Padgett, and O. Omidvar, "Overview of pulse coupled neural network (PCNN) special issue," *Ieee Transactions on Neural Networks*, vol. 10, no. 3, pp. 461-463, May1999.
- [46] J. L. Johnson, "The Signature of Images," *Ieee Int'l Conf. on Neural Nets*, Orlando: 1994.
- [47] J. Atmer, "Image signatures from PCNN using computers." Thesis, Royal Institute of Technology, Stockholm, 2003.
- [48] T. Lindblad and J. M. Kinser, *Image processing using pulse-coupled neural networks*. London: Springer, 1998.
- [49] Sobel and Feldman, "A 3x3 isotropic gradient operator for image processing," *Stanford Artificial Project*, 1968.
- [50] Ichikawa and Y.Ishii, "Retaining diversity of genetic algorithms for multivariable optimization and neural network learning," in *Int. Conf. Neural Networks-IEEE*, II ed San Francisco: 1993, pp. 1110-1114.
- [51] L. Lam, S. W. Lee, and C. Y. Suen, "Thinning methodologies - a comprehensive survey," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 869-885, Sept.1992.
- [52] H. Freeman, "Computer processing of line drawing images," in *Acm Computing Surveys*, 6(1) ed 1964, pp. 57-97.
- [53] R.Safae-Rad, B.Benhabib, K.C.Smith, and K.M.Ty, "Position, rotation, and scale-invariant recognition of 2-dimensional objects using a gradient coding scheme," in *Communications, Computers and Signal Processing* 1989.
- [54] R.Gonzalez and R.Woods, *Digital image processing*, Addison-Wesley Publishing Company, 1992, pp. 518-548.
- [55] P. Benioff, "Quantum-mechanical Hamiltonian models of Turing-machines," *Journal of Statistical Physics*, vol. 29, no. 3, pp. 515-546, 1982.
- [56] P. Benioff, "Quantum-mechanical models of Turing-machines that dissipate no energy," *Physical Review Letters*, vol. 48, no. 23, pp. 1581-1585, 1982.
- [57] P. Benioff, "The computer as a physical system - a microscopic quantum-mechanical Hamiltonian model of computers as represented by Turing-machines," *Journal of Statistical Physics*, vol. 22, no. 5, pp. 563-591, 1980.

- [58] P. Benioff, "Quantum-mechanical models of computers as represented by Turing-machines," *Bulletin of the American Physical Society*, vol. 25, no. 1, p. 32, 1980.
- [59] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6-7, pp. 467-488, 1982.
- [60] R. P. Feynman, "Quantum-mechanical computers," *Journal of the Optical Society of America B-Optical Physics*, vol. 1, no. 3, p. 464, 1984.
- [61] M.Pérez-Suárez and D.J.Santos, *Procesado de información con sistemas cuánticos* Servicio de Publicaciones de la Universidad de Vigo, 2006.
- [62] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge: Cambridge University Press, 2000.
- [63] P. A. M. Dirac, *The principles of quantum mechanics*, 4th ed. Oxford: Clarendon Press, 1958.
- [64] R. P. Feynman, A. J. G. Hey, and R. W. Allen, *Feynman lectures on computation*. Reading, Mass: Addison-Wesley, 1996.
- [65] T. Toffoli, "Bicontinuous extensions of invertible combinatorial functions," *Mathematical Systems Theory*, vol. 14, no. 1, pp. 13-23, 1981.
- [66] D. Deutsch, "Quantum-theory, the Church-Turing principle and the universal quantum computer," *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, vol. 400, no. 1818, pp. 97-117, 1985.
- [67] E. Bernstein and U. Vazirani, "Quantum complexity theory," *Siam Journal on Computing*, vol. 26, no. 5, pp. 1411-1473, Oct.1997.
- [68] A. Barenco, C. H. Bennett, R. Cleve, D. P. Divincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, "Elementary gates for quantum computation," *Physical Review A*, vol. 52, no. 5, pp. 3457-3467, Nov.1995.
- [69] P. Shor, "Algorithms for quantum computation: Discrete log and factoring.," in *35th Annual Symposium on Foundations of Computer Science 1994*, pp. 124-134.
- [70] K. Grover, "A fast quantum mechanical algorithm for database search," in *Twenty-Eighth Annual ACM Symposium on the Theory of Computing* Philadelphia: 1996, pp. 212-219.
- [71] M. e. al. Boyer , "Tight bounds on quantum search," in *PhysComp '96*, Los Alamitos: 1996.
- [72] J. E. Bresenham, "An incremental algorithm for digital plotting," *Communications of the Acm*, vol. 6, no. 7, p. 366, 1963.

- [73] W. K. Pratt, *Digital image processing PIKS inside*, 3rd ed. New York: Wiley, 2001.
- [74] H. Rosé, FraunHofer-quantum computing services, 2006
- [75] D. Durr and P. Hoyer, A Quantum algorithm for finding the minimum, arxiv.org/abs/quant-ph/9607014, 1999
- [76] B. Omer, "Classical concepts in quantum programming," *International Journal of Theoretical Physics*, vol. 44, no. 7, pp. 943-955, July2005.
- [77] J. S. Bell, "The paradox of Einstein, Podolsky and Rosen - Action at a distance in quantum-mechanics," *Speculations in Science and Technology*, vol. 10, no. 4, pp. 269-285, 1987.
- [78] R. L. Liboff, *Introductory quantum mechanics*, 3rd ed. Massachusetts: Addison-Wesley, 1998.
- [79] L. E. Ballentine, *Quantum mechanics a modern development*, 1st ed. Singapore: World Scientific, 2003.
- [80] E. Merzbacher, "Quantum mechanics," 2nd ed New York: Wiley, 1970.
- [81] A. O. Barut and S. Basri, "Path-integrals and quantum interference," *American Journal of Physics*, vol. 60, no. 10, pp. 896-899, Oct.1992.
- [82] T. V. Marcella, "Quantum interference with slits," *European Journal of Physics*, vol. 23, no. 6, pp. 615-621, Nov.2002.
- [83] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum mechanics*. New York etc.: John Wiley and Sons, 1977.
- [84] D. J. Griffiths, *Introduction to quantum mechanics*, Prentice Hall, Inc., 2004.
- [85] M. F. Crommie, C. P. Lutz, and D. M. Eigler, "Confinement of electrons to quantum corrals on a metal-surface," *Science*, vol. 262, no. 5131, pp. 218-220, Oct.1993.
- [86] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, arXiv:quant-ph/0001106v1, 2000
- [87] J. Song, Y. Xia, H. S. Song, J. L. Guo, and J. Nie, "Quantum computation and entangled-state generation through adiabatic evolution in two distant cavities," *Epl*, vol. 80, no. 6 2007.
- [88] S. B. Zheng, "Nongeometric conditional phase shift via adiabatic evolution of dark eigenstates: A new approach to quantum computation," *Physical Review Letters*, vol. 95, no. 8 Aug.2005.
- [89] S. Hameroff, "Consciousness, the brain, and spacetime geometry," *Cajal and Consciousness*, vol. 929, pp. 74-104, 2001.

- [90] S. Hameroff, A. Nip, M. Porter, and J. Tuszynski, "Conduction pathways in microtubules, biological quantum computation, and consciousness," *Biosystems*, vol. 64, no. 1-3, pp. 149-168, Jan.2002.
- [91] S. Hameroff, "Anesthesia, consciousness and hydrophobic pockets - a unitary quantum hypothesis of anesthetic action," *Toxicology Letters*, vol. 101, pp. 31-39, Nov.1998.
- [92] S. Hameroff, "Quantum computation in brain microtubules? The Penrose-Hameroff 'Orch OR' model of consciousness," *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, vol. 356, no. 1743, pp. 1869-1896, Aug.1998.
- [93] S. Hameroff and R. Penrose, "Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness," *Mathematics and Computers in Simulation*, vol. 40, no. 3-4, pp. 453-480, Apr.1996.
- [94] R. Penrose and M. Longair, *The large, the small and the human mind*. Cambridge: Cambridge University Press, 1999.