



UNIVERSIDADE DE ÉVORA

Escola de Ciências e Tecnologia

DEPARTAMENTO DE INFORMÁTICA

Técnicas de Visão Computacional para a deteção automática de padrões de fadiga

Pedro Miguel da Silva Roque

Orientação: José Miguel Gomes Saias

Mestrado em Engenharia Informática

Dissertação

Évora, 2013

Mestrado em Engenharia Informática

**Técnicas de Visão Computacional para a deteção
automática de padrões de fadiga**

Pedro Miguel da Silva Roque

Orientador

José Miguel Gomes Saias

Sumário

A fadiga é considerada uma das principais causas de acidentes automóveis em todo o mundo. Este facto levou vários investigadores a estudar e desenvolver tecnologias que permitam mitigar esse problema. Algumas dessas tecnologias têm conseguido relativo sucesso, estando já incluídas em alguns dos veículos de gama mais alta.

Uma das tecnologias mais estudadas utiliza a visão computacional para detetar mudanças no estado físico do condutor, tais como a abertura e orientação dos olhos e ainda a posição da cabeça e o estado da boca. Esta abordagem pode também ser aplicada noutras situações, como por exemplo em linhas de montagem industriais, onde os trabalhadores repetem durante horas os mesmos movimentos, ou em qualquer outro tipo de trabalho, cujas ações sejam repetidas e restringidas a um pequeno espaço físico.

Esta dissertação pretende demonstrar o estado da arte na deteção automática de padrões de fadiga, recorrendo exclusivamente a técnicas de visão computacional. Para tal, após a análise e estudo de vários trabalhos relacionados com este tema, foram implementados os métodos que, segundo os seus autores, obtiveram mais sucesso na deteção de padrões de fadiga.

O sistema desenvolvido recorre à análise do estado dos olhos e da boca e ainda da orientação dos olhos e da face da pessoa em questão, para detetar padrões de fadiga. Após a implementação foram realizados alguns testes e apresentados os respetivos resultados com vista a verificar a eficácia do sistema implementado, bem como das técnicas utilizadas.

*Computer Vision Techniques for the automatic detection
of fatigue patterns*

Abstract

Fatigue is considered to be a leading cause of car accidents throughout the world. This fact led several researchers to study and develop technologies to mitigate that problem. Some of these technologies have achieved relative success, and are currently included in high-end vehicles.

One of the most accepted technology uses computer vision to detect changes in the physical state of the driver, such as the eye opening and orientation, the position of the head and the state of the mouth. This approach can also be applied in other situations, for example in industrial assembly lines where workers repeat the same movements over and over, or any other type of job whose actions are repeated and restricted to a small physical space.

This dissertation aims to present the state of the art of fatigue patterns automatic detection, using only computer vision techniques. For this purpose, after the analysis and study of multiple works related to this topic, the most successful methods to detect fatigue patterns, according to their authors, were implemented in a prototype system.

The developed system uses the analysis of the state of the eyes and mouth and also the target person eyes and face orientation, to detect fatigue patterns. After the implementation phase, some tests were performed, in order to verify the effectiveness of the implemented system and of the used techniques.

Dedico esta dissertação ao meu filho, que nasceu durante o desenvolvimento da mesma

Agradecimentos

Em primeiro lugar quero agradecer a todos os Professores do curso de Engenharia Informática pelo conhecimento que me transmitiram ao longo de todo o mestrado, sem o qual não conseguiria realizar esta dissertação.

Em especial, quero agradecer ao Professor José Miguel Gomes Saias pela oportunidade de desenvolver a dissertação sobre este tema, e pela sua disponibilidade no apoio ao desenvolvimento e melhoramento desta dissertação.

À minha esposa e filho pela paciência e compreensão que sempre tiveram nas longas horas de trabalho.

À minha esposa, outros familiares e amigos por participarem nos testes ao sistema, e por estarem sempre presentes com uma palavra amiga.

A todos os meus mais sinceros agradecimentos.

Acrónimos

AECS *Average Eye Closure Speed*

avi *Audio Video Interleave*

AWAKE *System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation*

bpp *bits por pixel*

CEST *Conselho Europeu da Segurança dos Transportes*

EEG *Eletroencefalografia*

FMCSA *Federal Motor Carrier Safety Administration*

HARKEN *Heart And Respiration In-Car Embedded Non intrusive sensors*

LED *Light-Emitting Diode*

PERCLOS *PERcent of the Time Eyelids are CLOSed*

Ralph *Rapidly Adapting Lateral Position Handler*

RGB *Red, Green e Blue*

STASM *STephen Active Shape Models*

XML *Extensible Markup Language*

Conteúdo

Sumário	i
Abstract	iii
Lista de Conteúdo	xi
Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Âmbito e objetivos	1
1.2 Visão computacional	3
1.2.1 OpenCV	4
1.3 Contribuições	7
2 Estado da arte	9
2.1 Técnicas de medição de padrões de fadiga	9
2.1.1 Medição através do desempenho do condutor	10
2.1.2 Medição através do estado do condutor	12
2.1.3 Medição através do desempenho e estado do condutor	14
2.2 Visão Computacional na deteção de padrões de fadiga	15
2.2.1 Deteção e rastreamento das marcações de faixas de rodagem	15
2.2.2 Deteção do estado dos olhos	16
2.2.3 Deteção da orientação da face	19
2.2.4 Deteção da orientação do olhar	20

2.2.5	Deteção de bocejos	21
2.3	Conclusões do capítulo	22
3	Sistema proposto	23
3.1	Metodologia para a deteção de fadiga	23
3.2	Tecnologias utilizadas	27
3.3	Segmentação de vídeo	29
3.3.1	Deteção da face	30
3.3.2	Deteção dos olhos	31
3.3.3	Deteção das pupilas	32
3.3.4	Deteção do nariz	33
3.3.5	Deteção da boca	34
3.4	Deteção de gestos	34
3.4.1	Deteção do ato de pestanejar	34
3.4.2	Deteção do estado da boca	35
3.4.3	Deteção da inclinação do olhar e da cabeça	37
3.5	Sistema de classificação de gestos	39
3.5.1	Inicialização do sistema	39
3.5.2	Deteção de padrões de fadiga	40
3.6	Resumo do capítulo	42
4	Resultados e discussão	43
4.1	Metodologia para a execução de testes	43
4.2	Resultados dos testes com utilizadores	44
4.3	Discussão dos resultados	46
4.4	Resumo do capítulo	49
5	Conclusões	51
5.1	Balanço crítico	51
5.2	Trabalho futuro	52
	Bibliografia	55

Lista de Figuras

1.1	Exemplo de características Haar-like detetadas pelo OpenCV, de acordo com [35].	5
1.2	Exemplo da criação de uma matriz de Imagem Integral.	6
1.3	Exemplo da deteção da face, olhos, nariz e boca de pessoas e ainda da deteção de contornos utilizando a biblioteca OpenCV.	7
2.1	Exemplo da deteção de marcações de faixas de rodagem com recurso ao OpenCV, segundo [35].	11
2.2	Utilização de um equipamento de EEG para medição do cansaço de um condutor, segundo [17].	12
2.3	Diferença da intensidade média das linhas horizontais de pixels entre imagens digitais de olhos abertos (a) e fechados (b)), segundo [12].	17
2.4	Identificação do estado dos olhos, segundo [21].	18
2.5	Exemplo da deteção de características faciais com recurso à biblioteca STASM, segundo [26].	19
2.6	Captura e deteção da pupila numa situação em que os olhos estão semifechados, segundo [26].	20
2.7	Exemplos da deteção correta e incorreta das pupilas, segundo [23].	21
3.1	Primeira metade do fluxograma do sistema desenvolvido.	26
3.2	Segunda metade do fluxograma do sistema desenvolvido.	27
3.3	Exemplo do output do sistema em modo de depuração.	28
3.4	Exemplo da deteção de uma face em diversas posições.	31
3.5	Exemplo da deteção dos olhos em diferentes estados e posições.	32
3.6	Representação do processo de deteção das pupilas.	33
3.7	Exemplo da imagem de um nariz, detetado recorrendo ao OpenCV.	33

3.8	Exemplo do recorte da área da imagem digital que contém a boca.	34
3.9	Exemplo da distância entre as maiores quebras de intensidade da cor em imagens com os olhos fechados e abertos.	35
3.10	Exemplo de uma imagem digital que representa um bocejo.	36
3.11	Exemplo da detecção e contagem do número de contornos em imagens digitais que contêm a boca fechada ou aberta.	37
3.12	Exemplo da diferença da distância entre as pupilas e o topo da face, entre quando a cabeça está direita (a)) ou inclinada (b)).	38
3.13	Exemplo da diferença da distância entre o extremo superior dos olhos em relação ao extremo inferior do nariz, entre quando a cabeça está direita (a)) ou inclinada (b)).	39

Lista de Tabelas

4.1	Condições de luminosidade, ações efetuadas e aspetos relevantes, relativamente a cada teste.	45
4.2	Resultados individuais dos primeiros seis testes com utilizadores.	46
4.3	Resultados individuais dos últimos seis testes com utilizadores.	47
4.4	Resultados globais dos testes com utilizadores.	47

Capítulo 1

Introdução

Esta dissertação irá abordar os problemas causados pela fadiga humana, e analisar, implementar e testar algumas técnicas de visão computacional com vista a minimizar os mesmos.

Neste capítulo serão descritos o âmbito e os objetivos desta dissertação, bem como introduzidos alguns conceitos acerca de visão computacional. No capítulo 2 são apresentados alguns estudos e projetos desenvolvidos no âmbito deste tema, com especial ênfase nos trabalhos que envolvem técnicas de visão computacional. No capítulo 3 será descrito o desenvolvimento de um sistema de deteção automática de fadiga com recurso a técnicas de visão computacional e no capítulo 4 serão apresentados os resultados do desenvolvimento desse projeto. Por último, no capítulo 5 são apresentadas as conclusões que se puderam retirar ao longo do desenvolvimento desta dissertação.

De seguida serão apresentados o âmbito e os objetivos desta dissertação.

1.1 Âmbito e objetivos

A fadiga é um enorme problema que influencia grandemente as ações de uma pessoa que precisa de se concentrar para levar a cabo ações repetidas como conduzir veículos por longas distâncias, operar máquinas pesadas por longos períodos de tempo, ou mesmo trabalhar em linhas de produção industriais. Este facto agrava-se ainda mais quando se considera que algumas pessoas têm pouca resistência à fadiga e outras começam este tipo de ações já cansadas.

A fadiga é normalmente causada pela privação de sono, pela execução das mesmas tarefas

durante longos períodos de tempo (como por exemplo conduzir), por alguns tipos de medicação e ainda por alguns problemas de saúde como a narcolepsia.

Entre os principais efeitos da fadiga encontram-se o aumento do tempo de resposta e a diminuição da capacidade cognitiva, o que aumenta grandemente a probabilidade da ocorrência de erros humanos na ação que os mesmos estão a desempenhar. Este facto pode levar, por exemplo a que um condutor fatigado não consiga evitar um obstáculo na estrada e em casos extremos, pode mesmo levar um indivíduo a adormecer.

De acordo com [3], em 2006, cerca de 10% a 20% dos acidentes de trânsito na Europa foram causados pelo aumento do tempo de reação dos condutores devido à fadiga. Ainda de acordo com o mesmo autor cerca de 60% dos acidentes que envolveram transportes pesados foram causados por condutores com sintomas de fadiga. Também nos Estados Unidos, entre 3 de julho de 2005 e 31 de dezembro de 2007, cerca de 23% dos acidentes de transportes pesados tiveram a mesma causa [6].

Por este motivo algumas organizações expressaram já a sua preocupação relativamente a este problema. A *Federal Motor Carrier Safety Administration* (FMCSA) dos Estados Unidos avaliou algumas tecnologias de deteção de fadiga já existentes em 2009 [2] e concluiu que apesar da constante evolução tecnológica será necessário ainda mais trabalho para que se consiga criar um sistema eficaz ao ponto de se tornar comum. Também o Conselho Europeu da Segurança dos Transportes (CEST) demonstrou estar atento a este assunto publicando o livro [8], onde são apresentadas algumas formas de mitigar este problema.

Alguns fabricantes automóveis optam por incluir já em alguns dos seus veículos, dispositivos que visam detetar os sintomas de fadiga dos condutores e agir de modo adequado. Algumas marcas automóveis e os respetivos sistemas de deteção de fadiga são por exemplo, o Volvo *Driver Alert System*, o Mercedes-Benz *Attention Assist*, o Audi *Break Recommendation*, o Volkswagen *Driver Alert System* e o Ford *Driver Alert*.

Atualmente, este tema é alvo de diversos estudos e trabalhos pelo que existem vários tipos de tecnologias que visam tentar solucionar o problema. Uma das abordagens mais estudadas e propostas utiliza técnicas de visão computacional para extrair dados em tempo real que permitam detetar sintomas de fadiga no condutor.

Esta dissertação irá introduzir alguns conceitos e tecnologias de visão computacional bem como estudos e projetos já desenvolvidos acerca de sistemas de deteção automática de fadiga recorrendo a estas técnicas. Após essa introdução serão descritos e apresentados os resultados da criação de um sistema de deteção automática de fadiga desenvolvido ao longo desta dissertação, com o objetivo de detetar padrões de fadiga recorrendo apenas a técnicas de visão de computacional.

Na secção seguinte serão introduzidos alguns conceitos de visão computacional, bem como a inerente biblioteca OpenCV que será utilizada para desenvolver o sistema de deteção automática de fadiga proposto.

1.2 Visão computacional

Uma das tecnologias mais promissoras e utilizadas na detecção de fadiga recorre a técnicas de visão computacional, ou seja, recolhe, processa e analisa imagens digitais com vista a identificar nas mesmas possíveis padrões de fadiga. Um dos conceitos por detrás das imagens digitais consiste na álgebra de imagens. De acordo com [30] a álgebra de imagens consiste numa teoria matemática dedicada ao processamento e edição de imagens. Contudo, de acordo com o mesmo autor, dado tratar-se de um campo matemático, o mesmo nunca será um produto acabado e está em constante evolução, a fim de unificar o processamento de imagens e as tarefas de visão computacional.

Os principais objetos da álgebra de imagens são as imagens, os *templates* e a vizinhança. Destes, os objetos mais importantes são as imagens e os restantes podem ser vistos como um caso especial das mesmas. Cada imagem digital consiste numa matriz de valores em que cada posição da matriz corresponde à localização de um pixel ou ponto na imagem, e o valor correspondente é o valor daquele pixel. Um *template* é uma imagem cujos pixels são também imagens e uma vizinhança consiste numa imagem cujos valores dos pixels representam conjuntos de pontos. Assim, pode dizer-se que as imagens digitais são compostas de pixels e que cada pixel contém uma cor ou nível de cinzento.

No entanto, a matriz de uma imagem digital pode representar uma imagem de diversas formas. Pode por exemplo, representar a imagem em escala de cinzentos, onde o valor de cada pixel consiste apenas no nível de cinzento da mesma. Pode também representar a imagem a cores, contendo, por exemplo, o nível de vermelho, de verde e de azul (RGB)¹ de cada pixel. No entanto, isso implica que a mesma imagem a cores ou em escala de cinzentos ocupe diferentes quantidades de memória. As imagens a cores podem ocupar cerca de três vezes mais memória do que em níveis de cinzento. No entanto, estas três cores misturadas permitem representar qualquer tonalidade.

Cada pixel representa uma cor que é designada por Canal. Para além das cores vermelho, verde e azul, algumas imagens digitais contêm ainda o canal alfa que indica o nível de opacidade do pixel correspondente, onde um valor mínimo de opacidade significa que o pixel é transparente. Cada cor de um pixel pode também ser representada em diferentes quantidades de bits. Quanto mais bits representarem uma cor, maior será a gama de cores disponíveis. Este conceito é conhecido como bits por pixel (bpp) ou profundidade da cor. No entanto, quanto maior for a profundidade da cor, mais espaço de memória será necessário. Assim, pode concluir-se que cada imagem digital consiste numa matriz de valores, de modo que a maioria das operações realizadas sobre estas imagens é realizada por meio de operações matemáticas entre matrizes.

Um aspeto muito importante a considerar na manipulação de imagens digitais é a sua resolução. A resolução de uma imagem consiste na quantidade de pixels que a definem. Quantos mais pixels constituírem uma imagem digital, maior é o tamanho com que a

¹Do Inglês *Red, Green e Blue*.

mesma pode ser exibida sem perder detalhes. As cores de uma imagem também podem ser descritas de acordo com várias propriedades, sendo as principais o brilho e o contraste. O brilho consiste no quão clara ou escura é uma imagem, e o contraste consiste na diferença de brilho entre os lados escuros e claros da mesma imagem.

Na secção seguinte será introduzida a biblioteca de visão computacional OpenCV, que facilita muito a recolha, processamento e análise de imagens digitais em tempo real e que foi utilizada no desenvolvimento do sistema de deteção automática de fadiga proposto.

1.2.1 OpenCV

O OpenCV² é uma biblioteca de código aberto orientada para a visão computacional e desenvolvida pela Intel e outros programadores de código aberto. Esta biblioteca foi escrita em C e C++ e pode ser executada nos sistemas operativos Microsoft Windows, Linux, Mac OS, iOS e Android. De notar ainda que a mesma oferece interfaces não só em C e C++, mas também em Python e Java, existindo ainda alguns desenvolvimentos paralelos que disponibilizam interfaces em Ruby³, Matlab⁴ e C#/VB/VC++⁵.

Segundo [20], o OpenCV disponibiliza mais de 500 algoritmos otimizados para a análise de imagens e vídeo. Entre os quais, de acordo com [35] o OpenCV permite detetar objetos em imagens utilizando um método publicado por Paul Viola e Michael Jones em 2001 [33]. Este método utiliza quatro conceitos principais:

- Caraterísticas Haar-like;
- Imagem Integral e Imagem Integral Rodada;
- Método de aprendizagem automática AdaBoost;
- Classificador em cascata.

Para detetar caraterísticas Haar-like numa imagem, as variações de contraste entre grupos retangulares adjacentes de pixels são utilizadas para localizar áreas relativamente brilhantes ou escuras. Grupos adjacentes com uma variação específica de contraste definem uma caraterística Haar-like. Estas caraterísticas podem ser facilmente escaladas, aumentando ou diminuindo o tamanho do grupo de pixels a serem examinados, permitindo assim a deteção de objetos de diferentes tamanhos. Na figura 1.1 estão representados alguns exemplos de caraterísticas Haar-like detetadas pela biblioteca OpenCV.

As caraterísticas representadas na figura 1.1 podem ser utilizadas para detetar contornos (a), linhas (b) e pontos (c) de qualquer tamanho e em qualquer imagem. Um aglomerado

²A biblioteca OpenCV está disponível para download em <http://opencv.org/>.

³Mais informações disponíveis em <http://rubyforge.org/projects/opencv/>.

⁴Mais informações disponíveis em <http://www.cs.stonybrook.edu/~kyamagu/mexopencv/>.

⁵Mais informações disponíveis em http://www.emgu.com/wiki/index.php/Main_Page.

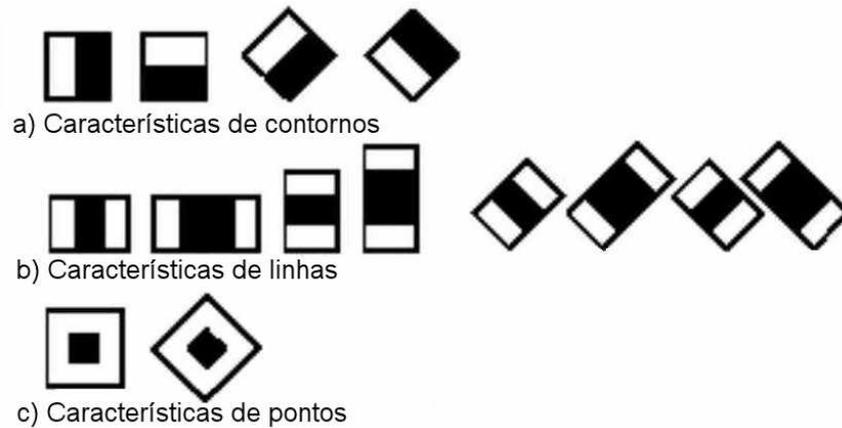


Figura 1.1: Exemplo de características Haar-like detetadas pelo OpenCV, de acordo com [35].

de várias destas características dispostas em posições específicas permitem definir qualquer objeto, que poderá deste modo ser procurado em qualquer imagem.

Para detetar a presença ou ausência de centenas destas características Haar-like numa imagem, [33] utiliza duas representações intermediárias da mesma, designadas por Imagem Integral e Imagem Integral Rodada, em conjunto com um filtro constituído por um conjunto de características Haar-like.

Através da imagem digital original na qual se pretende verificar se existe um determinado objeto é criada uma matriz que contém a soma do valor de contraste do pixel diretamente acima, à esquerda e superior esquerdo, juntamente com o valor do pixel atual. Esta matriz é designada por Imagem Integral.

Na figura 1.2 está representada a construção de uma Imagem Integral (b)) a partir do valor de contraste de cada pixel da imagem original (a)). Neste exemplo, o contraste total do grupo de pixels delineado seria de $62+10-20-27=25$ (c)). A primeira linha e coluna de zeros na Imagem Integral é adicionada de forma a permitir sempre a utilização da mesma fórmula no cálculo do contraste de qualquer zona da imagem.

Para uma área retangular da imagem integral cujos vértices superior esquerdo e inferior direito estão situados, respetivamente nas coordenadas (i_a, j_a) e (i_b, j_b) , a fórmula em questão pode ser expressa da seguinte forma:

$$\text{Contraste total} = (i_b, j_b) + (i_{a-1}, j_{a-1}) - (i_b, j_{a-1}) - (i_{a-1}, j_b) \quad (1.1)$$

Este método permite um cálculo simples do total de contraste de qualquer área retangular de uma imagem de forma a possibilitar a comparação desse valor com os valores que definem cada característica Haar-like.

A Imagem Integral pode então ser girada 45 graus para que se obtenha o valor total de

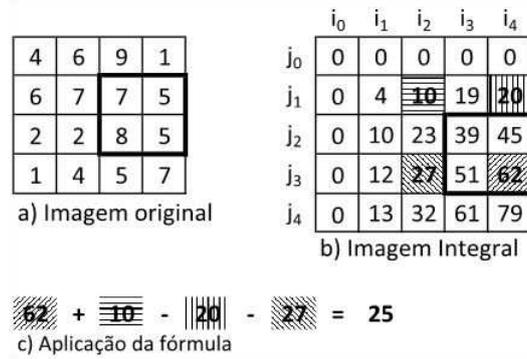


Figura 1.2: Exemplo da criação de uma matriz de Imagem Integral.

contraste dos pixels contíguos localizados num ângulo de 45 graus, para comparar com as características Haar-like rodadas conforme as representadas na segunda metade de cada linha da figura 1.1. Esses valores são armazenados noutra matriz designada por Imagem Integral Rodada.

Estas duas matrizes contêm os valores de contraste da imagem inteira, permitindo uma rápida deteção de características Haar-like. De acordo com [35], são necessários apenas dois passos para calcular as duas matrizes, um para cada uma. Assim, utilizando a Imagem Integral apropriada, aplicando a fórmula 1.1 e formando vários retângulos ligados, qualquer característica Haar-like de qualquer tamanho pode ser detetada.

Um componente chave para a deteção destas características consiste nos classificadores de Haar. De acordo com [33], esses classificadores foram criados recorrendo ao algoritmo de aprendizagem automática⁶ AdaBoost. Segundo [13], este algoritmo permite a combinação de várias características Haar-like semelhantes, atribuindo-lhes um peso proporcional à sua eficácia, de forma a obter um classificador Haar único mais poderoso que funciona como uma cadeia de filtros.

Cada região da imagem é então filtrada por essa cadeia de filtros e a partir do momento em que uma região dessas não passa num dos filtros, a mesma é marcada como não contendo o objeto definido pelo classificador Haar e é descartada, passando-se à região seguinte. Se uma região passa através de todos os filtros, é considerada como contendo o objeto definido pelo classificador Haar. Por esta razão esses classificadores são designados por Classificadores Haar em Cascata. Cada classificador consiste num ficheiro XML⁷ que contém todas as características Haar-like que representam o objeto.

Assim, é possível utilizar a biblioteca OpenCV em conjunto com vários filtros Haar em cascata para detetar diversos objetos em imagens digitais e em tempo real. Por exemplo, esta biblioteca já inclui ficheiros Haar em cascata que permitem a deteção da face, boca, nariz e olhos de pessoas conforme representado na figura 1.3. O OpenCV fornece também

⁶Do Inglês *machine learning*. Segundo [27], trata-se de um ramo de inteligência artificial dedicado ao desenvolvimento e estudo de sistemas capazes de aprender a partir de dados.

⁷Do Inglês *Extensible Markup Language*.

várias outras funções, tais como a detecção de contornos conforme exemplificado na mesma figura.



Figura 1.3: Exemplo da detecção da face, olhos, nariz e boca de pessoas e ainda da detecção de contornos utilizando a biblioteca OpenCV.

Na secção seguinte serão descritas as principais contribuições desta dissertação.

1.3 Contribuições

Esta dissertação apresenta um sistema de detecção automática de padrões de fadiga baseado exclusivamente em técnicas de visão computacional. Como contribuições principais desta dissertação encontram-se:

- Detecção de características com recurso à biblioteca de visão computacional OpenCV;
- Detecção de quatro padrões de fadiga distintos, recorrendo à captura de imagens a partir de uma única câmara de filmar;
- Teste e análise do desempenho das diferentes métricas usadas na avaliação dos níveis de fadiga da pessoa monitorizada;
- Criação de um sistema capaz de detetar padrões de fadiga em diversos ambientes de trabalho, como por exemplo, no habitáculo de um automóvel e nas linhas de montagem industriais.

No capítulo seguinte serão apresentados alguns estudos e projetos elaborados nos últimos anos, envolvendo sistemas de detecção automática de fadiga.

Capítulo 2

Estado da arte

Inúmeros projetos e estudos foram já desenvolvidos acerca deste tema. Na próxima secção serão abordadas algumas formas utilizadas por diferentes autores para detetar automaticamente padrões de fadiga, e na secção 2.2 será descrita a forma como alguns autores utilizam técnicas de visão computacional para ajudar na resolução do mesmo problema.

2.1 Técnicas de medição de padrões de fadiga

Para desenvolver um sistema que possa detetar a fadiga humana é essencial saber identificar sintomas ou efeitos da mesma. De acordo com [34], os sistemas de deteção automática de fadiga podem ser categorizados em três grandes grupos segundo a forma como medem a fadiga:

- Através do desempenho do condutor, mais especificamente da posição do carro na estrada, da sua distância ao veículo da frente, do controlo do volante, da aceleração e da velocidade;
- Através da medição do estado do condutor, avaliando se tem os olhos fechados ou abertos, a frequência e velocidade com que os fecha e reabre, a direção do olhar, a inclinação da cabeça, a frequência com que boceja ou, de uma forma mais intrusiva a frequência cardíaca e as ondas cerebrais;
- Através da medição simultânea do desempenho do condutor e do seu estado, fundindo assim os dois ramos anteriores.

Nas próximas três secções estão descritos cada um destes três ramos e são apresentados alguns trabalhos exemplificativos de cada um dos mesmos.

2.1.1 Medição através do desempenho do condutor

Neste tipo de abordagem, a fadiga é medida pelo comportamento do condutor ou do veículo. Aqui não é considerada a fadiga em si, mas o que ela pode causar. Como tal, são consideradas medidas como a posição do veículo em relação à faixa de rodagem, a sua distância ao carro da frente, a sua velocidade e a forma como o condutor roda o volante e acelera.

De acordo com [7], a deteção das marcações da faixa de rodagem é uma boa forma de avaliar o nível de fadiga dos condutores, pois permite detetar a posição do veículo relativamente às mesmas, e um condutor cansado não costuma conduzir no centro da faixa de rodagem durante longos períodos de tempo. Esta técnica é mais frequentemente implementada colocando uma câmara no lado exterior do veículo e utilizando técnicas de visão computacional para detetar a posição das marcações da faixa de rodagem. O desafio na implementação deste tipo de abordagem está inerente à largura variável das marcações das faixas de rodagem, ao desgaste das mesmas, ou mesmo à sua inexistência, e ainda com estradas pouco iluminadas e condições atmosféricas adversas.

No trabalho [4], foi desenvolvido um dispositivo de rastreamento de marcações das faixas de rodagem que funciona bem em condições de pouca iluminação, mas falha em estradas que não sejam planas ou com marcações de linhas desgastadas. Técnicas semelhantes são também utilizadas em sistemas de condução autónoma como o Ralph¹, apresentado por [28], que dirigiu de forma autónoma durante 96% de uma viagem de 4800km. Neste caso, a intervenção humana apenas foi necessária para mudar de faixa de rodagem ao ultrapassar outros veículos e para evitar a saída por escapatórias.

O sistema *Driver Alert System*² da Volvo, tal como o sistema *Driver Alert*³ da Ford utilizam tecnologias de deteção da faixa de rodagem para alertar os condutores aquando da ocorrência de estilos de condução anormais, que possam ser causados pela fadiga. Uma das tecnologias que pode ser utilizada para detetar as marcações da faixa de rodagem é a visão computacional através do OpenCV, conforme descrito na secção 1.2.1. Um exemplo da utilização desta tecnologia é o sistema desenvolvido por [36], que alcançou uma precisão de 90% na deteção de faixas de rodagem em tempo real. A figura 2.1 representa o resultado do trabalho deste autor, onde é visível a deteção das marcas da faixa de rodagem à esquerda (a vermelho) e à direita (a verde), e o centro dessa faixa de rodagem marcada com outra linha (azul) que deve ser o percurso a seguir por um condutor sem sintomas de fadiga.

¹Do Inglês *Rapidly Adapting Lateral Position Handler*.

²Mais informação acerca deste sistema pode ser encontrada em <https://www.media.volvocars.com/global/enhanced/en-gb/Media/Preview.aspx?mediaid=12130>.

³Mais informação acerca deste sistema pode ser encontrada em http://www.euroncap.com/rewards/ford_driver_alert.aspx.

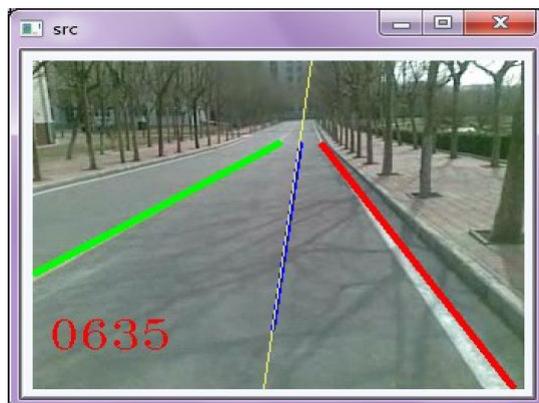


Figura 2.1: Exemplo da deteção de marcações de faixas de rodagem com recurso ao OpenCV, segundo [35].

De acordo com [7], um condutor sonolento tende a carregar no acelerador aumentando a sua velocidade e diminuindo assim a distância relativamente ao veículo da frente. Como tal, estes parâmetros podem também ser utilizados para detetar o nível de fadiga dos condutores. A velocidade do veículo pode ser obtida através do velocímetro do mesmo e a sua distância ao veículo da frente pode ser obtida com recurso a sensores de radar.

A análise dos movimentos do volante também pode ser eficaz para detetar o nível de fadiga dos condutores. O autor de [19] conseguiu uma eficácia de 86% na deteção de padrões de fadiga recorrendo apenas à monitorização dos movimentos do volante e distinguindo os que representam sinais de fadiga. Segundo este autor, um condutor sonolento tende a corrigir a sua trajetória mais vezes e com movimentos mais rápidos, resultando em frequentes e bruscas mudanças de direção.

O sistema *Driver Alert*⁴ da Volkswagen analisa os movimentos do volante para detetar e avisar condutores fatigados. Já o sistema *Attention Assist*⁵ da Mercedes-Benz, tal como o sistema *Break Recommendation*⁶ da Audi, utilizam a velocidade do veículo, a aceleração, a localização na faixa de rodagem e ainda os movimentos do volante para detetar estilos de condução anormais, possivelmente causados pela fadiga do condutor.

No entanto, para além das limitações já mencionadas relativamente à visibilidade das marcações das faixas de rodagem, estes sistemas necessitam ainda de um período inicial de treino para que se adaptem à forma de conduzir do condutor, de modo a que mais tarde possam detetar anomalias. Este processo não é, contudo eficaz caso o condutor inicie a viagem já fatigado, ou em viagens curtas, e não deteta os designados micro-sonos caso o condutor adormeça durante breves segundos numa estrada em linha reta e mantendo o

⁴Mais informação acerca deste sistema pode ser encontrada em <http://www.volkswagen.co.uk/new/ti-guan-gp/explore/interior/dashboard/driver-alert-system>.

⁵Mais informação acerca deste sistema pode ser encontrada em http://www.euroncap.com/rewards/mercedes_benz_attention_assist.aspx.

⁶Mais informação acerca deste sistema pode ser encontrada em <http://www.audi.co.uk/new-cars/a4/a4-allroad-quattro/driver-aids/driver-information-system.html>.

volante na mesma posição.

A próxima secção apresenta alguns sistemas de medição da fadiga do condutor que recorrem apenas a dados sobre o estado do condutor.

2.1.2 Medição através do estado do condutor

A fadiga também pode ser medida diretamente a partir do estado do condutor. Esta abordagem pode ser subdividida em dois tipos: intrusivos e não-intrusivos.

De acordo com [34], as abordagens intrusivas são as mais precisas. Estas incluem tecnologias que medem as ondas cerebrais, a frequência respiratória e os batimentos cardíacos. Mas estas são também as mais difíceis de aplicar e de serem aceites porque necessitam que dispositivos, como eléctrodos e/ou pulseiras (ou cintos) sejam colocados no condutor, o que pode causar algum desconforto e são menos práticas.

Alguns estudos com resultados promissores têm sido realizados utilizando a Eletroencefalografia (EEG) para detetar fadiga. Um exemplo é o estudo realizado por [16], que concluiu que as ondas cerebrais alfa diminuem um pouco e as ondas beta diminuem significativamente em qualquer pessoa que comece a ficar fatigada. Mas esta é também uma abordagem demasiado intrusiva e dispendiosa, pois precisa que um NeuroScan⁷ seja colocado na cabeça do condutor, conforme exemplificado na figura 2.2.



Figura 2.2: Utilização de um equipamento de EEG para medição do cansaço de um condutor, segundo [17].

Em alguns casos, dada a veracidade dos resultados da EEG, esta é utilizada para testar a eficácia de outros sistemas menos intrusivos.

Atualmente está a ser desenvolvido um projeto que visa construir um sistema que incorpore vários sensores não-intrusivos no banco e cinto de segurança do condutor, de forma a medir

⁷Dispositivo que permite a execução de Eletroencefalografias com vista a investigar e analisar o sistema nervoso central humano.

o seu ritmo cardíaco e a frequência respiratória para detetar sintomas de fadiga. Este projeto é designado por HARKEN⁸ e está prevista a sua conclusão até Maio de 2014.

Os métodos de deteção de fadiga mais propostos e estudados são os não-intrusivos, porque estes não interferem no quotidiano do utilizador, não necessitando de qualquer tipo de intervenção por parte do mesmo. Os sistemas não-intrusivos que medem a fadiga através do estado do condutor fazem-no recorrendo apenas a sinais visíveis. No entanto, este tipo de sistemas tem também os seus próprios problemas, relacionados com ambientes de baixa luminosidade e diferentes aparências dos condutores (como a utilização de óculos de sol). Os métodos não-intrusivos que medem a fadiga através do estado do condutor recorrem a fatores como os movimentos da cabeça, da boca, dos olhos e da pálpebra dos condutores.

De acordo com [7], o nível de fadiga de qualquer pessoa pode ser avaliado através da análise do movimento das suas pálpebras, mais especificamente, através da percentagem de fecho dos olhos (PERCLOS⁹) e da velocidade média de abertura e fecho do olhos (AECS¹⁰). PERCLOS é definida como a percentagem de um período de tempo predefinido (tipicamente 20% de um minuto), em que uma pessoa tem os seus olhos 80% ou mais fechados. De acordo com [37] e [21], PERCLOS é considerada a medição mais eficiente do nível de fadiga de um condutor, considerando apenas pistas visuais acerca do mesmo. AECS é definida como a quantidade de tempo que uma pessoa precisa para fechar e abrir totalmente os seus olhos quando pestaneja.

Para que essa tecnologia funcione, os olhos do condutor têm de ser detetados e rastreados, o que não é fácil em ambientes de fraca luminosidade. Em [38] sugere-se um sistema que utiliza luzes de infravermelhos para permitir a deteção e rastreamento dos olhos das pessoas em diferentes condições de iluminação. Isto é feito causando o efeito brilhante/escuro da pupila com luzes de infravermelhos permitindo assim extrair a posição dos olhos através da reflexão córnea. O mesmo estudo extrapolou ainda a orientação do olhar através da análise da posição relativa dos olhos e da respetiva pupila, de forma a detetar outro padrão de fadiga nos condutores. Segundo este autor, o olhar tende a ser estreito em condutores cansados.

Este autor utilizou ainda a orientação da cabeça como outro método para melhorar a deteção de padrões de fadiga. Esta foi possível após a deteção dos olhos e extrapolando a posição da face através da localização e tamanho dos mesmos. Esta técnica permite detetar movimentos da cabeça como inclinações súbitas, que são um bom indicativo de que a pessoa está cansada.

O trabalho em causa foi ainda mais longe na utilização de sinais visíveis para detetar padrões de fadiga nos condutores, mais especificamente através da deteção de bocejos recorrendo à análise da localização e forma da boca do condutor. De seguida, após a extração de todos estes sinais visuais, a informação contextual é fundida, permitindo que

⁸Do Inglês *Heart And Respiration In-Car Embedded Non intrusive sensors*. Mais informação sobre este projeto está disponível em <http://harken.ibv.org/>.

⁹Do Inglês *PERcent of the Time Eyelids are CLOSeD*.

¹⁰Do Inglês *Average Eye Closure Speed*.

o sistema decida se os mesmos indicam um condutor fatigado ou não.

Este sistema de monitorização da fadiga é um dos sistemas mais completos em tempo real e não-intrusivos já desenvolvido, utilizando a orientação do olhar e da cabeça, os movimentos da pálpebra e a forma da boca para detetar sintomas de fadiga. O autor utilizou duas câmaras equipadas com LEDs¹¹ de infravermelhos, uma com um amplo ângulo focada no rosto do condutor e outra com um ângulo estreito focada nos olhos.

Tal como na deteção de faixas de rodagem, alguns destes sistemas não-intrusivos também utilizam o OpenCV conforme descrito na secção 1.2.1. Alguns exemplos são [22] e [15], que utilizam o OpenCV para detetar sinais de fadiga nos condutores através dos seus olhos. Em [22] conseguiu-se 91,6% de taxa de precisão média e [15] obteve 100% de precisão ao testar com condutores relativamente estáticos.

A próxima secção irá introduzir a terceira e última abordagem de medição da fadiga.

2.1.3 Medição através do desempenho e estado do condutor

Será intuitivo concluir que um sistema que combine dados extraídos simultaneamente do estado e do desempenho do condutor será mais preciso na deteção de fadiga. Em [14] foram testados vários sistemas de deteção de fadiga e concluiu-se que um sistema de medida única não é sensível e confiável o suficiente para detetar eficazmente sintomas de fadiga. No entanto, este autor afirmou também que o maior desafio num sistema que utilize diversas métricas consiste em interpretar simultaneamente e eficazmente todos os dados extraídos dos vários sensores utilizados.

A União Europeia também contribuiu para esta abordagem através do desenvolvimento do projeto AWAKE¹², elevando a importância da utilização de vários sensores para medir os níveis de fadiga dos condutores. Segundo [5], neste projeto foram consideradas medições não só do estado do condutor, como os movimentos das pálpebras e a direção do olhar, mas também medições do desempenho do mesmo, através dos movimentos do volante, da posição na faixa de rodagem e ainda do uso do acelerador e do travão. Todos estes dados foram então combinados com uma avaliação do risco de tráfego naquele instante, obtida a partir de mapas de navegação digitais, dispositivos anti-colisão e leituras do velocímetro para estimar o nível de fadiga do condutor.

Na secção seguinte serão apresentados alguns trabalhos de autores que utilizaram técnicas de visão computacional para detetar padrões de fadiga em condutores de veículos.

¹¹Do Inglês *Light-Emitting Diodes*.

¹²Do Inglês *System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation*.

2.2 Visão Computacional na detecção de padrões de fadiga

Conforme foi mencionado nas secções anteriores, alguns sistemas de detecção automática de fadiga recorrem a técnicas de visão computacional. Estas são utilizadas em funções como a detecção e rastreamento das marcações de faixas de rodagem e da face, olhos e boca dos condutores.

Nas secções seguintes são abordadas algumas metodologias e técnicas utilizadas na realização destas funções.

2.2.1 Detecção e rastreamento das marcações de faixas de rodagem

A detecção e rastreamento das faixas de rodagem é normalmente conseguida através da colocação de uma câmara no lado exterior do veículo. Em [36] foi utilizada uma câmara e o OpenCV para capturar em tempo real a imagem da estrada na qual o veículo circulava. De seguida, utilizando o OpenCV, o autor converteu a imagem para escala de cinzentos e aumentou o seu contraste de forma a realçar as diferenças de tonalidades, como as linhas brancas no alcatrão preto. Desta forma, aplicando a transformada de Hough¹³ também, disponibilizada pelo OpenCV, o autor detetou as linhas presentes na imagem. De entre estas escolheu as duas maiores linhas verticais situadas nos extremos esquerdo e direito da imagem, que assumiu serem as marcações da faixa de rodagem, e traçou uma outra linha vertical no centro destas, que deveria ser a trajetória normal a seguir pelo condutor.

De acordo com [29], após a detecção das marcações da faixa de rodagem, trata-se apenas de calcular a distância do veículo em relação à marcação direita e à marcação esquerda e de comparar a evolução destas distâncias ao longo da viagem. Estas distâncias podem ser obtidas através das coordenadas dos pixeis em que as mesmas se encontram relativamente à linha central que entretanto foi desenhada através do OpenCV. Deste modo é possível constatar se o veículo se mantém no centro da faixa de rodagem, ou se o mesmo se aproxima gradualmente, ou subitamente de umas destas, ou ainda, se existem constantes e súbitas correções de trajetória ao longo da viagem, o que pode indicar que o condutor está fatigado.

Segundo [7], apesar desta ser uma das técnicas mais utilizadas nos sistemas de detecção de fadiga, tal como já foi mencionado na secção 2.1.1, o desafio na implementação deste tipo de abordagem está inerente à largura variável das marcações das faixas de rodagem, ao desgaste das mesmas, ou mesmo à sua inexistência, e ainda com estradas pouco iluminadas, condições atmosféricas adversas e níveis de trânsito variáveis. Ao contrário desta solução, que confia apenas na detecção e rastreamento das marcações de faixas de rodagem, com as limitações já mencionadas, existe um outro tipo de sistemas de detecção automática de fadiga, que com o mesmo nível de equipamentos consegue recorrer a várias métricas diferentes. São elas a orientação dos olhos e da face, a detecção de bocejos, do estado dos

¹³Consiste numa técnica matemática que permite detetar formas geométricas como retas, círculos e elipses em imagens digitais.

olhos, ou seja, se os mesmos estão abertos ou fechados e ainda da frequência e rapidez com que o condutor abre e fecha os olhos. Estas métricas serão abordadas nas secções seguintes.

2.2.2 Detecção do estado dos olhos

A deteção do estado dos olhos permite não só detetar se os olhos estão fechados ou abertos, como também calcular as métricas AECS e PERCLOS, descritas na secção 2.1.2, considerando a frequência com que o condutor pestaneja e o tempo que demora a fazê-lo.

Em [17] foi utilizado o OpenCV para detetar em tempo real o estado dos olhos do condutor, ou seja, se em cada instante os mesmos se encontravam fechados ou abertos. Para tal, o autor confiou no OpenCV para detetar e recortar a área da imagem que contém individualmente cada olho. De seguida constatou que analisando a intensidade média de cada linha horizontal de pixels das imagens recortadas era possível localizar verticalmente as sobrancelhas e as pestanas, quando os olhos estão fechados, ou as sobrancelhas e as pestanas superiores em conjunto com a pupila e a íris, quando estes estão abertos, dado que estas constituem as áreas mais escuras de cada uma dessas imagens.

Com esse conhecimento, conseguiu saber se os olhos estavam fechados ou abertos através da identificação da localização destas coordenadas, dado que a distância entre as sobrancelhas e as pestanas quando os olhos estão fechados é maior do que a distância entre as sobrancelhas e a pupila em conjunto com as pestanas superiores e a íris quando estes estão abertos. Este facto está representado na figura 2.3.

Este conhecimento permitiu ainda obter o nível de abertura dos olhos, analisando a distância entre essas coordenadas de pixels, pois estarão tanto mais afastadas quanto mais fechados estiverem os olhos.

Tal como [17], também [12] utilizou a diferença de intensidade média de cada linha horizontal de pixels das imagens que continham os olhos para identificar o estado dos olhos.

Em [17] foi ainda mencionada a possibilidade de detetar o estado dos olhos recorrendo a detetores de arestas, nomeadamente ao *Canny Edge Detector* ou ao *Sobel Edge Detector*. Ambos os algoritmos permitem identificar e enumerar as arestas presentes numa imagem, como por exemplo nas imagens que contêm os olhos do condutor. Segundo este autor, seguindo este método é possível diferenciar se os olhos representados na imagem estão abertos ou fechados, dado que a imagem que contêm os olhos abertos irá conter mais arestas do que aquela que contêm os olhos fechados.

Em [21] foi desenvolvido um sistema que permite medir a fadiga de uma pessoa através das medidas PERCLOS e AECS. Para tal, o autor criou um algoritmo que começava por converter a imagem da face da pessoa para binário (preto e branco), recortando de seguida a área normalmente correspondente aos olhos, que segundo o mesmo se situa entre os 7/24 e 7/12 da face, medindo do topo para a base da mesma. De seguida simplesmente analisava a

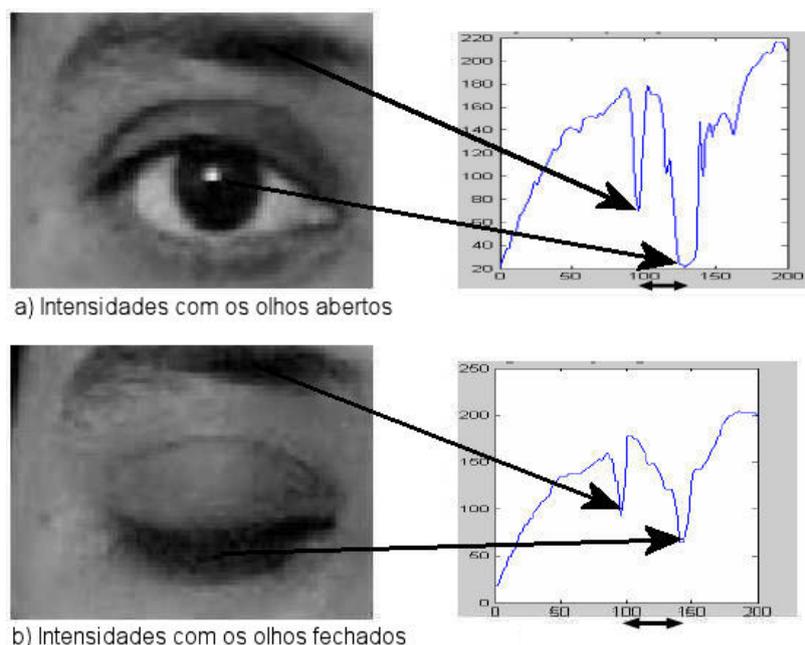


Figura 2.3: Diferença da intensidade média das linhas horizontais de pixels entre imagens digitais de olhos abertos (a)) e fechados (b)), segundo [12].

quantidade de pixels brancos contidos na área dos olhos, considerando que quanto maior a quantidade, mais abertos estariam os olhos. Na figura 2.4 está representada a diferença da quantidade de pixels brancos contidos na área dos olhos, consoante o estado dos mesmos.

De acordo com [9], o tempo médio que uma pessoa demora a pestanejar está compreendido entre 302,7ms e 317,3ms. De acordo com este mesmo autor, este curto intervalo de tempo impede que alguns algoritmos mais eficazes na detecção do ato de pestanejar sejam utilizados em sistemas de tempo real, obrigando à utilização de um algoritmo adequado à capacidade computacional existente e ao tempo de detecção requerido em detrimento de algoritmos mais eficazes.

Este autor utilizou as bibliotecas OpenCV e STASM¹⁴ para detetar as pupilas. De acordo com [25], a STASM foi criada em C++ e necessita do OpenCV para funcionar, permitindo detetar algumas características faciais como as extremidades das sobrancelhas, dos olhos, da boca e ainda as pupilas e as narinas, conforme está representado na figura 2.5.

Contudo, de acordo com o criador desta biblioteca, a mesma foi desenhada para ser aplicada em fotografias do tipo passaporte, não sendo tão eficaz em faces que não estejam totalmente verticais, nem quando as mesmas possuem características pouco comuns como a boca aberta.

Após a detecção das pupilas e devido às questões de desempenho já mencionadas, [9] optou por detetar se os olhos estavam fechados ou abertos através da análise da intensidade dos

¹⁴Do Inglês *STephen Active Shape Models*.

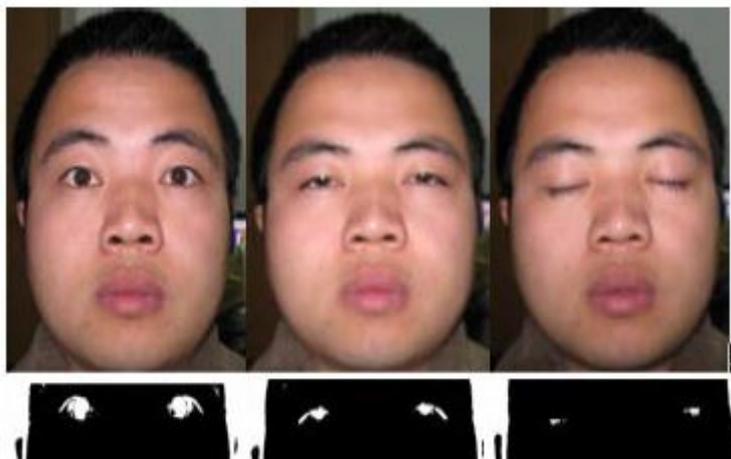


Figura 2.4: Identificação do estado dos olhos, segundo [21].

pixels da área da imagem que deve conter as pupilas. Segundo este autor quando os olhos estão abertos, a intensidade destes pixels apresenta uma simetria horizontal correspondente à forma circular do olho, enquanto que quando os olhos estão fechados, esta simetria não existe.

Outra técnica utilizada na detecção das pupilas recorre a luzes de infravermelhos para causar o efeito brilhante/escuro das mesmas, conforme mencionado na secção 2.1.2. Foi esse o método utilizado por [11], que após capturar a imagem dos olhos com o efeito brilhante/escuro e de as converter para escala de cinzentos, aplicou um algoritmo específico de forma a encontrar uma área com um determinado tamanho e nível de brilho correspondente às pupilas.

Esse algoritmo começa por converter a imagem de uma escala de cinzentos para binário e depois de aplicar um filtro morfológico e de descartar todas as áreas brancas com uma área inferior a um determinado número de pixels, seleciona a restante, que deverá corresponder à pupila. Esta técnica permite ainda detetar quando os olhos estão semifechados, pois ao invés de a pupila ter uma forma circular, esta tem uma forma semicircular conforme se pode observar na figura 2.6. Nessa figura podemos ver em (a) a imagem original já convertida para escala de cinzentos, em (b) a imagem binária (preto e branco), em (c) a imagem após a aplicação do filtro morfológico e em (d) a identificação da pupila após a exclusão das áreas brancas de dimensões reduzidas.

Tal como [11], também [38] e [24] utilizaram luzes de infravermelhos para possibilitar a detecção das pupilas do condutor. Este método é muito eficaz em ambientes com fraca iluminação, contudo a utilização de luzes de infravermelhos apontadas diretamente aos olhos de pessoas gera alguma controvérsia quanto à possibilidade de causar danos nos olhos. No entanto, autores como [18] afirmam que as luzes de infravermelhos normalmente utilizadas neste tipo de sistemas (com uma largura de banda situada entre os 760 e 1400nm) não causam qualquer tipo de dano ao sistema ocular das pessoas.

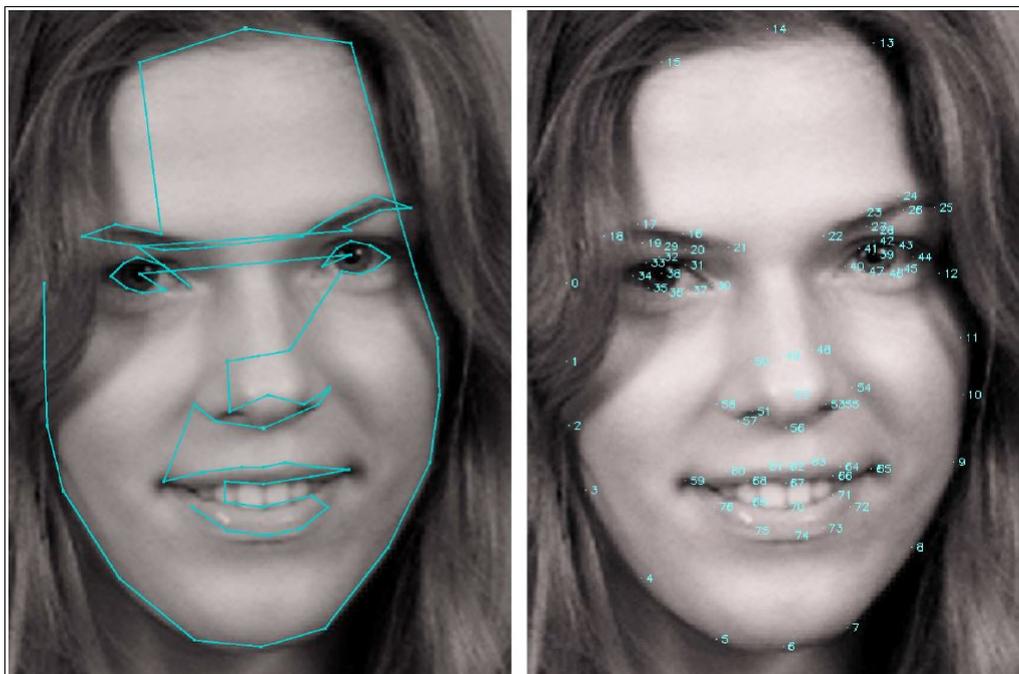


Figura 2.5: Exemplo da detecção de características faciais com recurso à biblioteca STASM, segundo [26].

Após a identificação da área da pupila, o diâmetro da mesma permite identificar se o condutor tem os olhos bem abertos ou semifechados, pois o diâmetro desta será inferior neste último caso.

Em [31] foi considerado também o nível de abertura dos olhos como forma de medir a fadiga do condutor. Este autor começou por calcular a média das projeções horizontais das primeiras 100 imagens dos olhos dos condutores. Esse valor foi de seguida utilizado como consistindo no valor médio que iria ser obtido nas imagens seguintes dos olhos do condutor em que estes estivessem abertos. De seguida, sempre que a correlação da projeção horizontal das imagens seguintes com a média obtida anteriormente fosse inferior a 1 significava que os olhos não estavam totalmente abertos, e quanto maior a diferença, mais fechados estavam os olhos.

Na secção seguinte será apresentada a forma como a visão computacional pode ser utilizada para detetar a orientação da face de uma pessoa.

2.2.3 Detecção da orientação da face

Segundo [9], a detecção da orientação da face de uma pessoa pode ser estimada, primeiramente através da detecção da face e olhos da mesma e comparando de seguida a posição dos olhos relativamente à face. Ou seja, comparando a distância de cada olho às extremidades da face. Deste modo, recorrendo à biblioteca OpenCV para detetar a face e os olhos e

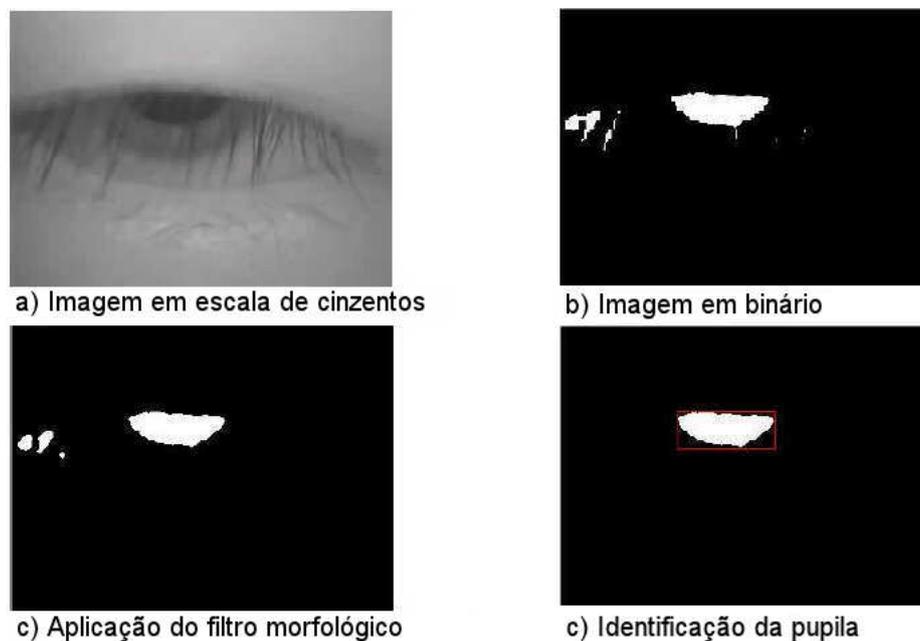


Figura 2.6: Captura e detecção da pupila numa situação em que os olhos estão semifechados, segundo [26].

calculando a distância de cada olho às extremidades da face, este autor conseguiu detetar a orientação da face da pessoa em análise.

Na secção seguinte serão descritas algumas técnicas de visão computacional para detetar a orientação do olhar.

2.2.4 Detecção da orientação do olhar

Para detetar a orientação do olhar é necessário detetar não só os olhos, mas também a pupila dos mesmos. Desta forma, comparando a distância das pupilas relativamente às extremidades dos olhos é possível extrapolar a orientação do olhar.

No trabalho descrito em [9], conforme já foi mencionado na secção 2.2.2, foi utilizado o OpenCV e a biblioteca STASM para detetar as pupilas.

Em [23] também foi desenvolvido um sistema para identificar a localização das pupilas com o objetivo de detetar a orientação do olhar da pessoa em questão. Para tal, este autor após detetar a face e os olhos da pessoa em questão aplicou as transformadas de Hough orientadas à deteção de círculos nas áreas da imagem digital correspondentes aos olhos. Deste modo, conseguiu localizar corretamente as pupilas, com exceção de algumas situações em que as mesmas estavam parcialmente cobertas pelas pálpebras, conforme está representado na figura 2.7.

Na secção seguinte será descrito o modo de detetar quando uma pessoa está a bocejar,

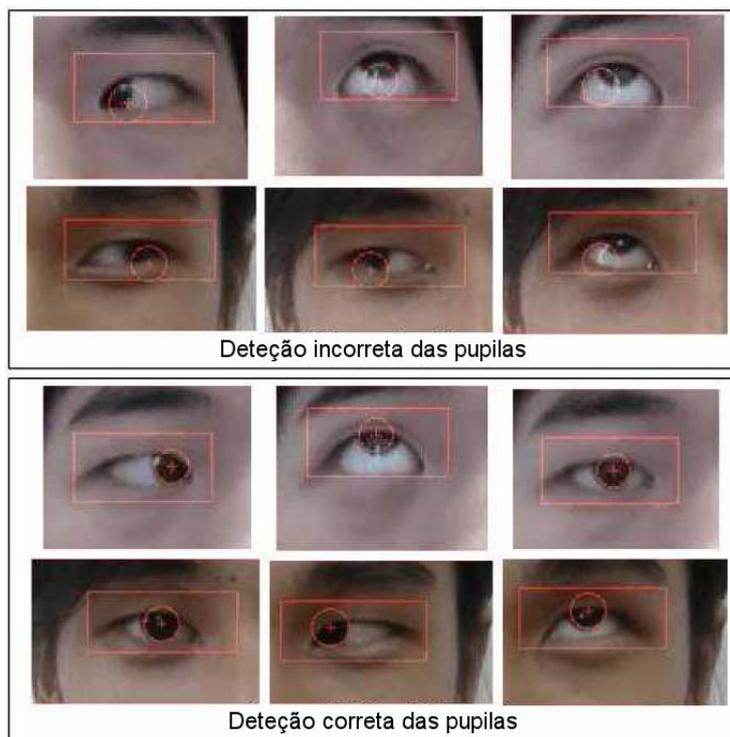


Figura 2.7: Exemplos da deteção correta e incorreta das pupilas, segundo [23].

recorrendo a técnicas de visão computacional.

2.2.5 Deteção de bocejos

De acordo com [24], quando uma pessoa boceja, abre mais a boca e durante muito mais tempo do que quando fala, pelo que tendo este facto em consideração torna-se possível distinguir estas duas ações.

Para detetar se o condutor tinha a boca aberta ou fechada, [1] recorreu à comparação dos níveis de intensidade dos histogramas da área da imagem correspondente à boca. Segundo este autor, quando a boca está aberta, a área da mesma contém mais pixels escuros do que quando está fechada, pelo que analisando esta ocorrência e tendo em consideração o intervalo de tempo durante o qual o condutor abre a boca, torna-se possível detetar quando o mesmo boceja.

Outra técnica que também pode ser utilizada para a deteção de bocejos baseia-se no trabalho [17]. O autor menciona a possibilidade de detetar se os olhos de uma pessoa estão abertos ou fechados de acordo com o número de arestas identificadas na imagem dos mesmos. De modo similar, este princípio pode também ser aplicado para detetar se uma pessoa tem a boca fechada ou aberta, pois o número de arestas é maior quando a boca está aberta do que quando está fechada.

Na secção seguinte serão descritas as conclusões retiradas ao longo deste capítulo.

2.3 Conclusões do capítulo

Verifica-se então, que de todos os métodos de deteção automática de fadiga que são atualmente estudados e analisados, os mais propostos são os não-intrusivos, por serem mais práticos e como o próprio nome indica, menos (ou nada) intrusivos. Destes, alguns são já implementados em automóveis e recorrem a métricas como a análise dos movimentos do volante, das alterações de velocidade e da posição do automóvel na faixa de rodagem.

Contudo, todos estes métodos restringem-se a serem utilizados apenas em automóveis, de forma a detetarem padrões de fadiga dos respetivos condutores exclusivamente durante o ato de condução. Pelo contrário, o ideal será um sistema que possa ser aplicado a qualquer pessoa e ao longo da execução de vários tipos de funções. Podendo desta forma ser introduzido não só em automóveis, mas também em diversas outras atividades do quotidiano, como linhas de produção industriais, e qualquer outra que requeira total concentração por parte do seu executor.

Para tal será necessário construir um sistema que meça padrões de fadiga da pessoa em questão, independentemente das ações que a mesma esteja a executar. Deste modo, este sistema deverá recorrer apenas a medições que possam ser feitas diretamente na pessoa, como o estado dos olhos e da boca, a direção do olhar e a posição da cabeça. O desenvolvimento de um sistema de deteção automática de fadiga com estes requisitos será apresentado no capítulo seguinte.

Capítulo 3

Sistema proposto

Neste capítulo será descrito o desenvolvimento de um sistema de detecção automática de fadiga, recorrendo exclusivamente a técnicas de visão computacional. O sistema utiliza uma única câmara para tentar detetar padrões de fadiga medindo a frequência com que a pessoa em questão pestaneja e quanto tempo demora a fazê-lo, se a mesma boceja e se a direção do seu olhar ou a posição da sua cabeça sugerem que está a adormecer.

Na secção seguinte será descrita a metodologia empregue para a detecção de fadiga no sistema proposto.

3.1 Metodologia para a detecção de fadiga

Nas figuras 3.1 e 3.2 está representado o fluxograma do sistema desenvolvido. De uma forma simplificada, a metodologia escolhida para a detecção de fadiga compreende os seguintes passos:

- 1) O sistema verifica se existe alguma face humana próxima do centro da imagem digital capturada pela câmara. Caso exista, mas não esteja suficientemente centrada, indica em que sentido a mesma deve ser movida, para que tal aconteça;
- 2) São capturados alguns parâmetros acerca de determinadas características da pessoa em questão, para que alguns gestos possam depois ser corretamente detetados;
- 3) O sistema utiliza a câmara para capturar a imagem da pessoa cujos sinais de fadiga devem ser detetados;
- 4) De seguida, tenta detetar e recortar dessa imagem a face da pessoa em questão:

- a) Caso a face não seja detetada com sucesso, verifica à quanto tempo a mesma não é detetada:
 - i) Caso a face não seja detetada com sucesso há mais de 30s salta para o passo 2);
 - ii) Caso a face tenha sido detetada há menos de 30s, salta para o passo 3).
- b) Caso a face seja detetada com sucesso, tenta detetar e recortar da mesma a área que contém os olhos:
 - i) Caso nenhum dos olhos seja detetado com sucesso avança para o passo iii);
 - ii) Caso algum dos olhos, ou ambos sejam detetados com sucesso, testa o seu estado:
 - (1) Caso os olhos detetados estejam fechados, verifica se os mesmos estão fechados há mais do que 317.3ms (AECS):
 - (a) Caso estejam fechados há mais do que 317.3ms, sinaliza que foi detetado um padrão de fadiga através da métrica AECS e avança para o passo (2);
 - (b) Caso não estejam fechados há mais do que 317.3ms, avança para o passo (2).
 - (2) Testa se durante o último minuto os olhos estiveram fechados durante mais do que 12s (PERCLOS):
 - (a) Caso durante o último minuto os olhos tenham estado fechados durante mais do que 12s, sinaliza que foi detetado um padrão de fadiga através da métrica PERCLOS, e salta para o passo iii);
 - (b) Caso durante o último minuto os olhos não tenham estado fechados durante mais do que 12s, avança para o passo iii).
 - iii) Recorta da imagem da face a área correspondente à boca, e testa a ocorrência de um bocejo:
 - (1) Caso seja detetado um bocejo, sinaliza que foi detetado um padrão de fadiga através da deteção de bocejos, e salta para o passo iv);
 - (2) Caso não seja detetado um bocejo avança para o passo iv).
 - iv) Verifica se anteriormente foi detetado algum olho:
 - (1) Caso anteriormente um, ou ambos os olhos tenham sido detetados, tenta localizar o nariz:
 - (a) Caso o nariz seja detetado, tenta detetar as pupilas:
 - (i) Caso uma pupila, ou ambas as pupilas sejam detetadas com sucesso, verifica se as mesmas evidenciam que a direção do olhar está demasiado baixa:

- [1] Caso as pupilas detetadas estejam demasiado baixas, é interpretado como a pessoa estando a adormecer e, como tal, sinaliza que foi detetado um padrão de fadiga através da inclinação da cabeça, e salta para o passo v);
 - [2] Caso as pupilas não estejam demasiado baixas, salta para o passo v).
- (ii) Caso nenhuma das pupilas seja detetada com sucesso, verifica apenas se a área da imagem detetada como correspondendo à zona onde estão localizados os olhos está demasiado baixa:
- [1] Caso a área dos olhos esteja demasiado baixa, é interpretado como a pessoa estando a adormecer e, como tal, sinaliza que foi detetado um padrão de fadiga através da inclinação da cabeça, e salta para o passo v);
 - [2] Caso a área dos olhos não esteja demasiado baixa, salta para o passo v).
- (b) Caso o nariz não seja detetado, salta para o passo v).
- (2) Caso anteriormente nenhum olho tenha sido detetado avança para o passo v);
- v) Verifica se foi detetado algum padrão de fadiga:
- (1) Caso tenha sido detetado algum padrão de fadiga, um alarme sonoro é reproduzido, e salta para o passo 3);
 - (2) Caso não tenha sido detetado nenhum padrão de fadiga, salta para o passo 3).

De notar que, alterando um, ou dois parâmetros num ficheiro de configuração, o sistema pode ser executado em seis modos diferentes, sendo eles:

- O modo de funcionamento normal, como descrito anteriormente;
- O modo de depuração, que é idêntico ao modo anterior, mas que ao longo da sua execução, descreve na linha de comandos os gestos que o sistema está a detetar, e exhibe, em tempo real, as imagens que estão a ser processadas, bem como as deteções efetuadas. Na figura 3.3 está representado um exemplo do output do sistema quando executado neste modo;
- O modo de teste, que consiste na execução normal do sistema, mas durante um período de tempo predefinido e com algumas adições com vista a registar a execução do teste¹;

¹Este modo de funcionamento será descrito na secção 4.1

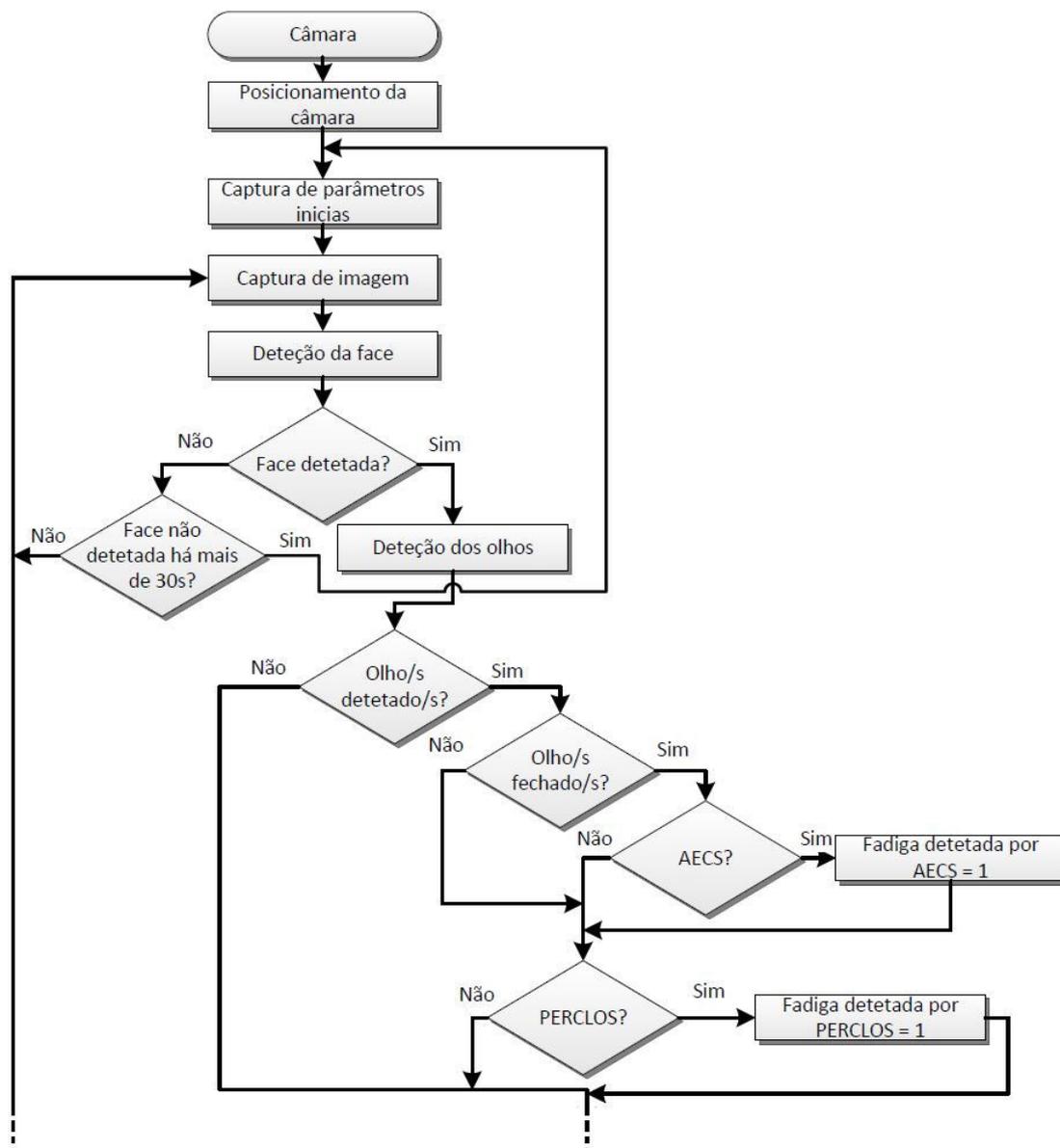


Figura 3.1: Primeira metade do fluxograma do sistema desenvolvido.

- O modo de calibração, que em vez de utilizar uma câmara como fonte de vídeo, utiliza um ficheiro de vídeo “avi”². Desta forma, é possível aperfeiçoar e parametrizar especificamente o sistema para a função e posição da pessoa filmada. Torna ainda possível utilizar vídeos para testar o desempenho do sistema;
- O modo de depuração em simultâneo com o modo de teste;
- O modo de depuração em simultâneo com o modo de calibração.

²Do Inglês *Audio Video Interleave*. Trata-se de um formato que combina áudio e vídeo, sendo um dos mais populares e compatíveis a nível mundial.

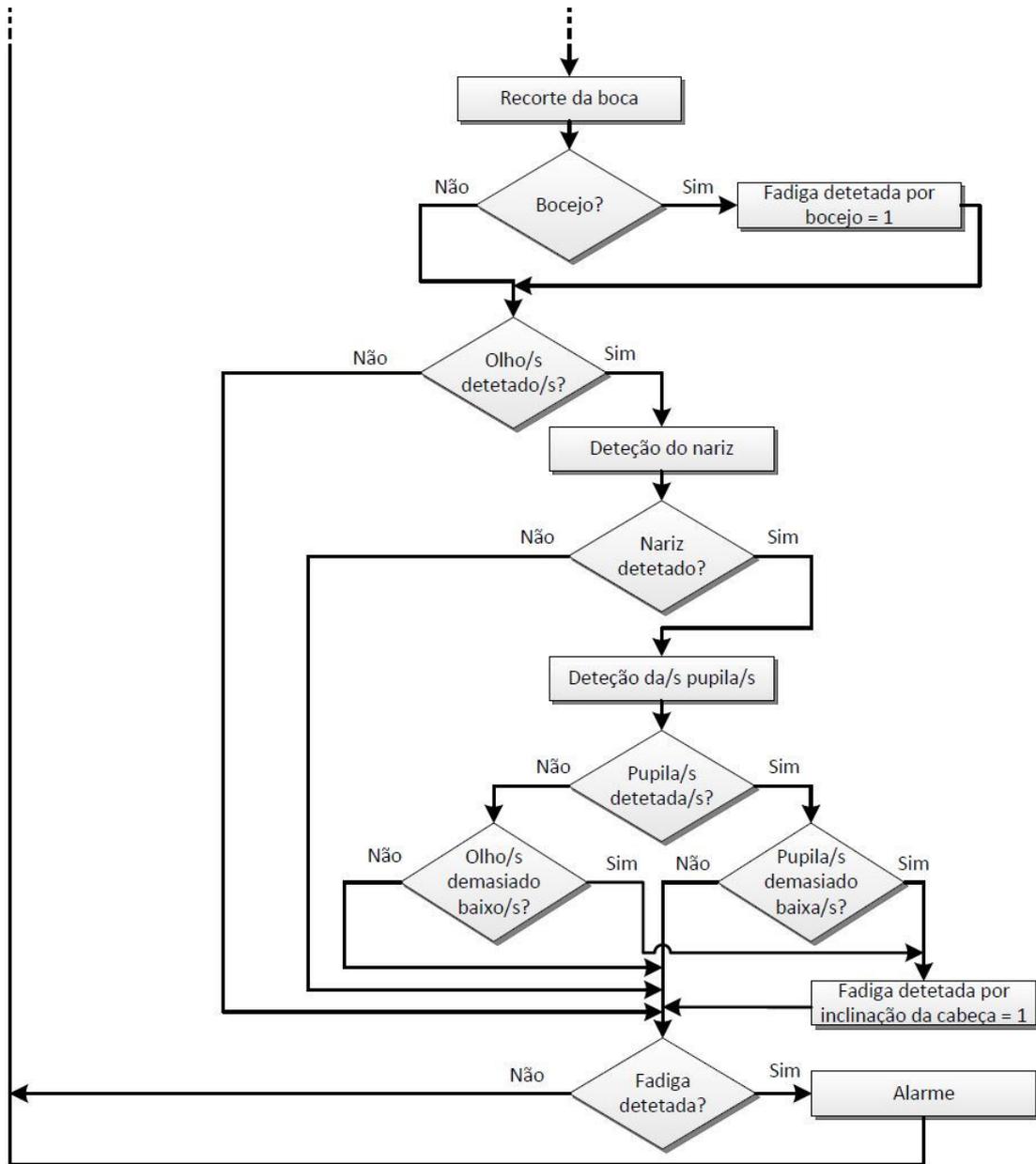


Figura 3.2: Segunda metade do fluxograma do sistema desenvolvido.

Na secção seguinte serão apresentadas as tecnologias utilizadas na implementação do sistema proposto.

3.2 Tecnologias utilizadas

O sistema desenvolvido depende grandemente da eficácia dos algoritmos responsáveis pela deteção dos olhos e, principalmente, da face da pessoa em questão. Por este motivo, a sua escolha tem um grande peso na eficiência final do sistema.



Figura 3.3: Exemplo do output do sistema em modo de depuração.

De realçar ainda que, o objetivo do sistema em questão consiste em detetar padrões de fadiga, numa pessoa e em tempo real. Pelo que, para além da deteção da face e dos olhos da pessoa, é ainda fundamental o acesso a ferramentas de visão computacional que permitam a manipulação de imagens digitais de uma forma eficaz e eficiente.

Por todos estes motivos, foi utilizada a biblioteca OpenCV, descrita na secção 1.2.1. Mais concretamente, esta biblioteca foi utilizada na execução de tarefas como:

- Captura de vídeo a partir de uma câmara;
- Conversão de imagens digitais para uma escala de cinzentos;
- Deteção da face, dos olhos e do nariz da pessoa capturada no vídeo;
- Recorte da área da imagem correspondente à face, olhos, nariz e boca da pessoa;
- Conversão de imagens digitais para uma escala binária (preto e branco);
- Deteção e enumeração de contornos nas imagens;
- Deteção e localização de contornos circulares nas imagens;

- Adição de texto às imagens de vídeo capturadas;
- Exibição de vídeo.

Como já foi referido na secção 1.2.1, o OpenCV está disponível em diversas linguagens de programação, entre as quais Python³. Esta é uma linguagem de programação de alto nível que possibilita a implementação de complicados algoritmos de uma forma rápida e eficaz. Esta característica, em conjunto com o OpenCV traz grandes vantagens na implementação do sistema proposto, conforme se pode constatar em [32]. Outra característica também muito vantajosa na utilização de Python consiste no facto desta ser executável nos sistemas operativos mais usados a nível mundial. Nomeadamente em Microsoft Windows, Linux/Unix e Mac OS X, sendo ainda portátil para as máquinas virtuais de Java e .NET.

De realçar ainda que, tanto a biblioteca de software OpenCV como a linguagem de programação Python podem ser utilizadas para fins pessoais, ou comerciais, sem qualquer encargo. Este facto, em conjunto com o uso do sistema operativo Linux possibilitou a criação de um sistema de deteção automática de fadiga sem qualquer custo em termos de software.

O sistema foi implementado no ambiente de desenvolvimento Microsoft Visual Studio Professional 2012 update 3, com o suplemento Python Tools for Visual Studio 1.5, de forma a possibilitar a utilização deste ambiente de desenvolvimento para a implementação em Python. A versão da biblioteca OpenCV utilizada foi a 2.4.5, e em termos de Python foi utilizada a distribuição Enthought Canopy⁴, que inclui o interpretador 2.7.3, bem como diversas bibliotecas de software.

O funcionamento do sistema proposto foi testado nos sistemas operativos Microsoft Windows 7 e Ubuntu 13.04. Para implementação e testes, em termos de hardware, foi utilizado um portátil equipado com um processador Intel i7-2630QM com 8GB de RAM e uma placa gráfica Nvidia GeForce GTX 560M com 2GB de memória. Para captura de vídeo foi utilizada uma câmara Trust SpotLight Webcam Pro⁵.

Na secção seguinte serão descritas as tarefas de segmentação de vídeo realizadas pelo sistema.

3.3 Segmentação de vídeo

A segmentação de vídeo consiste no processo de conseguir localizar e/ou extrair de uma imagem digital um conjunto de pixels com determinadas características. No sistema proposto, este processo foi utilizado para localizar e extrair de imagens digitais, o conjunto

³Mais informação sobre esta linguagem de programação pode ser encontrada em <http://www.python.org/>.

⁴Esta distribuição Python pode ser obtida em <https://www.enthought.com/downloads/>.

⁵Mais informação acerca desta câmara pode ser obtida em <http://www.trust.com/products/product.aspx?artnr=16428>.

de pixels que contém a face da pessoa em questão, os seus olhos, o seu nariz e a sua boca. Desta forma, a segmentação tem um papel fundamental, pois, por exemplo é impossível verificar se a boca e os olhos estão abertos ou fechados, se não for possível localizar na imagem digital o conjunto de pixels que constituem a boca e os olhos. Para esse efeito, a biblioteca OpenCV disponibiliza algumas ferramentas com o objetivo de realizar estes processos de segmentação.

Nas secções seguintes serão descritos os processos de segmentação de imagens utilizados no sistema proposto.

3.3.1 Detecção da face

A correta deteção da face da pessoa em questão nas imagens capturadas pela câmara assume no sistema um papel fundamental, pois sem a mesma, torna-se impossível detetar sinais de fadiga na pessoa.

O OpenCV inclui já todas as ferramentas necessárias para uma correta identificação de faces em imagens digitais, pelo que foi esse o método utilizado. Para tal, esta biblioteca inclui alguns filtros (classificadores Haar em cascata) especializados na deteção de faces de pessoas. Filtros esses, que foram criados através do algoritmo de aprendizagem automática AdaBoost, e que são utilizados por esta biblioteca no processo de deteção de características em imagens digitais, conforme descrito na secção 1.2.1.

A biblioteca OpenCV disponibiliza vários filtros especializados na deteção de faces de pessoas, uns mais indicados em certas circunstâncias do que outros. Por exemplo, o filtro “haarcascade_frontalface_default.xml” é normalmente o mais utilizado, contudo, se a face da pessoa estiver em perfil, o filtro “haarcascade_profileface.xml” é mais eficaz.

Por este motivo, e dada a importância da correta localização da face, optou-se pela utilização dos cinco filtros disponibilizados pelo OpenCV. De modo a que, se utilizando o primeiro filtro não for detetada nenhuma face, será utilizado o filtro seguinte, e procede-se deste modo até que seja encontrada uma face. Se após a utilização dos cinco filtros, tal não acontecer, a imagem digital atual é descartada, passando-se à seguinte.

Na figura 3.4 estão representadas algumas faces detetadas pelo sistema desenvolvido. Como se pode observar, utilizando os cinco filtros em cadeia, o sistema consegue detetar com sucesso as faces em diversas posições, o que não seria possível se se recorresse apenas a um dos cinco filtros.

De notar que a área da imagem digital que contém a face está ligeiramente expandida no limite inferior. Essa adição foi feita de modo a que, quando a pessoa em questão bocejasse, a parte inferior da boca não ficasse excluída da face, conforme por vezes acontecia se se confiasse unicamente na área indicada pelo OpenCV. Este assunto será descrito na secção 3.3.5.



Figura 3.4: Exemplo da detecção de uma face em diversas posições.

Após a detecção da área da imagem digital que contem a face, essa zona é recortada e utilizada nas tarefas seguintes, como a detecção dos olhos, que está descrita na próxima secção.

3.3.2 Detecção dos olhos

A detecção dos olhos é fundamental para que o sistema funcione em pleno, pois sem essa detecção não podem ser calculadas as métricas AECS e PERCLOS, bem como a detecção da orientação do olhar. Por este motivo foram feitos vários testes às ferramentas disponibilizadas pelo OpenCV para a detecção de olhos em imagens digitais:

1. Foi testada a detecção de ambos os olhos ao mesmo tempo, recorrendo aos filtros específicos para o efeito, e verificou-se que os mesmos tinham um bom desempenho em faces semelhantes às existentes em fotografias do tipo passaporte. Contudo, se a face estive inclinada, ou em perfil a detecção não era feita, ou estava incorreta, confundindo algumas vezes as narinas com os olhos;
2. Para melhorar a detecção de ambos os olhos, optou-se pelo uso do mesmo esquema de filtros em cascata, conforme foi utilizado na detecção da face. Para esse efeito foram utilizados os três filtros disponibilizados pelo OpenCV, em que o último da cadeia, é orientado à detecção dos olhos quando cobertos por óculos;
3. Para mitigar as hipóteses de uma detecção incorreta, a área da face onde se tenta localizar os olhos foi reduzida para os $2/5$ superiores, resolvendo assim o problema da confusão das narinas com os olhos;
4. Para minimizar o problema da falta de detecção dos olhos quando a face está em perfil ou inclinada, optou-se por tentar a detecção dos mesmos individualmente. Utilizando para o efeito os dois filtros disponibilizados pelo OpenCV, um para cada olho;
5. Para reduzir as detecções erradas, também aqui se reduziu a área de procura. A mesma foi limitada verticalmente à zona compreendida entre os $1/10$ e os $2/5$ da face (do limite superior, para o limite inferior). Horizontalmente, foi limitada dos 0 aos $3/5$ (da esquerda para a direita) para o olho direito, e dos $2/5$ até ao limite da face, para o olho esquerdo;

6. Por último, para ter mais hipóteses de detetar os olhos, o sistema começa por tentar deteta-los individualmente, e caso não tenha sucesso passa à deteção de ambos em simultâneo. Caso não consiga detetar nenhum dos olhos, as métricas AECS e PERCLOS, bem como a orientação do olhar e da face não serão consideradas na imagem atual.

De notar que, desta forma, o sistema vai sempre localizar a área de cada olho em separado, independentemente da deteção dos olhos ser feita individualmente ou em conjunto. Isto possibilita que, numa imagem possa ser detetado apenas um olho e que as métricas AECS e PERCLOS, e a deteção da orientação do olhar possam ser feitas baseando-se apenas no olho detetado. Nestas situações, se a deteção fosse exclusivamente baseada na deteção de ambos os olhos em simultâneo, nenhum dos olhos seria detetado, e estas métricas não poderiam ser calculadas, podendo, inadvertidamente, ignorar-se algum padrão de fadiga.

Na figura 3.5 estão representados alguns exemplos da deteção de olhos em diferentes estados e posições. Como se pode verificar na segunda linha de imagens, quando a face está em perfil, apenas é detetado um dos olhos.

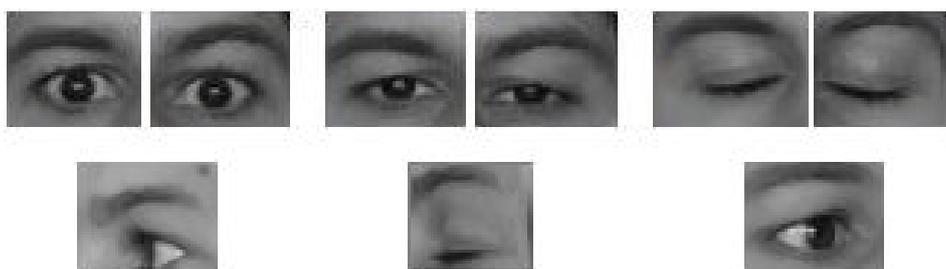


Figura 3.5: Exemplo da deteção dos olhos em diferentes estados e posições.

Na secção seguinte será descrito o processo de deteção de pupilas.

3.3.3 Deteção das pupilas

A deteção das pupilas é principalmente utilizada em sistemas de rastreio do olhar, ou seja, como forma de detetar o local para onde uma pessoa está a olhar. Nos sistemas de deteção automática de fadiga, o rastreio do olhar é útil para detetar as situações em que uma pessoa começa a adormecer, e a direção do seu olhar começa a descer. Foi neste sentido que a deteção das pupilas foi utilizada no sistema proposto.

A deteção das pupilas foi baseada no trabalho [23], e foi efetuada recorrendo a algumas ferramentas disponibilizadas pelo OpenCV. O processo de deteção das pupilas, implementado no sistema proposto, está representado na figura 3.6, e pode ser descrito nos seguintes passos:

1. A imagem digital que contém a área de um olho (a)) é recortada de forma a conter

pouco mais do que a área onde é provável que se encontre a pupila (b)). Isto é feito de forma a reduzir as hipóteses de uma deteção incorreta da pupila;

2. A imagem digital que contém o recorte da área de um olho (b)) é convertida para binário (branco e preto), conforme representado na imagem c);
3. Seguidamente é aplicado o detetor de arestas Canny, de modo a criar uma nova imagem digital constituída apenas pelas arestas da anterior. A esta, é aplicada a transformada de Hough orientada a círculos, de forma a localizar um círculo, que deve consistir na pupila. A pupila detetada neste exemplo, está circunscrita a branco na imagem d).

O detetor de arestas Canny é, no OpenCV, parte integrante do método que aplica a transformada de Hough. Segundo [10], o detetor de arestas Canny e a deteção de círculos baseada na transformada de Hough são dois dos algoritmos mais utilizados para os respectivos fins, sendo dos mais eficientes.

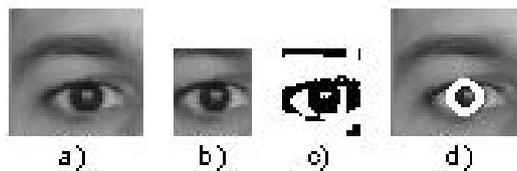


Figura 3.6: Representação do processo de deteção das pupilas.

De notar ainda, que embora este processo seja orientado à deteção de pupilas, no exemplo da figura 3.6, o mesmo detetou a íris e não a pupila. Este facto deveu-se à cor escura da íris em questão, o que torna a sua diferenciação relativamente à pupila muito difícil. Contudo, para o objetivo de detetar se a orientação do olhar de uma pessoa está demasiado baixa, é indiferente detetar a pupila, ou a íris, pelo que esse facto foi ignorado.

Na secção seguinte será descrito o processo de deteção do nariz.

3.3.4 Deteção do nariz

A deteção do nariz foi implementada, recorrendo ao único filtro que o OpenCV disponibiliza para o efeito. Contudo, apesar da utilização de apenas um único filtro, este mostrou-se bastante fiável, pelo que não foi tentada qualquer outro tipo de abordagem. Na figura 3.7 está representada a imagem de um nariz, detetado recorrendo ao OpenCV.



Figura 3.7: Exemplo da imagem de um nariz, detetado recorrendo ao OpenCV.

Na secção seguinte será descrito o processo de deteção da boca.

3.3.5 Detecção da boca

Tal como para a deteção da face, dos olhos e do nariz, o OpenCV também disponibiliza alguns filtros para a deteção da boca. Contudo, os mesmos não se apresentaram suficientemente eficientes para serem utilizados no sistema proposto. Alguns problemas encontrados, relacionaram-se com a posição e estado da boca. Por exemplo, se a face estivesse consideravelmente inclinada para cima, em vez da boca, por vezes era detetado o queixo. No entanto, o maior entrave à sua utilização consistiu no facto de que, por vezes, durante um bocejo, a área da imagem digital detetada como contendo a boca, consistia apenas na metade inferior, ou superior, da boca. Este facto fazia com que qualquer algoritmo de deteção de bocejos não obtivesse os resultados pretendidos, pelo que optou-se por uma abordagem diferente.

Conforme está exemplificado na figura 3.8, em vez de confiar a deteção da boca ao OpenCV, a mesma foi considerada como consistindo na metade inferior da face. Desta forma, sempre que a face seja detetada, a boca também será.



Figura 3.8: Exemplo do recorte da área da imagem digital que contém a boca.

Na secção seguinte serão descritas as técnicas utilizadas na deteção dos gestos que podem corresponder a padrões de fadiga.

3.4 Deteção de gestos

A deteção da face, olhos, nariz, boca e pupilas, conforme descrito nas secções anteriores, têm o objetivo de permitir a deteção de gestos que possam evidenciar sinais de fadiga.

Nas secções seguintes serão descritas as técnicas utilizadas no sistema proposto, para detetar esses gestos.

3.4.1 Deteção do ato de pestanejar

Para detetar quando uma pessoa pestaneja é necessário, não só detetar e rastrear os seus olhos, como detetar quando os mesmos estão fechados ou abertos.

A deteção e rastreamento dos olhos foi efetuada conforme descrito na secção 3.3.2, ou seja, utilizando o OpenCV.

Após essa deteção foi utilizada a técnica descrita em [17] e [12] para diferenciar o estado dos olhos. Ou seja, foi analisada a intensidade média de cada linha horizontal de pixels que constituíam as imagens digitais dos olhos, para localizar verticalmente as sobrancelhas e as pestanas, quando os olhos estão fechados, ou as sobrancelhas e as pestanas superiores em conjunto com a pupila e a íris, quando estes estão abertos, dado que estas constituem as áreas mais escuras de cada uma dessas imagens.

Dado que a distância entre as sobrancelhas e as pestanas quando os olhos estão fechados é maior do que a distância entre as sobrancelhas e a pupila em conjunto com as pestanas superiores e a íris quando estes estão abertos, tornou-se possível detetar o estado dos olhos. Este facto está representado na figura 3.9, onde podemos verificar que a distância entre os dois mínimos de intensidade das linhas horizontais de pixels da imagem digital que contém um olho fechado (linha tracejada), é superior à mesma distância na imagem digital que contém um olho aberto (linha preenchida).

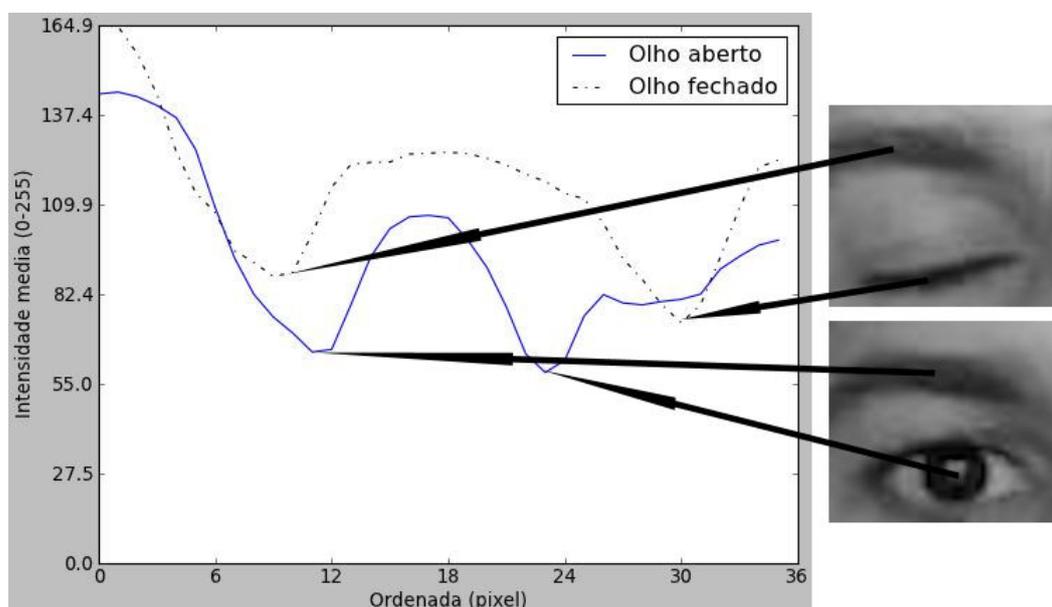


Figura 3.9: Exemplo da distância entre as maiores quebras de intensidade da cor em imagens com os olhos fechados e abertos.

Deste modo, efetuando estes cálculos sempre que um, ou ambos os olhos sejam detetados, torna-se possível identificar o estado dos mesmos.

Na secção seguinte será descrita a técnica implementada para a deteção do estado da boca.

3.4.2 Deteção do estado da boca

Como já foi referido na secção 3.3.5, a boca foi considerada como estando localizada na metade inferior da imagem digital que contém a face. Deste modo, sempre que a face seja detetada, também a boca será.

Depois de obter a imagem digital da boca da pessoa em questão, importa detetar o estado da mesma, ou seja, se esta se encontra aberta, ou fechada. Para esse efeito, inicialmente foi testada a técnica mencionada por [1] e descrita na secção 2.2.5. Esta técnica é semelhante à utilizada no sistema proposto, para a deteção do estado dos olhos, ou seja, é baseada nas diferenças de intensidades entre os diferentes estados da boca. Segundo este autor, quando a boca está aberta, a correspondente imagem digital contém mais pixels escuros do que quando está fechada. Contudo esta técnica foi testada e não obteve resultados satisfatórios, resultando em diversas deteções de estado erradas, pelo que foi abandonada.

O motivo que levou à obtenção de resultados insatisfatórios está relacionado com o facto de que quando a boca está aberta, embora apareça na área da mesma um conjunto mais escuro de pixels, também aparece um conjunto de pixels mais claro, que corresponde aos dentes. Este facto tem especial incidência durante um bocejo, como é visível na figura 3.10, dado que, durante o mesmo, normalmente, os dentes são bem visíveis. Deste modo a área mais escura que corresponde ao interior da boca acaba por ser atenuada pela área mais clara que corresponde aos dentes, não existindo desta forma, uma evidente mudança de intensidade da cor entre a boca fechada e um bocejo.



Figura 3.10: Exemplo de uma imagem digital que representa um bocejo.

Tentou-se ainda localizar o lábio superior, e o lábio inferior, detetando na imagem digital da boca os dois contornos de maior comprimento. Contudo, nem sempre os lábios eram detetados como um único contorno, pelo que nem sempre eram estes os dois maiores contornos. De realçar ainda, que por vezes o maior contorno consistia no queixo, pelo que também esta técnica foi abandonada.

Outra técnica testada consistiu em medir a distância entre os olhos e o limite inferior da boca. Pois, quando a boca está aberta, esta distância é maior do que quando a mesma está fechada. Contudo, conforme já foi referido na secção 3.3.5, por vezes, através do OpenCV, a boca era detetada como consistindo apenas na metade inferior, ou superior da mesma, pelo que esta métrica foi abandonada. Posteriormente, a boca passou a ser considerada como a metade inferior da face, de forma a corrigir esse problema. Contudo, esta métrica não poderia ser utilizada, dado que deste modo, a área da boca teria sensivelmente o mesmo tamanho, indiferentemente do seu estado, logo a distância do seu limite inferior relativamente aos olhos também seria sempre semelhante.

Por este motivo, foi utilizada uma técnica mencionada em [17] e descrita na secção 2.2.5. Ou seja, foi analisado o número de contornos existentes na imagem digital que continha a boca, e verificou-se que o seu número aumentava consideravelmente sempre que a boca estava aberta, relativamente a quando a mesma se encontrava fechada.

A detecção dos contornos foi realizada utilizando dois métodos disponibilizados pelo OpenCV:

1. Em primeiro lugar foram detetadas as arestas recorrendo ao detetor de arestas Canny;
2. De seguida estas arestas foram interligadas, dando origem aos contornos, através de um método específico do OpenCV para detecção de contornos.

Após a detecção dos contornos da imagem digital que contém a boca, é possível contar os mesmos através da estrutura de dados resultante da detecção de contornos do OpenCV.

Na figura 3.11 estão representadas quatro imagens digitais. As imagens a) e c) representam uma boca no estado aberta, e fechada, respetivamente. As imagens b) e d) correspondem às imagens a) e c), respetivamente, após a detecção de contornos. Na mesma figura está ainda representado um gráfico que demonstra a diferença da quantidade de contornos existentes em cada uma dessas imagens. É clara a diferença do número de contornos detetados na imagem digital que contém a boca aberta (cerca de 600 contornos), em relação à imagem digital que contém a boca fechada (cerca de 310 contornos).

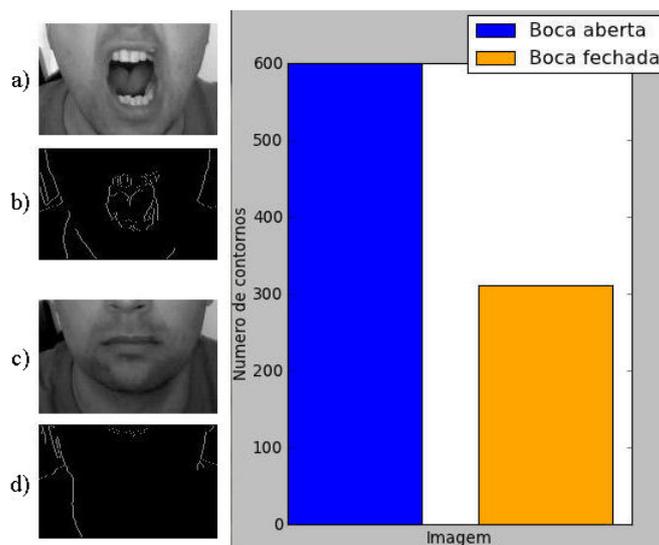


Figura 3.11: Exemplo da detecção e contagem do número de contornos em imagens digitais que contêm a boca fechada ou aberta.

Na secção seguinte será descrita a técnica aplicada para efetuar a detecção da inclinação do olhar.

3.4.3 Detecção da inclinação do olhar e da cabeça

Nos sistemas de detecção automática de fadiga, a inclinação do olhar e da cabeça têm especial relevância para detetar quando a pessoa em questão está a adormecer. Mais

concretamente quando essa pessoa está a deixar descair a cabeça e/ou o olhar.

Ao comparar imagens digitais de faces direitas e inclinadas, verificou-se que a distância das pupilas e dos olhos, relativamente ao topo da face era menor quando a cabeça estava inclinada para baixo, relativamente a quando a mesma estava direita. Como tal, para detetar a orientação do olhar, foi medida a distância das pupilas relativamente ao topo da face. Caso nenhuma das pupilas fosse detetada com sucesso, por sua vez, era medida a distância dos olhos relativamente ao topo da face.

Esta diferença de distâncias está representada na figura 3.12. Como se pode observar nessa figura, em a), a distância das pupilas relativamente ao topo da face é superior à correspondente distância em b).

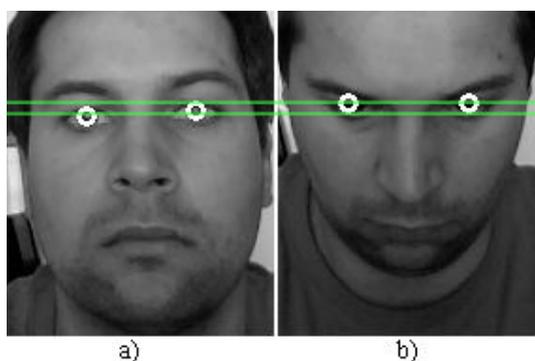


Figura 3.12: Exemplo da diferença da distância entre as pupilas e o topo da face, entre quando a cabeça está direita (a)) ou inclinada (b)).

Contudo, após alguns testes, verificou-se que o local mais indicado para a colocação da câmara seria aproximadamente à altura dos ombros da pessoa em questão, ou ligeiramente abaixo. Isto, porque é normalmente a esta altura que a pessoa realiza as suas ações, seja durante a condução de um automóvel, ou numa linha de montagem industrial, pelo que é para essa área que a face deve estar orientada. Deste modo será colocando-se também nessa zona que a câmara conseguirá capturar mais eficazmente a face. Constatou-se então, que ao colocar a câmara nesta localização, a diferença de distâncias representada na figura 3.12 deixava de existir, pelo que foi procurada uma abordagem diferente.

Foi então, que se verificou que a distância do extremo superior dos olhos e das pupilas, relativamente ao extremo inferior do nariz aumentava conforme a pessoa em questão inclinasse a cabeça para baixo. Este facto está representado na figura 3.13. Como tal, foi este o método utilizado para detetar a inclinação da cabeça e do olhar.

Na secção seguinte será descrita a forma como todos os gestos detetados pelo sistema proposto são utilizados para detetar padrões de fadiga.

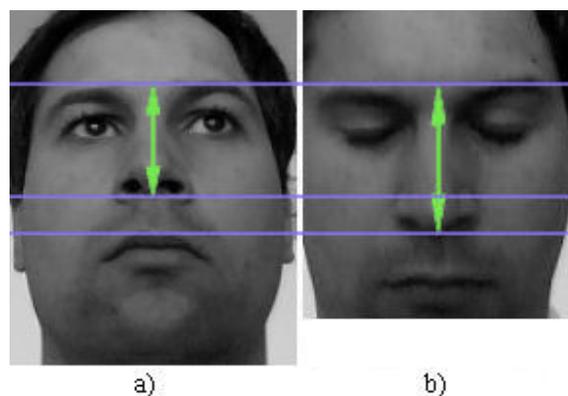


Figura 3.13: Exemplo da diferença da distância entre o extremo superior dos olhos em relação ao extremo inferior do nariz, entre quando a cabeça está direita (a)) ou inclinada (b)).

3.5 Sistema de classificação de gestos

O sistema proposto começa por ajustar a posição da câmara, de forma a capturar corretamente a face da pessoa a monitorizar. De seguida, passa à captura de alguns parâmetros necessários para a deteção de alguns gestos.

Na secção seguinte serão descritos estes dois passos iniciais.

3.5.1 Inicialização do sistema

Quando o sistema é iniciado, o mesmo começa por verificar se a câmara está orientada para uma localização que lhe permite capturar corretamente a face de uma pessoa. Caso tal não seja possível, o sistema informa em que sentido se deve movimentar, ou orientar a câmara de forma a corrigir a área capturada pela mesma. Apesar de, nesta fase o sistema apenas analisar se a face se encontra próxima do centro da imagem capturada, o mesmo obtém o melhor desempenho quando a câmara é posicionada aproximadamente ao nível dos ombros da pessoa a monitorizar.

Só após a deteção da face na zona central das imagens capturadas pela câmara, é que o sistema avança para o passo seguinte. Este consiste na captura de alguns parâmetros necessários para que o sistema possa detetar alguns gestos. Estes parâmetros são capturados durante o período de tempo predefinido num ficheiro de configuração, de forma a que o mesmo possa ser facilmente alterável. Os parâmetros capturados são:

1. O valor médio da distância entre os dois mínimos de intensidade das linhas horizontais de pixels da imagem digital que contêm cada olho;
2. O número médio de contornos existente nas imagens digitais que contêm a boca;

3. O valor médio da distância entre o extremo superior dos olhos e o extremo inferior do nariz;
4. O valor médio da distância entre o centro das pupilas e o extremo inferior do nariz.

Depois da obtenção destes parâmetros iniciais, sempre que o valor da distância entre os dois mínimos de intensidade das linhas horizontais de pixels da imagem digital que contêm cada olho seja maior que uma predefinida percentagem do valor encontrado no ponto 1, considera-se que a pessoa em questão tem os olhos fechados. A percentagem aqui mencionada é facilmente alterável num ficheiro de configuração.

De igual forma, também ao longo da execução do sistema, sempre que o número de contornos existente nas imagens digitais que contêm a boca for superior a uma predefinida percentagem do valor encontrado no ponto 2, considera-se que a pessoa em questão tem a boca aberta. Também esta percentagem é facilmente alterável num ficheiro de configuração.

Sempre que durante o funcionamento do sistema, a distância entre o extremo superior dos olhos e o extremo inferior do nariz for superior a uma predefinida percentagem do valor encontrado em 3, considera-se que a pessoa em questão tem a cabeça inclinada para baixo. De igual forma, sempre que a distância entre o centro das pupilas e o extremo inferior do nariz for superior a uma predefinida percentagem do valor encontrado em 4, considera-se que a pessoa em questão tem a cabeça inclinada para baixo. Estes dois valores percentuais são também alteráveis no mesmo ficheiro de configuração.

De notar que, a captura dos parâmetros iniciais é repetida sempre que o sistema não conseguir detetar uma face nas imagens capturadas durante mais do que 30s. Este procedimento visa reorientar a câmara e reobter os parâmetros iniciais adequados quando a pessoa monitorizada é substituída, como acontece frequentemente, por exemplo, numa linha de montagem industrial. O intervalo de tempo aqui mencionado (30s) pode ser alterado no ficheiro de configuração.

Na secção seguinte será descrita a forma como os padrões de fadiga são detetados.

3.5.2 Detecção de padrões de fadiga

Após ser possível ao sistema detetar quando a pessoa em questão tem os olhos fechados ou abertos, a boca fechada ou aberta, e ainda a inclinação do olhar ou da cabeça, é necessário utilizar este conhecimento para detetar padrões de fadiga.

Relativamente ao estado dos olhos, este foi utilizado para calcular as métricas AECS e PERCLOS. Deste modo, sempre que a pessoa monitorizada fecha os olhos durante mais do que 317,3ms⁶ seguidos, esta métrica indica que a pessoa apresenta sinais de fadiga.

⁶Intervalo de tempo indicado por [9], como sendo o máximo que uma pessoa sem fadiga demora a pestanejar.

De notar que este intervalo de tempo é alterável num ficheiro de configuração incluído no sistema.

Quanto à métrica PERCLOS, a mesma relata um estado de fadiga quando a pessoa em questão tem os olhos pelo menos 80% fechados, durante mais do que 20% de um minuto. Uma vez mais, esta percentagem e intervalo de tempo são alteráveis no mesmo ficheiro de configuração.

Relativamente ao estado da boca, o mesmo foi utilizado para detetar bocejos. Um bocejo é considerado um padrão de fadiga, e o sistema considera a deteção de um bocejo, sempre que a pessoa em questão tem a boca muito aberta durante mais do que 1,5s. Também este intervalo de tempo é alterável no mesmo ficheiro de configuração.

Relativamente à inclinação do olhar e da cabeça, sempre que a pessoa monitorizada tem a cabeça inclinada para baixo durante mais do que 1,5s, é considerado que a pessoa adormeceu, ou que está a adormecer e como tal, é considerado um sinal de fadiga. Uma vez mais, também este intervalo de tempo é alterável no ficheiro de configuração.

Por último, nesse mesmo ficheiro de configuração, é possível atribuir um peso percentual a cada métrica de fadiga aqui mencionada, ou seja, AECS, PERCLOS, bocejos e inclinação da cabeça. Quer isto dizer que, apenas alterando o ficheiro de configuração é possível:

1. Detetar fadiga através de uma única métrica, colocando a mesma com um peso de 1.00 (que corresponde a 100%), e as restantes com um peso de 0.00 (que corresponde a 0%);
2. Detetar fadiga através de algumas métricas, colocando as mesmas com um peso de 1.00 (que corresponde a 100%), e as restantes com um peso de 0.00 (que corresponde a 0%);
3. Detetar fadiga através de todas as métricas, atribuindo a todas um peso de 100%;
4. Detetar fadiga através da combinação de diversas métricas, colocando no peso destas, valores entre os 0.00 e os 1.00. Deste modo, só quando a soma dos pesos das métricas através das quais se detetou fadiga, for superior a 1.00, é que o sistema informa que foi detetada fadiga.

No sistema proposto, sempre que é detetada fadiga, é reproduzido um som de alarme e na linha de comandos é descrito qual das métricas despoletou o alarme.

Deste modo, o sistema consegue, em tempo real, detetar padrões de fadiga através de cinco métricas diferentes. Caso, por algum motivo, alguma destas métricas não se aplique devido à ação que a pessoa em questão está a desempenhar, a mesma pode ser desativada, atribuindo-lhe um peso de 0.00. Poderá ser o caso, por exemplo, quando a pessoa está a desempenhar uma função que a obrigue a dirigir constantemente o seu olhar para uma zona baixa. Como por exemplo, um operador de gruas, cujo habitáculo esteja localizado

no topo da mesma. Neste caso não faria sentido avisar o operador da grua sempre que o mesmo inclinasse a sua cabeça para baixo, pelo que teria de se desativar a deteção de fadiga através da inclinação da cabeça.

Na secção seguinte será feito um breve resumo deste capítulo.

3.6 Resumo do capítulo

Desta forma, foi construído um sistema de deteção automática de fadiga não-intrusivo, recorrendo apenas a técnicas de visão computacional para recolher e analisar dados acerca do estado da pessoa cujos sinais de fadiga se pretendem detetar. Apesar do mesmo utilizar apenas uma câmara para recolher esses dados, consegue detetar cinco padrões de fadiga diferentes. Nomeadamente o intervalo de tempo que uma pessoa demora a pestanejar (AECS) e a frequência com que o faz (PERCLOS), bocejos, e a inclinação descendente exagerada da cabeça e do olhar.

O sistema proposto é ainda parametrizável de forma a poder ser facilmente adaptado a diversos ambientes de funcionamento. Pode inclusive, alertar sempre que deteta um dos cinco padrões de fadiga, quando detetar apenas alguns destes, ou ainda, quando detetar uma determinada combinação dos mesmos.

Este sistema é também facilmente integrável em qualquer computador da atualidade, dado que pode ser executado nos sistemas operativos Microsoft Windows, Linux/Unix e Mac OS X, sendo ainda portátil para as máquinas virtuais de Java e .NET.

No capítulo seguinte serão apresentados os resultados obtidos na realização de alguns testes ao sistema aqui apresentado.

Capítulo 4

Resultados e discussão

Para além dos inúmeros testes realizados ao sistema proposto durante o seu desenvolvimento, após a sua finalização foram efetuados vários testes com diversas pessoas, com vista a avaliar o desempenho do mesmo.

Na secção seguinte será descrita a metodologia utilizada para a execução de testes ao sistema, com várias pessoas.

4.1 Metodologia para a execução de testes

Para a execução de testes com utilizadores, foi criado um modo de execução específico, designado por modo de testes. Neste modo, para além das tarefas normais que o sistema desempenha, descritas no capítulo 3, são também efetuadas outras, nomeadamente:

- Reprodução de uma música relaxante, com o objetivo de induzir sono;
- Controlo do tempo de execução do teste. O teste terá a duração que for introduzida num campo do ficheiro de configuração;
- As imagens de cada teste, capturadas pela câmara e processadas pelo sistema são guardadas em disco e através destas é gerado um ficheiro de vídeo “avi”. Deste modo, cada teste fica registado num ficheiro de vídeo;
- Por último, nas imagens do ficheiro de vídeo em que foi detetado algum padrão de fadiga, é descrito textualmente qual o padrão detetado. Esta descrição é repetida em algumas das imagens seguintes, de forma a que a mensagem possa ser lida enquanto se visiona o vídeo do teste efetuado.

Após a implementação do sistema, foi pedido a algumas pessoas que desempenhassem algumas ações em locais onde o sistema proposto estava a ser executado em modo de teste.

Deste modo, foram realizados doze testes, de cinco minutos cada, com cinco pessoas diferentes e em locais com condições de luminosidade variadas. As ações realizadas variaram entre a simulação da condução de um automóvel¹, a resolução de um cubo de Rubik e a realização de algumas tarefas num computador portátil. De notar que, as duas últimas ações mencionadas tiveram o objetivo de simular qualquer outra ação realizada, respetivamente, à altura de uma mesa de refeições ou à altura de uma secretária. Podendo essas ações consistir, por exemplo, no trabalho realizado numa linha de montagem, ou na monitorização de imagens de videovigilância.

Na tabela 4.1 foram anotadas não só as condições de luminosidade do local onde cada teste foi realizado, mas também as ações efetuadas pela pessoa monitorizada e ainda alguns aspetos físicos e gestuais das mesmas, que se verificaram influentes no desempenho do sistema.

Na secção seguinte são apresentados os resultados obtidos através dos testes efetuados.

4.2 Resultados dos testes com utilizadores

Nas tabelas 4.2 e 4.3 estão apresentados os resultados obtidos em cada um dos testes efetuados com utilizadores. Estes resultados foram extraídos, visualizando o vídeo gerado pelo sistema em cada um dos testes. Deste modo, foi possível comprovar se as anotações inscritas nas imagens de vídeo estavam de acordo com os gestos realizados no mesmo instante, pelas pessoas monitorizadas.

Para cada teste foi medido o desempenho individual de cada métrica, e ainda o desempenho global das cinco métricas em conjunto. Estes dois valores percentuais estão anotados, respetivamente, nas duas últimas colunas das tabelas 4.2 e 4.3. Para medir o desempenho individual de cada métrica foram contabilizados o número de alarmes corretos, incorretos e em falta ocorridos ao longo de cada teste:

- Os alarmes corretos correspondem àqueles que foram lançados corretamente pelo sistema, tendo os correspondentes gestos sido observados ao visualizar o respetivo vídeo de teste;
- Os alarmes incorretos correspondem àqueles que foram lançados pelo sistema, mas que os correspondentes gestos não foram observados ao visualizar o vídeo de teste;
- A falta de alarmes corresponde às situações em que os gestos observados nos vídeos de teste indicaram padrões de fadiga que deveriam ser detetáveis pelas respetivas métricas, mas que não foram detetados, e que como tal, o alarme não foi acionado.

¹Dentro de um automóvel, mas com este estacionado.

Teste	Luminosidade	Ação	Aspetos relevantes
1	Média, natural	Tarefas no portátil	
2	Alta, natural	Condução de um automóvel	A luz incidia mais na metade esquerda da face
3	Média, artificial	Tarefas no portátil	Bocejos pouco definidos
4	Média, artificial	Tarefas no portátil	As mãos eram muitas vezes colocadas na face
5	Média, artificial e natural	Tarefas no portátil	Cara escurecida e a luz incidia mais na metade esquerda da face
6	Média, artificial	Resolução do cubo de Rubik	Cara sombreada, principalmente a área dos olhos
7	Alta, natural	Resolução do cubo de Rubik	Olhos sempre semifechados
8	Média, natural	Tarefas no portátil	
9	Baixa, natural	Resolução do cubo de Rubik	Olhos sempre semifechados, cara escurecida, e bocejos muito pouco definidos
10	Baixa, natural	Condução de um automóvel	
11	Alta, natural	Tarefas no portátil	Pessoa usa óculos
12	Média, natural	Tarefas no portátil	

Tabela 4.1: Condições de luminosidade, ações efetuadas e aspetos relevantes, relativamente a cada teste.

Como se pode verificar nas tabelas 4.2 e 4.3, a eficácia global do sistema no conjunto de testes variou entre os 0,0% e os 96,9%. Já a eficácia de cada uma das métricas variou da seguinte forma:

- AECS - Entre 0,0% e 96,4%;
- PERCLOS - Apenas um resultado com 100%;
- Bocejos - Entre 0,0% e 100%;
- Orientação da cabeça/olhar - Entre 0,0% e 100%.

Na tabela 4.4 estão resumidos os resultados obtidos em todos os testes, representando assim a eficácia do sistema na globalidade dos testes efetuados. Como se pode verificar nesta tabela, a eficácia de cada métrica variou entre os 31,3% e os 100%. E a eficácia total do sistema foi de 62,7%.

Teste	Métrica	Alarmes corretos	Alarmes incorretos	Alarmes em falta	Eficácia por métrica (%)	Eficácia global (%)
1	AECS	21		1	95,5%	81,3%
	PERCLOS					
	Bocejos	4		4	50,0%	
	Orientação da cabeça/olhar	1		1	50,0%	
2	AECS	17	5	6	60,7%	59,0%
	PERCLOS					
	Bocejos	5	4	1	50,0%	
	Orientação da cabeça/olhar	1			100,0%	
3	AECS	15	2		88,2%	71,4%
	PERCLOS					
	Bocejos			4	0,0%	
	Orientação da cabeça/olhar					
4	AECS	6	14		30,0%	27,3%
	PERCLOS					
	Bocejos		1		0,0%	
	Orientação da cabeça/olhar		1		0,0%	
5	AECS	11	15		42,3%	39,3%
	PERCLOS					
	Bocejos			2	0,0%	
	Orientação da cabeça/olhar					
6	AECS			13	0,0%	12,5%
	PERCLOS					
	Bocejos	2		1	66,7%	
	Orientação da cabeça/olhar					

Tabela 4.2: Resultados individuais dos primeiros seis testes com utilizadores.

Na secção seguinte serão discutidos os resultados aqui apresentados.

4.3 Discussão dos resultados

Conforme foi apresentado na secção anterior, verificou-se que a eficácia do sistema proposto variou muito, não só entre os vários testes efetuados, mas também entre as várias métricas de fadiga implementadas.

Estas variações deveram-se, não só às diferentes condições em que os testes foram efetuados, mas também à forma como cada uma das pessoas monitorizadas efetua alguns dos gestos que o sistema precisa identificar para conseguir detetar padrões de fadiga. Alguns dos fatores que influenciaram negativamente os resultados foram:

- A face da pessoa monitorizada não ser iluminada de forma igual em toda a sua área, como aconteceu nos testes dois e cinco. Este facto levou a que alguns gestos só

Teste	Métrica	Alarmes corretos	Alarmes incorretos	Alarmes em falta	Eficácia por métrica (%)	Eficácia global (%)
7	AECS	3	3	1	42,9%	55,6%
	PERCLOS					
	Bocejos	7	2	2	63,6%	
	Orientação da cabeça/olhar					
8	AECS	27	3		90,0%	81,0%
	PERCLOS	1			100,0%	
	Bocejos	4			100,0%	
	Orientação da cabeça/olhar	2		5	28,6%	
9	AECS			1	0,0%	0,0%
	PERCLOS					
	Bocejos			4	0,0%	
	Orientação da cabeça/olhar			3	0,0%	
10	AECS	21	2	7	70,0%	71,1%
	PERCLOS					
	Bocejos	6	1		85,7%	
	Orientação da cabeça/olhar			1	0,0%	
11	AECS			8	0,0%	31,3%
	PERCLOS					
	Bocejos	5		3	62,5%	
	Orientação da cabeça/olhar					
12	AECS	27	1		96,4%	96,9%
	PERCLOS					
	Bocejos	3			100,0%	
	Orientação da cabeça/olhar	1			100,0%	

Tabela 4.3: Resultados individuais dos últimos seis testes com utilizadores.

Métrica	Alarmes corretos	Alarmes incorretos	Alarmes em falta	Eficácia (%)	Eficácia do sistema (%)
AECS	148	45	37	64,3%	62,7%
PERCLOS	1	0	0	100,0%	
Bocejos	36	8	21	55,4%	
Orientação da cabeça/olhar	5	1	10	31,3%	

Tabela 4.4: Resultados globais dos testes com utilizadores.

fossem detetados quando a face estava orientada numa posição que minimizasse as diferenças de iluminação da mesma;

- Algumas pessoas quando bocejam apenas abrem ligeiramente a boca, como foi o caso das pessoas monitorizadas nos testes três e nove. Este facto torna difícil diferenciar bocejos relativamente ao estado normal da boca;
- Algumas pessoas tendem a colocar a mão à frente da boca enquanto bocejam, como

foi o caso da pessoa monitorizada no teste quatro. Este facto impossibilita a deteção da face, e conseqüentemente da boca e de bocejos, e leva ainda muitas vezes a deteções erradas, o que conduz à captura incorreta de parâmetros, e conseqüentemente ao incorreto funcionamento do sistema;

- Em alguns testes, a face e sobretudo a área dos olhos da pessoa monitorizada estava demasiado escura, como foi o caso nos testes cinco, seis e nove. Este facto conduziu a algumas deteções erradas, e sobretudo à incorreta identificação de gestos. Por exemplo, no caso da área dos olhos estar sombreada, as zonas mais escuras dos mesmos tendem a ser as próprias sombras, e não as sobrancelhas e as pestanas, impossibilitando assim a deteção do estado dos olhos. Também no caso da boca estar demasiado escura, o algoritmo de deteção de contornos utilizado não apresenta uma diferença significativa no número de contornos entre quando a boca está fechada, ou aberta, dificultando assim a deteção de bocejos;
- Algumas pessoas tendem também a manter os olhos semifechados quando estão a fixar alguma coisa, conforme aconteceu nos testes sete e nove. Este facto faz com que seja difícil diferenciar o estado dos olhos;
- Se durante o período de captura de parâmetros, a pessoa monitorizada efetuar alguns gestos que não se irão repetir ao longo da ação que irá realizar, os parâmetros obtidos poderão não permitir detetar os gestos necessários à deteção de alguns padrões de fadiga. Este facto verificou-se no teste quatro;
- A pessoa monitorizada no teste onze utiliza óculos, o que dificulta a deteção dos olhos, apesar da utilização de um filtro específico para o efeito, e dificulta ainda mais a identificação do estado destes, originando a ineficácia das métricas AECS e PERCLOS.

De entre as métricas implementadas, verificou-se que a que apresentou a menor eficácia foi a responsável por medir a orientação da cabeça/olhar. Este facto deveu-se à sensibilidade da mesma relativamente ao posicionamento da câmara em relação à face da pessoa monitorizada, e ainda à quantidade e intensidade dos movimentos que a pessoa exercia ao longo do teste.

Pelo contrário, a métrica PERCLOS obteve uma eficácia de 100%, estando assim de acordo com a opinião dos autores de [37] e [21], que consideram esta, a métrica de padrões de fadiga mais eficiente, extraída através de técnicas de visão computacional.

De notar ainda que, em condições de iluminação adequadas e quando a pessoa monitorizada não possui qualquer adereço facial e desempenha uma ação que exige pouco movimento, conforme ocorreu no teste 12, o sistema apresenta um excelente desempenho, atingindo uma eficácia de 96,9% no teste mencionado.

Na secção seguinte será feito um breve resumo deste capítulo.

4.4 Resumo do capítulo

Neste capítulo foram apresentados e discutidos os resultados obtidos nos testes ao sistema proposto, recorrendo à ajuda de alguns utilizadores.

De um modo geral, os resultados obtidos foram satisfatórios, contudo estes variaram muito de acordo com as condições de luminosidade do ambiente onde o sistema foi executado, bem como com as diferentes formas como as várias pessoas executaram determinados gestos, e ainda com alguns adereços faciais utilizados pelas mesmas pessoas.

No capítulo seguinte são apresentadas as conclusões que se puderam retirar ao longo do desenvolvimento desta dissertação, bem como algumas alterações e melhorias que poderiam ser aplicadas ao sistema proposto de forma a ultrapassar, ou pelo menos minimizar alguns dos problemas aqui expostos.

Capítulo 5

Conclusões

Na secção seguinte são apresentadas as conclusões retiradas ao longo do desenvolvimento desta dissertação, e na secção 5.2 são indicados quais seriam os passos a seguir como forma de melhorar o sistema de deteção automática de fadiga proposto.

5.1 Balanço crítico

Existem diversos estudos e projetos elaborados em torno de sistemas de deteção de fadiga, mas nenhum deles é perfeito. Os sistemas mais precisos são os mais intrusivos, que dependem de medições constantes do ritmo cardíaco, ou das ondas cerebrais do condutor, mas devido ao desconforto que podem causar, estes são também os menos desenvolvidos e propostos. Contudo, alguns projetos como o HARKEN estão a ser desenvolvidos para limitar este desconforto e manter a precisão das medidas. Se esta tecnologia for bem sucedida poderá ser um grande passo nesta área.

Por outro lado, os sistemas que dependem apenas de pistas visuais como a deteção das marcações das faixas de rodagem, da orientação do olhar e de expressões faciais, são os sistemas com maior aceitação, sendo já possível adquirir veículos com alguns destes sistemas instalados. Contudo, estes sistemas não conseguem atingir os mesmos níveis de precisão que os seus congéneres mais intrusivos.

No capítulo 4 foi demonstrado o nível de eficácia do sistema de deteção automática de fadiga implementado ao longo do desenvolvimento desta dissertação. Esse sistema é baseado exclusivamente em técnicas de visão computacional, apresentando uma eficácia global razoável, mas sofrendo dos mesmos problemas que qualquer sistema de visão computacional, ou seja:

- Perde a sua eficácia perante ambientes com fraca luminosidade;
- É sensível a alterações dos padrões das características a detetar, como por exemplo a deteção de olhos, quando cobertos por óculos.

Contudo, em ambientes bem iluminados e monitorizando pessoas sem adereços faciais e cuja função que realizam não exige movimentos muito amplos, o sistema proposto apresentou uma eficácia entre os 81,0% e os 96,9%. De realçar ainda que, ao contrário dos sistemas atualmente utilizados em algumas marcas automóveis, este não se restringe apenas a ser aplicado durante a condução, mas sim em qualquer ambiente de trabalho, desde que seja possível capturar permanentemente a face da pessoa através de uma câmara de vídeo.

De entre as cinco métricas de padrões de fadiga implementadas, a mais eficaz foi a PERCLOS, contudo, esta métrica implica que durante o último minuto a pessoa monitorizada tenha tido os olhos fechados durante pelo menos 12s. Isto pode significar que, por exemplo uma pessoa possa ter os olhos fechados durante 12s seguidos e que só após esses 12s o alarme de fadiga seja ativado. Se esta for a situação durante a condução de um automóvel, ao fim de 12s pode ser tarde demais. Por este motivo, é totalmente recomendável a utilização desta métrica em conjunto com outras, como por exemplo, com a AECS, que foi a segunda mais eficaz nos testes efetuados, e muitos dos respetivos algoritmos são partilhados entre ambas.

Na secção seguinte são indicadas algumas melhorias que poderiam ser aplicadas ao sistema desenvolvido, de forma a melhorar o seu desempenho.

5.2 Trabalho futuro

O sistema desenvolvido não permite continuar a detetar a face da pessoa caso esta saia do alcance da câmara. Este facto podia ser minimizado substituindo a câmara utilizada por outra com a capacidade de mudar automaticamente a sua orientação. Deste modo, sempre que a face da pessoa monitorizada se aproximasse de um dos extremos da área capturada pela câmara, esta poderia reorientar-se de forma a voltar a colocar a face da pessoa no seu centro de visão, acompanhando assim os movimentos da pessoa em questão. Em casos extremos poderiam ainda ser utilizadas mais do que uma câmara colocadas em locais estratégicos, e alternando a fonte de vídeo entre estas, de acordo com a localização da pessoa a monitorizar.

Outro aspeto que poderia ser melhorado seria a sensibilidade do sistema relativamente à falta de iluminação da face da pessoa monitorizada. Isso poderia ser feito recorrendo a câmaras equipadas com LEDs de infravermelhos que permitem obter imagens a branco e preto mais nítidas, mesmo em condições de muito baixa luminosidade.

De forma a melhorar o desempenho global do sistema, poderia ser utilizado o modo de

pesos por métrica consoante o tipo de ação a ser efetuada pela pessoa monitorizada e ainda de acordo com as condições de luminosidade do respetivo local. Os respetivos pesos seriam atribuídos automaticamente pelo sistema, após o período inicial de captura de parâmetros, e de acordo com os mesmos.

A tecnologia está em constante evolução e a visão por computador é um dos campos que mais tem progredido na atualidade. Esse facto, com a ajuda de algumas orientações já existentes para construir este tipo de sistemas, vai com certeza permitir o desenvolvimento de sistemas de deteção de fadiga muito mais precisos e impercetíveis para os condutores, podendo eventualmente salvar-lhes a vida.

Bibliografía

- [1] Shabnam Abtahi. Driver drowsiness monitoring based on yawning detection. Master's thesis, University of Ottawa, Ottawa, Canada, 2012.
- [2] Federal Motor Carrier Safety Administration. *An evaluation of emerging driver fatigue detection measures and technologies*. U.S. Department of Transportation Federal Motor Carrier Safety Administration Office of Analysis Research and Technology, Washington, D.C., 2009.
- [3] Luis Miguel Bergasa, Jesús Nuevo, Miguel Ángel Sotelo, Rafael Barea, and María Elena López Guillén. Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems*, 7(1):63–77, 2006.
- [4] Massimo Bertozzi and Alberto Broggi. Gold: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE transactions on image processing*, 7:62–81, 1998.
- [5] S. Boveris. Driver fatigue monitoring technologies and future ideas. In *in Proc. AWAKE Road Safety Workshop*, Balocco, Italy, 2004.
- [6] Eun-Ha. Choi, National Center for Statistics, and Analysis (U.S.). *Sampling design used in the National Motor Vehicle Crash Causation Survey [electronic resource] / Eun-Ha Choi ... [et al.]*. NHTSA's National Center for Statistics and Analysis, Washington, D.C., 2008.
- [7] R. C. Coetzer and G. P. Hancke. Driver fatigue detection : A survey. In *AFRICON Conference*, 2009.
- [8] European Transport Safety Council. *The Role of Driver Fatigue in Commercial Road Transport Crashes*. European Transport Safety Council, 2001.
- [9] T. Danisman, I.M. Bilasco, C. Djeraba, and N. Ihaddadene. Drowsy driver detection system using eye blink patterns. In *Machine and Web Intelligence (ICMWI), 2010 International Conference on*, pages 230–233, 2010.

- [10] E.R. Davies. *Computer and Machine Vision: Theory, Algorithms, Practicalities*. Academic Press. Elsevier, 2012.
- [11] Lifang Deng, Xingliang Xiong, Jin Zhou, Ping Gan, and Shixiong Deng. Fatigue detection based on infrared video pupillography. In *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, pages 1–4, 2010.
- [12] Mandalapu Sarada Devi and Preeti R. Bajaj. Driver fatigue detection based on eye tracking. In *Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology, ICETET '08*, pages 649–652, Washington, DC, USA, 2008. IEEE Computer Society.
- [13] D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Always learning. Pearson Education, Limited, 2011.
- [14] Anneke Heitmann, Rainer Guttkuhn, Acacia Aguirre, Udo Trutschel, and Martin Moore-Ede. Technologies for the monitoring and prevention of driver fatigue. In *PROCEEDINGS of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Lexington, MA 02421, USA, 2001.
- [15] Shifeng Hu, Zuhua Fang, Jie Tang, Hongbing Xu, and Ying Sun. Research of driver eye features detection algorithm based on opencv. In *Intelligent Systems (GCIS), 2010 Second WRI Global Congress on*, volume 3, pages 348–351, dec. 2010.
- [16] Budi Thomas Jap, Sara Lal, Peter Fischer, and Evangelos Bekiaris. Using eeg spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, 36(2, Part 1):2352 – 2359, 2009.
- [17] Muhammad Fahad Khan and Farhan Aadil. Efficient car alarming system for fatigue detection during driving. *International Journal of Innovation, Management and Technology*, 3(4), August 2012.
- [18] Nikolaos Kourkoumelis and Margaret Tzaphlidou. Eye safety related to near infrared radiation exposure to biometric devices. *ScientificWorldJournal*, 11:520–8, 2011.
- [19] Jarek Krajewski, Martin Golz, Sebastian Schnieder, Thomas Schnupp, Christian Heinze, and David Sommer. Detecting fatigue from steering behaviour applying continuous wavelet transform. In *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research*, MB '10, pages 24:1–24:4, New York, NY, USA, 2010. ACM.
- [20] Robert Laganière. *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Publishing, may 2011.
- [21] Liying Lang and Haoxiang Qi. The study of driver fatigue monitor algorithm combined perclos and aecs. In *CSSE (1)*, pages 349–352. IEEE Computer Society, 2008.

- [22] Xing Li, Guang Han, Guangteng Ma, and Yanshan Chen. A new method for detecting fatigue driving with camera based on opencv. In *Wireless Communications and Signal Processing (WCSP), 2011 International Conference on*, pages 1–5, nov. 2011.
- [23] D.B.B. Liang and Lim Kok Houi. Non-intrusive eye gaze direction tracking using color segmentation and hough transform. In *Communications and Information Technologies, 2007. ISCIT '07. International Symposium on*, pages 602–607, 2007.
- [24] Weiwei Liu, Haixin Sun, and Weijie Shen. Driver fatigue detection through pupil detection and yawing analysis. In *Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on*, pages 404–407, 2010.
- [25] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. *ECCV*, 2008. <http://www.milbo.users.sonic.net/stasm>.
- [26] Stephen Milborrow. *Stasm 4*, May 2013. Document Version 1 for Stasm Version 4.0.0.
- [27] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.
- [28] Dean Pomerleau. Ralph: Rapidly adapting lateral position handler. In *IEEE Symposium on Intelligent Vehicles*, pages 506 – 511, September 1995.
- [29] Mahdi Rezaei and Reinhard Klette. Simultaneous analysis of driver behaviour and road condition for driver distraction detection. *International Journal of Image and Data Fusion*, ahead-of-p:1–20, 2011.
- [30] Gerhard X. Ritter and Joseph N. Wilson. *Handbook of Computer Vision Algorithms in Image Algebra*. CRC Press, Inc., Boca Raton, FL, USA, 2nd edition, 2000.
- [31] Mohamad Hoseyn Sigari. Driver hypo-vigilance detection based on eyelid behavior. *Advances in Pattern Recognition, International Conference on*, 0:426–429, 2009.
- [32] J. Solem. *Programming Computer Vision with Python: Tools and Algorithms for Analyzing Images*. O'Reilly and Associate Series. O'Reilly Media, Incorporated, 2012.
- [33] Paul Viola and Michael Jones. Robust real-time object detection. In *International Journal of Computer Vision*, 2001.
- [34] Qiong Wang, Huan Wang, Chunxia Zhao, and Jingyu Yang. Driver fatigue detection technology in active safety systems. In *International Conference on Remote Sensing, Environment and Transportation Engineering*, 2011.
- [35] Phillip Ian Wilson and Dr. John Fernandez. Facial feature detection using haar classifiers. *Journal of Computing Sciences in Colleges*, pages 127–133, 2006.
- [36] Wu Ye, Shan Yuetian, Xu Yunhe, Wang Shu, and Zhuang Yuchen. The implementation of lane detective based on opencv. In *Intelligent Systems (GCIS), 2010 Second WRI Global Congress on*, volume 3, pages 278–281, dec. 2010.

- [37] Shanshan Zhang, Fuqiang Liu, and Zhipeng Li. An effective driver fatigue monitoring system. In *MVHI*, pages 279–282. IEEE Computer Soceity, 2010.
- [38] Zhiwei Zhu, Qiang Ji, and Peilin Lan. Real time non-intrusive monitoring and prediction of driver fatigue. *IEEE Trans. Veh. Technol*, 53:1052–1068, 2004.