



FACULDADE DE CIÊNCIAS E TECNOLOGIA DA UNIVERSIDADE
DE COIMBRA

LICENCIATURA EM ENGENHARIA BIOMÉDICA

“SISTEMA PARA CALIBRAÇÃO DE DISPARO DE LASER ATRAVÉS DE LÂMPADA DE FENDA”

RELATÓRIO DA DISCIPLINA DE PROJECTO

REALIZADO NO



AIBILI – ASSOCIAÇÃO PARA INVESTIGAÇÃO BIOMÉDICA E INOVAÇÃO EM
LUZ E IMAGEM

POR

Lionel Monteiro Morgado

Coimbra, Setembro de 2007

“SISTEMA PARA CALIBRAÇÃO DE DISPARO DE LASER ATRAVÉS DE LÂMPADA DE FENDA”

RELATÓRIO DA DISCIPLINA DE PROJECTO

POR

Lionel Monteiro Morgado

Orientador:

Prof. Dr. António Miguel Morgado

Supervisor:

Eng.º Rui Bernardes

Resumo

O projecto “Sistema para calibração de disparo de laser através de lâmpada de fenda”, realizado no ano lectivo de 2006/2007, teve como objectivo implementar um sistema para a detecção da localização do disparo de um laser através de uma lâmpada de fenda *TAKAGI S10-15* equipada com câmara *Sony 3 CCD ExwaveHAD* e com *frame grabber Coreco Imaging PC-RGB*.

A falta de imagens contendo um guia laser, que serviria como referência na determinação da posição do disparo, levou a que o *software* de detecção tivesse de ser construído com base em princípios e suposições teóricas, não tendo sido testado num ambiente real de aplicação.

Face a esta situação foram alterados os objectivos iniciais do projecto.

Dos novos objectivos surgiu um módulo de segmentação das imagens adquiridas através do biomicroscópio com lâmpada de fenda.

Por concluir, ficou a implementação de uma aplicação em *software* existente, para segmentação em contínuo das imagens adquiridas pelo sistema disponibilizado para o projecto.

Os módulos de *software* desenvolvidos mostram capacidade para serem implementados em aplicações em tempo-real.

Índice

1. Introdução	5
1.1. Objectivos e limitações.....	5
1.2. Estrutura do documento.....	7
2. O olho humano	8
3. Fotocoagulação laser da retina	11
4. Visão computacional e sistemas de inspecção	13
4.1. Pré-processamento da imagem.....	15
4.1.1. Filtragem.....	15
4.1.2. Detecção de arestas.....	17
4.1.2.1. Operador <i>Canny</i>	17
4.1.3. Aumento do contraste.....	19
4.2. Segmentação da imagem.....	19
4.2.1. <i>Thresholding</i>	20
4.2.2. Relações básicas entre pixels.....	21
4.2.2.1. Vizinhança de um pixel.....	21
4.2.2.2. Adjacência, conectividade, regiões e fronteiras.....	22
4.2.3. <i>Region growing</i>	24
4.2.4. Segmentação da cor.....	24
4.2.4.1. Efeitos da cor da iluminação em imagens a cor.....	28
4.3. Processamento de imagens binárias.....	30
4.3.1. Processamento morfológico das imagens.....	30
4.3.2. Descritores de objectos.....	33
4.4. Classificação de objectos.....	34
5. Aquisição das imagens do fundo ocular	36
5.1. Lâmpada de fenda <i>TAKAGI S70-15</i>	36
5.2. Câmara <i>3CCD SONY ExwaveHAD DXC-390</i>	39
5.3. PC e <i>frame grabber CORECO IMAGING PC-RGB</i>	40

6. Detecção do guia laser	42
6.1. Princípios matemáticos.....	42
6.1.1. Desvio padrão.....	42
6.1.2. Variância.....	43
6.1.3. Covariância.....	43
6.1.4. A matriz de covariância.....	44
6.1.5. Valores próprios.....	45
6.1.6. Análise em componentes principais.....	45
6.1.6.1. Aplicação do método para construção de modelos de imagens.....	46
6.1.7. Decomposição em valores singulares.....	48
6.1.7.1. SVD e normas de matrizes.....	50
6.1.7.2. SVD e ordem matricial.....	52
6.1.8. PCA e SVD.....	53
6.1.9. Transformada de Hough.....	55
6.2. O algoritmo de detecção do guia laser.....	57
6.2.1. Aumento do contraste do guia laser.....	59
6.2.2. Detecção de arestas.....	64
6.2.3. Detecção do fundo e eliminação de arestas da fronteira.....	65
6.2.4. Aplicação da transformada de <i>Hough</i> circular.....	66
6.2.5. Determinação da posição do guia laser.....	67
6.2.6. Resultados da detecção do guia laser.....	67
7. Segmentação das imagens adquiridas	69
7.1. Princípios matemáticos.....	72
7.1.1. Classificadores.....	72
7.1.1.1. Análise discriminante.....	73
7.1.1.1.1. Análise discriminante <i>stepwise</i>	73
7.1.1.1.2. Análise canónica.....	76
7.1.1.1.3. Interpretação das funções discriminantes.....	77
7.1.1.1.4. Condições assumidas.....	77
7.1.1.1.5. Predição e classificação de casos.....	79
7.1.1.1.6. Análise discriminante linear de <i>Fisher</i>	81

7.1.1.1.7. Selecção de modelos e o dilema do deslocamento-variância.....	83
7.2. Segmentação estática e segmentação dinâmica.....	85
7.2.1. Segmentação estática.....	85
7.2.2. Segmentação dinâmica.....	86
7.3. Análise estatística das imagens.....	86
7.3.1. Recolha de dados.....	86
7.3.2. Interpretação dos resultados e escolha de um modelo.....	89
7.4. Algoritmo de segmentação	93
7.4.1. Segmentação da cor.....	95
7.4.2. Segmentação morfológica.....	100
7.4.3. Resultados da segmentação.....	103
8. Implementação do algoritmo de segmentação no <i>Visionário</i>	105
9. Conclusões	108
Referências bibliográficas	110
ANEXOS A – Análise discriminante linear de <i>Fisher</i>	
ANEXOS B - Tempos de processamento nos algoritmos desenvolvidos	

Lista de símbolos e siglas

x	Operador de multiplicação
\oplus	Dilatação morfológica
\ominus	Erosão morfológica
\circ	Abertura morfológica
\bullet	Fecho morfológico
\otimes	Convolução
R	Canal vermelho do espaço de cor RGB
G	Canal verde do espaço de cor RGB
B	Canal azul do espaço de cor RGB
r	Canal normalizado de vermelho
g	Canal normalizado de verde
b	Canal normalizado de azul
θ	Ângulo de tom em graus no espaço HSI
CCD	<i>Charge-coupled device</i>
DSP	<i>Digital signal processor</i>
HSI	Sistema de cor: tom, saturação e intensidade média
ICA	<i>Independent Component Analysis</i>
PCA	<i>Principal Component Analysis</i>
RAM	<i>Random Access Memory</i>
RGB	Sistema de cor: vermelho, verde e azul
SVD	<i>Singular Value Decomposition</i>
XOR	Operação lógica exclusiva-Or

1. Introdução

Este documento visa a exposição do trabalho desenvolvido pelo seu autor, no âmbito da disciplina de projecto da licenciatura em Engenharia Biomédica.

O projecto decorreu no Centro de Novas Tecnologias para a Medicina (CNTM) da Associação para Inovação Biomédica e Investigação em Luz e Imagem (AIBILI) durante o ano lectivo de 2006/2007.

Este projecto encontra-se inserido num outro maior, que visa a criação de um sistema automático para fotocoagulação laser da retina em pacientes humanos.

A construção do mesmo foi dividida em vários módulos, que se pretende virem mais tarde a fazer parte do mesmo sistema.

O trabalho inicialmente proposto diz respeito a um deles: o módulo de calibração do disparo laser.

Com o decorrer do trabalho os objectivos foram modificados, tendo sido atribuídas novas tarefas.

Aqui será exposto o trabalho desenvolvido e os seus fundamentos teóricos.

1.1. Objectivos e limitações

Como já referido, os objectivos desta disciplina de projecto foram alterados com o decorrer do trabalho.

Definiu-se como objectivo inicial a implementação de um sistema de calibração que permitisse a determinação, em imagens do fundo ocular captadas através de um biomicroscópio com lâmpada de fenda, do local onde se encontra dirigido o laser para aplicação de fotocoagulação.

Para a aplicação de laser em retinas com patologia, é necessário identificar o alinhamento do laser com o sistema de visualização do instrumento. Isto é conseguido através de um modo de teste que permite emitir luz usando o mesmo sistema óptico.

A posição do disparo laser é dada pela posição de um guia laser no momento do disparo. Este guia laser, emitido sobre a retina no modo de teste, surge usualmente como um ponto luminoso de forma circular com aproximadamente um décimo do diâmetro de um disco óptico humano.

Este tipo de aplicação torna-se extremamente útil para a gestão da terapêutica, pois o processo de fotocoagulação pode envolver centenas de disparos muitas das vezes distribuídos por várias sessões. A acrescentar a isto temos ainda o facto de muitas das vezes o efeito dos disparos laser não ser imediatamente visível, sendo necessário recorrer a técnicas invasivas de fluorescência.

A preservação da informação relativa aos disparos efectuados é de extrema importância, pois permite avaliar a eficácia da terapêutica realizada e planear novas intervenções cirúrgicas.

A falta de imagens do fundo ocular contendo o referido guia laser, impediu que se progredisse no refinamento e teste do módulo de detecção do guia laser. Assim, decidiu-se alterar o objectivo inicial do projecto e criar um novo objectivo. Como segundo objectivo decidiu-se criar um algoritmo para segmentação da imagem adquirida através do biomicroscópio com lâmpada de fenda. Este algoritmo de segmentação tem como objectivo a rejeição de elementos que possam surgir na imagem e que não pertençam ao fundo ocular, dotando a imagem de características que melhorem a sua visualização. A implementação do mesmo permite ainda, restringir a imagem inicial a uma região de menores dimensões para o posterior processamento por parte de outros módulos a implementar no sistema. Isto traduz-se num aumento da velocidade de processamento dos módulos seguintes, como sejam o módulo de detecção do guia laser, e outros módulos que se encontram paralelamente em desenvolvimento, como o módulo de seguimento dos movimentos da retina (*retinal tracking*).

É de referir que os módulos foram desenvolvidos tendo em conta o facto de se destinarem a ser inseridos num sistema para fotocoagulação da retina, ainda em desenvolvimento, e que como tal, serão utilizados, em simultâneo com outros processos.

Sabendo-se que o tempo médio de reacção de um indivíduo a um estímulo é de cerca de 0.2 segundos, e tomando este como um tempo próximo do tempo mínimo entre dois disparos consecutivos, a velocidade de processamento dos algoritmos desenvolvidos foi um factor determinante nas escolhas feitas durante o trabalho desenvolvido. Se por um lado uma análise mais aprofundada das imagens permite obter melhores resultados, por outro, a lentidão do processo pode comprometer a sua utilidade. Assim, teve-se sempre em conta o compromisso entre a velocidade de processamento e o aprofundamento do mesmo.

Com o decorrer do projecto, decidiu-se ainda criar uma nova funcionalidade para segmentação e visualização em contínuo das imagens de vídeo adquiridas, num programa já existente: o *Visionário*.

No final, foram então considerados os seguintes objectivos:

- Construção de um módulo para detecção de um guia laser em imagem do fundo ocular;
- Construção de um módulo de segmentação automática das imagens do fundo ocular adquiridas através de uma câmara de vídeo ligada a um biomicroscópio com lâmpada de fenda;
- Implementação de módulo para segmentação e visualização em contínuo das imagens de vídeo adquiridas, num sistema já existente, o “Visionário”.

1.2. Estrutura do documento

O documento encontra-se dividido segundo várias secções, consoante os tópicos abordados.

Começa-se por fazer uma introdução aos principais temas associados ao trabalho proposto: o olho humano e suas principais características anatómicas e fisiológicas, a fotocoagulação laser da retina, e por fim os principais conceitos relacionados com a visão computacional, o processamento de imagem e os sistemas de inspecção.

Posto isto, segue-se para a descrição propriamente dita do trabalho desenvolvido. Esta parte encontra-se dividida em três secções principais, de acordo com os objectivos propostos. Ao longo do documento são apresentados os resultados obtidos, acabando com uma visão geral do trabalho feito, e daquilo que ficou por concretizar.

2. O olho humano

O olho humano é constituído por três camadas concêntricas. A camada exterior, ou cobertura fibrosa, inclui a córnea, com o seu epitélio (a conjuntiva), e a esclerótica. A camada média, ou camada vascular, inclui a íris e a coróide. A camada interior é a camada neural ou retina. [1, 2]

O olho consegue distinguir dois aspectos da luz, o seu brilho (ou luminância) e o seu comprimento de onda (ou cor). A luz entra no olho através da córnea e passa através de uma série de fluidos e estruturas que constituem o chamado meio dióptrico (córnea, humor aquoso, cristalino e humor vítreo), e incide sobre fotoreceptores de um epitélio especializado, a retina. Estes fotoreceptores incluem os cones e os bastonetes. Os bastonetes operam melhor em condições de iluminação reduzida (visão escotópica), enquanto que os cones operam melhor em condições de forte iluminação (visão fotópica), sendo os últimos, os responsáveis pela percepção da cor. [1]

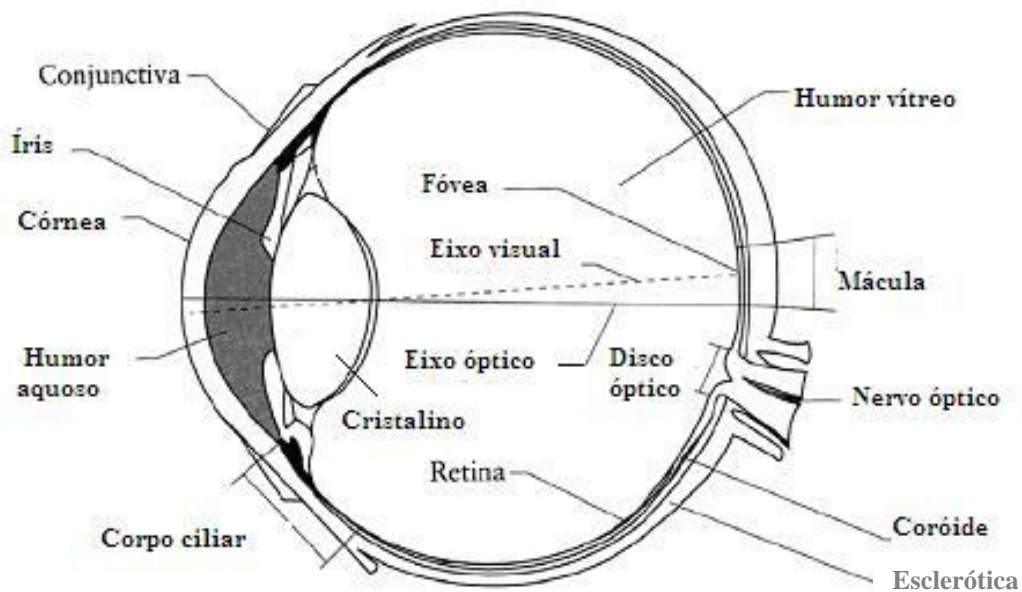


Figura 1 – Esquemática geral da anatomia do olho humano.

A informação visual é originada na retina através de interneurónios, cabendo aos axónios das células ganglionares o transporte dos sinais de saída para o cérebro. Estes axónios formam, por sua vez, o nervo óptico. [1]

Examinando o interior do olho, pode-se ver o disco óptico, que corresponde à cabeça do nervo óptico. No disco óptico a retina não contém fotoreceptores. Este facto leva a que nesta zona a luz não possa ser convertida em sinais eléctricos, criando um buraco na nossa visão. Dai o disco óptico ser muitas vezes chamado de ponto cego.

Normalmente, não nos apercebemos da existência do ponto cego, uma vez que um olho cobre o ponto cego do outro, e o cérebro preenche a informação em falta. Os vasos retinianos brotam do disco óptico em direcção à periferia da retina. Normalmente os vasos terminam um ou dois milímetros antes da periferia da retina. Dois dos maiores vasos que saem do disco óptico encontram-se localizados aproximadamente no centro desta, constituindo o arco vascular.

A mácula é uma zona oval próxima do centro da retina. Tem um diâmetro de aproximadamente 1.5 mm e é histologicamente definida como tendo duas ou mais camadas de células ganglionares. O seu reconhecimento é facilitado pelo facto de se encontrar entre estes dois grandes vasos sanguíneos. A mácula tem as dimensões aproximadas do disco óptico e apresenta uma coloração mais escurecida que o resto da retina. É também a região mais espessa da retina, tendo aproximadamente 0,22 milímetros de espessura, exceptuando a fóvea que tem apenas 0,1 milímetros de espessura. A espessura e a coloração escurecida da mácula devem-se em especial a pigmentos granulados, sendo que as outras partes da retina têm menos quantidade destes. Por sua vez, a fóvea é uma pequena área circular (diâmetro de aproximadamente 15 µm) na mácula, onde a densidade de cones e células ganglionares é muito maior que noutras zonas da retina.

Enquanto que a perda de visão periférica pode por vezes passar despercebida, a danificação da mácula resulta em perda de visão central que é facilmente perceptível, uma vez que esta estrutura é especializada para a alta acuidade na visão. A destruição progressiva da mácula é uma doença severa denominada de degeneração macular.

Na terminologia da visão, o fundo ocular refere-se à região da parte interior do olho onde se situam estruturas essenciais para a captação da imagem pelo indivíduo, como a retina, a mácula e o disco óptico.

[3]

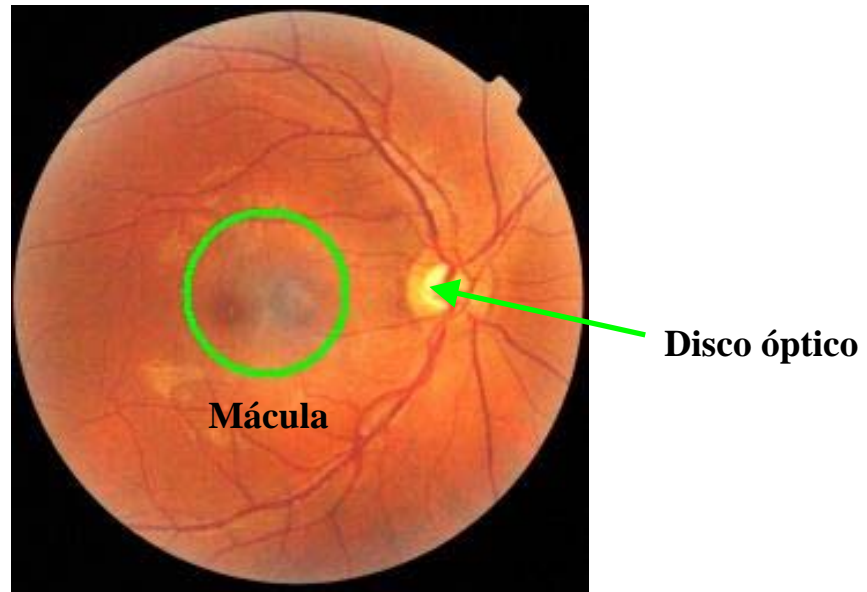


Figura 2 – Imagem do fundo ocular obtida com uma câmara de fundo, onde estão assinalados o disco óptico e a mácula.

3. Fotocoagulação laser da retina

O conceito de fotocoagulação da retina foi introduzido por Meyer-Schwicherath para tratamento da retinopatia diabética nos anos 50.

Na fotocoagulação laser da retina, o oftalmologista dispara um feixe de luz laser nas zonas da retina afectadas, que por efeito térmico vai actuar sobre a lesão, evitando a progressão da doença e o aparecimento de lesões mais severas.

O arco fotocoagulador de xénon utilizado para este propósito na altura era grande, policromático, ineficiente, e difícil de operar, o que levou à procura dum melhor método de tratamento. Uma melhoria da técnica foi conseguida com a introdução dos lasers de rubi, árgon e cripton, acoplados a lâmpadas de fenda, com um braço articulado contendo espelhos. Uma lente de contacto, um feixe apontador, e um manípulo móvel eram usados para posicionar o feixe laser na retina. Na década de 1970, estas inovações permitiram que se criassem pontos laser individuais de dimensões, potência, e duração variável, com um elevado grau de precisão e exactidão. [4]

As variantes técnicas desenvolvidas a partir deste equipamento foram refinadas e mostraram ser eficazes na prevenção de deslocamentos da retina, retinopatia diabética, Degeneração Macular Relacionada com a Idade, entre outros. [5]

A perda da visão deixou em muitos casos de ser uma inevitabilidade. Passou a ser possível impedir este mal maior, ao travar a evolução de doenças da retina para estádios mais avançados, através de uma actuação precocemente.

Apesar de tudo, poucas foram as mudanças introduzidas nos dispositivos usados em fotocoagulação da retina nas décadas seguintes.

Um paciente ao qual é aplicado fotocoagulação pode receber centenas ou mesmo milhares de disparos laser, que podem ser distribuídos ao longo de várias sessões, com vários minutos de duração, durante um período que pode ir de 2 a 4 semanas. Para além de ser um processo moroso, exige uma certa cooperação do paciente e alguma perícia por parte do cirurgião para que o processo seja bem executado, e para que se evitem complicações, que podem advir por exemplo de um disparo acidental no centro da fóvea. [4]

O laser azul-verde de árgon (488 e 514.5 nm) é o mais usado nos dias de hoje. Este induz a absorção de luz pela melanina do epitélio pigmentado, e nos vasos sanguíneos pela hemoglobina, provocando um aquecimento que induz queimaduras,

cicatrizando o tecido. Após intervenções de fotocoagulação com lasers de árgon é usual verificar-se o aumento do tamanho das cicatrizes. Este facto pode levar à deterioração da acuidade visual. Os disparos segundo micropulsos de baixa potência com díodos laser (810 nm) apresentam resultados semelhantes aos obtidos com a aplicação de lasers de árgon, provocando contudo menores danos secundários, já que a sua acção se foca no epitélio pigmentado da retina e não nos coriocalpares ou na parte exterior da retina. Desta forma, os danos provocados pela transmissão de calor da área afectada para as regiões em redor são eliminados. È contudo, apontada alguma dificuldade no uso deste tipo de laser, uma vez que as queimaduras induzidas não são imediatamente visíveis, sendo necessário recorrer a métodos de fluorescência para se determinar o efeito dos disparos executados. [6,7]

A necessidade de dispositivos mais precisos, seguros e eficazes, tem levado a um aumento de interesse pela construção de sistemas automatizados.

Uma vez que o fundo ocular é acessível através de imagens, estes sistemas podem ser construídos usando os princípios básicos por trás de qualquer sistema de inspecção baseado no processamento de imagens digitais. [5]

4. Visão computacional e sistemas de inspecção

A visão é considerada por muitos como o mais avançado dos nossos sentidos, tornando assim a imagem, um dos mais importantes contentores de informação.

Ao contrário dos seres humanos, que se encontram limitados à banda visível do espectro electromagnético, as máquinas podem cobrir a região que vai desde as ondas gama até às ondas rádio. Desta forma, podem operar em fontes que os humanos não se encontram habituados a associar com imagens. Neste grupo incluem-se os ultra-sons, a microscopia por electrões e imagens geradas por computadores. Assim, o processamento digital de imagens engloba um vasto e variado campo de aplicações.

As fronteiras do processamento digital de imagem não se encontram bem delineadas, havendo uma grande discordância entre autores na definição do conceito. Este é muitas vezes confundido com conceitos como análise de imagem e visão computacional. Segundo *Gonzalez e Woods*, a visão por computador tem como objectivo derradeiro “o uso de computadores para simular a visão humana, incluindo a aprendizagem e a capacidade para inferir e agir com base em estímulos visuais”. Por outro lado, a análise de imagem (também chamada de entendimento da imagem) encontra-se entre o processamento de imagens e a visão computacional. [8]

Muitas vezes o processamento de imagem é definido como uma disciplina em que ambas a entrada e saída do processo são imagens, o que é claramente redutor.

Exposto isto, torna-se evidente a confusão gerada entre os conceitos.

Assim, uma definição mais ajustada para processamento de imagem será a de qualquer forma de processamento de informação para a qual a entrada é uma imagem, tal como uma fotografia ou uma *frame* de um video, não sendo a saída necessariamente uma imagem, mas por exemplo um conjunto de características da imagem.

Muitas técnicas de processamento de imagem envolvem o tratamento da mesma como um sinal bidimensional e aplicam técnicas padrão de processamento de sinal.

A visão computacional é frequentemente utilizada em sistemas industriais de controlo de qualidade, uma vez que apresenta vantagens em relação à visão humana, como: rapidez, exactidão e baixos custos. O seu desempenho pode ainda ser melhor que o da visão humana, quando é necessário extrair informação precisa de imagens de uma forma rápida e/ou repetitiva. As máquinas podem ainda ser utilizadas em ambientes perigosos ou inacessíveis ao ser humano. A visão por computador pode ser utilizada

para aplicações de segurança ou vigilância, como por exemplo, na detecção e reconhecimento de faces e sistemas de monitorização de intrusos. Na medicina, pode ser aplicada na detecção de lesões e anormalidades, ou em terapêuticas com recurso a radiação na fase de planeamento da zona alvo onde se deve aplicar o tratamento, de forma a não se danificar tecido circundante.

Num sistema típico de inspecção uma câmara encontra-se ligada a um computador com *software* adequado para a execução de tarefas de visão computacional. A câmara capta imagens do item ou cena e envia-as ao computador. Aqui, ocorre normalmente uma fase de pré-processamento que visa o aumento de qualidade das imagens. Após este pré-processamento, as imagens são segmentadas para que os objectos de interesse sejam separados dos restantes e do fundo. De seguida, os objectos são classificados segundo classes pré-definidas. Os resultados da classificação podem ser usados para se decidir que operações realizar sobre o objecto na imagem.

Um sistema típico de visão computacional extrai informação quantitativa de uma imagem, recorrendo aos seguintes passos:

- Adquire uma imagem contendo a informação requerida. Isto é feito com iluminação, óptica, uma câmara CCD e um digitalizador. As imagens resultantes são armazenadas numa memória especial de alta velocidade chamada *frame buffer*.
- Processa a imagem para reduzir as distorções introduzidas pela aquisição da imagem. Isto poderia incluir correcção para as lentes e distorção geométrica da câmara e correcções para variações na iluminação.
- Processa as imagens para reduzir o ruído e salientar características (detalhes dos objectos) que contêm a informação desejada. Um exemplo, pode ser o aumento da intensidade das arestas como objectivo de se facilitar a sua medição.
- Extrai as características da imagem, como por exemplo a posição e dimensão de objectos, posição de arestas, cor e orientação dos objectos.
- Reduz os dados por processamento do conjunto de características. Isto é feito para agrupar características, por exemplo diferentes arestas do mesmo objecto, bem como para remover características produzidas devido a ruído.
- Devolve os resultados ou toma uma decisão.

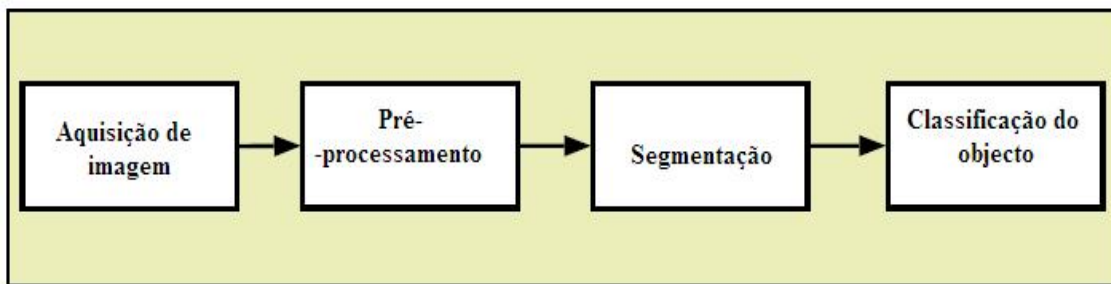


Figura 3 – Etapas de um sistema típico de visão computacional

Um sistema de visão computacional em tempo-real é rápido o suficiente para acompanhar o processo para o qual está a ser usado. Assim, dependendo da tarefa a que se destina, a sua velocidade de processamento pode ir de milésimas de segundo a alguns segundos.

As aplicações em tempo-real de menor rapidez podem ser executadas com um *frame grabber* (aquisição de imagem e *frame buffer*) e um computador de uso genérico rápido. À medida que uma maior rapidez se torna uma exigência, é necessário recorrer a *hardware* especial para se manter a operação em tempo real. Isto pode ocorrer para elevadas quantidades de dados por unidade de tempo, processamento extensivo e tempos de processamento muito curtos. [9]

4.1. Pré-processamento da imagem

Pretende-se com esta etapa aumentar a qualidade da imagem por redução do ruído presente, evidenciando características que são usadas na etapa de segmentação. Duas técnicas típicas que são usadas no pré-processamento são a filtragem e o aumento de contraste.

4.1.1. Filtragem

Uma imagem digital em tons de cinzento pode ser definida como uma função $f(x,y)$, onde x e y são coordenadas espaciais discretas e f é o nível de cinzento discreto da imagem. Assim, uma imagem digital consiste num conjunto finito de elementos com

um valor finito de níveis de cinzento. Os elementos discretos de uma imagem digital são chamados de elementos da imagem ou pixels.

A filtragem é uma técnica comum, usada para aumentar a qualidade da imagem ou adaptá-la para posterior processamento. Através da filtragem pode-se salientar fronteiras de objectos nas imagens ou eliminar ruído. A filtragem pode ser executada tanto no domínio espacial como no das frequências. A filtragem espacial opera directamente sobre os pixels que compõem as imagens, enquanto que a filtragem no domínio das frequências exige uma conversão inicial para este domínio. Esta conversão do domínio espacial para o domínio das frequências pode ser feito através de uma transformada discreta de Fourier. Após filtrada, a imagem deve ser retransformada para o domínio espacial, através de uma transformada inversa de Fourier.

A filtragem espacial é baseada numa pequena máscara que é simplesmente movida de pixel em pixel na imagem. Em cada pixel, o resultado da filtragem é calculado através de uma relação matemática previamente definida. Em geral, a máscara de um filtro pode ter as dimensões $m \times n$, onde a largura $m = 2a+1$, a altura $n = 2b+1$, e a e b são inteiros não negativos. Para filtragem linear espacial, como por exemplo suavização e salientação de arestas, a imagem final g em cada pixel da imagem (x, y) pode ser calculada como a soma dos produtos dos coeficientes de filtragem w e os correspondentes pixels da imagem f como

$$g(x, y) = \sum_{s=-bt}^b \sum_{t=-a}^a w(t, s) f(x+t, y+s)$$

É de referir que os pixels que se encontram mais próximos do bordo da imagem que a na vertical da imagem e b na horizontal, não podem ser filtrados, uma vez que a localização dos pixels na equação deve ser não-negativa. A solução para este problema pode passar por omissão destes pixels marginais, resultando uma imagem de menores dimensões que a original, ou então atribuindo-lhes o valor de 0.

Os filtros de suavização podem ser usados para suavizar imagens ou reduzir o seu ruído. O resultado de uma suavização corresponde simplesmente à média dos pixels abrangidos pela máscara. Em alternativa, pode-se recorrer a um filtro de mediana, que para além de eliminar eficazmente certo tipo de ruídos aleatórios, não borrata tanto as imagens como um filtro de média. Um filtro de mediana limita-se a atribuir como resultado a mediana dos pixels abrangidos pela sua máscara. Por exemplo, para um

filtro 3 x 3 que englobe os valores 30, 35, 70, 30, 35, 40, 35, 30 e 5, os valores são dispostos por ordem crescente, segundo 5, 30, 30, 30, 35, 35, 35, 40 e 70, sendo a mediana o valor 35, pois este encontra-se a meio dos valores abrangidos.

4.1.2. Detecção de arestas

Com a detecção de arestas pretende-se marcar os pontos da imagem digital onde há uma transição abrupta dos níveis de intensidade dos pixels. Estas variações abruptas reflectem muitas vezes mudanças importantes nas cenas ou imagens. São exemplos, as descontinuidades devido à profundidade, à orientação das superfícies, às propriedades dos materiais ou devido a variações na iluminação.

A detecção de arestas reduz significativamente o conjunto de dados, eliminando informação irrelevante, e preservando informação estrutural considerada importante. Há vários métodos para detecção de arestas, mas a maioria pode ser dividida segundo dois grupos: *search-based* e *zero-crossing based*. Os métodos *search-based* detectam arestas procurando um máximo ou mínimo local da primeira derivada da imagem, usualmente o máximo direccional local do gradiente da magnitude. Os métodos *zero-crossing based* procuram os zeros na segunda derivativa da imagem com o objectivo de encontrar arestas, usualmente os zeros do Laplaciano ou os zeros de uma expressão diferencial não-linear. [8]

4.1.2.1. Operador *Canny*

Actualmente, as técnicas mais usadas para detecção de arestas fazem uso do operador *Canny* ou variações do mesmo. A sua criação resultou da procura de um algoritmo que tivesse um bom desempenho em situações reais, marcando apenas uma vez, o máximo de arestas reais (boa detecção), o mais próximo possível destas na imagem (boa localização), e desprezando o efeito do ruído na imagem real, o tanto quanto possível, por forma a se evitar a criação de falsas arestas (resposta minimalista).

O filtro optimizado pode ser aproximado através da primeira derivada de uma função gaussiana:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

A primeira etapa da aplicação do operador de *Canny* consiste numa suavização por convolução da imagem com um filtro gaussiano. De seguida, calcula-se o gradiente da imagem suavizada, por convolução da mesma com a derivativa da gaussiana em ambas as direcções vertical e horizontal.

Esta operação é descrita por:

$$\begin{aligned} I'(x, y) &= g(k, l) \otimes I(x, y) \\ &= \sum_{k=-N}^N \sum_{l=-N}^N g(k, l) I(x-k, y-l) \end{aligned}$$

, onde $g(k, l)$ é o modelo de convolução, $I(x, y)$ é a imagem original, $I'(x, y)$ é a imagem filtrada, e $2n+1$ é a dimensão da máscara de convolução.

Segue-se uma etapa denominada de supressão não-máxima, onde se procuram os valores máximos na direcção do gradiente, e se eliminam todos os outros, reduzindo as arestas falsas. O máximo local é encontrado por comparação do pixel com os vizinhos ao longo da direcção do gradiente. Isto permite que as arestas sejam finas, e constituídas por apenas um pixel de largura

Segue-se a aplicação de um *threshold* ou valor limiar de corte. Em vez de usar apenas um limiar estático, o filtro *Canny* introduz a limitação por histerese, que permite alguma adaptação ao conteúdo local da imagem. Há dois níveis de *threshold*, um alto t_a e um baixo t_b , em que $t_a > t_b$. Os pontos com valores acima de t_a são automaticamente considerados como arestas. Traçando o contorno da aresta, os pixels vizinhos com valores de gradiente menores que t_a e maiores que t_b podem ser considerados como pertencentes a uma aresta. Esta metodologia permite ultrapassar problemas relacionados com descontinuidades das arestas, uma vez que permite preservar as arestas menos acentuadas, mantendo a supressão de algum ruído. No entanto, este processo de histerese torna o algoritmo mais lento, apesar de mais eficiente.

O desempenho do algoritmo de *Canny* depende fortemente do ajuste de dois parâmetros: σ , que é o desvio padrão do filtro gaussiano, e os valores de *threshold*, t_a e t_b . Quanto maior o valor de σ , maior é a dimensão do filtro gaussiano. Isto implica uma

maior suavização da imagem, necessária para redução do ruído, bem como para detecção de arestas mais largas. Contudo, um valor mais elevado de σ , implica uma maior imprecisão na localização da aresta e uma eliminação das arestas menos pronunciadas. Assim, este parâmetro pode ser ajustado para diferentes ambientes com diferentes níveis de ruído. [10]

No que toca ao limiar, a utilização de dois limiares com histerese, permite maior flexibilidade do que no caso em que se usa apenas um, continuando contudo a persistir os problemas típicos da sua aplicação. Um valor muito elevado do limiar pode levar à perda de informação importante. Por outro lado, um limiar demasiado baixo pode levar a que seja considerada informação irrelevante. É pois difícil definir um valor que funcione bem em todas as imagens. [8, 10, 11]

4.1.3. Aumento do contraste

A fraca iluminação, problemas relacionados com a abertura das lentes ou uma gama desajustada dos sensores de imagem, podem gerar imagens com baixo contraste.

O aumento de contraste de uma imagem pode ser conseguido com técnicas como o *contrast-stretching*, em que os níveis de cinzento dos pixels são “esticados” para novos valores. Um caso especial de *contrast-stretching* é a equalização do histograma, em que os níveis de cinzento são “esticados” para que toda a escala de cinzento disponível seja utilizada.

O aumento de contraste local é também uma técnica de melhoramento do contraste baseada no histograma. Nesta técnica, a imagem é dividida em sub-blocos e a equalização do histograma é feita separadamente para cada bloco. Após a equalização do histograma, a imagem é construída por interpolação dos sub-blocos, a fim de suavizar as regiões de transição.

Qualquer destes tipos de melhoramento do contraste pode ser feito de forma independente para cada canal de cor.

4.2. Segmentação da imagem

A segmentação é referida como o processo de partição de uma imagem digital em várias regiões ou conjuntos de pixels.

Através da segmentação, uma imagem é dividida em objectos ou regiões. Esta divisão pode ser mais ou menos complexa e ter mais ou menos passos de acordo com os objectos de interesse que se pretende isolar. A segmentação de objectos não triviais é considerada uma das tarefas mais difíceis no processamento de imagens.

Os algoritmos de segmentação são baseados em descontinuidade ou similaridade dos valores dos pixels. Numa abordagem baseada em descontinuidades, uma imagem é dividida com base em mudanças abruptas na intensidade dos pixels, tais como arestas, pontos, grupos contínuos de pixels e linhas. Numa abordagem de similaridade, um conjunto de critérios pré-definidos é usado na decomposição da imagem. *Thresholding*, *region growing* e *region splitting e merging* são exemplos de técnicas de segmentação baseadas na similaridade.

Há técnicas mais complexas que no entanto não serão abarcadas neste documento, uma vez que saem fora do âmbito desta disciplina de projecto.

4.2.1. Thresholding

Esta é uma das técnicas mais importantes e utilizadas, pois é intuitiva e simples de implementar.

Esta técnica assenta no princípio que os objectos podem ser separados do seu fundo com base no seu nível de cinzento. No *thresholding* os pixels com níveis de intensidade abaixo de determinado valor definido são considerados como pertencentes ao fundo, enquanto que os restantes são considerados como pertencentes à cena.

O *thresholding* pode ser considerado como uma operação que envolve testes da variável T :

$$T = [x, y, p(x, y), f(x, y)]$$

, onde f é o nível de cinzento do pixel na localização (x, y) e $p(x, y)$ é uma propriedade local do pixel. A imagem $g(x, y)$ à qual é aplicado o *thresholding*, pode ser definida como:

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) > T \\ 0 & \text{se } f(x, y) \leq T \end{cases}$$

Quando T depende apenas de $f(x, y)$, o *threshold* é denominado de global. Se T depender de $f(x, y)$ e de $p(x, y)$, então é denominado de local. Se também depender de x e y então o *threshold* é dito dinâmico ou adaptativo.

Esta técnica apresenta contudo problemas quando aplicada a objectos ou regiões com iluminação não-uniforme.

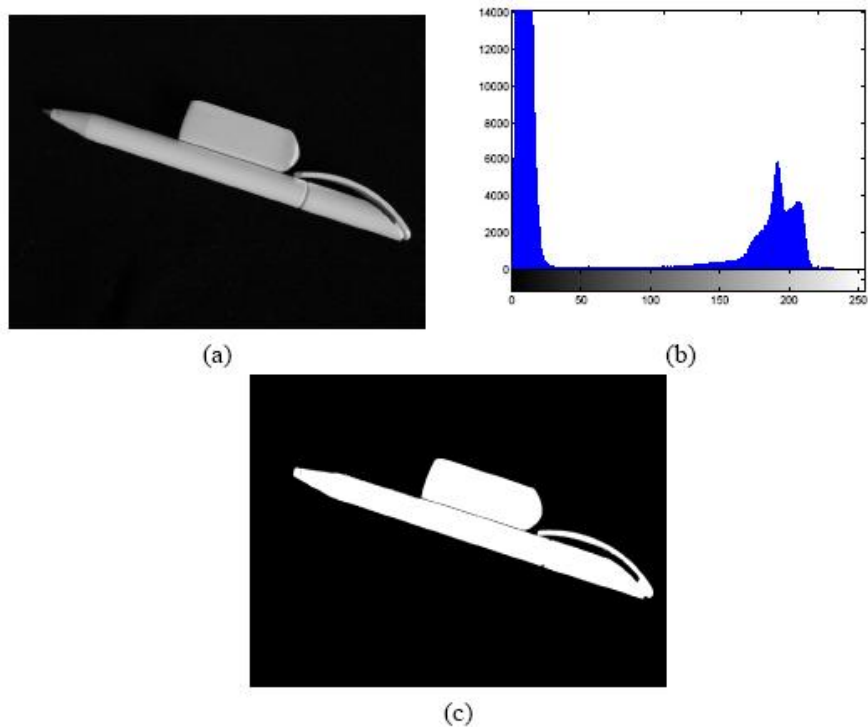


Figura 4 – Exemplo de *thresholding*: (a) imagem original em tons de cinzento; (b) histograma da imagem original; (c) resultado da aplicação de um valor de *threshold* ou limiar de corte, de 100.

4.2.2. Relações básicas entre pixels

Apesar da sua simplicidade, as ideias expostas nesta secção são de extrema importância definindo conceitos fundamentais para o processamento de imagem digital. Aqui são expostos os princípios básicos de relacionamento entre pixels.

4.2.2.1. Vizinhança de um pixel

Sendo que uma imagem pode ser descrita como uma função $f(x, y)$, um pixel p nas coordenadas (x, y) tem quatro vizinhos horizontais e verticais cujas coordenadas são dadas por

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1).$$

Este conjunto de pixels chamados de *4-neighbors* (4-vizinhos) de p , é denotada por $N_4(p)$. Cada pixel é uma unidade de distância de (x, y) , e alguns dos pixels de p incidem fora da imagem digital se (x, y) for na margem da mesma.

Os quatro vizinhos diagonais de p têm coordenadas

$$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

, e são denotados por $N_D(p)$. Estes pontos, conjuntamente com os *4-neighbors*, são chamados de *8-neighbors* de p , denotados por $N_8(p)$. Da mesma forma que no caso anterior, alguns pontos incidem fora da imagem se (x, y) se encontrar na fronteira da imagem.

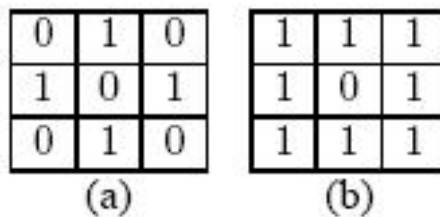


Figura 5 – Representação das vizinhanças de pixels: (a) *4-neighbors*; (b) *8-neighbors*.

4.2.2.2. Adjacência, conectividade, regiões e fronteiras

A conectividade entre pixels é um conceito fundamental que simplifica a definição de numerosos conceitos, tais como região e fronteira. A fim de se estabelecer se dois pixels se encontram ligados, deve ser determinado se eles são vizinhos e se o seu nível de cinzento satisfaz um critério específico de similaridade. Por exemplo, no caso de uma imagem contendo pixels com 0s e 1s, dois pixels podem ser *4-neighbor*, mas diz-se que eles estão ligados apenas se os seus valores forem iguais.

Seja V o conjunto de níveis de cinzento usados para definir a adjacência, e considerando uma imagem binária, os pixels com valor 1 são denotados por $V = \{1\}$.

Numa imagem em tons de cinzento a ideia de adjacência é similar, mas o conjunto V contém mais elementos. Por exemplo na adjacência de pixels com valores entre 0 e 255, o conjunto V pode ser qualquer valor entre estes 256 valores.

Consideram-se três casos de adjacência:

- a) adjacência-de-4 – dois pixels p e q com valores de V são adjacentes-de-4 se se encontrarem no conjunto $N_4(p)$.
- b) adjacência-de-8 - dois pixels p e q com valores de V são adjacentes-de-8 se se encontrarem no conjunto $N_8(p)$.
- c) adjacência-de- m (adjacência misturada) - dois pixels p e q com valores de V são m -adjacentes se:
 - i) q se encontra em $N_4(p)$, ou
 - ii) q se encontra em $N_D(p)$ e o conjunto $N_4(p) \cap N_4(q)$ não tem pixels com valores de V .

A adjacência misturada é uma modificação da adjacência-de-8. É introduzida para eliminar ambiguidades que surgem muitas vezes quando utilizando a adjacência-de-8.

Um caminho ou curva (digital) do pixel p com coordenadas (x, y) até ao pixel q com coordenadas (s, t) é uma consequência de pixels distintos com coordenadas

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

, onde $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, e os pixels (x_i, y_i) e (x_{i-1}, y_{i-1}) são adjacentes para $1 \leq i \leq n$. Neste caso, n é o comprimento do caminho. Se $(x_0, y_0) = (x, y)$, o caminho é fechado.

Se S um subconjunto de pixels numa imagem, diz-se que dois pixels p e q estão ligados em S se existir um caminho entre eles constituído apenas por pixels de S . Para qualquer pixel p em S , o conjunto de pixels que se encontram ligados a este em S são chamados de componentes ligados de S . Se apenas tiver um componente ligado, então o conjunto de S é chamado de conjunto ligado.

Seja R um subconjunto de pixels numa imagem, chamamos R uma região da imagem se R é um conjunto ligado. A fronteira ou contorno da região R é o conjunto de pixels na região que tem um ou mais vizinhos que não se encontram em R .

Normalmente, quando se refere a uma região, refere-se a um subconjunto da imagem, e quaisquer pixels na fronteira da região que possam coincidir com a margem da fronteira são incluídos implicitamente como parte da região de fronteira.

4.2.3. Region growing

Este método consiste no agrupamento de um conjunto ou subconjunto de pixels numa região maior de acordo com um determinado critério, partindo de certos pontos iniciais. As regiões onde se encontram estes pontos iniciais vão aumentando por anexação de regiões de pixels vizinhos com características semelhantes aos pixels já incluídos na região. A anexação é especificada por uma regra de conectividade, como por exemplo a *4-neighbour* ou a *8-neighbour*.

Este processo pára quando não há mais pixels a corresponder às características especificadas, como seja a cor, a textura ou o nível de cinzento. Pode-se ainda definir outras características para se melhorar o desempenho desta técnica, como o tamanho das regiões, a forma, e a semelhança entre pixels candidatos. [8]

4.2.4. Segmentação da cor

A colorimetria é a ciência que estuda a medida das cores, e que desenvolve métodos para a quantificação da cor, ou seja para o desenvolvimento de valores numéricos da cor.

Uma imagem a cores é constituída normalmente por três canais de cor que podem ser considerados como imagens separadas em níveis de cinzento. Desta forma, a segmentação de imagens pode ser resumida a uma segmentação de imagens em níveis de cinzento. Os resultados das segmentações separadas são depois associados naquela que será a imagem colorida final.

Vários espaços de cor têm sido desenvolvidos de acordo com o fim a que se destinam. A escolha de um espaço de cor é por vezes difícil e deve reflectir a aplicação a que se destina.

No sistema de cor RGB, tem-se três canais distintos que especificam as quantidades de vermelho (R), verde (G) e azul (B).

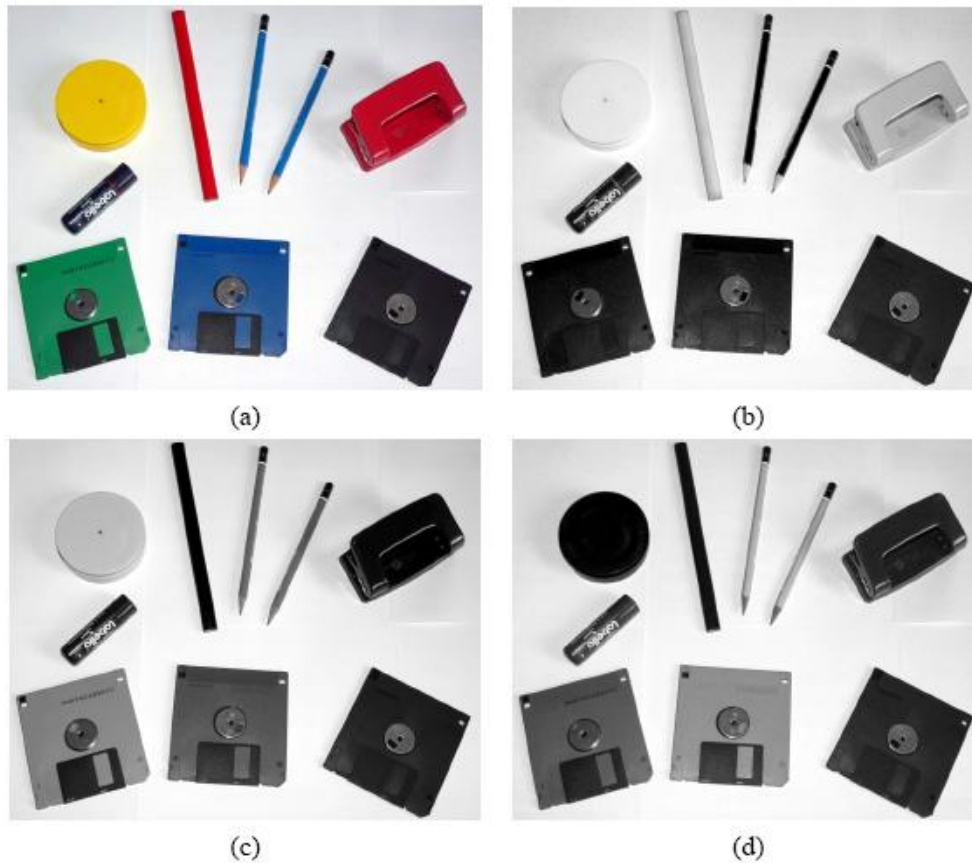


Figura 6 – Canais de cor na imagem RGB: (a) imagem RGB original; (b) Canal vermelho (R) da imagem original; (c) canal verde (G) da imagem original; (d) canal azul (B) da imagem original.

Um problema comum no processamento de imagem está relacionado com a correlação entre as variáveis. Por exemplo, em imagens naturais, R, G e B tendem a estar correlacionadas uma vez que uma variação da intensidade de iluminação e direcção (sombreado) produz efeitos semelhantes nos três canais de cor. Desta forma, torna-se necessário normalizar as variáveis de forma a eliminar este efeito que pode muitas vezes ser indesejável.

A conversão das componentes absolutas em valores relativos, permite eliminar a informação relativa à intensidade, passando estas novas variáveis a referir a proporção de cor independentemente da sua intensidade, segundo

$$r = \begin{cases} \frac{R}{S} & \text{se } S > 0 \\ 0 & \text{se } S = 0 \end{cases} \quad g = \begin{cases} \frac{G}{S} & \text{se } S > 0 \\ 0 & \text{se } S = 0 \end{cases} \quad b = \begin{cases} \frac{B}{S} & \text{se } S > 0 \\ 0 & \text{se } S = 0 \end{cases}$$

, com $S = R + G + B$.

Os canais podem tornar-se ainda mais decorrelacionados por recurso a técnicas como a análise em componentes principais (PCA) e a análise em componentes independentes (ICA).

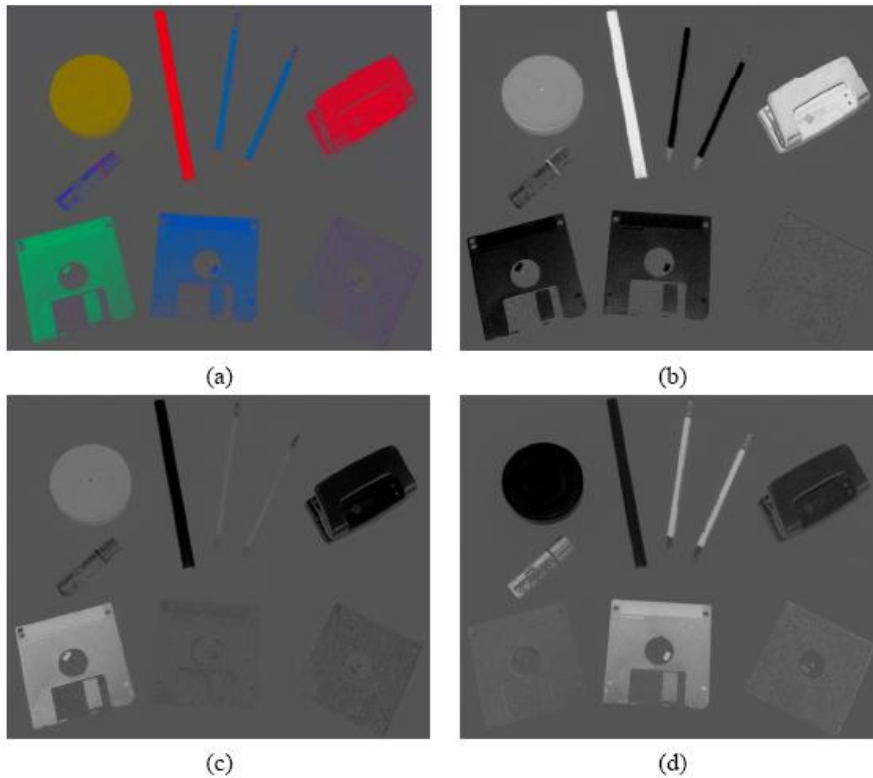


Figura 7 – Canais normalizados de RGB: (a) imagem RGB normalizada; (b) Canal de vermelho normalizado; (c) canal de verde normalizado; (d) canal de azul normalizado.

Depois de convertidas as componentes de cor para os valores relativos, pode-se segmentar a imagem recorrendo a técnicas como o *thresholding*, isto para cada canal de cor separadamente. Como resultado obtêm-se máscaras binárias, ou seja, matrizes com o mesmo número de linhas e colunas da imagem inicial, mas que em lugar dos níveis de intensidade, contêm valores lógicos (0 ou 1) que discriminam se o pixel correspondente obedece à condição especificada (ex: intensidade do pixel maior que 45). As máscaras para os diferentes canais são posteriormente combinadas recorrendo a operadores lógicos. A máscara pode ser por fim aplicada à imagem, pela sua multiplicação com cada matriz de cor da imagem inicial, preservando-se apenas as regiões de interesse.

[8, 12]

Há ainda sistemas de cores mais intuitivos, como o sistema HSI, que mais se aproximam da forma como o ser humano experiencia as cores. Estas três letras correspondem à primeira letra dos seguintes termos anglo-saxónicos *hue*, *saturation* e *intensity*, que são as 3 características da cor descritas pelo sistema, às quais corresponde respectivamente a terminologia portuguesa: tom, saturação e intensidade.

O tom é fisicamente o intervalo de longitude de onda entre o qual se pode escrever uma determinada cor. Na prática, é a característica que faz com que se possa reconhecer uma cor como sendo vermelha, uma outra como sendo azul, e assim por diante.

A saturação de uma cor é o seu grau de pureza. Uma cor é tanto mais saturada quanto menor for o seu conteúdo de branco ou cinza. As cores da natureza são sempre mais ou menos saturadas. As cores mais saturadas são aquelas que não são originárias de pigmentos, mas sim de fenómenos interferenciais.

A intensidade, ou luminosidade de uma cor é a característica que faz com que ela apareça mais clara do que uma outra, independente da sua saturação.

É possível converter o sistema RGB no sistema HSI através de várias relações matemáticas.

Para cada pixel, pode-se obter o valor de H a partir do sistema RGB através da seguinte relação matemática:

$$H = \begin{cases} \theta & \text{se } B \leq G \\ 360 - \theta & \text{se } B > G \end{cases}$$

com

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{\left[(R - G)^2 + (R - B)(G - B) \right]^{\frac{1}{2}}} \right\}$$

A saturação, por sua vez, estabelece com o sistema RGB a relação:

$$S = 1 - \frac{3}{R + G + B} [\min(R, G, B)]$$

E finalmente, para cada pixel tem-se a intensidade média:

$$I = \frac{(R + G + B)}{3}$$

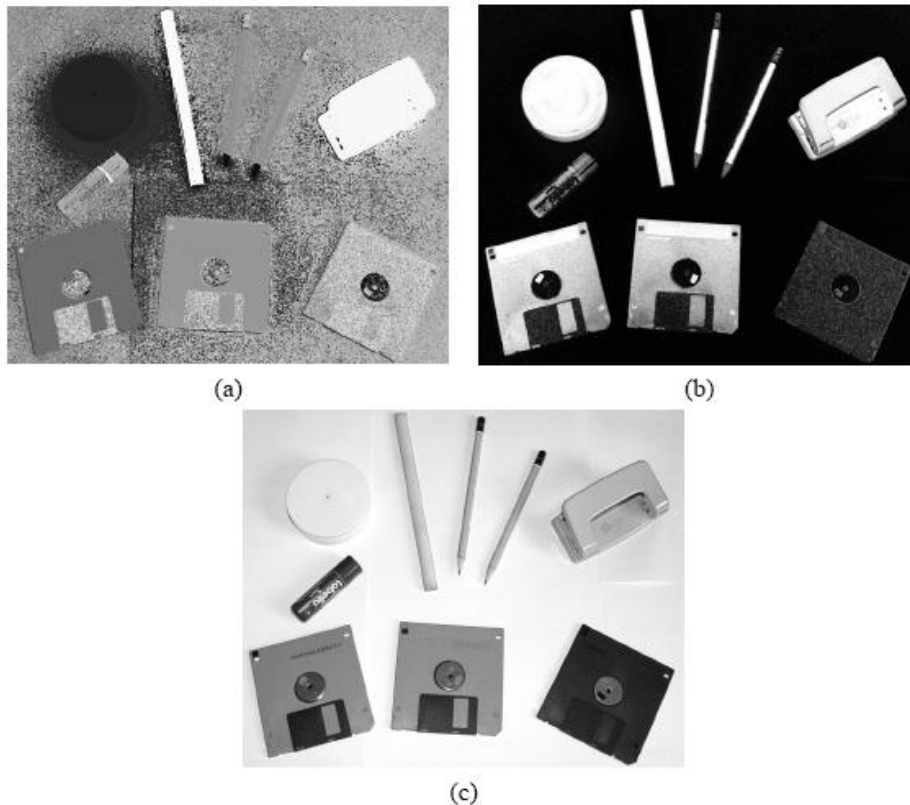


Figura 8 – Canais de cor na imagem HSI: (a) canal de tom; (b) canal de saturação; (c) canal de intensidade média

4.2.4.1. Efeitos da cor da iluminação em imagens a cor

Quando uma câmara captura uma imagem de uma cena, regista a luz reflectida por diferentes superfícies na cena. As câmaras a cor registam as intensidades relativas e as cores da luz reflectida. Assim, a imagem captada pela câmara é uma função da iluminação incidente na cena, das propriedades reflexivas de diferentes superfícies e das características da câmara. Desta forma, se a iluminação incidente numa superfície muda, então a aparência dessa superfície na imagem captada também muda. Por exemplo, no caso de uma superfície perfeitamente branca iluminada com uma luz cujo pico do

espectro de iluminação está mais próximo da região correspondente ao vermelho, a luz reflectida terá também esta característica, surgindo na imagem uma superfície avermelhada.

Um dispositivo de captura de imagem tem normalmente um ou vários sensores calibrados de forma a responderem a diferentes comprimentos de onda do espectro de luz visível. As câmaras convencionais têm geralmente três sensores, para captura de regiões do espectro vermelho (R), verde (G) e azul (B), integrando a energia luminosa recebida nos respectivos comprimentos de onda por forma a criar uma resposta.

Numa imagem digital, a imagem de saída é uma matriz de pixels, em que cada pixel resulta duma combinação da resposta de cada sensor centrada na mesma localização. A aparência da cor e a intensidade dos diferentes objectos são representadas pelos valores das respostas dos sensores. A resposta dos sensores é função da luz reflectida pelas diferentes superfícies. Assim, a cor e intensidade dos objectos é função da iluminação incidente. Os sistemas de processamento digital de imagem que se baseiam na aparência dos objectos nas imagens de entrada são seriamente afectados por este fenómeno pois a aparência é influenciada pelas propriedades da iluminação incidente. Interessa aqui as alterações da cor dos objectos nas imagens de acordo com a cor da iluminação. O processo de compensação das mudanças na aparência das cenas coloridas devido à cor da iluminação é chamada de constância de cor, e é exibida pelo sistema visual humano até um certo grau. Deste processo deriva uma descrição das cenas independente da iluminação.

É descrita na literatura a possibilidade de aproximar o mapeamento de uma imagem de uma cena adquirida sob um determinado iluminante à correspondente imagem adquirida sob um iluminante diferente usando uma transformação matricial. Para tal pode-se recorrer a matrizes 3×3 que melhor se adaptem ao efeito. [13]

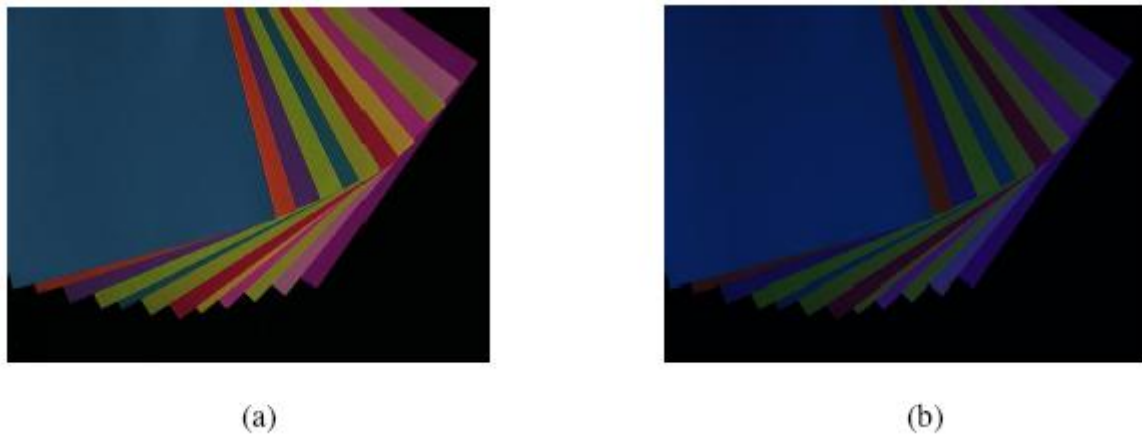


Figura 9 – Mudança da cor devido a mudanças de iluminação: (a) imagem de um conjunto de papéis coloridos, tirada sob um iluminante para o qual a câmara estava bem calibrada; (b) o mesmo conjunto de papéis em (a) mas fotografados sob uma luz azulada.

4.3. Processamento de imagens binárias

O resultado da segmentação de imagens é normalmente uma imagem binária onde os objectos ou regiões de interesse são marcados com 1s e os fundos com 0s. A imagem binária contendo informação dos objectos pode ser chamada como uma máscara binária, e permite que se continue um processamento mais aprofundado da imagem apenas nas regiões desejadas. Este tipo de objectos tem ainda utilidade na determinação de características morfológicas como a dimensão e forma, bem como na medição de características como a intensidade média entre outras.

A seguir são descritas algumas técnicas morfológicas e de medição.

4.3.1. Processamento morfológico das imagens

No processamento morfológico de imagens binárias destacam-se duas operações: a dilatação e a erosão.

Estas operações são baseadas num denominado elemento estruturante, que é uma máscara que se move ao longo da imagem num processo semelhante a uma filtragem.

Na dilatação, esta máscara verifica para cada pixel no qual é centrada o seu valor. Se este for 1, então a máscara converte todos os pixels abrangidos por si em 1s. Este processo é representado por $A \oplus B$, em que A é a imagem em processamento e B é o elemento estruturante.

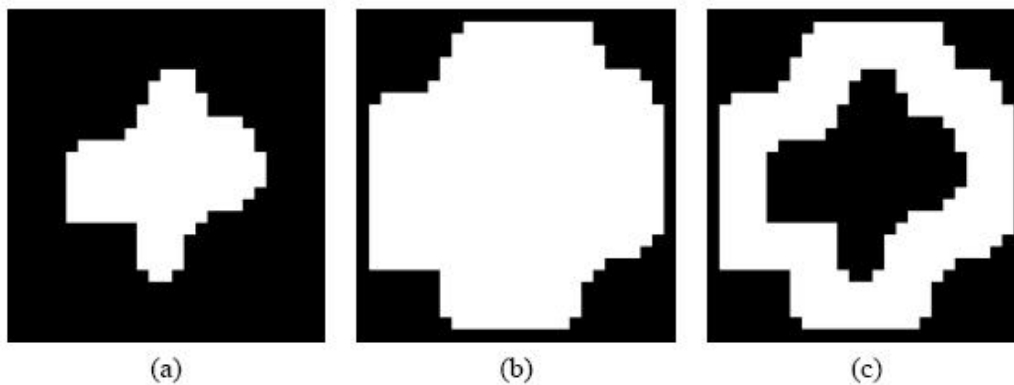


Figura 10 – Exemplo de dilatação morfológica: (a) imagem binária original de 27X28 pixels; (b) imagem dilatada com uma máscara de 9 x 9; (c) imagem contendo a fronteira resultante da aplicação da operação XOR entre (a) e (b).

Na erosão, ocorre o processo oposto, ou seja, caso o pixel central da máscara seja 0, aos pixels abrangidos pela mesma é atribuído o valor de 0 na imagem final. Este processo pode ser representado por $A \ominus B$, em que A é a imagem em processamento e B é o elemento estruturante.

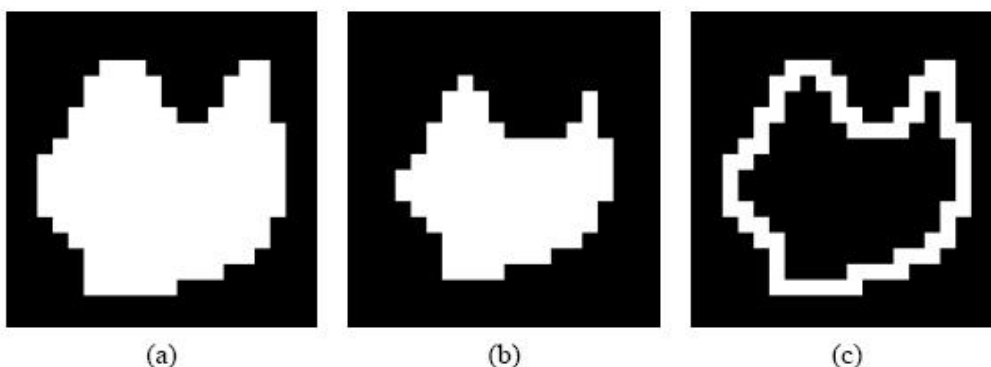


Figura 11 - Exemplo de erosão morfológica: (a) imagem binária original de 20X20 pixels; (b) imagem erodida com uma máscara de 9X9; (c) imagem contendo a fronteira resultante da aplicação da operação XOR entre (a) e (b).

O uso destas duas técnicas em conjunto permite criar aberturas e fechamentos morfológicos.

O procedimento de abertura suaviza a forma de um objecto, eliminando pequenas protusões, e removendo istmos estreitos. Esta operação consiste numa erosão seguida de uma dilatação, e pode ser representada por:

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

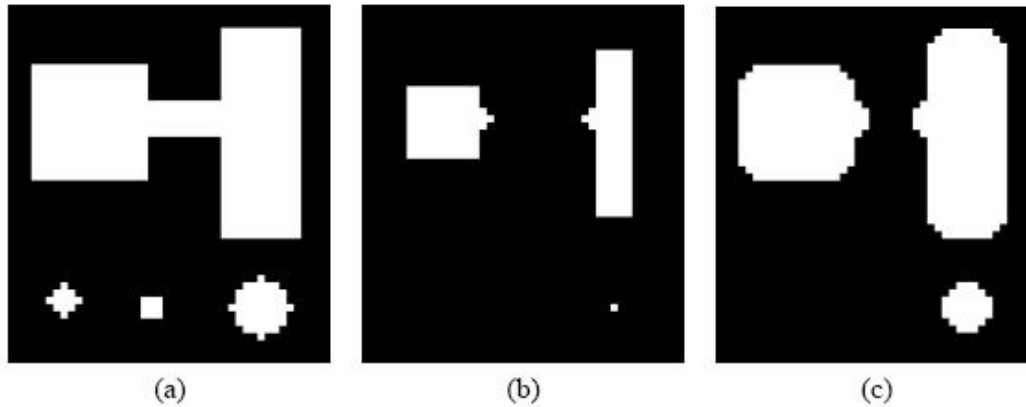


Figura 12 – Exemplo de abertura morfológica: (a) imagem original de 44 x 49 pixels; (b) resultado da erosão de (a); (c) resultado da dilatação de (b) com o mesmo elemento estruturante usado na erosão de (a).

Por sua vez, o fecho morfológico, também permite suavizar as formas de um objecto, bem como fechar pequenas aberturas dentro do mesmo e no seu contorno. Consiste numa dilatação seguida de uma erosão, e pode ser simbolizada por:

$$A \bullet B = (A \oplus B) \ominus B$$

O fecho morfológico funciona de forma oposta à abertura, ligando partes próximas ou separadas de um objecto ou região. [8]

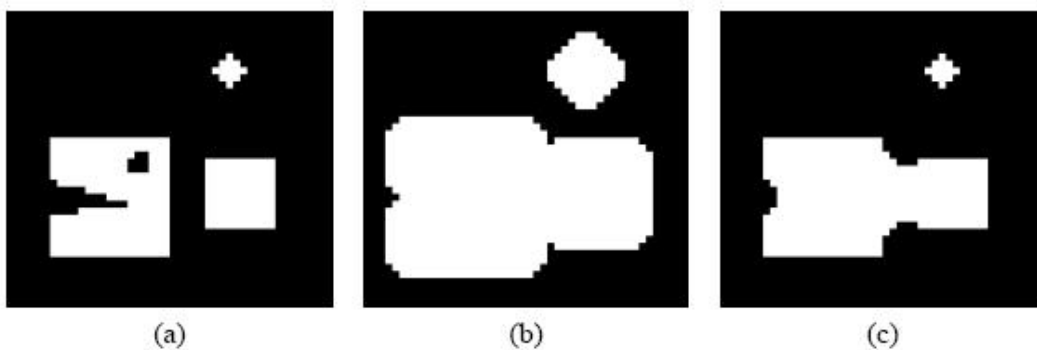


Figura 13 - Exemplo de fecho morfológico: (a) imagem original de 46 x 42 pixels; (b) resultado da dilatação de (a); (c) resultado da erosão de (b) com o mesmo elemento estruturante usado na erosão de (a).

4.3.2. Descritores de objectos

Em aplicações de visão por computador é muitas vezes necessário medir ou calcular determinadas propriedades de um ou vários objectos separados numa imagem. Após a segmentação da imagem, e a etiquetagem de cada região numa máscara binária diferente, as propriedades dos objectos podem ser calculados.

O tamanho de um objecto pode ser determinado pela simples contagem dos pixels da máscara binária do objecto. A intensidade média de um objecto, pode ser medida somando os níveis de cinzento da imagem original aos quais corresponde o valor de 1 na máscara binária, e dividindo pelo número total de pixels envolvidos no cálculo.

Há propriedades que podem descrever a forma de um objecto, como a compactação e a taxa entre a altura e a largura de um objecto. A compactação de uma região traduz o quão juntos os pixels dessa região se encontram uns dos outros. Uma estimativa simples de compactação pode ser feita por divisão do número de pixels com valor 1 na máscara binária, pelo número total de pixels contidos num rectângulo que englobe a região na sua totalidade.

A taxa entre o comprimento e a largura de uma região permite determinar o alongamento de uma região. A uma taxa de 1 corresponde uma forma circular ou quadrada. Se a taxa for menor ou maior que 1, o objecto em causa é oval ou rectangular.

[8, 12]

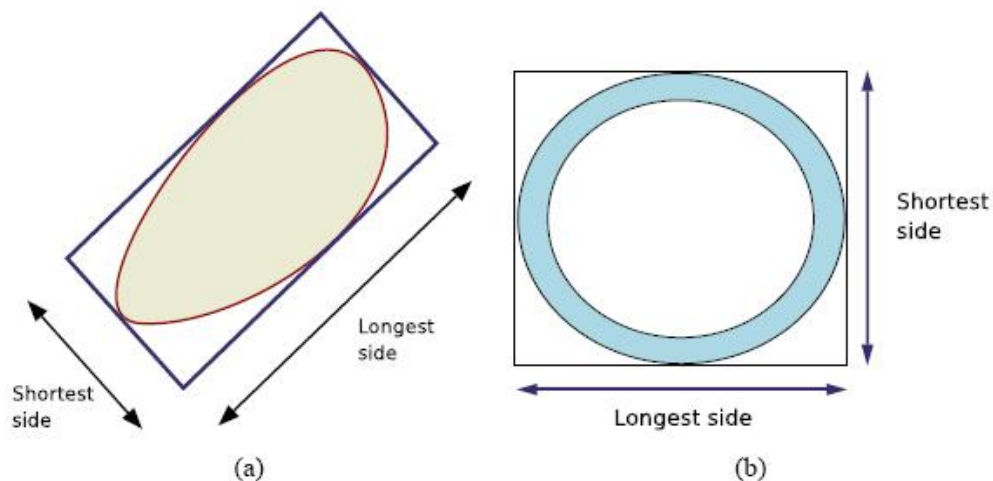


Figura 14 – Compactação dos objectos e alongamento: (a) objecto com grande compactação e alongamento; (b) objecto com uma baixa compactação e alongamento.

4.4. Classificação de objectos

Uma vez que o objectivo da visão por computador é obter informação de um objecto, o último passo da abordagem consiste usualmente na classificação dos mesmos. Neste passo os objectos encontrados na segmentação da imagem são classificados como pertencentes a alguma das classes, ou rejeitados caso isto não se verifique. A informação da classe pode ser posteriormente usada por exemplo quando decidindo se o objecto na imagem é válido ou deve ser rejeitado por defeitos na qualidade.

Há várias técnicas para classificação dos objectos. Contudo, todos os métodos de classificação necessitam de uma base de dados de conhecimento, ou seja, um conhecimento prévio do domínio do problema. Primeiro, o conhecimento deve ser provido por um especialista ou especialistas do domínio, e é frequentemente chamado de grupo de referência ou *ground truth*. O especialista no domínio classifica os objectos num conjunto de imagens. Estes dados, são posteriormente utilizados para a construção de um modelo de classificação. Primeiro, o conjunto de imagens fornecido e o seu grupo de referência são divididos em duas partes separadas: um conjunto para ser usado pelos métodos de classificação, e outro para testar os métodos desenvolvidos.

É possível desenvolver um classificador baseado em regras. Isto significa que algumas condições podem classificar os objectos candidatos, segundo diferentes classes. As regras usam informação dos objectos, medida na imagem segmentada. Apesar dos classificadores baseados em regras serem fáceis de desenvolver, podem ser difíceis de aplicar a problemas com domínios mais complexos.

A rede neuronal artificial é uma técnica de classificação largamente utilizada, uma vez que aprende a partir dos dados de entrada e não necessita que sejam definidas regras dependentes do domínio. O treino de uma rede artificial é também simples, uma vez que a rede actua como uma caixa negra. Isto quer dizer que os pesos na rede neuronal são mudados automaticamente durante o processo de treino. Há dois tipos de redes neuronais que podem ser divididas segundo o tipo de processo de aprendizagem: os mapas auto-organizativos (SOM) são redes neuronais não supervisionadas, enquanto que as redes neuronais multicamada *feedforward* necessitam de um treino supervisionado. As redes neuronais multicamada *feedforward* são treinadas dando vectores característicos de treino e as classes correspondentes à rede. Se a saída da rede não corresponder à classe respectiva, o processo de treino ajusta os pesos na rede. O processamento é continuado até que a rede corresponda correctamente às entradas dadas. Os mapas auto-organizativos não necessitam de conhecimento prévio das classes

dos objectos mas ajustam os pesos da rede somente a partir dos vectores de entradas característicos usados no treino. A característica de caixa negra das redes neuronais é também uma desvantagem, uma vez que não é possível provar formalmente que uma rede neuronal funciona correctamente em todas as situações. Outra desvantagem das redes neuronais diz respeito ao facto de que estas classificam os objectos de acordo com as classes mais próximas, mas não expressam a fiabilidade com que esses objectos pertencem a essas classes.

Os classificadores estatísticos são também utilizados em aplicações de classificação. Os classificadores estatísticos usam funções de probabilidade para classificar objectos candidatos. A ideia base de um classificador estatístico é de seleccionar a classe à qual o objecto pertence com maior probabilidade. É também possível rejeitar um objecto se ele não pertencer a nenhuma das classes consideradas com uma certa probabilidade. Há ainda outras técnicas de classificação como os classificadores estruturais, que não serão aqui referidos. [8, 12]

5. Aquisição das imagens do fundo ocular

O trabalho no âmbito deste projecto foi desenvolvido tendo em atenção as características especiais do *hardware* disponibilizado.

Para este projecto foi disponibilizado o seguinte *hardware*:

- Lâmpada de fenda *TAKAGI S70-15*;
- Adaptador *TAKAGI S10-15*;
- Câmara *3CCD SONY ExwaveHAD DXC-390*;
- PC;
- *Frame grabber Coreco Imaging PC-RGB*;

5.1. Lâmpada de fenda *TAKAGI S70-15*

Para captar imagens do fundo ocular é utilizada instrumentação especial, sendo comum em oftalmologia a utilização de biomicroscópios com câmaras acopladas.

As fotografias do fundo ocular são úteis na monitorização do progresso de várias doenças oculares revelando mudanças nas estruturas do olho.

O microscópio com lâmpada de fenda é essencialmente uma peça de equipamento simples cujo potencial se encontra mal explorado. [14]

No decorrer deste projecto foi utilizado o biomicroscópio *TAKAGI S70-15*. Este é um microscópio do tipo galileano, que consiste num sistema de iluminação e num sistema de observação binocular, o qual quando bem alinhado resulta num foco coincidente da fenda e do microscópio. [15]

O sistema de iluminação é basicamente um pequeno foco projector que emite uma imagem da abertura de uma fenda iluminada. Esta parte do sistema é flexível permitindo ao feixe de fenda assumir várias formas e tamanhos.



Figura 15 – Biomicroscópio com lâmpada de fenda.

O biomicroscópio com lâmpada de fenda *TAKAGI S-70* contém cinco filtros (transparente, absorvente de calor, ultra-violeta, absorvente de vermelho e um de azul cobalto) que podem ser mudados consoante as conveniências. É ainda possível controlar a intensidade luminosa da lâmpada. [15]

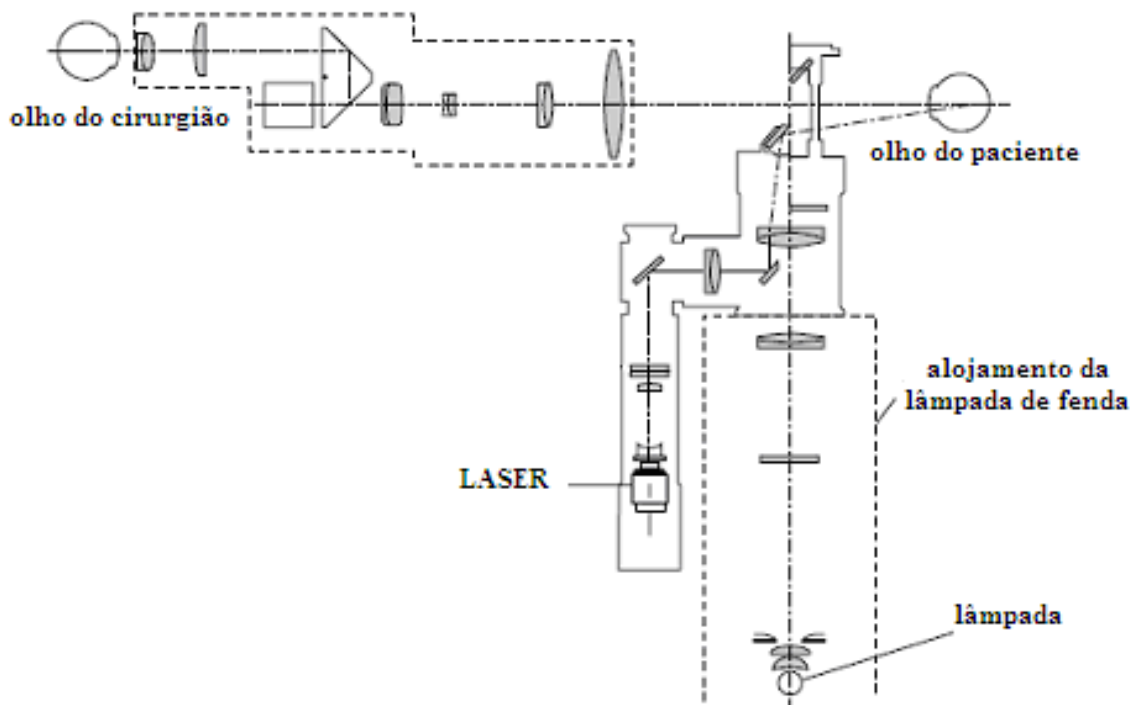


Figura 16 - Esquema típico do sistema óptico de uma lâmpada de fenda.

Por seu lado, o sistema de observação é constituído por um microscópio binocular com oculares convergentes ajustáveis. Permite ampliações de 6.3 X a 40 X variando as oculares e as objectivas do biomicroscópio.

Normalmente o sistema de iluminação e de observação estão focados no mesmo ponto (acoplado). Pode contudo ser vantajoso para determinadas ocasiões focar a fenda luminosa num ponto diferente do ponto de observação (desacoplamento). [25]

O sistema permite ainda a junção de uma fonte luminosa exterior à luz emitida pela lâmpada de fenda, podendo-se assim realizar intervenções como a fotocoagulação por disparo de laser. [23]

Na literatura é ainda referido que o laser pode ser orientado pela rotação de um espelho segundo dois eixos, e que características como a potência e duração do disparo podem também ser ajustados.

Quando utilizada esta montagem, pode-se ainda usar no olho do paciente uma lente com três espelhos. A lente é feita com 3 espelhos separados em intervalos de 120 graus em redor de uma lente central. Estes espelhos têm diferentes inclinações para que o cirurgião possa ver diferentes zonas do fundo ocular.

A lente de 3 espelhos é colocada pelo cirurgião em redor do olho, e este observa o fundo ocular através da lente central ou através de um dos espelhos. Da mesma forma, o laser pode ser reflectido por um dos espelhos antes de passar pela lente central. Um laser de baixa potência, denominado de guia laser, marca a posição para onde o disparo é efectuado caso o cirurgião pressione um pedal ou interruptor que activa o laser de maior potência. [17, 19]

Para se obter uma vista mais abrangente do interior ocular pode ainda ser necessário recorrer a métodos de dilatação da pupila.

A montagem do dispositivo garante o sistema com uma alta mobilidade, podendo-se movimentar o sistema de iluminação independentemente do sistema de observação. Assim, pode-se posicionar a fenda luminosa em regiões de maior interesse dentro da imagem captada, o que se revela de extrema utilidade no processo de disparo laser, pois permite partir de uma imagem global para zonas mais específicas.

O sistema contém várias articulações, que permitem ao conjunto uma grande combinação de posicionamentos e orientações. Estas articulações permitem a movimentação conjunta do sistema de observação e do sistema de iluminação ou apenas do sistema de iluminação.

O sistema tem ainda um manípulo ou *joystick*, com um interruptor. Este interruptor é usualmente utilizado numa montagem que permite activar o disparo laser quando pressionado.

O *joystick* está ainda associado à mobilidade do sistema.

O sistema de observação + iluminação tem 3 graus de liberdade:

- Aproximação e afastamento ao examinando por translação usando o *joystick*;
- Varrimento paralelo ao plano focal por translação usando o *joystick*;
- Varrimento por rotação do braço de suporte do sistema óptico + iluminação, segundo um eixo vertical e perpendicular à base do biomicroscópio.

O sistema de iluminação da lâmpada de fenda *TAKAGI S70-15* tem 6 graus de liberdade:

- Aproximação e afastamento ao examinando por translação usando o Joystick;
- Varrimento paralelo ao plano focal por translação usando o joystick;
- Inclinação do sistema luminoso, sendo um eixo horizontal paralelo ao plano focal do sistema de observação;
- Varrimento por rotação do sistema luminoso segundo um eixo vertical perpendicular à base do biomicroscópio;
- Rotação do feixe de fenda segundo um eixo perpendicular ao plano de observação e centrado no centro do feixe;
- Varrimento por rotação do braço de suporte do sistema óptico + iluminação, segundo um eixo vertical à base do biomicroscópio.

[15]

5.2. Câmara 3CCD SONY ExwaveHAD DXC-390

Esta câmara de vídeo a cores é constituída por 3 CCDs (Charged-coupled devices), tem elevada taxa de sinal-ruído e uma resolução de 800 linhas de TV. Apresenta ainda uma placa de DSP (*digital signal processor*) de 10 bits que permite algum tratamento de imagem.

Encontra-se ligada ao biomicroscópio através de um adaptador *CCD TV TAKAGI S10-15*, permitindo-lhe captar imagens do seu sistema de observação. [17]



Figura 17 – Adaptador CCD-TV Takagi S10-15.

5.3. PC e *frame grabber* Coreco Imaging PC-RGB

Para a realização deste projecto foi disponibilizado um PC com o sistema operativo *Microsoft Windows 2000*, processador de 1.6 GHz, 512 MB de RAM, e com *frame grabber Coreco Imaging PC-RGB*.

Esta placa recebe e transfere imagens para o computador hospedeiro, permitindo libertar recursos do CPU. Pode adquirir imagens monocromáticas ou em modo RGB. Os dados são armazenados na placa numa memória segundo *frames*. Permite ainda a sobreposição não-destrutiva de imagens como notas, janelas ou alvos, e a sincronização da resposta da placa com um sistema de interrupção. Na sua electrónica conta com um ADC *flash* de 24 bits. A taxa de entrada é ajustável até uma frequência máxima de 30 Hz. [18]



Figura 18 – *Frame grabber Coreco Imaging PC-RGB.*

6. Detecção do guia laser

Uma vez que a capacidade computacional é limitada, foi necessário introduzir um artifício que permitisse poupar alguns recursos. Desta forma pensou-se em implementar um sistema de interrupção para sinalização do disparo laser. Quando o interruptor para disparo do laser é premido, o *frame grabber* recebe um sinal a partir da porta série do computador, adquirindo uma imagem.

Assim, o módulo de detecção do guia laser foi construído partindo do princípio de que quando uma interrupção é induzida, este está sempre presente na imagem analisada.

O facto do guia laser ser um ponto luminoso de forma circular, com dimensões conhecidas foi também tido em consideração.

6.1. Princípios matemáticos

O algoritmo de detecção do guia laser faz uso de várias técnicas matemáticas.

Nesta secção pretende-se apresentar os princípios utilizados e os conceitos por trás destes.

6.1.1. Desvio padrão

O desvio padrão é uma medida da dispersão dos dados em relação à sua média.

Para um conjunto de dados $X = [X_1 X_2 \dots X_n]$, com n elementos, a média

corresponde a $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$. Este valor dá uma espécie de ponto médio dos dados. No

entanto, podemos ter dados muito diferentes entre si, e entre a própria média, assim, torna-se útil saber como os dados se encontram dispersos entre si. Esta relação é dada pelo desvio padrão através de

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}}$$

6.1.2. Variância

A variância é outra medida da dispersão dos dados, e é de facto similar ao desvio padrão, sendo calculada por

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

6.1.3. Covariância

As duas últimas medidas introduzidas são puramente unidimensionais. O desvio padrão e a variância apenas podem calcular esta relação para cada dimensão independentemente das restantes. Contudo, muitos conjuntos de dados têm mais do que uma dimensão, sendo de grande utilidade uma ferramenta que permita medir de forma similar as relações entre as várias dimensões.

Esta ferramenta é a covariância. Esta é sempre medida entre duas dimensões. Desta forma, se se calcular a covariância entre uma dimensão e ela própria, obtém-se a variância.

Por exemplo, para um conjunto de dados tridimensional x , y e z , pode-se obter a relação entre x e y , x e z , e entre y e z .

Sabendo que a variância pode também ser representada por

$$s^2 = \text{var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X}).(X_i - \bar{X})}{(n-1)}$$

Então para a covariância temos

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X}).(Y_i - \bar{Y})}{(n-1)}$$

, com $Y = [Y_1 Y_2 \dots Y_n]$.

Da informação dada pela covariância, aquela considerada mais importante diz respeito ao sinal do resultado. Se o resultado for positivo, temos uma indicação que ambas as dimensões aumentam no mesmo sentido, se o resultado for negativo, uma das dimensões aumenta e a outra diminui. Caso a covariância seja nula, então as variáveis são independentes e não estão relacionadas entre si.

A importância da covariância está no facto de poder ser aplicada a grandes conjuntos de dados e variáveis, que deixam de ser visualizáveis a partir da 3ª dimensão.

6.1.4. A matriz de covariância

Relembrando que a covariância é sempre calculada entre matrizes de duas dimensões, se tivermos um conjunto de mais do que duas dimensões, então para um conjunto de dados de dimensão n temos $\frac{n!}{(n-2)!*2}$ valores diferentes de covariância.

Uma maneira prática de se obter todos os valores possíveis de covariância entre as diferentes dimensões, é matricialmente. Desta forma, para n dimensões temos

$$C^{n*n} = (c_{i,j}, c_{i,j} = \text{cov}(\text{Dim}_i, \text{Dim}_j))$$

, onde C^{n*n} é uma matriz quadrada com n linhas e n colunas, e Dim_x é a x -ésima dimensão. Traduzindo a equação anterior para uma forma mais agradável e compreensível tem-se

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix}$$

É de salientar que a diagonal principal da tabela contém as variâncias para cada uma das dimensões, e que a matriz é simétrica em relação a essa diagonal já que $\text{cov}(x, y) = \text{cov}(y, x)$.

6.1.5. Valores próprios

Como sabido, duas matrizes podem ser multiplicadas desde que tenham dimensões idênticas. Quando se multiplica um vector por uma matriz à sua esquerda, como resultado obtém-se um novo vector que é transformado em relação à sua posição original. Assim, a matriz pode ser vista como uma matriz de transformação. É a natureza desta transformação que os vectores próprios revelam.

Apenas as matrizes quadradas podem ter vectores próprios, e nem todas têm. Contudo, para uma matriz de dimensões $n \times n$, com vectores próprios, é possível calcular n vectores próprios.

Os vectores próprios podem ainda ser multiplicados por constantes, sem que se altere o seu significado, uma vez que ao se fazer isto, apenas se está a redimensionar o vector, tornando-o mais ou menos longo, e preservando-se a sua direcção. Por fim, os vectores próprios são todos ortogonais, formando ângulos rectos entre si. Assim, é possível representar os dados com base nestes vectores, em vez de se usar coordenadas como x, y, z , etc.

[19, 20, 25]

6.1.6. Análise em componentes principais

A análise em componentes principais, mais conhecida como PCA (abreviatura do anglo saxónico *Principal Component Analysis*), é um dos métodos de álgebra linear mais utilizados. Devido à sua simplicidade e natureza não-paramétrica, o método de PCA tem aplicação na análise de dados em áreas tão distintas como as neurociências e a computação gráfica.

Este método de natureza estatística consiste numa transformação linear que altera os dados para um novo sistema de coordenadas por forma a que a maior variância devido a qualquer projecção dos dados coincida com a primeira coordenada (chamada primeira componente principal), a segunda maior variância coincida com a segunda coordenada, e por aí fora. Estatisticamente, é um método que simplifica um conjunto de dados, por redução da sua dimensionalidade, preservando as características dos dados que mais contribuem para a sua variância, através das componentes principais de ordens

inferiores e ignorando as de ordens maiores. Estas componentes de ordens inferiores muitas vezes contêm os aspectos dos dados considerados “mais importantes”. [8]

Esta técnica também conhecida como transformada de *Karhunen-Loève* ou transformada de *Hotelling*, é o melhor transformador linear quando se pretende manter o subespaço com maior variância. No entanto, esta característica torna o método dependente de uma grande carga computacional, quando comparado com outras transformações lineares. É ainda de referir que os seus vectores de base dependem do conjunto de dados. [22]

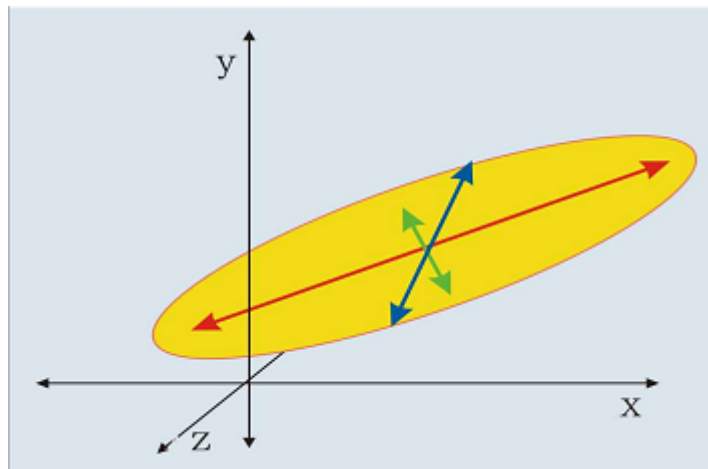


Figura 19 – Representação dos eixos de maior variância para um conjunto de pontos distribuídos segundo um elipsóide num espaço tridimensional.

6.1.6.1. Aplicação do método à construção de modelos de imagens

A aplicação desta técnica ao conjunto de dados fornecidos por uma imagem envolve os passos que se seguem.

a) Formatar os dados

Esta técnica é aplicável a um conjunto de dados $X = [x_1 \ x_2 \ x_3 \ \dots \ x_n]^T$, em que $x_1, x_2, x_3 \dots x_n$ correspondem a observações distintas da variável X .

Quando aplicada a imagens, devemos considerar cada uma das imagens como um análogo de X , e os seus pixels como sendo observações.

Normalmente, uma imagem corresponde a uma matriz de $N \times M$ pixels, a qual pode ser expressa por um vector de dimensões $N \times M$

$$X = [x_1 \ x_2 \ x_3 \ \dots \ x_{N \times M}]$$

, resultante da colocação sucessiva das linhas da imagem frente a frente, da primeira para a última. Os valores do vector corresponderão então aos valores dos pixels da imagem segundo uma escala de intensidade. Desta forma para L imagens obteremos uma matriz do género

$$\text{MatrizImagens} = [X_1 \ X_2 \ X_3 \ \dots \ X_L]^T$$

b) Subtrair a média

Para que a PCA funcione correctamente é necessário subtrair a média dos dados segundo cada dimensão, ou seja, a cada valor de X_l é subtraída a média de todos os valores de X_l . Isto deve ser feito para as L colunas da *MatrizImagem* por forma a que fiquem centradas no valor 0.

c) Calcular a matriz de covariância

Este segue o cálculo já atrás referido.

d) Calcular os valores e vectores próprios da matriz de covariância

Os vectores próprios contêm informação acerca dos padrões nos dados, permitindo extrair linhas que caracterizem os dados.

e) Escolher componentes e formar um vector característico

É nesta fase que ocorre a compressão dos dados e a redução da sua dimensionalidade. O vector próprio com o maior valor próprio, é designado de componente principal do conjunto de dados, e descreve a relação mais importante entre as várias dimensões dos dados.

Uma vez determinados os vectores próprios, estes são ordenados por ordem decrescente, e conseqüentemente, de significância. As componentes de menor ordem podem ser descartadas, reduzindo-se assim a dimensionalidade dos dados. Esta rejeição implica uma perda de informação, que contudo, pode não ser significativa desde que os valores próprios sejam pequenos.

Assim, segue-se a construção de uma matriz com os vectores próprios que pretendemos manter, o chamado vector característico

$$VectorCaracterístico = (vectorprop_1 \ vectorprop_2 \dots \ vectorprop_n)$$

f) Obter o novo conjunto de dados

Esta é a última etapa da PCA, e também a mais simples, e resume-se ao cálculo

$$MatrizFinal = VectorCaracterístico^T * MatrizImagens$$

A *MatrizFinal* contém os dados finais que procuramos distribuídos ao longo das suas colunas.

Esta matriz dá os dados iniciais apenas em termo dos vectores escolhidos. Os dados iniciais representados por x e y , passam a ser representados em termos dos vectores próprios mantidos na transformação.

Basicamente transformam-se os dados para que passem a ser expressos de acordo com os padrões internos, onde estes padrões são linhas que descrevem de forma aproximada a relação entre os dados. A utilidade deste procedimento está ligada à classificação dos pontos como uma combinação da contribuição de cada uma dessas linhas. Agora, os valores dos dados dizem-nos exactamente onde (ex: acima/abaixo) os pontos se situam em relação à linha de tendência.

6.1.7. Decomposição em valores singulares

Segue-se a introdução a um técnica muito utilizada na computação e análise de matrizes, a decomposição em valores singulares. Por uma questão de simplificação, será

adoptada a sigla SVD (do anglo-saxónico *Singular Value Decomposition*) para a sua referência.

Para além de ser uma técnica robusta aos erros numéricos de computação, a SVD permite expor a estrutura geométrica de uma matriz, que é um aspecto importante para muitos cálculos matriciais. Uma matriz pode ser descrita como a transformação de um espaço vectorial para outro. Os componentes da SVD quantificam as alterações resultantes entre a geometria implícita nos espaços vectoriais.

Dois vectores x e y são considerados ortogonais se $x^T y = 0$.

Sendo A uma matriz quadrada tal que as suas colunas sejam vectores ortogonais de comprimento 1, ou seja $x^T x = 1$, então A é uma matriz ortogonal e $A^T A = I$, é a matriz identidade.

A SVD de uma matriz A de dimensões $M \times N$ é qualquer factorização da forma

$$A = U \Sigma V^T$$

, onde U é uma matriz ortogonal de $M \times M$ elementos, V é uma matriz ortogonal de $N \times N$ elementos e Σ é uma matriz diagonal de $M \times N$ elementos, com o elemento $s_{ij} = 0$ se $i \neq j$ e $s_{ii} = s_i \geq 0$. Por conseguinte, pode ser demonstrado que existem matrizes U e V tais que $s_1 \geq s_2 \geq \dots \geq s_N \geq 0$. As quantidades s_i são denominadas de valores singulares de A , e as colunas de U e V são chamadas de vectores singulares esquerdo e direito, respectivamente.

Dependendo dos valores de M e N , a decomposição e descrita pode ser dada de uma forma mais compacta. Em particular, se $M > N$, as últimas $M-N$ colunas de U e as últimas $M-N$ linhas de Σ podem ser omitidas. Nesta abordagem, referida na linguagem anglo-saxónica como decomposição *economy-size*, a matriz $U \in \mathfrak{R}^{M \times N}$ será uma matriz ortogonal em coluna e $\Sigma \in \mathfrak{R}^{N \times N}$ uma matriz diagonal.

A partir da decomposição em valores singulares, pode-se calcular muitas propriedades da matriz original. O espaço nulo da matriz A é o conjunto de x para o qual $Ax = 0$, e a distância de A é o conjunto b para o qual $Ax = b$ como solução de x . Sejam u_j e v_j as colunas de U e V respectivamente, então a decomposição de $A = U \Sigma V^T$ pode ser escrita como

$$A v_j = s_j u_j, j = 1, 2, \dots, N.$$

Se $s_j = 0$, então $Av_j=0$ e v_j faz parte do espaço nulo de A enquanto que se $s_j \neq 0$, então, u_j está ao alcance de A . Consequentemente, podemos construir bases para vários subespaços vectoriais definidos por A . Um conjunto de vectores v_1, v_2, \dots, v_k num espaço vectorial V forma uma base de V se todos os vectores x em V poderem ser expressos como combinações lineares dos mesmos de exactamente uma maneira. Sendo V_0 o conjunto de colunas v_j para as quais $s_j = 0$, e V_1 as restantes colunas v_j , de forma similar, seja U_1 o conjunto de colunas u_j para as quais $s_j \neq 0$, e sendo U_0 as restantes colunas u_j , incluindo aquelas com $j > n$, então se k é o número de valores singulares não-nulos, há k colunas em V_0 , $N-k$ colunas em V_1 e U_1 , e $M-N+k$ colunas em U_0 . Cada um destes conjuntos forma uma base para os subespaços vectoriais de A .

1. V_0 é uma base ortogonal do espaço nulo (A);
2. V_1 é uma base ortogonal para o complemento ortogonal do espaço nulo (A);
3. U_1 é uma base ortogonal para a distância (A);
4. U_0 é uma base ortogonal para o complemento ortogonal da distância (A).

[20]

6.1.7.1. SVD e normas de matrizes

Um conceito familiar de comprimento em duas dimensões é a distância euclideana da origem ao ponto especificado pelas coordenadas do vector $\{x_1, x_2\}$. Esta distância é calculada pela fórmula $(x_1^2 + x_2^2)^{\frac{1}{2}}$.

No caso geral de N dimensões, o comprimento ou norma de um vector x é definido por

$$\|x\| = (x_1^2 + x_2^2 + \dots + x_N^2)^{\frac{1}{2}} = (x^T x)^{\frac{1}{2}}.$$

Quando um vector é multiplicado por uma matriz A , o comprimento da matriz resultante Ax muda de acordo com a matriz A . Se A é ortogonal, o comprimento é preservado. De outra forma, a quantidade $\frac{\|Ax\|}{\|x\|}$ mede o quanto a matriz foi “esticada”.

Assim, calcular a norma da matriz significa intuitivamente encontrar o factor máximo de distensão da matriz. Usando a SVD, este cálculo é simplificado.

A norma euclideana de uma matriz, muitas vezes referida como a norma L_2 é definida da seguinte forma. Seja x um vector de dimensão N , e A uma matriz de dimensões $M \times N$, então

$$\|A\|_E = \max_{\|x\|=1} \left\{ \frac{\|Ax\|}{\|x\|} \right\}$$

Uma alternativa para a norma de A é a norma de *Forbenius*, que é a norma euclidiana de um vector construído por empilhamento das colunas de A num vector $M \times N$. A norma de *Forbenius* é então

$$\|A\|_F = \left(\sum \sum |a_{ij}|^2 \right)^{\frac{1}{2}}.$$

Dada a SVD esta norma é facilmente calculada.

Seja $U\Sigma V^T$ a SVD da matriz A de dimensões $M \times N$, onde $\{s_1, s_2, \dots, s_k\}$, com $k \leq N$ são os valores singulares não-nulos em Σ . Então,

$$\|A\|_E = s_1$$

$$\|A\|_F = \left(\sum_{l=1}^k s_l^2 \right)^{\frac{1}{2}}$$

Voltando a noções já apresentadas, foi mencionado que a multiplicação de um vector x por uma matriz A , efectivamente “estica” o vector. Esta interpretação geométrica é melhor vista em termos dos valores singulares de A . O conjunto de vectores x de N elementos para os quais $\|x\|=1$ define um círculo unitário. A multiplicação destes vectores pela matriz $M \times N$ resulta num conjunto de vectores M -dimensionais $b = Ax$ de comprimentos variáveis. Geometricamente este conjunto define um elipsóide k -dimensional incluído num espaço M -dimensional, onde k é o número de valores singulares não-nulos. O comprimento dos eixos do elipsóide corresponde aos valores singulares de A , sendo no caso geral, o eixo maior e o eixo

menor, dados por s_{max} e s_{min} , respectivamente. Intuitivamente, os valores singulares de uma matriz descrevem a extensão da distorção provocada pela multiplicação do vector original pela matriz. A magnitude dos valores singulares pode ser usada para salientar as dimensões do vector que são mais afectadas, e de certa forma mais importantes. [22]

6.1.7.2. SVD e ordem matricial

A noção de ordem é fundamental para aplicação de muitos teoremas que envolvem álgebra linear. No entanto, se a matriz é deficitária (ou quase) de ordem então das pequenas perturbações dos valores da matriz (erros de arredondamento ou dados difusos) resulta numa matriz com ordem total. Assim, a determinação da ordem de uma matriz não é uma tarefa trivial. A SVD fornece uma definição prática de ordenação, permitindo também quantificar a noção de quase deficiência de ordem.

A definição familiar de ordem corresponde ao número de colunas de uma matriz linearmente independentes.

Seja A uma matriz com SVD $U\Sigma V^T$. Uma vez que a multiplicação por matrizes ortogonais preserva a independência linear, a ordem de A é precisamente a ordem da matriz diagonal Σ , ou equivalentemente, o número de valores singulares não-nulos. Se A é aproximadamente deficiente de ordem, ou seja, singular, então os valores singulares são pequenos. Para além do mais, supondo $\text{ordem}(A) = m$ e querendo aproximar A por uma matriz B de menor ordem k , então pode-se usar os valores singulares de A para calcular a matriz com a melhor aproximação, e determinar se a aproximação é única. Sejam s_i as entradas diagonais de Σ , e sendo u_i e v_i os vectores coluna de U e V respectivamente, então

$$\min_{\text{RANK}(B)=k} \|A - B\|_E = \|A - A_k\|_E = s_{k+1}$$

, onde $A_k = \sum_{i=1}^k s_i u_i v_i^T$. A solução $B = A_k$ é única quando $s_{k+1} < s_k$.

Vê-se assim que a SVD de A produz uma sequência de aproximações a A de sucessivas ordens $A_i = U\Sigma_i V^T$, onde Σ_i é a versão de ordem i de Σ obtida por atribuição do valor 0 ao último $m-1$ valor singular T_0 . Também A_i é a melhor i -ésima aproximação a A segundo a distância euclidiana.

O uso da SVD na aproximação matricial tem várias vantagens práticas. Primeiro, aplicações em que se encontram erros de arredondamento ou dados difusos usam tipicamente a ordem efectiva da matriz, ou seja, o número de valores singulares maiores do que ϵ , onde ϵ reflecte a exactidão dos dados. Assim, as decisões são tomadas apenas com base na negligibilidade de alguns valores singulares, em vez de vectores ou conjuntos de vectores. Em segundo, o armazenamento de uma aproximação a uma matriz resulta muitas vezes em poupança de memória. Note-se que uma matriz A pode ser definida como

$$A = s_1 u_1 v_1^T + s_2 u_2 v_2^T + \dots + s_m u_m v_m^T$$

Cada produto externo $u_1 v_1^T$ é uma matriz simples de ordem 1, e pode ser armazenada em $M+N$ membros, por oposição às dimensões de $M \times N$ da mesma matriz. Em adição, a multiplicação de $u_1 v_1^T$ com um vector x requer $M+N$ operações, em vez de $M \times N$. [22, 24, 30, 21]

6.1.8. PCA e SVD

Uma vez obtida a matriz de covariância, durante o processo de PCA pode-se então seguir para o cálculo dos valores próprios e dos vectores próprios.

Este passo é realizado recorrendo à SVD.

Tendo formas de cálculo semelhantes, torna-se óbvio que a SVD e a PCA estão intimamente relacionadas.

Para uma matriz X de dimensões $M \times N$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix}$$

, onde cada linha de X corresponde a todas as medições segundo um determinado tipo x_i , e cada coluna de x contém as diferentes medições de uma específica dimensão.

Chega-se assim a uma definição de matriz de covariância C_x :

$$C_X = \frac{1}{N-1} XX^T$$

, onde $\frac{1}{N-1}$ é um factor de normalização.

Voltando à matriz $M \times N$ original X , e definindo uma nova matriz Y de dimensões $N \times M$

$$Y = \frac{1}{\sqrt{N-1}} X^T$$

, em que cada coluna de Y tem média zero. A definição de Y torna-se clara pela análise de $Y^T Y$.

$$\begin{aligned} Y^T Y &= \left(\frac{1}{\sqrt{N-1}} X^T \right)^T \left(\frac{1}{\sqrt{N-1}} X^T \right) = \\ &= \frac{1}{N-1} X^{TT} X^T = \\ &= \frac{1}{N-1} XX^T \end{aligned}$$

$$Y^T Y = C_X$$

A construção de $Y^T Y$ iguala a matriz de covariância de X . Sabe-se já de trás que as componentes principais de X são os vectores próprios de C_X .

Sendo que

$$A = U \Sigma V^T$$

, multiplicando A com A^T pela esquerda vem:

$$\begin{aligned} A^T A &= (U \Sigma V^T)^T U \Sigma V^T = \\ &= (V \Sigma U^T) U \Sigma V^T \end{aligned}$$

, como $U^T U = I$, então

$$A^T A = V \Sigma^2 V^T$$

Assim, calculando a SVD de Y , obtém-se a matriz V cujas colunas contêm os vectores próprios de $Y^T Y = C_X$.

Sabendo-se que as colunas de V são as componentes principais de X , torna-se possível calcular a PCA de uma forma mais simples poupando tempo de cálculo e recursos computacionais.

[21, 22]

6.1.9. Transformada de *Hough*

A transformada de *Hough* é uma técnica que pode ser usada para isolar características de uma forma particular numa imagem. Uma vez que é necessário que as características desejadas sejam especificadas de uma forma paramétrica, a transformada de *Hough* clássica é usualmente mais usada na detecção de curvas regulares como linhas, círculos e elipses.

Esta transformada baseia-se no princípio de que por qualquer ponto podem passar infinitas linhas, cada uma com uma orientação diferente. Com esta transformada pretende-se saber qual destas linhas passa mais vezes em determinadas características presentes na imagem, ou seja, determinar qual das linhas tem maior correspondência com uma estrutura na imagem.

Num ponto (x_i, y_i) passam infinitas linhas que seguem a equação $y_i = m x_i + b$. Para um segundo ponto (x_j, y_j) , vai-se então obter uma linha que vai interceptar a linha associada com (x_i, y_i) em (m', b') , onde m' é o declive e b' a intercepção da linha contendo (x_i, y_i) e (x_j, y_j) no plano xy .

Na transformada de *Hough* subdivide-se o espaço dos parâmetros segundo células acumuladoras. A célula nas coordenadas (i, j) com o valor do acumulador $A(i, j)$, corresponde ao quadrado associado com as coordenadas (m_i, b_i) do espaço dos parâmetros.

Para cada ponto (x_k, y_k) no plano da imagem, resolve-se para cada subdivisão o valor de b segundo os valores do eixo do parâmetro m usando a equação $b = -x_k m + y_k$.

Os bs resultantes são depois aproximados ao valor mais próximo permitido no eixo dos bs . Se um valor de mp tem como solução bq , então $A(p,q) = A(p,q)+1$. No final deste processo vão-se obter valores que correspondem ao número de pontos no plano xy que estão na linha $y = m_i x + b_j$. A eficácia da colinearidade dos pontos é definida pela quantidade de subdivisões no plano ab .

Ao usar-se a equação $y = mx + b$ na representação de uma linha, vai-se ter um problema quando esta se aproximar da posição vertical, pois m aproxima-se do infinito. Desta forma, uma maneira de se ultrapassar esta adversidade é usando um sistema de coordenadas esféricas para descrever os pontos, recorrendo então à representação

$$r = x \cdot \cos\theta + y \cdot \sin\theta.$$

Com este sistema de coordenadas o processo de construção de tabelas de acumuladores é semelhante ao processo anteriormente descrito, sendo que a grande diferença consiste nas representações obtidas. Para esta situação vai-se obter um conjunto de curvas sinusoidais no plano θr , em vez dos segmentos de recta obtidos no plano mb . Tal como no caso anterior, para Q pontos colineares segundo a equação $r_i = y_i \cdot \cos\theta_j + y_i \cdot \sin\theta_j$, vai-se obter Q curvas sinusoidais que se interceptam no ponto (r_i, θ_j) no espaço dos parâmetros. Incrementando θ e resolvendo para r , obtem-se q entradas para o acumulador $A(i, j)$ associadas com a célula determinada por (r_i, θ_j) .

O ângulo θ vai de -90° a $+90^\circ$ medidos segundo o eixo dos xx .

Até agora foi apenas referida a aplicação a linhas, no entanto esta transformada é aplicável a todas as formas que sigam equações do género $g(v, c) = 0$, onde v é um vector de coordenadas e c é um vector de coeficientes. Neste grupo incluem-se os círculos e as elipses.

Como sabido, a um círculo corresponde à equação

$$(x-c_1)^2 + (y-c_2)^2 = r^2.$$

Na aplicação matemática da transformada de *Hough* a círculos são usados 3 parâmetros, que originam acumuladores 3-D do tipo $A(i, j, k)$. O processo consiste em incrementar c_1 e c_2 para satisfazer r , e actualizar o acumulador. Induz-se então que a

complexidade da transformada de *Hough* varia com o número de coordenadas e coeficientes numa dada representação de uma função.

A abordagem computacional baseada na transformada de *Hough* segue então os seguintes procedimentos:

1. Cálculo das arestas de uma imagem e sua conversão em elementos binários, por aplicação de um limiar às arestas detectadas, ou seja, constrói-se uma nova imagem em que aos pixels das arestas detectadas é atribuído o valor de 1 caso tenham um valor acima de um limiar definido e 0 caso contrário;
2. Especificação das subdivisões no plano $r\theta$;
3. Projecção da forma geométrica procurada na célula acumuladora;
4. Verificação das células acumuladoras com maiores contagens;
5. Verificação da relação entre pixels numa célula escolhida (em especial continuidade). O conceito de continuidade está neste caso relacionado com a distância entre pixels desconectados, mas que podem ser significativos para a detecção caso a falha entre pixels esteja dentro de valores aceitáveis.

[8, 26, 27]

6.2. O algoritmo de detecção do guia laser

O algoritmo de detecção do guia laser foi construído com base na transformada de *Hough*.

A capacidade para reconhecer formas parcialmente deformadas ou ocluídas, a robustez na presença de outras estruturas na imagem e a tolerância ao ruído, são características que fazem deste num bom método de reconhecimento, tendo sido preteridos outros métodos como por exemplo o *feature matching*. No entanto, este é um método que requer uma computação relativamente extensa e alguma memória para armazenamento de dados, tornando-se assim consideravelmente lento.

Para se aumentar a velocidade do processamento, pensou-se em recorrer à redução da dimensão da imagem. No entanto, o facto do guia laser ter dimensões muito reduzidas, podendo desaparecer com o redimensionamento, afastou esta possibilidade.

O módulo de detecção do guia laser foi desenvolvido na linguagem de programação *Matlab*, tendo sido aproveitadas algumas funções já existentes no *Matlab 7.2 R2006a*. Este módulo foi subdividido segundo vários sub-módulos, para se facilitar

a sua construção e teste. Esta estruturação teve também como objectivo tornar o código facilmente legível, encontrando-se largamente comentado.

Segue-se um esquema para exposição de todos os módulos construídos e das funções utilizadas.

O módulo principal foi denominado de *laser_guide*, ao qual estão ligados vários outros módulos, que serão descritos mais adiante.

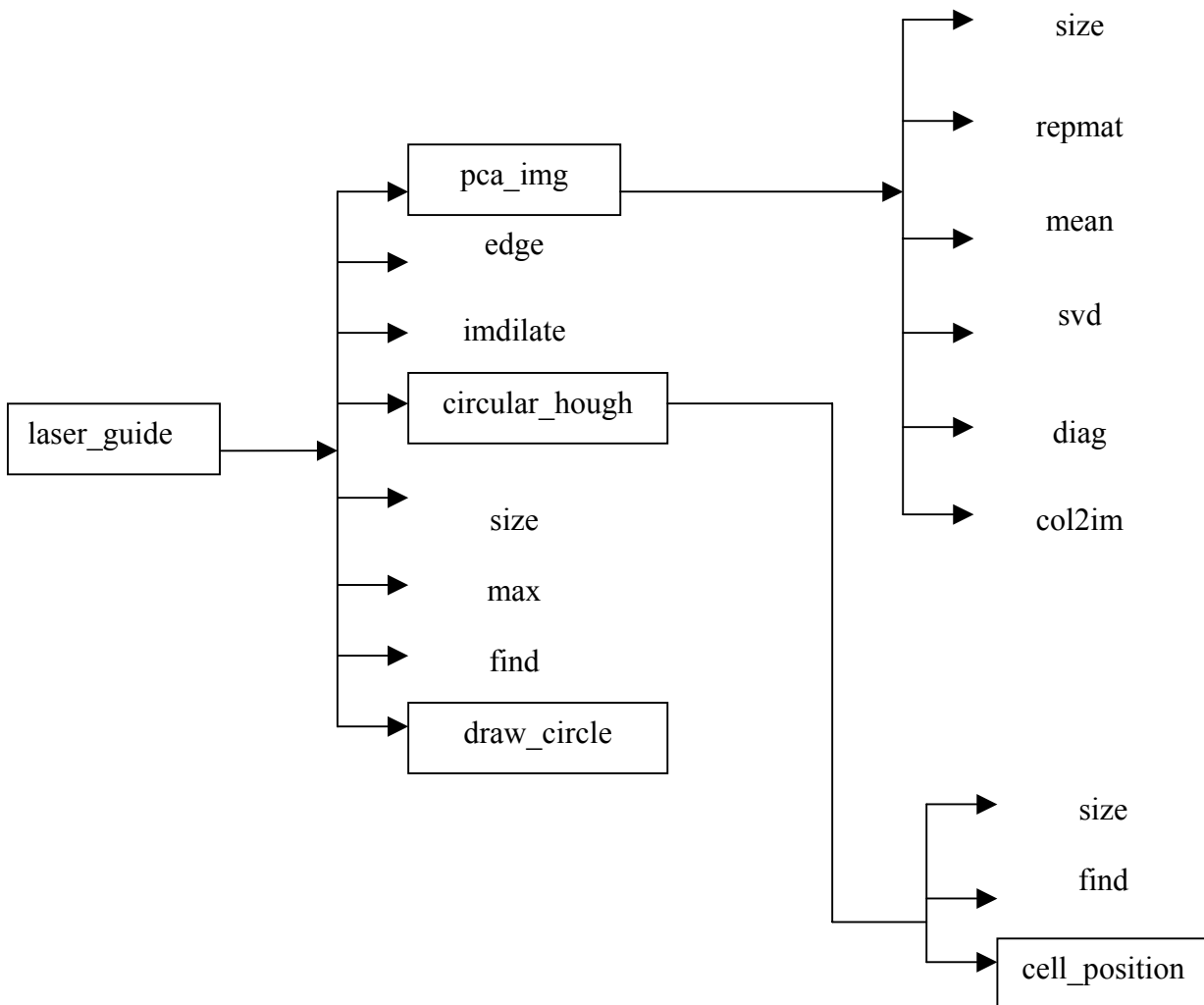


Diagrama 1 - Os vários módulos produzidos e funções utilizadas. As funções cujo nome não se encontra dentro de uma caixa, são funções já existentes no *Matlab 7.2 R2006a*, e descritas na ajuda do *software*.

A descrição do algoritmo desenvolvido, é acompanhada de imagens ilustrativas do processamento executado em cada passo. Uma vez que não foram disponibilizadas imagens com o guia laser, o algoritmo foi desenvolvido com base em pressupostos teóricos. Usando um programa de edição de imagem foi marcado um círculo vermelho em imagens do fundo ocular, de forma a simular o guia laser. Aquando da criação do círculo teve-se ainda em atenção o tamanho do guia laser em relação aos restantes elementos do fundo ocular presentes na imagem.

6.2.1. Aumento do contraste do guia laser

A imagem a cores, previamente segmentada, é sujeita a uma análise em componentes principais para que se obtenha uma imagem onde o ponto luminoso do guia laser tenha mais contraste e seja mais visível.

Esta análise em componentes principais é realizada pelo algoritmo definido no módulo *pca_img*.



Figura 20 – Imagem do fundo ocular segmentada, com um modelo do guia laser marcado. Esta imagem foi segmentada com o algoritmo de segmentação descrito mais à frente, e vai ser usada ao longo desta secção para exemplificar o resultado dos vários passos do processamento executado pelo presente módulo em discussão. Pretende-se simular o guia laser através do círculo vermelho, tendo-se em consideração a sua cor e o seu tamanho relativamente ao disco óptico.

Partindo da imagem a cores RGB, previamente segmentada, começa-se por construir uma matriz a partir das matrizes dos três canais constituintes da nossa imagem inicial. Para a análise que se segue, considera-se cada canal de cor da imagem como variáveis independentes de um espaço tridimensional. Desta forma, torna-se necessário construir uma nova matriz, em que cada coluna contenha a informação de cada canal de cor e as linhas contenham os casos particulares, ou seja, o valor do respectivo canal para cada pixel.

Cada uma das matrizes correspondentes a R, G e B, é transformada num vector linha, por colocação das sucessivas linhas da matriz frente a frente.

Para uma matriz $M \times N$, com L_1, L_2, \dots, L_M , as linhas dessa matriz, sendo que cada linha tem N elementos, o vector linha corresponde a

$$V_L = [L_1, L_2, \dots, L_M]$$

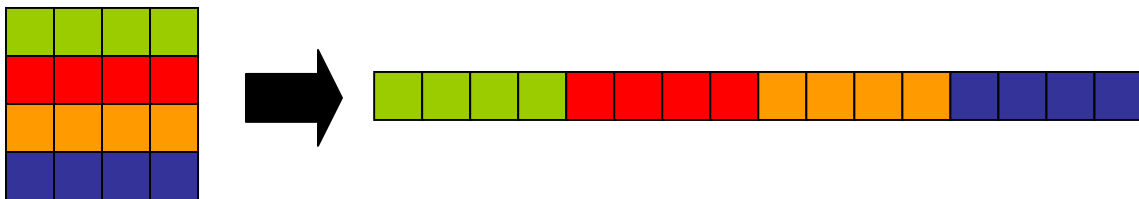


Figura 21 – Modelo que demonstra o primeiro passo da formatação das matrizes de cada canal de cor: colocação das linhas da imagem frente a frente.

Os três vectores linha obtidos são transpostos, transformando-se em colunas que e são inseridas numa nova matriz, que desta forma passa a conter a intensidade de cada canal de cor segundo as colunas.

Para cada uma das colunas desta nova matriz calcula-se a sua média, e subtrai-se esse valor a cada pixel cada respectiva coluna, para que as colunas fiquem centradas em 0.

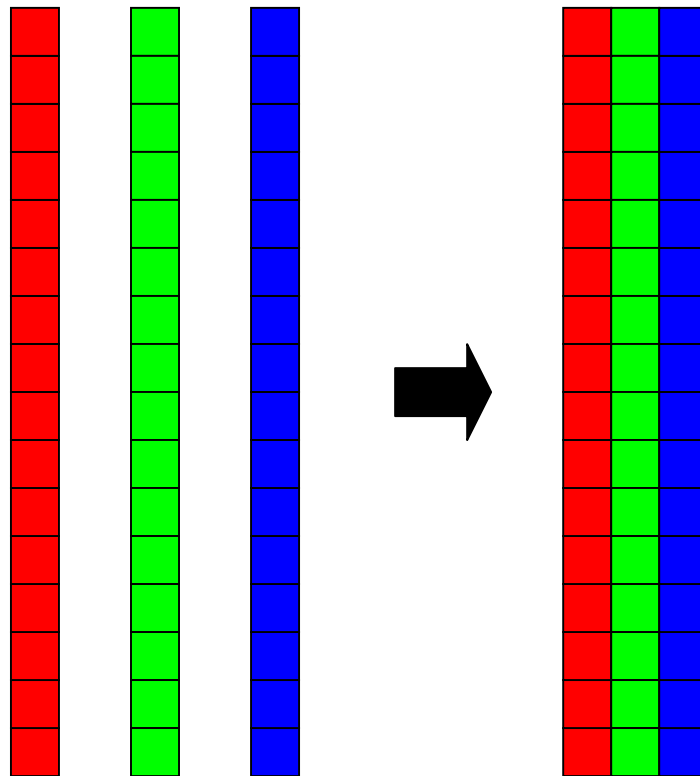


Figura 22 – Modelo descritivo da construção da matriz com os dados dos três canais de cor segundo as colunas.

Nesta fase da PCA seguir-se-ia o cálculo da matriz de covariância desta última matriz obtida. No entanto, e uma vez que não se tem interesse em prolongar cálculos desnecessários, segue-se a implementação descrita na secção 6.1.8.

Desta forma, o próximo passo consiste na multiplicação dos elementos da matriz pela constante $\frac{1}{N-1}$.

A matriz resultante é sujeita à decomposição em valores singulares descrita na secção 6.1.7, resultando as matrizes U , Σ e V .

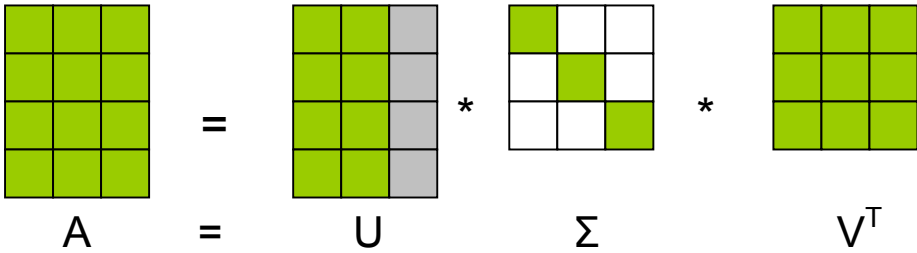


Figura 23 – SVD: esquemática da matriz inicial e das resultantes da sua decomposição.

Sabendo que a matriz V contém os vectores próprios da matriz de covariância segundo as suas colunas, é seleccionada a coluna mais à esquerda da matriz V , que correspondente à componente principal.

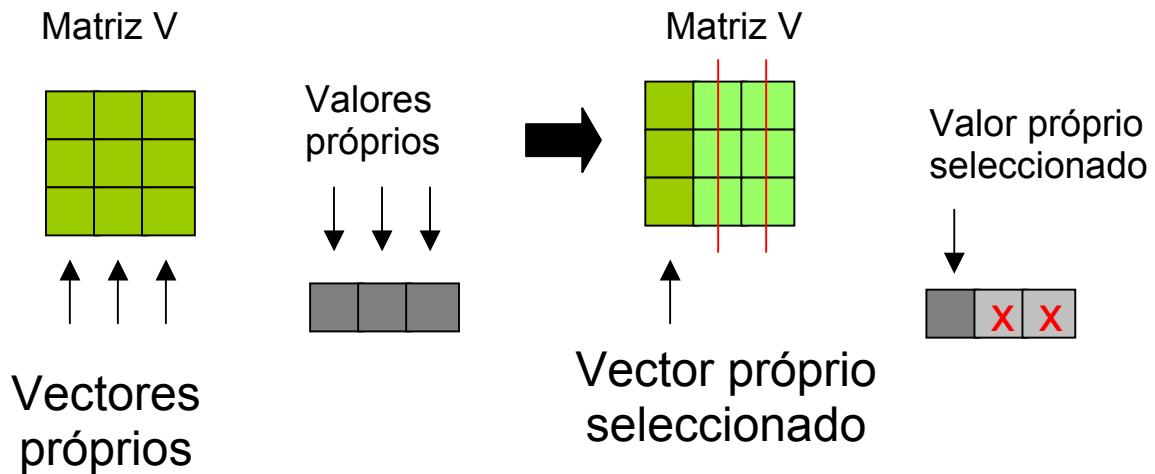


Figura 24 – Ilustração do processo de escolha da componente principal a partir da matriz V .

De seguida calcula-se a multiplicação matricial entre a transposta do vector coluna de V escolhido e a transposta da matriz com os dados ajustados, originando-se um novo vector linha.

Transpondo o vector linha, resulta um vector coluna, que através da função *col2im*, é reestruturado segundo uma matriz com o número de linhas e colunas da imagem original.

De todo este processo, obtém-se uma redução dimensional de três matrizes iniciais, correspondentes cada uma delas a um espaço de cor diferente, para apenas uma matriz, pela projecção dos dados segundo um novo sistema de coordenadas definido pela componente principal e pelo eixo ortogonal a esta.

Na prática, obtém-se uma imagem em níveis de cinzento, onde o contraste do ponto luminoso do guia laser é aumentado.

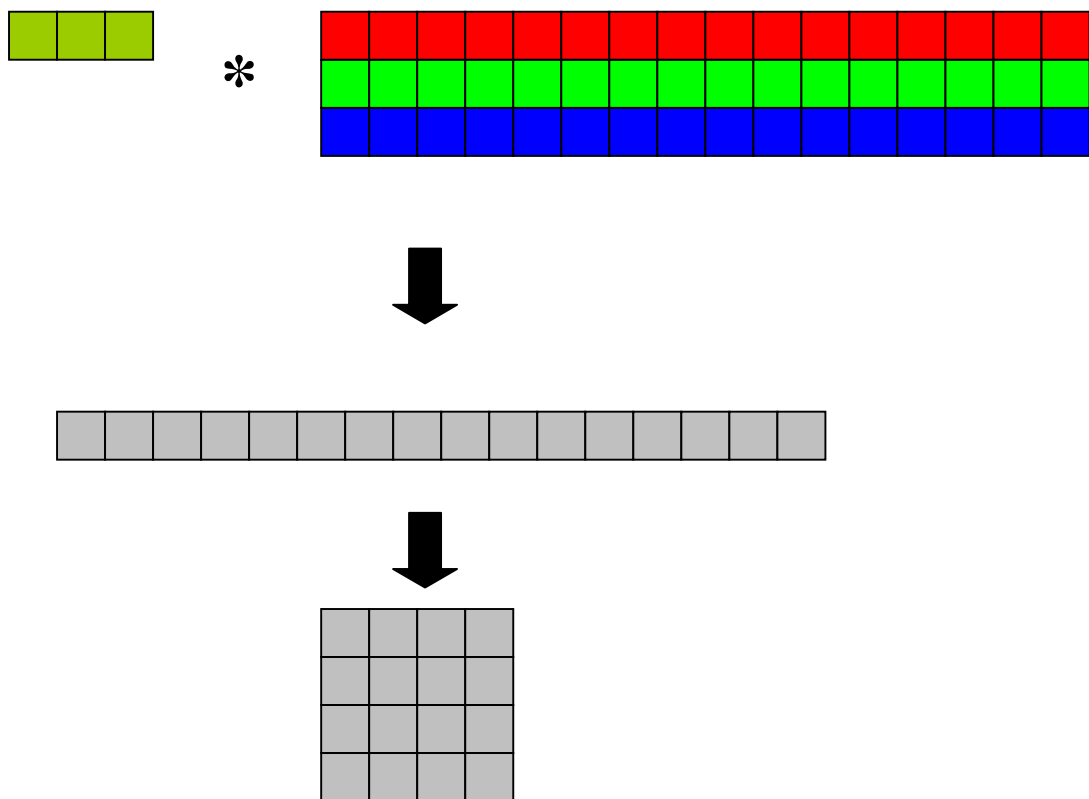


Figura 25 – O processo de obtenção da nova imagem em níveis de intensidade a partir da componente principal.

O algoritmo pode ser descrito da seguinte forma:

1. A informação dos três canais de cor (R, G e B) da imagem original é separada para diferentes variáveis;
2. Cria-se uma matriz cujas colunas contêm a informação dos canais de cor.
3. É calculada a média de cada coluna da matriz construída em 2;
4. Subtrai-se a cada elemento da matriz obtida em 2 a média da respectiva coluna e divide-se pelo número total de linhas da matriz obtida em 2 menos 1;
5. Executa-se a decomposição em valores singulares da matriz obtida em 4;
6. Escolhe-se a primeira coluna da matriz V, que é a componente principal;
7. Multiplica-se a transposta do vector obtido em 6 pela transposta da matriz obtida em 2.
8. Converte-se a matriz linha obtida em 7 para uma matriz com o mesmo número de linhas e colunas que a imagem original.



Figura 26 – Imagem resultante da PCA.

6.2.2. Detecção de arestas

Para a determinação das arestas, recorre-se a um operador *Canny*, sendo definidos o desvio padrão da gaussiana do filtro usado pelo operador, e os valores inferior e superior de aceitação das arestas.

Da aplicação do operador de *Canny* resulta uma imagem binária com as dimensões da imagem original, em que os pixels pertencentes a uma aresta são marcados com o valor de 1 e os restantes têm o valor de 0.

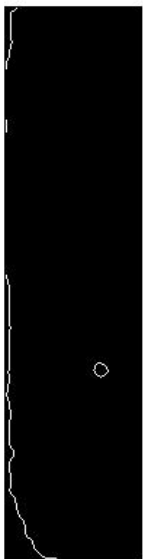


Figura 27 – As arestas detectadas a partir da imagem em níveis de cinzento, obtida com a PCA.

6.2.3. Detecção do fundo e eliminação de arestas da fronteira

As imagens são matrizes, e como tal têm sempre uma conformação quadrada ou rectangular. Assim, muitas são as vezes em que as fronteiras, da região segmentada não coincidem com os limites da matriz, resultando zonas de transição abrupta entre a imagem segmentada e o fundo, cujos pixels têm valor nulo. Estas transições abruptas, vão pois ser detectadas como arestas na imagem. Uma vez que estas arestas de fronteira não correspondem àquilo que se procura, e que são constituídas por um elevado número de pontos que não tem interesse incluir no próximo passo do processamento da imagem, é de todo o interesse a sua eliminação, por forma a aumentar a velocidade do processamento que se segue.

Para eliminar estes pontos de forma rápida e simples, cria-se uma máscara binária com as mesmas dimensões da imagem inicial, recorrendo a um método de *threshold*, em que os pixels da imagem inicial com valor nulo, ou seja, rejeitados pela segmentação inicial, assumem o valor de 1 e os restantes de 0.

A imagem resultante é dilatada usando-se uma máscara de 3 x 3.

Atribui-se o valor 0 a todos os elementos localizados numa posição à qual corresponde um elemento com valor 1 na máscara de bordo, eliminando-se desta forma as arestas detectadas na fronteira da região.



Figura 28 – (a) máscara de bordo; (b) a matriz com as arestas que não são eliminadas pela máscara (a).

6.2.4. Aplicação da transformada de *Hough* circular

Nesta fase vai-se aplicar uma transformada de *Hough* adaptada à detecção de círculos.

Antes de mais começa-se por construir a matriz acumuladora. Esta é constituída por R matrizes, com o mesmo número de linhas e colunas da imagem original, sendo os seus elementos inicializados a 0. R é o número de raios diferentes considerados na análise. Para que as projecções sejam executadas de forma correctas e tenham significado segundo as coordenadas da imagem, neste passo é necessário ter em atenção o facto de que o *Matlab* tem um sistema de coordenadas diferente para imagens e para matrizes genéricas.

Uma vez determinadas as arestas da imagem original, é necessário saber as coordenadas dos pontos que as constituem, uma vez que é sobre eles que incide a transformada de *Hough*.

Para cada ponto com valor 1 encontrado, são determinados os pontos constituintes do círculo de raio R_i centrado nesse ponto, ou seja, todos os pontos à distância R_i do ponto em análise.

Uma vez que a matriz acumuladora tem um número discreto de divisões, o círculo sofre uma quantização. Assim, é somada uma nova acumulação numa determinada célula acumuladora de cada vez que é projectado um novo ponto nessa célula, ou caso esta seja a mais próxima do ponto.

No final, é devolvida a célula acumuladora.

O algoritmo pode ser descrito da seguinte forma:

1. Criam-se R matrizes ou células acumuladoras, com as dimensões da imagem original.
2. Procuram-se as coordenadas dos pontos constituintes das arestas.
3. Para cada raio
 - 3.1. São calculadas as coordenadas de início do círculo centrado no ponto
 - 3.2. São calculadas as coordenadas de fim do círculo centrado no ponto.
 - 3.3. Para os pontos situados no eixo entre o início e o fim são calculados os restantes pontos do círculo.
 - 3.4. É marcada uma nova acumulação na célula acumuladora mais próxima do ponto projectado.



Figura 29 – Imagem no chamado espaço de *Hough*.

6.2.5. Determinação da posição do guia laser

Procura-se o valor máximo de acumulação. A partir deste obtêm-se as coordenadas da respectiva célula na matriz acumuladora. Esta célula é considerada o centro do guia laser procurado.

6.2.6. Resultados da detecção do guia laser

Adicionou-se ao módulo, código que permitisse visualizar os resultados da detecção. Assim, este devolve a imagem inicial com uma circunferência a verde com o raio do guia laser, em redor do ponto detectado.

O algoritmo desenvolvido executa a detecção em aproximadamente 0.2 segundos, tendo-se assim atingido a velocidade de processamento pretendida.

O facto de se ter usado uma linguagem interpretada de programação, leva ainda a pensar que será possível obter uma maior velocidade de processamento recorrendo a linguagens de mais baixo nível como C/C++.

Não havendo imagens com o guia laser, e verificando-se o seu bom desempenho em imagens marcadas com um modelo artificial, optou-se por não se prolongar o trabalho neste campo, já que não faria sentido testar exaustivamente o modelo teórico.

Adicionou-se ainda a este módulo a capacidade para guardar as coordenadas do centro do guia laser. Estas são guardadas num ficheiro de texto com o nome *laser_shots_coordinates* na directoria c:\ . Para cada disparo são guardadas em linhas separadas, a data do disparo e as coordenadas do guia laser detectado.

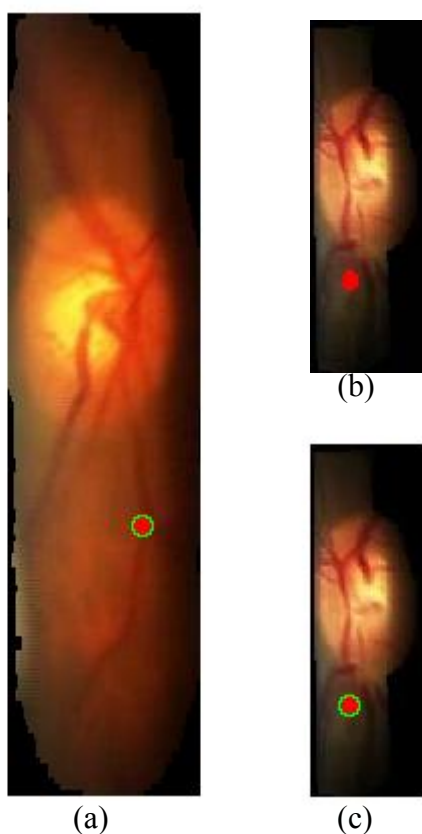


Figura 30 – Imagens resultantes da detecção do guia laser, em que a circunferência a verde se encontra centrada no círculo detectado: (a) imagem utilizada ao longo da apresentação do algoritmo; (b) outro exemplar em que o modelo do guia laser foi marcado; (c) resultado do processamento de (b) com o algoritmo desenvolvido.

7. Segmentação das imagens adquiridas

A falta de imagens contendo o guia laser, impediu que se prosseguisse com os testes do algoritmo de detecção, para avaliação do seu desempenho em casos reais.

Isto levou a que se tivesse que alterar os objectivos iniciais do projecto.

Decidiu-se desta forma, criar um módulo que permitisse a segmentação de imagens do fundo ocular adquiridas através do *hardware* disponibilizado.

Este módulo pode ser visto como pertencente a uma fase de pré-processamento da imagem onde se pretende detectar o guia laser, estando desta forma ligado ao trabalho há desenvolvido.

A segmentação das imagens adquiridas é um passo extremamente importante. A identificação da região de interesse, permite eliminar todos os elementos indesejados, advindo daí a possibilidade de restringir o processamento seguinte a uma menor região. A importância deste facto é evidente, já que o tempo de processamento para determinação do guia laser está fortemente dependente das dimensões da imagem, uma vez que tem que recorrer a métodos iterativos.

O algoritmo de segmentação foi desenvolvido tendo também em vista a criação de um módulo que permitisse a visualização em contínuo das imagens segmentadas do fundo ocular. É assim de toda a importância eliminar todos os reflexos e artefactos, preservando ao máximo a região da imagem correspondente ao fundo ocular. No entanto, e como é sabido, um processamento mais aprofundado implica necessariamente um maior consumo de tempo. Como o tempo de processamento assume um papel relevante, muitas das técnicas e procedimentos normalmente utilizadas no processamento de imagens do fundo ocular, nomeadamente as baseadas na análise de histogramas para aumento da qualidade da imagem, deixam de poder ser utilizadas neste caso, sendo necessário adaptar o processamento aos recursos computacionais disponíveis.

Apesar de ser uma ferramenta bastante difundida, o biomicroscópio com lâmpada de fenda fornece imagens com uma utilidade clínica limitada, quer no que toca ao diagnóstico, ao tratamento ou a fotodocumentação, comparativamente a uma câmara de fundo. A qualidade da imagem é reduzida devido a grandes variações de contraste, a um reduzido campo de visão, e a reflexos especulares provenientes da córnea, esclerótica e das lentes usadas para o exame. O facto do paciente não estar

completamente imóvel, é também um factor importante, pois este ao mudar a sua distância em relação ao sistema de observação do biomicroscópio, vai também influenciar a quantidade de luz reflectida pelos seus olhos, mudando a aparência dos objectos ou cenas na imagem capturada, e introduzindo também outros elementos indesejados como pupilas, pele da pálpebra, pestanas, etc. Como já referido, a fenda luminosa pode também assumir várias formas e dimensões, e surgir em várias zonas da imagem.

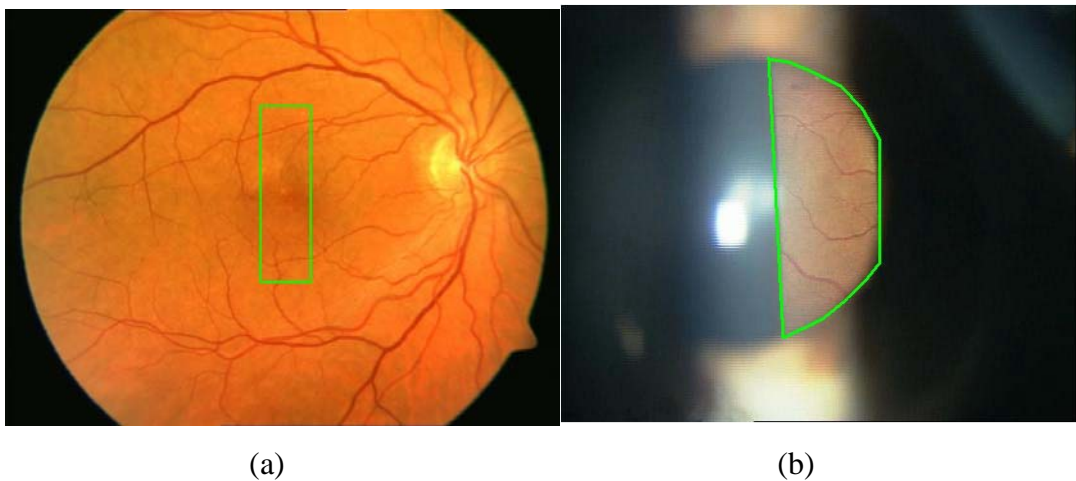


Figura 31 – Alguns problemas associados às imagens adquiridas através de um biomicroscópio com lâmpada de fenda: (a) uma imagem obtida com uma câmara de fundo, onde foi marcado um rectângulo a verde, representando o campo de visão típico de um biomicroscópio com lâmpada de fenda; (b) uma imagem adquirida com o *hardware* disponibilizado, onde a região de interesse foi delimitada a verde. É ainda possível ver nesta imagem vários reflexos, e pedaços de íris.

Nas imagens captadas há um grande conjunto de factores em constante mudança, tornando-se difícil definir condições que ajudem ao seu processamento.

Enquanto que o interesse e esforço têm sido dirigidos para a análise, extracção de características, detecção de mudanças e tratamentos guiados em imagens do fundo ocular de boa qualidade, a análise computadorizada de imagens obtidas com este tipo de instrumento tem recebido pouca atenção. Daí ser um assunto mal documentado. [28]

De entre uma panóplia de elementos que podem estar presentes nas imagens, definiu-se um pequeno conjunto que engloba os mais comuns.

No fundo ocular há também vários elementos com características muito distintas e que se pretende preservar.

É sabido que a coloração do fundo ocular varia consideravelmente de pessoa para pessoa, e de acordo com a posição no fundo ocular individual. Esta diferença na coloração tem sido atribuída à quantidade de pigmentos oculares presentes. No entanto, se esta gama de coloração fosse representada num espaço de cor tridimensional (sendo RGB as coordenadas) poder-se-ia ver que ocupa um volume bem definido por eixos fisiologicamente significativos. Estes eixos representariam a concentração de melanina no epitélio pigmentado da retina, sangue na coróide e melanina na coróide. Este volume descreve a coloração de um tecido ocular. [29]

Face a este cenário, a informação colorimétrica parece ser o único ponto de partida para a segmentação. Procurou-se então gerar um modelo capaz de reconhecer todas as cores que se podem encontrar em fundos oculares.

Elemento a preservar	Tamanho	Cor	Forma	Observações
Disco óptico	Grande	Amarelada	Redondo/oval	Contornos bem definidos Bom reflector
Drusas	Variável	Amarelada	Irregular	Bom reflector
Mácula/Fóvea	Variável	Acinzentada	Circular/oval	Mau reflector
Vasos sanguíneos	Variável	Avermelhada	Variável	

Tabelas 1 - Elementos mais comuns que se pretende preservar nas imagens captadas, e suas principais características.

Elemento a eliminar	Tamanho	Cor	Forma	Observações
Reflexos especulares	Variável	Variável mas com grandes contribuições de G e B	Irregular	Contornos de definição variável
Reflexos difusos	Variável	Variável	Irregular	Intensidade luminosa muitas vezes inferior à região de interesse
Íris	Grande	Variável	Circular	Surge perifericamente em relação à imagem do fundo ocular

Tabela 2 - Elementos mais comuns que se pretende eliminar das imagens e suas principais características.

7.1. Princípios matemáticos

O algoritmo de segmentação das imagens adquiridas faz uso de várias técnicas matemáticas.

Uma parte significativa do trabalho desenvolvido, no que toca a este tópico, consistiu na recolha e análise de dados, com o objectivo de se obter um modelo que permitisse classificar os elementos da imagem de forma rápido e eficaz, para que os resultados da classificação possam ser utilizados na segmentação da imagem adquirida.

Nesta secção pretende-se então apresentar princípios e conceitos envolvidos no trabalho desenvolvido.

7.1.1. Classificadores

Uma capacidade essencial dos sistemas automáticos inteligentes é a capacidade de reconhecer objectos. O sistema de reconhecimento de padrões envolve três fases: extracção de características dos objectos a observar, selecção das características mais descritivas e construção de um classificador. Na primeira fase é feita uma recolha dos dados por um processo de medição. Na segunda, procura-se reduzir a dimensão dos dados preservando a informação essencial para a decisão. Por fim, temos o classificador que contém um conjunto de funções descritoras dos nossos grupos, havendo assim uma função por cada grupo. Calculando o valor das funções para cada objecto observado, obtem-se assim a sua classificação, que corresponde ao maior valor obtido.

De uma forma generalista, há dois tipos possíveis de abordagens: a clássica ou estatística, em que se assume um modelo das distribuições probabilísticas dos dados, e uma mais recente, em que se segue uma estratégia de aprendizagem indutiva em que o algoritmo “aprende” o modelo a partir dos dados. Para esta última abordagem temos o exemplo das redes neuronais e das árvores de decisão.

Há três factores importantes para o processo de aprendizagem: a complexidade amostral, a complexidade computacional e o desempenho do algoritmo. [30]

Consoante os casos, são ainda consideradas duas estratégias diferentes para classificação de regiões individuais: baseadas em regiões ou baseadas numa classificação dos pixels. Nas técnicas baseadas em regiões, são analisados conjuntos de

dados com medições de características de uma região, enquanto que numa classificação baseada nos pixels, é obtida informação contextual ou não-contextual sobre cada pixel individual.

[21]

7.1.1.1. Análise discriminante

A análise por funções discriminantes é usada para determinar quais as variáveis que melhor discriminam entre dois ou mais grupos.

Computacionalmente, a análise por funções discriminantes é muito semelhante à análise de variâncias (ANOVA). A ideia base deste método consiste em determinar quando é que grupos diferem no que respeita à média de uma variável, usando posteriormente essa variável para prever a pertença a um grupo. Para que se entenda esta metodologia, deve antes de mais ser claro que se as médias para as variáveis entre grupos distintos forem muito distintas, então pode-se dizer que essas variáveis são boas discriminadoras dos diferentes grupos.

No caso de termos apenas uma variável, o teste final de significância dessa variável para discriminar grupos é o teste F . F é essencialmente calculado como a taxa da variância inter-classes dos dados sobre a média da variância intra-classe.

[21, 31]

7.1.1.1.1. Análise discriminante *stepwise*

De uma forma simplificada, procura-se um modelo que permita prever a que classe um caso pertence. Na análise discriminante *stepwise*, há variáveis que podem ou não ser incluídas na predição de pertença a um grupo, construindo-se um modelo passo-a-passo ou *stepwise*. O procedimento *stepwise* é “guiado” pelo valor de teste F . O valor de F para uma variável indica a sua significância estatística na discriminação entre grupos, ou seja, é uma medida da extensão para a qual a variável faz uma contribuição unitária para a predição da pertença a um grupo. Para se chegar a uma conclusão, pode-se seguir as seguintes metodologias:

- a) Análise *forward stepwise* – em cada passo as variáveis são revistas e avaliadas para se determinar qual a que mais contribui para a discriminação entre classes. Essa variável é depois incluída no modelo, e o processo recomeça;
- b) Análise *backward stepwise* – neste caso, todas as variáveis são incluídas no modelo e então, passo-a-passo, a variável com menor contribuição para a predição é eliminada. Desta forma, num modelo bem sucedido, apenas são mantidas as variáveis mais importantes, ou seja, aquelas que mais contribuem para a discriminação entre grupos.

Usualmente são incluídas diversas variáveis num estudo de forma a se determinar quais as que melhor fazem uma discriminação entre grupos.

Na análise discriminante linear procura-se a maior separabilidade entre as classes para se criar uma região de decisão entre as diferentes classes. [32]

Considerando a função discriminante $g(x)$, esta divide o espaço de entrada em duas categorias de regiões dependendo se $g(x) > 0$ ou $g(x) < 0$.

Ao caso mais simples corresponde a equação:

$$g(x) = w \cdot x + w_0$$

, onde w é o vector de peso e w_0 é o chamado pólo ou zero na origem.

As funções discriminantes podem ser generalizadas por exemplo para superfícies de decisão quadráticas do género

$$g(x) = w_0 + \sum_{i=1} w_i x_i + \sum_{i=1} \sum_{j=1} w_{ij} x_i x_j$$

Desta relação chega-se à conclusão que $g(x) = 0$ define uma superfície de decisão. Para o caso linear, esta região é definida por um hiperplano, ou seja, um subespaço linear, afim ou projectivo de codimensão 1. Supondo que x_1 e x_2 são pontos desse hiperplano, então a sua diferença será um vector localizado nesse hiperplano

$$w \cdot x_1 + w_0 = w \cdot x_2 + w_0 \Leftrightarrow w \cdot (x_1 - x_2) = 0$$

Então o vector de peso w é normal a qualquer vector que passe pelo hiperplano. Desta forma w determina a orientação do hiperplano, O vector normal w aponta para a região onde $g(x) > 0$.

Considerando x como um ponto no hiperplano, se projectarmos x no vector de peso normalizado $x.w/|w|$ temos a distância normal do hiperplano à origem igual a:

$$w.x / |w| = -w_0 / |w|$$

Desta forma, o pólo ou zero na origem determina a posição do hiperplano. Pode-se da mesma forma mostrar que a distância normal ao hiperplano é dada por $|g(x)| / |w|$.

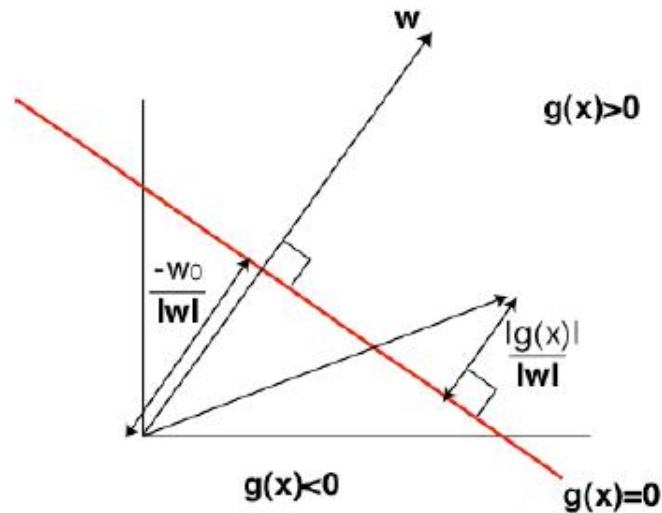


Figura 32 – Representação gráfica de alguns de conceitos ligados à análise discriminante.

Supondo que há n classes, há várias maneiras de definir regras para classes discriminantes. Uma maneira que evita regiões não definidas pode ser descrita como:

$$g_i(x) = w_i \cdot x + w_{i0}, i = 1, \dots, n$$

Face ao tipo de classificação que é feito, é demonstrável que este classificador divide os dados da análise em regiões convexas simples interconectadas, sendo incapaz

de discernir conjuntos separados de características pertencentes à mesma classe e apresentando dificuldades no caso de uma classe ter vários modelos probabilísticos como base. [21, 31]

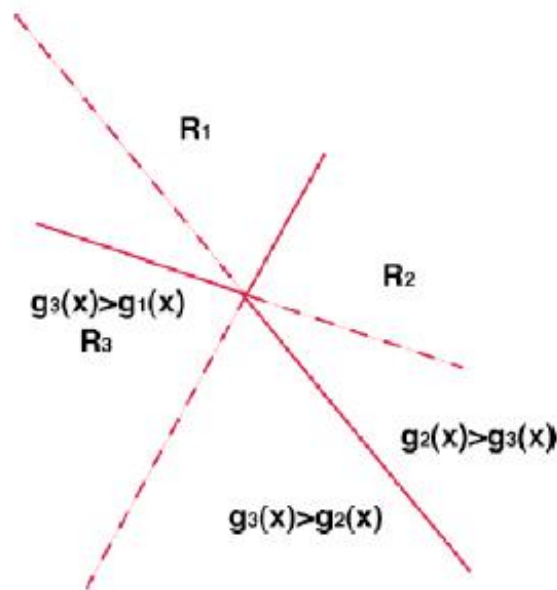


Figura 33 – Aspecto da divisão de um espaço bidimensional por várias funções discriminantes.

7.1.1.1.2. Análise canónica

Quando se executa uma análise discriminante não se tem obrigatoriamente de especificar a maneira como as várias classes devem ser combinadas para se obterem diferentes funções discriminantes. Em vez disso, pode-se determinar automaticamente uma combinação ótima de variáveis de forma a que a primeira função forneça a descrição mais global dos grupos, a segunda a melhor, etc. As funções criadas são também independentes e ortogonais, ou seja, as suas contribuições para a discriminação dos grupos não se sobrepõem. Computacionalmente, é executada uma análise de correlação canónica que vai determinar as sucessivas funções e raízes canónicas ou valores próprios associados às respectivas funções canónicas. O número máximo de funções é igual ao menor número de entre: o número de classes menos uma, ou o número de variáveis da análise. [21]

7.1.1.1.3. Interpretação das funções discriminantes

Como já referido, vai-se obter os coeficientes w para cada variável em cada função discriminante, que podem ser interpretados da seguinte forma: quanto maior o coeficiente normalizado, maior é a contribuição da respectiva variável para a discriminação entre grupos. Contudo, estes coeficientes não dizem a qual dos grupos as respectivas funções discriminam. Pode-se identificar a natureza da discriminação para cada função discriminante/canónica observando as médias das funções através dos grupos.

A discriminação dos grupos pode também ser analisada através da visualização de gráficos das classificações para os vários grupos.

Outra forma de determinar quais as variáveis que determinam uma função discriminante particular é observando a estrutura factorial. Os coeficientes da estrutura factorial são as correlações entre as variáveis no modelo e as funções discriminantes. Estas correlações podem ser vistas como a carga das variáveis em cada função discriminante.

O mais importante a reter é que os coeficientes das funções discriminantes denotam a contribuição parcial de cada variável, enquanto que os coeficientes de estrutura denotam a correlação simples entre as variáveis e as funções. Caso se pretenda saber a significância das funções discriminantes, deve-se analisar a matriz de estrutura factorial, caso se pretenda saber a contribuição individual de cada variável para a função discriminante, então deve-se prestar atenção aos coeficientes da função discriminante.

7.1.1.1.4. Condições assumidas

Como previamente mencionado, a análise de funções discriminantes é computacionalmente semelhante à análise *MANOVA*, sendo aplicáveis à primeira, as condições impostas à segunda.

São estas então:

a) Distribuição normal – é assumido que os dados (para as variáveis) representam uma amostra de uma distribuição normal multivariada. Contudo, a violação da normalidade não é fatal para a análise, sendo ainda utilizáveis os resultados relativos

aos testes de significância. O tipo de distribuição dos dados pode ser determinado através de testes específicos;

b) Homegeneidade das variâncias/covariâncias – é assumido que as matrizes de variância/covariância das variáveis são homogêneas ao longo dos grupos. Os pequenos desvios não induzem uma margem de erro na análise, contudo, convém verificar sempre os valores da variância intra-classe e as matrizes de correlação. Para esta análise pode-se realizar o teste *Box M* descrito na análise *ANOVA/MANOVA*. Este teste é particularmente sensível a desvios da normalidade multivariada, sendo como tal um parâmetro ao qual não se deve dar demasiada importância.

c) Correlação entre médias e variâncias – a maior ameaça para os testes de validade da significância ocorre quando as médias ao longo dos grupos estão correlacionadas com as variâncias. Intuitivamente, se houver uma grande variabilidade num grupo com uma média particularmente elevada, então essas médias elevada não são fiáveis. Contudo, os testes de significância, na sua globalidade, são baseados na variância média ao longo de todos os grupos. Assim, os testes de significância para médias significativamente mais elevadas (com maior variância) são baseados nas relativamente mais pequenas variâncias globais, resultando significância estatística errônea. Na prática, este padrão pode ocorrer se uma das classes em estudo tiver alguns *outliers* extremos, que tenham um grande impacto nas médias, e que aumentem a variabilidade. A prevenção deste problema torna necessária a verificação da estatística descritiva, ou seja, as médias e os desvios padrão ou variâncias para tal correlação.

d) Valor de tolerância – Outra suposição, que é tida à partida, é que as variáveis não são totalmente redundantes. Os cálculos envolvidos na análise discriminante, incluem uma inversão da matriz de variância/covariância das variáveis do modelo. Se uma das variáveis for completamente redundante, então não é possível inverter esta matriz. Isto pode acontecer, por exemplo, para uma variável que é a soma de várias outras. Para se evitar esta situação, é necessário verificar os chamados valores de tolerância para cada variável. Este valor de tolerância é calculado como 1 menos a raiz quadrada da respectiva variável, estando todas as outras variáveis incluídas no modelo corrente. Então, é a proporção de variância que é única para a respectiva variável. De uma forma geral, o valor de tolerância para uma variável redundante tende para 0.

7.1.1.1.5. Predição e classificação de casos

Após construído um modelo, interessa saber qual a capacidade do modelo para reconhecer a que grupo pertence um caso particular. Esta capacidade é avaliada através da classificação preditiva das classes.

Como é óbvio, se a estimativa da capacidade de reconhecer os casos, for executada com dados usados na construção dos modelos, então obter-se-á uma classificação melhor, do que no caso em que isto não se verifique. Em geral, é mais fácil prever um acontecimento já ocorrido (predição *post hoc*), do que prever um acontecimento novo, futuro e como tal desconhecido (predição *a priori*). Assim, torna-se desaconselhável criar intervalos de confiança para determinar correctamente o sucesso de futuras classificações, com base em dados a partir dos quais se criaram as funções discriminantes. Os intervalos de predição devem ser criados com base num novo conjunto de dados.

As funções de classificação servem para determinar a que classe é que um caso tem maior tendência de pertença. Há tantas funções de classificação quanto classes, e permitem determinar pontuações para a pertença a uma classe. A sua formulação matemática é

$$S_i = w_{i1} \cdot x_1 + w_{i2} \cdot x_2 + \dots + w_{in} \cdot x_n + c_i$$

, em que o índice i denota a classe, os índices $1, 2, \dots, n$ dizem respeito às n variáveis, c_i é a constante para a classe i , w_{ij} é o peso da variável j no cálculo do valor de classificação para a classe i ; x_j é o valor observado para o caso respectivo da variável j , e S_i é o resultado da classificação.

Apesar da sua semelhança matemática com as funções discriminantes, estes dois tipos de função não devem ser confundidas.

As funções de classificação podem também ser usadas para calcular directamente o valor de classificação para algumas novas observações.

Uma vez calculados os valores de classificações para um caso, é fácil classificá-lo. Normalmente, um caso é classificado como pertencente à classe com o maior valor de classificação. No entanto, esta classificação pode por vezes falhar. Assim, é de todo o

interesse saber a probabilidade de uma previsão falhar ou não. Esta é conhecida como probabilidade *posterior*. A probabilidade *posterior* é também passível de cálculo, e está relacionada com a distância *Mahalanobis*.

A distância *Mahalanobis* é uma medida da distância entre dois pontos num espaço definido por duas ou mais variáveis correlacionadas. Para variáveis descorrelacionadas, a distância *Mahalanobis* é semelhante à distância euclidiana. No entanto, quando as variáveis estão correlacionadas, os eixos de coordenadas que definem os pontos, podem ser pensados como não sendo ortogonais, deixando a distância euclidiana de ser uma medida apropriada ao contrário da distância *Mahalanobis*.

Para cada classe da amostra pode-se determinar a localização do ponto que representa as médias para todas as variáveis no espaço multivariado definido pelas variáveis no modelo. Estes pontos são chamados os centróides das classes. Para cada caso pode-se então calcular a distância *Mahalanobis* ao centróide da classe. Desta forma, classifica-se o caso como pertencente à classe mais próxima, ou seja, àquela com uma menor distância *Mahalanobis*.

Esta última afirmação não é contudo totalmente correcta, pois como já referido, é assumida uma distribuição normal multivariada em redor de cada centróide. Uma vez que se calcula a distância para cada caso tendo conhecimento prévio dos valores para o caso nas variáveis do modelo, estas probabilidades são designadas de probabilidades posteriores. Em resumo, a probabilidade posterior é a probabilidade baseada no conhecimento prévio dos valores de outras variáveis, de que o respectivo caso pertence a um grupo particular.

Há ainda um outro factor adicional que deve ser considerado na classificação de casos. Por vezes, quando se sabe em avanço que um grupo tem mais observações que outro, a probabilidade *a priori* de um caso pertencer a esse grupo é maior. A possibilidade de especificar diferentes probabilidades *a priori*, por forma a compensar as diferenças de tamanho entre os grupos, permite ajustar a classificação dos casos e o cálculo das probabilidades posteriores, afectando de forma significativa a exactidão da predição. [31]

7.1.1.1.6. Análise discriminante linear de *Fisher*

Esta variante da análise discriminante linear assenta na ideia de que de um espaço de entrada impraticavelmente enorme, se pode definir um subespaço (hiperplano) que preserve as diferenças entre classes tão bem quanto possível podendo-se assim criar um espaço de decisão mais reduzido. De uma forma simplista, tenta encontrar características que sejam distintivas.

No caso de dois grupos, a análise de funções discriminantes pode ser comparada (e é análoga) a uma regressão múltipla. A interpretação dos resultados é directa e é semelhante à lógica da regressão múltipla: as variáveis com maiores coeficientes de regressão são aquelas que mais contribuem para a predição de pertença a um grupo.

De seguida será apresentado como exemplo o caso mais simples, em que se consideram apenas duas classes.

A medida da separação entre as projecções é a diferença entre as médias

$$|w.(m_a - m_b)|$$

, onde

$$m_a = \frac{1}{N} \sum_{i=1}^N x, \text{ somando sobre os } N x \text{ da classe } a$$

$$m_b = \frac{1}{N} \sum_{i=1}^N x, \text{ somando sobre os } N x \text{ da classe } b$$

, onde w é o vector unitário desconhecido ao longo da linha discriminante.

Para se otimizar a discriminação, é de todo o interesse ter uma diferença entre as médias das classes superior às suas variações individuais.

Para se aumentar esta separação entre os grupos, não se pode simplesmente redimensionar w , uma vez que os dados podem conter ruído que será também reajustado.

Para uma classe pode-se definir uma medida da dispersão para as amostras projectadas. Tomando como exemplo a classe a , tem-se:

$$\sum_{y \in \text{classe } a}^N (y - \tilde{m}_a)^2$$

, onde \tilde{m}_a é a média da amostra dos pontos da classe a projectados na linha discriminante e $y = w \cdot x$.

A dispersão total S é definida pela soma da dispersão de todas as classes (neste caso a e b) por

$$S = \sum_{y \in \text{classe } a} (y - \tilde{m}_a)^2 + \sum_{y \in \text{classe } b} (y - \tilde{m}_b)^2$$

Dividindo S pelo número total de pontos, obtém-se uma estimativa da variância dos dados combinados ao longo do eixo de projecção.

Tem-se assim os fundamentos básicos por trás do discriminante linear de *Fisher*.

Partindo daqui, interessa pois saber qual o maior w para J

$$J(w) = \frac{|\tilde{m}_a - \tilde{m}_b|^2}{S} = \frac{|w \cdot (m_a - m_b)|^2}{S}$$

Pretende-se então maximizar a diferença entre as médias das classes projectadas, enquanto se minimiza a dispersão dos dados na linha projectada.

Pode-se demonstrar que $S = w^T \cdot S_w \cdot w$ onde S_w mede a variação intra-classe, sendo também conhecido como matriz de dispersão intra-classe:

$$S_w = \sum_{i=1}^2 \sum_{x \in \text{class } i} (x - m_i)(x - m_i)^T$$

Para o numerador a medida da variação inter-classe é a matriz de dispersão inter-classes

$$S_B = (m_1 - m_2) \cdot (m_1 - m_2)^T$$

, com a diferença entre as médias projectadas dada por

$$|\tilde{m}_a - \tilde{m}_b|^2 = w^T S_B w$$

Calculando w que maximiza a função de critério $J(w) = \frac{w^T \cdot S_B \cdot w}{w^T \cdot S_w \cdot w}$ obtém-se

então $w = S_w^{-1} \cdot (m_a - m_b)$.

Desta forma, transforma-se um problema bidimensional num problema unidimensional. [21, 31]

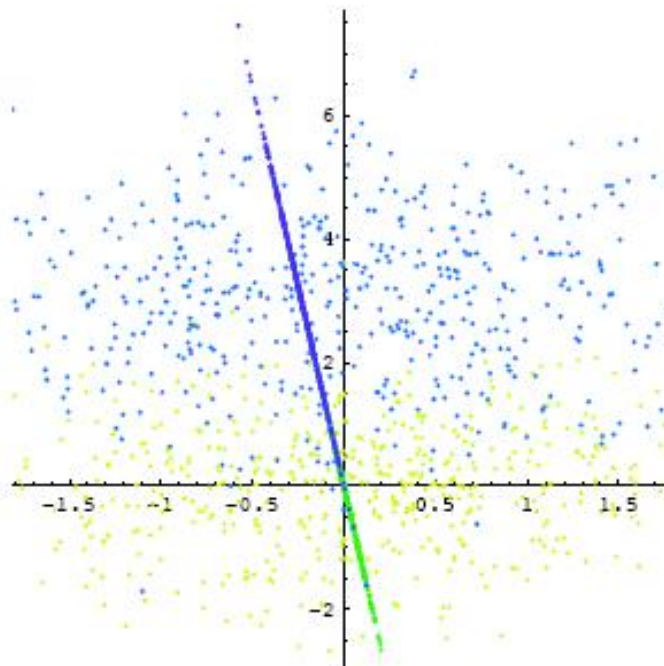


Figura 34 – Representação gráfica da redução dimensional através da análise de *Fisher*: caso bidimensional, em que o problema é reduzido para um problema unidimensional.

7.1.1.1.7. Selecção de modelos e o dilema do deslocamento-variância

O método descrito anteriormente parte do princípio que os dados têm uma separação linear. No entanto, muitos são os casos em que isto não acontece, sendo necessário recorrer a superfícies de decisão mais complexas.

Tendo em conta o processo de regressão linear, o modelo resultante pode ser resultado de uma adaptação ao conjunto de dados analisado.

Há assim um problema relacionado com a escolha de uma função que mantenha uma boa relação entre as variáveis, e que ao mesmo tempo seja genérica o suficiente para se adaptar a novos casos. Por outro lado, esta deve ser parametrizada de forma a permitir um cálculo simples.

Há assim, uma desvantagem suportada pela flexibilidade fornecida por graus de liberdade adicionais para adaptação aos dados. Pode-se assim obter parâmetros muito variados para diferentes conjuntos de dados, a partir do mesmo processo de análise. O facto de se obterem diferentes parâmetros de adaptação (como por exemplo o declive da linha de regressão) de cada vez, quer dizer que apesar de se poder adaptar correctamente o modelo aos dados, são introduzidas variações entre a média de adaptação global, e a adaptação para um conjunto individual de dados. Este problema poderia ser superado pela análise de um conjunto de dados muito grande, no entanto o conjunto exigido cresce exponencialmente com a ordem de adaptação.

Por outro lado, se a função é restritiva, como é o caso de linhas rectas que passam pela origem, então, obtém-se o mesmo ajuste para diferentes conjuntos de dados, uma vez que o que tem que se ajustar é um parâmetro, o declive. O problema aqui é que o ajuste apenas pode ser bom se uma linha recta passar pela origem. Caso não seja uma linha recta, por exemplo, há um erro fixo dado pelo pólo que não desaparece por muitos dados que se recolham. Este é referido como o dilema do deslocamento-variância.

O ajuste com menos parâmetros pode não ser tão sensível à inevitável variabilidade do conjunto de dados do conjunto de treino, mas pode dar grandes constantes de erro ou pólos se os dados não se adaptarem ao modelo subjacente.

Não há forma de dar a volta a este problema, e as redes neuronais não são excepção.

È assim necessário estudar o problema que se está a tentar modelar e escolher o modelo mais adequado. Porque uma linha recta, ou em maiores dimensões, um modelo linear é o modelo matematicamente mais simples, faz sentido começar-se sempre por ele. Mais tarde pode-se escolher outro tipo de modelos, caso se descubra que o conjunto de dados em análise segue um modelo não-linear.

[21, 32]

7.2. Segmentação estática e segmentação dinâmica

Considere-se uma imagem I , e cada pixel $i \in I$. Um espaço de cor C é o espaço de cor de uma imagem captada por uma câmara a cores. O espaço da cor é tipicamente representado por 3 componentes (ex: RGB, HSV, etc), sendo que diferentes espaços de cor apresentam diferentes propriedades de segmentação da cor. Considerando ainda $c(i)$ a cor de um pixel num dado espaço de cor i , no caso de um espaço tridimensional $c(i)$ é constituído por 3 componentes.

A segmentação da cor pode ser definida como a tarefa de atribuir a um pixel da imagem uma classe, dentro de um conjunto de classes de cor pré-definidas. As classes de cor podem ser vistas como uma partição da totalidade do grupo de cor C , $CC = \{C_1, \dots, C_n\}$, tal que $C_k \in C$ e $K_1 \neq K_2 \Rightarrow C_{K_1} \cap C_{K_2} = \emptyset$. É conveniente considerar uma cobertura completa das classes de cor, ou seja $C = \bigcup_{k=1}^n C_k$, possivelmente definindo uma classe de cor especial que pode incluir cores indefinidas ou desinteressantes.

Desta forma, a segmentação da cor pode ser obtida pela avaliação da função $f_{CC} : I \rightarrow \{1, \dots, n\}$

$$f_{CC}(i) = k, \text{ tal que } c(i) \in C_k$$

Conhecendo com antecedência as classes de cor procuradas, a segmentação da cor pode seguir duas abordagens, que podem ser classificadas como estática e dinâmica. [23, 33]

7.2.1. Segmentação estática

As abordagens baseadas numa segmentação estática, definem de forma estática um conjunto de classes de cor antes da execução da tarefa, não sendo alteradas durante a operação. As definições estáticas de CC podem ser obtidas através da definição de *thresholds* no espaço de cor, contudo sofrem de alguns tipos de limitação: são sensíveis a variações da iluminação, a representação explícita de tabelas de cor é um processo consumidor de tempo (a não ser que sejam usados métodos de aprendizagem

automática), e os *thresholds* estáticos podem não ser suficientes para dividir o espaço de cor. [33]

7.2.2. Segmentação dinâmica

A segmentação dinâmica é obtida pela computação dinâmica das classes CC_t no tempo t para cada imagem I_t , possivelmente considerando classes de cor prévias CC_{t-1} .

Assim, o processo é adaptativo, uma vez que se encontra relacionado com a imagem actual e é executado durante a execução da tarefa de segmentação. Desta forma, é mais robusto ao ruído e a variações da intensidade luminosa e não requer uma calibração intensiva. Contudo, as abordagens dinâmicas podem ser computacionalmente dispendiosas e de difícil aplicação em sistemas que se querem rápidos e que necessitam de processar uma grande quantidade de dados em curtos espaços de tempo, com recursos computacionais limitados. [33]

7.3. Análise estatística das imagens

Na procura de um modelo matemático que permitisse discriminar o conjunto de dados colorimétricos, foi necessário executar várias análises estatísticas.

7.3.1. Recolha de dados

As imagens utilizadas foram retiradas de sequências de vídeo gravadas no CNTM da AIBILI. Os vídeos foram obtidos com o *hardware* utilizado neste projecto, e o *software IFC Camera Configurator 5.9.0.10* da *Dalsa Coreco*.

No total foram gravados cerca de 10 vídeos com imagens do fundo ocular de diferentes pacientes, dos quais foram utilizadas cerca de 35 imagens para criação dos modelos e cerca de 15 imagens para avaliação da predição dos mesmos.

É ainda de extrema importância referir que aquando da captação dos vídeos foi utilizado um iluminante de luz branca.

Para extracção de imagens dos vídeos foi utilizada uma livreria do *Matlab* própria para o efeito.

As *frames* foram seleccionadas tendo em conta que se procurava um conjunto que fosse o mais variado possível, no que toca à coloração dos elementos do fundo ocular, aos vários tipos de reflexos e aos artefactos mais frequentes. Este critério teve por base o objectivo de se construir um conjunto que englobasse o maior número de casos possíveis para classificação, permitindo assim uma análise não só dos elementos do próprio indivíduo (intrasujeito), mas também de diferentes sujeitos (intersujeito).

Estas imagens foram guardadas e posteriormente sujeitas a uma segmentação manual.

Partindo do conhecimento nulo acerca das imagens captadas, decidiu-se distinguir dois grupos para análise. Um dos grupos diz respeito àquilo que se pretende manter na nossa imagem, ou seja, o fundo ocular com todos os seus constituintes, e o outro foi definido pelos vários elementos da imagem que se pretende rejeitar, como os reflexos, íris e artefactos.

Construiu-se um *software* que permite carregar as imagens, seleccionar as regiões de interesse e guardar essa informação segundo dois tipos de imagens: uma contendo a região seleccionada (fundo ocular) sobre um fundo preto, e outra contendo a região não seleccionada com a região seleccionada marcada a preto. A gravação das imagens segmentadas para além de permitir ter esta informação sempre disponível, permitiu também interpretar os resultados das análises feitas com base nos dados introduzidos.

Criou-se outro *software* para guardar a informação contida nos pixels das imagens guardadas, segundo ficheiros de texto, de forma a poder ser lida por outros programas, nomeadamente de análise estatística.

A análise estatística foi utilizada como ferramenta de estudo das imagens, para posterior obtenção de um modelo matemático que permitisse a classificação dos pixels.

Relembra-se que a rapidez do algoritmo é um factor fundamental para a utilidade prática do mesmo. Assim, recorreu-se a um método estatístico que permitisse descrever os conjuntos de dados através de relações matemáticas simples, e como tal computacionalmente menos dispendiosas. Por estes motivos escolheu-se a análise discriminante linear de *Fisher*.

A grande quantidade de dados contidos em cada uma das imagens, levou a que fosse também necessário prestar atenção à quantidade de variáveis usadas pelo modelo. Assim, começou-se por efectuar análises com o mínimo de variáveis possível, de forma a se poder construir um modelo simples. O número de variáveis foi aumentado com o

avançar do estudo estatístico, até um ponto considerado o mais indicado segundo o fim a que este se destinava.

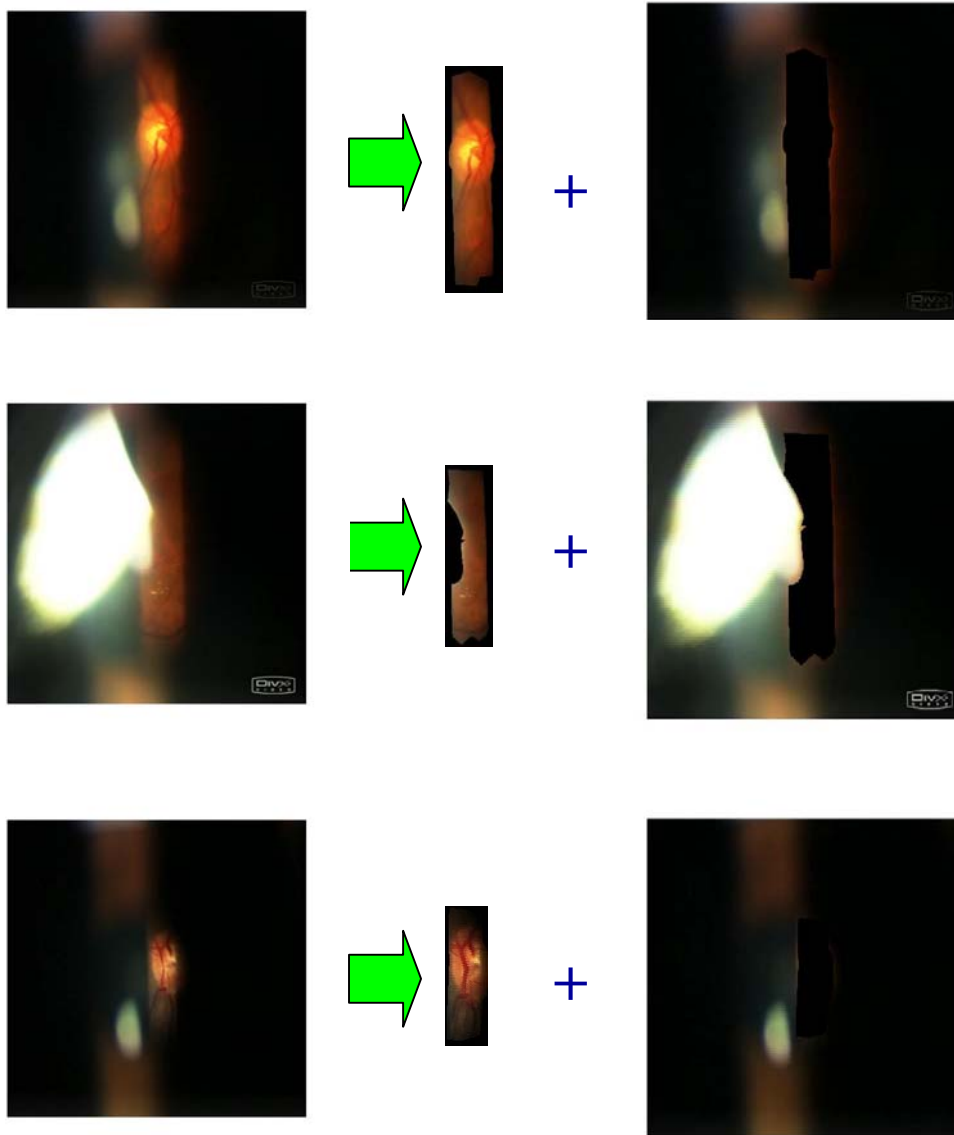


Figura 35 – Alguns exemplos das imagens analisadas, e respectivas imagens resultantes da segmentação manual.

Para a análise estatística foram utilizados dois tipos de *software*: uma *toolbox* do *Matlab* e um programa especializado para o tratamento estatístico de dados, o *SPSS*.

7.3.2. Interpretação dos resultados e escolha de um modelo

Tal como se esperava os canais G e B apresentaram uma elevada correlação, impossibilitando a criação de modelos contendo ambas as variáveis. Este facto definiu os moldes do estudo estatístico realizado.

A tabela que se segue faz um resumo das variáveis analisadas e dos resultados obtidos. Em anexo encontram-se mais dados relativos ao estudo estatístico efectuado.

Variáveis analisadas	Variáveis excluídas	Predição correcta dos elementos a rejeitar (%)	Predição correcta dos elementos a preservar (%)	Predição correcta global (%)
R, G e B	G	97,3	60,3	94,71
	B	99,4	77,6	97,87
R e Br		99,2	79,7	97,85
LogR, logG e logB		95,4	83,7	94,58
LnR, lnG e lnB		95,4	83,7	
R, G e Br		97,8	59,9	95,15
r, g e b		94,6	90,0	94,28
r, g, b e Br		93,8	92,6	93,72
R, r, g, b e Br		99,2	79,3	97,8
B, r, g, b e Br		99,1	80,6	97,8
G, r, g, b e Br		98,9	81,1	97,67
R, G e B		99,3	74,7	97,56

Tabela 3 - Resultados dos cálculos da predição para os modelos obtidos com a análise estatística efectuada.

R – canal vermelho; G – canal verde; B – canal azul; logR – logaritmo de base 10 de R; logG – logaritmo de base 10 de G; logB – logaritmo de base 10 de B; lnR – logaritmo de base neperiana de R; lnG – logaritmo de base neperiana de G; lnB – logaritmo de base neperiana de B; r – canal R normalizado; g – canal G normalizado; b – canal B normalizado; Br – intensidade media dos canais de cor RGB.

O tratamento estatístico teve apenas em consideração a cor individual de cada pixel, desprezando as vizinhanças. Desta forma, tornou-se necessário segmentar os vídeos captados, usando os modelos obtidos, com a finalidade de visualizar os resultados da sua aplicação.

È de referir que apenas 7% dos dados utilizados na análise foram classificados manualmente como pertencentes a retina, assim, o valor de predição real dos modelos não corresponde ao cálculo simples da média entre as duas taxas de predição.

Através do estudo efectuado foi possível determinar que, em imagens do fundo ocular predominam pixels com contribuições mais significativas de R e G, enquanto que os reflexos apresentam contribuições mais elevadas de G e B, relativamente a R. Isto sob um iluminante branco.

De entre todos os modelos obtidos salientam-se quatro, que melhor servem o objectivo pretendido: os modelos com as variáveis RG, RB, rgbBr, e o modelo com as variáveis RGB.

O primeiro é aquele que melhor elimina os elementos indesejáveis, permitindo restringir ao máximo a região de interesse, no entanto, tem o inconveniente de eliminar estruturas de coloração amarelada como o disco óptico e drusas, já que não tem em consideração a contribuição de B. A preservação destes elementos é importante, pois são efectivamente elementos do fundo ocular, podendo também servir como pontos de referência para o seguimento dos movimentos da retina. O grande pólo associado à sua equação matemática, leva a que este modelo apresente também uma fraca tolerância em imagens com uma iluminação muito baixa. O mesmo acontece como o modelo RB, que apresenta um desempenho semelhante, pecando pela ausência da variável G, que permite eliminar reflexos com valores mais elevados desta variável.

Para efeitos de detecção do guia laser, o modelo RG mostra ser um bom modelo de segmentação, pois permite um processamento rápido, restringindo as imagens do fundo ocular apenas aos elementos essenciais: a retina com os vasos sanguíneos e o guia laser. Contudo, este modelo rejeita estruturas como o disco óptico e drusas.

Uma vez que se pretende uma segmentação mais precisa, aplicável a um sistema para visualização de imagens em contínuo, continuou-se o estudo no sentido de melhorar o desempenho já conseguido. Este facto levou à procura de um modelo que englobasse informação de todos os canais de cor. Surgiu o modelo rgb.

Este é de todos os modelos estudados até à altura o mais conservador, reconhecendo a maior quantidade de pixels do fundo ocular. Contudo, preserva muitos vestígios luminosos de baixa intensidade, que formam regiões extensas sem informação útil. Este modelo mostrou-se também menos eficaz que os outros modelos, na eliminação de íris com tons acastanhados.

Na generalidade, as colorações encontradas no fundo ocular vão de tons mais avermelhados, passando por tons alaranjados, até tons amarelados. O castanho resulta de uma combinação de r, g e b, que se situa neste espaço entre o cor-de-laranja e o amarelo. Assim, a coloração acastanhada, é reconhecida pelo modelo como sendo

pertencente ao fundo ocular. A análise de imagens individuais, mostrou que esta cor é também uma das cores de entre as que se pode encontrar num fundo ocular.

O facto da maioria da população portuguesa ter olhos castanhos, faz com que na maioria dos casos as regiões da imagem preservadas após a segmentação sejam muito extensas, o que exige um maior processamento por parte dos módulos de detecção do guia laser e do módulo de seguimento ocular, abrandando estes processos.

Assim, tentou-se eliminar estas íris indesejáveis sempre que possível, já que não fazem parte da informação visual que se quer obter.

Quando se pretende aceder ao fundo ocular, a fenda luminosa emitida pela lâmpada de fenda é direccionada para a região interior à íris, a pupila. A luz reflectida pelo fundo ocular, volta a passar pela pupila, surgindo assim a íris numa região periférica em relação ao fundo ocular, e tendencialmente com uma menor intensidade luminosa, isto para as imagens adquiridas

Tendo em conta as imagens analisadas, torna-se evidente a influência da intensidade, nos modelos obtidos, explicando-se através desta o comportamento dos modelos RG e RB. Nestes a informação de intensidade associada às variáveis permite rejeitar grande parte dos pixels pertencentes a íris castanhas.

Uma análise mais pormenorizada revelou que nos dados classificados manualmente, havia grupos colorimetricamente semelhantes, mas que foram separados em grupos diferentes. É o caso dos pixels da mácula, que podem apresentar uma coloração acinzentada, comum em muitos reflexos. Este é mais um caso em que se mostra necessário saber a intensidade dos pixels relativamente a todo o conjunto.

O modelo rgbBr, veio apoiar esta ideia, uma vez que se verificou um aumento da correcta taxa de predição dos pixels pertencentes ao fundo ocular, e uma diminuição da taxa de reconhecimento de elementos a eliminar em comparação com o modelo rgb, devido à informação de intensidade média fornecida pela variável Br.

Ponderou-se então a construção de um modelo com as variáveis R, G e B.

Como já referido, não foi inicialmente possível construir um modelo com estas variáveis devido à grande correlação entre G e B.

Aumentou-se a significância dos dados relativos ao fundo ocular na análise estatística, por redução da quantidade dos dados respeitantes aos pixels que se pretende rejeitar. Desta forma conseguiu-se então criar o modelo procurado.

A sua predição e a análise dos vídeos segmentados, mostraram que o modelo tem um comportamento intermédio entre os modelos RG e RB. Apresentou uma maior

eficácia que o modelo RB, e maior tolerância a estruturas como a mácula, disco óptico e drusas que o modelo RG. No entanto, revelou ser um pouco mais tolerante no que toca a íris acastanhadas, do que este último modelo.

Apesar de apresentarem valores bastante elevados de predição correcta, nenhum dos modelos individuais tem um comportamento totalmente satisfatório. Isto deve-se ao facto dos modelos serem generalista, procurando responder de forma correcta ao máximo possível de pontos.

Ponderou-se ainda a utilização de outro tipo de modelos matemáticos, no entanto esta ideia foi abandonada face ao seu elevado peso computacional.

Algoritmos criados mais tarde revelaram que de facto é necessário ter em linha de conta a intensidade média dos canais de cor da imagem do fundo ocular, ou seja, é necessário fornecer dados relativos a toda a região de interesse, por forma a se saber se o pixel em análise tem maior ou menor intensidade que a globalidade dos pixels do fundo ocular.

Quando os vídeos foram adquiridos a intensidade luminosa não foi calibrada, por isso as várias imagens usadas na análise estatística apresentam diferentes valores de intensidade. Através da análise individual das imagens verificou-se que para uma iluminação mais elevada, o fundo ocular vai surgindo com cores cada vez mais próximas de reflexos. B assume um valor mais elevado relativamente aos outros canais de cor, distorcendo a coloração.

O modelo rgb, é aquele que realmente faz uma discriminação da cor, daí se ter obtido com ele um dos valores mais elevados de predição correcta para os pixels da retina. O modelo rgbBr vem acrescentar um factor de correcção relativamente à intensidade média, que permite incluir a mácula no grupo de elementos a preservar. Observando os vídeos segmentados com o modelo rgbBr, verifica-se que muitos dos pixels não pertencentes ao fundo ocular, e que não foram eliminados contêm um valor de intensidade muito baixo. Estes são reflexos difusos, elimináveis por *thresholding* (passa-alto) dos canais de cor, caso a intensidade luminosa das imagens fosse constante.

A calibração da intensidade luminosa da fenda, ao permitir aplicar à imagem um *threshold* genérico, tornaria os modelos rgbBr e RGB, bons modelos de segmentação, mesmo tendo em conta as variações de iluminância (intensidade luminosa por unidade de superfície) na região de interesse, que não se podem eliminar.

Apesar de tudo, a calibração da intensidade luminosa não foi um factor disponibilizado nos vídeos utilizados.

7.4. Algoritmo de segmentação

O algoritmo de segmentação pode ser dividido em duas partes. A primeira diz respeito à segmentação de cor através de modelos lineares de classificação, obtidos com a análise estatística, e a segunda parte diz respeito a um processamento morfológico que visa eliminar pixels ou conjuntos de pixels remanescentes da execução da primeira parte.

Apesar de não se ter obtido um modelo linear de classificação óptimo, estes mostraram ser aplicáveis e capazes de uma boa segmentação.

Não havendo um modelo totalmente satisfatório, o estudo foi direccionado para o desenvolvimento de uma solução adaptável às condições disponibilizadas, ou seja, intensidade luminosa muito variável entre imagens.

A informação de intensidade média da região de interesse associada à informação de cor é a chave para discernir entre fundo ocular ou outros.

Havendo a possibilidade de se captar fundos oculares muito iluminados, resta referir que é de todo o interesse que sejam também preservados, já que contêm informação relevante. Para além do mais, é previsível que o uso de um guia laser venha também introduzir reflexos, sendo necessário que o algoritmo de segmentação lhes apresente alguma tolerância para que este não seja eliminado. Quando o interruptor do biomicroscópio é premido, verifica-se também um aumento da intensidade luminosa da imagem.

As imagens obtidas com o biomicroscópio, têm uma intensidade muito variável e contêm uma grande e imprevisível quantidade de reflexos. Desta forma, não é possível determinar uma intensidade média do fundo ocular, sem antes se eliminarem os reflexos mais óbvios e significativos.

O facto da fenda poder assumir dimensões muito variadas impede que esta média seja calculada usando conhecimentos sobre a posição e forma da região de interesse, à semelhança do que é feito com imagens adquiridas com câmaras de fundo, já que o tamanho da região de interesse na imagem varia consoante as dimensões da fenda e a focagem do sistema de observação. A acrescentar a isto há ainda o facto de uma fenda de menores dimensões abranger uma zona mais reduzida do fundo ocular, resultando daí que a imagem possa ser constituída maioritariamente por elementos mais

reflectores como o disco óptico, ou menos reflectores como a mácula, que influenciam fortemente esse cálculo.

Sabe-se que sob iluminação branca o disco óptico e as drusas surgem com uma coloração amarelada. Os reflexos podem assumir qualquer forma, tamanho ou cor, no entanto os mais frequentes assumem valores muito elevados de B e/ou G em relação a R, sendo eficazmente eliminados pelos modelos RG e RB ou pelo modelo RGB.

Face à importância da velocidade de execução da tarefa de segmentação, não é possível aprofundar muito o processamento na imagem. Assim, desenvolveu-se um algoritmo que recorre a uma pós-segmentação para melhorar os resultados de uma primeira segmentação genérica, usando modelos semelhantes aos RG, RB e RGB.

Muitas vezes recorre-se ao *region growing* para melhoramento da segmentação das imagens, no entanto, esta técnica interactiva é demasiado lenta para a aplicação que se pretende criar.

Apesar do que possa parecer à primeira vista, a pós-segmentação faz sentido.

A imagem final tem que ser apresentada no espaço RGB, pois este é o espaço de cor mais natural para o ser humano. O facto de se poder utilizar este espaço nos modelos de segmentação é de grande importância, pois permite evitar cálculos de transformação noutros espaços de cor. Por outro lado, o tempo necessário para execução de uma divisão ou uma multiplicação é cerca de oito vezes superior ao tempo necessário para executar uma adição ou uma subtracção. Assim, uma vez que no caso do modelo rgbBr, é necessário recorrer a divisões entre matrizes para transformação do espaço de cor, e sendo que o uso de cálculos exponenciais é ainda mais lento, a combinação de vários modelos semelhantes aos modelos RG, RB e RGB, permite introduzir uma correcção à segmentação conseguida com estes, num tempo de processamento inferior ao tempo de processamento exigido para a aplicação de qualquer outro modelo mais complexo, obtendo-se um melhor resultado no final.

Esta abordagem tem contudo um senão: para que a pós-segmentação funcione correctamente é necessário que haja realmente na imagem retina reconhecível pelo primeiro modelo de segmentação, e que esta represente uma porção significativa dos pixels não eliminados. Só assim as médias dos canais podem reflectir a sua coloração.

O algoritmo foi ainda pensado de forma a permitir a eliminação das tão indesejadas íris acastanhadas. Estas surgem nas imagens numa região periférica em relação ao fundo ocular, apresentando muitas vezes nas zonas de transição com o fundo ocular, uma região onde R desce abaixo da média da cena iluminada, e B assume

valores relativamente mais elevados. Separando a íris do fundo ocular, torna-se possível através de técnicas morfológicas simples, eliminá-la conjuntamente com os vestígios remanescentes.

7.4.1. Segmentação da cor

Neste algoritmo são conjugados vários modelos a fim de se melhorar o desempenho do conjunto.

O grande tamanho das imagens (aproximadamente 512 x 512 pixels) faz destas um conjunto de dados bastante extenso (512 linhas x 512 colunas x 3 canais de cor = 786432 pontos). O primeiro passo do algoritmo de segmentação consiste na criação de uma imagem redimensionada a partir da imagem RGB inicial, sendo que a imagem inicial é sempre preservada. O processamento de imagens de menores dimensões permite aumentar a velocidade do algoritmo. Ao se reduzir a imagem inicial para uma imagem com 131 linhas e 131 colunas, os dados são diminuídos aproximadamente 45 vezes (cerca de 15 vezes para cada um dos três canais de cor). O redimensionamento é feito usando o método *nearest neighbor*.

A redimensionalização da imagem leva cerca de 1/3 do tempo total do processamento, no entanto permite reduzir significativamente o tempo gasto noutros passos do algoritmo, possibilitando a execução de uma análise muito mais aprofundada num menor espaço de tempo.

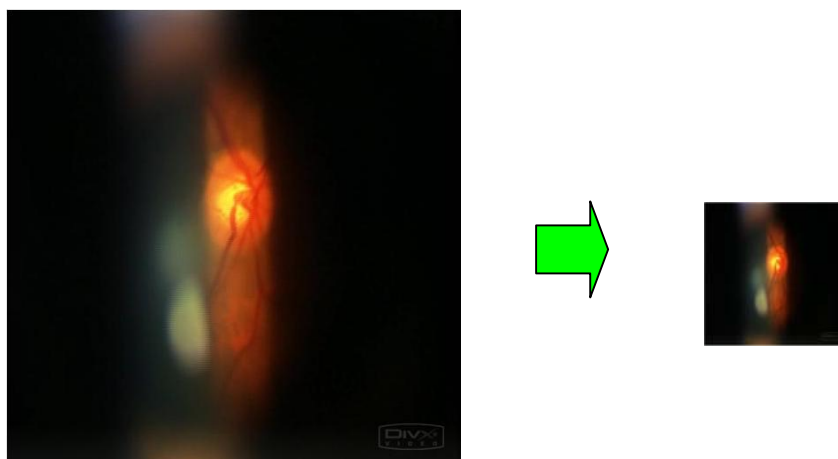


Figura 36 – Primeiro passo do algoritmo de segmentação: redução das dimensões da imagem original.

Segue-se a aplicação dos modelos discriminantes lineares. Como já referido, os canais de cor vão ser variáveis de equações lineares.

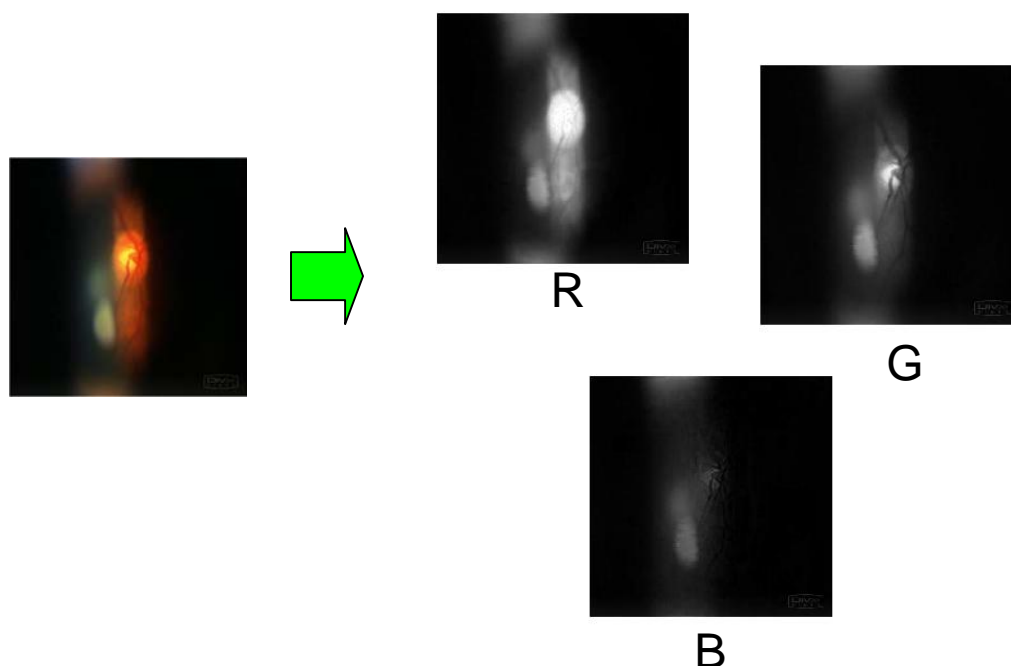


Figura 37 – A imagem original e os seus três canais de cor.

Esta primeira segmentação da cor é generalista, ou seja, visa eliminar o máximo de reflexos e artefactos, em qualquer condição de iluminação, preservando o tanto quanto possível o fundo ocular e seus constituintes. Este passo serve apenas para que se possa calcular as médias de R, G e de B do fundo ocular, que nos dão informação acerca da iluminação da cena. Estes dados serão posteriormente utilizados para que se possa executar uma segmentação mais minuciosa, discernindo entre o fundo ocular e verdadeiros reflexos. Como já referido, os reflexos podem assumir qualquer cor, mas em última instância a sua intensidade relativamente ao conjunto permite distingui-los daquilo que pretendemos preservar.

A utilização passo-a-passo de várias equações, permite melhorar o desempenho individual de cada modelo na classificação de certas características aumentando a sensibilidade do conjunto.

Nesta primeira fase são aplicadas duas equações. A primeira usando as variáveis R e B, e a segunda as variáveis R e G.

A primeira equação visa a identificação de reflexos em que a relação de R com B, é probabilisticamente mais frequente em reflexos e artefactos que num fundo ocular.

A equação é definida por:

$$g(x) = R - 0.8837 \times B - 46.904$$

Esta equação permite aplicar um *threshold* superior e um inferior que elimina pixels com valores muito baixos de intensidade e alguns reflexos que saturam os sensores da câmara.

O resultado da equação é uma matriz bidimensional com o mesmo número de linhas e colunas que a imagem redimensionada. Esta matriz contém as classificações para cada pixel. Os pixels com classificação maior que 0 são então considerados como pertencentes ao fundo ocular e os restantes rejeitados. Desta comparação resulta uma máscara binária em que os pixels com classificação menor ou igual a 0 assumem o valor 0, e o valor 1 caso contrário.

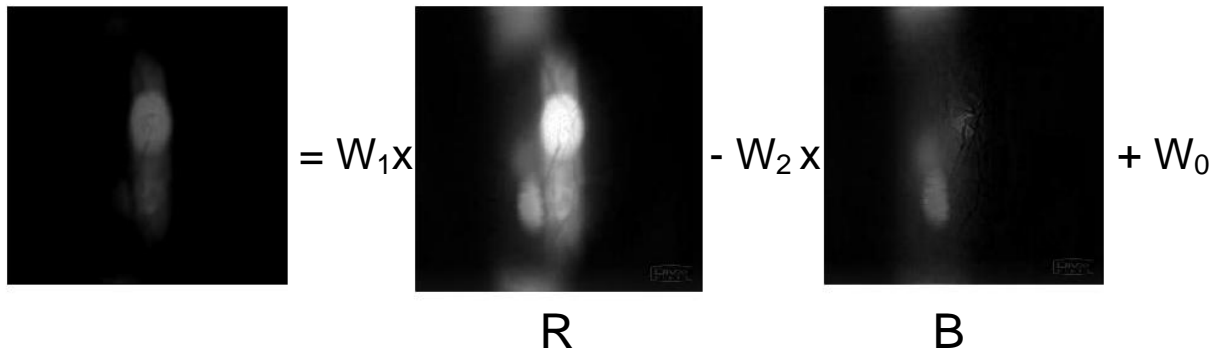


Figura 38 – A aplicação do modelo RB e o seu resultado. Na imagem do resultado da classificação os valores iguais ou menores que 0 surgem a preto.

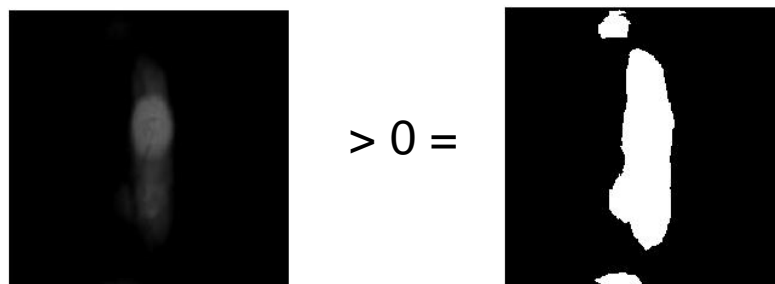


Figura 39 – Máscara binária resultante do *thresholding* da matriz que contém as classificações.

A segunda equação, é definida por

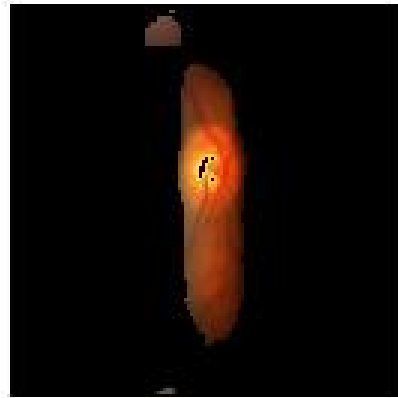
$$g(x) = R - 0.986 \times G - 29.888$$

, e vem introduzir informação do canal G, que permite eliminar pixels em que a relação entre R e G é estatisticamente mais frequente em reflexos que na retina. Esta equação define um limite superior próximo de $R=G+30$, e um limite inferior próximo de $R>30$.

A multiplicação da máscara binária obtida com a equação anterior, pela matriz R, permite desta forma atribuir $R = 0$ aos pixels que se pretende eliminar. Assim, estes pixels assumem uma classificação negativa em que R é dominante, pois G é sempre maior ou igual a 0. Desta forma, para que os resultados da aplicação do primeiro modelo sejam passados à análise seguinte, basta apenas usar como variável R na nova equação, a matriz resultante da multiplicação do canal R da imagem inicial pela máscara binária.



(a)



(b)

Figura 40 – Segmentação inicial: (a) imagem que resulta da aplicação do primeiro modelo; (b) imagem resultante da aplicação do segundo modelo.

Eliminados todos os reflexos, é agora possível calcular uma média dos canais de cor. Estes valores vão permitir que se execute uma nova segmentação adaptada às

condições de iluminação da imagem. Desta forma, preservam-se todos os elementos do fundo ocular que se encontrem dentro de uma determinada gama de intensidade.

A média de intensidade de cada canal para o fundo ocular é calculada através de

$$I_c = \frac{\sum_{i=1}^{m \times n} i_c}{\sum_{i=1}^{m \times n} x_i}$$

, onde I_c é a intensidade média do canal c , i_c é a intensidade do pixel i , x_i é o valor do pixel i da máscara binária, e m e n são as dimensões da máscara.

Uma vez obtida informação acerca da intensidade dos canais de cor e da coloração do fundo ocular, pode-se agora partir para uma segmentação de cor mais refinada.

Esta segmentação envolve a aplicação de duas equações lineares que utilizam as variáveis R, G, B e as médias da intensidade de cada canal de cor.

A segmentação que diz respeito às variáveis R e B, é uma boa segmentação generalista no que toca a estas variáveis. Para não se perder esta informação, e aproveitando cálculos já efectuados, vai-se utilizar a matriz de R obtida a partir da aplicação do primeiro modelo de segmentação.

Os dois modelos de segmentação que se seguem funcionam como *thresholds* inferior e superior respectivamente.

O primeiro modelo é definido por

$$g(x) = (R - 0.2 \times I_R) - (I_B \times 3.5/2 - G) \times 0.806 - B \times 0.687 + 12.225$$

, e o segundo por

$$g(x) = (R - I_R \times 0.15) - (G - 1.3 \times I_B) \times 0.787 - (B - 0.6 \times I_B) \times 1.153$$

Da sua aplicação resultam duas máscaras binárias que são conjugadas numa só, em que os pixels com valor 1 correspondem aos pixels não eliminados pelos modelos de classificação utilizados.

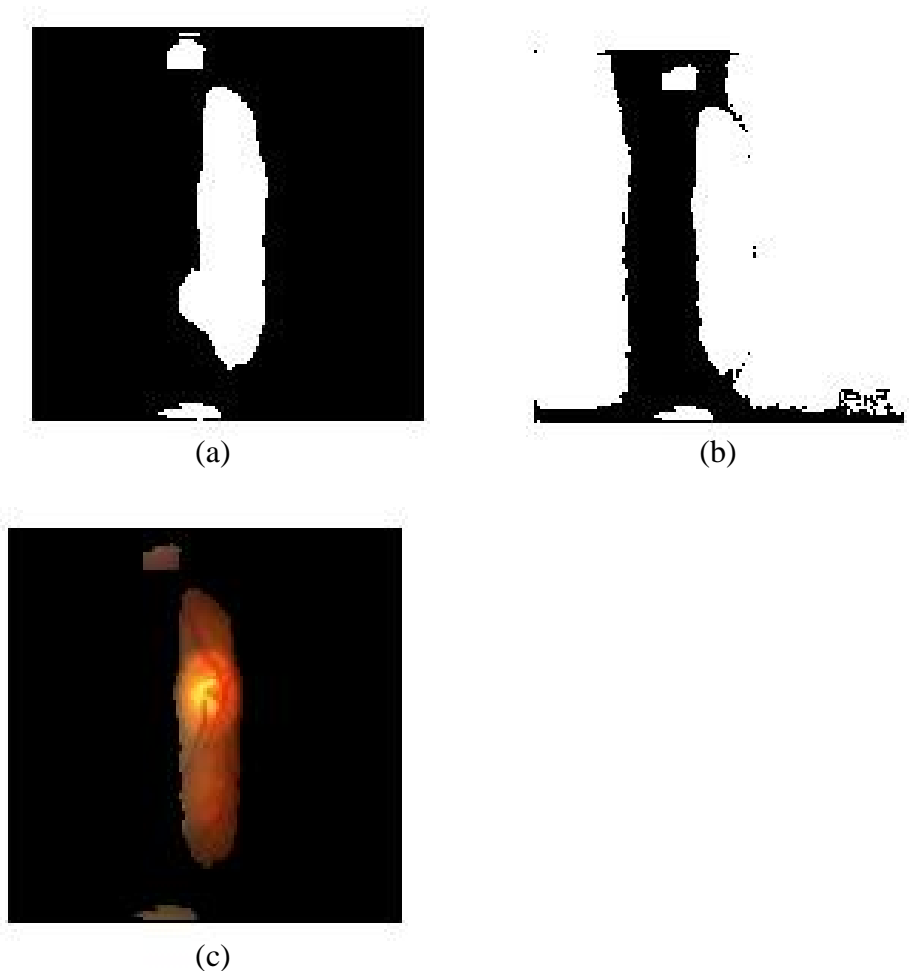


Figura 41 - Pós-processamento e seu resultado: (a) máscara binária resultante da aplicação da primeira equação de pós-segmentação; (b) máscara binária resultante da aplicação da segunda equação de pós-segmentação; (c) imagem resultante da segmentação da cor.

7.4.2. Segmentação morfológica

Após a segmentação da cor, a imagem pode preservar ainda vários grupos de pixels isolados, sendo comum a preservação de vestígios pertencentes a íris ou pedaços de pele, por parte dos modelos de cor utilizados.

A eliminação destes grupos de pixels indesejados, é aqui feita com recurso a informação de intensidade e morfologia dos grupos.

Considerou-se que o grupo de pixels pertencentes ao fundo ocular é constituído por um conjunto de pixels contínuo e verticalmente alongado.

Caso haja vestígios de pele ou íris, estes grupos surgem numa posição periférica à região de interesse e com uma intensidade mais baixa, já que a fenda luminosa é apontada para a região interior à íris.

A presença de uma grande quantidade de grupos residuais com poucos pixels pode tornar o processo de análise lento. Uma forma rápida que se encontrou para a sua eliminação, foi pelo cálculo da quantidade de pixels por coluna. Foram então somadas as colunas da máscara binária obtida no final da pós-segmentação da cor, e definidas como limites laterais da região a processar, a primeira e últimas colunas com mais do que 1/6 do número total de linhas da imagem. Este passo permite definir uma janela com 1/3 da imagem total ou menos na maioria dos casos (a fracção típica da imagem iluminada pela fenda). Isto com um tempo de processamento desprezável. Caso a imagem resultante tenha menos que 5% do número total de colunas da imagem redimensionada o processamento termina, devolvendo uma imagem com as dimensões da imagem inicial com $R=G=B=0$.

Se esta condição não se verificar, então as regiões contidas na imagem são sujeitas a uma análise mais aprofundada.

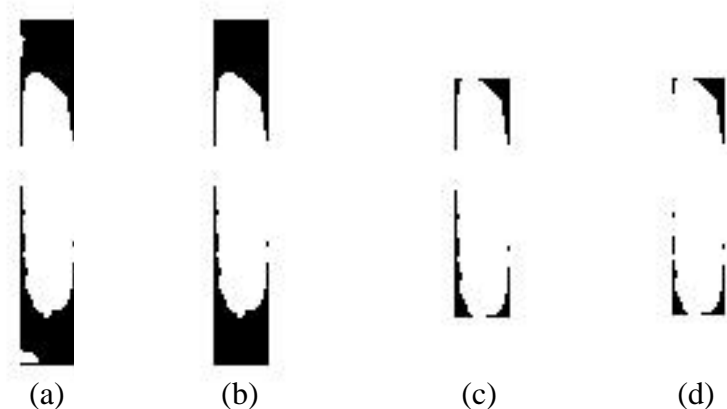


Figura 42 – Máscaras binárias resultantes do processamento morfológico: (a) ajuste horizontal da janela à cena iluminada pela fenda; (b) após rejeição de grupos de pixels pela sua análise individual; (c) ajuste vertical da janela ao grupo preservado; (d) ajuste horizontal da janela ao grupo preservado.

Segue-se uma análise individualizada dos grupos de pixels preservados. Para tal é necessário identificar os grupos da máscara binária redimensionada no passo anteriormente. Para tal, aplicou-se uma etiquetagem *4-neighbor*, que é a mais simples e rápida.

É também criada uma imagem de intensidade média, somando as três matrizes dos canais de cor da imagem colorimetricamente segmentada e horizontalmente delimitada.

Para cada uma das regiões detectadas são somados os correspondentes pixels da matriz de intensidade, e determinado o valor máximo do somatório. De cada vez que se encontra um novo máximo este é guardado, conjuntamente com a etiqueta da região à qual pertence.

No final é apenas preservado o grupo ao qual corresponde o maior valor calculado.

Este método permite obter um valor que é influenciado não só pela dimensão do grupo na vertical, mas também pela intensidade dos pixels. Desta forma, o algoritmo torna-se mais robusto que uma simples medição do comprimento da região ou o cálculo da intensidade média das colunas, conseguindo preservar regiões pertencentes ao fundo ocular de menores dimensões que outros vestígios, mas que devido a serem o alvo principal da iluminação apresentam uma intensidade luminosa superior. Por outro lado, este algoritmo dificilmente é enganado por pequenos grupos de elevada intensidade média que eventualmente não tenham já sido eliminados.

Segue-se um ajuste vertical da janela ao conjunto de pixels preservado, definindo como margens verticais, a primeira e a última linha da máscara binária, que contêm mais que $1/5$ do número total de pixels de uma linha.

De novo executa-se um ajuste horizontal da janela. Desta vez, pretende-se ajustar a janela às margens horizontais do grupo.

Uma vez adaptada a janela à região de interesse, interessa redimensionar a máscara binária de forma a se poder devolver a imagem segmentada nas suas dimensões originais.

Começa-se por definir uma nova matriz com o mesmo número de linhas e colunas da imagem original redimensionada, e projecta-se a janela com a máscara binária final nessa matriz. Esta matriz é redimensionada segundo as linhas e as colunas ao tamanho da imagem inicial, e multiplicada por cada canal de cor da imagem original. Desta forma, os pixels preservados pela segmentação mantêm o seu valor, enquanto que os rejeitados assumem o valor 0.

No final, é devolvida a imagem segmentada nas suas dimensões originais, pronta para ser visualizada. São também devolvidos os valores das linhas e colunas que

permitem definir uma janela em redor da região de interesse, para o caso do processamento sobre esta ainda não ter terminado.

7.4.3. Resultados da segmentação

Da aplicação do algoritmo desenvolvido aos vídeos iniciais, obteve-se a sua versão segmentada. Apesar dos resultados poderem ser visualizados através destes, seleccionou-se um conjunto de imagens para se apresentar aqui o melhoramento conseguido com a segmentação dinâmica, face a uma segmentação estática com apenas um dos melhores modelos individuais obtidos, o modelo RGB.

O algoritmo dinâmico apresenta uma maior tolerância a variações da intensidade luminosa, conseguindo preservar elementos do fundo ocular, que os modelos individuais rejeitam.

No entanto, o algoritmo apresenta ainda problemas de reconhecimento quando surgem na imagem dois factores associados: grupos de pixels com uma coloração não eliminável com os modelos gerais, e ausência de porções do fundo ocular com dimensões significativas para o cálculo da intensidade média dos canais de cor. Apesar destes serem ainda passíveis de eliminação, optou-se por não aprofundar o processamento de forma a não se aumentar o tempo do mesmo.

O tempo total de processamento do algoritmo de segmentação (segmentação da cor e segmentação morfológica) é de aproximadamente 0.1 segundos para imagens RGB de 514 x 513 pixels.

O facto de se ter usado uma linguagem interpretada de programação, leva ainda a pensar que será possível obter uma maior velocidade de processamento recorrendo a linguagens de mais baixo nível como C/C++.

É também de referir que o algoritmo desenvolvido poderá ter o seu desempenho comprometido em imagens obtidas com uma iluminação que não seja branca, sendo necessário ajustar a cor destas imagens através de algoritmos de constância de cor.

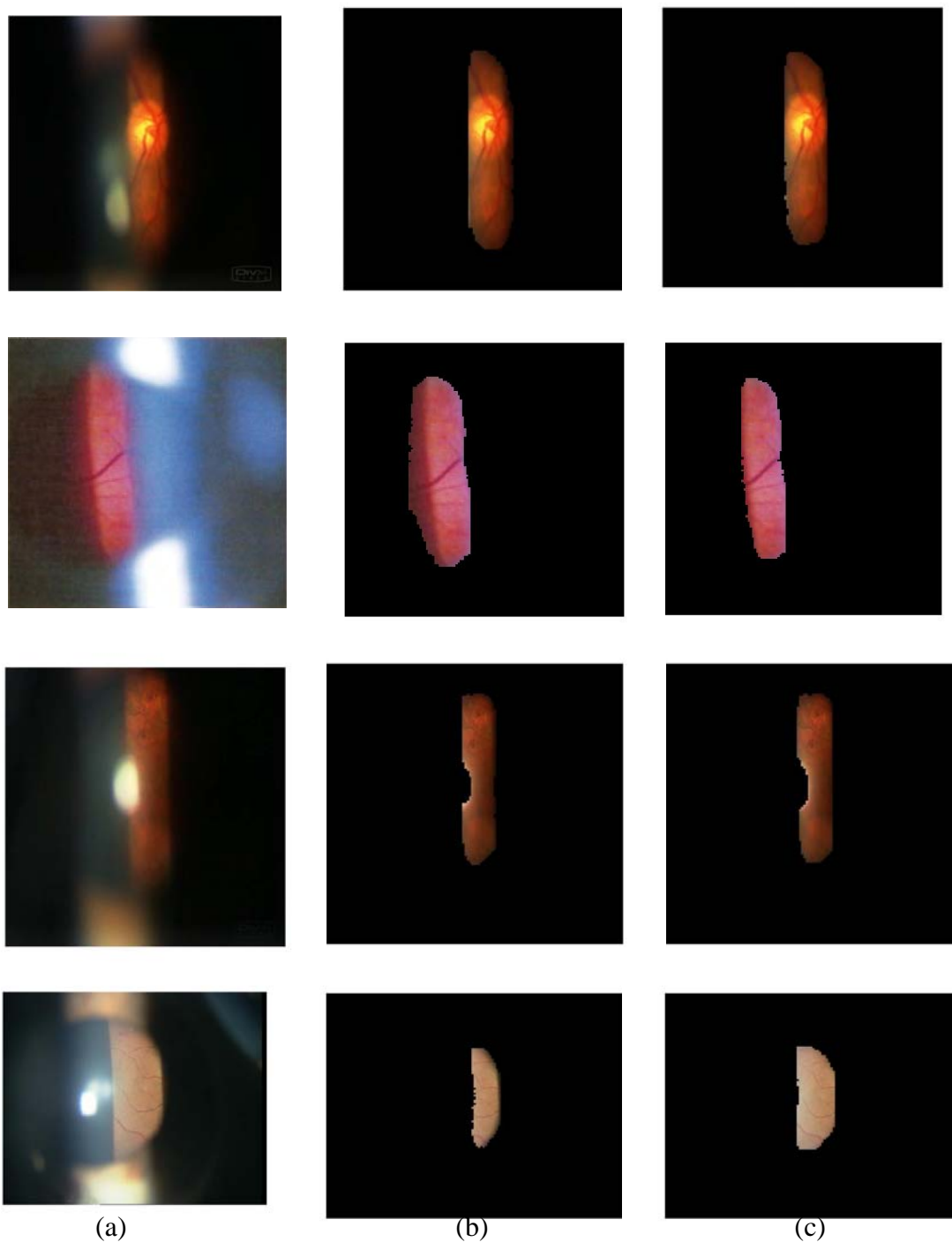


Figura 43 – Exemplos da segmentação conseguida: (a) coluna contendo as imagens originais; (b) coluna contendo as imagens resultantes da segmentação de cor estática com o modelo RGB; (c) coluna contendo as imagens resultantes da segmentação de cor dinâmica.

8. Implementação do algoritmo de segmentação no *Visionário*

Nesta etapa pretendeu-se incluir uma nova funcionalidade no *software Visionário*.

O *software Visionário* surgiu como resultado de um projecto de desenvolvimento de uma aplicação de captura, manipulação e armazenamento de imagem, adaptado ao *hardware* indicado para este projecto.

Este *software* foi construído em *Visual C++*, e faz uso de várias bibliotecas já existentes. Entre estas encontra-se o *kit* de desenvolvimento *IFC – Imaging Foundation Classes* e o *kit* de desenvolvimento *CxImage*.

O *kit* de desenvolvimento *IFC* da *Coreco Imaging* consiste num conjunto de classes em *C++* para desenvolvimento de aplicações usando a placa de captura de imagem *Frame Grabber Coreco Imaging PC-RGB*. O correcto funcionamento do *Visionário* implica a presença deste. Sendo que foi utilizada a versão 5.5 deste *kit*. [34]

O *kit* de Desenvolvimento *CxImage* é um conjunto de classes *opensource* para manipulação de imagem.

Nesta fase pretendeu-se implementar uma funcionalidade que permitisse segmentar as imagens captadas, com um dos modelos discriminantes obtidos, em modo contínuo pela câmara e visualizar os resultados.

Uma vez que a linguagem de programação *C++* constituiu uma novidade no âmbito da licenciatura, foi necessário adquirir os devidos conhecimentos de programação antes de se executar esta tarefa.

O facto de se ter que implementar um novo módulo em *software* já existente obrigou também a que se estudasse o mesmo.

O *Visionário* tem um módulo para captura e visualização de imagem em directo. Este módulo faz uso de funções da biblioteca *IFC – Imaging Foundation Classe* e de classes disponibilizadas pela *Microsoft* para desenvolvimento de aplicações, as *Microsoft Foundation Classes*.

O pouco tempo disponível para a realização da tarefa e a falta de conhecimentos, levou a que se tivesse de começar a tarefa quase em simultâneo com o estudo.

Tentou-se então adicionar a nova funcionalidade neste módulo, com os recursos disponíveis.

Adicionou-se ao interface gráfico um novo botão que permite activar a nova funcionalidade.

No entanto, apenas foi incluída uma funcionalidade que é activada quando é premido o botão criado, e que permite: interromper a ligação estabelecida com a câmara de video, configurar a câmara para o modo de aquisição de imagens no espaço de cor RGB, captar uma fotografia para o *frame buffer*, estabelecer uma nova ligação com a câmara e mostrar a fotografia guardada no *frame buffer* numa janela própria para o efeito.

Ainda se chegou a implementar o modelo de segmentação, contudo, o facto das livrarias usadas não estarem documentadas com um nível de profundidade desejado, de não se poder “abrir” as funções (o que facilitaria a compreensão do seu funcionamento interno), e a falta de experiência na manipulação do *frame buffer* e ponteiros em C++, levaram a que não se conseguisse mostrar a imagem segmentada final.

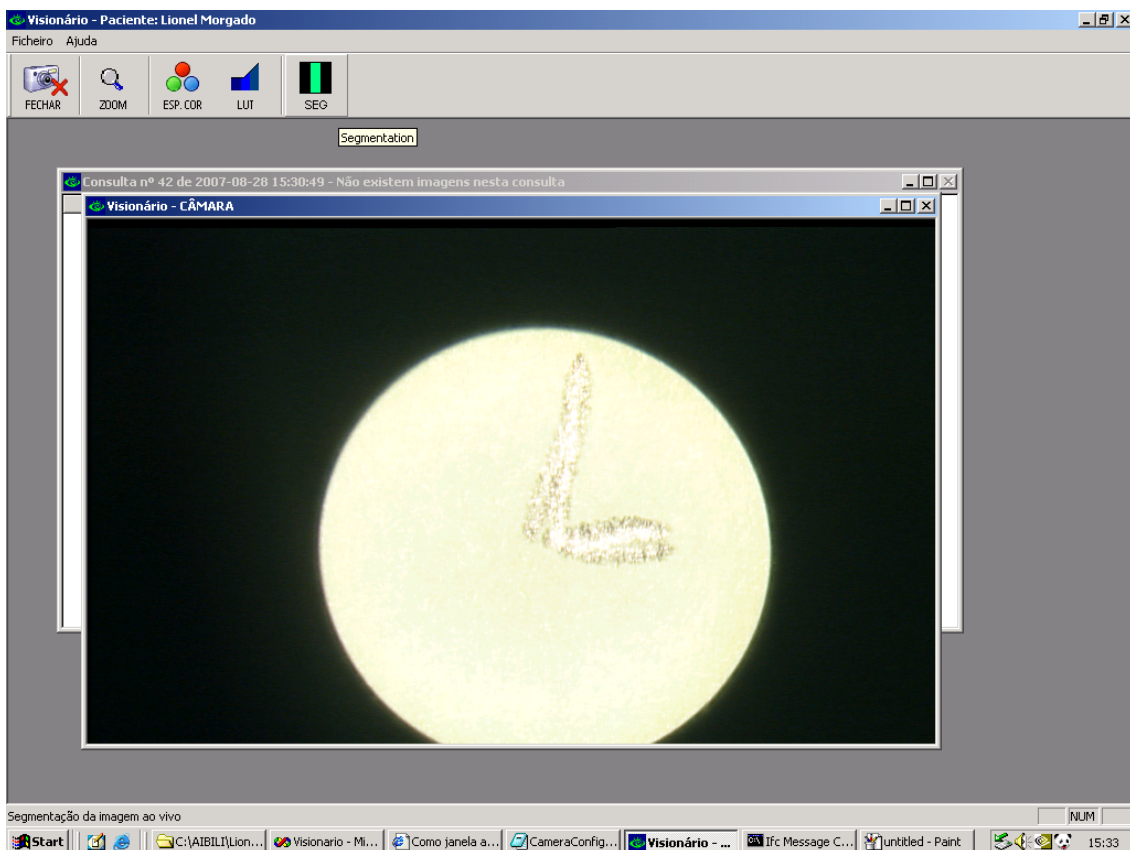


Figura 44 – Vista geral da do *Visionário* com o novo botão e com uma imagem captada.

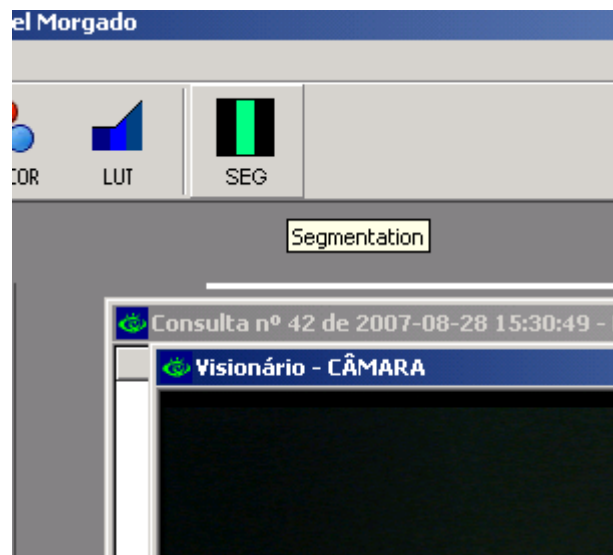


Figura 45 – Pormenor do botão implementado no *Visionário*.

9. Conclusões

O objectivo inicial deste projecto era desenvolver um sistema para determinação da posição de disparos laser em imagens do fundo ocular obtidas através de um biomicroscópio com lâmpada de fenda, sendo que esta posição pode ser determinada através da posição de um guia laser em imagens do fundo ocular no momento do disparo.

Não havendo imagens com o referido guia laser, construiu-se um algoritmo de detecção baseado em modelos e suposições teóricas. O teste do algoritmo em imagens com um modelo do guia laser marcado artificialmente, mostrou que a detecção é eficaz, e que foi atingida uma velocidade de processamento considerável (aproximadamente 0,2 segundos). Resta contudo estudar a eficácia do mesmo em casos reais.

Através do estudo estatístico foi possível obter modelos de classificação para o fundo ocular com taxas de predição bastante elevadas. As técnicas de processamento morfológico utilizadas, permitem ainda aumentar a qualidade das imagens segmentadas, através da eliminação de pequenas regiões que não são classificadas de forma correcta pelos modelos utilizados. Desta forma, o algoritmo de segmentação apresenta um desempenho bastante satisfatório face à fraca qualidade dos vídeos analisados. De todas as formas, a averiguação da verdadeira eficácia do algoritmo terá que passar por um estudo que englobe mais imagens, e de diferentes pacientes.

O algoritmo de segmentação das imagens do fundo ocular adquiridas através do biomicroscópio com lâmpada de fenda apresenta também uma velocidade de processamento considerável (aproximadamente 0,1 segundos). Apesar de se ter atingido a meta estabelecida em relação a este parâmetro, a velocidade de processamento encontra-se ainda um pouco afastada de um valor desejável de aproximadamente 20 a 30 imagens por segundo. O aumento da sua velocidade de processamento passará pela sua implementação em linguagens de mais baixo nível como C/C++, e pela intervenção directa sobre o *hardware* (câmara de vídeo, sistema de iluminação, *frame grabber*, etc).

Para se atingir estas velocidades de processamento foi necessário ter em atenção todo o tipo de dados envolvidos nos cálculos, e o tipo de cálculos efectuados, procurando-se sempre implementar alternativas tendencialmente mais simples matematicamente, mas capazes de executar os processamentos requeridos de forma

rápida. Assim, pesaram-se sempre as vantagens de técnicas mais lentas mas com melhor desempenho, com as vantagens de técnicas mais rápidas, mas com pior desempenho.

Os módulos criados advêm de um estudo bastante exaustivo, ao longo do qual foram criados vários algoritmos para teste e estudo de diferentes abordagens e técnicas. Estes não são aqui apresentados, já que são apenas parte do processo de procura das melhores soluções para os problemas propostos.

Apesar do grande trabalho, esforço e sacrifício dispendidos ao longo de todo o ano lectivo, o fim do período disponível para realização das tarefas propostas, levou a que não se conseguisse terminar a implementação no *Visionário* do modelo de segmentação e visualização em contínuo das imagens adquiridas, tal como seria desejável.

Durante o decorrer da disciplina de projecto, foi ainda criado *software* para segmentação manual de imagens, conversão de imagens em ficheiros de texto, manipulação e criação de vídeos.

Apesar daquilo que não foi conseguido, e das dificuldades encontradas ao longo da disciplina de projecto, faz-se um balanço bastante positivo, tanto a nível dos conhecimentos adquiridos, como do trabalho desenvolvido.

Referências bibliográficas

- [1] R.M. Berne, M.N. Levy, B.M. Koeppen, B.A. Stanton. *Physiology*, Mosby 2004;
- [2] P. Posel, E. Schulte, *Sobotta*. Marbán 2004;
- [3] Grand, Y.; *Óptica fisiológica- Tomo I: el ojo como instrumento optico* ; Asociación Amigos Escuelas Universitarias de Óptica; 1991;
- [4] Blumenkranz, M. S.; Yellachich, D. ; Andersen, D. E. ; Wiltberger, M. W. ; Mordaunt, D. ; Marcellino, G. R. ; Palanker, D. ; *Semiautomated Patterned Scanning Laser For Retinal Photocoagulation*; *Retina, The Journal Of Retinal And Vitreous Disease*, Volume 23, Numero 3, 2006;
- [5] S. Bonnel, S. Mohand-Said, J-A. Sahel. *The aging of the retina*; 2003;
- [6] G. Blankenship, *Red krypton and blue-green argon panretinal laser photocoagulation for proliferative diabetic retinopathy: a laboratory and clinical comparison*. Tr. Am. Ophth. Soc. vol. LXXXIV, 1986;
- [7] M. Laursen, F. Moeller, B. Sander, A. Sjoelie, *Subthreshold micropulse diode laser treatment in diabetic macular oedema*. Br. J. Ophthalmol. 2004;
- [8] Gonzalez, R.; Woods, R.; *Digital Image Processing*, Prentice-Hall Inc, 1992;
- [9] Dawson, B. ; *Tutorials - Applying Realtime Machine Vision*; Coreco Imaging;
- [10] Neoh, Hong Shan; Hazanchuk, Asher; *Adaptative edge detection for Real-Time Video Processing using FPGAs*;
- [11] Neoh, H. S. ; Hazanchuk, A.; *Adaptive Edge Detection for Real-Time Video Processing using FPGAs*; Altera, The Programmable Solutions Company, 2005;
- [12] Kuivalainen, M.; *Retinal Image Analysis Using Machine Vision*, 2005;
- [13] Gabbur, P. ; *Quantitative Evaluation of Feature Sets, Segmentation Algorithms and Color Constancy Algorithms Using Word Prediction*; 2003;
- [14] Asmuth, J. ; Madjarov, B. D. ; Sajda, P. ; Berger, J. W. ; *Mosaicking and enhancement of slit lamp biomicroscopic fundus images*; Br. J. Ophthalmol.; 2001;
- [15] TAKAGI, *Instruction Manual: Slit Lamp microscope SM-70, SM-30, SM-10*;
- [16] <http://www.academy.org.uk/lectures/eperjesi5.htm>;
- [17] manual - SONY, 3CCD Color Video Camera ExwaveHAD DXC-390/390P;
- [18] lâmina - FRAME GRABBER CORECO PC-RGB;
- [19] Smith, L.I. ; *A tutorial on Principal Component Analysis*; 2002;
- [20] Kuttler, K. ; *Introduction To Linear Algebra*; 2007;
- [21] Meyer-Bäse, A. ; *Pattern Recognition for Medical Imaging*; Elsevier; 2004;

- [22] Shlens, J. ; *A Tutorial on Principal Component Analysis*, 2005;
- [23] Iocchi, L. ; *Robust Color Segmentation through Adaptive Color Segmentation Transformation*; Robocup Symposium 2006;
- [24] Leach, Sónia; *Singular Value Decomposition – A Primer*;
- [25] Martins, M. T. F. O. ; *Álgebra Linear e Geometria Analítica*; 2002;
- [26] Chen, T. ; Chung, K. ; *An Efficient Randomized Algorithm for Detecting Circles*; Computer Vision and Image Understanding; 2001;
- [27] Cooper, G. R. J. ; Cowan, D. R. ; *The detection of circular features in irregularly spaced data*; Elsevier; 2003;
- [28] Madjarov, B. D. ; Berger, J. W. ; *Automated real time extraction of fundus images from slit lamp fundus biomicroscope video image sequences*; Br. J. Ophthalmol.; 2000;
- [29] Preece, J. S. ; Claridge, E. ; *Physics Based Medical Image Understanding of the Coloration of the Ocular Fundus with Application to Detection of Diabetic Retinopathy*;
- [30] Santo, C. S. L. E.; *Complexidade de Sistemas de Classificação*;
- [31] Tatsuoka, M. M. ; *Multivariate Analysis : Techniques for Educational and Psychological Research*; Macmillan Publishing Company ; 1988;
- [32] Vapnik; V. ; N. ; *The nature of statistical learning*; Springer-Verlag; 2005;
- [33] Li, B. ; Hu, H. ; Spacek, L. ; *An Adaptive Color Segmentation Algorithm for Sonny Legged Robots*; 21.^a Conferência Internacional IASTED; 2003;
- [34] Coreco Imaging; *IFC-SDK – Software Manual*; 2001;
- [35] Hekmat, S. ; *C++ Essencials* ;PragSoft Corporation ; 2005 ;
- [36] Johnson, R. ; *MATLAB Programming Style Guidelines*; 2002;

ANEXOS A

Análise discriminante linear de *Fisher*

Discriminant

Warnings

All-Events Stacked Histogram is no longer displayed.
 Number of Data Values requested exceeded maximum allowed (10000000).
 This constraint is not executed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	3704188	100.0
Excluded	0	.0
Missing or out-of-range group cases	0	.0
All have one missing discriminating variable	0	.0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total	3704188	100.0

Group Statistics

Class	N	Mean	Std. Deviation	Unweighted Weighted	Weighted
R	1193082	1193082	3699200	3699200.0	3699200.0
G	1212171	1212171	3699200	3699200.0	3699200.0
B	1193082	1193119	3699200	3699200.0	3699200.0
Total	3598335	1193117	3699200	3699200.0	3699200.0
R	30024	24024	8000	8000.0	8000.0
G	3126	4024	8000	8000.0	8000.0
Total	1193082	1193117	3699200	3699200.0	3699200.0
R	1212171	1212171	3699200	3699200.0	3699200.0
G	1193082	1193082	3699200	3699200.0	3699200.0

Tests of Equality of Group Means

	Value/ Lambda	F	df1	df2	Sig.
R	.000	20274.183	1	3704188	.000
G	.002	6997.697	1	3704188	.000
B	.007	24427.204	1	3704188	.000

Post Hoc Multiple Comparisons

Comparison	R	G	B
R	1.000	.999	.999
G	.999	1.000	.999
B	.999	.999	1.000

Analysis 1

Box's Test of Equality of Covariance Matrices

LOG Determinants

Class	Rank	Log Determinant
1	2	13.747
2	2	13.480
3	2	14.253

The rank and natural logarithms of determinants provided are those of two-group covariance matrices.

Test Results

Box's M	1911428
Approx. df1	3
df2	1.5E+011
Sig.	.000

Tests null hypothesis of equal population covariance matrices.

Variables Failing Tolerance Test^a

Minimum Covariance Variance	Tolerance	Minimum Tolerance
1	14712.337	.080
2		.000

All variables passing the tolerance criteria are entered simultaneously.
 a. Minimum tolerance level is .001.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2347	100.0	100.0	.999

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Hypothesis	Value of Lambda	Chi-Square	df	Sig.
1	.002	1189437.992	2	.000

Standardized Canonical Discriminate Function Coefficients

Function	Coefficient
1	1
R	-11.333
G	11.436

Structure Matrix

	Function
B	.187
G ^a	.152
R	.101

Point-to-within-groups correlations between discriminating variables and standardized canonical discriminant functions
 Variables ordered by absolute size of correlation within function.
 a. This variable not used in the analysis.

Functions at Group Centroids

Class	Function
0	.097
1	-5.513

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	3764190
Excluded	0
Missing or out-of-range group codes	0
At least one missing discriminating variable	0
Use in Output	3764190

Prior Probabilities for Groups

Class	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	.500	3095200	3695200.0
1	.500	64990	64990.000
Total	1.000	3764190	3764190.0

Classification Function Coefficients

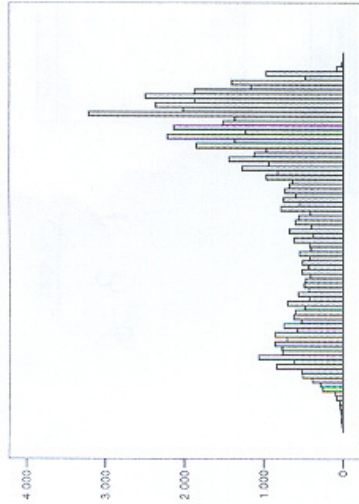
	Class	
	0	1
R	.033	.565
B	-.022	.557
(Constant)	-1.585	-17.850

Fisher's linear discriminant functions

Separate-Groups Graphs

Canonical Discriminant Function 1

Class = 1



Classification Results^a

Original	Count	%	Predicted Group Membership		Total
			Class 0	Class 1	
0	3597749	101.451	3095200	101.451	3695200
1	25815	39.175	64990	2.7	64990
	%		97.3	60.3	100.0
			39.7	60.3	100.0

a. 96.6% of original grouped cases correctly classified

Handwritten notes:
 Fisher's = 0.565xR + 0.033x B - 17.850
 Fisher's > Fisher's → Fisher's → Fisher's

Discriminant

	Notes
Output Created	08-MAR-2007 12:40:40
Comments	C:\ABILI\Sandrina\Projects & Studies\CNTM 004A_Laser_Guided_Photocoagulation\Segmentacao
Input	<none> <none>
Data	
Filter	
Weight	
Split File	
N of Rows in Working Data File	7553177
Definition of Missing Handling	User-defined missing values are treated as missing in the analysis phase.
Cases Used	In the analysis phase, cases with no user- or system-missing values for any predictor variable are used. Cases with user-, system-missing, or out-of-range values for the grouping variable are always excluded.
Syntax	DISCRIMINANT /CRITERIA=LOGR LOGG LOGB /MARSALES=LOGR LOGG LOGB /ANALYSIS=ALL /PRIORS=EQUAL /STATISTICS=MEAN STDDEV UNIVF BOXM COEFF CORR TABLE /CLASSIFY=NONMISSING /POOLED.
Resources	Elapsed Time 0:02:45.68

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	6425113	85,1
Excluded	0	,0
Missing or out-of-range group codes	1128064	14,9
At least one missing discriminating variable	0	,0
Both missing or out-of-range group codes and at least one missing discriminating variable	1128064	14,9
Total	7553177	100,0

Variáveis:
logR (logaritmo de base 10 do canal vermelho), logG (logaritmo de base 10 do canal verde) e logB (logaritmo de base 10 do canal azul)

Group Statistics

Class	Mean	Std. Deviation		Valid N (Listwise)	
		Unweighted	Weighted	Unweighted	Weighted
0	1,9528	,77607	,77607	6287608	6287608,0
	1,9818	,70997	,70997	6287608	6287608,0
	1,9342	,78317	,78317	6287608	6287608,0
1	1,8889	,38634	,38634	137505	137505,00
	1,6106	,34955	,34955	137505	137505,00
	1,2569	,40689	,40689	137505	137505,00
Total	1,9516	,76984	,76984	6425113	6425113,0
	1,9738	,70824	,70824	6425113	6425113,0
	1,9197	,78319	,78319	6425113	6425113,0

Tests of Equality of Group Means

	Wilks'	Lambda	F	df1	df2	Sig.
logR	1,000	,659214	1	6425111	,000	,000
logG	,984	,37386751	1	6425111	,000	,000
logB	,984	,102234,19	1	6425111	,000	,000

Pooled Within-Groups Matrices

Correlation	logR	logG	logB
logR	1,000	,981	,979
logG	,981	1,000	,982
logB	,979	,982	1,000

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	3	-10,111
1	3	-8,969
Pooled within-groups	3	-9,967

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	771692,3
F	128069,4
Approx. df1	6
df2	3,1E+011
Sig.	,000

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Unweighted	Weighted
0	6287608	6287608.0
1	137505	137505.00
Total	6425113	6425113.0

Classification Function Coefficients

	Class	
	0	1
logR	-.24,065	2,674
logG	65,287	61,134
logB	-.32,178	-.55,501
(Constant)	-10,778	-.17,952

Fisher's linear discriminant functions

Classification Results^a

Original	Count	%	Predicted Group Membership		Total
			0	1	
0	5927960	92.4	289518	6287608	
1	224333	16.3	119072	137505	
%			4.6	100.0	
			16.3	63.7	100.0

a. 95.1% of original grouped cases correctly classified.

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	333.2	100.0	100.0	.500

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.750	1846575.630	3	.000

Standardized Canonical Discriminant Function Coefficients

Function	1
logR	-.5162
logG	.735
logB	4.545

Structure Matrix

Function	1
logB	.219
logG	.132
logR	.018

Podiet within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Functions at Group Centroids

Class	Function
0	.085
1	-.3,902

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	7553177
Excluded	0
Missing or out-of-range group codes	1128064
At least one missing discriminating variable	6425113
Used in Output	

Discriminant

Notes	
Output Created	08-MAR-2007 12:38:06
Comments	C:\ABLU\Sandrina\Projects & Studies\CNTM_004A_Laser_Guided_Phocoagulation\SegmentacaoRetina2007-03-08_Output1.EXPLORE.spv
Data	<none>
Filler	<none>
Weight	<none>
N of Rows in Working Data File	7553177
Definition of Missing	User-defined missing values are treated as missing in the analysis.
Cases Used	In the analysis phase, cases with no information specified for a predictor variable are used. Cases with user-system-missing, or out-of-range values for the grouping variable are always excluded.
Missing Value Handling	
Syntax	DISCRIMINANT /GROUPS=Class(0 1) /VARIABLES=lnR lnG lnB /ANALYSIS=ALL /PRINT=CRITERIA /STATISTICS=MEAN STDDEV UNIVF BOXM COEFF CORR TABLE /CLASSIFY=NONMISSING POOLED
Resources	Elapsed Time 0:02:33.62

Group Statistics

Class		Mean	Std. Deviation	Valid N. (listwise)	
				Unweighted	Weighted
0	lnR	4.4564	1.78698	6287608	6287608.0
	lnG	4.5632	1.63478	6287608	6287608.0
	lnB	4.4537	1.80332	6287608	6287608.0
1	lnR	4.3723	88959	137505	137505.00
	lnG	3.7085	80487	137505	137505.00
	lnB	2.8942	93689	137505	137505.00
Total	lnR	4.4937	1.77263	6425113	6425113.0
	lnG	4.5449	1.62618	6425113	6425113.0
	lnB	4.4204	1.80335	6425113	6425113.0

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
lnR	1.000	659.214	1	6425111	.000
lnG	.994	37386.751	1	6425111	.000
lnB	.984	102234.19	1	6425111	.000

Pooled Within-Groups Matrices

	lnR	lnG	lnB
Correlation	lnR 1.000	.981	.979
	lnG .981	1.000	.992
	lnB .979	.992	1.000

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	3	-5.107
1	3	-3.965
Pooled within-groups	3	-4.963

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	771962.3
F	128609.4
df1	6
df2	3.1E+011
Sig.	.000

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.333 ^a	100.0	100.0	.500

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.750	184.667 ^a	3	.000

Standardized Canonical Discriminant Function Coefficients

Function	Function
1	1
lnR	-.5162
lnG	.736
lnB	4.345

Structure Matrix

Function	Function
1	1
lnR	.219
lnG	.132
lnB	.018

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Functions at Group Centroids

Class	Function
0	.085
1	-3.902

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	7553177
Excluded	0
Missing or out-of-range group codes	1128004
At least one missing discriminating variable	6425113
Used in Output	

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Prior	Weighted
0	.500	6287608
1	.500	137505
Total	1,000	6425113

Classification Function Coefficients

	Class	
	0	1
lnR	-10.451	1.161
lnG	28.358	26.550
lnB	-13.975	-24.104
(Constant)	-10.778	-17.962

Fisher's linear discriminant functions

Classification Results^a

Original	Class	Predicted Group Membership		Total
		0	1	
Count	0	5997990	288618	6287608
	1	22433	115072	137505
%	0	95.4	4.6	100.0
	1	16.3	83.7	100.0

a. 95.1% of original grouped cases correctly classified.

$$F_{\text{score}} = 1.161 \times \ln(R) + 26.550 \times \ln(G) - 24.104 \times \ln(B) - 17.962$$

Discriminant

Warnings

All Groups Stacked Histogram is no longer displayed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	1459622	100.0
Excluded	0	.0
Missing or out-of-range group codes	0	.0
At least one missing discriminating variable	0	.0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total	1459622	100.0

Group Statistics

Class	Mean	Std. Deviation	Valid N (listwise)	
			Unweighted	Weighted
0	157.977	119.9904	1479680	1475660.0
Bril	54.875	38.2745	1479680	1475660.0
1	86.261	59.2056	19842	19942.000
Bril	49.205	26.6935	19842	19942.000
Total	157.023	119.6676	1499622	1498622.0
Bril	54.800	38.1461	1499622	1498622.0

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
R	.985	7100.439	1	1499620	.000
Bril	1.000	434.893	1	1499620	.000

Pooled Within-Groups Matrices^a

	R	Bril
Covariance	14252.859	4503.006
Bril	4503.036	1454.934
Correlation	1.000	.989
Bril	.989	1.000

a. The covariance matrix has 1499620 degrees of freedom.

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	2	13.006
1	2	12.668
Pooled within-groups	2	13.038

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	54596.696
F	18194.906
df1	3
df2	1.4E+010
Sig.	.000

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	122 ^a	100.0	100.0	.330

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.981	172818.761	2	.000

Standardized Canonical Discriminant Function Coefficients

Function	
1	6.709
Bril	-6.585

Structure Matrix

Function	
1	.197
Bril	.049

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

Functions at Group Centroids

Class	Function
0	.041
1	-3.010

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Variáveis:
R (canal vermelho), Br (intensidade média dos canais de cor RGB)

Classification Processing Summary

Processed	1499622
Excluded	0
Missing or out-of-range group codes	0
At least one missing discriminating variable	0
Used in Output	1499622

Prior Probabilities for Groups

Class	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	.500	1479680	1479680.0
1	.500	15942	15942.000
Total	1.000	1499622	1499622.0

Classification Function Coefficients

	Class	
	0	1
R	-.038	-.208
Br1	.154	.081
(Constant)	-.1951	-.8425

Fisher's linear discriminant functions

Classification Results^a

Original	Count	Predicted Group Membership		Total
		Class 0	Class 1	
0	1381018	98662	1479680	1479680
1	4855	15087	19642	19642
%	0	93.3	6.7	100.0
	1	24.3	75.7	100.0

a. 93.1% of original grouped cases correctly classified.

Variáveis:
R (canal vermelho), G (canal verde), B (canal azul) e Bril (intensidade média dos canais de cor RGB)

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	2	12,605
1	2	14,750
Pooled within-groups	2	13,646

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	Approx. df1	Approx. df2	Sig.
13005249	4336409	3	.000
	6.3E+012		

Tests null hypothesis of equal population covariance matrices.

Variables Failing Tolerance Test^a

Variable	Within-Groups Variance	Tolerance	Minimum Tolerance
B	14250,490	.002	.001
Bril	14046,644	.000	.000

All variables passing the tolerance criteria are entered simultaneously.

a. Minimum tolerance level is .001.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1,735 ^a	100,0	100,0	,796

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,366	134,60702	2	,000

Standardized Canonical Discriminant Function Coefficients

Function	Coefficient
1	
R	-15,136
G	15,202

Discriminant

Warnings

All-Groups Stacked Histogram is no longer displayed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Included		
Missing or out-of-range group codes	0	,0
At least one missing discriminating variable	0	,0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
Total	13379278	100,0

Group Statistics

Class	Mean	Std. Deviation	Valid N. (listwise)	
			Unweighted	Weighted
0				
R	158,192	118,7865	12847200	12847200
G	158,041	118,7387	12847200	12847200
B	156,640	121,1549	12847200	12847200
Bril	157,610	120,2042	12847200	12847200
1				
R	129,222	61,9117	432078	432078,00
G	71,398	45,6179	432078	432078,00
B	40,607	37,7456	432078	432078,00
Bril	80,403	44,5902	432078	432078,00
Total				
R	157,257	118,4639	13379278	13379278
G	155,243	119,0636	13379278	13379278
B	152,893	121,1249	13379278	13379278
Bril	155,116	119,3018	13379278	13379278

Tests of Equality of Group Means

Wilks' Lambda	F	df1	df2	Sig.
R	,988	25043,849	1	13379276
G	,983	225147,15	1	13379276
B	,971	395039,52	1	13379276
Bril	,987	177437,45	1	13379276

Pooled Within-Groups Matrices^a

	R	G	B	Bril
Covariance	14012,215	13846,591	14062,095	14009,095
	13546,591	13841,539	14079,957	13991,459
	14062,095	14079,957	14250,450	14132,965
	14009,095	13891,459	14132,965	14046,644
Correlation	1,000	,963	,995	,999
	,968	1,000	,999	1,000
	,985	,969	1,000	,999
	,989	,989	1,000	1,000

a. The covariance matrix has 13379276 degrees of freedom.

Analysis 1

Variáveis:
 R (canal vermelho), G (canal verde), B (canal azul) e Bril (intensidade média dos canais de cor RGB)

Structure Matrix

	Function
B ^a	.123
G	.098
Bril ^b	.085
R	.033

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
 Variables ordered by absolute size of correlation within function
 a. This variable not used in the analysis.

Functions at Group Centroids

Class	Function
0	.241
1	-.710

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	13379278
Excluded	0
Used in Output	13379278

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Prior	Unweighted
0	.500	12847200
1	.500	432078
Total	1,000	13379278

Classification Function Coefficients

	Class
R	.002
G	.010
(Constant)	-1.589

Fisher's linear discriminant functions

Classification Results^a

Original	Count	%	Predicted Group Membership		Total
			Class 0	Class 1	
0	12871950	96.6	75250	12847200	12947700
1	96599	99.4	6	432078	432078
Total			22.4	77.6	100.0

a. 98.7% of original grouped cases correctly classified.

Discriminant

Warnings

All-Group Starček Histogram is no longer displayed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	13379278	100.0
Excluded	0	.0
Missing or out-of-range group codes	0	.0
At least one missing discriminating variable	0	.0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total	13379278	100.0

Group Statistics

Class	Mean	Std. Deviation	Valid N (listwise)	
			Unweighted	Weighted
0	155.192	119.7955	12847200	12847200
Bril	157.610	120.3042	12847200	12847200
G	159.041	119.7367	12847200	12847200
B	156.640	121.0349	12847200	12847200
1	120.222	61.9117	432078	432078
Bril	84.403	44.5902	432078	432078
G	71.389	45.0179	432078	432078
B	40.607	37.7456	432078	432078
Total	157.257	118.4839	13379278	13379278
Bril	155.116	119.3018	13379278	13379278
G	155.243	119.0636	13379278	13379278
B	152.893	121.1249	13379278	13379278

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
R	.998	25043.849	1	13379276	.000
Bril	.997	177437.45	1	13379276	.000
G	.993	225147.15	1	13379276	.000
B	.971	395039.52	1	13379276	.000

Pooled Within-Groups Matrices^a

	R	Bril	G	B
Covariance				
R	14012.215	14009.095	13946.591	14062.005
Bril	14009.095	14046.644	13991.459	14132.905
G	13946.591	13991.459	13941.439	14078.957
B	14062.005	14132.905	14079.957	14250.490
Correlation				
R	1.000	.999	.998	.995
Bril	.999	1.000	1.000	.999
G	.998	1.000	1.000	.999
B	.995	.999	.999	1.000

a. The covariance matrix has 13379276 degrees of freedom.

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	2	12.341
1	2	14.048

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	11460578
F	3820186
df1	3
df2	6.5E+012
Sig.	.000

Tests null hypothesis of equal population covariance matrices.

Variables Failing Tolerance Test^a

Within-Groups Variance	Tolerance	Minimum Tolerance
G	13941.539	.000
B	14250.490	.000

All variables passing the tolerance criteria are entered simultaneously.

a. Minimum tolerance level is .001.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.790 ^a	100.0	100.0	.801

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-Square	df	Sig.
1	.358	13729420.0	2	.000

Standardized Canonical Discriminant Function Coefficients

	Function
R	-.16516
Bril	18.575

Structure Matrix

	Function
B ^a	.129
G ^a	.096
Bril	.086
R	.032

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

a. This variable not used in the analysis.

Functions at Group Centroids

Class	Function
0	.744
1	-.7324

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	13379278
Excluded	Missing or out-of-range group codes
	At least one missing discriminating variable
Use in Output	0

Prior Probabilities for Groups

Class	Pror.	Cases Used in Analysis
0	.500	Unassigned
1	.500	Weighted
Total	1.000	13379278

Classification Function Coefficients

	Class	
	0	1
R	.025	1.209
Brl	-.013	-1.200
(Constant)	-1.590	-30.556

Fisher's linear discriminant functions

Classification Results^a

Original	Class	Predicted Group Membership		Total
		0	1	
Count	0	12843565	103035	12947200
	1	87860	344209	432078
%	0	99.2	.8	100.0
	1	20.3	79.7	100.0

a. 96.6% of original grouped cases correctly classified.

Discriminant

Warnings

All-Groups Stacked Histogram is no longer displayed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	13379278	100.0
Excluded		
Missing or out-of-range group codes	0	.0
At least one missing discriminating variable	0	.0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total	13379278	100.0

Group Statistics

Class		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0	Brl	157.610	120.2042	12947200	12947200
	R	158.192	119.7995	12947200	12947200
	G	158.041	119.7387	12947200	12947200
1	B	156.640	121.1549	12947200	12947200
	Brl	80.403	44.5902	432078	432078.00
	R	129.222	61.9117	432078	432078.00
Total	G	71.398	45.6179	432078	432078.00
	B	40.807	37.7456	432078	432078.00
	Brl	155.116	119.3018	13379278	13379278
Total	R	157.257	116.4689	13379278	13379278
	G	156.243	119.0636	13379278	13379278
	B	152.893	121.1249	13379278	13379278

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Brl	.967	177437.45	1	13379276	.000
R	.968	25043.849	1	13379276	.000
G	.963	225147.15	1	13379276	.000
B	.971	395039.52	1	13379276	.000

Pooled Within-Groups Matrices^a

Covariance		Brl	R	G	B
Brl	R	14006.644	14009.095	13891.459	14132.965
R	G	14009.095	14012.215	13946.591	14062.095
G	B	13891.459	13946.591	13941.539	14079.957
B		14132.965	14062.095	14079.957	14250.490
Correlation		Brl	R	G	B
Brl	R	1.000	.999	1.000	.999
R	G	.999	1.000	.998	1.000
G	B	1.000	.998	1.000	.999
B		.999	.999	.999	1.000

a. The covariance matrix has 13379276 degrees of freedom.

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	2	12.341
1	2	14.048
Pooled within-groups	2	13.253

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	11460578
Approx. df1	3620186
df2	3
Sig.	6.5E-072
	.000

Tests null hypothesis of equal population covariance matrices.

Variables Failing Tolerance Test^a

	Within-Groups Variance	Tolerance	Minimum Tolerance
G	13941.539	.000	.000
B	14250.490	.000	.000

All variables passing the tolerance criteria are entered simultaneously.
a. Minimum tolerance level is .001.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.790*	100.0	100.0	.801

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.358	137.29420	2	.000

Standardized Canonical Discriminant Function Coefficients

Function	1
Brl	.18,575
R	-.18,516

Structure Matrix

Function	1
B+	.129
G+	.096
Brl	.066
R	.032

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

a. This variable not used in the analysis.

Functions at Group Centroids

Class	Function
0	.244
1	-.7,324

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	13379278
Excluded	0
Used in Output	13379278

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Prior	Weighted
0	.500	12947200
1	.500	432078
Total	1,000	13379278

Classification Function Coefficients

	Class	
	0	1
Brl	-.013	-1,200
R	.025	1,209
(Constant)	-1,590	-30,556

Fisher's linear discriminant functions

Classification Results^a

Original	Count	Predicted Group Membership		Total
		0	1	
0	12943665	103635	12840020	12947200
1	87889	844209	432078	432078
%	0	99.2	.8	100.0
	1	20.3	79.7	100.0

a. 88.6% of original grouped cases correctly classified.

Discriminant

Warnings

All-Groups Stacked Histogram is no longer displayed.

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	13379278	100.0
Excluded	0	.0
Missing or out-of-range group codes	0	.0
At least one missing discriminating variable	0	.0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total	13379278	100.0

Group Statistics

Class	R	Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0	Brl	158,182	118,7865	12847200	12847200
	R	157,610	120,2042	12947200	12947200
1	Brl	129,222	61,9117	432078	432078,00
	R	80,403	44,5902	432078	432078,00
Total	R	157,257	118,4839	13379278	13379278
	Brl	155,116	119,3018	13379278	13379278

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
R	.986	25643.849	1	13379276	.000
Bril	.987	177437.46	1	13379276	.000

Pooled Within Groups Matrices^a

	R	Bril
Covariance	14012.215	14009.095
Correlation	1.000	.999

a. The covariance matrix has 13379276 degrees of freedom.

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Class	Rank	Log Determinant
0	2	12.341
1	2	14.048

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M	11460578
F	3820186
df1	3
df2	6.5E+012
Sig.	.000

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.750 ^a	100.0	100.0	.801

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.358	13729420.0	2	.000

Standardized Canonical Discriminant Function Coefficients

	Function
R	1
Bril	-18.516
	18.575

Structure Matrix

	Function
R	.086
Bril	.032

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions. Variables ordered by absolute size of correlation within function.

Functions at Group Centroids

Class	Function
0	-.244
1	-7.524

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	13379278
Excluded	Missing or out-of-range values
	Not in missing discriminating variable
Used in Output	13379278

Prior Probabilities for Groups

Class	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	.500	12947200	12947200
1	.500	432078	432078.00
Total	1.000	13379278	13379278

Classification Function Coefficients

	Class	
	0	1
R	.025	1.209
Bril	-.013	-1.200
(Constant)	-1.990	-30.956

Fisher's linear discriminant functions

Variáveis:
 R (canal vermelho normalizado), G (canal verde normalizado) e B (logaritmo canal azul normalizado)

Functions at Group Centroids

Class	Function
0	-.089
1	3.452

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed	13282741
Excluded	0
Missing or out-of-range group codes	0
At least one missing discriminating variable	0
Used in Output	13282741

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Unweighted	Weighted
0	500	12947200
1	500	335541
Total	1,000	13282741

Classification Function Coefficients

	Class	
	0	1
R	.460	.742
G	.784	.764
B	.464	.293
Blgh	-.011	-.021
(Constant)	-.27,042	-.34,357

Fisher's linear discriminant functions

Classification Results^a

Original	Count	Class		Total
		0	1	
0	12141308	805892	12947200	
1	24713	310228	335541	
%		93.8	6.2	100.0
		7.4	92.6	100.0

a. 99.7% of original grouped cases correctly classified.

Discriminant

26

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	16405605	100,0
Excluded	0	,0
Missing or out-of-range group codes	0	,0
At least one missing discriminating variable	0	,0
Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
Total	16405605	100,0

Group Statistics

Class	Mean	Std. Deviation	Valid N (listwise)	Unweighted	Weighted
0	29,384	11,9376	15906560	15906560	15906560
Rn	35,515	9,8452	15906560	15906560	15906560
Gn	28,055	10,2135	15906560	15906560	15906560
Bn	160,370	118,8314	15906560	15906560	15906560
Brilh	160,284	118,7646	15906560	15906560	15906560
R	158,855	120,1438	15906560	15906560	15906560
G	159,945	119,0711	15906560	15906560	15906560
B	56,606	11,1446	489045	489045	489045,00
Rn	28,374	4,6038	489045	489045	489045,00
Gn	14,368	7,6591	489045	489045	489045,00
Bn	119,979	56,7957	489045	489045	489045,00
Brilh	65,577	43,8666	489045	489045	489045,00
R	38,206	35,9092	489045	489045	489045,00
G	74,269	41,6555	489045	489045	489045,00
B	30,212	12,7888	16405605	16405605	16405605,00
Rn	35,288	9,6045	16405605	16405605	16405605,00
Gn	27,639	10,4140	16405605	16405605	16405605,00
Bn	158,141	117,6333	16405605	16405605	16405605,00
Brilh	157,403	118,3175	16405605	16405605	16405605,00
R	155,124	120,3260	16405605	16405605	16405605,00
G	157,339	118,3891	16405605	16405605	16405605,00
B					

Tests of Equality of Group Means

Wilks' Lambda	F	df1	df2	Sig.
,967	2525966,8	1	16405603	,000
,984	260763,34	1	16405603	,000
,949	880718,99	1	16405603	,000
,997	57246,531	1	16405603	,000
,981	315891,65	1	16405603	,000
,969	518622,37	1	16405603	,000
,985	257443,96	1	16405603	,000

Pooled Within-Groups Matrices^a

Covariance	Rn	Gn	Bn	Brilh	R	G	B
Rn	141,949	-55,215	31,572	584,912	583,301	552,378	570,203
Gn	-55,215	94,625	-31,658	-382,990	-369,993	-375,029	-375,317
Bn	31,572	-31,658	102,927	725,898	737,766	751,019	737,108
Brilh	584,912	382,990	725,898	13789,487	13734,791	13947,366	13776,062
R	583,301	369,993	737,766	13734,791	13734,498	13967,436	13764,368
G	552,378	375,029	751,019	13847,366	13867,436	14034,676	13901,622
B	570,203	375,317	737,108	13776,062	13764,368	13901,622	13799,437
Correlation	Rn	1,000	-.476	,281	,403	,391	,407
	Gn	-.476	1,000	-.321	-.325	-.325	-.328
	Bn	,281	-.321	1,000	,609	,621	,618
	Brilh	,403	-.325	,609	1,000	,998	,999
	R	,391	-.325	,621	,998	1,000	,999
	G	,407	-.325	,625	,999	,999	1,000
	B	,318	-.328	,618	,999	,999	1,000

a. The covariance matrix has 16405603 degrees of freedom.

Classification Statistics

Classification Processing Summary

Processed	16405605
Excluded	0
Missing or out-of-range group codes	0
At least one missing discriminating variable	0
Used in Output	16405605

Prior Probabilities for Groups

Class	Prior	Cases Used in Analysis
0	,500	15906560
1	,500	489045
Total	1,000	16405605

Classification Function Coefficients

	Class	
	0	1
Rn	-.429	,437
Gn	,739	,794
Bn	,440	,506
Brilh	,025	1,046
R	-.034	-1,085
(Constant)	-.25521	-.55806

Fisher's linear discriminant functions

Classification Results*

Original	Class	Predictor Group Membership		Total
		0	1	
Count	0	15778053	128507	15906560
%	0	103446	385587	489045
	1	90.2	.8	100.0
		20.7	79.3	100.0

a. 98.6% of original grouped cases correctly classified.

Discriminant

Analysis Case Processing Summary

Unweighted Cases	Valid	N	Percent
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
Total		16405605	100.0

Group Statistics

Class	Mean	Std. Deviation	Valid N. (listwise)	
			Unweighted	Weighted
0	28.384	11.9376	15506560	15906560
	35.515	9.8452	15906560	15906560
	28.056	10.2135	15906560	15906560
	160.370	118.8374	15906560	15906560
	180.284	118.7646	15906560	15906560
	159.945	116.0711	15906560	15906560
	158.855	120.1438	15906560	15906560
1	56.608	11.1446	499045	499045
	28.374	4.6038	499045	499045
	14.368	7.6991	499045	499045
	119.979	56.7957	499045	499045
	65.677	43.8666	499045	499045
	74.259	41.6585	499045	499045
	46.203	35.9092	499045	499045
Total	30.212	12.7986	16405605	16405605
	55.286	9.8045	16405605	16405605
	27.639	10.4140	16405605	16405605
	159.141	117.6333	16405605	16405605
	157.403	118.3175	16405605	16405605
	157.339	118.3891	16405605	16405605
	155.124	120.3260	16405605	16405605

Tests of Equality of Group Means

Wilks' Lambda	F	df1	df2	Sig.
Gn	964	260765.34	1	16405603
Bn	949	880716.99	1	16405603
Brh	897	57246.531	1	16405603
R	.981	315891.66	1	16405603
B	.885	257443.96	1	16405603
G	.989	518622.37	1	16405603

Pooled Within-Groups Matrices*

Covariance		Rn	Gn	Bn	Brh	R	B	G
Rn	141.949	-55.215	31.672	584.912	563.301	570.203	562.378	
Gn	-55.215	94.625	-31.658	-382.990	-389.993	-375.317	-376.029	
Bn	31.672	-31.658	107.927	725.888	737.766	737.108	751.019	
Brh	584.912	-382.990	725.888	13769.487	13794.791	13776.082	13847.365	
R	563.301	-389.993	737.766	13794.791	13764.368	13764.368	13857.436	
B	570.203	-375.317	737.108	13776.082	13764.368	13764.368	13801.622	
G	562.378	-375.029	751.019	13847.366	13857.436	13801.622	14034.676	
Correlation		Rn	1.000	.425	.403	.407	.391	
Gn	-.478	1.000	.321	-.335	-.328	-.325		
Bn	-.321	.321	1.000	.609	.621	.618		
Brh	.425	-.335	.609	1.000	.969	.985		
R	.403	-.328	.621	.969	1.000	.999		
B	.407	-.325	.618	.985	.999	1.000		
G	.391	-.325	.625	.985	.999	.999	1.000	

a. The covariance matrix has 16405603 degrees of freedom.

Classification Statistics

Classification Processing Summary

Processed	16405605
Excluded	Missing or out-of-range group codes
	At least one missing discriminating variable
Used in Output	16405605

Prior Probabilities for Groups

Class	Prior	Case Used in Analysis
0	.500	Unweighted
1	.500	Weighted
Total	1.000	15906560
		499045.00
		16405605

Classification Function Coefficients

	Class	
	0	1
Rn	.429	.437
Gn	.739	.784
Bn	.440	.505
Brhh	.025	1.046
R	-.034	-1.065
(Constant)	-25.521	-55.806

Fisher's linear discriminant functions

Classification Results^a

	Original	Predicted Group Membership		Total
		Class 0	Class 1	
Count	0	15778033	128507	15906560
%		103448	395597	499045
		90.2	.8	100.0
		20.7	79.3	100.0

a. 98.6% of original grouped cases correctly classified

Discriminant

Analysis Case Processing Summary

Valid	Excluded	N	Percent
Missing or out-of-range group codes		16405603	100.0
At least one missing discriminating variable		0	.0
Both missing discriminating variables		0	.0
out-of-range group codes and at least one missing discriminating variable		0	.0
Total		16405603	100.0

Group Statistics

Class	Rn	Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0		29.384	11.9376	15906560	15906560
	Gn	35.515	9.8452	15906560	15906560
	Bn	28.066	10.2135	15906560	15906560
	Brhh	180.370	118.8314	15906560	15906560
	B	159.945	119.0711	15906560	15906560
	C	158.855	120.1438	15906560	15906560
	R	160.284	118.7048	15906560	15906560
1		58.606	11.1448	499045	499045
	Gn	29.374	4.8038	499045	499045
	Bn	14.388	7.8591	499045	499045
	Brhh	119.979	56.7957	499045	499045
	B	74.259	41.6555	499045	499045
	G	36.206	35.9082	499045	499045
	R	65.577	43.8565	499045	499045
Total		30.212	12.7988	16405603	16405603
	Gn	35.258	9.8045	16405603	16405603
	Bn	27.639	10.4140	16405603	16405603
	Brhh	159.141	117.6333	16405603	16405603
	B	157.339	118.3891	16405603	16405603
	G	155.124	120.3280	16405603	16405603
	R	157.403	118.3175	16405603	16405603

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Rn	.867	252.9596	1	16405603	.000
Gn	.984	280763.34	1	16405603	.000
Bn	.949	880718.99	1	16405603	.000
Brhh	.897	57246.631	1	16405603	.000
B	.885	257443.96	1	16405603	.000
G	.959	518622.37	1	16405603	.000
R	.881	315591.65	1	16405603	.000

Pooled Within-Groups Matrices^a

	Rn	Gn	Bn	Brhh	B	G	R
Covariance	141.949	-55.215	31.572	594.912	570.203	552.378	563.301
	Gn	-55.215	84.625	-382.990	-375.317	-375.029	-369.995
	Bn	31.572	-31.658	102.927	726.898	737.108	751.019
	Brhh	594.912	-382.990	726.898	13789.487	13847.366	13734.701
	B	570.203	-375.317	737.108	13776.052	13796.437	13901.822
	G	552.378	-375.029	751.019	13847.366	13901.822	14034.676
	R	563.301	-369.995	737.108	13764.368	13867.436	13754.188
Correlation	1.000	-.476	.281	.425	.457	.467	.403
	Gn	1.000	-.321	-.336	-.328	-.323	-.325
	Bn	.261	-.321	1.000	.018	.025	.021
	Brhh	.425	-.335	.018	1.000	.965	.966
	B	.457	-.328	.018	.965	1.000	.999
	G	.467	-.325	.025	.965	.999	1.000
	R	.403	-.325	.021	.969	.999	1.000

a. The covariance matrix has 16405603 degrees of freedom.

Classification Statistics

Classification Processing Summary

Processed Excluded	16405605
Missing or out-of-range group codes	0
At least one missing discriminating variable	0
Used in Output	16405605

Prior Probabilities for Groups

Class	Cases Used in Analysis	
	Unweighted	Weighted
0	.500	15906560
1	.500	499045
Total	1.000	16405605

Classification Function Coefficients

	Class	
	0	1
Rn	436	429
Gn	737	762
Bn	434	460
Br1h	-.001	1.250
B	-.008	-1.266
(Constant)	-25.487	-55.005

Fisher's linear discriminant functions

Classification Results*

Original	Count	Predicted Group Membership		Total
		0	1	
0	15755021	1506389	15906560	100.0
1	988649	422096	499045	100.0
%		89.1	39	
		19.4	80.6	

a. 98.5% of original grouped cases correctly classified.

Discriminant

Analysis Case Processing Summary

Unweighted Cases	N		Percent
	16405605	0	
Missing or out-of-range group codes	0	0	100.0
At least one missing discriminating variable	0	0	.0
Both missing and out-of-range group codes and at least one missing discriminating variable	0	0	.0
Total	0	0	100.0

Group Statistics

Class	Mean	Std. Deviation	Valid N. (listwise)	
			Unweighted	Weighted
0	29.384	11.9376	15906560	15906560
Gn	35.515	9.8452	15906560	15906560
Bn	26.056	10.2135	15906560	15906560
Br1h	160.370	118.8314	15906560	15906560
G	158.855	120.1438	15906560	15906560
B	155.945	119.0711	15906560	15906560
R	160.284	118.7645	15906560	15906560
1	56.606	11.1446	499045	499045
Gn	28.374	4.6038	499045	499045
Bn	14.368	7.6591	499045	499045
Br1h	119.979	56.7557	499045	499045
G	36.206	35.0052	499045	499045
B	74.259	41.6555	499045	499045
R	65.577	43.8696	499045	499045
Total	30.212	12.7896	16405605	16405605
Rn	35.298	9.8045	16405605	16405605
Gn	27.630	10.4140	16405605	16405605
Br1h	159.141	117.6333	16405605	16405605
G	155.124	120.3260	16405605	16405605
B	157.339	119.3891	16405605	16405605
R	157.403	118.3175	16405605	16405605

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Rn	.867	2525966.8	1	16405603	.000
Gn	.984	260765.34	1	16405603	.000
Bn	.949	880718.99	1	16405603	.000
Br1h	.987	57246.531	1	16405603	.000
G	.969	518622.37	1	16405603	.000
B	.885	257443.96	1	16405603	.000
R	.881	315991.65	1	16405603	.000

Pooled Within-Groups Matrices*

Covariance	Rn					
	Rn	Gn	Bn	Br1h	G	R
Rn	141.949	-55.215	31.672	594.912	552.378	563.301
Gn	-55.215	94.625	-31.659	-382.950	-375.029	-375.317
Bn	31.672	-31.653	102.927	725.888	751.019	737.108
Br1h	594.912	-382.960	725.888	13785.467	13847.366	13776.052
G	552.378	-375.029	751.019	13847.366	14034.678	13901.622
B	570.203	-375.317	737.108	13776.052	13901.622	13796.437
R	563.301	-369.983	737.108	13867.436	13754.791	13754.488
Correlation	1.000	.476	.261	.425	.391	.407
Gn		1.000	-.321	-.335	-.328	-.325
Bn		.285	1.000	.609	.625	.618
Br1h		.425	-.335	1.000	.965	.969
G		.391	.609	.965	1.000	.999
B		.407	.625	.965	.999	1.000
R		.403	.328	.618	.999	1.000
R			-.325	.621	.999	1.000

a. The covariance matrix has 16405603 degrees of freedom.

Classification Statistics

Classification Processing Summary

Processed	16405605
Excluded	0
Used in Output	16405605

Prior Probabilities for Groups

Class	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	.500	15906560	15906560
1	.500	48945	48945.00
Total	1.000	16405605	16405605

Classification Function Coefficients

	Class	
	0	1
Rn	.441	.433
Gn	.737	.713
Bn	.429	.466
Brh	-.018	.622
G	.009	-.634
(Constant)	-25.500	-52.234

Fisher's linear discriminant functions

Classification Results^a

Original	Class	Predicted Group Membership		Total
		0	1	
Count	0	15731603	174967	15906560
	1	94233	404812	499045
%	0	98.9	1.1	100.0
	1	18.9	81.1	100.0

a. 98.4% of original grouped cases correctly classified.

Variáveis:
R (canal vermelho), G (canal verde) e B (canal azul).

Clas * new_class Crosstabulation

		new_class		Total
		0	1	
Clas	0	14319030	107850	14426880
	1	111939	329880	441819
Total		14430969	437730	14868699

Correct %
99,3
74,7
98,5

ANEXOS B

Tempos de processamento nos algoritmos
desenvolvidos

Algoritmo de detecção do guia laser

Processamento	Tempo (s)
PCA	0.06
Detecção de arestas	0.1
Aplicação da máscara de bordo	0.03
Transformada de <i>Hough</i>	0.05
Determinação das coordenadas do centro do guia laser	Desprezável
Salvar coordenadas do guia laser detectado	Desprezável
Marcação do guia laser na imagem original	0.02
TOTAL	0.26

Algoritmo de segmentação das imagens adquiridas

Processamento	Tempo (s)
Redimensionar	0.016
Segmentação da cor:	
Segmentação de cor estática	0.015
Determinação das médias de R, G e B	Desprezável
Segmentação de cor dinâmica	0.016
Segmentação morfológica:	
Ajuste horizontal da janela à fenda	Desprezável
Análise dos grupos de pixels	0.03
Ajuste vertical da janela ao grupo de pixels	Desprezável
Ajuste horizontal da janela ao grupo de pixels	Desprezável
Aplicação da máscara binária aos canais de cor	0.08
TOTAL	0.16