

DISSERTAÇÃO DE MESTRADO

**SISTEMA BASEADO EM VÍDEO PARA DETECÇÃO
DE SONOLÊNCIA EM MOTORISTAS**

Kédson Lopes de Queiroz

Brasília, outubro de 2011

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO

**SISTEMA BASEADO EM VÍDEO PARA DETECÇÃO
DE SONOLÊNCIA EM MOTORISTAS**

Kédson Lopes de Queiroz

*Relatório submetido ao Departamento de Engenharia
Elétrica como requisito parcial para obtenção
do grau de Mestre em Engenharia Elétrica*

Banca Examinadora

Prof. Ícaro dos Santos, Ph.D ENE/UnB
Orientador

Eduardo Jorge Valadares Oliveira, Dr MS
Examinador externo

Prof. Geovany Araújo Borges, Dr ENE/UnB
Examinador interno

Prof. João Yoshiyuki Ishihara, Dr ENE/UnB
Examinador interno

Dedicatória

Dedico este trabalho a Deus, que até aqui tem me guiado, concedendo-me uma ínfima porção de Sua infinita sabedoria, inserindo em minha vida ao tempo certo as pessoas certas nos lugares certos.

Kédson Lopes de Queiroz

Agradecimentos

Agradeço à minha família e a Débora Cristina que sempre estiveram ao meu lado, oferecendo-me suporte. A meu colega Oscar Gaidos pelas contribuições. Ao meu orientador Ícaro que me fez entender que as coisas conseguidas com esforço e luta são mais proveitosas, ensinando-me com paciência a confiar em suas direções. Em especial ao prof. Waslon Lopes que foi o primeiro a plantar a semente no solo fértil dos estudos na minha vida, investindo em mim e tornando-se um grande referencial.

Kédson Lopes de Queiroz

RESUMO

A sonolência em motoristas é uma das principais causas de acidentes de trânsito. O monitoramento de um motorista para detectar a seu nível de fadiga é um problema complexo que envolve elementos fisiológicos e comportamentais. Diferentes abordagens têm sido feitas e, entre elas, a visão computacional tem o potencial de monitoramento da pessoa atrás do volante sem interferir na sua condução. Uma estimativa do estado de alerta do motorista pode ser obtida através da análise de suas expressões faciais e do abrir e fechar dos seus olhos. Este trabalho demonstra o desenvolvimento de um dispositivo não invasivo, que utiliza a localização e o rastreamento do rosto e dos olhos para monitorizar o estado de alerta em motoristas. O sistema desenvolvido neste trabalho funciona com índices de desempenho superiores a 94% na detecção de fadiga.

Palavras-chave: Visão computacional, HaarCascade, detecção do piscar, monitoramento da fadiga em motoristas.

ABSTRACT

Sleepiness in drivers is a major cause of traffic accidents. The monitoring of a driver to detect the level of fatigue is a complex problem that involves physiological and behavioral elements. Different approaches have been made and among them, computer vision has the potential to monitor the person behind the wheel without interfering with your driving. An estimate of the alertness of the driver can be obtained by analyzing facial expressions and open and close your eyes. This work demonstrates the development of a noninvasive device that uses the location and tracking of face and eyes to monitor alertness in drivers. The system developed in this paper works with performance exceeding 94% in the detection of fatigue.

Keywords: Computer vision, Haar-Cascade, blink detection, monitoring of fatigue in drivers.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	CONTEXTUALIZAÇÃO	1
1.2	DEFINIÇÃO DO PROBLEMA	2
1.3	OBJETIVOS DO PROJETO	2
1.3.1	REQUISITOS	2
1.4	APRESENTAÇÃO DA DISSERTAÇÃO	3
2	REVISÃO BIBLIOGRÁFICA	4
2.1	INTRODUÇÃO	4
2.2	ESTUDOS ANTERIORES, TRABALHOS E PESQUISAS RELACIONADAS	4
3	CONFIGURAÇÕES DE HARDWARE E SOFTWARE	6
3.1	INTRODUÇÃO	6
3.1.1	O SISTEMA	6
3.2	RECURSOS DE HARDWARE	7
3.2.1	MICRO-COMPUTADOR	7
3.3	RECURSOS DE SOFTWARE	8
3.3.1	DESCRIÇÃO DOS RECURSOS DE VISÃO COMPUTACIONAL	8
3.3.2	DESIGN E DESCRIÇÃO DOS RECURSOS DE SOFTWARE	8
3.4	MÓDULO DE CAPTURA DE VÍDEO	11
3.5	MÓDULO DE DETECÇÃO DO ROSTO E DOS OLHOS	12
3.6	MÓDULO DE DETECÇÃO DO ESTADO DOS OLHOS	13
4	CAPTURA DE VÍDEO	15
4.1	INTRODUÇÃO	15
4.2	O SISTEMA DE CAPTURA DE VÍDEO	15
4.3	A RADIAÇÃO INFRAVERMELHA	16
4.3.1	ESPECTRO ELETROMAGNÉTICO E LUZ INFRAVERMELHA	16
4.3.2	O SISTEMA DE ILUMINAÇÃO INFRAVERMELHA PROXIMAL	17
4.3.3	FILTROS ÓPTICOS DE IV PROXIMAL	19
4.3.4	ADAPTAÇÃO DA CÂMERA PARA VISÃO NO INFRAVERMELHO PROXIMAL	20
4.4	VISÃO NO INFRAVERMELHO PROXIMAL	20
4.4.1	EFEITO DA LUZ INFRAVERMELHA NO OLHO HUMANO	22

4.4.2	SOFTWARE DE CAPTURA E GRAVAÇÃO DE VÍDEO.....	23
4.5	O SISTEMA DE CONTROLE DA LUMINOSIDADE.....	26
5	DETECÇÃO DO ROSTO E DOS OLHOS.....	29
5.1	INTRODUÇÃO	29
5.2	DETECÇÃO E RASTREAMENTO DO ROSTO	29
5.2.1	MEDIDAS DE DESEMPENHO NA DETECÇÃO E RASTREAMENTO DO ROSTO.....	31
5.3	DETECÇÃO E RASTREAMENTO DO OLHO.....	34
5.4	MÉTODO PARA ESTIMAR O CENTRO DOS OLHOS.....	34
6	DETECÇÃO DO ESTADO DOS OLHOS	38
6.1	INTRODUÇÃO	38
6.2	ESTIMAÇÃO DA FADIGA EM MOTORISTAS	38
6.2.1	MÉTODOS ALTERNATIVOS PARA ESTIMAR SONOLÊNCIA	39
6.3	OS ALGORITMOS USADOS PARA IDENTIFICAR O ESTADO DOS OLHOS.....	40
6.3.1	LIMARIZAÇÃO AUTOMÁTICA BASEADA EM HISTOGRAMA	41
6.3.2	TEMPLATE MATCHING	44
6.3.3	LIMARIZAÇÃO ELEVADA	47
6.3.4	RESULTADOS - ESTATÍSTICA DO ESTADO DE ALERTA EM MOTORISTAS.....	47
7	CONCLUSÕES E TRABALHOS FUTUROS	50
7.1	TRABALHOS FUTUROS.....	50
	REFERÊNCIAS BIBLIOGRÁFICAS	51
	ANEXOS.....	56
I	CÓDIGOS FONTE USANDO OPENCV EM C++	57
I.1	CONFIGURAÇÃO EM C++ DOS ALGORITMOS QUE CONTROLAM A EXPOSIÇÃO E GANHO	57
I.2	CONFIGURAÇÃO EM C++ DOS ALGORITMOS QUE LOCALIZAM E DETERMINAM OS CENTROS DOS OLHOS E A DISTÂNCIA ENTRE OS PONTOS ANTROPOMÉTRICOS	61
I.3	CONFIGURAÇÃO EM C++ DOS ALGORITMOS USADOS PARA DETERMINAR O ESTADO DOS OLHOS.....	62

LISTA DE FIGURAS

3.1	Esquemático mostrando a integração dos sistemas de hardware e software e a localização destes dentro de um veículo.	6
3.2	Fotos mostrando dimensões, sistema de alimentação e localização do FITPC dentro de um veículo de testes.	7
3.3	Módulos do sistema de captura, detecção do rosto, dos olhos e do estado dos olhos. ...	9
3.4	Diagrama geral dos módulos do sistema, mostrando a sequência de funcionamento, iniciando pelo módulo de captura, seguido pela detecção do rosto, dos olhos e do estado dos olhos.	10
3.5	Diagrama ilustra o software para gravar vídeo com diferentes CODECs e taxas de captura foi utilizado nos testes de campo para comparar os sistemas de iluminação e filtros IVP.	11
3.6	Diagrama ilustra a módulo da aplicação desenvolvida para detecção do rosto e dos olhos.	12
3.7	Diagrama ilustra módulo da aplicação desenvolvida para definição do estado dos olhos.	13
4.1	Módulos do sistema de captura e visão com luz infravermelha.	15
4.2	O espectro eletromagnético [1].	16
4.3	A Figura (a) mostra o efeito pupila clara e a Figura (b) pupila escura [2].	17
4.4	Posicionamento da fonte IVP e da câmera para produzir efeitos pupila brilhante e escura [3] [4][5].	18
4.5	Disposição do conjunto de LEDs e câmera com sensor CMOS.	18
4.6	(a) filtros ópticos IVP de 850 e 950 nm utilizados nos testes de campo. (b) gráficos da transmitância de (a).	19
4.7	A Figura (a) mostra o filtro que bloqueia luz IV que será substituído pelo filtro de baixo custo mostrado na Figura (b), que fica localizado à frente do sensor CMOS.	20
4.8	Câmeras de CFTV e placa de vídeo USB usados inicialmente na captura de imagens.	21
4.9	A Figura (a) mostra a câmera PHILIPS SPC1030NC e a Figura (b) a câmera Power-pack VX-88, em seus respectivos invólucros de plástico.	22
4.10	As Figuras de (a) a (f) mostram os filtros IVP de baixo custo e ópticos de 850 e 950 nm e a disposição do sistema de captura e iluminação utilizados durante os testes de campo em dois veículos diferentes.	24
4.11	Problemas relacionados a iluminação e obstruções dentro do ambiente veicular.	25

4.12	Localização do sistema de captura de vídeo para evitar obstruções das mãos e de acessórios do vestuário.....	25
4.13	Posicionamento recomendado para o sistema de captura de vídeo e da iluminação IVP dentro do veículo.	26
4.14	Parte do sistema de controle de luminosidade que identifica excesso e a falta de luminosidade.	27
5.1	Recurso de otimização para rastreamento e detecção da face.	32
5.2	as figuras de (a) a (d) mostram a detecção do rosto em tempo real, de quatro motoristas durante os testes de campo em dois veículos diferentes.	33
5.3	Recurso de otimização para rastreamento e detecção do olho.....	34
5.4	Localização do conjunto de pontos antropométricos faciais.....	35
5.5	As figuras de (a) a (d) mostram os testes de campo realizados em dois veículos diferentes, mostrando a detecção do rosto e olhos em tempo real de quatro motoristas.	37
6.1	Metodologia do PERCLOS.....	39
6.2	Duração de um piscar lento ilustrando a proposta de Davina [6].....	40
6.3	(A) Olho, (B) Histograma do olho, (C) Imagem binária (limiarizada) do globo ocular e pálpebras.	41
6.4	Demonstração dos três estados que o olho pode assumir: aberto, fechando (meio aberto) e fechado.....	41
6.5	Gráfico mostrando o motorista em estado de alerta.	42
6.6	Gráficos mostrando em 6.6(a) motorista entrando em fadiga e 6.6(b) motorista sonolento.....	42
6.7	Validação em laboratório do algoritmo baseado em limiarização automática utilizando iluminação IVP e filtro de baixo custo, estado de alerta.	43
6.8	Validação em laboratório do algoritmo baseado em limiarização automática utilizando iluminação IVP e filtro de baixo custo. Simulação de sonolência.....	43
6.9	Teste do algoritmo baseado em limiarização automática sem iluminação IVP, vídeo coletado no horário crepuscular, vê-se claramente o excesso de luminosidade descaracterizando o rosto do motorista.	43
6.10	Obtenção dos escores, proveniente do aumento e diminuição da similaridade.....	44
6.11	Gráfico mostrando o comportamento dos escores de correlação ao longo do tempo. ...	45
6.12	Teste em horário noturno do algoritmo baseado em Template Matching, identificando o olho fechado, com iluminação IVP, câmera localizada na tapa sol.	46
6.13	Teste do algoritmo baseado em Template Matching com iluminação IVP, horário noturno.	46
6.14	Teste do algoritmo baseado em Template Matching sem iluminação IVP, horário crepuscular.	47
6.15	Teste do algoritmo baseado em limiar alto no laboratório com filtro IVP de baixo custo, estado do olho fechado.....	48
6.16	Teste do algoritmo baseado em limiar alto no laboratório com filtro IVP de baixo custo, sonolência.	48

LISTA DE TABELAS

3.1	Configuração do Hardware	7
4.1	Tabela de transmissão óptica do filme fotográfico. [7].....	19
5.1	Quadro comparativo com desempenho da detecção do rosto.	31
5.2	Proporção de distâncias no rosto humano a partir de (D2, D3, D4 e D5) para D1.....	35
5.3	Resultado do rastreamento da face e dos olhos	36
5.4	Quadro comparativo com desempenho da detecção do rosto e dos olhos.....	36
6.1	Resultado da detecção do estado do olho e da fadiga, utilizando os 3 algoritmos propostos para o sistema.....	49

LISTA DE SÍMBOLOS

Siglas

ACGIH	American Conference of Governmental Industrial Hygienists
AdaBoost	Adaptive Boost
ANNT	Agência Nacional de Transportes Terrestres
BEIs	Biological Exposure Indices
BSD	Berkeley Software Distribution
CDROM	Compact Disk Read Only Memory
CEMSA	Centro de Estudos Multidisciplinares em Sonolência e Acidentes
CIC	Departamento de Ciência da Computação (UnB)
CMOS	Complementary metal-oxide semiconductor
CPU	Central Processing Unit
EOG	Eletrooculograma
FHWA	Federal Highway Administration
FT	Faculdade de Tecnologia
HD	Hard Disk
IPP	Integrated Performance Primitives
IV	Infravermelho
IVP	Infravermelho Proximal
LED	Light Emitting Diode
MB	Mega Bytes
MLL	Machine Learning Library
MS	Ministério da Saúde do Brasil
GHz	Giga Hertz
OpenCV	Open Source Computer Vision Library
PC	Personal Computer
PI	Controlador Proporcional e Integral
PERCLOS	Percentage of eyelids closure over time (Mallis, 1999)
RAM	Random Access Memory
RI	Região de Interesse
RIV	Radiação infravermelha
SDV	Standard deviation – Desvio Padrão
UI	Unidade de Intensidade
UnB	Universidade de Brasília
USB	Universal Serial Bus

Capítulo 1

Introdução

1.1 Contextualização

A análise de dados estatísticos revela que a sonolência ao volante é uma das principais causas de acidentes de trânsito, inclusive com motoristas profissionais. Segundo o Centro de Estudos Multidisciplinares em Sonolência e Acidentes (CEMSA), cerca de 30% dos acidentes de trânsito são causados por adormecimento ao volante e, entre 17 e 19% das mortes no trânsito são causadas por sonolência ao dirigir [8].

Outro dado relevante foi apresentado pela Operação Hypnos-PRF (dados de 2007), quando foi constatado que 87% dos motoristas trabalham de 12 a 20 horas por dia [9].

Segundo a Agência Nacional de Transportes Terrestres (ANNT), o Brasil possui uma frota de 1.030.445 caminhões e ônibus interestaduais, cerca de 1.300.000 caminhoneiros e uma frota de 100.000 ônibus urbanos. No total, são 208 empresas que prestam serviços interestaduais, 1390 empresas de ônibus urbanos e 140.816 empresas e cooperativas no ramo de transportes [10].

A economista e doutora em Engenharia de Transportes, Ieda Lima, apresentou, em maio de 2007, um estudo sobre os acidentes de trânsito no país envolvendo caminhões, durante o primeiro painel do VII Seminário Brasileiro do Transporte Rodoviário de Cargas, realizado na Câmara Federal, em Brasília, o qual mostrou o registro de um acidente com veículo de carga a cada cinco minutos e um custo total, com esses acidentes com caminhões no território nacional, calculado em R\$ 7,7 bilhões¹.

O cenário mundial não é diferente. Nos Estados Unidos, 58% dos acidentes são causados por fadiga do condutor². Os acidentes que envolvem ônibus são mais preocupantes, principalmente pela grande quantidade de passageiros envolvidos. Esses acidentes tendem a resultar em um número elevado de mortes; portanto, fica evidente a necessidade de se buscar um meio de reduzir a quantidade de acidentes que envolvem sonolência ao volante.

¹Fonte: Associação Brasileira de Concessionárias de Rodovias

²fonte: FHWA (*Federal Highway Administration*) EUA

1.2 Definição do problema

Nos últimos anos, muitos pesquisadores têm trabalhado em sistemas de detecção de distração em motoristas, principalmente sobre o aspecto da sonolência e são abordados no Capítulo 2. O monitoramento de um motorista para detectar a desatenção é um problema complexo, pois envolve elementos fisiológicos e comportamentais. Diferentes abordagens têm sido feitas e, entre elas, a visão computacional tem o potencial de monitoramento do indivíduo atrás do volante sem interferir na sua condução.

No entanto, a apesar de ser uma área de pesquisa ativa e da evolução no número de algoritmos propostos, poucos deles foram testados no ambiente veicular sob condições reais de funcionamento, e os que foram testados apresentam desempenho pouco satisfatório sob condições ambientais de luminosidade não controladas. Neste trabalho propõe-se desenvolver o protótipo de um dispositivo capaz de monitorizar o estado de alerta do motorista em ambiente real, sob diversas condições de luminosidade e funcionamento.

1.3 Objetivos do projeto

O objetivo principal do projeto não é propor novos algoritmos, mas adaptar e integrar recursos de visão computacional bem estabelecidos, confiáveis e de baixa complexidade computacional, com a finalidade de desenvolver um dispositivo capaz de detectar sonolência em motoristas, que funcione sob diversas condições de luminosidade e funcionamento, e possa ser integrado a um veículo de grande porte (caminhão ou ônibus). Além disso, o dispositivo deve ser competitivo no mercado, o que implica custo reduzido, robustez e confiabilidade.

1.3.1 Requisitos

Associar custo reduzido (competitividade) e segurança (confiabilidade) impõe uma série de restrições ao sistema de monitoramento de fadiga a ser desenvolvido. A configuração escolhida é a que utilize de componentes mais comuns disponíveis no mercado nacional com suporte e fabricação nos próximos 5 anos. Além disso, o sistema desenvolvido utilizou ferramentas de código livre e foi embarcado em uma plataforma gratuita - distribuição GNU/Linux. De uma forma simplificada, os requisitos mínimos do sistema são os seguintes: operação com apenas uma câmera (visão monocular); desenvolvimento em plataforma gratuita utilizando linguagem de alto desempenho (C, C++); operação não invasiva nem distrativa; operação em tempo real; operação robusta na presença de mudanças de iluminação; funcionamentos diurno e noturno; operação robusta na presença de oclusões parciais, seja pelas mãos do motorista ou pelo de acessórios do vestuário (ex. óculos de sol, bonés); operação segura com mínimo de tolerância a falsos positivos, e nenhum a falsos negativo.

1.4 Apresentação da dissertação

O Capítulo 2 contém uma breve revisão dos principais trabalhos relacionados à detecção e monitoramento da fadiga em motoristas. Esse capítulo tem como objetivo apresentar uma visão global das abordagens utilizadas em outras pesquisas comparando-as com o sistema proposto e mapeando as principais limitações de cada uma delas.

Em seguida, o Capítulo 3 apresenta a descrição geral do sistema de monitoramento proposto. A primeira parte aborda os recursos computacionais relacionados ao hardware, suas características funcionais de processamento. Na segunda parte, são descritos os recursos de software e a arquitetura dos módulos que compõem o sistema.

O Capítulo 4 descreve o sistema de captura de vídeo, o projeto e a construção dos recursos de iluminação infravermelha proximal e os cenários onde são utilizados.

O Capítulo 5 apresenta a modelagem do sistema responsável pela detecção e pelo rastreamento do rosto e dos olhos, explicando como é feita a otimização dos algoritmos, e compara seu desempenho.

O Capítulo 6 descreve os algoritmos adaptados para detecção do estado dos olhos e da sonolência. Finalmente, o Capítulo 7 contém as conclusões e as principais contribuições e limitações deste trabalho. As referências bibliográficas e um anexo contendo códigos fonte encerram este documento.

Capítulo 2

Revisão Bibliográfica

2.1 Introdução

Para aumentar a segurança no trânsito e reduzir o número de acidentes, diversas universidades, centros de pesquisa e empresas do ramo automobilístico (Toyota, Daimler Chrysler, Mitsubishi etc) e governos (União Europeia etc) estão contribuindo para o desenvolvimento de sistemas para monitorizar o comportamento de motoristas usando diferentes tecnologias. Nesse sentido, o uso de informações visuais para obter o estado de sonolência do motorista e entender seu comportamento é um campo de pesquisa ativa.

Essa é uma tarefa difícil, porque há muitos fatores envolvidos como por exemplo, mudanças nas condições de iluminação e uma variedade de possíveis posturas e expressões faciais. Considerando a iluminação, as principais vertentes de pesquisa podem ser divididas em duas partes: a primeira, fornece detalhes sobre os sistemas que funcionam com luz natural; enquanto a segunda, trata de sistemas que operam com a ajuda de sistemas de iluminação baseados em infravermelho proximal (IVP). As principais vantagens e desvantagens desses sistemas serão apresentadas na seção que segue.

2.2 Estudos anteriores, trabalhos e pesquisas relacionadas

Na última década, diversas técnicas têm sido utilizadas para desenvolver sistemas de monitoramento para detectar sonolência em motoristas (SMSM). Elas podem ser divididas em três categorias principais. A primeira categoria inclui os sistemas baseados em sinais biomédicos, tais como, as atividades cerebral, muscular e cardiovascular [11], [12] e [13]. Normalmente, esses sistemas utilizam eletrodos ligados ao corpo do motorista, o que, muitas vezes causa desconforto. Segundo [14], a maioria deles ainda está longe de ser efetivamente introduzida no mercado.

A segunda categoria inclui os sistemas baseados no desempenho do motorista, avaliam variações na posição lateral do veículo, velocidade, ângulo do volante e outros sinais provenientes de sensores ligados ao veículo [15], [16] e [17]. A vantagem dessa abordagem é que a aquisição do sinal é mais simples. Por essa razão, tais sistemas já estão disponíveis comercialmente [18], [19], [20],

mas, há poucos detalhes disponíveis sobre os mecanismos ou parâmetros utilizados nesses sistemas. Estes sistemas estão sujeitos a diversas limitações, tais como: tipo do veículo, a experiência do motorista, características geométricas, condições da estrada e assim por diante.

A limitação mais severa que deve ser considerada é que esses sistemas utilizam rotinas que demandam uma quantidade de tempo considerável para analisar o usuário e o seu comportamento. E portanto, não conseguem detectar “cochilos” que ocorrem quando um motorista sonolento adormece por alguns segundos em um segmento de estrada muito reta, sem alterar os sinais comportamentais de condução.

A terceira categoria é a que mais se assemelha ao sistema proposto neste trabalho. Inclui os sistemas baseados na análise visual do motorista, utilizando técnicas de processamento de imagem. A visão computacional é uma técnica naturalmente não intrusiva para monitoramento da sonolência do motorista: funciona a partir de imagens obtidas por câmeras localizadas na frente do usuário. Essa abordagem é eficaz devido ao fato que a ocorrência de sonolência é refletida através da aparência do rosto, da posição da cabeça e atividade dos olhos do motorista. Diferentes tipos de câmeras e algoritmos de análise têm sido relatados na literatura para essa abordagem: sistemas baseados em câmera no espectro visível [21], [22], sistemas baseados em câmeras adaptadas para infravermelho proximal [2], [3], [23], [24] e sistemas baseados em visão estéreo (2 câmeras) [25] [26].

Alguns deles são produtos comerciais como: *SmartEye* [25] *Seeing Machines DSS* [26], *Smart Eye Pro* [25] e *Seeing Machines Face API* [26]. No entanto, esses produtos comerciais ainda são limitados a ambientes bem controlados e alguns deles requerem inicialização manual e processos de calibração demorados e rígidos e não fazem parte da evolução dos fabricantes de automóveis, pois sua confiabilidade não é suficiente para as empresas fabricantes de veículos assumirem a responsabilidade de eventual acidente.

Ao instalar o sistema, os proprietários assumem a responsabilidade. Portanto, é uma demonstração que há ainda um longo caminho a fim de obter um produto robusto e comercialmente viável nessa categoria.

O sistema de monitoramento proposto neste trabalho não requer inicialização manual, nem é limitado a ambientes controlados, uma característica atraente que o diferencia dos demais sistemas de monitoramento de fadiga incorporados no mercado e na literatura. Além disso, a visão estéreo utilizada em outros sistemas adiciona calibração adicional, eleva o custo de produção e introduz limitações de hardware. Ademais, o objetivo do sistema proposto não é suplantiar sistemas de visão estéreo, os quais têm algumas vantagens, e sim, desenvolver um produto comercial com câmera única de baixo custo de produção.

Capítulo 3

Configurações de hardware e software

3.1 Introdução

Este capítulo aborda as especificações do hardware utilizado para embarcar o sistema, descreve os requisitos de software e as configurações consideradas em sua modelagem e mostra, em detalhes, os módulos para detecção e rastreamento do rosto e dos olhos e a monitorização do estado dos olhos.

3.1.1 O sistema

O sistema consiste em 3 partes, veja a Figura 3.1. A unidade de processamento analisa o fluxo de imagens fornecidas pela câmera IV para extrair informações sobre a pálpebra motorista e padrões de piscar e características face. O ponto de conexão é o lugar onde o motorista se posiciona a frente do volante a uma distancia pre-determinada (máximo 1,2 m) da camera e iluminação IV. O indicador informa o status do sistema e alto-falante fornece avisos sonoros para o motorista.

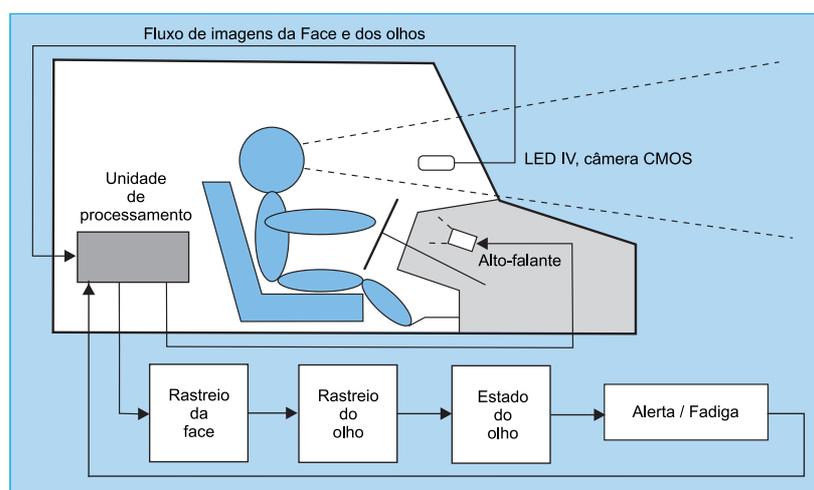


Figura 3.1: Esquemático mostrando a integração dos sistemas de hardware e software e a localização destes dentro de um veículo.

3.2 Recursos de hardware

3.2.1 Micro-Computador

Para embarcar o sistema, a plataforma computacional escolhida foi o FITPC2 [27], que é um micro-computador pequeno e compacto (Dimensões 104x96 mm, e altura 22.9 mm), e por isso pode ser inserido no ambiente veicular facilmente como visto na Figura 3.2, inserido no painel do carro. Sua placa mãe é baseada no processador Intel Atom Z530 (baixo custo, baixo consumo e bom desempenho) e *Chipset* US15W. Funciona com todos sistemas operacionais livres e comerciais. Tem características de baixo consumo de potência 9,6 W e fonte de 12 V. Assim, pode-se aproveitar a voltagem da bateria do carro, utilizando uma fonte regulada de voltagem (o FITPC2 precisa de 12 V constantes). O FITPC2 possui 6 portas USB que são utilizadas como fonte de tensão para circuito de iluminação infravermelho. Para emitir os alarmes, o FITPC2 possui dois canais de áudio de alta definição. A Tabela 3.1 expõe as características do FITPC que mais se destacam.

Tabela 3.1: Configuração do Hardware

Especificação do FITPC2					
CPU	Intel Atom 1.6 Ghz Z510	Storage	HD 160G	RAM	1G DDR2-533
Chipset	Intel US15W	Vídeo	256Mb	Fonte	12V DC
Portas USB	6 portos USB 2.0 Alta velocidade 480 Mbps				
Sist. Operacional	GNU/Linux Ubuntu 10.04				



Figura 3.2: Fotos mostrando dimensões, sistema de alimentação e localização do FITPC dentro de um veículo de testes.

3.3 Recursos de software

3.3.1 Descrição dos recursos de visão computacional

Como o sistema é baseado unicamente na análise visual das expressões faciais do motorista, focando principalmente os olhos, a escolha de um *framework* de visão computacional é indispensável ao desenvolvimento do sistema. A biblioteca de visão computacional escolhida, obviamente por ser *Open Source*, é o OpenCV – *Open Source Computational Vision* [28], [29] e [30] e foi desenvolvida inicialmente pela Intel em 2000 e atualmente mantida pela comunidade acadêmica, que também é responsável por sua atualização.

O OpenCV é liberado sob a licença BSD, seu uso é livre tanto para fins acadêmicos quanto comerciais, desde que, juntamente com os binários distribuídos comercialmente, conste uma aceitação dos termos propostos no contrato de licença liberada pela Intel [31].

A biblioteca é multi-plataforma e escrita em C e C++. Além de ser projetado para promover eficiência computacional, focando, principalmente, aplicações de tempo real, o OpenCV é escrito em C otimizado, de forma a aproveitar ao máximo, em termos de operações computacionais, os processadores multicore, principalmente os de arquitetura Intel, mas esse recurso só é possível se forem utilizadas bibliotecas adicionais da Intel, denominadas IPP – *Integrated Performance Primitives*, que são rotinas de baixo nível, otimizadas para diversas áreas de algoritmos, que vão desde redes até os sistemas inteligentes.

A biblioteca OpenCV contém mais de 300 funções, que incluem uma gama de ferramentas robustas de aprendizagem máquina MLL – *Machine Learning Library*, utilizadas, principalmente, em reconhecimento de padrões estatísticos, e sua aplicação abrange diversas áreas, tais como: visão computacional, incluindo a inspeção de produtos, imagens médicas, segurança, interface de usuário, calibração de câmeras, visão estéreo e robótica.

3.3.2 Design e descrição dos recursos de software

O sistema desenvolvido foi modularizado para que seja de fácil adaptação e manutenção, quanto ao surgimento de diversas situações tais como:

- **Alterações de iluminação:** Os veículos são utilizados em estradas abertas e rodovias e, por isso, estão sujeitos a mudanças na iluminação que não podem ser controladas.
- **Comportamento do condutor:** O condutor em geral, além da condução, realiza outras atividades (manipulação do rádio etc) que interferem no seu desempenho ao volante.
- **Contingência:** As imagens obtidas do motorista estão corrompidas por fatores diversos tais como óculos, lenços, chapéus, barbas, e assim por diante. A isso se acrescenta que o veículo está em constante movimento, sujeito a vibrações que geram distorções nas imagens.

Todas essas situações foram analisadas e levadas em consideração no desenvolvimento e implementação do sistema; portanto, este é dividido em várias etapas (módulos). Cada uma garante sua contribuição para a eficiência do sistema como um todo.

A Figura 3.4 mostra, de uma forma simplificada o diagrama geral do trabalho proposto, que consiste dos seguintes módulos:

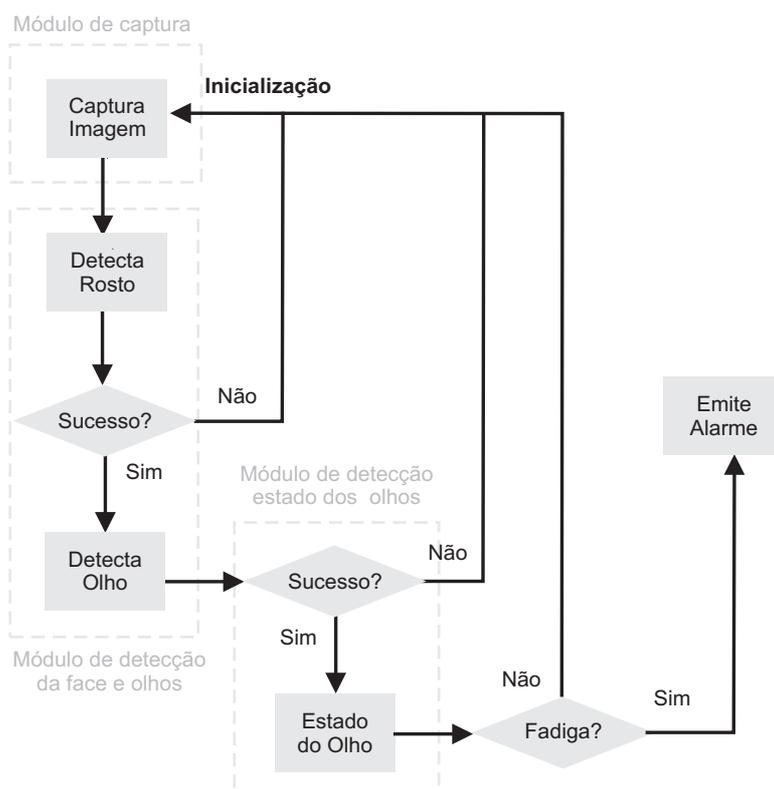


Figura 3.3: Módulos do sistema de captura, detecção do rosto, dos olhos e do estado dos olhos.

- Módulo de captura de vídeo;
- Módulo de detecção do rosto e dos olhos;
- Módulo de detecção do estado dos olhos.

A Figura 3.4 mostra o diagrama geral do conjunto de módulos que fazem parte do sistema, citando a sequência das funções do OpenCV responsáveis pela funcionalidade de cada um deles. Uma descrição detalhada das funções pode ser encontrada na documentação do openCV disponível no site [29]. Nas próximas seções cada módulo será descrito separadamente, em detalhes.

Capítulo 4

Captura de Vídeo

4.1 Introdução

Neste capítulo, é apresentado o módulo de captura de vídeo e, a partir dele, são expostas as soluções e os procedimentos adotados para que o sistema de monitoramento de fadiga proposto funcione de forma satisfatória em condições noturnas de iluminação reduzida ou com excesso de luminosidade em situações de incidência direta de luz solar nos horários crepusculares. Descreve, também, os métodos e instrumentos utilizados para diferenciar a aquisição de imagens, migrando do espectro de luz visível ao espectro infravermelho, utilizando recursos de iluminação e filtro infravermelho.

4.2 O sistema de captura de vídeo

O hardware do sistema de captura é dividido em duas partes interdependentes que interagem entre si, como ilustrado na Figura 4.1:

- Sistema de iluminação infravermelha proximal;
- Sistema de visão no infravermelho proximal.



Figura 4.1: Módulos do sistema de captura e visão com luz infravermelha.

Nas seções seguintes, serão apresentadas, em detalhes, as informações que contribuíram para a sua construção.

4.3 A radiação infravermelha

4.3.1 Espectro eletromagnético e luz infravermelha

O espectro eletromagnético é constituído por um conjunto de ondas que abrangem uma vasta gama de comprimentos, incluindo raios ultravioleta, infravermelhos e luz visível. A Figura 4.2 mostra graficamente esse fenómeno. A radiação infravermelha (RIV) tem um comprimento de onda que vai desde os 720 nm limitada pela cor vermelha na região visível do espectro, até os 15000 nm limitada pelas microondas. Para efeitos práticos, o infravermelho é subdividido em três regiões espectrais nos seguintes intervalos:

- **Infravermelho proximal:** entre 700 e 1500 nm;
- **Infravermelho médio:** 1500 e 15000 nm;
- **Infravermelho distal:** a partir de 15000 nm até 1 mm.

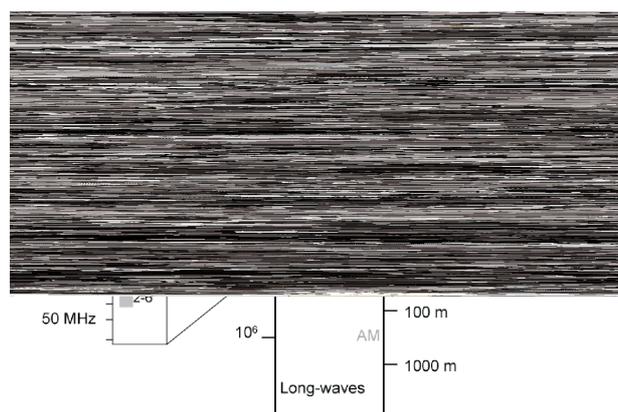


Figura 4.2: O espectro eletromagnético [1].

O sol é a principal fonte natural de radiação IV, que constitui 49% do espectro de emissão solar. E como veremos nos próximos capítulos, a luz solar é a principal fonte de ruído nos sistemas de

aquisição de vídeo em ambientes veiculares. As fontes artificiais de produção de infravermelho (IV) são os emissores não luminosos IV distais (IVDs) (resistências elétricas) e os emissores luminosos de IV proximais (IVPs) (LEDs ou lâmpadas).

4.3.2 O sistema de iluminação infravermelha proximal

A iluminação infravermelha proximal (IVP) é ideal para iluminar o cenário em condições de escuridão. Primeiramente, porque minimiza o impacto das mudanças na luz ambiente. Em segundo lugar, ao contrário do clarão luminoso de uma fonte de luz visível, a iluminação IV é invisível aos olhos e não desvia a atenção do motorista de sua tarefa [32]. Terceiro, produz o efeito de pupila brilhante, que constitui a base para um dos mecanismos utilizados para determinar o estado do olho. Conforme a necessidade de adequação da luminosidade IVP para condições de funcionalidade dentro dos veículos onde ocorreram os testes de campo, foram projetados e construídos dois tipos de conjuntos LEDs de iluminação IVP. O primeiro tipo de iluminador IVP é composto por dois conjuntos de LEDs IVP distribuídos simetricamente ao longo de duas distribuições circulares concêntricas.

Os diâmetros dos anéis de LEDs foram calculados empiricamente a fim de obter uma imagem escura da pupila se o anel externo estiver ligado, e uma imagem brilhante da pupila se o anel interno estiver ligado. A fim de maximizar o efeito de “pupila brilhante”, os LEDs no anel interno são o mais próximo possível das objetivas da câmera. Um exemplo do efeito da pupila clara/escura pode ser visto na Figura 4.3.

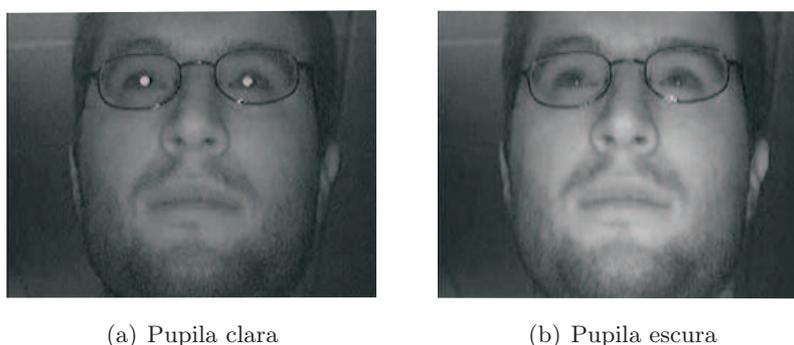


Figura 4.3: A Figura (a) mostra o efeito pupila clara e a Figura (b) pupila escura [2].

O valor do raio do anel externo é um compromisso entre a iluminação resultante para o ambiente, que melhora, uma vez que é aumentado, e o espaço disponível no painel do carro. A posição simétrica dos LEDs nos anéis em torno do eixo óptico da câmera cancelam o excesso de brilho gerado pelos LEDs. A configuração do anel interno obtém o efeito pupila brilhante, pois o centro do anel coincide com o eixo óptico da câmera, atuando como se houvesse apenas um LED localizado no eixo óptico da lente, conforme a Figura 4.4. O anel externo proporciona uma iluminação ambiente que é usada para aumentar o contraste. Apesar dos LEDs produzirem o efeito pupila escura, um brilho pode ser observado em cada pupila. Esta configuração de LEDs proporciona o efeito de pupila clara com e sem óculos, com lentes de contato, e ainda funciona, até certo ponto, com óculos de sol.

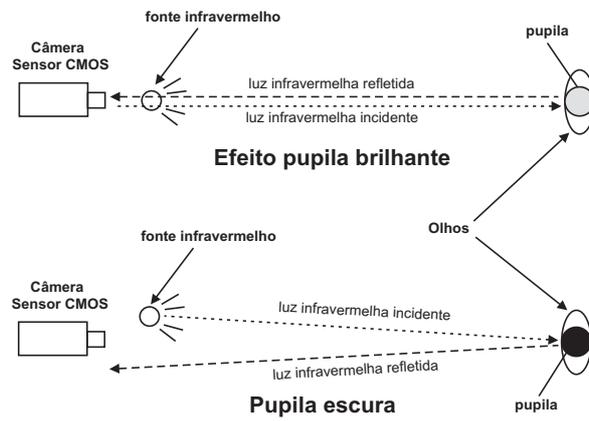


Figura 4.4: Posicionamento da fonte IVP e da câmera para produzir efeitos pupila brilhante e escura [3] [4][5].

O segundo tipo de iluminador IVP é composto por dois conjuntos de LEDs emissores de IVP distribuídos simetricamente ao longo de duas distribuições circulares bilaterais em relação à lente da câmera. Essa configuração favorece apenas a iluminação do ambiente e funciona bem para visualização da pupila escura com óculos de sol.

Em ambas as configurações, foram utilizados LEDs emissores de IVP de 850 e 950 nm, um conector USB usado como fonte de tensão e um potenciômetro de 500 ohms para controlar a potência da luz IVP irradiada. A necessidade do controle de potencia irradiada está diretamente relacionada com a segurança relativa à exposição dos olhos à iluminação IVP e será abordada com maiores detalhes na Subseção 4.4.1.

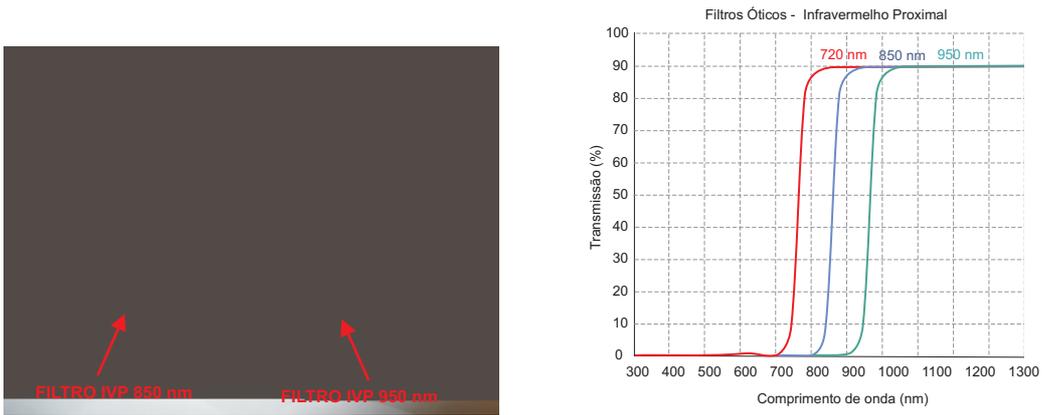
Os orifícios de maior raio no centro da placa servem para fixar a câmera. O conjunto da câmera e a configuração dos LED's podem ser observados na Figura 4.5. É importante salientar que, independentemente do tipo de distribuição, o conjunto de LEDs usados para iluminação IV e os filtros usados para bloquear a luz visível devem ter o mesmo comprimento de onda. A utilização de filtros IVP será abordada em detalhes na Seção 4.3.3.



Figura 4.5: Disposição do conjunto de LEDs e câmera com sensor CMOS.

4.3.3 Filtros ópticos de IV proximal

Os filtros ópticos mais comuns são os filtros de cor, aqueles que só deixam passar a luz de um determinado comprimento de onda. Filtros ópticos de IV são “filtros passa-longas” que transmitem apenas para comprimentos de onda maiores que 695 nm (Figura 4.6(b)). A frequência de corte no comprimento de onda é definida em 50% de transmitância máxima. Para que a câmera funcione de forma satisfatória, capturando a iluminação IVP emitida pela fonte luminosa de IVP, nos testes de campo foram utilizados filtros IV, da marca Kenko, adequados para cada comprimento de onda de cada conjunto de LEDs, isto é, com transmitância específica para os comprimentos de onda de 850 e 950 nm respectivamente, seu gráfico e pode ser visto na Figura 4.6(a).



(a) Filtros ópticos Kenko de 850 e 950 nm

(b) Gráfico da transmitância máxima dos filtros

Figura 4.6: (a) filtros ópticos IVP de 850 e 950 nm utilizados nos testes de campo. (b) gráficos da transmitância de (a).

4.3.3.1 Filtro IV de baixo custo

Um filtro IV de baixo custo pode ser feito com um filme fotográfico. Pode-se observar na tabela 4.1 da transmissão óptica do filme:

Tabela 4.1: Tabela de transmissão óptica do filme fotográfico. [7].

Comprimento de Onda (nm)	Luz visível < 700 nm	700	750	800	850	900	950	1000	1050	1100
Transmissão aproximada (%)	Abaixo de 3%	3	10	70	90	88	87	83	80	75

Ao analisar a Tabela 4.1, observa-se que, para o comprimento de onda de 850 nm, tem-se a melhor permissividade, portanto, baseado nessas características, decidiu-se por utilizar definitivamente o arranjo de LEDs e filtros IVP de 850 nm para a iluminação e visualização do ambiente. No entanto, o filme é delicado, pode sofrer facilmente riscos na superfície, também não é resistente a umidade, mas pode ser protegido com vidro ou plástico, e pode ser útil para diminuir os custos do equipamento em detrimento dos filtros ópticos que são dispendiosos.

4.3.4 Adaptação da câmera para visão no infravermelho proximal

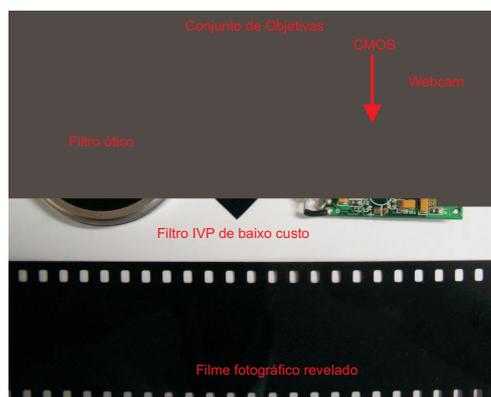
Câmeras utilizadas no módulo de captura possuem sensores CMOS sensíveis ao infravermelho proximal no comprimento de onda de até 1100 nm. Nelas são incorporados filtros de bloqueio IV para rejeitar a luz IV ambiente evitando que o excesso de luminosidade sature o sensor. O filtro de bloqueio IV foi substituído por outro filtro que bloqueia a luz visível e permite passar somente a luz IVP (Figura 4.7). No entanto, esse procedimento acarreta um fator negativo, no caso do ambiente veicular, durante os horários crepusculares; por causa da incidência de luz solar direta, o sensor CMOS tende a ficar saturado. O procedimento de substituição dos filtros, se resume a duas etapas:

- Retirar filtro bloqueador da luz IVP do conjunto de objetivas (transparente);
- Colocar filtro IVP bloqueador de luz visível.

Esse procedimento foi realizado nas duas câmeras utilizadas no projeto e pode ser visto na Figura 4.7.



(a) Filtro bloqueador de luz IVP



(b) Filtro IVP de baixo custo

Figura 4.7: A Figura (a) mostra o filtro que bloqueia luz IV que será substituído pelo filtro de baixo custo mostrado na Figura (b), que fica localizado à frente do sensor CMOS.

4.4 Visão no infravermelho proximal

Para que o sistema de monitoramento em vídeo seja efetivo, a câmera deve recolher uma amostra das imagens rápido o suficiente para capturar o fechamento palpebral durante um piscar de olhos espontâneo de uma pessoa em estado de alerta. Durante um piscar de olhos espontâneo típico, as pálpebras fecham 75%, por cerca de 150 ms, o que corresponde a uma largura de banda de aproximadamente 6 Hz [6]. Para atender ao requisito de amostragem de Nyquist, a câmera deve, amostrar as imagens a uma taxa de, no mínimo, 12 Hz, ou superior, para detectar as pálpebras fechadas durante um piscar de olhos típico espontâneo.

A largura de banda para transferência de dados da câmera para um computador é outro fator

que deve ser considerado para operação em tempo real. Embora o 480 Mbit/s de largura de banda da interface USB 2.0 seja suficiente para a transferência de dados de vídeo em tempo real, a natureza “*host-centric*” da interface USB requer um *overhead* de *hand-shaking* entre o computador host e a webcam, durante o qual alguns quadros de vídeo podem ser perdidos.

Considerando as tecnologias existentes no mercado, a natureza “*peer-to-peer*” da interface *FireWire* que não requer *hand-shaking* seria mais adequada do que a interface USB 2.0 para transferência de dados de vídeo em aplicações em tempo real. No entanto, a taxa de amostragem de 30 quadros por segundo através da interface USB 2.0 foi considerada adequada para operar um sistema de monitoramento e alerta baseado em vídeo em tempo real. Isso porque os movimentos das pálpebras observados durante os períodos de sonolência são relativamente muito mais lentos do que o piscar espontâneo de um indivíduo alerta.

Inicialmente, em caráter experimental, foram utilizados dispositivos para captura de vídeo considerados ótimos/ideais, que são as câmeras e a placa de captura utilizados em circuito fechado de televisão (CFTV) (Figura 4.8), ou seja, com especificações funcionais de desempenho, controle do hardware e qualidade de imagens superiores às do produto final (Webcams com interface USB 2.0).

Webcams são projetadas para propósitos gerais, tais como aplicações de *streaming* de vídeo e vídeo-conferência através da internet, e não são ideais para aplicações de visão computacional. No entanto, considerando a segurança e o desempenho, a qualidade da imagem de ‘baixo custo’ das webcams USB foi considerada suficiente para ser utilizada como mecanismo de captura de vídeo no desenvolvimento do algoritmo de detecção do rosto e dos olhos.

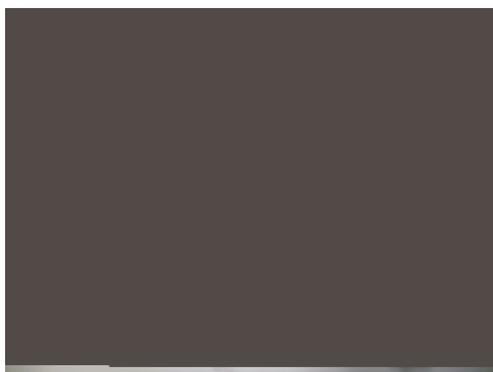


Figura 4.8: Câmeras de CFTV e placa de vídeo USB usados inicialmente na captura de imagens.

Atualmente, duas webcams são usadas para aquisição dos dados por vídeo no projeto. A câmera Powerpack VX-88 utiliza uma lente de vidro e um sensor CMOS que adquire imagens digitais com 24 bits de cor RGB e resolução VGA de 640 x 480 pixels com taxa de captura de até 60 quadros por segundo, através de uma interface USB 2.0. O sensor CMOS é um sensor do tipo 1/4 com *scan* progressivo e filtros de cores primárias RGB é sensível aos espectros visível e infravermelho, no entanto, seu *Firmware* não permite controle manual de ganho e exposição. A lente de vidro, com abertura de 60 graus, da Powerpack VX-88 permite a aquisição de imagens de melhor qualidade do que as lentes de plástico usadas em webcams comuns.

O recurso de balanceamento automático de branco da câmera compensa diferentes níveis de

iluminação do ambiente através do controle da taxa de exposição eletrônica, reduzindo a variação na intensidade da imagem, tornando a câmera robusta sob situações variadas de iluminação. Entretanto, a câmera não permite que o parametro de exposição e o ganho do sensor CMOS possam ser ajustados manualmente. Já a câmera PHILIPS SPC1030NC (como mostra a Figura 4.9(a)) possui as mesmas características funcionais de captura de imagens da Powerpack VX-88, desde que seu conjunto de lentes objetivas sejam substituído pelo da Powerpack VX-88 (que pode ser observada na Figura 4.9(b)), seu diferencial está no *Firmware* que permite controle manual dos parametros de ganho e exposição.



Figura 4.9: A Figura (a) mostra a câmera PHILIPS SPC1030NC e a Figura (b) a câmera Powerpack VX-88, em seus respectivos invólucros de plástico.

4.4.1 Efeito da luz infravermelha no olho humano

De acordo com alguns estudos, emissões contínuas e focalizadas de luz infravermelha podem superaquecer o tecido celular do olho, afetando a retina e, principalmente, a córnea, causando retinite¹ e até mesmo cegueira. Os comprimentos de onda de luz infravermelha que acarretam em maior risco de lesão ocular para os seres humanos são os de menor comprimento de onda.

Experimentos constatam que para um espectro entre 770 e 950 nm, o limite de exposição que produz lesões na retina (no olho de um primata) é de aproximadamente $30 W/cm^2$ por 1 segundo, $23 W/cm^2$ por 10 segundos, $20 W/cm^2$ por 100 segundos e $19 W/cm^2$ para 1000 segundos.

A *American Conference of Governmental Industrial Hygienists* (ACGIH) anunciou no *The Threshold Limit Values* (TLVs) e no *Biological Exposure Indices* (BEIs) que o limite de máximo de radiação emitido pela fonte de luz IVP sem causar danos à córnea nem ao cristalino e no comprimento de onda entre 700 nm e 3000 nm não deve exceder os $10 mW/cm^2$ de irradiância por um período superior a 1000 segundos [33] [34] [35].

Para o dispositivo de iluminação IVP desenvolvido para o projeto que utiliza 20 LEDs, foi estimada uma exposição de $0,95 mW/cm^2$ irradiância IVP, a uma distância de 80 cm dos LEDs emissores. Esse nível de irradiância foi calculado a partir da intensidade radiante, da eficiência dos

¹Retinite é a inflamação da retina no olho.

LEDs e do ângulo de irradiância, essas informações foram obtidas a partir do *datasheet* de dados do LED emissor de IVP.

4.4.2 Software de captura e gravação de vídeo

Com a finalidade de testar e comparar, de forma visual, os diferentes sistemas de iluminação e captura IVP (de 850 e 950 nm) e suas características de iluminação, posicionamento e funcionamento no ambiente veicular, foi desenvolvido um software capaz de:

- Capturar vídeo em tempo real com diferentes resoluções;
- Compactar vídeo com diferentes CODECs² e taxas de captura;
- Gravar e armazenar os vídeos compactados.

O diagrama de desenvolvimento da aplicação é explanado na Figura 3.4. A Figura 4.10, mostra imagens de vídeos coletados utilizando o software de captura em diversas situações reais de condução, em diferentes horários do dia, onde são testados diferentes filtros IVP (ópticos e de baixo custo), com e sem iluminação IVP.

Além dos filtros e iluminação IVP, também foram testadas diversas disposições do sistema de captura de vídeo dentro do veículo (no painel, retrovisor e barra de proteção). Na Figura 4.10(a) pode-se ver uma imagem coletada em horário crepuscular, utilizado um filtro óptico IVP de 850 nm. A Figura 4.10(b) é uma imagem coletada sem usar filtro IVP.

Em ambas situações, não foi utilizada a iluminação IVP, e observa-se que; nestas situações, a luz solar incidente no rosto do motorista, não provocou a perda significativa das características faciais.

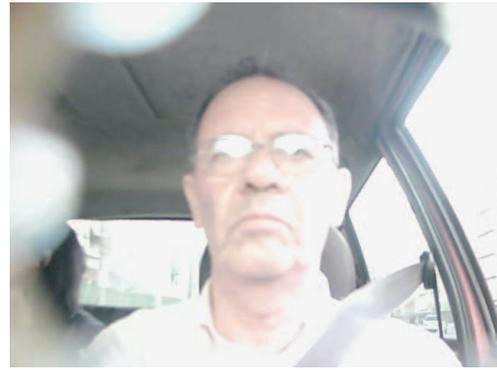
As Figuras 4.10(c) e (d) mostram a captura de vídeo em horário noturno. Em ambas situações foram utilizados filtros e iluminação IVP de 950 nm. No entanto, o que muda é a disposição do sistema de captura. Em (c) no painel atrás do volante, e em (d) no retrovisor.

As Figuras 4.10(e) e (f), mostram a utilização do filtro IVP de baixo custo. No entanto nestes casos, observou-se a perda de características faciais dos motoristas, foi provocada pelo excesso de luminosidade mediante a incidência de luz solar direta, que ocorre no horário crepuscular. Observa-se também, que o óculos escuro usado pela motorista é transparente a iluminação IVP.

²CoDec é o acrônimo de Codificador/Decodificador, dispositivo de hardware ou software que codifica/decodifica sinais.



(a) Captura diurna com luz solar incidente, filtro IVP de 850 nm, disposição de captura no painel atrás do volante.



(b) Captura diurna com luz solar incidente, sem filtro IVP, disposição de captura no painel atrás do volante.



(c) Captura noturna com iluminação e filtro IVP de 950 nm, disposição de captura no painel atrás do volante.



(d) Captura noturna com iluminação e filtro IVP de 950 nm, disposição de captura no retrovisor.



(e) Captura diurna sem iluminação IVP e com filtro IVP de baixo custo, disposição de captura no painel.



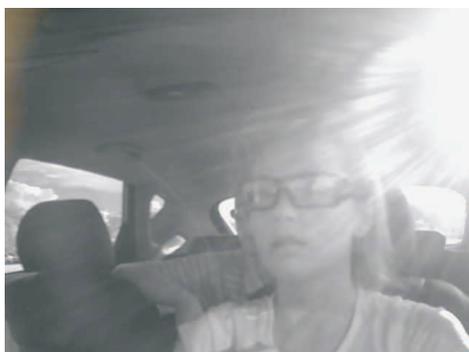
(f) Captura diurna com luz solar incidente, com filtro IVP de baixo custo, disposição de captura painel, com óculos escuros.

Figura 4.10: As Figuras de (a) a (f) mostram os filtros IVP de baixo custo e ópticos de 850 e 950 nm e a disposição do sistema de captura e iluminação utilizados durante os testes de campo em dois veículos diferentes.

A partir da análise dos vídeos coletados nos testes de campo, foram observadas e corrigidas diversas situações que dificultam a utilização de algoritmos para detecção de face e dos olhos, dentre as quais destacaram-se:

- Quantidade e disposição dos LEDs IVP e os tipos de filtros IVP a serem utilizados.
- Locais que produzem trepidação excessiva no sistema de captura do vídeo.
- Posicionamento dentro do veículo para conjunto de iluminação infravermelha e câmera de captura, visando evitar possíveis oclusões (mãos do motorista no volante) conforme mostrado na Figura 4.12. Os ângulos que evitam incidência direta de luz solar na câmera mostrados na Figura 4.11(a), e o uso de bonés/chapéus que venham a ocluir total ou parcialmente a região dos olhos do condutor do veículo (Figura 4.11(b)).

Os resultados da implementação destas medidas podem ser vistos nas Figuras 4.12 e 4.13.



(a) Luz solar incidente no sistema de captura localizado a esquerda do painel, observa-se a perda de detalhes na imagem e principalmente das características faciais, agravadas ainda mais pelo uso dos óculos escuros.



(b) Mostra o condutor a noite em uma situação de obstrução dos olhos por uso de acessórios do vestuário.

Figura 4.11: Problemas relacionados a iluminação e obstruções dentro do ambiente veicular.

Resultados – Soluções aplicadas.

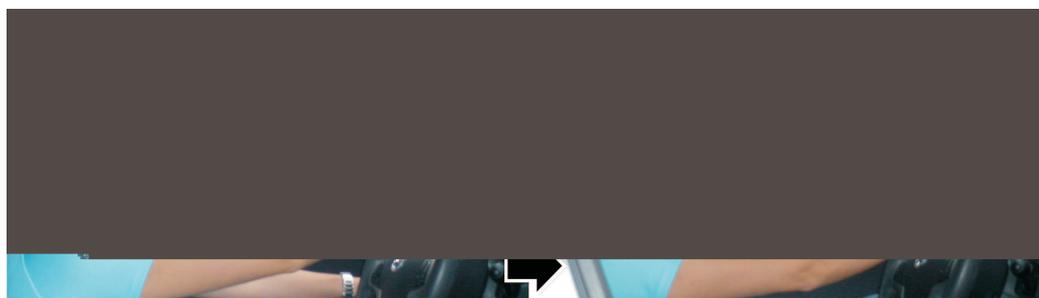


Figura 4.12: Localização do sistema de captura de vídeo para evitar obstruções das mãos e de acessórios do vestuário.

Resultados – Soluções aplicadas.



Figura 4.13: Posicionamento recomendado para o sistema de captura de vídeo e da iluminação IVP dentro do veículo.

4.5 O sistema de controle da luminosidade

Como o módulo de captura é baseado na adaptação/modificação da câmera para capturar a luz emitida pelos LEDs IVP, o procedimento de substituição dos filtros exacerbou a sensibilidade do sensor CMOS, principalmente com relação à variação da luz solar nos horários crepusculares. Os testes de campo demonstraram que essas variações, nas condições de iluminação deveriam ser superadas.

A solução adotada para transpor a dificuldade relativa ao excesso de luminosidade, foi modelada segundo a proposta por Vatani [36]. Na qual, foram utilizados dois *frameworks* associados, primeiro o DDX [37], é utilizado exclusivamente para controlar o *Firmware* de câmeras com interface *Firewire*, destinadas para uso específico em robótica e visão computacional (valor comercial elevado). Associado ao segundo *framework*, o OpenCV, que realiza o tratamento das imagens a fim de identificar as condições de luminosidade.

A solução que foi implementada é mais simples e menos dispendiosa que a sugerida por Vatani [36] tanto em termos de hardware quanto de software. No projeto atual, utilizou-se webcams USB comuns no mercado (PHILIPS SPC1030NC), ao invés de câmeras *Firewire*.

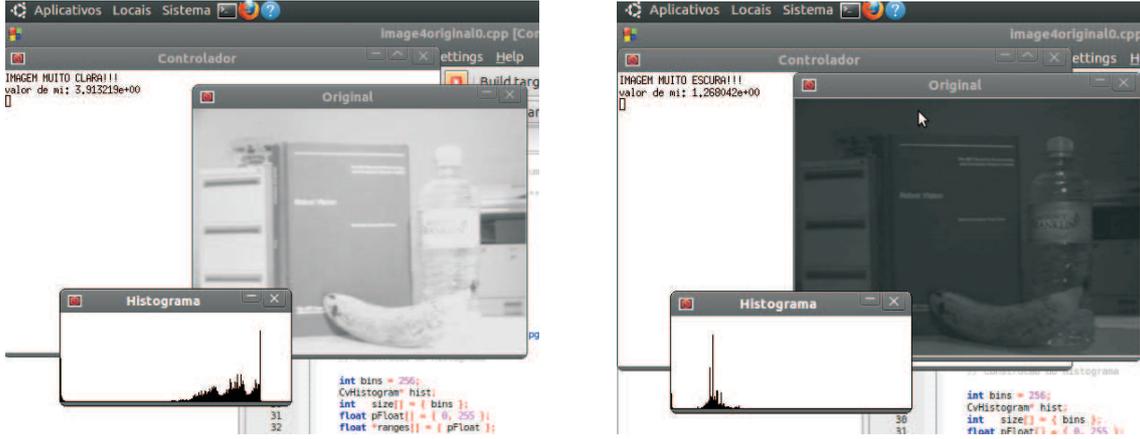
Já no quesito software, ao invés de utilizar 2 *Frameworks*, um para identificar luminosidade e outro para controlar a câmera (DDX associado ao OpenCV), o sistema desenvolvido utiliza apenas um, o V4L2 (*Video for Linux*), que é um *driver* de controle para câmeras nativo do Linux.

Uma vez definidos a câmera e os mecanismos de controle, primeiro foram implementadas metodologias que identificam o excesso e a falta de luminosidade no ambiente.

O histograma de uma imagem pode indicar a natureza das condições de iluminação e a exposição da imagem, e se está subexposta ou superexposta. O histograma foi dividido em 5 regiões, estas regiões representam a faixa de resposta dinâmica do sensor da câmera. As regiões esquerda representam as cores escuras enquanto as regiões direita representam as cores claras.

Uma imagem subexposta será inclinada para a esquerda enquanto uma imagem superexposta será inclinada para a direita no histograma. As Figuras 4.14(a) e 4.14(b) mostram uma imagem

subexposta e outra superexposta e o teste do histograma correspondente.



(a) Metodologia implementada para detectar EXCESSO de luminosidade. (b) Metodologia implementada para detectar FALTA de luminosidade.

Figura 4.14: Parte do sistema de controle de luminosidade que identifica excesso e a falta de luminosidade.

A partir da observação das figuras nota-se que detalhes da imagem desaparecem quando subexpostas ou superexpostas. Por isso, a imagem deve aparecer na região média do histograma. O histograma é apenas uma representação visual. Para usá-lo no controle de exposição automática ele precisa ser interpretado de outra forma. A forma escolhida de interpretá-lo e utilizá-lo, foi através do valor da média da amostra [38]. O valor da média da amostra (MSV) foi calculado a partir do histograma, e através dele se determina o equilíbrio da distribuição de tons na imagem:

$$\mu = \frac{\sum_{i=0}^4 (i+1)x_i}{\sum_{i=0}^4 x_i} \quad (4.1)$$

Em que x_i é a soma dos valores na região i . Nas Figuras 4.14(a) e 4.14(b), o valor de μ representa o MSV para cada histograma. Usando MSV, a imagem é corretamente exposta quando $\mu \approx 2,5$.

Para a câmera PHILIPS SPC1030NC os parâmetros de exposição e ganho, tem os intervalos de 1-5000 e 1-100, respectivamente, e para cada um deles foi implementado um controlador PI (proporcional-integral). O controlador PI foi utilizado por obter um melhor controle do erro estacionário [39]. As constantes do controlador proporcional utilizadas foram:

$$K_{exposicao} = 35 \cdot c + 20 = 0,55\% \quad (4.2)$$

$$K_{ganho} = 15 \cdot c + 10 = 0,25\% \quad (4.3)$$

O parâmetro c descreve quanto em percentagem da imagem é utilizado para o cálculo da exposição. No caso da implementação atual $c = 1$. É desejável manter o parâmetro de ganho mais baixo possível para minimizar o ruído na imagem. O controlador para este parâmetro foi mantido desligado, definido o parâmetro de ganho como 1 até que o parâmetro de exposição atingir máximo. Neste ponto, o controlador de ganho é ativado e o controlador para o parâmetro de exposição é

desligado. As configurações dos algoritmos para calcular o valor de μ e controlador PI, estão dispostas no Anexo I.

Atualmente, o sistema desenvolvido para controle de iluminação está sendo integrado aos demais módulos do sistema. A medida adotada para controlar a luminosidade não foi observada em nenhum dos sistemas que foram estudados, nem faz parte dos que são comercializados atualmente; portanto, é um diferencial que deve ser considerado nos termos de segurança, desempenho e até comercial.

Capítulo 5

Detecção do rosto e dos olhos

5.1 Introdução

O rosto e as características faciais têm atraído a atenção desde tempos imemoriais pela quantidade de informação visual que pode fornecer uma pessoa (humor, idade etc). Atualmente, graças aos avanços tecnológicos, a sua detecção e análise têm vários campos de aplicação. Por exemplo, os sistemas podem incluir a identificação de rostos [40],[41] [42], segurança e vigilância [43] sistemas, teleconferência [44], jogos [45], realidade virtual [46], interação homem-computador [47] [48], biometria [49] e outros. Neste capítulo, serão abordadas metodologias e ferramentas usadas para integração e otimização da detecção do rosto e dos olhos a fim de determinar o estado de alerta em motoristas.

5.2 Detecção e rastreamento do rosto

O algoritmo utilizado para detecção do rosto e dos olhos nos motoristas foi proposto inicialmente por Paul Viola [50] e melhorado por Rainer Lienhart [51]. Ele tem se destacado dos demais métodos pela robustez concernente a oclusões, rapidez e acurácia na detecção do rosto em tempo real. A biblioteca OpenCV já possui esse método implementado, podendo ser utilizado por meio da função `cvHaarDetectObjects`, essa função possui 5 parâmetros de entrada e como resultado de saída retorna uma sequência de retângulos, e cada um desses retângulos representa um rosto. Como entrada da função, temos:

- **image:** A imagem onde será feita a procura.
- **cascade:** O classificador do tipo `HaarCascade` (na forma de arquivo `XML`) lido pelo OpenCV.
- **storage:** Espaço alocado na memória para armazenar a sequência dos resultados da detecção do objeto(resultados em forma de retângulos).
- **scale factor:** O fator pelo qual a janela de procura da função será aumentada para percorrer novamente a imagem.

- **min neighbors:** Faz o agrupamento de retângulos (da região da face) vizinhos, sendo esse parâmetro utilizado para mesclar retângulos similares.
- **flags:** Faz um processamento específico disponibilizado pela biblioteca OpenCV.
- **min size:** Valor em pixels da menor janela de procura, largura e altura da janela. Como saída, temos os dados do vértice superior direito do retângulo, largura e altura.

Essa função pode ser utilizada para vários tipos de objetos, desde que o classificador seja treinado para tal. Nas próximas seções, será demonstrado, em detalhes, como otimizar esses parâmetros a fim melhorar a acurácia e reduzir o custo computacional durante a detecção e o rastreamento do rosto e dos olhos.

Uma vez que a aplicação é desenvolvida para monitoramento da fadiga em ambiente veicular, assumi-se que há apenas uma pessoa no vídeo; portanto, o sistema detecta apenas uma face. Para isso, o motorista deve estar a uma distância de, no máximo, 120 cm da câmera. Como a área da face será utilizada posteriormente para detecção e rastreamento dos olhos, selecionou-se um retângulo igual ou maior que a face real do motorista. Como parâmetros de entrada da função `cvHaarDetectObjects`, utilizou-se a seguinte configuração:

```
scale fator = 1.3
min neighbors = 4
min size = (40,40)
flag = CV_HAAR_FIND_BIGGEST_OBJECT
```

Configuração da função `cvHaarDetectObjects` do OpenCV em C++:

```
CvSeq* rosto = cvHaarDetectObjects( image, cascade,storage, 1.3, 4, CV_HAAR_FIND_BIGGEST_OBJECT, (40,40));
```

O `scalefator` indica quanto o retângulo pode ser maior que o rosto; `minsize` limita o tamanho mínimo para a área de busca do rosto em 40X40 pixels e a `flag` faz com que a função só retorne o maior retângulo encontrado. Assim, o método de escolha do rosto é realizado internamente pelo OpenCV. Normalmente, as detecções isoladas são falsas detecções, por isso faz sentido descartá-las. Também faz sentido, de alguma forma, mesclar as detecções múltiplas para cada região em uma única detecção.

O OpenCV realiza essas duas tarefas internamente antes de retornar a sua lista de rostos detectados. A etapa de fusão dos retângulos analisa as áreas da imagem que contêm uma grande quantidade de sobreposição de retângulos, em seguida, substitui os retângulos agrupados por um retângulo médio. O parâmetro `minneighbors` é um compromisso de escolha entre desempenho e acurácia do algoritmo, no OpenCV utilizando C++, o padrão para esse parâmetro é de três, o que significa fundir grupos de três ou mais e descartar os grupos com menos retângulos. No algoritmo proposto, o melhor valor a ser utilizado foi determinado de forma heurística, considerando-se acurácia *versus* desempenho, chegando ao valor de 4 retângulos.

Os dois classificadores utilizados para a detecção dos motoristas pertencem aos arquivos XML

que acompanham o OpenCV. Na primeira detecção, é utilizado um classificador treinado para detectar rostos na posição frontal (`haarcascade_frontalface_alt.xml`). Caso este não retorne nenhum rosto, é utilizado outro classificador para rostos em perfil (`haarcascade_polifileface.xml`), treinado com imagens diferentes do primeiro classificador. Essa medida melhora a acurácia na detecção da face do motorista, mas sob a penalidade de prejuízo ao desempenho, aumentando o custo computacional em aproximadamente 20 ms.

5.2.1 Medidas de desempenho na detecção e rastreamento do rosto

Com o intuito de mensurar o desempenho da etapa do algoritmo que realiza a detecção do rosto, foram realizados testes de campo com quatro motoristas com diferentes características físicas, em dois veículos diferentes, nos períodos diurnos e noturnos e em horários crepusculares. Os resultados podem ser observados na Figura 5.2 de (a) a (d). O algoritmo utilizado para rastrear o rosto quando aplicado em vídeos com resolução de 640X480 pixels e taxa de captura de 30 quadros por segundo o custo computacional foi de aproximadamente 80 ms para rastrear o rosto. Esse é um resultado que precisa ser otimizado/reduzido, considerando que a solução deverá ser embarcada em uma plataforma com recursos computacionais limitados (Atom de 1.6 GHz com 1G de RAM).

Para fins de comparação, o módulo de detecção do rosto foi testado em um computador adicional, um notebook Acer, processador Intel Dual Core de 1.8 GHz 3G de RAM, utilizando o mesmo sistema operacional GNU/Linux distribuição Ubuntu 10.04 com os seguintes resultados em média, dispostos na Tabela 5.1:

Tabela 5.1: Quadro comparativo com desempenho da detecção do rosto.

FITPC2 Atom Z510 1.6 Ghz – 1G de RAM		Laptop Acer Intel Centrino Dual core 1.7Ghz – 3G de RAM	
Captura	10 ms	Captura	5 ms
Equalização da imagem	20 ms	Equalização da imagem	7 ms
Primeira detecção do rosto	80 ms	Primeira detecção do rosto	40 ms
Total	100 ms	Total	52 ms

Deduziu-se que a maior parte do tempo gasto no processamento é exigida pelo classificador `HaarCascade` (80 ms). A importância relativa ao desempenho da primeira detecção do rosto refere-se à robustez do algoritmo mediante retomada da detecção durante uma perda de visualização do rosto devido a uma situação de oclusão total ou rotação angular. Podem-se otimizar os recursos de busca do rosto a fim reduzir o tempo (custo computacional) para detecção do rosto em um quadro. O método proposto é simples e eficaz na redução da complexidade computacional para detecção do rosto.

Para reduzir o custo computacional, foram feitas algumas melhorias através da implementação das seguintes medidas: reduziu-se a região em busca de toda a imagem, para uma região em torno do rosto detectado anteriormente.

Os dois parâmetros que mais influenciaram na redução do custo computacional do algoritmo de detecção de objeto baseado em `haarCascade`, foram, primeiramente, determinar uma Região

de Interesse (RI) da imagem (denominada RI_ROSTO) que é a área (região) de onde destaca-se o rosto; o outro, é o janela de busca mínima dentro desta região (denominada como JNL_MIN).

Inicialmente o JNL_MIN foi definido como um quadrado muito pequeno e o RI_ROSTO rastreando toda a imagem. O rosto é detectado facilmente porque o motorista não se move rapidamente e o tamanho do rosto não sofre variação drástica.

O RI_ROSTO foi definido como um quadrado com comprimento de lado $L \times 6/5$, concêntrico com rosto detectado anteriormente, e o JNL_MIN como um quadrado com comprimento de lado $L \times 9/10$ (L é o comprimento do lado do rosto detectado anteriormente).

Se nenhum rosto é detectado, o classificador vai percorrer RI_ROSTO e reduzir JNL_MIN até que o rosto seja encontrado ou limite da região de busca seja atingido. Este procedimento pode ser visto em detalhes na Figura 5.1.

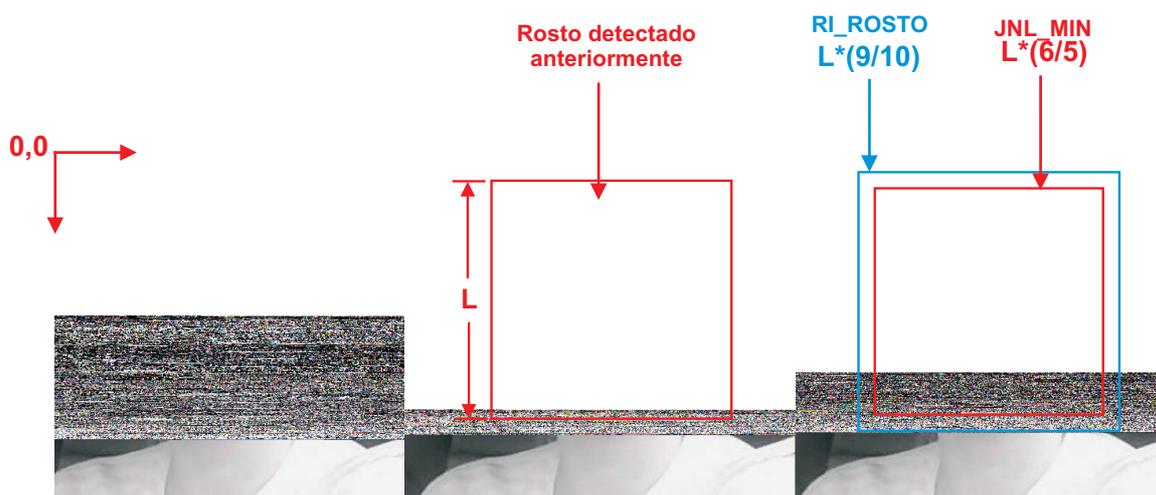
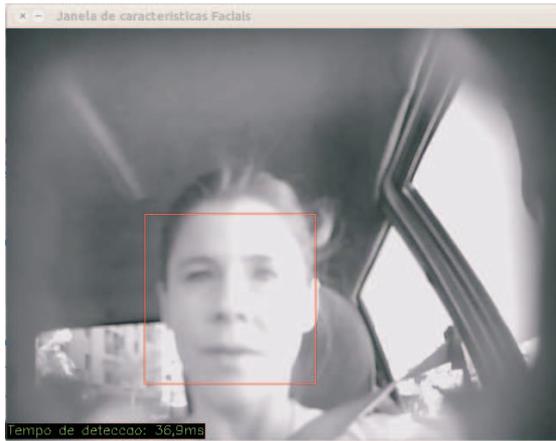


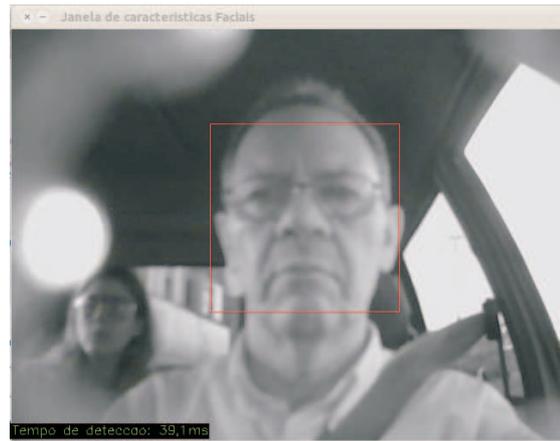
Figura 5.1: Recurso de otimização para rastreamento e detecção da face.

Usando esse método simples, reduziu-se o tempo de detecção do rosto de 80 ms para cerca de 40 ms. A Figura 5.2, mostra os resultados da detecção do rosto dos motoristas em ambiente real, nos horários diurnos (a), (b) e noturnos (c) e (d). A taxa de detecção utilizando as condições mencionadas foi de 98,8%.

Resultados - Detecção da face em motoristas



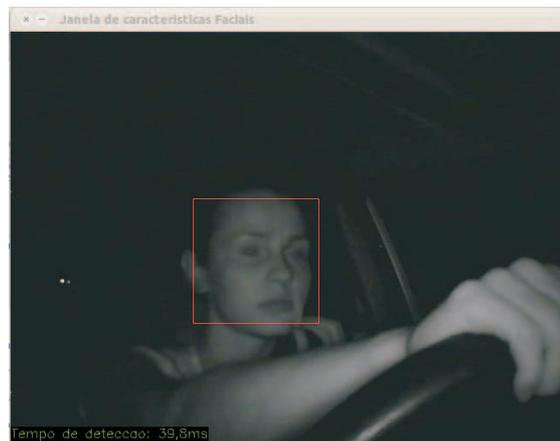
(a) Detecção da face em tempo real no turno vespertino(16 h), sem iluminação IVP, com o sistema de captura localizado na parte frontal painel (atrás do volante).



(b) Detecção do rosto em tempo real com o condutor usando óculos comum, no turno vespertino(17 h 30 min), sem iluminação IVP, com o sistema de captura localizado na parte fronta painel (atrás do volante).



(c) Detecção do rosto em tempo real com o condutor usando óculos comum, em horário noturno (22 h 25 min), o sistema de captura e iluminação IVP 850 nm localizado na parte frontal painel (atrás do volante).



(d) Detecção do rosto em tempo real em horário noturno (22 h 45 min). O sistema de captura e iluminação IVP 950 nm localizado no retrovisor.

Figura 5.2: as figuras de (a) a (d) mostram a detecção do rosto em tempo real, de quatro motoristas durante os testes de campo em dois veículos diferentes.

5.3 Detecção e rastreamento do olho

Diferentemente do rosto, na primeira fase do procedimento de detecção, os olhos não foram procurados em toda a imagem; optou-se por extrair uma RI (região de interesse) dos olhos direito e esquerdo de dentro do retângulo do rosto previamente detectado. Portanto, se o retângulo do rosto é $[0,0, largura, altura]$, define-se o retângulo de RI olhos como $[0, altura/4, largura/2, altura/4]$ este procedimento pode ser visto detalhadamente na Figura 5.3.

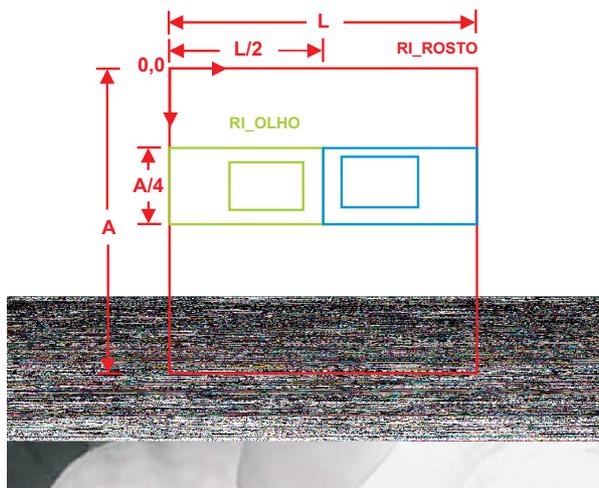


Figura 5.3: Recurso de otimização para rastreamento e detecção do olho.

Depois de extrair o RI olhos, detectam-se os olhos direito e esquerdo (dentro dos retângulos menores, vistos na Figura 5.3) utilizando o mesmo detector **HaarCascade** para detecção de objetos, só que desta vez, usando um classificador específico para cada olho, dentro da região de interesse. Na configuração da função, para detectar objetos, utilizou-se:

```
classificador do olho esquerdo = haarcascade_mcs_lefteye.xml  
classificador do olho direito = haarcascade_mcs_righteye.xml
```

```
scale fator = 1.1  
min neighbors = 3  
min size = (30,20)  
flag = CV_HAAR_FIND_BIGGEST_OBJECT
```

Configuração da função `cvHaarDetectObjects` do OpenCV em C++:

```
CvSeq* olho = cvHaarDetectObjects( image, cascade, storage, 1.1, 3, CV_HAAR_FIND_BIGGEST_OBJECT, (30,20));
```

5.4 Método para estimar o centro dos olhos

Visando obter maior robustez no caso da perda do rastreamento dos olhos e facilitar a detecção de características faciais marcantes para, a partir delas, estimar sua posição, partindo da

localização do centro dos olhos e vice-versa, baseado na pesquisa de Sayeed [52] e [53], um método foi acrescentado ao algoritmo. Esse método utiliza medidas antropométricas¹. No algoritmo implementado, ao invés de usar todos os marcos utilizados por Farkas [54], foi utilizado apenas um pequeno subconjunto de pontos. O modelo proposto foi desenvolvido a partir das constantes proporcionais da Tabela 5.4 utilizando a distância entre o centro dos olhos esquerdo e direito (P1 e P2) como parâmetro de princípio de medição. Segundo Sayeed [52], esse método é capaz de estimar o centro dos olhos em quase 99% dos casos em que é aplicado.

Tabela 5.2: Proporção de distâncias no rosto humano a partir de (D2, D3, D4 e D5) para D1.

Proporção	Descrição	Constante
D_2/D_1	Proporção da distância entre o centro do olho direito e o centro da sobrancelha direita com a distância entre os centros dos olhos	≈ 0.33
D_3/D_1	Proporção da distância entre o centro do olho esquerdo e o centro da sobrancelha esquerda para a distância entre os centros dos olhos	≈ 0.33
D_4/D_1	Proporção da distância entre o ponto médio dos centros dos olhos e a ponta do nariz com a distância entre os olhos	≈ 0.60
D_5/D_1	Proporção da distância entre o ponto médio do centro dos olhos e centro da boca para a distância entre os olhos	≈ 1.10



Figura 5.4: Localização do conjunto de pontos antropométricos faciais.

Uma vez que o rosto está alinhado e ajustado com o modelo antropométrico, o algoritmo processa a detecção dos centros das regiões com características faciais (P3, P4, P6, P7). Esse processo começa com a detecção do ponto médio dos dois centros dos olhos (P5), que serve como ponto de referência para a detecção de P6 e P7 e a distância D1 que é usada como o principal parâmetro para medir a localização de outras regiões com características faciais. Os pontos P3, P4, P7, P6 são estimados através do cálculo das distâncias D2, D3, D4 e D5, respectivamente (Figura 5.4) utilizando as constantes de proporcionalidade propostas pelo modelo antropométrico do rosto listadas na Tabela 5.4.

¹É a ciência biológica que lida com a medição do corpo humano e suas diferentes partes.

No Anexo I, estão dispostas as configurações das funções que localizam e determinam os centros dos olhos e a distância entre os pontos antropométricos, utilizando OpenCV em C++.

A Tabela 5.3 apresenta os resultados do rastreamento ocular em quatro vídeos de teste, mostrados na Figura 5.5.

A *Taxa de Acertos* no rastreamento ocular é definida pela Equação 5.1, que é (*Total de Quadros - Falha no Rastreamento*) dividido pelo *Total de Quadros*. O campo *Total de Quadros* significa o número total de quadros em cada vídeo. *Falha no Rastreamento* é a contagem de quadros onde houve falha durante o rastreamento do olho.

$$\text{Taxa de Acertos} = \frac{\text{Total de Quadros} - \text{Falha no Rastreamento}}{\text{Total de Quadros}} \quad (5.1)$$

Os índices de desempenho e acurácia podem ser vistos nas Tabelas 5.4 e 6.1. A taxa correta de rastreamento ocular é maior do que 98,3% em média.

Tabela 5.3: Resultado do rastreamento da face e dos olhos

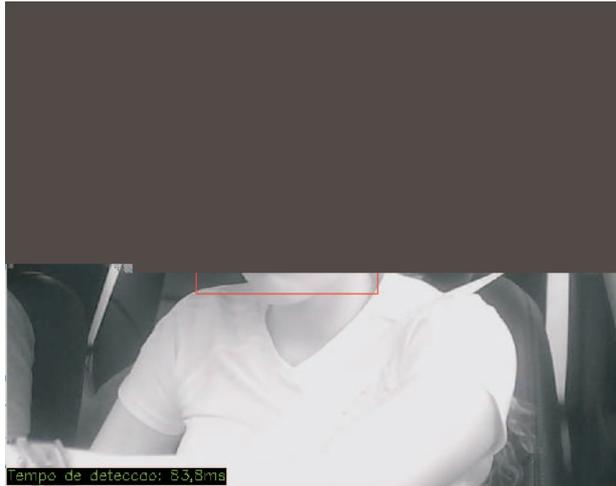
	Vídeo 1 (dia)	Vídeo 2 (noite)	Vídeo 3 (noite)	Vídeo 4 (dia)
Total de Quadros	2833	2845	2822	2855
Falha no Rastreamento	55	44	22	38
Taxa de Acertos	98,1%	98,5%	99,2%	98,7%
Média da Taxa de Acertos			98,6%	

Tabela 5.4: Quadro comparativo com desempenho da detecção do rosto e dos olhos.

FITPC2 Atom Z510 1.6 Ghz – 1G de RAM		Laptop Acer Intel Centrino Dual core 1.7Ghz – 3G de RAM	
Captura	10 ms	Captura	5 ms
Equalização da imagem	20 ms	Equalização da imagem	7 ms
Primeira detecção do rosto	40 ms	Primeira detecção do rosto	20 ms
Primeira detecção do olho	17 ms	Primeira detecção do olho	10 ms
Total	87 ms	Total	42 ms

A Figura 5.5, mostra detecção do rosto e dos olhos, em condições de pouca luminosidade dos horários noturnos Figura 5.5(b) e em condições de iluminação excessiva com luz solar incidente Figura 5.5(d).

Resultados - Detecção do rosto e dos olhos em motoristas



(a) Detecção em tempo real do rosto e dos olhos, no turno vespertino(18 h), sem iluminação IVP, com o sistema de captura localizado na barra lateral de proteção do carro.



(b) Detecção do rosto e dos olhos em tempo real, com o condutor usando óculos comum em horário noturno (22 h 25 min), o sistema de captura e iluminação IVP 850 nm localizado na parte frontal painel (atrás do volante).



(c) Detecção em tempo real do rosto e dos olhos, com o condutor usando óculos comum em horário vespertino (17 h). Sem iluminação IVP.



(d) Detecção do rosto e dos olhos em tempo real no turno vespertino (16 h), sem iluminação IVP, com o sistema de captura localizado na parte frontal painel (atrás do volante).

Figura 5.5: As figuras de (a) a (d) mostram os testes de campo realizados em dois veículos diferentes, mostrando a detecção do rosto e olhos em tempo real de quatro motoristas.

Capítulo 6

Detecção do estado dos olhos

6.1 Introdução

O piscar é uma ação subconsciente, cujo principal objetivo é manter a superfície de nossos olhos úmidas. A taxa do piscar de um indivíduo tem sido associada a diversas áreas de pesquisa. Um exemplo disso é a monitorização do nível de fadiga de um indivíduo [55] [56]. Estresse [57] e decepção [58] também podem ser aferidos através da taxa do piscar de um indivíduo. Estes são apenas alguns exemplos de como uma estimativa da taxa do piscar de um indivíduo pode ser aplicada. O objetivo principal deste capítulo é descrever o módulo do sistema que processa um sinal de vídeo a partir dele, rastreando o rosto e os olhos e, posteriormente, somente os olhos, analisa seu estado (aberto/fechado/piscar) e estima o nível de fadiga.

6.2 Estimação da fadiga em motoristas

A frequência e a duração do piscar do olho são os melhores indicadores da função oculomotora e servem como parâmetros seguros para mensurar o estado de alerta e a capacidade dos motoristas para reagirem aos estímulos do ambiente [59].

A redução ou cessação do piscar espontâneo do olho são considerados um dos primeiros sinais confiáveis de sonolência [60] [61]. A diminuição da amplitude, o aumento da duração e a diminuição na taxa do piscar dos olhos podem ser correlacionados com a deterioração do desempenho durante a execução de tarefas que exigem vigilância devido à sonolência [62] [63]. As mudanças no comportamento do piscar podem ser melhor observadas quando os dados são coletados através de vídeo ou de EOG (eletrooculograma).

O grau de fechamento dos olhos é uma das métricas mais importantes para a detecção de sonolência e dos cochilos. Percentagem do fechamento dos olhos (PERCLOS - *Percentage of Eye Closure*) é uma métrica para estimar o nível de sonolência com base na porcentagem de tempo que os olhos estão fechados em pelo menos 80% durante 1 (um) minuto [64] [65].

O PERCLOS é considerado a medida mais confiável e válida do estado de alerta de motorista pela *Federal Highway Administration* dos EUA [66] [67] e tornou-se uma métrica bem adaptada

para muitos sistemas de monitoramento da sonolência [4] [64] [68] [69]. A Figura 6.1 ilustra o princípio de medição do PERCLOS: t_1 a t_4 são usados para medir o valor do PERCLOS. A fórmula é:

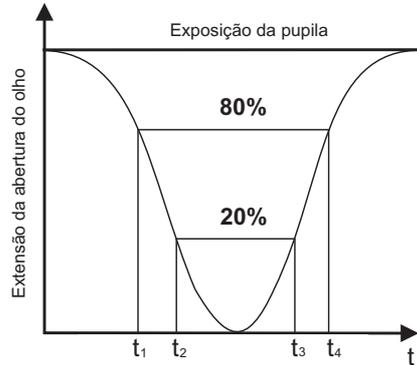


Figura 6.1: Metodologia do PERCLOS.

$$f = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \quad (6.1)$$

O f é o valor do PERCLOS, t_1 e t_2 são os tempos em que os olhos estão fechando a partir do maior grau de abertura, diminuindo para 80%, e de 80% para 20%; t_3 é o tempo de 20% a 20% em que o olho está fechado, e abrindo a partir de t_4 , que é 20% de abertura chegando até 80% onde é computada a quantidade de tempo gasto.

Por exemplo, em uma janela de 60 s (1 minuto), se ocorreram 16 encerramentos lentos da pálpebra, um a cada 3 segundos, então o valor PERCLOS é de 80%, indicando um nível de fadiga crítico. Quando esse método é utilizado para determinar o estado dos olhos, utilizando processamento de imagem, como abordado no Capítulo 5, define-se que o nível de exposição da pupila quando acima de 20% o olho é considerado aberto ou que está fechando.

O fechamento parcial da pálpebra também é um bom indicador de sonolência profunda [62] [70] [71]. Durante o início do sono, há uma tendência geral para os músculos, incluindo o músculo elevador da pálpebra superior, para relaxar. O resultado do relaxamento do músculo da pálpebra superior é um fechamento das pálpebras lento, considerado, normalmente, como “olhos caídos”.

Finalmente, a detecção do fechamento ocular completo por um período prolongado de tempo pode ser usado para identificar cochilos durante a condução de veículos. Esse e outros fatores serão considerados nos métodos alternativos para estimar e determinar a fadiga das seções seguintes.

6.2.1 Métodos alternativos para estimar sonolência

Considerando o custo computacional e a complexidade da implementação do algoritmo tradicional para determinar o PERCLOS em uma plataforma com recursos computacionais limitados, foram considerados dois métodos derivados das sugestões de pesquisa de dois autores. O primeiro método foi proposto por Hongbiao Ma [72], que considera o fechamento da pálpebra (ou sua distância) como sendo o melhor parâmetro para monitorizar o estado de alerta em motoristas. Ele

propõe um método para mensurar a distância da pálpebra, primeiro limiarizando de forma automática da região do olho delineando suas principais características (contornos e região da íris), e, em seguida, mensurando a distância entre os contornos, ou seja, a distância palpebral. Esse processo é descrito nas seções a seguir.

O segundo método é mais simplificado e foi proposto por Horng [73], que define em seu algoritmo, que uma vez identificado o estado do olho, o nível de alerta do motorista também poderá ser calculado, considerando que, se os olhos permanecem fechados por mais de cinco quadros consecutivos, a uma taxa de 30 quadros por segundo, o condutor é considerado sonolento.

Esse mecanismo alternativo de mensurar fadiga em motoristas pode ser considerado válido, pois, segundo a pesquisa de Davina [6], um piscar de olhos lento pode durar entre 200 e 400 ms e, mediante essa informação, considerando uma taxa de captura de 30 quadros por segundo, tem-se que um quadro dura aproximadamente 30 ms e, portanto, a duração de um piscar lento poderá durar aproximadamente entre 8 e 10 quadros. Esse método está ilustrado na Figura 6.2, observa-se nos quadros de 5 a 13 um piscar lento.

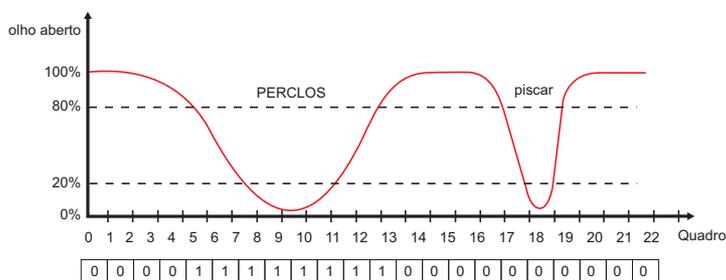


Figura 6.2: Duração de um piscar lento ilustrando a proposta de Davina [6].

Nas próximas seções, será demonstrada a implementação dos algoritmos propostos para detecção do estado dos olhos em conjunto com métodos alternativos para mensurar a fadiga.

6.3 Os algoritmos usados para identificar o estado dos olhos

No Capítulo 5, foi demonstrado o funcionamento dos algoritmos utilizados para detectar e rastrear o rosto e os olhos. O principal recurso desses algoritmos é a utilização de classificadores independentes para o rosto e para cada olho (direito e esquerdo). Após uma etapa de otimização onde foram utilizados recursos de localização geométrica (definição de áreas mais relevantes de busca e pontos antropométricos), os classificados foram aplicados em áreas específicas da imagem, denominadas RI (Regiões Importantes).

Dentro dessas áreas, o recurso mais importante que deverá ser identificado e classificado são os olhos. Uma vez separada essa região, os algoritmos utilizados para identificação do estado do olho atuaram somente nessa região específica, ou seja, na imagem do olho propriamente dita. Para determinar a melhor técnica para identificar o estado do olho (aberto/fechado), foram implementados e testados três algoritmos baseados em três metodologias distintas: limiarização automática baseada em histograma, *Template Matching* e limiarização elevada. Nas próximas seções, estão

dispostos os três algoritmos propostos para determinar o estado dos olhos.

6.3.1 Limiarização automática baseada em histograma

6.3.1.1 Limiar para delinear o olho

Uma vez delimitada a região dos olhos (Figura 6.3 (A)), é necessário usar um algoritmo que determine um limiar automático de forma a delinear o contorno dos olhos, da íris e das pálpebras. Para tanto, é necessário calcular o valor do histograma da imagem do olho (Figura 6.3 (B)), onde H é o índice com valor máximo no histograma (o eixo x é o nível de cinza de 0 a 255, e o eixo y é o número de pixels com cada nível de cinza, o valor máximo será coordenada- x com maior valor de y), sugerido por Hongbiao Ma [72], $H \times 2/3$ foi determinado como bom valor de limiar para delinear o segmento de contorno dos olhos (Figura 6.3 (C)).

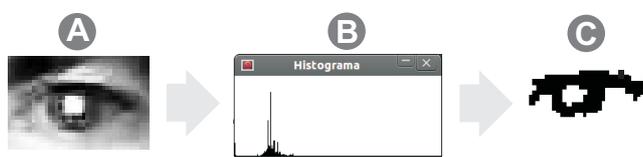


Figura 6.3: (A) Olho, (B) Histograma do olho, (C) Imagem binária (limiarizada) do globo ocular e pálpebras.

O algoritmo para calcular o valor máximo do histograma usado na limiarização da imagem do olho e a configuração de suas funções, utilizando o OpenCV, está disposto no Anexo I.

6.3.1.2 Cálculo da distância da pálpebra

Distância da pálpebra é definida como sendo a distância entre a pálpebra superior e a inferior. Essa distância é relativamente grande quando o motorista está alerta (Figura 6.4 (A)) e tornando-se pequena (Figura 6.4 (B) e (C)) a medida que o motorista vai ficando sonolento ou fadigado.



Figura 6.4: Demonstração dos três estados que o olho pode assumir: aberto, fechando (meio aberto) e fechado.

Uma vez que imagem do olho está limiarizada e sem ruído, uma boa estimativa da distância da pálpebra é computar o número de pixels pretos da imagem do olho. Este valor é armazenado um vetor com 100 posições, indicando a janela onde será analisado o estado do olho. A configuração das funções do OpenCV no algoritmo, para calcular o número de pixels da imagem do olho, está disposta no Anexo I.

6.3.1.3 Detecção de fadiga

Através do método acima, é obtida a distância da pálpebra. Ao analisar a tendência de mudança na distância da pálpebra, pode-se decidir se o motorista está alerta ou sonolento. Supõe-se que o motorista esteja alerta no início (Figura 6.5) e a distância da pálpebra é considerada como valor normal e ideal.

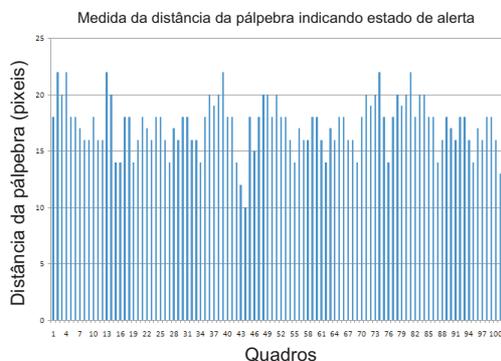
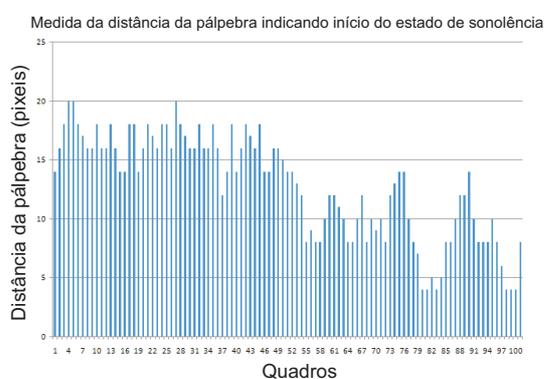
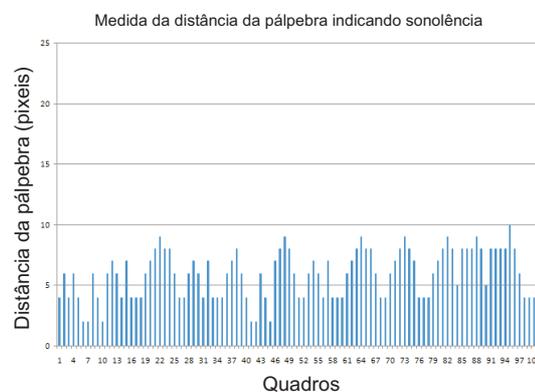


Figura 6.5: Gráfico mostrando o motorista em estado de alerta.

A distância da pálpebra é grande na maior parte do tempo. O encurtamento ocasional dessa distância denota o piscar de olhos. A média móvel da distância da pálpebra é calculada para 100 quadros contínuos. Analisando os dados coletados e os gráficos, conforme a sugestão de [72] quando a média móvel cai abaixo de um limiar de 60% do valor normal o motorista é considerado entrando em estado de fadiga (Figura 6.6(a)). Permanecendo abaixo do limiar de 60%, o motorista é considerado sonolento (Figura 6.6(b)) e então é emitido um aviso sonoro.



(a) Motorista entrando em fadiga



(b) Motorista sonolento

Figura 6.6: Gráficos mostrando em 6.6(a) motorista entrando em fadiga e 6.6(b) motorista sonolento.

A Figura 6.7 e Figura 6.9, mostram a funcionalidade do algoritmo de limiarização automática em ambiente laboratorial. A Figura 6.8, mostra detecção do olho fechado com vídeo coletado em ambiente veicular. A configuração das funções do OpenCV no algoritmo para determinar a fadiga baseado na distância palpebral pode ser vista no Anexo I.



Figura 6.7: Validação em laboratório do algoritmo baseado em limiarização automática utilizando iluminação IVP e filtro de baixo custo, estado de alerta.



Figura 6.8: Validação em laboratório do algoritmo baseado em limiarização automática utilizando iluminação IVP e filtro de baixo custo. Simulação de sonolência.

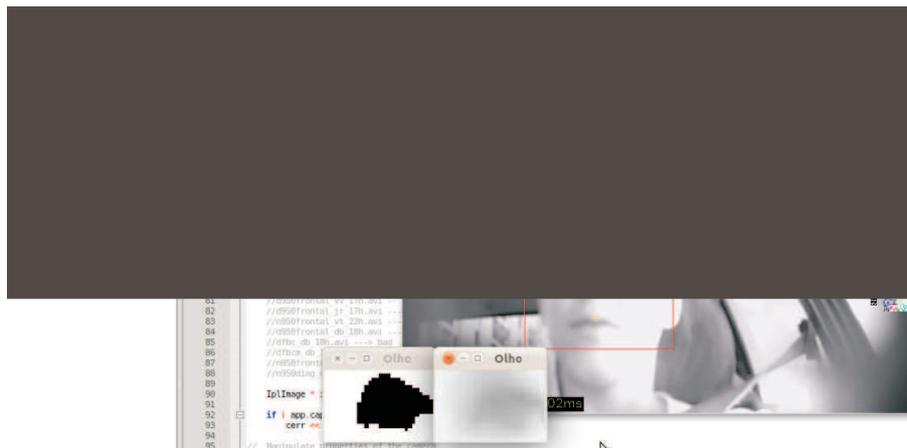


Figura 6.9: Teste do algoritmo baseado em limiarização automática sem iluminação IVP, vídeo coletado no horário crepuscular, vê-se claramente o excesso de luminosidade descaracterizando o rosto do motorista.

6.3.2 Template Matching

Neste algoritmo, a detecção do estado dos olhos e a análise de duração desse evento são baseados na observação e comparação dos escores de correlação obtidos durante rastreamento da RI do olho. A funcionalidade desse algoritmo é condicionada ao armazenamento das imagens usadas como modelos dos olhos (templates). E para que esse modelo seja salvo, o motorista deverá olhar continuamente para câmera durante 10 segundos. Este intervalo de tempo é suficiente para que o algoritmo salve uma imagem do olho que será usada posteriormente como um modelo de olho aberto (ilustrado na Figura 6.10).

Como os olhos fecham durante o piscar, a semelhança com o modelo (template) do olho aberto diminui ao se fechar os olhos. Da mesma forma, ele recupera a sua similaridade com o modelo quando os olhos ficam completamente abertos novamente. A Figura 6.10 ilustra o aumento e diminuição na similaridade, que corresponde diretamente aos escores de correlação entre o modelo, e a imagem da região do olho que esta sendo processada no quadro atual, durante a captura de vídeo.

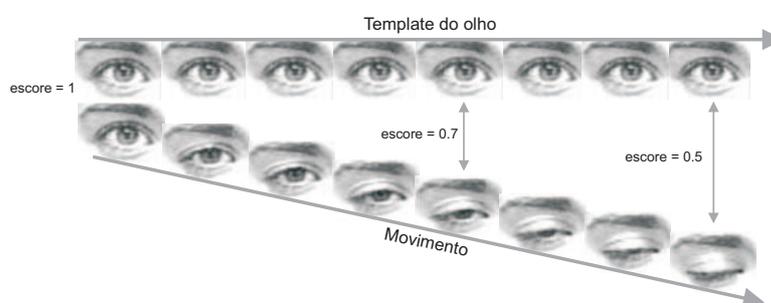


Figura 6.10: Obtenção dos escores, proveniente do aumento e diminuição da similaridade.

Grauman [74] observou que as pontuações/valores dos escores determinam uma forma de onda clara e são úteis para determinar um limiar que será utilizado para classificar o estado dos olhos (abertos/fechado) a cada quadro do vídeo. O gráfico da Figura 6.11 mostra um exemplo de dados coletados indicando quando houve um piscar mais longo seguido por três piscadas curtas e, finalmente, três piscadas longas. Um exame atento dos escores de correlação ao longo dos quadros revela limites claros que permitem a detecção de piscadas. Quando o olho está aberto, os escores de correlação são muito altos, cerca de 0,85 a 1,0. Com o piscar, a pontuação cai para valores de 0,7 a 0,55.

Finalmente, é importante a notar que nesse exemplo o gráfico não contém pontuações abaixo de 0,45. Pontuações nessa faixa normalmente indicam que o rastreador perdeu a localização do olho. E, nesses casos, o sistema deve ser reinicializado para realocar e acompanhar a nova posição do olho. Uma vez conhecidas as escalas de escores de correlação e o que significam, pode-se dividir a implementação do algoritmo nas seguintes etapas:

- Criação online do Template dos olhos;
- Obtenção dos escores de correlação e do estado dos olhos;

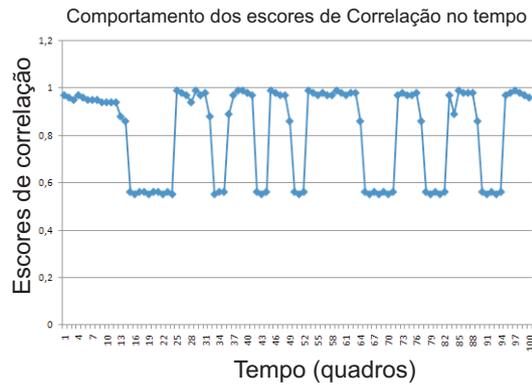


Figura 6.11: Gráfico mostrando o comportamento dos escores de correlação ao longo do tempo.

- Aferir a condição de fadiga.

6.3.2.1 Criação do online do Template

Para armazenar o modelo do olho aberto, durante o período em que o motorista olha fixamente para câmera, foi inserido um pequeno atraso de espera na aplicação, porque, se o motorista piscar o olho durante a operação de salvar o template, é razoável esperar um momento para que ele abra novamente os olhos e o sistema capture nova imagem do olho. A configuração das funções do OpenCV no algoritmo que salva o template imagem do olho pode ser vista no Anexo I.

6.3.2.2 Obtenção dos escores de correlação, estado dos olhos e fadiga

Para obter os escores de correlação, foi utilizado método nativo do OpenCV e foi implementado em C++ através da função:

```
cvMatchTemplate(imgOriginal, imgTemplate, imgResult, CV_TM_CCORR_NORMED);
```

A instrução `cvMatchTemplate` é responsável pelo cálculo de correlação usando `imgOriginal`(imagem original), `imgTemplate`(modelo) e coloca o mapa de correlação em `imgResult`.

Os cálculos utilizados para determinar o mapa e os escores de correlação dependem do último parâmetro, `CV_TM_CCORR_NORMED`. OpenCV oferece seis diferentes métodos de cálculo dos coeficientes de correlação:

```
CV_TM_SQDIFF, CV_TM_SQDIFF_NORMED, CV_TM_CCORR, CV_TM_CCORR_NORMED, CV_TM_CCOEFF, CV_TM_CCOEFF_NORMED
```

Os cálculos com índice `_NORMED` (de normalizado) associado retornam valores de 0 até 1,0; os outros métodos retornam valores maiores. Já o índice `SQDIFF` retorna 0 se o modelo e o template corresponderem. Os outros dois (`CCORR` `CCOEFF`) também são baseados em correlação e retornam 1,0 em caso de correspondência do modelo.

6.3.2.3 Detecção de fadiga

Para detectar a fadiga, este algoritmo utiliza a condição alternativa sugerida por [73], e feito da seguinte forma: nos quadros onde os valores do escore de correlação estão entre 0,8 e 0,55 o olho é considerado fechado. Nos quadros com valores maiores que 0,8, considera-se olho aberto. Quadros com valores inferiores a 0,55 são perdidos e sinalizam para reinicialização do rastreamento do rosto e dos olhos. Ao computar uma sequência de 8 (oito) quadros com valores de escore entre 0,8 e 0,55, emite-se o alarme indicativo de fadiga.

O testes realizados para validar o algoritmo baseado em *Template Matching*, podem ser vistos nas na Figura 6.12, Figura 6.13 e Figura 6.14, que mostram a identificação do estado do olho fechado, em diversas condições de iluminação e posicionamento do conjunto de iluminação IVP e captura. Que incluem, condições noturnas utilizando iluminação IVP (Figuras 6.12 e Figura 6.13 e diurnas no horário crepuscular sem iluminação IVP (Figura 6.14). No Anexo I, estão dispostas as configurações das funções do OpenCV determina o estado do olho a partir dos escores de correlação.



Figura 6.12: Teste em horário noturno do algoritmo baseado em Template Matching, identificando o olho fechado, com iluminação IVP, câmera localizada na tapa sol.



Figura 6.13: Teste do algoritmo baseado em Template Matching com iluminação IVP, horário noturno.

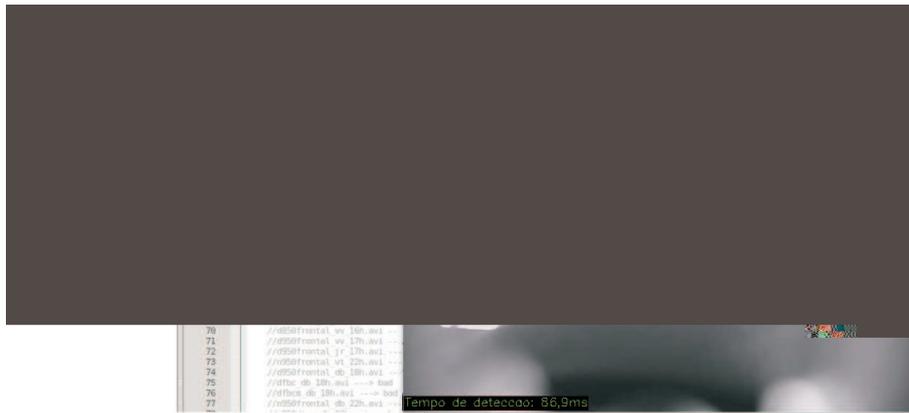


Figura 6.14: Teste do algoritmo baseado em Template Matching sem iluminação IVP, horário crepuscular.

6.3.3 Limiarização elevada

O terceiro algoritmo que foi implementado é derivado da propriedade de reflexão dos raios infravermelhos pela córnea. Essa propriedade faz com que o olho, quando iluminado por uma luz infravermelha, apareça na imagem de uma câmera com um círculo de alta intensidade na região da córnea. Uma característica tão marcante na imagem pode ser usada para diminuir a complexidade computacional em algoritmos para determinar o estado dos olhos.

A proposta de decisão do algoritmo baseia-se em aplicar um limiar alto à imagem. Assim, quando a imagem do olho é transformada em imagem binária e é aplicado a ela um valor de limiar muito alto, torna-se uma imagem preta, com apenas a região de córnea em branco, e pode ser observado na extrema direita da Figura 6.15. Para olho fechado, a região estaria inteiramente preta, pode ser visualizado na Figura 6.16. Depois, calcula-se a intensidade média da imagem. Se a média do valor da intensidade for zero, o olho está fechado. Caso contrário, quando a intensidade média é diferente de zero, a córnea é detectada e o olho está aberto.

6.3.3.1 Detecção de fadiga

Neste algoritmo, o estado de alerta pode ser analisado utilizando o mesmo método alternativo para determinar fadiga sugerido por [73] que foi utilizado no algoritmo baseado em *Template Matching*, ou seja, monitorizando o estado dos os olhos, quando permanecerem fechados por 8 ou mais quadros consecutivos é emitido o alerta de fadiga. No Anexo I, estão dispostas as configurações da funções do OpenCV em C++ no algoritmo que afere a fadiga da limiarização elevada.

6.3.4 Resultados - Estatística do estado de alerta em motoristas

A estatística para validar a acurácia dos algoritmos propostos considera o estado dos olhos, a emissão de alertas e os falsos positivos e negativos. Os 4 motoristas que participaram dos testes de campo estavam em estado de alerta, independentemente do horário de coleta dos dados. E, portanto, a maioria dos vídeos armazenados é efetiva apenas para validar o estado do olho (aberto



Figura 6.15: Teste do algoritmo baseado em limiar alto no laboratório com filtro IVP de baixo custo, estado do olho fechado.



Figura 6.16: Teste do algoritmo baseado em limiar alto no laboratório com filtro IVP de baixo custo, sonolência.

ou fechado), impossibilitando sua utilização na coleta de dados investigativos apurados da fadiga. Entretanto, os sujeitos simularam situações de fadiga as quais foram corretamente detectadas pelos algoritmos.

Mesmo assim, e considerando que, futuramente, serão realizados testes de campo no ambiente veicular definitivo (veículos de grande porte) e para onde a aplicação se destina, considerando a diferença substancial entre os dois ambientes, um ambiente de teste foi montado em laboratório, utilizando o mesmo hardware e captura e iluminação IVP (na maioria dos testes, a disposição dos LEDs IVP favoreceu a reflexão da pupila).

A Tabela 6.1 mostra os resultados do teste com cada algoritmo separadamente, onde *Olhos Fechados* representa o número de vezes em que ocorreram piscadas (abertura e fechamento dos olhos). *Simulação de Fadiga* é o número de vezes em que foi simulada a fadiga através de um piscar de olhos prolongado. *Alarmes Corretos* é número de vezes que foram emitidos alarmes que detectaram corretamente a simulação de fadiga. *Falso Positivo* é o número de alarmes emitidos erroneamente. *Falso Negativo* é o numero de vezes em que a simulação de fadiga não foi detectada. A taxa de acertos é definida pela Equação 6.2.

$$\text{Taxa de Acertos} = \frac{\text{Alarmes Corretos}}{\text{Fadiga}} \quad (6.2)$$

Tabela 6.1: Resultado da detecção do estado do olho e da fadiga, utilizando os 3 algoritmos propostos para o sistema.

	Limiar automático	Limiar alto	Template Matching
Total de Quadros	3850	3850	3850
Olhos Fechados	78	83	55
Simulação de Fadiga	35	52	28
Falso Positivo	7	0	5
Falso Negativo	1	0	1
Alarmes Corretos	33	51	27
Taxa de Acertos	94,3%	98,1%	96,4%

A análise dos dados revela que o algoritmo baseado em limiarização automática apresentou um desempenho inferior aos demais, mas que pode ser justificado pelo uso de apenas uma disposição de LEDs durante a maioria dos testes no laboratório desse algoritmo.

A disposição os LEDs que foi utilizada, favorece a reflexão da córnea, criando uma área clara no centro da íris prejudicando a funcionalidade da parte do algoritmo que calcula o número de pixels e, a partir desse valor estima a distância da pálpebra.

Esse problema poderá ser contornado utilizando-se uma disposição de LEDs que favoreça a iluminação do ambiente, associada a uma nova ferramenta para controle de luminosidade que está em fase de implementação e testes. Um fator importante também deve ser considerado. Mesmo apresentando um desempenho inferior aos demais, o algoritmo baseado em limiarização automática é o unico capaz de monitorizar e predizer o estado de alerta do motorista.

Capítulo 7

Conclusões e trabalhos futuros

Os objetivos desta dissertação foram: (I) desenvolver um sistema capaz de detectar fadiga em motoristas integrando ferramentas de visão computacional; (II) embarcar o sistema em um dispositivo de baixo custo de produção; (III) inserir e testar o dispositivo em ambiente veicular.

Conforme demonstrado nos Capítulos anteriores, os algoritmos para detectar e rastrear rosto e olhos foram implementados e testados com desempenho satisfatório, maior que 98% para a face e 95% para os olhos, em situações reais de funcionamento em veículos de passeio. Também foram implementados e testados três algoritmos que identificam o estado dos olhos e aferem o nível de fadiga em motoristas, com índices de sucesso maiores que 94%.

A implementação de mais de um algoritmo pode ser justificada pelo acréscimo de segurança ao sistema, uma vez que o desempenho desses algoritmos pode ser afetado pelas condições de luminosidade e disposição dos sistemas de captura e iluminação IVP dentro do veículo de destino (ônibus e caminhões). Mais de uma opção de algoritmo oferece maior grau de sucesso em superar dificuldades relacionadas a incertezas de um ambiente veicular onde o sistema ainda não foi testado.

7.1 Trabalhos futuros

A distração de motoristas ao conduzir veículos é um fator causador de um número significativo de acidentes no trânsito (por exemplo, conversando ao celular) e sua monitorização dever ser considerada em qualquer projeto que envolva dispositivos de alerta comportamental para motoristas. Adicionalmente, em caráter embrionário, foi testada uma metodologia para aferir o grau de distração do motorista conversando por meio de uma ferramenta que monitoriza os movimentos da boca, e pode ser adicionada ao sistema atual.

Finalmente, os algoritmos implementados associados ao controle de luminosidade, juntamente com o mecanismo que monitoriza índice de distração do motorista, provê ao dispositivo, um diferencial inexistente nos demais sistemas encontrados no mercado e na literatura.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] KEINER, L. O especto infravermelho. *Coastal Carolina University*.
- [2] BERGASA, L.; NUEVO, J.; GUILL, M. Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems*, v. 7(1), p. 63–77, 2006.
- [3] JI, Q.; YANG, X. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, v. 8, p. 357–377, 2002.
- [4] JI, Q.; ZHU, Z.; LAN, P. Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Transactions on Vehicular Technology*, v. 53(4), p. 1052–1068, 2004.
- [5] JIN, K. et al. Real-time pupil detection based on three-step hierarchy. paper presented at the signal processing. *Paper presented at the Signal Processing Proceedings. 7th International Conference on, ICSP 04*, 2004.
- [6] BRISTOW, D.; FRITH, C.; REES, G. Two distinct neural effects of blinking on human visual processin. *NeuroImage*, v. 27(1), 2005.
- [7] FILTRO IV de baixo custo. Disponível em: <http://www.northcountryradio.com/Articles/irfltr.htm> – Acessado em: 06/07/2011.
- [8] CEMSA. Disponível em: <http://www.cemsa.org.br/midiartigos.html> – Acessado em: 08/07/2011.
- [9] HYPNOS PRF. Disponível em: <http://www.forumdaseguranca.com/site/ver.php?codigo=185> – Acessado em: 08/07/2011.
- [10] ANTT. Disponível em: <http://www.antt.gov.br/relatorioanual/relatorioanual.asp> – Acessado em: 08/07/2011.
- [11] ORON-GILAD, T.; RONEN, A.; SHINAR, D. Alertness maintaining tasks amts while driving. *Accident Analysis & Prevention*, v. 40, p. 851–860, 2008.
- [12] PAPADELIS, C. et al. Monitoring sleepiness with on board electrophysiological recordings for preventing sleep deprived traffic accidents. *Clinical Neurophysiology*, v. 118(9), p. 1906 – 1922, 2007.
- [13] FABER, J. Detection of different levels of vigilance by eeg pseudo spectra. *NN World*, v. 14(3), p. 285–290, 2004.

- [14] WRIGHT, N. et al. A review of in-vehicle sleepiness detection devices. *TRL Limited*, Report PPR157, 2007.
- [15] WAKITA, T.; OZAWA, K.; MIYAJIMA, C. Driver identification using driving behavior signals. *IEICE*, v. 89, p. 1188–1194, 2006.
- [16] TAKEI, Y.; FURUKAWA, Y. Estimate of drivers fatigue through steering motion. *IEICE*, v. 2, p. 1765–1770, 2005.
- [17] MCCALL, J.; TRIVEDI, M.; RAO, B. Lane change intent analysis using robust operators and sparse bayesian learning. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR05, 2005.
- [18] VOLVO car corporation. Disponivel em: <http://www.media.volvocars.com> – Acessado em: 06/07/2011.
- [19] MERCEDES-BENZ. Disponivel em: <http://www.emercedesbenz.com/> – Acessado em: 05/07/2011.
- [20] LEXUS. Disponivel em: <http://www.testdriven.co.uk/lexus-ls-600h/> – Acessado em: 04/07/2011.
- [21] DORAZIO, T.; LEO, M.; GUARAGNELLA, C. A visual approach for driver inattention detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, v. 40(8), p. 2341–2355, 2007.
- [22] SUZUKI, M.; YAMAMOTO, N.; NAKANO, T. Measurement of drivers consciousness by image processing a method for presuming driveras drowsiness by eye-blinks coping with individual differences. *Systems Man and Cybernetics*, v. 4, p. 2891–2896, 2006.
- [23] BRETZNER, L.; KRANTZ, M. Towards low-cost systems for measuring visual cues of driver fatigue and inattention in automotive applications. *Vehicular Electronics and Safety*, p. 161–164, 2005.
- [24] HEINZMANN, J.; TATE, D.; SCOTT, R. Using technology to eliminate drowsy driving. *PE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production*, v. 8(5), p. 161–164, 2008.
- [25] SMART eye. Disponivel em: <http://www.smarteye.se> – Acessado em: 04/07/2011.
- [26] SEEINGMACHINES. Disponivel em: <http://www.seeingmachines.com> – Acessado em: 08/07/2011.
- [27] FITPC2. Disponivel em: <http://www.fit-pc.com/web/> – Acessado em: 03/07/2011.
- [28] OPENCV. Disponivel em: <http://sourceforge.net/projects/opencvlibrary/> – Acessado em: 08/07/2011.
- [29] OPENCV. Disponivel em: <http://sourceforge.net/projects/opencvlibrary/> – Acessado em: 08/07/2011.

- [30] BRADSKI, G.; KAEHLER, A. Learning opencv computer vision with the opencv library. *O'Reilly*, 2008.
- [31] INTEL Corporation. Disponivel em: <http://www.intel.com> – Acessado em: 05/07/2011.
- [32] KOONS, D.; FLICKER, M. *IBM Blue Eyes Project*. Disponivel em: <http://almaden.ibm.com/cs/BlueEyes/index.html> – Acessado em: 05/07/2011.
- [33] SLINEY, D.; WOLBARST, M. *Safety with Lasers and Other Optical Sources*. [S.l.]: New York: Plenum Press, 1980. 147 p.
- [34] MATTHES, R. Statement on light-emitting diodes (leds) and laser diodes: Implications for hazard assessment. *ICNIRP Health Physics Society*, v. 78(6), p. 744–752, 2000.
- [35] SLINEY, D. Statement on light emitting diodes (leds) and laser diodes: Implication for hazard assessment. *ICNIRP Health Physics*, v. 78(6), p. 744–752, 2000.
- [36] VATANI, N.; ROBERTS, J. Automatic camera exposure control. *CSIRO ICT Centre Australia*.
- [37] ELLIOT, D. Ddx video – a light weight video framework for autonomous robotic platforms. *In Proceedings of Australasian Conference on Robotics and Automation*, 2005.
- [38] SHIRVAIKAR, M. An optimal measure for camera focus and exposure. *In Proceedings of the IEEE*, 2004.
- [39] JANNERUP; SOERENSEN. Introduktion til regulering-steknik. *Polyteknisk Forlag*, 2004.
- [40] TURK, M.; PENTLAND, A. Face recognition using eigen faces. *IEEE The Media Laboratory (MIT)*, 2001.
- [41] GARCIA, G.; ZIKOS, G.; TZIRITAS, G. Wavelet packet analysis for face recognition. *Elsevier Image and Vision Computing*, v. 18, p. 289–297, 2000.
- [42] LIU, C. Gabor-based kernel pca with fractional power polynomial models for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligent*, v. 26(5), p. 572–582, 2004.
- [43] SMITH, P.; SHAH, M.; LOBO, N. Determining driver visual attention with one camera. *IEEE Transactionson Intelligent Transportation Systems*, v. 4(4), 2003.
- [44] PHUNG, S.; CHAI, D.; BOUZERDOUM, A. A universal and robust human skin color model using neural networks. *IEEE International Joint Conference on Neural Networks*, v. 4, p. 2844–2849, 2001.
- [45] BRADSKI, G. Computer vision face tracking for use in a perceptual user interface. *Intel Technical Journal*, v. 2, p. 1–15, 1998.
- [46] KIM, I.; LEE, S.; KWON, Y. 3d tracking of multi-objects using color and stereo for hci. *IEEE International Conferences on Image Processing*, v. 3, p. 278–281, 2001.

- [47] SANTANA, M. *Sobre la Deteccionen Tiempo Real de Caras en Secuencias de Video. Una Aproximacion Oportunista*. Tese (Doutorado) — Universidad de las Palmas de Gran Canaria, Diciembre 2002.
- [48] BERGASA, L. *Seguimiento Facial Mediante Vision Artificial Orientado ala Ayuda ala Movilidad*. Tese (Doutorado) — Universidad de Alcala, Escuela Politecnica, Departamento de Electronica, Diciembre 2002.
- [49] LI, S.; CHU, R.; ZHANG, L. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Recognition Analysis and Machine Intelligence*, v. 29(4), p. 627–639, 2001.
- [50] VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition*, p. 511–518, 2001.
- [51] LIENHART, R.; MAYDT, J. An extended set of haar-like features for rapid object detection. *in IEEE ICIP 2002*, p. 900–903, 2002.
- [52] SOHAIL, A.; BHATTACHARYA, P. Detection of facial feature points using anthropometric face model. p. 660, 2006.
- [53] SOHAIL, A.; BHATTACHARYA, P. Localization of facial feature regions using anthropometric face model. *In: I International Conference on Multidisciplinary Information Sciences and Technologies*, 2006.
- [54] FARKAS, L. *Anthropometry of the head and face*. Raven Press, New York, 1994.
- [55] BARBATO, G.; FICCA, G.; CASIELLO, M. Effects of sleep deprivation on spontaneous eye blink rate and alpha eeg power. *Biological Psychiatry*, v. 38, p. 340–341, 1995.
- [56] SEKI, M.; SHIMOTANI, M.; NISHIDA, M. A study of blink detectio using bright pupils. *JSAE Review*, v. 19, p. 58–61, 1998.
- [57] TSUBOTA, K. et al. Abnormal blinking of newscasters. *The Lancet*, v. 354(9175), p. 308, 1999.
- [58] FUKUDA, K. Eye blinks: new indices for the detection of deception. *International Journal of Psychophysiology*, v. 40, p. 239–245, 2001.
- [59] HORNE; BAULK. Awareness of sleepiness when driving. *International Journal of Psychophysiology*, 2004.
- [60] LAL, S.; CRAIG, A. A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, v. 55(3), p. 173–194, 2001.
- [61] SANTAMARIA, J.; CHIAPPA, K. The eeg of drowsiness in normal adults. *Journal of Clinical Neurophysiology*, v. 4, p. 327–382, 1987.
- [62] MORRIS, T.; MILLER, J. Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology Psychophysiology of Workload*, v. 42(3), p. 343–360, 1996.

- [63] ORDEN, K. V.; JUNG, T.; MAKEIG, S. Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, v. 52(3), p. 221–240, 2000.
- [64] MALLIS, M. Evaluation of techniques for drowsiness detection: Experiment on performance-based validation of fatigue-tracking technologies. *Drexel University*, 1999.
- [65] WIERWILLE, W.; ELLSWORTH, L. Evaluation of driver drowsiness by trained raters. *Accident Analysis & Prevention*, v. 26(5), p. 571–581, 1994.
- [66] JI, Q.; BEBIS, G. Visual cues extraction for monitoring drivers' vigilance. *Paper presented at the First HONDA Initiation Grant Forum: The HIG Symposium*, 1999.
- [67] WIERWILLE, W.; ELLSWORTH, L. Research on vehicle-based driver status performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness. *DOT HS 808*, p. 247, 1994.
- [68] GRACE, R.; BYRNE, V.; BIERMAN, D. A drowsy driver detection system for heavy vehicles. *The AIAAIEEE SAE Digital Avionics Systems Conference*, 1998.
- [69] JI, Q.; YANG, X. Real time visual cues extraction for monitoring driver vigilance. *Unpublished manuscript*, 2001.
- [70] LAL, S.; A.CRAIG. Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology*, v. 39(3), p. 313–321, 2002.
- [71] PEIRIS, M.; JONES, R.; DAVIDSON, P. Frequent lapses of responsiveness during an extended visuomotor tracking task in non-sleep-deprived subjects. *Journal of Sleep Research*, v. 15, p. 291–300, 2006.
- [72] MA, H. et al. A fast method for monitoring driver fatigue using monocular camera. *Tsinghua University, Beijing*, 2008.
- [73] HORNG, W. et al. Driver fatigue detection based on eye tracking and dynamic template matching. *Sensing and Control*, p. 7–12, 2004.
- [74] GRAUMAN, K.; BETKE, M.; G.BRADSKI. Communication via eye blink and eye-brow raises: Video-based human-computer interfaces. *Universal Access In The Information Society*, v. 2(4), p. 359–373, 2003.