



FACULDADE DE TECNOLOGIA SENAI CIMATEC
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL E TECNOLOGIA INDUSTRIAL
Doutorado em Modelagem Computacional e Tecnologia Industrial

Tese de Doutorado

**Modelo computacional para navegação de cadeira de
rodas baseada no elance entre os algoritmos de
classificador em cascata e modelos de forma ativa
modificado**

Apresentada por: Marcos Batista Figueredo
Orientador: Roberto Luiz Souza Monteiro

29 de Dezembro de 2016

Marcos Batista Figueredo

Modelo computacional para navegação de cadeira de rodas baseada no elance entre os algoritmos de classificador em cascata e modelos de forma ativa modificado

Tese de Doutorado apresentada ao Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial, Curso de Doutorado em Modelagem Computacional e Tecnologia Industrial do Faculdade de Tecnologia SENAI CIMATEC, como requisito parcial para a obtenção do título de **Doutor em Modelagem Computacional e Tecnologia Industrial**.

Área de conhecimento: Interdisciplinar

Orientador: Roberto Luiz Souza Monteiro

Faculdade de Tecnologia SENAI CIMATEC

Salvador
Faculdade de Tecnologia SENAI CIMATEC
Vol.1 2016

Ficha catalográfica elaborada pela Biblioteca da Faculdade de Tecnologia SENAI CIMATEC

F475m Figueiredo, Marcos Batista

Modelagem computacional para navegação de cadeira de rodas baseada no elance entre os algoritmos de classificador em cascata e modelos de forma ativa modificado / Marcos Batista Figueiredo. – Salvador, 2016.

101 f. : il. color.

Orientador: Prof. Dr. Roberto Luis Souza Monteiro.

Tese (Doutorado em Modelagem Computacional e Tecnologia Industrial) – Programa de Pós-Graduação, Faculdade de Tecnologia SENAI CIMATEC, Salvador, 2016.
Inclui referências.

1. Visão computacional. 2. Active Shape Models. 3. Classificador cascata. 4. Cadeira de rodas. 5. Detecção facial. I. Faculdade de Tecnologia SENAI CIMATEC. II. Monteiro, Roberto Luis Souza. III. Título.

CDD: 620.00113

Nota sobre o estilo do PPGMCTI

Esta tese de doutorado foi elaborada considerando as normas de estilo (i.e. estéticas e estruturais) propostas aprovadas pelo colegiado do Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial e estão disponíveis em formato eletrônico (*download* na Página Web http://ead.fieb.org.br/portal_faculdades/dissertacoes-e-teses-mcti.html ou solicitação via e-mail à secretaria do programa) e em formato impresso somente para consulta.


Ressalta-se que o formato proposto considera diversos itens das normas da Associação Brasileira de Normas Técnicas (ABNT), entretanto opta-se, em alguns aspectos, seguir um estilo próprio elaborado e amadurecido pelos professores do programa de pós-graduação supracitado.

Faculdade de Tecnologia SENAI CIMATEC


Doutorado em Modelagem Computacional e Tecnologia Industrial

A Banca Examinadora, constituída pelos professores abaixo listados, leu e aprovou a Tese de doutorado, intitulada "Modelo para navegação de cadeira de rodas baseada em classificadores em cascata e modelos de aparência ativa", apresentada no dia 29 de dezembro de 2016, como parte dos requisitos necessários para a obtenção do Título de Doutor em Modelagem Computacional e Tecnologia Industrial.

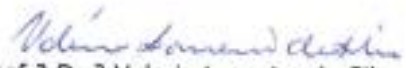
Orientador:


Prof. Dr. Roberto Luiz Souza Monteiro
SENAI CIMATEC

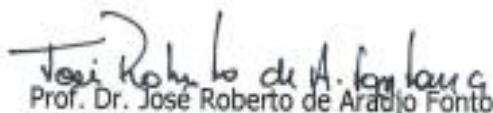
Membro Interno:


Prof. Dr. Hernane Borges de Barros Pereira
SENAI CIMATEC


Membro Interno:


Prof.ª Dr.ª Valeria Loureiro da Silva
SENAI CIMATEC

Membro Externo:


Prof. Dr. José Roberto de Araújo Fontoura
UNEB

Membro Externo:


Prof. Dr. Eduardo Manuel de Freitas Jorge
UNEB

Dedico este trabalho a todas as pessoas que contribuiriam direta ou indiretamente a ele:
a minha amada esposa Soraia aos nossos queridos filhos Guilherme e Gustavo, que sempre me deram energia para construir este caminho; A minha família, que sempre esteve e sempre estará ao meu lado, com especial carinho a Jussara e Jandiara, minhas queridas irmãs, sem vocês não chegaria até aqui. Ao meu querido Egberto Melo, uma grande fonte de inspiração, meus sobrinhos Egberto, Flávia e Juliana, a minha sogra Maria Rita, pelas correções a apoio e minha cunhada Suzana pela ajuda nos momentos difíceis. A toda família Argoinense que estão presentes em todos os momentos da minha vida. E agradecer em especial aos meus avós Jandira e Arlindo, que, onde quer que estejam, saibam que continuo amando vocês e com muitas saudades.

Agradecimentos

Ao Prof. Dr. Roberto Monteiro, mestre e amigo, pelos ensinamentos e orientações sempre objetivas e principalmente pela confiança e incentivo para realização desta pesquisa.

Aos Prof. Dr. Hernane Pereira, Dr. Marcelo Moret, Dra. Lynn Alves e Dr. Renelson Sampaio por fundamentar os meus primeiros passos na área de pesquisa com muita paciência e dedicação, o que possibilitou a busca por novos horizontes, meu sincero reconhecimento.

Aos amigos do curso de Modelagem computacional pelo aprendizado diário, compartilhamento de ideias e bons momentos. Aos amigos Jeferson, Paulo, Portugal pelas discussões, auxílio, correções e brincadeiras. Foram momentos que guardarei eternamente. Em especial ao amigo e parceiro Alexandre a quem dedico meu apreço por todas as horas que interagimos em prol deste trabalho e de nossa amizade. A todos os colaboradores do sistema FIEB que ajudaram no meu percurso. Aos professores do Doutorado que estiveram sempre dispostos a colaborar, as secretárias pela atenção e ajuda na burocracia. Ao sistema FIEB, na figura do Prof. Dr. Alex Sales, que proporcionou o auxílio financeiro para a realização deste curso.

A toda minha família que sempre me apoiou de forma sábia, humilde e amorosa ensinando-me a vencer em cada momento de dificuldade, E neste momento especial, minha sincera gratidão pelo apoio incansável.

Salvador, Brasil
29 de Dezembro de 2016

Marcos Batista Figueredo

Resumo

A mobilidade é um importante meio de interação social que, além de permitir a realização de diversas tarefas do cotidiano, estabelece uma conexão do paciente com o universo social e do trabalho. No Brasil mais de 300.000 pessoas não possuem qualquer tipo de movimento nos membros inferiores e superiores, as chamadas paraplegias ou tetraplegias e a maioria desse grupo, utiliza cadeira de rodas para exercer sua cidadania. O deslocamento da maioria das cadeiras é realizado pela força humana e em algumas delas se utiliza algum tipo de suporte tecnológico que permita a navegação pelo paciente. Diversas pesquisas buscam aprimorar estes suportes de forma a tornarem a navegação simples e eficiente, mas, em geral, as soluções apresentadas possuem uma grande quantidade de sensoriamento, intrusividade e alto custo. Pacientes acometidos por esta(s) deficiência(s) possuem apenas os movimentos da face para estabelecer qualquer tipo de interação Homem×Máquina(HMI). Assim, propomos o desenvolvimento de um modelo computacional que permita a navegação de uma cadeira de rodas utilizando expressões faciais. Ao contrário dos trabalhos anteriormente propostos que utilizam (HMI), sugerimos um modelo que se baseia em duas expressões faciais: a pose da cabeça e o fechamento dos olhos, e utilizando apenas um sensor de entrada e uma câmera USB. O modelo converte as expressões faciais em comandos simples para a navegação da cadeira e utilizando-se duas técnicas para a interpretação destas expressões: a primeira os classificadores em cascata e a segunda os modelos de forma ativa. Na primeira, apresentamos um classificador capaz de detectar o fechamento dos olhos do paciente a uma distância de 50cm da câmera e na segunda utilizamos o casamento entre a resposta do modelo de forma ativa com o coeficiente de correlação de Pearson que permite estabelecer a posição cabeça do paciente. Nossos experimentos mostram que o modelo tem excelente acurácia e precisão e um desempenho robusto na detecção dos olhos fechados e estimação da pose, contornando muito bem problemas naturais do reconhecimento de padrões como oclusão e iluminação. A resposta do modelo alcançou 98% de acerto médio com uma taxa de falsos positivos na casa de 2%.

Palavras-chave: Visão Computacional, *Active Shape Models*, Classificador Cascata, Cadeira de rodas, Detecção Facial

Abstract

Mobility is an important means of social interaction that, besides allowing the accomplishment of several daily tasks, establishes a connection of the patient with the social and work universe. In Brazil More than 300,000 people do not have any type of movement in the lower and upper limbs, the so-called paraplegias or tetraplegias and most of that group, use wheelchair to exercise their citizenship. The displacement of most of the chairs is performed by human force and in some of them some kind of technological support is used that allows the patient to navigate. Several researches seek to improve these supports in order to make the navigation simple and efficient, but, in general, the presented solutions have a great amount of sensing, intrusiveness and high cost. Patients affected by this deficiency (s) have only the movements of the face to establish any type of Human Machine interaction (HMI). Thus, we propose the development of a computational model that allows the navigation of a wheelchair using facial expressions. In contrast to previously proposed studies (HMI), we suggest a model that is based on two facial expressions: the pose of the head and the closing of the eyes, using only an input sensor and a USB camera. The model converts facial expressions into simple commands for navigating the chair and using two techniques for the interpretation of these expressions: first the cascade classifiers and the second the models in an active way. In the first one, we present a classifier capable of detecting the closure of the patient's eyes at a distance of 50 cm from the camera, and in the second, we used the marriage between the model's response actively and Pearson's correlation coefficient, which allows us to establish the head position of the patient. Our experiments show that the model has excellent accuracy and precision and robust performance in closed eye detection and pose estimation, bypassing very well the natural problems of pattern recognition such as occlusion and illumination. The model response achieved 98 % average hit with a false positive rate in the house of 2 %.

Sumário

1	Introdução	1
1.1	Definição do problema	4
1.2	Objetivo	7
1.2.1	Objetivo Geral	7
1.2.2	Objetivos Específicos	7
1.3	Importância da Pesquisa	7
1.4	Motivação	8
1.5	Limites e Limitações	9
1.6	Hipóteses	10
1.7	Aspectos Metodológicos da Pesquisa	10
1.8	Contribuições	11
1.9	Organização da tese	12
2	Fundamentação Teórica	13
2.1	Introdução	13
2.2	Estado da Arte	13
2.2.1	Projetos de CRs	13
2.3	Análise comparativa dos Projetos de CR	16
2.3.1	Detecção do estado dos olhos	19
2.3.2	Detecção do Movimento da Cabeça	23
2.4	Análise comparativa dos métodos	25
2.5	Conclusões	27
3	Modelo Proposto	29
3.1	Definições iniciais - Pré-processamento	29
3.1.1	Definição de Imagem Digital	29
3.1.2	Conversão em escala cinza	30
3.1.3	Filtro de Retina	30
3.1.4	Extração da região de interesse	31
3.1.5	Extração da região dos olhos	31
3.1.6	Distância de Mahalanobis	33
3.2	Técnica para Detecção dos Olhos Fechados	33
3.2.0.1	Características de Haar	34
3.2.0.2	Imagem Integral	35
3.2.0.3	Algoritmo de Aprendizagem - AdaBoost	37
3.2.0.4	Um classificador em cascata	39
3.3	Técnica para detecção dos movimentos da cabeça	40
3.3.1	Modelo de Pontos Distribuídos - PDM	42
3.3.1.1	Etiquetagem do conjunto de treino	42
3.3.1.2	Alinhamento das formas	43
3.3.1.3	Análise de Componentes Principais	44
3.3.2	Aparência dos níveis de cinza	45
3.3.3	Localização dos pontos característicos em uma nova imagem	46
3.3.4	Coefficiente de correlação de Pearson	46
3.4	Arquitetura da cadeira	47

3.4.1	Entradas do usuário	48
3.4.2	Sensores/Dispositivos	49
3.5	Arquitetura do Modelo Computacional	49
3.6	Limitações do Modelo	53
3.7	Conclusão	55
4	Resultados e discussão	57
4.1	Preliminares	57
4.1.1	Tempo de resposta do modelo	57
4.1.2	Modelo Matemático da CR	58
4.1.3	Gráfico ROC	60
4.1.3.1	Acurácia	61
4.1.3.2	Sensibilidade	62
4.1.3.3	Precisão	62
4.1.3.4	Especificidade	62
4.1.4	Conjunto de positivos para o classificador em cascata	62
4.1.5	Conjunto de negativas para o classificador em cascata	62
4.1.6	Samples	63
4.1.7	Parâmetros de Treinamento	63
4.1.8	Conjunto de treino do ASM	64
4.1.9	Software que implementa o modelo	66
4.2	Resultados dos testes do Classificador em cascata	67
4.3	Resultados dos testes no Detector dos Movimentos da cabeça	71
4.4	Resultados da integração das técnicas	76
4.5	Análise dos Resultados	84
5	Conclusões	87
5.1	Trabalhos Futuros	90
	Referências	92

Lista de Tabelas

1.1	Exemplos de projetos de Cr's	3
1.2	Projetos de CR e seus componentes	6
2.1	Projetos de CR e suas características	17
2.2	Alguns trabalhos relevantes na área de detecção de objetos	22
2.3	Análise geral dos métodos de detecção de olhos fechados	26
2.4	Análise geral dos métodos de detecção de movimentos da Cabeça	27
3.1	Tabela com dados para extração da região dos olhos em que $W = 800$ que corresponde a 200 <i>pixels</i> de olho a olho	32
4.1	Matriz de confusão para construção do gráfico ROC	61
4.2	Possibilidades para o detector após o treinamento	61
4.3	Parâmetros que podem ser alterados nos classificadores	64
4.4	Apresentação dos parâmetros e dos classificadores	68
4.5	Resultados geral dos testes realizados	81
4.6	Tabela Comparativa entre nosso projeto e alguns projetos de CR	84
4.7	Comparativo entre nossa abordagem e os métodos de detecção de olhos fechados	85
4.8	Análise geral dos métodos de detecção de movimentos da Cabeça	86

Lista de Figuras

1.1	Modelo comercial aplicado para pessoas com tetraplegia	2
1.2	Projeto Smart Chair que utiliza uma cadeira altamente sensorizada	2
1.3	Projeto LURCH	3
1.4	Projeto IntellWhells	3
1.5	Diagrama dos elementos que compõem a base de um sistema de navegação baseado em visão computacional	5
1.6	Roteiro metodológico	11
2.1	Imagens utilizadas para treinamento do modelo	21
3.1	Aplicação do Filtro de Retina a imagem do olho	31
3.2	Exemplo de extração da face numa imagem complexa	31
3.3	Extração do Olho	32
3.4	Pré-processamento da informação de entrada do modelo	33
3.5	Ilustração de uma janela de Haar	34
3.6	Algumas características de Haar e a janela de detecção	35
3.7	Representações da Imagem Integral, (a) região de cálculo, (b) Soma das áreas A-B-C+D, (c) região racionada de cálculo, (d) Soma rápida A-B-C+D	36
3.8	Representação do cálculo da imagem Integral, observamos que a soma da região em (a) é igual a sete em (b) representamos como $108-73-80+52$. . .	36
3.9	Modelo do Classificador em Cascata	40
3.10	Etiquetagem realizada em uma das imagens	43
3.11	Vizinhança dos pontos marcados em tons de cinza de uma amostra	45
3.12	Hardware da cadeira	48
3.13	Hardware da cadeira em blocos	48
3.14	Arquitetura do modelo	51
3.15	Rotação captada pelo detector	53
3.16	Dimensões mínimas da cadeira para utilização do modelo proposto	55
3.17	Diagrama interação Homem Maquina	55
4.1	Modelo matemático da cadeira	59
4.2	Saída do simulador, em Matlab, e suas entrada obtidas do software c++	60
4.3	Conjunto de imagens positivas com dimensão de 24×24 pixel	62
4.4	Conjunto de imagens negativas com dimensão variadas	63
4.5	Parcela do conjunto de imagens utilizado para treinar o ASM	65
4.6	Rotações da face utilizadas como templete	65
4.7	Rotações da face em 3 dimensões	66
4.8	Arquivo com trecho dos códigos do software	67
4.9	Gráfico ROC com os melhores classificadores	68
4.10	Comparação da resposta do classificador a imagens com ou sem oclusão	69
4.11	Leitura do estado do olho tomado com 10 amostras	70
4.12	Gráfico ROC das amostras	71
4.13	Exemplo de pontos marcados na posição central	71
4.14	Teste inicial realizado com uma única imagem avaliando o modulo	72
4.15	Comparação do conjunto CL com os templates	73

4.16	Comparação do conjunto CC com os templates	73
4.17	Comparação do conjunto CD com os templates	73
4.18	Rotação da cabeça amostra 1	74
4.19	Rotação da cabeça amostra 2	75
4.20	Rotação da cabeça amostra 3	75
4.21	Rotação da cabeça amostra 4	76
4.22	Comportamento da nevegação no simulador	77
4.23	Caminho percorrido pela cadeira: cenário 1	77
4.24	Caminho percorrido pela cadeira: cenário 2	78
4.25	Caminho percorrido pela cadeira: cenário 3	78
4.26	Caminho percorrido pela cadeira: cenário 4	79
4.27	Caminho percorrido pela cadeira: cenário 5	79
4.28	Caminho percorrido pela cadeira: cenário 6	80
4.29	Caminho percorrido pela cadeira: cenário 7	80
4.30	Cenário 1: deslocamento horizontal	81
4.31	Cenário 2: deslocamento ao redor de um obstáculo	82
4.32	Cenário 3: deslocamento com desvio de obstáculo	82
4.33	Cenário 4: deslocamento em L a esquerda	83
4.34	Cenário 5: deslocamento em L a direita	83
4.35	Cenário 5: deslocamento em L a direita	83
4.36	Análise da resposta do detector	86

Lista de Algoritmos

1	Algoritmo básico do Adaboost	38
2	algoritmo para localização dos pontos característicos numa nova imagem .	46

Lista de Siglas

PPGMCTI ..	Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial
CR	Cadeira de Rodas
CRI	Cadeira de Rodas Inteligente
IBGE	Instituto Brasileiro de Geografia e Estatística
NAVECHAIR	Software para navegação de CR baseado em visão computacional
EOG	Eletroculografia - Método de estudo da função retiniana
EEG	Eletroencefalografia - Método para registrar a atividade elétrica do cérebro
EMG	Eletromiografia Método para medir a atividade elétrica das células musculares
HCI	Interação Homem Computador
HMI	Interface Homem Máquina
ECG	Atividade elétrica do coração
BCI	Interface Cérebro Computador
SVM	Support Vector Machine
TM	Template Matching
PCA	Principal Component Analysis
FACS	Facial Action Coding System
MATLAB ...	Matrix Laboratory
OpenCV	Open Source Computer Vision
IDE	Integrated Development Environment
CPU	Central Processing Unit
SO	Sistema Operacional
RGB	Color System, Red, Green, Blue
AdaBoost ...	Adaptative Boosting
JPG	Joint Photographic Experts Group
ONG's	Organizações Não-Governamentais
AVI	Audio Video Interleave
FPS	Imagens por segundo
ROI	region of interest -Região de interesse
PDM	Point Distribution Model - Modelo de pontos distribuídos
FPS	Frames per second - Quadros por segundos
LK	algoritmo de Lucas-Kanade

Lista de Símbolos

A^T Matriz transposta de A

Introdução

Para muitas pessoas uma cadeira de rodas(CR) representa a sua conexão com o mundo, podendo representar um abrigo passageiro ou permanente de sua vida e com ela se deslocando, trabalhando ou realizando sua socialização diária. Para muitos, se não todos, a forma como a cadeira auxilia nas tarefas faz com que ela se torne uma extensão de si próprio, um espaço pessoal e para alguns completamente incorporada a sua imagem corporal.

As CR passaram de um equipamento com característica de pouco conforto e de difícil dirigibilidade a um patamar de versatilidade, agilidade e segurança. Sua linha temporal destaca avanços constantes e contínuos detre eles a introdução de conceitos da tecnologia assistiva, o desenvolvimento de materiais mais fortes, leves e a incorporação maciça de tecnologia computacional.

No Brasil, segundo o último censo realizado pelo [IBGE \(2011\)](#), existem 24,5 milhões de pessoas com algum tipo de deficiência, seja ela mental, visual, motora, entre outras. Desse total 4.3 milhões, o que equivale a 2,5% da população, possuem severas limitações que impedem o indivíduo de executar tarefas básicas do cotidiano.

As deficiências podem ter origem genética em consequência de doenças, pelo uso de drogas, por abusos ou traumas sofridos além de lesões causadas por acidentes. O [IBGE \(2011\)](#) estima que 7% da população com algum tipo de deficiência não possui qualquer tipo de movimento nos membros inferiores e superiores.

A maioria desse grupo, utiliza uma cadeira de rodas para exercer sua socialização diária além de tarefas comuns como ir à mesa para as refeições ou se deslocar ao sanitário. Esse deslocamento, para a maioria das cadeiras, é realizado pela força humana, ou seja, alguém realiza o deslocamento do indivíduo pelos cômodos da residência ou pelo ambiente externo.

A preocupação da comunidade acadêmica acerca do tema da acessibilidade centra-se numa área de cunho interdisciplinar que se chama de Tecnologia Assistiva(TA) a qual definimos como todo um conjunto de ações que permitem às pessoas com necessidades especiais tornarem suas vidas o mais comum possível, promovendo a independência, autonomia e qualidade de vida do usuário.

Os trabalhos de [Braga \(2012\)](#), [Halawani et al. \(2012\)](#), [Song et al. \(2011\)](#), [Zhao et al.](#)

(2012) indicam a atenção dada em realizar ações tecnológicas de baixo custo, que tornem mais independente a vida cotidiana dessas pessoas. Estes estudos indicam ainda que essas ações promovem ao paciente a oportunidade de realizarem diversas atividades com domínio e independência.

Durante a última década muitos projetos têm computadores embarcados, como o projeto *Smart Chair*, apresentado na Figura 1.2. No entanto, o desenvolvimento de produtos tem enfatizado o conceito de larga produção ignorando o problema mais difícil que é criar produtos de acesso universal. Na Figura 1.1 podemos observar uma ação deste tipo que utiliza a boca do paciente para realizar a navegação. Isso é especialmente verdadeiro em relação a tecnologia de apoio onde interfaces de seleção são muitas vezes inadequadas para pessoas com severas deficiências.



Figura 1.1: Modelo comercial aplicado para pessoas com tetraplegia.
Fonte: (CARBON, 2016)

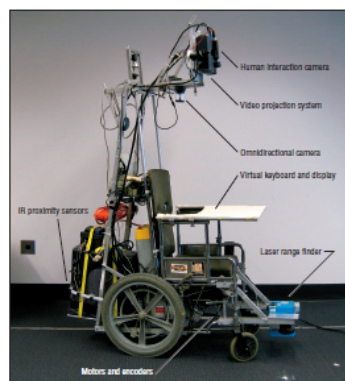


Figura 1.2: Projeto Smart Chair que utiliza uma cadeira altamente sensorizada.
Fonte: (PARIKH *et al.*, 2007)

Segundo Bradski & Keahler (2008), Braga *et al.* (2009), Liu & Zhao (2011), CRs podem ser capazes de entender comandos de voz, locomoção autônoma, desvio de obstáculo, adaptação ao ambiente, suporte a algoritmos de visão computacional entre outras características. Seus movimentos podem ser autônomos, semi-autônomos ou controlado

exclusivamente pelo usuário. Podem possuir um ou mais sensores e quase todas tem um controle convencional (*joystick*), dois destes projetos pode ser observados na [Tabela 1.1](#):

Tabela 1.1: Exemplos de projetos de Cr's

Projeto	Descrição
 <p data-bbox="437 696 799 728">Figura 1.3: Projeto LURCH</p> <p data-bbox="437 757 788 788">Fonte: (BONARINI <i>et al.</i>, 2012)</p>	<p data-bbox="847 488 1230 734">Projeto LURCH/2009 que visa o desenvolvimento de uma cadeira de rodas autônoma capaz de evitar obstáculos, auto localizar-se explorar ambientes internos de uma forma segura</p>
 <p data-bbox="496 1115 745 1178">Figura 1.4: Projeto IntellWhells</p> <p data-bbox="496 1207 745 1238">Fonte: (BRAGA, 2012)</p>	<p data-bbox="847 810 1230 1205">Projeto IntellWhells/2013 possibilita que os utilizadores possam escolher entre vários modos de comando e até combiná-los. Entre as opções já disponíveis existem os comandos de voz, movimentos de cabeça ou o <i>brain computer interface</i>, que permitirá dirigir a cadeira através dos pensamentos.</p>

Dentre o universo das deficiências físicas, as mais severas ao ser humano são a paraplegia e a tetraplegia. Essas enfermidades são complexas tanto do ponto de vista orgânico como social e psicológico. No âmbito orgânico resulta, de acordo com o nível e grau da lesão, em déficit motor sensitivo (superficial e profundo), disfunções vasomotoras, alterações esfinterianas e disfunção sexual. Pode ter origem traumática, por acidente automobilístico, ferimento por arma de fogo, mergulho em águas rasas, quedas ou causas não traumáticas como tumores, malformações, alterações vasculares, processos infecciosos, degenerativos ou compressivos (COOK; HUSSEY, 1995; CRUZ; IOSHIMOTO, 2010; MALLIN, 2004; MELO, 2009).

Na lesão medular são observadas três fases:

1. Choque medular;
2. Retorno da atividade medular;
3. Ajustamento.

Na primeira, ocorre anestesia superficial ou profunda dos membros além de alterações va-

somotoras, atonia da musculatura, arreflexia tendinosa e cutânea, vasodilatação parálitica anidrose, disfunção vesical e intestinal. Na segunda, ocorre a reorganização funcional das estruturas medulares abaixo do nível da lesão e na terceira, o paciente tenta se adaptar a nova condição de vida. Nesse momento, o processo de reabilitação é fundamental. Na tetraplegia, existe o comprometimento dos membros superiores, tronco e membro inferiores, já na paraplegia, ocorre o comprometimento dos membros inferiores e/ou tronco.

Segundo o IBGE (2011) as pessoas portadoras de tetraplegia, em sua maioria, permanecem em casa e na sua cama isoladas do convívio social. As desvantagens em relação a pessoas com paraplegia, por exemplo, são inúmeras como: impossibilidade de manipulação de dispositivos que exijam atuação de ordem física, com comer, escrever ou manipular cadeiras de rodas.

Um CR convencional exige a atuação do usuário para locomoção enquanto a motorizada dispensa esse esforço ao utilizar um motor elétrico, entretanto, há uma pequena atuação física nessas cadeiras, pois o usuário deve comandar a navegação através do *joystick*, localizado em algum ponto do braço da cadeira. Por essa razão, as cadeiras motorizadas não atendem a um portador de tetraplegia. Para Kaplan *et al.* (1989) a qualidade de vida relacionada à saúde refere-se ao impacto sobre três funções: mobilidade, atividade física e social.

Com base nesta realidade propusemos um projeto de cunho interdisciplinar (SmartWheels) que visa dotar uma cadeira de rodas comum com elementos que tornem tanto possível como viável a mobilidade, mesmo que parcial, de indivíduos com lesões na medula que utilizem como controle apenas as expressões faciais focando a detecção nos classificadores em cascata.

1.1 Definição do problema

Diversos modelos de navegação de CR são apresentados na literatura, entre eles destacamos os trabalhos de (BRAGA *et al.*, 2008; DYMOND; POTTER, 1992; BASU; PENTLAND., 1996; ADACHI *et al.*, 1998; KUNO *et al.*, 1999; BERGASA *et al.*, 2000; TAYLOR; NGUYEN, 2003; NGUYEN *et al.*, 2004; KING *et al.*, 2005; MANOGNA *et al.*, 2010; ZHAO *et al.*, 2012). Todos destacaram o uso eficiente de sensores e acoplamentos ao usuário, sendo que muitos ressaltam que, apesar da boa resposta do seu modelo, um ponto fraco é a alta intrusividade do mesmo.

Muitos não são específicos para indivíduos com tetraplegia ou paraplegia dificultando assim adaptação do usuário além de requererem um período de treinamento, tipicamente difíceis de configurar segundo Bailey *et al.* (2007). Nestas circunstâncias, é muito difícil

que nosso público alvo utilize, de modo satisfatório, as CRs propostas.

Outro ponto importante é que os projetos apresentados dependem da aquisição de diversos equipamentos, entre eles câmeras especiais, capacetes para detecção de ondas cerebrais, entre outros, que tornam o custo da cadeira elevado, como analisado em Braga (2012).

Como uma alternativa, Arai & Mardiyanto (2011) propõe o uso de conceitos da visão computacional, pois depende exclusivamente de um tripé: câmera, computador e atuador, visto na Figura 1.5. Estes conceitos norteiam diversos trabalhos, entre eles Perez *et al.* (2012), Halawani *et al.* (2012) e utilizam expressões faciais como forma de interação entre a cadeira e o usuário.

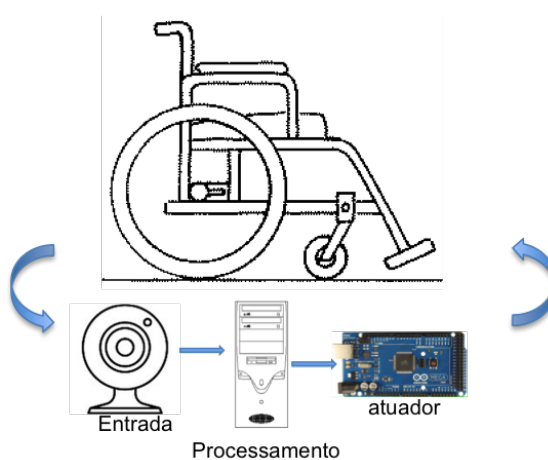


Figura 1.5: Diagrama dos elementos que compõem a base de um sistema de navegação baseado em visão computacional

Fonte: Elaborada pelo autor

Neste contexto, a navegação de uma CR por indivíduos com tetraplegia ou paraplegia, devido as suas limitações físicas, só pode ocorrer, se a extração das informações ocorrerem por expressões faciais. Dentre as diversas expressões faciais descritas em (EKMAN; ROSENBERG, 2005), os modelos de navegação de CR utilizam, em geral, a rotação da cabeça (DYMOND; POTTER, 1992), (BERGASA *et al.*, 2000), abrir e fechar a boca (BONARINI *et al.*, 2012), movimento dos olhos (ARAI; MARDIYANTO, 2011; DUONG *et al.*, 2012).

A Tabela 1.2 apresenta uma lista de projetos para navegação e nela percebemos um grande usos de equipamentos para sensoriamento ou recolha de informações. No caso da recolha de informações nota-se o uso de capacetes, óculos e demais equipamentos considerados intrusivos (FIGUEREDO, 2011). Em geral, necessita de treinamento prévio e não são portáteis.

Tabela 1.2: Projetos de CR e seus componentes

Fonte: Elaborada pelo autor

Autor	Entrada	Imagem	Análise do desempenho
(HALAWANI <i>et al.</i> , 2012)	Duas câmeras, sendo uma acoplada ao usuário; GPS, sensor ultrassom.		96% acerto na navegação indoor.
(PARIKH <i>et al.</i> , 2007)	Câmera omnidirecional, câmera simples, Laser, Sistema de Projeção, GPS.		97% acerto na navegação indoor
(KUNO <i>et al.</i> , 2000.)	2 Câmeras, Sensor ultrassom.		acerto de 96% trajetória.
(REIS <i>et al.</i> , 2009)	4 sonares; joystick; câmeras frontais; 1 controle ps4, microfone e sensores diversos.		98% acerto para navegação e erro máximo de 1% na trajetória.
(BONARINI <i>et al.</i> , 2012)	Duas câmeras; Eletro miographic; sensores diversos.		97% acerto trajetória.
(PEREZ <i>et al.</i> , 2012)	Infravermelho, ultrassom, sensor magnético, laser, acelerômetro encode e câmeras.		98% acerto trajetória.

Face o exposto a pergunta que vem a tona é se é possível criar um modelo computacional de navegação para uma cadeira de rodas com desempenho superior ou igual a 97% na navegação e com erro mínimo de 2% na trajetória, baseado simplesmente na detecção de movimentos da cabeça e dos olhos. A essência desse movimento não deve ser simplesmente a guiagem do veículo, mas o entendimento que, ao longo do percurso, o usuário pode modificar, parar ou voltar ao ponto de partida tomando para si o controle total do equipamento.

1.2 *Objetivo*

De acordo com os pressupostos apresentados anteriormente, os objetivos deste trabalho são:

1.2.1 *Objetivo Geral*

Desenvolver um modelo computacional para o gerenciamento dos movimentos básicos de uma Cadeira de Rodas motorizada utilizando expressões faciais.

1.2.2 *Objetivos Específicos*

1. Estudar e desenvolver uma metodologia de navegação adequada ao contexto do público alvo;
2. Construir um classificador em cascata capaz de identificar com precisão maior que 95% o fechar dos olhos de um paciente;
3. Construir um detector capaz de identificar a rotação da cabeça com acerto igual ou superior a 97%;
4. Tornar o modelo auto-configurável a qualquer usuário, sem a necessidade de treinamento tornando assim a navegação na forma sente-use;
5. Desenvolver um simulador que permita testar a metodologia desenvolvida;
6. Validar o modelo computacional num protótipo físico da CR.

1.3 *Importância da Pesquisa*

Diversos estudos ([MALLIN, 2004](#); [ANDRADE](#); [PEREIRA, 2009](#); [CRUZ](#); [IOSHIMOTO, 2010](#); [VARELA et al., 2013](#)) apontam que os recursos de Tecnologia Assistiva (TA) no Brasil ainda são pouco explorados e expressam a urgência em desenvolver pesquisas que abram possibilidades nesse campo principalmente devido ao aumento populacional em indivíduos que apresentam necessidades desses recursos.

Além disso [Andrade & Pereira \(2009\)](#) e [Cook & Hussey \(1995\)](#) apontam que a qualidade de vida dos pacientes é muito comprometida se esta preocupação não se traduzir em

propostas da academia que possam ser aplicadas pela indústria e vice-versa.

Frente ao panorama aqui apresentado, julgamos relevante o estudo de um modelo que permita a navegação de uma CR por um paciente com tetraplegia que vise permitir um esforço biomecânico mínimo aliado a um bom desempenho na condução. Consideramos especialmente relevante esta condição, pois impacta diretamente na autoestima e na capacidade de interação social, comprometida pelo trauma ou doença acometida ao paciente.

Pretende-se também apresentar à comunidade científica a viabilidade da utilização de recursos mínimos da visão computacional para a navegação da CR com pouca intrusividade.

Assim, a importância desta pesquisa orbita tanto na possibilidade do estudo da convergência e desempenho de algoritmos de visão computacional, apresentando uma alternativa aos projetos de CR existentes, como na apresentação de uma solução que melhore e transforme a qualidade de vida dos pacientes com tetraplegia e sua interação com o meio.

1.4 Motivação

Diversos ambientes da vida cotidiana, sejam eles públicos, privados ou mais especificamente residenciais não estão planejados com o objetivo de auxiliar pessoas com deficiências motoras, no entanto a sociedade vem se preocupando com esta realidade sobretudo na força de leis como:

- Constituição Federal de 5 de outubro de 1988, capítulo *VII* intitulado “Da família, da criança, do adolescente e do idoso”.
- Lei *n*° 8.213 de 24 de julho de 1991 intitulada “Da habilitação e da Reabilitação Profissional”
- Lei *n*° 13.146, de 6 de julho de 2015 intitulada Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência).

Diversas outras leis e ou dispositivos legais, federais, estaduais e municipais vem criando uma rede de obrigações aos legisladores e à sociedade com um todo. Não temos com reproduzir aqui, em sua íntegra, mas as medidas aqui expostas vislumbram o caráter integrador que permeiam as ações. Segundo dados do [IBGE \(2011\)](#) em municípios de até 20 mil habitantes 16.3% são portadores de deficiência e com mais de 500 mil temos 13%. Além disso, 17.1% é negro, 17.5% é indígena e menos de 14% é branco. Sendo que as deficiências físicas acometem 230.000 mil pessoas.

Destacamos ainda que a proporção de pessoas portadoras de deficiência aumenta com a idade prejudicando cerca de 54% com idade superior a 65 anos. Considerando o envelhecimento da estrutura da população brasileira, esse número tende a aumentar e crescer a demanda por estas tecnologias.

A inserção no mercado de trabalho também é extremamente baixa, cerca de 4.1%, haja vista que portadores de tetraplegia não perdem a capacidade mental podendo exercer diversas funções do seu ambiente residencial que contribuem com a sociedade.

Essas informações constataam que, apesar da edição de leis, os portadores de deficiência são, em sua grande maioria, excluídos do convívio social e laboral. De forma geral, a deficiência significa limitar o crescimento pessoal e profissional de maneira violenta, afastando o indivíduo de diversas oportunidades de realização pessoal, afetiva entre outras.

Dentre as deficiências físicas, a paraplegia e a tetraplegia podem ser consideradas as mais severas por imporem ao paciente uma série de impedimentos. Ambas resultam de lesão na medula óssea e tem como consequências, por exemplo, espasticidade, que mantém os músculos ativos através de movimentos involuntários, falha no sistema excretor, aparecimento de escaras devido a falta de atividade muscular entre outras.

Sobre diversos aspectos, o portador de tetraplegia está em desvantagem em relação a uma pessoa em plena capacidade física e as cadeiras de rodas convencionais exigem atuação física do usuário para locomoção. Já as cadeira automatizadas utilizam um *joystick* localizado na ponta de algum braço da cadeira para realizar a navegação. Por esse motivo ambas não atendem a este público.

1.5 Limites e Limitações

O estudo aqui realizado propõe que seja desenvolvido um modelo computacional para a navegação de uma CR por pacientes com tetraplegia/paraplegia. Apesar deste fato, a utilização de pacientes não foi possível e isto se constitui uma limitação do trabalho. A utilização de pacientes iria requerer, inicialmente, o desenvolvimento de um simulador e posteriormente ajustes físicos na CR para implantação do modelo.

A navegação da CR, com o modelo implementado, só é possível em ambientes que exista controle de iluminação e com pisos regulares. Além disso, não foram implementadas ações que parem imediatamente a CR em situações de risco, como por exemplo, passagem de animais a frente da cadeira. As ações podem ser feitas com uma câmera frontal que capte todos os eventos que ocorrem a frente do indivíduo, e esta é uma limitação deste trabalho.

1.6 *Hipóteses*

Os sistemas mecânicos e ergonômicos das cadeiras de rodas com propulsão manual e ou automática possuem baixa ou nenhuma tecnologia que atenda de forma imediata e com baixo custo as necessidades de usabilidade do paciente em suas atividades diárias.

Levando em consideração aspectos de intrusividade, ergonomia, design e baixo custo pretendemos demonstrar a viabilidade das seguintes hipóteses:

- 1^a Hipótese** → É possível se incorporar tecnologias não intrusivas a uma cadeira de rodas comum que permitam a um indivíduo com tetraplegia assumir o controle do seu equipamento;
- 2^a Hipótese** → Ao desenvolver e analisar um novo classificador em cascata com o objetivo de estabelecer o movimento de “seguir em frente” e “parar” acreditamos que este classificador terá um desempenho de detecção da ordem de 97%;
- 3^a Hipótese** → Ao integrar um modelo de distribuição de pontos ao coeficiente de correlação de Pearson com o objetivo de realizar a rotação da cadeira utilizando o movimento da cabeça, acreditamos que a eficiência de detecção desta rotação será de 97%;
- 4^a Hipótese** → O casamento entre os modelo de detecção para olhos fechados e rotação de cabeça terá a mesma eficiência que a navegação realizada pelo controlador manual da cadeira com erro médio de 2% .

1.7 *Aspectos Metodológicos da Pesquisa*

A transcurso metodológico deste trabalho está representado na [Figura 1.6](#). Nela, inicialmente, busca-se analisar um grande número de propostas de CR com alguma forma de interação por visão computacional, entender seus modelos, pressupostos, método de testes, resultados, ergonomia, intrusividade, limitações de cada projeto.

O segundo passo (Observação) determina quais algoritmos de visão computacional podem, de fato, realizar a navegação de uma CR a luz das limitações do paciente com tetraplegia. Identificando de que forma estes algoritmos atuam, quais são seus parâmetros principais e quais ajustes ótimos podem ser aplicados no problema.

Na terceira etapa (PROBLEMA), mergulha-se na implementação computacional do modelo. Em consonância com a etapa anterior, esta implementação irá garantir ou refutar

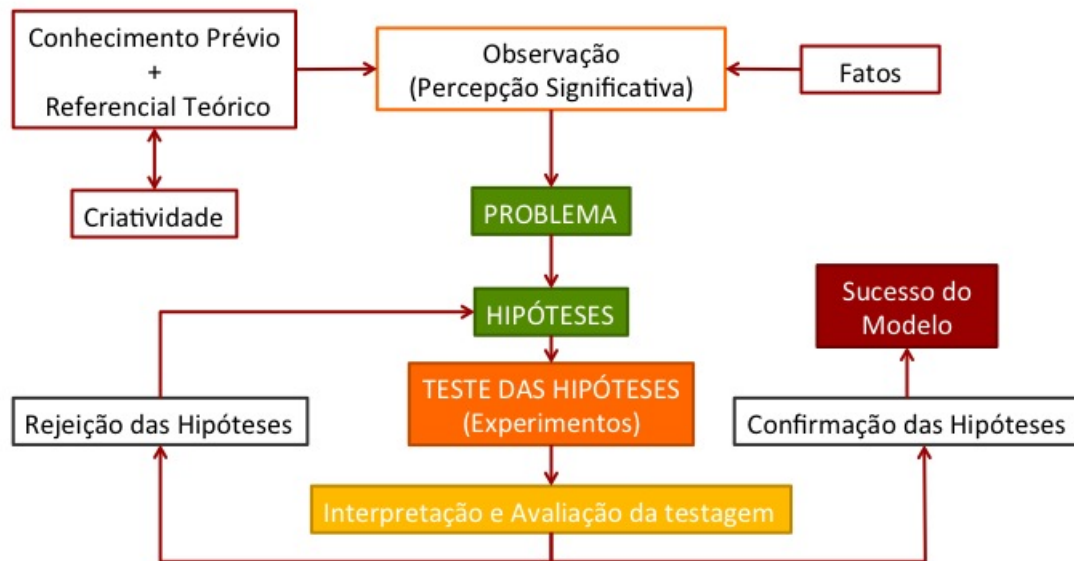


Figura 1.6: Roteiro metodológico

as hipóteses da pesquisa. Após a implementação e análise, o software criado, NaveChair, passará por testes para validar ou não as hipóteses e, conseqüentemente, o modelo.

1.8 Contribuições

A principal contribuição deste trabalho é a proposição de um modelo capaz de realizar a navegação de um CR apenas utilizando um Classificador em cascata e Modelos de Forma Ativa.

A consequência direta desta ação representa outra contribuição pois validado o modelo pode-se, no futuro, propor ações que melhorarem as condições de vida dos pacientes com tetraplegia

Outra contribuição deste trabalho é a análise do enlace entre dois paradigmas distintos de visão computacional, sendo um baseado em aparência e outro em características invariantes, que vão requerer treinamentos diferenciados bem como um conjunto de dados específico.

O conjunto de dados utilizados para treinamento neste trabalho recolheu 10.000 imagens de olhos fechados de indivíduos com diversas etnias. Além disso, foram utilizadas 70 imagens para construção do Modelo de Forma Ativa, 7.000 imagens negativas, 9000 de teste. Este conjunto de dados representam uma outra contribuição deste trabalho, pois em futuras implementações do modelo já se dispõe de um conjunto específico para tal.

Os requisitos de projeto levantados por este trabalho poderão ser diretrizes para desenvolvedores de CR aperfeiçoarem seus produtos e adequarem a realidade de pessoas com severas dificuldades físicas. Inicialmente atendendo às necessidades de locomoção dos pacientes e em um segundo momento, às prescrições médicas de um programa de promoção da saúde.

Além disso o trabalho contribui com:

- Diretrizes e formas para um produto viável comercialmente;
- A integração do Active Shape Models ao coeficiente de correlação de Pearson;
- Determinação de variáveis ideais para um classificador em cascata que vise a detecção de objetos de dimensão reduzida;
- Identificação da influência dos parâmetros de formação dos classificadores bem como a relação entre a complexidade e a eficiência/eficácia do detector.

1.9 Organização da tese

Para apresentação do trabalho proposto, esta tese encontra-se organizada da seguinte forma:

No CAPÍTULO 2 apresenta-se a fundamentação teórica que se divide na revisão da literatura de interesse desta pesquisa, com projetos de CRs que utilizam aspectos tratados no escopo do trabalho, estado da arte da detecção de olhos e detecção do movimento da cabeça. Há, neste capítulo uma análise comparativa ente os projetos atuais e seus pontos fracos e fortes e comparamos com nosso projeto. No CAPÍTULO 3 exibi-se o modelo proposto, sua arquitetura e suas limitações, já no CAPÍTULO 4, expõem-se experimentos integrando os módulos tanto na cadeira física como no universo computacional Nesse capítulo os resultados são comentados e obtêm-se as primeiras conclusões. Finalmente no CAPÍTULO 5 são feitas as conclusões e a análise das hipóteses bem como comentários gerais sobre o trabalho e a tecitura de novos objetivos para futuros trabalhos.

Fundamentação Teórica

2.1 Introdução

O objetivo deste capítulo é apresentar os principais trabalhos de investigação desenvolvidos no âmbito do escopo deste trabalho que venham a fundamentar o desenvolvimento do modelo computacional para navegação da CR. Neste capítulo retratar-se-a, também um resumo comparativo entre estes projetos e suas características. Inicialmente expomos alguns projetos de CR para depois focar em trabalhos ligados à detecção do estado dos olhos e na detecção do movimento da cabeça. Apresentamos um subcapítulo no qual fazemos uma análise comparativa entre os modelos de CR e, por fim, nossas impressões sobre os trabalhos estudados e as consequências, influências diretas e indiretas, contribuições e subsídios para o trabalho.

2.2 Estado da Arte

A mobilidade é um importante meio de interação social, lazer, atividade física e atividades laborais. Comum e corriqueiro para uma pessoa normal a mobilidade torna-se, em certos momentos, quase impossível a um indivíduo com severas deficiências. Ao compensar esta carência com tecnologia, torna-se possível uma ampla gama de atividades e, por consequência, uma melhora significativa na qualidade de vida do indivíduo. A mobilidade pode ser auxiliada por diversos meios como: bengalas, andadores, carrinho e cadeiras de rodas sejam elas manuais ou elétrica. Diversos projetos de CR propõem as mais diversas abordagens que visam dotar o paciente de um certo grau de mobilidade. Nas sessões a seguir apresentamos alguns destes projetos.

2.2.1 Projetos de CRs

Existem mais de 35 projetos de CRs espalhados pelo mundo ([BRAGA, 2012](#)). Esse tema tem despertado o interesse por diversos fatores, entre eles o fato de que a Organização Mundial de Saúde (OMS) estimar que cerca de 2% da população mundial, ou aproximadamente de 130 milhões de pessoas, vive com deficiências físicas. Esse número crescente é devido à problemas físicos, congênitos, doenças, acidentes, guerras, e do envelhecimento da população como consequência do aumento da expectativa de vida. No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) estima que 4% da população brasileira

tem algum tipo de deficiência.

Todos esses projetos variam existentes em diversos aspectos. Centraremos nosso olhar sobre os projetos que versam dentro do universo de pessoas com tetraplegia ou qualquer tipo de deficiência que impeça os movimentos do corpo.

Inicialmente destacamos o projeto IntellWheels (BRAGA *et al.*, 2008; BRAGA *et al.*, 2009), que tem como objetivo principal a criação de uma plataforma de desenvolvimento para cadeiras de rodas inteligentes, facilitando o design, teste de novas metodologias e de técnicas relativas às CRIs (BRAGA, 2012). Esta plataforma leva em consideração aspectos de baixo custo, conforto e ergonomia. O projeto dispõe de um modelo físico de uma cadeira que utiliza diversos agentes como comandos por voz, visão computacional e mapeamento e localização. Permite que possam ser realizados estudos da eficiência de qualquer modelo a ser implementado para uma CRI, utilizando seus diversos agentes.

No trabalho de Dymond & Potter (1992) os autores propõem uma cadeira que se vale dos movimentos da cabeça para realizar a navegação utilizando um sistema de vídeo. Estes movimentos são extremamente naturais para todos e essencialmente o único para pessoas com tetraplegia ou quadriplegia. A ideia do trabalho é tentar mapear o desejo do usuário e realizá-lo sem a necessidade de um movimento totalmente guiado pela cabeça. Utiliza um capacete com sensores e mapeia 10 movimentos básicos garantindo um erro de menos de 3%. Estuda também a variação total do ângulo que a cabeça do paciente pode girar o que fica em torno de 69° a 233° horizontal e 10° na vertical.

Seguindo o conceito de desejo do usuário, Adachi *et al.* (1998) tentam também predizê-lo utilizando apenas as informações da face por meio de detecção das expressões faciais, aplicando filtros de suavização e tomando por base 10 pontos característicos ao redor dos olhos, boca e maçãs da face e por diferença entre os quadros. Isso permite indicar a localização desses pontos e assim determinar a direção que o usuário deseja guiar a cadeira. Os testes realizados foram numa CRI física, propondo também uma maior utilização de sensores integrados ao modelo para aumentar a acurácia dos dados.

Kuno *et al.* (1999) utilizam os gestos faciais integrados a uma certa inteligência da cadeira para gerar o movimento, tentando assim compreender as intenções do usuário, determinando o melhor modo de navegação, o que deixa, na verdade, o controle do movimento da cadeira inteiramente autônomo, tendo o usuário apenas a necessidade de indicar a direção, reduzindo assim a operação por parte do mesmo. A cadeira também dispõe de um reconhecedor facial para localizar e se ajustar à melhor posição para que o usuário sente na cadeira, facilitando assim a operação de sair e sentar na cadeira. Começa, mas não aprofunda pontos considerados anteriormente como:

- movimentos involuntários durante a direção;
- movimentos rápidos durante a direção;
- desatenção ao realizar a direção.

Tendo como base um sistema de geração de comandos por movimentos faciais e utilizando um rastreador 2D e um detector de cor além de um detector difuso, [Bergasa et al. \(2000\)](#) trabalham com uma arquitetura específica para tentar definir comandos que gerem movimento para guiar uma CRI:

- Giro da cabeça na horizontal deslocamento direita, esquerda;
- Giro da cabeça na vertical aumentar diminuir a velocidade;
- Manter o olho fechado ou aberto função desliga e liga a cadeira.

Os autores não levam em consideração os aspectos tratados em [Kuno et al. \(1999\)](#), mas destacam que o modelo revela-se pouco intrusivo, robusto e com pouca influência de aspectos de iluminação. Destaca que o modelo necessita de um largo treinamento do usuário e que a câmera deve ficar em torno de 80cm do usuário e que o ambiente testado foi controlado.

No trabalho de [Barea et al. \(2002\)](#) aplica-se um modelo de navegação baseado em EOG, na qual foi definido como uma técnica para medir o movimento ocular, através da análise do potencial de ação gerado por um dipolo elétrico formado entre a córnea e a retina, através de repolarizações e despolarizações (trocas iônicas na membrana celular) das células da retina e da córnea. O sistema consiste em uma cadeira de rodas elétrica padrão com um computador de bordo, sensores e uma interface de usuário gráfica executada pelo computador que apresenta as respostas obtidas do usuário em tempo real e navega a cadeira. Esta técnica também foi usada por ([CHACKO et al., 2013](#))

Em [Taylor & Nguyen \(2003\)](#) e [Nguyen et al. \(2004\)](#) retomam-se as ideias de sensores colocados na cabeça do usuário para detecção do movimento em um sistema desenvolvido para plataforma livre. Com uma implementação mais simples, discute alguns comandos. Depende fortemente da integridade do conjunto de treino e da topologia da rede neural escolhida, já que o modelo se baseia em redes neurais para o aprendizado e determinação do movimento da cadeira. Utiliza a tecnologia DIMM-PC para controle da cadeira e não testou em pacientes. Este mesmo conceito foi revisto em [King et al. \(2005\)](#) e tendo o modelo obtido índices de acerto entre 98% e 99%, fazendo uso de um acelerômetro.

Trabalhos como o de [Urbano et al. \(2008\)](#), desconsideram o uso de visão computacional e se contrapõem aos modelos estudados até então, por que propõe o uso do *joystick*

e baseia-se numa arquitetura que permite utilizar pequenos gestos que são readquiridos pelo paciente após um processo fisioterápico. Os movimentos são detectados por sensores que identificam para onde o usuário deseja direcionar o controle da CR. O trabalho adapta uma cadeira comercial e demonstra resultados positivos, mas os paciente estudados demonstraram cansaço após um certo tempo de operação da CRI.

([MANOGNA *et al.*, 2010](#)) criaram um sistema para controlar diretamente a rotação do motor com base no movimento da cabeça. O dispositivo é fixado na cabeça do paciente e, com base em sensores EMG, produz movimento da CRI. O acelerômetro percebe o movimento e o decodifica para 4 movimentos básicos, necessitando de auxílio de comandos por voz e ultra-som para evitar colisões. O mesmo tipo de estudo foi realizado por [Wei *et al.* \(2010\)](#) que avaliaram um equipamento similar em um ambiente fechado. [Wei *et al.* \(2010\)](#) sugere o uso do mesmo tipo de sensor, mas segue 3 etapas: 1. EMG e aquisição da imagem, 2. Análise do ambiente pela cadeira e 3. Interface homem-máquina que interpreta os comandos.

Os estudos de [Manogna *et al.* \(2010\)](#), [Wei *et al.* \(2010\)](#) foram atualizados com o estudo de [Lee & Samaras \(2011\)](#) que detecta o abrir e fechar da boca para movimentos extras e estuda o problema da ergonomia da cabeça. Define angulações mínimas e máximas para o indivíduo e o tempo total de uso da cadeira sem cansar o paciente.

[Zhao *et al.* \(2012\)](#) descrevem o algoritmo de Lucas-Kanade(LK) para detecção do movimento facial e discute a eficiência do método, sua precisão e tempo de reação. Demonstra sua robustez, com taxa de reconhecimento satisfatória. A fraqueza do modelo está nas mudanças abruptas da cabeça. Destacamos que este método já foi discutido em [Liu & Zhao \(2011\)](#). Em ambos os casos foi obtido uma acurácia de 97% e 96% respectivamente.

[Jain & Joshi \(2014\)](#) Controla a cadeira por movimentos da língua utilizando um conjunto de sensores magnéticos montados em um molde bucal que mede o campo magnético gerado por um pequeno ímã permanentemente fixado na língua. Os sinais do sensor são transmitidos através de uma ligação sem fios e processados para controlar a cadeira de rodas motorizada. Como este projeto é bastante intrusivo foi proposto por [Porthiyas *et al.* \(2014\)](#) um modelo que se utiliza o sopro do paciente para guiá-lo de um ambiente para outro em um hospital. A cadeira tem diversos sensores e atuadores que levam o paciente de um cômodo a outro sem qualquer interferência do paciente.

2.3 Análise comparativa dos Projetos de CR

Segundo [Braga \(2012\)](#), existem mais de 400 projetos espalhados pelo mundo que buscam dotar CR de algum tipo de inteligência. Estes projetos, conforme analisamos anteri-

ormente, são os mais diversificados possíveis e utilizam isoladamente, ou em conjunto, conceitos de Interação Homem Computador (HCI) ou Interface Homem Máquina (HMI).

Todas as interações que ocorrem por sinais biológicos são definidas com HCI e geralmente utilizam equipamentos acoplados ao usuário, como sensores de ondas cerebrais, oculares e de sinais elétricos do corpo. Podem utilizar um tela de retorno ao usuário das informações ou não e ainda podem se caracterizar pelo uso exclusivo ou não de acelerômetros são acoplados diretamente no usuário, ou em acessórios.

Já os modelos que possuem algum tipo de sinal não biológico são inseridos no universo da HMI, que trazem características do modelos HCI, mas sua forma de entrada difere, apesar de alguns modelos utilizarem sensores que são instalados na língua do paciente. As abordagens analisadas são todas focadas em pacientes que não possuem mobilidade nos membros inferiores e superiores. A [Tabela 2.1](#) apresenta uma visão geral dos modelos apresentados:

Tabela 2.1: Projetos de CR e suas características
 Fonte: Elaborada pelo autor

AUTOR	ENTRADAS											MODELOS								
	Sonar	Infravermelho	Scanner laser	Câmera	Sensor contato/colisão	Odometria	Bussola	GPS	Sensor acelerômetro	Giroscópios	Microfone	Joystick/Game Control	Touch screen	HCI			HMI			
														EOG	EMG	EEG	Acelerômetros	Movimento da Cabeça	Movimento dos olhos	Bio-Sinal
(DYMOND; POTTER, 1992)				✓				✓			✓						✓			
(ADACHI <i>et al.</i> , 1998)	✓	✓						✓			✓						✓	✓		
(KUNO <i>et al.</i> , 2000.)	✓			✓				✓									✓	✓	✓	✓
(BAREA <i>et al.</i> , 2002)				✓							✓	✓		✓			✓		✓	✓
(TAYLOR; NGUYEN, 2003)	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓					✓			
(PARIKH <i>et al.</i> , 2007)	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓				✓		✓	✓
(URBANO <i>et al.</i> , 2008)	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓							✓	
(MANOGNA <i>et al.</i> , 2010)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓		✓			
(LEE; SAMARAS, 2011)	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓				✓	✓		✓	
(BRAGA, 2012)	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓								✓
(HALAWANI <i>et al.</i> , 2012)	✓			✓				✓	✓											✓
(BONARINI <i>et al.</i> , 2012)	✓	✓		✓												✓				
(PEREZ <i>et al.</i> , 2012)	✓	✓		✓				✓	✓			✓					✓			
(CHACKO <i>et al.</i> , 2013)	✓									✓		✓	✓					✓		
(JAIN; JOSHI, 2014)																			✓	
(PORTHIYAS <i>et al.</i> , 2014)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓							✓	

Nos modelos baseados em HCI, os paciente assumem o controle da sua CR por meio da comunicação entre sua atividade cerebral (EEG), batimentos cardíacos (ECG), atividade muscular (EMG) e movimento ocular (EOG) individualmente ou coletivamente com um computador que envia os sinais para a cadeira que os interpreta e aciona os comandos.

Conforme descrito em (BONARINI *et al.*, 2012) e (PEREZ *et al.*, 2012) o eletrocefalograma (EEG) registra os sinais elétricos cerebrais do couro cabeludo, onde o sinal cerebral

origina-se de potenciais pós-sinápticos, agregados no córtex, e transfere através do crânio para o couro cabeludo. Um BCI é responsável pela extração dos dados e convertê-los em comandos de controle da cadeira usando processamento de sinais. As técnicas EEG são de baixa intrusividade e custo, mas tem grandes desafios no processamento dos sinais e reconhecimento de padrões, uma vez que tem relação sinal-ruído ruim e uma resolução limitada para determinadas faixas de frequência(ONER; HU, 2013)

O EMG mede as correntes elétricas que são geradas nos músculos durante a sua contração, observado se a soma de todos os potenciais de ação que ocorrem ao redor do local onde o eletrodo é posicionado. Em quase todos os casos, a contração muscular provoca um aumento da amplitude global do EMG. Os sinais EMG podem ser usados para uma variedade de aplicações, incluindo clínicas, HCI e jogos de computador.

Eles são fáceis de adquirir e de acurácia relativamente alta que outros bio-sinais (MANNING *et al.*, 2010). Por outro lado, os sinais EMG são facilmente susceptíveis ao ruído causados por radiação eletromagnética, artefatos de movimento e a interação de diferentes elementos, sendo necessário um pré-processamento que filtre ruídos indesejados. Os sinais de EMG também têm assinaturas diferentes dependendo da idade, desenvolvimento muscular, caminhos de unidade motora, camada de gordura da pele e estilos de gesto. As aparências externas dos gestos de dois indivíduos podem parecer idênticas, mas os sinais EMG característicos são diferentes.

A Eletroculografia (EOG) é uma técnica utilizada para medir o movimento ocular voluntário ou reflexos. Ela registra a diferença de potencial elétrico entre a retina e a córnea em função do tempo. O sinal é captado por eletrodos colocados ao redor do olho e em outros pontos estratégicos que registram a diferença de potencial entre a retina e a córnea(BAREA *et al.*, 2002). Os modelos baseados em EOG possuem bom desempenho alta acurácia(CHACKO *et al.*, 2013).

CRs que se baseiam em acelerômetros utilizam um sensor, em si, ou aparelhos/controles de games que dispõe deste recurso para captar a inclinação, em particular o movimento da cabeça. Quando é detectada uma inclinação os valores são registrados pelo sensor e enviados ao computador ou diretamente ao micro-controlador da cadeira. Dependendo da direção o computador/micro-controlador interpreta a informação e aplica o comando como “vá para esquerda”. Apesar de ser apenas um sensor, a utilização continua o elevou a o status de modelo que pode trabalhar sozinho ou em conjunto em geral possui boa resposta (KUNO *et al.*, 1999).

Modelos que utilizam Bio-Sinal retiram informações da face do indivíduo quase sempre de maneira invasiva, como o movimento da língua (JAIN; JOSHI, 2014) e os por controle de voz usam um microfone e comandos treinados em um reconhecedor de expressões

(BRAGA, 2012)

Finalmente temos os modelos que se baseiam no movimento da cabeça ou dos olhos. Estes apresentam diversas técnicas de captação da informação, entre elas: acelerômetros, sensores de deslocamento, simples webCams, câmeras térmicas, 3D, com visão noturna, omnidirecionais, sensores EOG ou EMG. Após a captação da informação é feito o tratamento dela que, em geral, ocorre num computador que envia as informações ao controlador.

Nosso modelo se inclui num paradigma HMI, com subcategoria Movimentos dos Olhos e da Cabeça utilizando como entrada uma câmera. Já que desejamos criar um modelo computacional de navegação para uma cadeira de rodas com acurácia superior ou igual a 97% na navegação, com erro mínimo de 2% na trajetória, baseado simplesmente na detecção de movimentos da cabeça e dos olhos.

Separadamente ou em conjunto estes modelos já demonstraram sua acurácia, (DYMOND; POTTER, 1992; ALGHOWINEM *et al.*, 2013; ALSHAQAQI *et al.*, 2013; ARAI; MARDIYANTO, 2011; BATUR; HAYES, 2005; BEYMER, 1994; BAZIN *et al.*, 2009; BIRKBECK; JAGERSAND, 2004), em problemas específicos de reconhecimento de padrões ou em trabalhos relacionados a CRs. Existem diversas formas de tratamento do sinal recebido pela câmera na próxima seção, discutiremos a qual a forma ideal para este tratamento tanto da detecção de olhos como o movimento da cabeça.

2.3.1 Detecção do estado dos olhos

A detecção dos olhos é um tema de pesquisa na área do reconhecimento de padrões que tem por objetivo central encontrar elementos do mundo físico, tais como rostos, lápis, canetas, dentre outras, em imagens digitais. Comumente este processo é utilizado em diversas tarefas como a análise clínica, segurança, trânsito, traumatologia, reabilitação entre outras.

Embora existam muitos algoritmos diferentes para executar a detecção dos olhos, cada um tem suas próprias fraquezas e forças. Alguns usam tons de pele, outros contornos e outros são ainda mais complexos, envolvendo técnicas como redes neurais e Wavelets. Estes algoritmos sofrem do mesmo problema: alta taxa de falsos positivos, alto custo computacional e, considerando uma imagem como uma coleção de intensidades de luz, a difícil e demorada análise das imagens obtidas.

Os falsos positivos representam um dos “calcanhares de Aquiles” para os algoritmos de detecção, em particular, a detecção de olhos, pois, acrescenta um grau de dificuldade a mais devido a sua natureza dinâmica. Alguns desses impedimentos podem ser tratados

com técnicas de processamento e outros ainda não. Modelos menos susceptíveis a falsos positivos podem ser simplesmente chamados de robusto e a busca por um modelo dessa natureza é alvo deste trabalho. Podemos ainda expandir a ideia de robustez se conseguirmos superar os seguintes problemas:

1. A variância de escala: a mesma face pode ser apresentada para o sistema de diferentes escalas, ou seja, a distância entre o rosto e a câmera influencia o reconhecimento;
2. A variação de posição: a mesma face pode ser captada sob diferentes perspectivas e orientações, ou seja, de perfil, frontal, com rotações e translações;
3. Variação de iluminação: uma mesma imagem da face captada sob diferentes condições de iluminação tais como posição e força da fonte de luz pode causar ambiguidades;
4. Expressões emocionais e detalhes faciais: sorrir, fazer caretas, barbas e bigodes, usar óculos, entre outras características podem interferir na operação;
5. Ruídos gerados por granulações na imagem, defeitos de hardware, oclusão da câmera e objetos que venham interferir na precisão da imagem.

Os problemas expostos acima não contemplam todas as limitações que vem a surgir no processo de detecção, pois outros entraves acontecem e somam-se a eles como falhas de hardware e software. Estes tópicos revelaram-se presentes na maioria das pesquisas dentre as quais podemos destacar as de ([ALSHAQAQI *et al.*, 2013](#); [BUDZAN](#); [WYÅijGOLIK, 2013](#); [GONG](#); [XIANG, 2011](#); [HALAWANI *et al.*, 2012](#)).

Segundo [Jenkins \(2008\)](#) os métodos baseados em aparência são os mais eficientes para detecção de características e expressões faciais devido a sua natureza estatística e de comparação direta mais próxima da experiência humana na detecção de objetos, algo que foi analisado também em [Figueredo \(2011\)](#), além de se revelar menos susceptíveis a falha de invariância de escala e complexidade do mundo.

Dentre os diversos métodos de detecção baseados na aparência, notadamente, o modelo de [Viola & Jones \(2001\)](#) e o que obtêm resultados mais satisfatórios atualmente e é amplamente utilizado pelo seu grau de robustez. Segundo [Schmidt & KasiÅĐski \(2007\)](#), o modelo se utiliza de um conjunto de positivos que são treinados por meio de pesos e, dessa forma, procura compreender numa imagem de entrada a correspondência de luz similar, ou seja, localizar na imagem objetos similares ao treinado.

Segundo [Jolliffe \(2002\)](#) o modelo de [Viola & Jones \(2001\)](#), funciona muito bem na detecção facial e já se tornou padrão por sua alta taxa de acerto e baixa taxa de falsos positivos.

É adaptativo e amplamente utilizado pelo seu grau de robustez e rapidez. Em nossa implementação a detecção é baseada em uma cascata de classificadores de Haar (*Haar Cascade Classifier-HCC*), com parâmetros avaliados e um banco de positivos com cerca de 10000 imagens frontais de olhos fechados cuja parcela de imagens é representada na [Figura 2.1](#).



Figura 2.1: Imagens utilizadas para treinamento do modelo

Fonte: Elaborada pelo autor

O controle do estado da CR, baseado na abertura e fechamento dos olhos, é um método pouco utilizado para controle das mesmas, conforme o estado da arte analisado, mas largamente aplicado em pesquisas ligadas a fadiga de motoristas para evitar acidentes de trânsito.

[Eriksson & Papanikotopoulos \(1997\)](#), [Eriksson & Papanikolopoulos \(2001\)](#) descreve um sistema que localiza e rastreia os olhos por operações algébricas sobre a imagem da face extraída pelo método de [Wei *et al.* \(2010\)](#), o que notadamente é utilizado também em ([LIN *et al.*, 2002](#); [JI; YANG, 2002](#); [CHAU; BETKE, 2005](#); [ARAI; MARDIYANTO, 2011](#); [FAZLI; ESFEHANI, 2012](#); [ALSHAQAQI *et al.*, 2013](#); [SAINI; SAINI, 2014](#)). Todos os olhos são localizados a partir da detecção facial que diminui o espaço de busca e facilita de forma significativa a tarefa de detecção, para em seguida realizar a detecção. Alguns utilizam filtros que melhoram a condição de detecção.

Percebe-se que a evolução da utilização de visão computacional para movimentação de CRs foi gradual. Também foi quase constante a utilização de acelerômetros ou câmeras na cabeça do indivíduo, que torna o equipamento bastante intrusivo e faz necessário treinamento específico para o uso do hardware e, portanto, de difícil adaptação. As taxas de acerto sempre oscilam sempre acima de 90%, porém destes citados, nenhum se propõe a analisar a influência dos parâmetros de criação do detector. A sua maioria não dispõe de banco de dados específico para o fim e em geral utiliza do mesmo banco padrão disponível na web.

A [Tabela 2.2](#) apresenta os trabalhos relevantes no estudo de detecção do estado dos olhos, em geral muito estudado na análise da fadiga de motoristas.

Os métodos baseados no conhecimento utilizam o conhecimento humano sobre o que se quer detectar e tenta descrever estes aspectos computacionalmente de tal forma que se possa localizar numa imagem tais características. Os métodos baseados em características invariantes buscam encontrar características estruturais que existem mesmo quando a condição pose, de ponto de vista, ou de iluminação variam e, em seguida, usá-los para

Tabela 2.2: Alguns trabalhos relevantes na área de detecção de objetos

Abordagem baseada em:	Trabalhos
Conhecimento	Salehi <i>et al.</i> (2012), Sebari & He (2013), Rautaray & Agrawal (2015)
Características invariantes	Ghimire & Lee (2013), Cootes <i>et al.</i> (1995), Toth & Madabhushi (2012)
Templetes	Cootes <i>et al.</i> (2001), Brunelli & Poggio (1993), Kossafi <i>et al.</i> (2014)
Aparência	Jia <i>et al.</i> (2012), Viola & Jones (2001), Guo <i>et al.</i> (2012)

localização.

Nos métodos baseados em “templates” vários padrões são armazenados para que, posteriormente, sejam capazes de descrever as novas imagens de entrada. Geralmente a correlação entre eles é gerada a partir do cálculo da distância entre os pontos.

Nos métodos baseados na aparência, o algoritmo “aprende” o que vem a ser o objeto a ser detectado por meio de um conjunto de treinamento que deve representar a total variabilidade das informações. Este modelo se contrapõe aos anteriores no seu grau de velocidade e aplicabilidade.

Alguns trabalhos utilizam técnicas mescladas, ou seja, misturam todos os já discutidos e obtêm resultados similares como os de (FAZLI; ESFEHANI, 2012; ALSHAQAQI *et al.*, 2013; SAINI; SAINI, 2014; ERIKSSON; PAPANIKOTOPOULOS, 1997; ERIKSSON; PAPANIKOLOPOULOS, 2001).

Outros propõem a utilização de hardware específico para obter uma taxa de acerto e controle acima de 98%, destacamos o de Lin *et al.* (2002) que utiliza um hardware a 6cm do olho para detecção da pupila. Neste caso o equipamento é um óculos, também proposto por Ji & Yang (2002) com hardware diferente.

Nesta linha, os esforços centram-se em métodos com o mínimo de intrusão como os de Arai & Mardiyanto (2011), Zhua & Jib (2005) que utilizam o filtro de Kalman para rastreamento do olho enquanto e Budzan & WyÅijgolik (2013) apresenta um algoritmo para a detecção e localização do rosto e dos olhos por imagens térmicas. O algoritmo utilizam métodos baseados no conhecimento e operações morfológicas, particularmente uma modificação da transformada de Hough.

Para este estudo, adotamos a abordagem baseada exclusivamente em algoritmos de visão computacional sem a necessidade de hardware específico, ou seja, apenas utilizando uma câmera simples frontal ao indivíduo.

2.3.2 Detecção do Movimento da Cabeça

A capacidade de entendimento do movimento de uma cabeça humana é trivial ao olhar humano, pois permite inferir as intenções de outras pessoas que estão nas proximidades e compreender uma forma não-verbal importante de comunicação. Também pode ser considerado uma forma natural, simples e eficaz de apontar para um objeto. Assim, a detecção do movimento da cabeça permite uma boa fonte de informações para o processo de HMI, controlando a ação de alguns dispositivos ou sinais de controle.

Do ponto de vista computacional, mais precisamente em visão computacional, estimar a “pose” é o processo de inferir a orientação de uma cabeça humana a partir de imagens digitais.

Muitos processos foram propostos por diversos pesquisadores para a realização desta tarefa. Procuraram tratar a representação da imagem, baseada em pixel, numa informação que traduzisse o movimento da cabeça e, como outras etapas do processamento imagens, demonstrando um boa tolerância à invariância inerente ao estudo de imagens.

Uma gama cada vez mais crescente de tecnologias tem repousado seus esforços no aprimoramento e criação de novos algoritmos capazes de realizar a detecção do movimento da cabeça. Isso se dá principalmente pela crescente demanda de tecnologias voltadas à pessoas com deficiência, aplicações de realidade virtual, teleconferências

Durante a última década diversos trabalhos tem sido propostos para detectar o movimento da cabeça, seja para aplicações comerciais ou acadêmicas. Todas as abordagens utilizam um método específico ou uma integração entre vários métodos. Necessariamente, a análise destes permite dizer que o problema de pesquisa, seja ele comercial ou acadêmico, interpõe qual método é mais ou menos eficiente.

O estudo de [Basu & Pentland. \(1996\)](#) determina que a construção de um modelo 3D elipsoidal da cabeça permitindo a interpretação do fluxo desse movimento estabelece uma metodologia mais eficaz que a abordagem 2D. O trabalho contribui com a regularização desse fluxo destacando que este modelo é fixo porém estável. Isso significa que o modelo ignora outras informações da imagem considerando apenas a região da cabeça, podendo ser estendido a modelos não elipsoidais.

Considerando um filtro de Kalman para prever ou prever o possível movimento da cabeça [Kiruluta et al. \(1997\)](#) procura diminuir a ação do movimento deste membro supondo que, com poucos movimentos, a cadeira já pode seguir uma rota estimada por esta posição, diminuindo assim o esforço do usuário. Sugere que a predição pode ajudar inclusive no processamento pois não seria necessário utilizar todos os frames do movimento e sim

uma análise da velocidade em torno da ação somado à predição já estudada. Utiliza um capacete para realizar a recolha dos dados do usuário e predizer o movimento.

Liao & Cohen (2005) aborda o reconhecimento de gestos faciais em ambientes interativos e se propõe a reconhecer os gestos por meio de pontos de interesse na face do indivíduo. Ele inicialmente utiliza o *Active Shape Models* para extrair características da face e, em seguida, monta um modelo em 3D que busca as características baseados no *Facial Action Coding System* FACS(EKMAN; ROSENBERG, 2005).

Liu *et al.* (2012) apresenta uma técnica na qual estima a pose em relação aos quadros adjacentes e subsequentes. Os pesquisadores utilizam o algoritmo *Scale Invariant Characteristic Transform*-(SIFT) com descritores que são usados para combinar os pontos característicos correspondentes entre duas imagens adjacentes. Depois de combinar os pontos característicos correspondentes, o ângulo da cabeça é encontrado usando a geometria espacial.

Murphy-Chutorian & Trivedi (2010) apresentam um sistema para detecção do movimento da cabeça voltado para a distração e desatenção dos motoristas. O procedimento para a estimativa da pose utiliza um algoritmo visual 3-D. O sistema é composto por três módulos interconectados que detectam a cabeça do indivíduo, fornecendo estimativas iniciais da pose e, continuamente, rastreando a sua posição e orientação em seis graus de liberdade. O módulo de detecção da cabeça é constituído por um conjunto de classificadores em cascata usando Haar-wavelet.

Kupetz *et al.* (2010), propõe que seu sistema de controle traduza a posição da cabeça do usuário em velocidade e controle direcional para uma cadeira de rodas. O movimento da cabeça foi medido utilizando uma câmera infravermelha colocada na parte de trás da cabeça do usuário. O sistema de controle inclui um modo de espera que é ativado pressionando a cabeça para trás contra o encosto, que ativa o sistema de travagem e sua desativação.

Sriteerakul *et al.* (2011) propõem um método para controle das alterações da pose da cabeça pela textura em um espaço de orientação. Em outras palavras, o estado atual da nova pose pode ser rastreado pelo emprego de uma comparação entre locais binários e a sua textura. Várias texturas são estimadas a partir da rotação da cabeça anterior com vários ângulos conhecidos. Assim, a pose anterior pode ser considerada como um estimador. Para aumentar a precisão de rastreamento, vários avaliadores, cada um com diferentes posições da cabeça, são implantados para produzir um resultado combinado. O método funciona bem em sequências de imagens com diferentes condições de luz.

O modelo de Song *et al.* (2011) é baseado no algoritmo de detecção AdaBoost. Os pes-

quisadores propõem um algoritmo de detecção de movimento da cabeça que não depende da identificação biométrica específica e de acompanhamento. Ele usa método de classificação recursiva para detectar ações de abertura e fechamento da boca como também os movimentos da cabeça. As combinações da pose da cabeça associados aos da boca são mapeados e utilizados para realizar eventos de mouse como mover, arraste, clicar e assim por diante.

[Liu & Zhao \(2011\)](#) utilizam o LK para determinar o padrão dos movimentos da cabeça, realizando o treinamento dos classificadores com uma variação do AdaBoost chamado de GentleBoost. Utilizou-se também a diferença de coordenadas da narina em uma imagem ou um quadro para identificar o movimento da cabeça.

[Zhao & Yan \(2011\)](#) propõem um sistema de detecção baseado em rede neurais. Os faces são detectadas utilizando o método de detecção da pele YCbCr, e então rotula-se k pontos-chave manualmente que apresentam diferentes orientações. As coordenadas e texturas locais destes k pontos-chave são obtidos para compor os vetores de características de entrada das redes neurais. Ao treinar 1.300 conjuntos de dados, os resultados mostraram que o método baseado em rede neural pode estimar a orientação da cabeça na velocidade correta de 90%.

[Qin & Shelton \(2016\)](#) apresentam o casamento de duas importantes tarefas de visão computacional: rastreamento multi-alvo e estimação da pose em vídeos de vigilância. Três componentes são modelados em uma formulação probabilística, fornecendo soluções eficazes para mitigar ambiguidades visuais no rastreamento multi-câmera e da pose, usando uma única câmera de monitoramento.

[Liu et al. \(2016\)](#) propõem um método para estimar a pose da cabeça com rede neurais. O método é tratado com um problema de regressão. Uma rede neural é treinada para aprender características de cabeça e resolver este problema de regressão. Para fornecer os dados da posição da cabeça, foi gerado um conjunto de dados, com informações acerca de diversas informações como variação do sexo, idade, raça e expressão. O conjunto de dados inclui 74000 imagens de poses a partir de 37 modelos.

2.4 Análise comparativa dos métodos

Um número muito grande de técnicas podem ser utilizadas para a detecção tanto do estado dos olhos como do movimento da cabeça. Podem ser empregados métodos como o PCA, Redes Neurais, SVM, Filtro de Kalman, TM entre outros. Podemos ainda utilizar a mistura de diferentes técnicas, que somadas, produzem um respostas efetiva ao problema estudado.

Cada um dos trabalhos analisados na [Tabela 2.3](#) e [Tabela 2.3](#) possuem pouca ou nenhuma interseção, mas, de maneira geral, podem ser apresentados os dados relativos à Acurácia ([Equação 4.6](#)), a taxa de falsos positivos, muito presente principalmente nos trabalhos de detecção de olhos, o tempo de resposta (T_r) ([Equação 4.1](#)) em segundos(s) do modelo e a possibilidade ou não de uma detecção eficiente (sensibilidade) independente da variação da escala, pose e iluminação.

Tabela 2.3: Análise geral dos métodos de detecção de olhos fechados

Trabalho	Análise					Taxa de FP
	Acurácia	Resposta (s)	Varição Escala	Varição Pose	Varição iluminação	
Abordagem baseada no conhecimento						
Salehi <i>et al.</i> (2012)	97%	0.3	✓	✓	✓	N/A
Sebari & He (2013)	96%	0.1	✓	✓	✓	2%
Rautaray & Agrawal (2015)	97%	0.1	✓	✓	✓	2%
Murawski & Rózanowski (2013)	70%	0.3	✗	✗	✗	N/A
Lethaus <i>et al.</i> (2011)	93%	00.5	✓	✗	✗	3%
Wu & Trivedi (2010)	92.43%	N/A	✓	✓	✓	7.26%
Reinders (1997)	98%	0.2	✓	✓	✓	2%
Isiaka & M (2014)	100%	N/A	✗	✗	✗	N/A
Abordagem baseada características invariantes						
Ghimire & Lee (2013)	98.3%	0.2	✓	✓	✓	2%
Cootes <i>et al.</i> (1995)	92.45%	0.2	✓	✓	✓	2%
Toth & Madabhushi (2012)	91.92%	0.2	✓	✓	✓	2%
Xu & Lee (2013)	98.3%	00.1	✓	✓	✓	3.5%
Krafka <i>et al.</i> (2016)	98.4%	0.005	✗	✗	✗	3.8%
Parada <i>et al.</i> (2015)	98.75%	0.003	✓	✓	✓	0.5%
Wood <i>et al.</i> (2015)	99.23%	0.005	✓	✓	✓	2.3%
Abordagem baseada Templates						
Cootes <i>et al.</i> (2001)	96.4%	✗	✗	✗	✗	N/A
Brunelli & Poggio (1993)	95.32%	✗	✗	✗	✗	N/A
Kossaifi <i>et al.</i> (2014)	92%	✗	✗	✗	✗	N/A
El-Sayed & Khafagy (2014)	100%	✗	✗	✗	✗	N/A
Li <i>et al.</i> (2001)	95%	✗	✗	✗	✗	N/A
Rajpathak <i>et al.</i> (2009)	96%	✗	✗	✗	✗	N/A
Hemlata & Motwani (2012)	90%	✗	✗	✗	✗	N/A
Abordagem baseada na Aparência						
Jia <i>et al.</i> (2012)	95%	0.005	✓	✓	✓	4.1%
Guo <i>et al.</i> (2014)	94%	0.03	✓	✓	✓	2%
Schmidt & Kasiński (2007)	98%	0.003	✓	✓	✓	1%
Eriksson & Papanikolopoulos (2001)	96%	0.01	✓	✓	✓	3%
Bacivarov <i>et al.</i> (2008)	95.5%	N/A	✓	✗	✗	N/A
Soetedjo (2012)	91%	N/A	✓	✓	✓	0.5%
Liu <i>et al.</i> (2012)	97.67%	0.289	✗	✗	✗	5.05%
Zhu <i>et al.</i> (2002)	91.5%	✓	✓	✓	✓	N/A
Abordagem baseada métodos híbridos						
Hassaballah & Ido (2009)	97.8%	0.3	✗	✗	✗	0.4%
Majumder <i>et al.</i> (2013)	96%	0.2	✗	✗	✓	N/A
Jin <i>et al.</i> (2006)	90.75%	0.023	✗	✗	✗	N/A
Chennubhotla & Jepson (2004)	95%	0.01	✗	✗	✗	0.1%

Tabela 2.4: Análise geral dos métodos de detecção de movimentos da Cabeça

Trabalho	Acurácia	Análise	
		Resposta (s)	Erro
Abordagem baseada em algoritmos de visão computacional			
Siriteerakul <i>et al.</i> (2011)	95%	0.2	3%
Zhao <i>et al.</i> (2012)	92.86%	0.1	5%
Zhao & Yan (2011)	90%	0.1	4%
Basu & Pentland. (1996)	93%	N/A	0.2%
Liao & Cohen (2005)	96%	N/A	0.1%
Liu <i>et al.</i> (2012)	95%	N/A	3%
Murphy-Chutorian & Trivedi (2010)	94%	N/A	2%
Liu & Zhao (2011)	93%	N/A	4%
Zhao <i>et al.</i> (2011)	90%	N/A	5%
Qin & Shelton (2016)	96%	N/A	3%
Abordagem baseada Acelerômetros e Giroscópios			
Kiruluta <i>et al.</i> (1997)	98%	N/A	1%
Kupetz <i>et al.</i> (2010)	97%	N/A	3%
Song <i>et al.</i> (2011)	96%	0.1	2%
Nguyen <i>et al.</i> (2004)	93.75%	N/A	2%
King <i>et al.</i> (2005)	99.05%	N/A	1%

2.5 Conclusões

O objetivo deste capítulo foi apresentar os principais trabalhos de investigação desenvolvidos no âmbito do escopo deste trabalho que fundamentam o desenvolvimento do modelo computacional para navegação da CR.

Os projetos de CR apresentados na [Tabela 2.1](#) revelaram que, em sua maioria, apresentam a utilização de algum tipo de sensor que norteia a navegação e muitos empregam como método de entrada a detecção do movimento da cabeça e a captação de alguma informação dos olhos. Estes projetos tornam a CR extremamente invasiva, como alguns exemplos descritos na [Tabela 1.2](#). Buscando a redução desta invasividade, índices de acurácia superiores a 97% e a facilidade de aprendizagem para o uso da cadeira, a navegação pode receber uma enorme contribuição no campo da visão computacional.

Os algoritmos que podem trabalhar, em conjunto, para navegação de uma CR propõem a detecção do estado dos olhos e o movimento da cabeça como meios que podem apre-

sentar resultados promissores. As tabelas 2.4 e 2.3, apresentam variações de acurácia que oscilam na casa de 94.8% para detecção dos olhos e 94.6% para movimento da cabeça. Esses resultados revelam que os métodos empregados podem sofrer aprimoramentos; antes as dificuldades apresentadas, majorarem estes resultados para índices maiores ou iguais a 97%. Essa meta é estabelecida tanto pela possibilidade tecnológica atual como pela resposta deste método esperada pelo usuário final que, em fato, deveria ser de 100%.

No entanto pesquisas desta natureza sofrem desafios na apresentação de métodos robustos, pois dificilmente são apresentados dados sobre o tempo de resposta entre o usuário, CPU e atuadores, comportamento quanto a variações de pose, iluminação, oclusão e escala. Não testam os métodos utilizando base de dados com alto grau de variação dos diversos ruídos que ocorrem na captação dessas imagens. Além disso, os métodos de detecção do movimento da cabeça apresentam, em particular, o uso intenso de hardware específico, que torna o paciente parte da máquina e não usuário da mesma.

Modelo Proposto

Após a análise do estado da arte, identificamos como atuam os algoritmos de detecção propostos por Viola & Jones (2004) para objetos e por Cootes *et al.* (2001) para formas e notamos que possuem muita flexibilidade nos ajustes e na otimização da variância de escala e pose. Assim, o objetivo deste capítulo é apresentar a arquitetura do modelo computacional proposto para a navegação de uma CR.

Iniciamos apresentando alguns pressupostos que formam uma base para o modelo e, em seguida, apresentamos separadamente as técnicas e de que forma as integramos e melhoramos nossas limitações e conclusões.

3.1 Definições iniciais - Pré-processamento

A fase de aquisição da informação da face merece uma atenção especial em qualquer modelo de visão computacional (GONZALEZ *et al.*, 2004), uma vez que a sua coleta é oriunda de dispositivos diversos, realizada de forma manual ou automática e em muitos casos sem qualquer padronização. Sendo assim, os resultados podem ser imagens com fundos distintos, diferentes níveis de luminosidade, tonalidade, posicionamento distinto, sombra exacerbada, entre outros ruídos que aumentam a taxa de falso positivo de qualquer modelo. Para tentar minimizar estes defeitos, pode-se aplicar técnicas de pré-processamento sobre a base de dados a fim de eliminar ou enfatizar as características desejadas.

A rotina de pré-processamento consiste, portanto, em uma série de operações que são aplicadas à imagem no nível de intensidade do *pixel*. Sendo que o objetivo desta rotina é a obtenção da mais regular das características da região facial e dos olhos, eliminando possíveis distorções.

3.1.1 Definição de Imagem Digital

Uma imagem digital é definida com uma função $f(x, y) = z$, sendo que os valores de x e y definem a posição na malha e o valor de z que representa o nível de cinza ou grau de brilho, comumente chamado de *pixel*. O nível de cinza corresponde a uma unidade de informação em termos da quantidade de energia refletida e ou emitida pelo objeto

imageado, em diferentes comprimentos de onda no espectro eletromagnético. [Rosot et al. \(2001\)](#) afirmam que na maioria dos sistemas, um valor de $z = 0$ representa o preto absoluto e $z = 255$ o branco absoluto. Imagens em cores utilizam bandas distintas de frequência com isso uma imagem no padrão *RGB* utiliza três bandas: Vermelha (R-red), Verde (G-green) e Azul (B-Blue). Em todo nosso trabalho utilizamos imagens e escala cinza e portanto a imagem é um conjunto tal que $z \in [0, 255]$.

3.1.2 Conversão em escala cinza

Uma imagem digital em escala RGB pode ser convertida para a uma outra escala de cor, chamada escala de cinza. Na nova imagem o valor de cada pixel representa uma única amostra de um espaço de cores, variando entre o preto como a menor intensidade e o branco como maior intensidade. Imagens em nível de cinza são diferentes de imagens binárias em preto e branco, que contém apenas duas cores. Imagens em nível de cinza podem contar com diversos tons de cinza em sua composição ([GONZALEZ et al., 2004](#)).

Segundo [Bradski & Keahler \(2008\)](#) dada uma imagem f em escala de cores *RGB* e as constantes: $a = 0.2126$, $b = 0.71526$ e $c = 0.0722$ dizemos que sua correspondente em escala cinza é dada pela equação 3.1

$$z_{(x,y)} = aR_{(x,y)} + bB_{(x,y)} + cG_{(x,y)} \quad (3.1)$$

com $x \in$ largura da imagem, $y \in$ comprimento da imagem.

3.1.3 Filtro de Retina

O filtro de retina conforme [Benoit et al. \(2010\)](#), é uma técnica de processamento para imagem pertencente a “abordagem baseada na visão biológica”, cuja ideia básica é copiar o sistema visual humano (HVS) modelando algumas de suas partes, a fim de desenvolver módulos com processamento de imagem a baixo nível.

Este algoritmo, proposto por [Meylan et al. \(2007\)](#), baseia-se na análise de luminância e a sua valorização. Assume-se que a percepção da cor está relacionada com as taxas de intensidade da luz refletida em faixas de comprimentos de onda específicos computadas entre as áreas adjacentes. Como consequência, esse algoritmo é dedicado a aplicações de cor possuindo assim três efeitos importantes: cancelamento de ruídos em sinais de alta frequência, detalhamento do sinal e redução da luminância. Compressão da luminância logarítmica, permitindo um maior nível de detalhes da imagem mesmo em condições de

pouca luz, como vemos na [Figura 3.1](#).



Figura 3.1: Aplicação do Filtro de Retina a imagem do olho

Fonte: Elaborada pelo autor

3.1.4 Extração da região de interesse

A detecção de uma ROI tem por objetivo a diminuição da área de busca para execução de alguma tarefa. Ela pode ser realizada por meio de operações algébricas, técnicas de segmentação ou ainda algoritmos de detecção para objetos. Neste trabalho optamos pelo uso do algoritmo de [Viola & Jones \(2004\)](#), que é capaz de detectar faces com acurácia e precisão acima de 98% e baixa taxa de falsos positivos com um baixo custo computacional ([BEYMER, 1994](#)). A [Figura 3.2](#) apresenta um exemplo da precisão e acurácia do método quando submetido a uma imagem complexa e a extração da respectiva ROI, observamos em particular, que o método aceita bem a variação de pose.

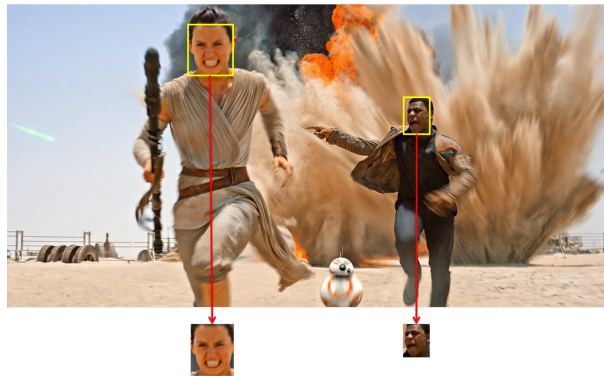


Figura 3.2: Exemplo de extração da face numa imagem complexa

Fonte: starWars.com

3.1.5 Extração da região dos olhos

Apesar da ROI diminuir a região de busca ela ainda pode ser melhorada. Diversos elementos que estão contidos na face não são de interesse do modelo e portanto podem ser descartados. Isso melhora a acurácia, precisão e tempo de resposta do modelo e a ope-

ração em si, por ser estritamente baseada em matemática básica, não influencia no T_r . A Figura 3.3 e a Tabela 3.1 apresentam esta operação que foi estritamente baseada no trabalhos Ekman & Rosenberg (2005).

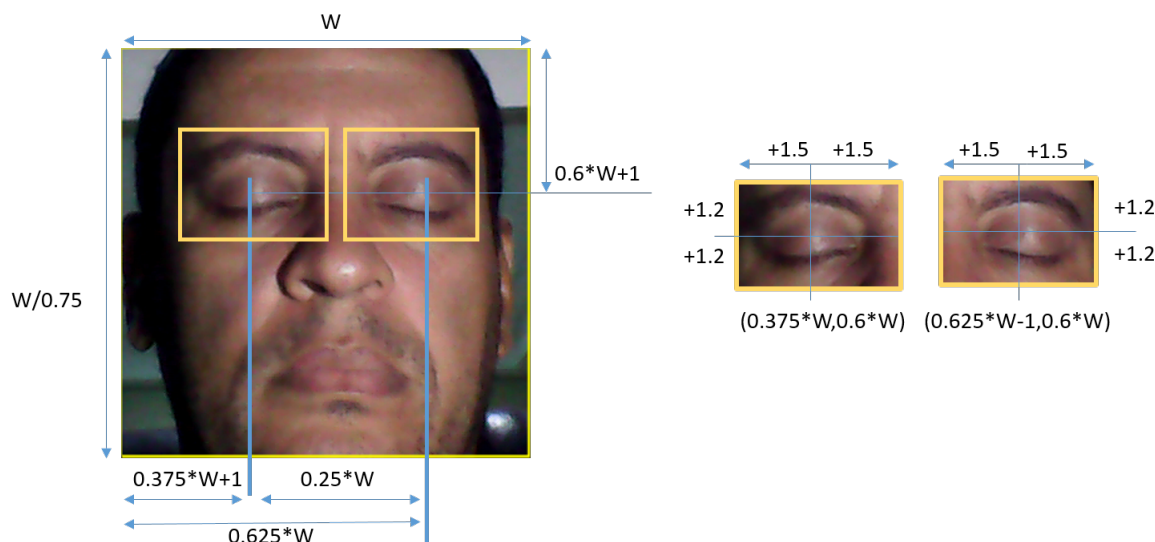


Figura 3.3: Extração do Olho

Fonte: Elaborada pelo autor

Parâmetros	Valor
Largura da Imagem	W
Altura da Imagem	$W/0.75$
coordenada y dos olhos	$0.6*W$
coordenada x do olho direito	$0.375*W$
coordenada x do olho esquerdo	$(0.625*W)-1$
distância entre os olhos	$0.25*W$

Tabela 3.1: Tabela com dados para extração da região dos olhos em que $W = 800$ que corresponde a 200 *pixels* de olho a olho

As etapas de pré-processamento podem ser resumidas na Figura 3.4. Nela apresentamos a imagem de entrada, obtida pela webCam em escala RGB com dimensão 800×600 a qual é convertida para escala cinza, usando o filtro Equação 3.1, com os parâmetros descritos e, em seguida, aplicamos o filtro de retina que efetivamente melhora a definição do contorno da imagem. Logo após é aplicado a operação de extração da face usando o classificador padrão de faces disponível no pacote de distribuição da biblioteca OpenCV. E, finalmente, usando as operações numéricas descritas Subseção 3.1.5 que resulta numa região da face que conterà os olhos. Essa operação pode ser realizada dessa forma devido a natureza específica no nosso problema de pesquisa, visto que o paciente tetraplégico sempre está em posição frontal na cadeira.



Figura 3.4: Pré-processamento da informação de entrada do modelo

Fonte: Elaborada pelo autor

3.1.6 Distância de Mahalanobis

A distância de Mahalanobis é uma medida baseada nas correlações entre variáveis com as quais distintos padrões podem ser identificados e analisados (MAHALANOBIS, 1936). Formalmente seja um conjunto de médias $\mu = (\mu_1, \mu_2, \dots, \mu_p)^T$ e S a matriz de covariância para um vetor $x = (x_1, x_2, \dots, x_p)^T$ é definimos a distância de Mahalanobis como:

$$D_M(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)} \quad (3.2)$$

O uso da distância de Mahalanobis propostos por Cootes *et al.* (1995) se dá pela sua invariância de escala.

3.2 Técnica para Detecção dos Olhos Fechados

Similar ao que utilizamos na Subseção 3.1.4, para a detecção de um olho fechado utilizaremos o algoritmo de Viola & Jones (2004). A técnica é baseada nas características de Haar (Figura 3.6), que determinam as intensidades dos pixels em regiões definidas. Como este algoritmo funciona muito bem para a detecção de face, (LAKSHMANAN *et al.*, 2001; BLACK; YACOOB, 1997), partimos do suposto que ele, após o treinamento, irá detectar o objeto olho fechado com acurácia acima de 97%.

O algoritmo de (VIOLA; JONES, 2004) utiliza a combinação de quatro conceitos-chave:

- Características retangulares, chamadas de recursos de Haar;
- Imagem Integral;
- Algoritmo de aprendizagem - AdaBoost;
- Um classificador em cascata;

3.2.0.1 Características de Haar

As características de Haar codificam a existência de contrastes entre as regiões orientadas da imagem. Um conjunto desses recursos são utilizados para codificar os contrastes exibidos por um rosto humano. Na [Figura 3.5](#), apresentamos um exemplo de utilização deste recurso para detecção facial. Essas características são chamadas de Haar, pois seu conceito é semelhante aos coeficientes da wavelet de Haar, definida em uma janela de detecção com $W \times H$ pixels conforme a fórmula:

$$F = \sum_{i=1}^N \omega_i \text{recSum}(r_i) \quad (3.3)$$



Figura 3.5: Ilustração de uma janela de Haar

Fonte: Elaborada pelo autor

Em que ω_i é arbitrariamente escolhido como fator de peso e $\text{recSum}(r_i)$ é a soma das intensidades dos pixels dado por:

$$\text{recSum}(r_i) = \sum_{(x,y) \in N} I(x,y) - \sum_{(x,y) \in B} I(x,y) \quad (3.4)$$

O retângulo r_i é descrito como uma função com cinco parâmetros: $r = (x, y, w, h, \varphi)$ exemplificada na qual x, y são as coordenadas da posição superior do conjunto de pixels, w e h definem as dimensões do retângulo e $\varphi = \{0^\circ, 45^\circ\}$ representa o grau de rotação.

[Viola & Jones \(2001\)](#), relatam o fato de que a escolha do uso de características, ao invés de modelos baseados em estatísticas de pontos da imagem pixel é importante devido aos benefícios do domínio de conhecimento *ad hoc*, ao qual se podem extrair conhecimentos encobertos nas imagens, dificilmente encontradas em um conjunto finito de treinamentos.

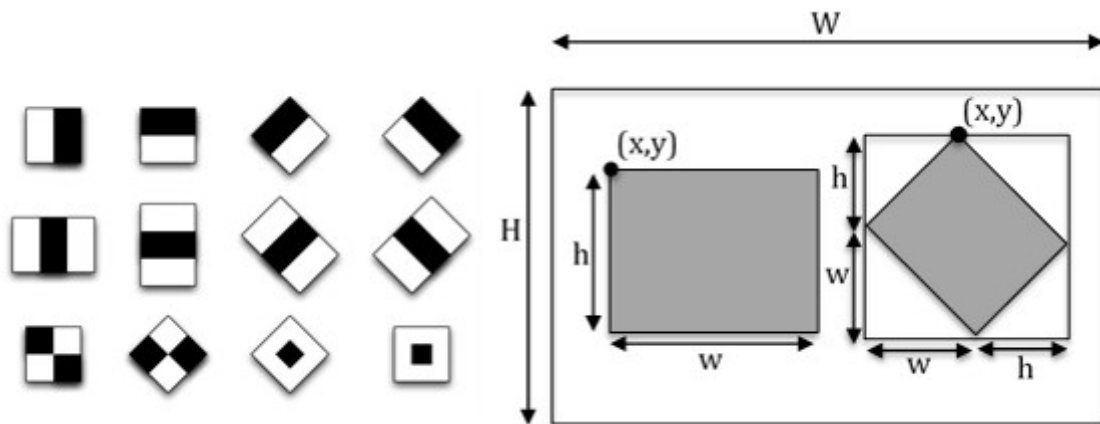


Figura 3.6: Algumas características de Haar e a janela de detecção

Fonte: Elaborada pelo autor

No caso da detecção de olhos fechados este fato é utilizado para a representação das informações aproximadas e, também, relacionadas a planos de fundo da imagem de teste. Este conhecimento torna-se bastante fino com relação ao olho aberto e fechado. A utilização entre os dois dificilmente é encontrada em outros modelos com abordagem baseada em aparência.

De modo geral, portanto, as características nada mais são do que avaliações para um conjunto de informações da intensidade de luz em um pixel. Esse processo consiste na soma da intensidade dos pixels de regiões brancas das características, subtraído a soma da intensidade do restante cinza da imagem. Os resultados são usados como valor da característica de uma determinada localização e podem ser combinados para a formação de hipóteses fracas nas imagens.

Tipicamente o modelo adota os retângulos vistos na [Figura 3.6](#) e para determinar a presença ou ausência de centenas de características de Haar em cada posição de imagem e com escalas diferentes.

3.2.0.2 *Imagem Integral*

O procedimento de detecção mais simples consiste em varrer a imagem de entrada I com S janelas usando K características em uma região W . Assim utilizando $S = 24 \times 24$ podemos gerar em cada região W cerca de 160000 características, o que torna os cálculos extremamente lentos. Com o intuito de otimizar o cálculo das características retangulares [Viola & Jones \(2001\)](#) usaram uma técnica chamada de imagem integral.

As imagens integrais criam, a partir da imagem original, uma nova representação da

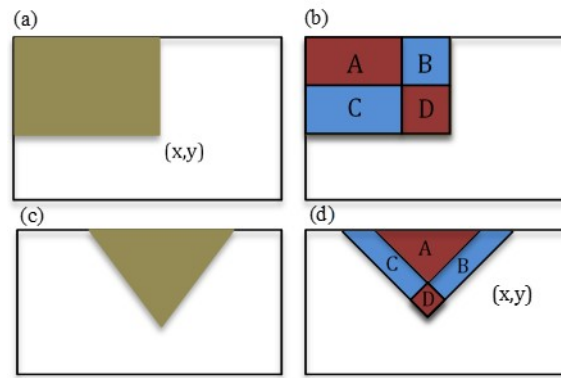


Figura 3.7: Representações da Imagem Integral, (a) região de cálculo, (b) Soma das áreas A-B-C+D, (c) região racionada de cálculo, (d) Soma rápida A-B-C+D

Fonte: Elaborada pelo autor

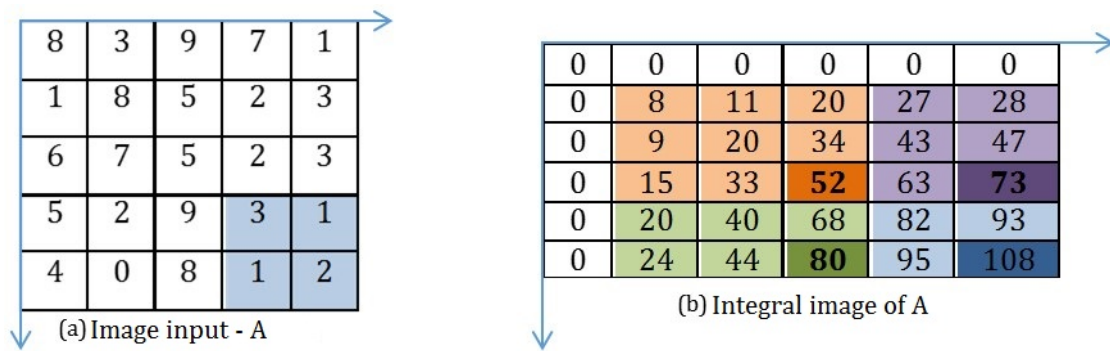


Figura 3.8: Representação do cálculo da imagem Integral, observamos que a soma da região em (a) é igual a sete em (b) representamos como 108-73-80+52

Fonte: Elaborada pelo autor

imagem. Simplesmente, soma-se os valores de cada pixel a esquerda e acima, inclusive. A ideia de usar essa representação é aumentar a velocidade na extração de características, pois qualquer retângulo de uma imagem pode ser calculado por meio desta ideia. São necessários apenas, quatro índices para calcular qualquer retângulo e como consequência imediata, precisa-se de apenas uma passagem para obter dados desejados em sub-regiões de uma imagem, vide [Figura 3.7,3.8](#).

A ideia em usar essa representação é aumentar a velocidade na extração de características, pois qualquer retângulo de uma imagem pode ser calculado por meio da imagem integral.

Formalmente uma imagem integral I é uma representação intermediária para uma imagem e contém a soma dos pixels, em escala cinza, da imagem N com altura y e largura x , ou seja:

$$I(x, y) = \sum_{x'=0}^x \sum_{y'=0}^y N(x', y') \tag{3.5}$$

A imagem integral é calculada de forma recursiva, através das fórmulas:

$$I(x, y) = I(x, y - 1) + I(x - 1, y) + N(x, y) - I(x - 1, y - 1) \quad (3.6)$$

com $I(-1, y) = I(x, -1) = I(-1, -1) = 0$, necessitando apenas de uma varredura sobre os dados de entrada. Esta representação intermediária $I(x, y)$ permite o cálculo de um valor recursivo do retângulo em (x, y) com altura e largura (h, w) utilizando quatro referências, assim:

$$F(x, y, h, w) = I(x, y) + I(x + w, y + h) + I(x, y + h) - I(x + w, y) \quad (3.7)$$

3.2.0.3 Algoritmo de Aprendizagem - AdaBoost

A técnica de [Viola & Jones \(2001\)](#) utiliza algoritmos de aprendizagem para que a implementação da detecção sempre aprimore a acurácia de uma classificação, preservando e melhorando tudo aquilo que já aprendeu, ou seja, prover a melhor estimativa possível para a classe de estudo.

Como uma imagem produz uma enorme quantidade de dados a serem traduzidos para obter o resultado desejado, selecionar os dados de quem possua maior relevância para esta inferência é um fator limitante para a performance do sistema como um todo. Para melhorar este quadro, [Viola & Jones \(2004\)](#) propõem combinar classificadores fracos podendo eventualmente disponibilizar um classificador forte. Em outras palavras, o classificador forte produzido pelo algoritmo de aprendizagem é simplesmente um conjunto de características com pesos que classifica de forma precisa dois conjuntos de imagens pré-rotuladas as quais as características com pesos maiores são boas para a classificação de exemplos definidos como parte de um certo conjunto.

Um classificador fraco h_j é uma simples estrutura contendo um vetor de característica f_j , um limiar θ_j e uma paridade p_j . A ideia básica desse classificador é encontrar um limiar que melhor separe o valor de uma característica entre as imagens definidas como positivas ou negativas.

$$\begin{cases} 1 & \text{se } p_j f_j(x) < p_j \theta_j \\ 0 & \text{Caso contrário} \end{cases} \quad (3.8)$$

Para que complete a tarefa de classificação é importante descobrir um limiar ótimo θ_j . Este limiar não é definido de modo claro e é assumido apenas, que favorece o valor de menor erro de predição. Logo, temos como consequência vários modelos que procuram descobrir um bom limiar.

O problema Boosting nasceu na área de computação conhecida com *machine learning* que informalmente pode ser exposto da seguinte forma: suponha que exista um método de classificação que é ligeiramente melhor do que uma escolha aleatória, para qualquer distribuição em X . Chamado de *weak learner* ou classificador fraco. A existência de um classificador fraco implica na existência de um classificador forte, *strong learner*, com erro pequeno sobre todo o espaço X

Em estatística, consiste em perguntar se dado um método razoável de estimação é possível obter um método próximo de ótimo.

Algorithm 1 Modelo básico do algoritmo AdaBoost

Sejam:

Amostras: $x_1 \dots x_n$

Saídas esperadas: $y_1 \dots y_n$ com $y \in \{-1, 1\}$

Pesos Iniciais: $w_{1,1} \dots w_{n,1}$ definidos como $\frac{1}{n}$

Função Erro: $E(f(x), y, i) = e^{-y_i f(x_i)}$

Classificadores Fracos(CF): $h : x \rightarrow [-1, 1]$

para $t = 1, 2, \dots, T$ faça

Escolha $h_t(x)$ tal que minimize ϵ_t , em que $\epsilon_t = \sum_{\substack{i=1 \\ h_t(x_i) \neq y_i}}^N w_{i,t}$

Tome: $\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right)$

Atualize $F_t(x) = F_{t-1}(x) + \alpha_t h_t(x)$

Atualize os pesos: $w_{i,t+1} = e^{-y_i \alpha_t h_t(x_i)} \quad \forall i$

Normalize: $w_{i,t+1}$, fazendo $\sum_i w_{i,t+1} = 1$

Este problema foi resolvido por [Schapire \(1990\)](#), que apresentou um algoritmo que transforma um classificador fraco num classificador forte. A partir de então, foram desenvolvidos vários algoritmos dentro do contexto *boosting*, sendo um dos mais recentes e bem sucedido o algoritmo conhecido com AdaBoost que é oriundo do fato de que o *boosting* gera em cada passo uma distribuição sobre as observações da amostra, dando maior peso (maior probabilidade de estar na amostra perturbada) às observações classificadas incorretamente no passo anterior.

Nesse sentido o AdaBoost está focado nas classificações ruins, ou então, nos dados de

difícil classificação, e esta é a característica principal deste algoritmo: minimizar o erro sobre um conjunto de treinamento. Uma das vantagens do adaboost [Nock & Nielsen \(2006\)](#), [Gao & Zhou \(2010\)](#), é a existência de outros parâmetros, além dos turnos T , para que melhore o aprendizado.

O resultado, após sucessivas iterações do algoritmo, de um conjunto de hipóteses com pesos, em que aqueles que possuem menores erros de classificação ganham maior importância, é chamado de hipótese forte ou classificador forte.

Muitas variações deste algoritmo foram apresentadas na literatura para resolver problemas diversos. Entre eles destacamos os problemas binários, ou seja, que existem somente duas classes para que as amostras sejam classificadas. Temos os *Discrete AdaBoost*, *Real AdaBoost*, *Gentle AdaBoost* e o *Modest AdaBoost* ([FREUND; SCHAPIRE, 1999](#); [SCHAPIRE, 1990](#))

3.2.0.4 Um classificador em cascata

Aumentar a velocidade de uma tarefa de classificação, em geral, resulta num aumento dos erros associados. Porém, para surtir este efeito teríamos que reduzir o número de avaliação dos classificadores fracos, o que resultaria numa perda na acurácia do sistema.

Por isso Viola e Jones ([VIOLA; JONES, 2004](#)), propõem uma árvore degenerativa de decisão, *decision stump*, cuja estrutura contém o encadeamento de classificadores do mais genérico ao mais específico, segundo o qual os primeiros níveis da cascata não são muito precisos, apesar de conseguir classificar uma grande quantidade de amostras com uma pequena quantidade de características.

O uso da cascata caracteriza-se pelo fato de que, em uma imagem durante uma tarefa de detecção, a maioria das sub janelas analisadas pelo classificador são rejeitadas. Por essa razão, uma generalização nos primeiros estágios deve ser suficientemente alta para que evite a passagem para estágios subsequentes, sub janelas classificadas como falso positivo [Horton et al. \(2007\)](#), conforme descrevemos na [Figura 3.9](#).

A eficiência do classificador em casca depende de dois conjuntos: o primeiro é o conjunto de positivos bem estruturado e que apresente todas as possíveis variações do objeto que se deseje avaliar. O segundo é um bom conjunto de parâmetros de classificação.

Apesar dos estudos de ([BLEDSOE, 1966](#); [CHELLAPPA et al., 1995](#); [ETHEVE et al., 2006](#); [FIGUEREDO, 2011](#); [GONG; XIANG, 2011](#); [HJELMAS, 2001](#); [KANADE, 1973](#); [KHAN; ALIZAI, 2006](#)) demonstrarem que a utilização de uma cascata de classificadores

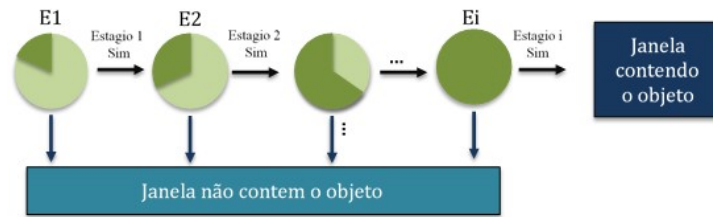


Figura 3.9: Modelo do Classificador em Cascata

Fonte: Elaborada pelo autor

bem treinados venha a ser extremamente eficiente quanto à detecção de objetos, o processo de treinamento para geração da cascata é extremamente demorado. Viola & Jones (2001) afirmam que para construção do seu detector de faces usando 32 estágios, 4297 características, 5000 imagens de faces (24×24) e 10.000 imagens que não contem nenhuma face, o tempo total de treinamento foi de 2 semanas numa máquina com clock 466MHz. Sendo esta uma das desvantagens da técnica.

3.3 Técnica para detecção dos movimentos da cabeça

A capacidade de entendimento do movimento de uma cabeça humana é trivial ao olhar humano, pois permite inferir as intenções de outras pessoas que estão nas proximidades e compreender uma forma não-verbal importante de comunicação. Também pode ser considerado uma forma natural, simples e eficaz de apontar para um objeto. Assim, a detecção do movimento da cabeça permite uma boa fonte de informações para o processo de interação Homen \times Máquina, controlando a ação de alguns dispositivos ou sinais de controle.

Do ponto de vista computacional, mais precisamente em visão computacional, estimar a “pose” é o processo de inferir a orientação de uma cabeça humana a partir de imagens digitais.

Muitos processos foram propostos, por diversos pesquisadores, para realização desta tarefa, que procurara tratar a representação da imagem, baseada em pixel, numa informação que traduzisse o movimento da cabeça e, como outras etapas do processamento de imagens, demonstrando um boa tolerância à invariância inerente ao estudo de imagens.

Uma gama cada vez mais crescente de tecnologias tem repousado seus esforços no aprimoramento e criação de novos algoritmos capazes de realizar a detecção do movimento da cabeça. Isso se dá principalmente pela crescente demanda de tecnologias voltadas à pessoas com deficiência, aplicações de realidade virtual, teleconferências.

O controle da direção em uma cadeira de rodas tem sido usado de diversas formas como destacado em (YI *et al.*, 2015; JAFFE, 1982; MANOGNA *et al.*, 2010; ZHAO *et al.*, 2012). Frequentemente a solução proposta envolve o uso de algum hardware específico para captar a rotação da cabeça.

Aqui apresentamos um casamento de duas técnicas que viabilizam o uso da pose para navegação da CR sem o uso de hardware específico. Utilizamos marcos faciais que compõe um modelo deformável ou flexível e, para avaliar a posição atual da cabeça em relação a câmera, utilizamos a correlação normalizada que, basicamente, compara os pontos atuais com os pontos de 7 amostras.

Notadamente, o modelo utilizado para a determinação dos pontos característicos foi o *Active Shape Model-ASM*, apresentado por Cootes *et al.* (1995). O modelo é estatístico e frequentemente utilizado em diversas aplicações (CHEN *et al.*, 2012; BIRKBECK; JAGERSAND, 2004; HOMMEL; HANDMANN, 2011). Nele a forma do objeto é definida por um conjunto de pontos (marcos) e as formas deste são limitadas pelo modelo de distribuição de pontos (PDM) e pelas variações dos níveis de cinza dos pontos característicos. (COOTES *et al.*, 1995) ainda utilizam a distância de Mahalanobis entre os marcos e a média entre estes para detectar a melhor posição para cada um dos referenciais.

O ASM utiliza um conjunto de treinamento, definido previamente, com base no objeto e suas possíveis deformações, rotações e translações para que o modelo consiga localizar um objeto em outra fonte. De posse destes dados é possível realizar uma busca numa imagem qualquer de tal forma a localizar objetos similares. A quantidade de pontos neste conjunto de treinamento pode variar de acordo com a especificidade do objeto, mas um número excessivo de pontos não significa a garantia de uma convergência segura como vemos no experimento realizado com imagens da face treinadas com diferentes conjuntos de pontos.

O ASM pode ser dividido em 3 etapas (COOTES *et al.*, 1995):

1. Construção de um conjunto com modelos de pontos característicos;
2. Detecção dos níveis de cinza de cada ponto;
3. Localização dos pontos característicos em uma nova imagem.

A primeira etapa destina-se à aquisição da informação relevante das imagens. Obtendo-se assim, um conjunto de parâmetros necessários para a modelação de uma forma média. Estes parâmetros são obtidos após uma análise estatística da informação recolhida em que é estudada tanto as posições como a intensidade da vizinhança de cada ponto. A variedade e variabilidade deste pontos constitui a base da técnica e a obtenção destes pontos é

chamada de Modelo de Pontos Distribuídos (PDM). A análise estatística destina-se a diminuir a grande quantidade de informação proveniente das imagens de treino.

3.3.1 Modelo de Pontos Distribuídos - PDM

Com a utilização do PDM espera-se obter uma forma característica da imagem que definirá o ponto do modelo ser capaz de determinar esse mesmo conjunto de características em outra imagem. Este processo demanda algumas etapas as quais temos:

3.3.1.1 Etiquetagem do conjunto de treino

Nesta etapa, determina-se a localização dos pontos mais adequados nas imagens do conjunto de treino. É realizada manualmente e considerando locais que melhor descrevem o objeto que, no nosso estudo, será uma face. Este conjunto é formado por pontos-chave da face que são exatamente ao redor da face, dos olhos, nariz e boca que são pontos de elevado valor de curvatura, vértice e bifurcação, [Figura 3.10](#). É fundamental que os pontos se posicionem o mais rigorosamente possível no contorno que pretendem descrever, de modo a que representem corretamente essa estrutura.

O número n de pontos que descreve cada contorno é arbitrário. No entanto, uma vez escolhido, esse número tem de ser o mesmo para todos os contornos das imagens. Para que seja possível capturar as variações da forma dos contornos é necessário a correspondência entre as posições dos marcos de imagem para imagem. A figura [Figura 3.10](#) representa um exemplo desta etiquetagem. Nela percebemos que a estrutura da face é determinada por um ponto/marco m com $m = 1, \dots, n$ e assim sucessivamente nas outras imagens. E assim teremos ao final da coleta de informações uma coleção de pontos/marcos.

Um contorno i com $i = 1, \dots, N$ presente em cada uma das N , imagens de treino, e delimitado por um conjunto de n pontos etiquetados

Essa etapa é realizada em todo o conjunto de treino o que equivale a $N = 150$ imagens e $n = 68$ pontos marcados em cada. Em que o ponto (x_{ij}, y_{ij}) pertence a j -ésima coordenada da forma da i -ésima imagem.

$$x_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{ik}, y_{ik}, \dots, x_{in-1}, y_{in-1})^T \quad (3.9)$$

onde o tamanho de x_i é $2n$ e $1 \leq i \leq N$.

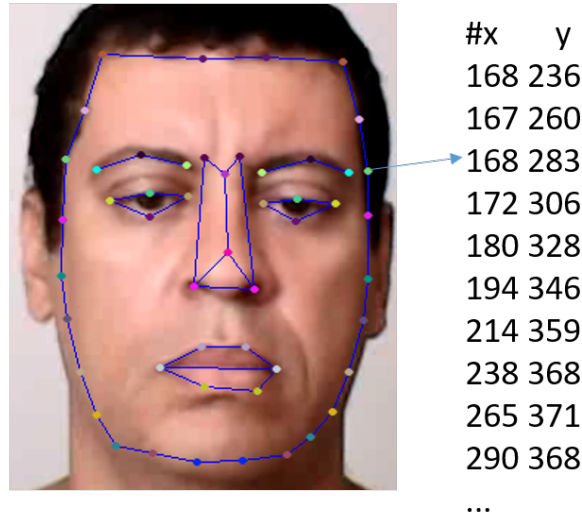


Figura 3.10: Etiquetagem realizada em uma das imagens

Fonte: Elaborada pelo autor

3.3.1.2 Alinhamento das formas

O conjunto de contornos resultantes do treino não se encontram, em geral, alinhados. Para poder comparar os pontos equivalentes de distintos contornos, estes têm que ser alinhados relativamente a um sistema de eixos (COOTES *et al.*, 1995). Este alinhamento é feito executando a minimização das distâncias entre os marcos e, em seguida, por rotações, translações e redimensionamento (escala). De modo que as distâncias Euclidianas entre marcos correspondentes de diferentes contornos sejam os menores possíveis. Iterativamente, utilizamos as seguintes forma:

1. Calcular o centro de gravidade das amostras e colocar todos os demais pontos com o mesmo centro por meio de uma translação;
2. Tomar um objeto com estimativa inicial da forma média e redimensioná-lo de forma que $|\bar{x}| = 1$;
3. Fazer \bar{x}_0 como estimativa inicial;
4. Alinhar todos os pontos com a forma média $D = \sum_{i=1}^{N_s} |x_i - \bar{x}|^2$ sendo D o mínimo;
5. Re-estimar a forma média a partir das formas anteriores;
6. Alinhar a nova forma média, \bar{x} , com \bar{x}_0 e redimensioná-la de forma que $|\bar{x}| = 1$;
7. Repetir os passos anteriores até que o valor da forma média convirja.

A forma média é dada por $\bar{x} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i$ e os modos de variação, podem ser encontrados

aplicando uma análise de componentes principais aos desvios da média, o que consiste calcular a matriz S dada por

$$S = \frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})(x_i - \bar{x})^T$$

Sendo que a variação dos pontos são descritos pelos vetores próprios de S e p_i , tal que:

$$Sp_i = \lambda_i p_i$$

onde λ_i é o i -ésimo valor próprio de S ($\lambda_i \geq \lambda_{i+1}$) e $p_i^T p_i = 1$. O que nos leva a interpretar que cada objeto do conjunto de treino pode ser descrito pela forma média e pela combinação dos primeiros t vetores próprios obtidos:

$$x = \bar{x} + pb$$

onde $P = (p_1, p_2 \dots p_t)$ é a matriz dos primeiros t vetores próprios e $b = (b_1, b_2, \dots, b_t)$ é o vetor de pesos de cada vetor próprio. Os vetores próprios são ortogonais e portanto $P^T P = I$ e:

$$b = P^T (x - \bar{x}) \quad (3.10)$$

Como a variância de b_i sobre o conjunto de treino é dada por λ_i , os seus limites são da ordem

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i} \quad (3.11)$$

Estes valores na [Equação 3.11](#) são propostos por [Cootes et al. \(1995\)](#), mas podem ser ajustados. Neste trabalho não alteramos este valor.

3.3.1.3 Análise de Componentes Principais

Para reduzir a dimensionalidade dos dados adquiridos na fase de treino, aplica-se a Análise de Componentes Principais (PCA) ao conjunto de vetores x_i que representam as várias formas alinhadas. O espaço multidimensional que contém a informação relativa ao marcos de todos os contornos é então reduzida a um espaço de dimensão menor, mas com toda a informação essencial para utilização dos mesmos. A descrição mais detalhada do PCA pode ser vista em ([COOTES et al., 1995](#); [HOTELLING, 1933](#); [DUNTEMAN, 1989](#); [JOLLIFFE, 2002](#)).

3.3.2 Aparência dos níveis de cinza

Além da informação relativa a posição dos pontos/marcos o ASM modela também as intensidades em torno dos marcos/pontos. Esta informação é utilizada para detecção da região da imagem que mais se assemelha com os pontos/marcos treinados. Em cada ponto da forma os níveis de cinza descrevem a aparência típica da vizinhança onde está localizada o contorno do ponto [Figura 3.11](#). Examinamos as regiões de níveis de cinza próximas. (COOTES *et al.*, 1995) descrevem que usando contornos $1D$, ou seja, para cada um dos pontos característicos j na imagem i do conjunto de treinamento extraímos os níveis de cinza g_{ij} de tamanho n_p que formam o contorno, sendo este como o centro j . Assim, para cada um dos pontos característicos j na imagem i temos os níveis de cinza que podemos definir como $g_{ij} = (g_{ij0}, g_{ij1}, \dots, g_{ijn_{p-1}})^T$, que podem ser normalizados como:

$$y_{ij} = \frac{dg_{ij}}{\sum_{k=0}^{n_p-2} |dg_{ijk}|} \quad (3.12)$$

Dos quais podemos tomar a média das derivadas normalizadas para cada ponto característico e assim tomarmos a matriz de covariância das derivadas, resultando assim num modelo de nível de cinza em torno de qualquer ponto característico j representado por \bar{y}_j e a covariância deste C_{yj} , através do conjunto de treinamento.

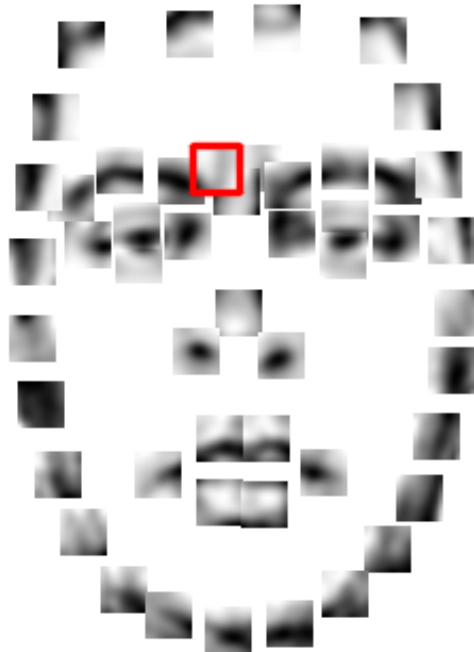


Figura 3.11: Vizinhança dos pontos marcados em tons de cinza de uma amostra

Fonte: Elaborada pelo autor

3.3.3 Localização dos pontos característicos em uma nova imagem

De posse de uma nova imagem contendo o elemento treinado podemos utilizar o [algoritmo 2](#) para localizar a posição dos pontos característicos numa nova imagem.

Algorithm 2 algoritmo para localização dos pontos característicos numa nova imagem

Entrada: $P, \bar{x}, \bar{y}_j, C_{y_j}$

Saída: Localização dos pontos característico numa nova imagem

- 1 *Localizar a face da imagem de entrada usando o Algoritmo de Viola-jones*
Localizar a forma média \bar{x} , no centro da imagem facial
 $b = 0$
repita
 - 2 **para** $j = 1$ *to* n **faça**
 - 3 Calcular os níveis de cinza na nova imagem
 $f_j = (f_{j0}, f_{j1}, \dots, f_{jn_{s-1}})^T$
 Calcular as derivadas de f_j :
 $df_j = (f_{j1} - f_{j0}, f_{j2} - f_{j1}, \dots, f_{jn_{s-2}} - f_{jn_{s-1}})^T$
 Normalizar as derivadas de f_{y_j}

$$f_{y_j} = \frac{df_j}{\sum_{k=0}^{n_{s-2}} |df_{jk}|}$$

 Minimizar a distância entre df_j e y_j usando a distância de Mahalanobis.
 - 4 **fim**
 - 5 Obter o vetor de movimento e calcular o menor deles para deslocar os pontos
 Movimentar os pontos até a posição correta
 Determinar $b = P^T(xf - \bar{x})$
 Limitar b baseado nos autovalores obtidos no conjunto de treinamento e atualizar b
 Estimar $xf = \bar{x} + Pb$
 - 6 **até** distância seja mínima;
-

Finalmente, após a obtenção dos pontos característicos da imagem de entrada, pode-se calcular a correlação normalizada entre todas as imagens de formação e identificar a posição da cabeça. Nota-se que a resposta dos modelos é simétrica e evidencia a simetria vertical visual da cabeça. Além disso, a queda do vetor de resposta do seu pico é monótona, gradual e com um baixo desvio padrão ($\sigma = 0,0431$).

3.3.4 Coeficiente de correlação de Pearson

A correlação ou o quanto uma variável ou conjunto de dados está relacionado com outro é uma medida estatística que pode ser dada de diversas formas. Neste trabalho, utilizamos a medida mais comum de correlação: Correlação de Pearson ou, simplesmente, *Pearson Product Moment Correlation-PPMC*. Sendo esta representada pela letra ρ em que $-1 \leq \rho \leq 1$, nos apresenta a relação linear entre estes conjuntos. Temos também que:

- $\rho = 1$ significa uma correlação perfeita entre os conjuntos;
- $\rho = 0$ não é possível identificar o tipo de correlação;
- $\rho = -1$ significa uma correlação negativa perfeita entre os conjuntos.

O coeficiente de correlação de Pearson é calculado da seguinte forma:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.13)$$

onde $A = \{x_1, x_2, \dots, x_n\}$ e $B = \{y_1, y_2, \dots, y_n\}$ são os valores tomados em cada conjunto e

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.14)$$

e

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (3.15)$$

Apos o cálculo das informações podemos obter as seguintes interpretações:

- $\rho \geq 0.9$ positivo ou negativo indica uma correlação muito forte;
- $0.7 \leq \rho < 0.9$ positivo ou negativo indica uma correlação forte;
- $0.5 \leq \rho < 0.7$ positivo ou negativo indica uma correlação moderada;
- $0.3 \leq \rho < 0.5$ positivo ou negativo indica uma correlação fraca;
- $0 < \rho < 0.3$ positivo ou negativo indica uma correlação desprezível.

Em nosso trabalho adotamos o seguinte padrão de correlação:

- $0.7 \leq \rho < 1$ positivo ou negativo indica uma correlação forte;
- $\rho < 0.7$ positivo ou negativo indica uma correlação fraca.

3.4 Arquitetura da cadeira

Apesar do trabalho orbitar na avaliação da eficiência do modelo para detecção e envio de informações para a cadeira, procuramos, nesta sessão, descrever o hardware que serve de base para um futuro protótipo funcional da cadeira.



Figura 3.12: Hardware da cadeira

Fonte: Elaborada pelo autor

A arquitetura do protótipo em que será implementado o modelo está demonstrada na [Figura 3.12](#). Baseamos nosso trabalho na cadeira comercial *freedom carbon*, produzido pela *Freedom Veículos Elétricos Ltda*, que possui duas rodas traseiras motorizadas com dois movimentos frente e ré, utilizando baterias de 12V/45Ah, um joystick tradicional e um módulo de controle.

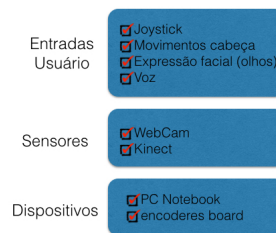


Figura 3.13: Hardware da cadeira em blocos

Fonte: Elaborada pelo autor

O hardware da cadeira possui três blocos: entradas do usuário, sensor, e dispositivos conforme vemos na [Figura 3.13](#)

3.4.1 Entradas do usuário

De maneira geral, o projeto deverá incorporar vários tipos de entradas do usuário, que possibilitarão a integração da cadeira com um controle para decisões inteligentes, responsáveis por ações de movimento, conflitos, parada emergencial entre outros. Mas, no estágio atual, a nossa cadeira conta apenas com a câmera frontal, microfone acoplado ao *kinect* e o *joystick*, Sendo a terceira apenas considerada para comparação com o modelo computacional. Destacamos ainda que:

- O joystick: vem de fábrica para navegação da cadeira em velocidade constante controlada pelo usuário e com giro em torno do eixo com 360° de liberdade. Além de permitir giros simples e bem controlados.
- Câmera frontal: responsável pela captação de duas informações: detecção do fechamento dos olhos por um período não natural e rotação da cabeça em relação a uma linha ao eixo que liga a cadeira à lente da câmera.
- Kinect: colocado na parte superior da cadeira, acima da cabeça do usuário é responsável pelo comando de parada brusca e análise da navegação do usuário. A parada brusca funciona como um botão de pânico, acionado por um comando de voz “Kinect Pare” e pelo posicionamento da cadeira em relação ao ambiente, ou seja, a sua centralização em corredores como também a análise da possibilidade de choque.

3.4.2 Sensores/Dispositivos

Uma das propostas gerais do projeto envolve tanto o baixo custo como a utilização mínima de equipamentos acoplados a cadeira. Assim, dispomos apenas de câmera acoplada ao *laptop* e um sensor kinect (Figura 3.13) colocado na parte superior da cadeira que associados ou independentemente, permitem ao modelo obter diversas informações do ambiente, capazes de determinar ações da cadeira, como por exemplo, a passagem repentina de alguém em frente à mesma ou a proximidade de alguma obstáculo, o sensor também permite a parada total por voz.

Na comunicação entre a cadeira e a placa de comando utilizamos um *laptop* com processador de 1.6GHz dual-core Intel Core i5 (Turbo Boost up to 2.7GHz) com 3MB e 4GB de memória. A placa microcontroladora é uma Arduino UNO, ATmega328, que dispõe de 14 pinos digitais de entrada/saída, 6 entradas analógicas e conexão USB.

3.5 Arquitetura do Modelo Computacional

Nosso modelo computacional, busca utilizar duas ações do usuário. A primeira que, a partir do estado olho (aberto, fechado), informa à cadeira duas possíveis ações:

- 1) seguir em frente;
- 2) parar.

No qual cada comando é acionado quando o outro está inativo, ou seja, o estado atual

for seguir o seguinte será parar e vice-versa. A caracterização dessa ação se dá num movimento de fechamento do olhos de maneira não natural, ou seja, permanecer com o olho fechado, por um período de tempo superior a 2000 milissegundos.

Segundo [Ekman & Rosenberg \(2005\)](#) o olho humano demora 280 milissegundos para piscar e portanto existe uma diferenciação razoável entre o ato natural de piscar e a ação de controle do modelo.

A segunda indica que a cadeira deve iniciar uma rotação que vai variar de 0° a 90° e o sentido dessa rotação (direita ou esquerda) é dado pela posição da cabeça do usuário.

Para que este movimento se torne um comando para a cadeira o indivíduo deve rotacionar a cabeça para a direita ou esquerda e permanecer com ela nesta posição por mais de 2 seg. Após este período de tempo na posição a cadeira interpreta essa ação e “girar” a CR para esquerda ou direita e “Para o giro” quando a cabeça retornar à posição centralizada, simplificamos esses conceitos na [Figura 3.14](#).

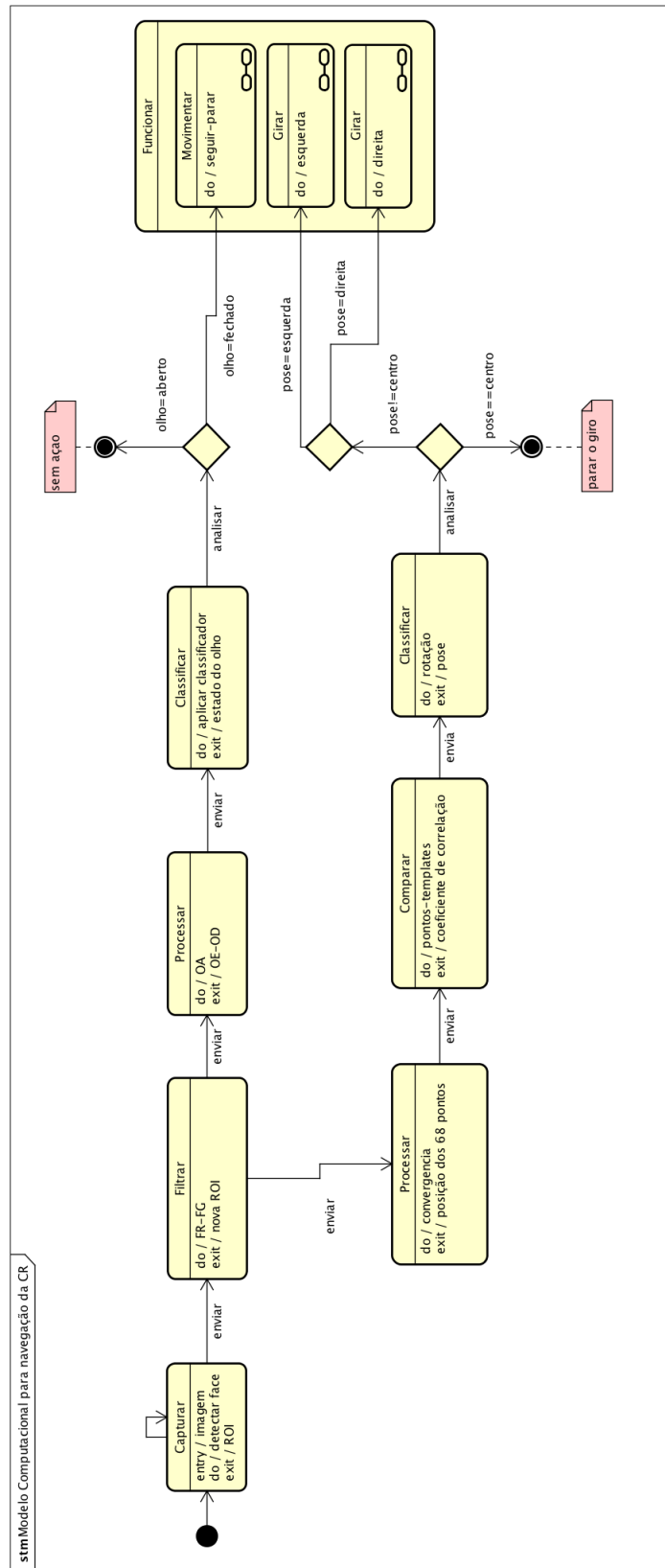


Figura 3.14: Arquitetura do modelo

Fonte: Elaborada pelo autor

O modelo apresentado na [Figura 3.14](#), analisa 28 imagens de entrada por segundo. Cada imagem é capturada com uma resolução de 800×600 no padrão RGB a uma distância de 50cm do usuário, sendo a câmera ajustada num ângulo de aproximadamente 36° .

A operação de detectar face visa a redução do espaço de busca e é realizada utilizando a técnica de [Viola & Jones \(2004\)](#) ao qual aplicamos um classificador incorporado a biblioteca OpenCV. Ao detectar a face a técnica apresenta uma região retangular que pode ser extraída da imagem de entrada. Esta nova imagem, que chamamos simplesmente de ROI, possui uma dimensão menor que a original, mas variável, pois depende de aspectos particulares da formação do rosto do paciente.

A ROI é filtrada, com objetivo de reduzir e melhorar ainda mais o espaço de busca. O filtro que converte a imagem em tons de cinza (FG-filtro conversão escala cinza) que varia de 0 a 255 sendo 0 preto e 255 branco. Em seguida, acontece uma filtragem para remoção de ruídos, notadamente, sombras e/ou falta de iluminação adequada ($<200\text{lux}$) utilizando o filtro de retina ([BENOIT *et al.*, 2010](#)) (FR-Filtro de Retina).

A ROI é enviada, simultaneamente, para extração da região do olho por meio de operações algébricas - OA ([EKMAN; ROSENBERG, 2005](#)) e determina o tipo de convergência da posição da cabeça. Estas etapas são realizadas no computador pelo poder de processamento e rapidez.

Na extração dos olhos, após o recorte da região de interesse, temos duas novas imagens que são normalizadas na dimensão 24×24 e na detecção do movimento da cabeça. A imagem de entrada não passa por esta normalização, pois necessitamos da informação da face completa e assim utilizamos a ROI por completa. Nesse momento, o Modelo de Forma Ativa se revela extremamente eficiente na variação de escala.

Ao iniciar a ação de seguir em frente, a velocidade do movimento varia de $[0, 1\text{km/h}]$ permanecendo constante na velocidade máxima e sendo reduzida a 0.5km/h nas rotações.

As lesões medulares entre as vértebras C1 e C3, preservam a inervação dos músculos esternocleidomastoideo, porção superior do trapézio e elevador da escápula, permitindo controle cefálico e sustentação da cabeça ([CRUZ; IOSHIMOTO, 2010](#)). Caso venha a ocorrer comprometimento do diafragma torna os pacientes dependentes do respirador, no entanto o tetraplégico com nível em C4, após superar a fase aguda, pode apresentar diafragma e trapézio funcionantes, tendo condições de manter-se livre de assistência respiratória e portanto capazes de rotacionar a cabeça.

A análise dessas características permite estabelecer que o indivíduo com tetraplegia preserva, mesmo que limitado, o movimento de girar a cabeça. Assim, o modelo utiliza desta

característica para realizar o movimento de rotação da cadeira em torno do seu eixo num ângulo de rotação mínimo confortável ao usuário de 12° em relação ao eixo central, este ângulo foi estabelecido a partir de testes que descreveram o ângulo mínimo de detecção da rotação.

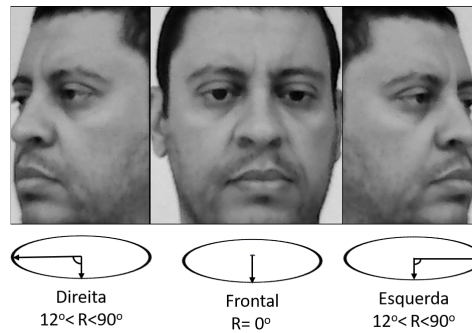


Figura 3.15: Rotação captada pelo detector

Fonte: Elaborada pelo autor

3.6 Limitações do Modelo

As expressões faciais, ao nível inconsciente, oferecem diversas informações sobre o estado emocional do indivíduo e, muitas vezes, apenas através de uma simples micro expressão. Já ao nível consciente, obtivemos um desejo expresso e este objeto pode vir a ser um comando para a navegação.

O modelo deverá ser capaz de identificar este desejo e expressá-lo como um comando de maneira única, ou seja um por vez e não sua associação, assim um indivíduo deve, por exemplo, indicar o desejo de iniciar o movimento e depois indicar o desejo de rotacionar ou não a CR. Esse fator é limitante, do ponto de vista tecnológico do nosso trabalho, pois seria ideal que os comandos agissem de forma independente e analisados em módulos independentes.

Outro fato de limitação reside na necessidade de distinção entre uma expressão intencional ou involuntária. Será avaliada por um determinado lapso mínimo e máximo de tempo. Esta característica elimina as expressões resultantes por exemplo dos olhos fechados uma vez que o contexto de aplicação é dinâmico exigindo um controlo reativo por parte do utilizador. A natureza dinâmica dessa ação pode variar em cada paciente e portanto, necessita de calibração do modelo particularmente.

Existem múltiplas variantes que se apresentam dentro das lesões alvo deste trabalho além da incapacidade motora. Por exemplo, vítimas de acidentes podem apresentar deformações no rosto que tornam as expressões em voga pouco identificáveis, ou os que sofreram

de queimaduras dos tecidos ao ponto de mudarem de pigmentação ou de forma. Além destes, existem ainda as enfermidades degenerativas dos tecidos musculares, sistema nervoso. Paralisia parcial ou total dos músculos da face, motivada sobretudo por doenças como a paralisia de bell, a trombose, o tumor cerebral ou a doença de Parkinson ou neurológicas como a esclerose múltipla. Este trabalho limita sua adaptação o modelo para um paciente que possua a o movimento de rotação da cabeça funcional. nos demais casos os casos o objeto natural de extração da característica se perde ou torna extremamente variante o lapso de tempo o que inviabiliza nossa abordagem.

O modelo que será proposto necessita de que o meio físico de aplicação da CR esteja em conformidade com regras mínimas de acessibilidade estando elas atendidas a possibilidade de falha no sistema é conseqüentemente reduzida a patamares obtidos nos resultados computacionais, não atendidas temos fatores limitantes ao modelo. Além de questões do dimensionamento espacial da residência a eliminação de obstáculos e ponto crucial para que a navegação ocorra satisfatoriamente. Destacamos como fatores limitadores do modelo as condições de piso e iluminação.

O piso caso não esteja regular, poderá causar vibrações na cadeira que a fará modificar sua rota. Apesar do indivíduo poder corrigir sua rota a todo momento esta ação pode gerar cansaço e frustração.

A iluminação deve ser observada e estabilizada para que o detector não sofra com o surgimento de sombras na face do paciente. Boa parte dos ruídos de iluminação são tratados pelo filtro de retina, mas ainda sim temos suscepção a falha enquanto com uma normalização da iluminação temos uma resposta muito boa.

As dimensões mínimas que este trabalho adotará será de $80cm \times 120cm$ com um diâmetro espacial de $150cm$ a partir do seu centro de área livre [Figura 3.16](#).

Sempre que possível, existe acesso a níveis superiores por meio de rampas cujo grau de elevação não deve ser superior a $8,33\%$, com largura mínima de $100cm$.

Padrões que diferem do especificado devem ser evitados porque podem causar danos tanto ao paciente como a própria CR. A iluminação desta área deve ter origem na parte superior e inferior de cada etapa da rampa. As portas do ambiente devem possuir, em seu conjunto, $100cm$ com uma largura. a largura de corredores dever ser de no mínimo $120cm$ e caso existam duas cadeiras no ambiente, $180cm$. A largura mínima para a transposição de obstáculos isolados com extensão acima de $40cm$ deve ser de $90cm$.

O modelo deveria estar preparado para enfrentar qualquer umas das situações comuns a um sistema com interação homem máquina, conforme descrito na [Figura 3.17](#),mas por

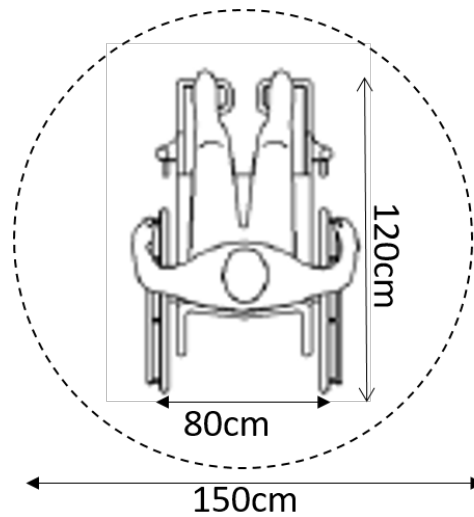


Figura 3.16: Dimensões mínimas da cadeira para utilização do modelo proposto
 Fonte: Elaborada pelo autor

questões de tempo e projeto todas estas iterações não podem ser contempladas.

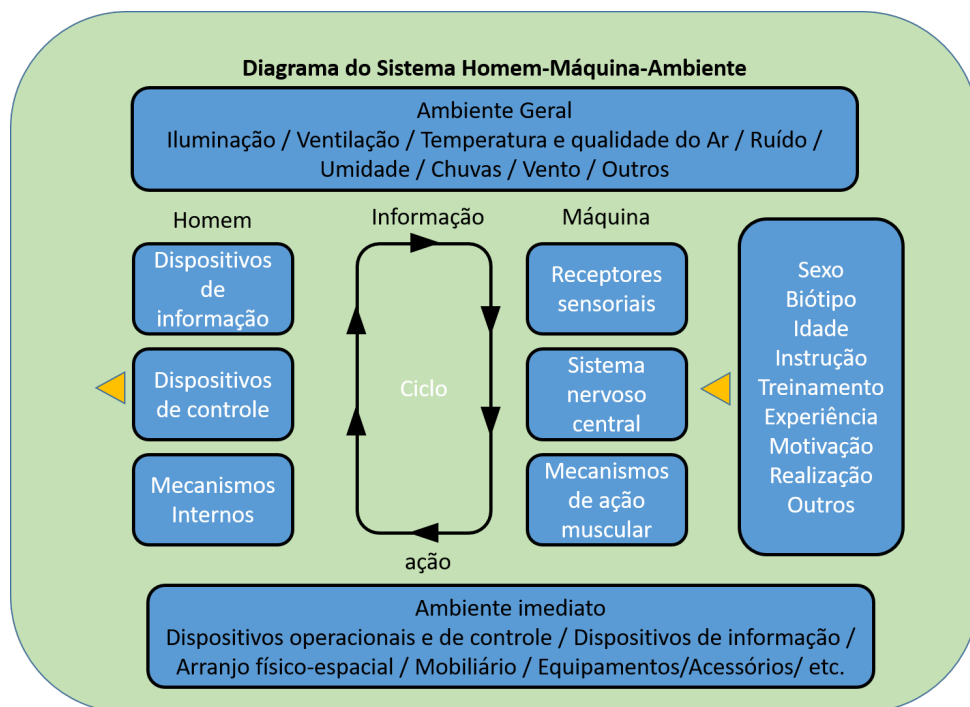


Figura 3.17: Diagrama interação Homem Máquina
 Adaptado de: (FILHO, 2010)

3.7 Conclusão

Neste capítulo apresentou-se o conjunto solução proposto para o problema de pesquisa, que se baseia em supostos de visão computacional, notadamente a área de reconhecimento

de padrões. Expôs-se os tipos de técnicas utilizados, suas particularidades e características que as tornam intrínsecas ao problema. Apresentamos o conjunto de hardware que será alvo do modelo.

Apontamos as dificuldades advindas em um sistema para o reconhecimento de padrões e do espaço físico necessário para ocorrer a mesma, destacando ainda a reação em tempo útil, com a formulação de uma métrica para determinar este lapso de tempo máximo aceitável em função de algumas propriedades do sistema. Apresentamos o diagrama básico da arquitetura ([Figura 3.14](#)) com suas características

Resultados e discussão

Neste capítulo é constituído, inicialmente, por uma seção que tem por objetivo apresentar a metodologia utilizada na realização dos testes práticos, com uma descrição detalhada que permita sua reprodução. Caracterizamos os cenários de teste, com seus objetivos e para finalizar, os resultados alcançados serão apresentados em conjunto com uma breve interpretação sobre seu significado e significância para validação do modelo.

4.1 Preliminares

Inicialmente destacaremos os elementos que nortearam a obtenção dos resultados. Descrevendo detalhadamente o conjunto de treino usado em cada técnica, os parâmetros alterados e em particular nossas escolhas. Destacamos, também, os elementos e fórmulas necessárias para análise das técnicas e conseqüentemente validação do modelo, nos limites de pesquisa escolhidos. Apresentamos também os resultados obtidos dos testes realizados, quando da integração das técnicas e discutimos estes ao longo da seção.

4.1.1 Tempo de resposta do modelo

No processamento de imagem o tempo de resposta necessário para a execução da maioria dos algoritmos como também seu envio ao hardware é crucial. Principalmente em se tratando da locomoção de indivíduos, a complexidade pode ser maior ou menor, sendo que em alguns casos inviabilizam a sua utilização em tempo real. Temos assim, que levar em conta que a simplificação destes processos tem de ser uma das diretivas principais do projeto. Essa simplificação contudo, não pode ser feita de maneira a acarretar perda de precisão dos resultados obtidos.

Assim, um compromisso entre tempo e a qualidade da resposta do modelo é uma das essências para o sucesso do modelo.

Como métrica de valor aceitável(BRAGA, 2012),(REIS *et al.*, 2009) podemos considerar as seguintes variáveis:

V_{max} - Velocidade máxima da cadeira em condições de segurança;

D_s - Distância máxima que a cadeira pode percorrer sem controle do paciente num certo limite de tolerância;

T_a - Tempo de resposta dos atuadores mecânicos/eletrônicos quando a cadeira se desloca a velocidade máxima;

T_e - Tempo que o paciente leva para, a partir da expressão, realizar uma ação;

T_{ri} - Tempo médio de reação do paciente entre a ocorrência de um elemento imprevisto;

De posse das variáveis o tempo de resposta será dado por:

$$T_r = \frac{D_{seg}}{V_{max}} - T_a - T_e - T_{ri} \quad (4.1)$$

Considerando que a cadeira terá como $V_{max} = 0.5m/s$ e $D_s = 1ms$ podemos estabelecer que como tempos ideais $T_a = 0.5s$, $T_e = 0.1s$ e $T_{ri} = 0.5s$ teremos como resposta $T_r = 0.9$.

Para o cenário definido o tempo máximo de processamento associado a um comando seria de $0.9s$. Em fato nosso propósito é que esta observação seja a menor possível. Devido ao altíssimo grau de impossibilidades que acomete o paciente com tetraplegia o ambiente no qual a cadeira está inserida deve possuir condições de circulação ideais, este fator caso não seja verificado é limitante ao modelo.

4.1.2 Modelo Matemático da CR

O modelo descreve a posição do veículo ao longo do tempo em função das velocidades, geometria e tipo de roda. Uma CR pode ser assim modelada com um robô não-holonômico de duas rodas traseiras motorizadas e duas dianteiras não motorizadas, a qual tem os movimentos governados pelas equações:

$$\begin{cases} \dot{x} &= v \cos \theta \\ \dot{y} &= v \sin \theta \\ \dot{\theta} &= \omega \end{cases} \quad (4.2)$$

Em que v é um vetor de entrada no qual $v = [v, \omega]^T$ e:

v : representa a velocidade linear da cadeira;

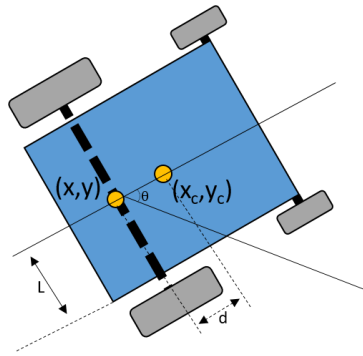


Figura 4.1: Modelo matemático da cadeira

Fonte: Elaborada pelo autor

θ : a velocidade angular da cadeira;

(x_c, y_c) : as coordenadas do centro do eixo da CR;

(x, y) : as coordenadas do centro da roda localizado a uma distância d .

ω : é o ângulo de rotação da cadeira em torno do seu eixo;

Podemos descrever a [Equação 4.2](#) em sua forma matricial da seguinte forma:

$$\dot{q} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -d \sin(\theta) \\ \sin(\theta) & d \cos(\theta) \end{pmatrix} \begin{pmatrix} v \\ \omega \end{pmatrix} \quad (4.3)$$

que pode ser simplificada em:

$$\dot{q} = Jv \quad (4.4)$$

Considerando a matriz J sempre inversível e $d \neq 0$, podemos construir o seguinte modelo linear:

$$\dot{q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.5)$$

Assim, qualquer vetor de entrada u pode ser mapeado a um vetor de entrada v . De posse destas informações, criamos em MatLab o software, cuja saída é representada na [Figura 4.2](#), que utiliza as informações do software NaveChair e as interpreta similarmente ao protótipo físico, com as restrições inerentes a um modelo puramente matemático, como deslizamento, choques, irregularidades do piso, elevações etc.

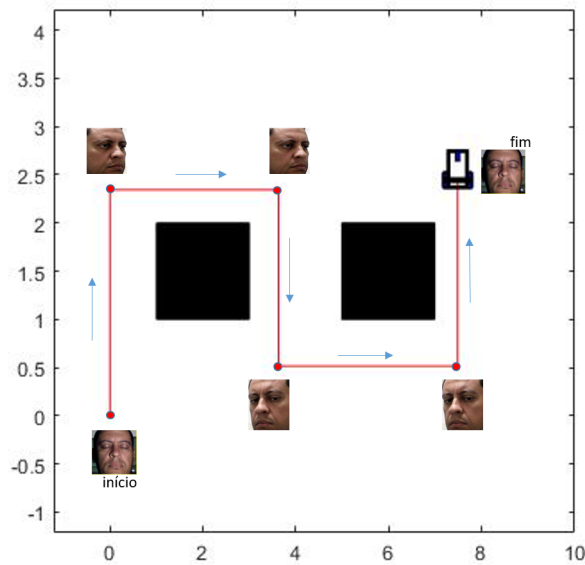


Figura 4.2: Saída do simulador, em Matlab, e suas entrada obtidas do software c++
 Fonte: Elaborada pelo autor

4.1.3 Gráfico ROC

Um *receiver operating characteristic* (ROC) ou curva ROC, é uma representação gráfica que ilustra o desempenho de um sistema de classificação binário como o seu limiar de discriminação é variada. A curva é traçada a partir da taxa de verdadeiros positivos (VR) contra a taxa de falsos positivos (FP).

A taxa de verdadeiros positivos também é conhecida como sensibilidade, ou seja, a capacidade do modelo classificar um evento. A taxa de falsos positivos também é conhecida como especificidade que avalia a capacidade do modelo prever um indivíduo como um evento dado que ele realmente não o é.

A análise ROC fornece ferramentas para selecionar modelos possivelmente ótimos e descartar subótimos independentemente (e antes de especificar) o contexto de custo ou a distribuição de classe. Análise ROC está relacionada de forma direta e natural para análise de custo / benefício de diagnóstico na tomada de decisão.

Para nossa avaliação temos as métricas Acurácia, Precisão, Sensibilidade Especificidade, com a seguinte matriz de confusão descrita na tabela [Tabela ??](#) e com resultados previstos na [Tabela 4.2](#) (SPIEGEL *et al.*, 2004).

A matriz de confusão nos permite determinar que as seguintes situações podem ocorrer:

Valor Predito	Valor Observado(Valor Verdadeiro)		
		Y=1	Y=0
	Y=1	VP(Valor Positivo)	FP(Falso Positivo)
Y=0	FN(Falso Negativo)	VN(Verdadeiro Negativo)	

Tabela 4.1: Matriz de confusão para construção do gráfico ROC

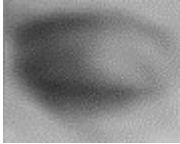
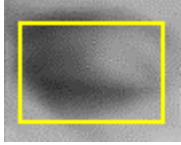
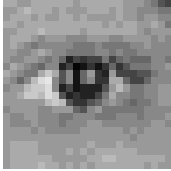
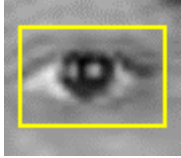


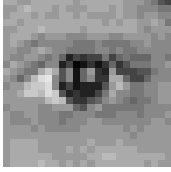
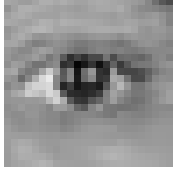
Análise	Imagem de Teste	Resultado do detector
Verdadeiro Positivo	 <p>Imagem de teste com o elemento de detecção presente</p>	 <p>Detector encontra o objeto</p>
Falso Positivo	 <p>Imagem de teste não possui o objeto de busca</p>	 <p>Detector localiza um falso objeto.</p>
Falso Negativo	 <p>Imagem de teste possui o objeto de detecção</p>	 <p>Detector não encontra o objeto de detecção</p>
Verdadeiro Negativo	 <p>Objeto de detecção não está presente</p>	 <p>Detector não localiza nenhum objeto na imagem</p>

Tabela 4.2: Possibilidades para o detector após o treinamento

4.1.3.1 Acurácia

Representa a porcentagem de amostras positivas e negativas classificadas corretamente sobre a soma de amostras positivas e negativas.

$$A = \frac{VP + VN}{VP + FN + FP + VN} \tag{4.6}$$

4.1.3.2 *Sensibilidade*

Porcentagem de amostras positivas classificadas corretamente sobre o total de amostras positivas

$$P = \frac{VP}{VP + FN} \quad (4.7)$$

4.1.3.3 *Precisão*

Porcentagem de amostras positivas classificadas corretamente sobre o total de amostras classificadas como positivas.

$$P = \frac{VP}{VP + FP} \quad (4.8)$$

4.1.3.4 *Especificidade*

Porcentagem de amostras negativas identificadas corretamente sobre o total de amostras negativas.

$$P = \frac{VN}{VN + FP} \quad (4.9)$$

4.1.4 *Conjunto de positivos para o classificador em cascata*

É um conjunto de imagens que contém o objeto de interesse. Neste trabalho utilizamos um conjunto de dados obtidos pelo autor com 10000 imagens, das quais apresentamos na

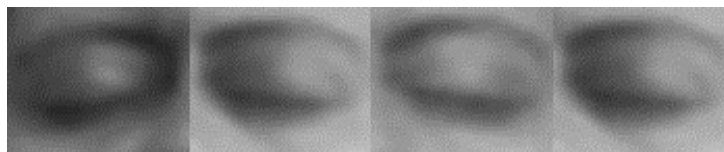


Figura 4.3: Conjunto de imagens positivas com dimensão de 24×24 pixel

Fonte: Elaborada pelo autor

4.1.5 *Conjunto de negativas para o classificador em cascata*

Estas se opõem completamente a qualquer imagem do grupo de positivas. Todas com dimensões variadas e maiores que as positivas. Uma amostra desses dados é apresentado na [Figura 4.4](#)



Figura 4.4: Conjunto de imagens negativas com dimensão variadas

Fonte: Elaborada pelo autor

4.1.6 Samples

Samples é uma imagem na qual o fundo é uma imagem negativa com imagens positivas incorporada em posições aleatórias sobre ele. É representado por um arquivo VEC. O número e a dimensão do conjunto de amostras desempenha um papel importante no processo de formação.

4.1.7 Parâmetros de Treinamento

Parâmetros definidos para a formação do classificador. Em suma temos 11 parâmetros com variadas possibilidades de configuração, o que torna a definição de quais utilizar um processo laborioso e lento, são eles:

1. Número de positivos usados (nPos);
2. Número de negativas usadas (nNeg);
3. Número de estágios (nStage) - Representa o número de etapas do classificador em cascata que pode variar de 1 a $n > 1$. O equilíbrio entre número de estágios e tempo de processamento depende necessariamente do grau de precisão que se deseja ao classificador;
4. Tipo de característica (Ftype) - Em nosso trabalho usamos características de Haar, mas é possível usar na classificação, as técnicas: Gradientes de Histogramas(HOG), padrões binários locais (LPB);
5. Dimensões das amostras;
6. Tipo de Boosted (Btype) - Podemos usar AdaBoost Discreto, Read AdaBoost (RAB), LogitBoost (LB) e Gentle AdaBoost (GAB);
7. Taxa de acerto mínimo(MHRate) - Definido como a razão entre o número de objetos detectados para um total de objetos presentes;

8. Taxa de falsos positivos (MFRate) taxa de falsos alarmes decide o número máximo permitido de falsas detecções presentes no processo de formação;
9. Peso (WTRate) - Após estabelecido, depois de cada fase da cascata as amostras com o menor peso são ignorados;
10. Contador fraco (Wcount) - Número máximo de árvores fracas envolvidas em todas as etapas do treinamento;
11. Memória que está sendo usada para armazenar as características e mais tarde os parâmetros do classificado, em Mb.

Dos parâmetros descritos, realizamos modificações nos seguintes:

Tabela 4.3: Parâmetros que podem ser alterados nos classificadores

Parâmetros	Classificador 1	Classificador 2	Classificador 3	Classificador 4
MinHitRate	[0.9, 0.999]	[0.9, 0.999]	[0.9, 0.999]	[0.9, 0.999]
MaxFalseAlarm	[0.1, 0.5]	[0.1, 0.5]	[0.1, 0.5]	[0.1, 0.5]
Nstages	[20, 50]	[20, 50]	[20, 50]	[20, 50]
Btype	GAB	RAB	LB	DAB
WTRate	[0.95, 0.98]	[0.95, 0.98]	[0.95, 0.98]	[0.95, 0.98]
Wcount	100	100	100	100

Definidos os parâmetros, podemos obter um conjunto de informações acerca do classificador treinado com a análise das seguintes características:

1. Tempo de treinamento (TT): Apresenta o tempo necessário para o classificador ser treinado;
2. O tempo de detecção (DT): O tempo necessário para o classificador treinado para detectar olhos fechados na imagem.

4.1.8 Conjunto de treino do ASM

Para o treinamento do ASM utilizamos 6000 imagens rotadas utilizando 150 faces com uma rotação da cabeça entre $0 \leq \theta \leq 80$, com 68 pontos característicos, [Figura 4.5](#). também incluímos amostras de pessoas com tom de pele claro e escuro, com algum tipo de oclusão, cabelos curtos e longos, com e sem óculos e com iluminação controlada e constante utilizando a base de dados AR. A AR face database ([MARTÍNEZ; BENAVENTE, 1998](#)) contém 4000 imagens em escala RGB com resolução de 768×576 de 126 faces. A base AR serve também para análise da resposta da técnica e neste caso selecionamos 50 faces aleatórias deste banco para determinar se o detector encontra o objeto desejado. Esta

quantidade é somada ao nosso banco que foram retiradas, do conjunto de treino, a mesma quantidade de imagens.



Figura 4.5: Parcela do conjunto de imagens utilizado para treinar o ASM

Fonte: Elaborada pelo autor

Caso o ângulo de rotação em torno do eixo y seja superior a 20 graus temos um giro da cabeça para a direita e inferior a 20 graus para a esquerda. Como nosso modelo integra o ASM com o coeficiente de correlação de Pearson devemos escolher quais as imagens serão definidas como templates. Na Figura 4.6 definimos que $L1, L2, L3$ estabelecem que o paciente está olhando para a esquerda, C define olhando para o centro e $R3, R2, R1$ olhando para a direita. Como o usuário deve permanecer por mais de 2 segundos nesta posição o modelo tem a liberdade de escolher uma correlação mais forte entre três opções. A posição central, por não ser acionadora de nenhuma comando, foi utilizada apenas para efeito de comparação do estado da cabeça.

Ao estabelecer o comando, este é enviado a cadeira que aciona a motorização com um giro de até 90° em seu eixo. Este processo é repetido até o usuário encerrar o movimento e retornar a cabeça à posição central.

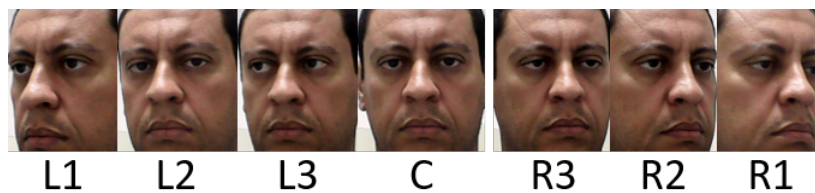


Figura 4.6: Rotações da face utilizadas como template $T(x, y)$

Fonte: Elaborada pelo autor

Dada uma imagem $I(x, y)$ e um template $T(x, y)$ com as mesmas dimensões, a correlação normalizada entre eles é definida na Equação 4.10 com base na equação 3.13.

$$VC(I, T) = \frac{\sum_{x,y} (I - \bar{I}) (T - \bar{T})}{\sqrt{\sum_{x,y} (I - \bar{I})^2 \sum_{x,y} (T - \bar{T})^2}} \quad (4.10)$$

em que, $\bar{I} = \frac{1}{n} \sum_{x,y} I$ e $\bar{T} = \frac{1}{n} \sum_{x,y} T$

A estimativa da pose da cabeça humana revelou-se um tema desafiador, muito relacionada ao processo de interação homem-máquina, à detecção de fadiga e a produtos recreativos entre outros. Em nosso escopo buscamos um modelo estável que retorne cerca de 98% de acerto na detecção da posição da cabeça. Basicamente a pose pode ser descrita em três dimensões, descritas na [Figura 4.7](#). Utilizamos somente a rotação em torno do eixo $z = Yaw$

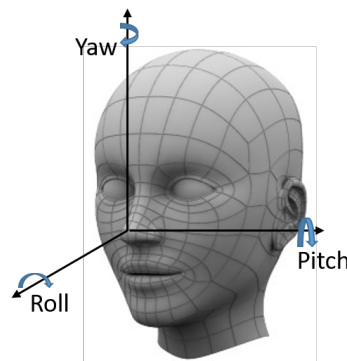


Figura 4.7: Rotações da face em 3 dimensões

Fonte: Elaborada pelo autor

4.1.9 Software que implementa o modelo

Para [Furtado \(2002\)](#) e [Gonzalez *et al.* \(2004\)](#) a simulação, gerada pelo software que implementa um modelo computacional, é um processo que requer planejar um sistema real e coordenar experimentos com a finalidade de entender seu desempenho e/ou determinar estratégias para sua ação.

O software que implementa o modelo computacional, [Figura 4.8](#), descrito na [Figura 3.14](#), chamado NaveChair foi implementado em C++ utilizando a biblioteca OpenCV para as rotinas de processamento de imagens. O software não possui interface gráfica, pois será implementado direto na cadeira, que não dispõe de equipamentos de saída de informações. O software também se comunica a um modelo computacional de CR implementado em MatLab, devido a sua integração com C++ e rápida interpretação das informações.

acima, com um número satisfatório de recursos. Para este problema de detecção foram utilizados quatro detectores que foram construídos variando os seus parâmetros conforme descrito na [Tabela 4.3](#), em que: MinHitRate representa a taxa de acerto mínimo desejado para cada fase do classificador; MaxFalseAlarm é a taxa de alarme falso máximo desejado para cada fase do classificador; Nstages é o número de estágios da cascata; Btype representa o tipo de boost utilizado (Type of boosted classifiers: DAB - Discrete AdaBoost, RAB - Real AdaBoost, LB - LogitBoost, GAB - Gentle AdaBoost.); WTRate é a linha de corte e o peso utilizado no boost; Wcount -contagem máxima de árvores falsas para todas as fases da cascata.

Tabela 4.4: Apresentação dos parâmetros e dos classificadores

Parâmetros	Classificador 1	Classificador 2	Classificador 3	Classificador 4
MinHitRate	0.999	0.987	0.985	0.999
MaxFalseAlarm	0.5	0.5	0.4	0.5
Nstages	20	25	20	25
Btype	GAB	RAB	LB	DAB
WTRate	0.97	0.98	0.95	0.95
Wcount	100	100	100	100
Tempo processamento	4 dias	5 dias	4 dias	7 dias

Os parâmetros escolhidos na [Tabela 4.4](#), são ditos os melhores, pois obtiveram uma taxa de positivos acima de 96%. As demais variações destes parâmetros obtiveram retorno de positivos e ou falso positivos fora do limiar desta pesquisa e então foram descartados. Dos classificadores que treinamos, quatro obtiveram taxas dentro do limiar esperado e dentre eles a [Figura 4.9](#) apresenta a comparação entre eles.

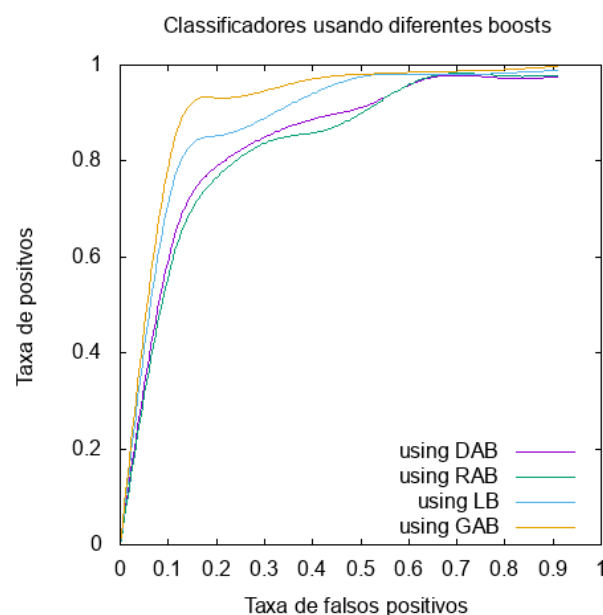


Figura 4.9: Gráfico ROC com os melhores classificadores

Fonte: Elaborada pelo autor

A curva que utiliza o boost GAB apresentou uma taxa de falsos positivos mais baixa (cerca de 1.5%) entre os classificadores e a taxa de positivos acima de 97%, utilizando como base de testes 100 imagens da base AR e nossa base. Este resultado permitiu a escolha do classificador 1 como ideal para implementação no modelo.

Utilizamos duas formas de análise do classificador criado. A primeira utilizou imagens da base AR e imagens diversas da web que continham pessoas com olhos fechados. Em seguida utilizamos vídeos diversos que continham pessoas abrindo e fechando os olhos além de vídeos do próprio autor. Todas as imagens e vídeos tinha boas condições de luz (acima de 200 Lux) e cada vídeo tinha, em média, 1 minuto, operando com quadros de tamanho 640×320 , com 23 FPS. Os vídeos e as imagens utilizadas para os testes apresentam ligeiros elementos de oclusão e iluminação assim realizamos um teste em separado que compara a resposta do detector à imagens com e sem estes elementos, que está representado na [Figura 4.10](#). A importância deste teste reside no fato de alguns pacientes usarem óculos, mas ressaltamos que o modelo não se adapta a óculos de sol que tenham bloqueio total dos olhos.

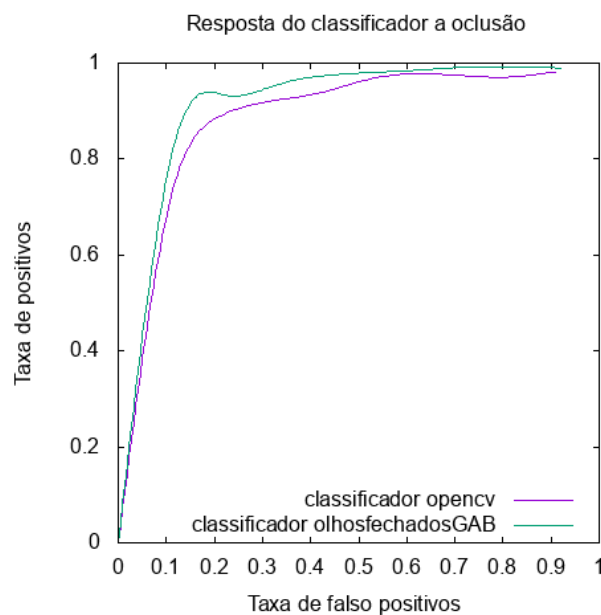


Figura 4.10: Comparação da resposta do classificador a imagens com ou sem oclusão

Fonte: Elaborada pelo autor

A direção da câmera não sofreu problemas com trepidação para o teste em questão, mas prevemos a possibilidade de tremulação da imagem conforme as ideias de ([MATSUSHITA et al., 2006](#)) e ([GRUNDMANN et al., 2011](#)).

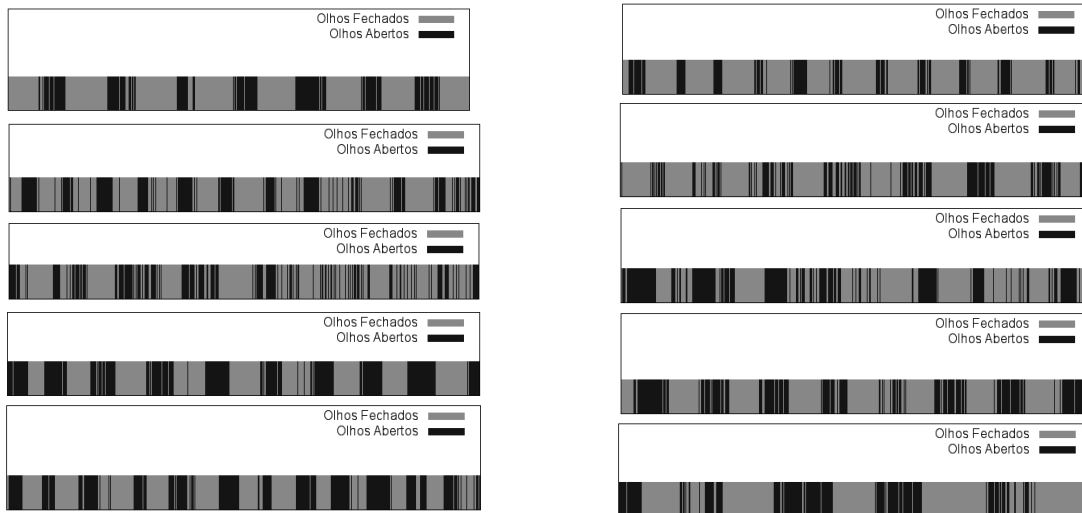


Figura 4.11: Leitura do estado do olho tomado com 10 amostras

Fonte: Elaborada pelo autor

A Figura 4.11 utilizou um fluxo de vídeo com 1min cada para 30 indivíduos diferentes revelando uma taxa de acerto da ordem de 98%, na qual as faixas cinza claro indicam este sucesso e as pretas a não detecção do estado fechado ou seja aberto. O modelo lê um frame por vez e determina o estado do olho naquela imagem, ou seja, detecta uma intensidade de luz similar a que foi treinada. Durante o processo de fechamento dos olhos o modelo terá de tratar o limiar entre o fechamento e a abertura dos mesmos.

Apresentamos também o resultado da acurácia do classificador treinado na Figura 4.12, este resultado apresenta o comportamento dos positivos em reação aos falsos positivos e nos permitiu avaliar a precisão do classificador.

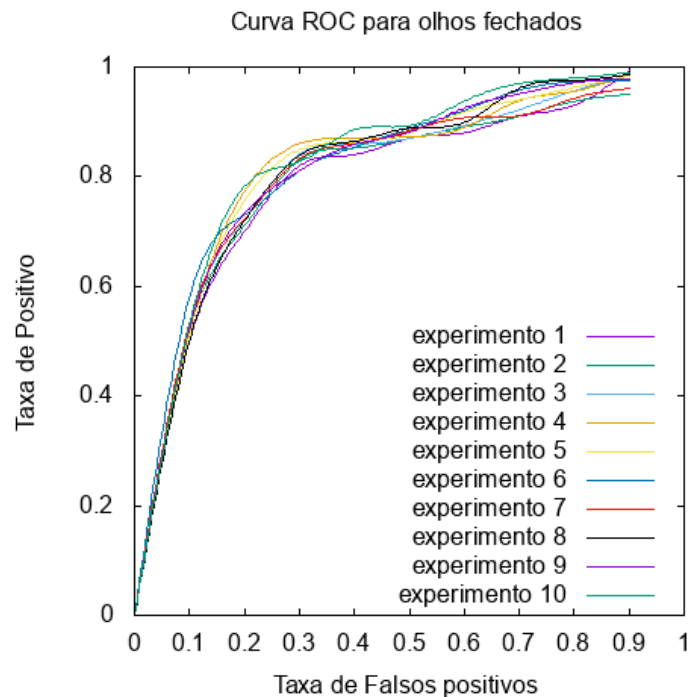


Figura 4.12: Gráfico ROC das amostras

4.3 Resultados dos testes no Detector dos Movimentos da cabeça

Treinamos o detector com 150 imagens uma rotação da cabeça entre $0 \leq \theta \leq 80$, com 68 pontos característicos, que foram marcados manualmente seguindo a ordem da [Figura 4.13](#).

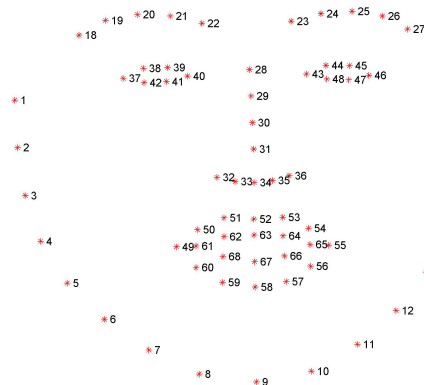


Figura 4.13: Exemplo de pontos marcados na posição central

Fonte: Elaborada pelo autor

Para realização dos testes, as imagens obtidas incluíram pessoas com tom de pele claro e escuro, com algum tipo de oclusão, cabelos curtos e longos sem controle de iluminação. As imagens em vídeo, obtidas, possuem uma resolução 800×600 pixels, que representa basicamente os tipos de câmera mais comuns. Estas imagens passam pelo mesmo pré-

processamentos, descrito em [section 3.1](#).

A seleção dos templates que servirão de base para a comparação da técnica

$T(x, y)$, Na [Figura 4.14](#) apresentamos o primeiro teste realizado e os resultados obtidos.

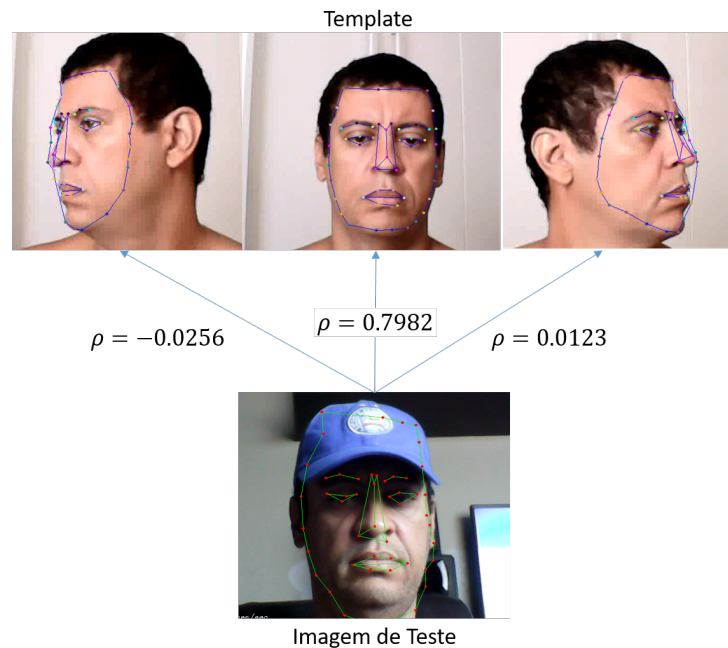


Figura 4.14: Teste inicial realizado com uma única imagem avaliando o módulo

Fonte: Elaborada pelo autor

Caso o ângulo de rotação em torno do eixo y seja superior a 20 graus temos um giro da cabeça para a direita e inferior a 20 graus para a esquerda. Estes dados são enviados a uma cadeira que aciona a motorização bem como realiza um giro de até 90° em seu eixo. Este processo é repetido até o usuário encerrar o movimento e retornar a cabeça à posição central.

Após treinamento e carregamento dos templates no software utilizamos 9000 imagens disponíveis na web das quais selecionamos três conjuntos (CL,CC,CD) contendo 3000 imagens cada, com rotações a esquerda, direita e centro. As imagens possuem diversas resoluções e todas no padrão de cor RGB e com indivíduos de diversas idades, sexos e etnia além de algumas com oclusão básica (óculos e bonés). A ideia do teste é verificar o nível de falsos positivos e determinar a acurácia e precisão do detector. Os resultados são apresentados na [Figura 4.17](#), [Figura 4.16](#), [Figura 4.15](#).

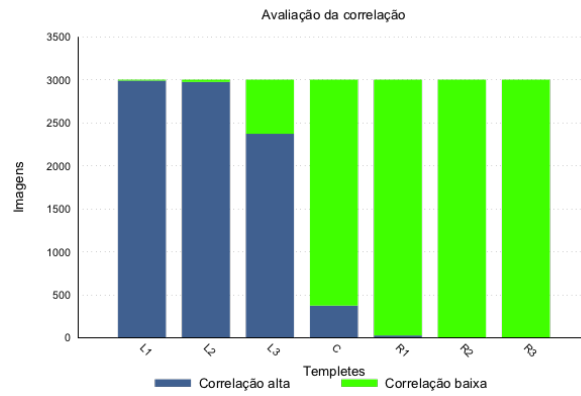


Figura 4.15: Comparação do conjunto CL com os templates
 Fonte: Elaborada pelo autor

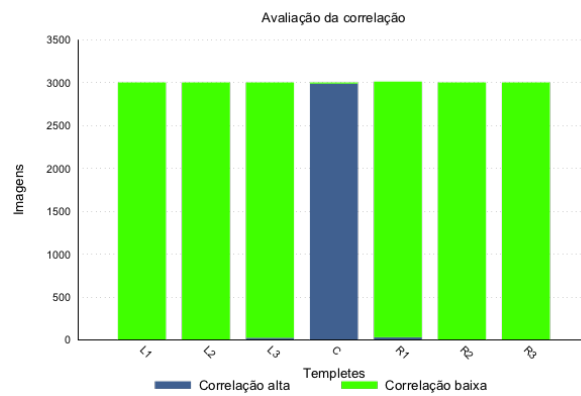


Figura 4.16: Comparação do conjunto CC com os templates
 Fonte: Elaborada pelo autor

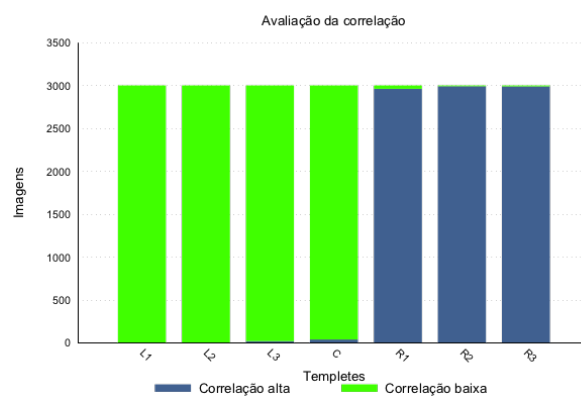


Figura 4.17: Comparação do conjunto CD com os templates
 Fonte: Elaborada pelo autor

Na [Figura 4.15](#) verificamos que a acurácia chegou a casa do 99.1% com um número de falsos positivos em torno de 3% em cada amostra. Observamos que os templates L3 e C estão com a correlação muito próxima o que causa um aumento dos falsos positivos.

Na [Figura 4.16](#) temos a melhor comparação, a posição central retorna $A = 99,8\%$ e com um número de falsos positivos de menos de 1%. Diferente do que obtivemos nos testes anteriores a detecção do giro da cabeça a direita, [Figura 4.17](#), retornou 99,6% de acerto e com taxa de falsos positivos também abaixo de 1%. Assim na média temos que a técnica apresentada tem uma boa acurácia $A = 99.5\%$ com a taxa de falso positivos $< 1\%$ e $P = 99.8\%$

Analizamos, em paralelo, o comportamento da técnica ao ser utilizada uma sequência de imagens. A comparação anterior usa 9000 imagens no total mas em termos de vídeo temos um fluxo de 23 frames por segundo. Este, portanto, torna-se uma representação mais robusta, em termos dos dados de entrada. As imagens foram obtidas por uma webCam comum a uma distância fixa da câmera de aproximadamente 80cm com iluminação constante e os resultados são apresentados na [Figura 4.18](#), [Figura 4.19](#), [Figura 4.20](#), [Figura 4.21](#).

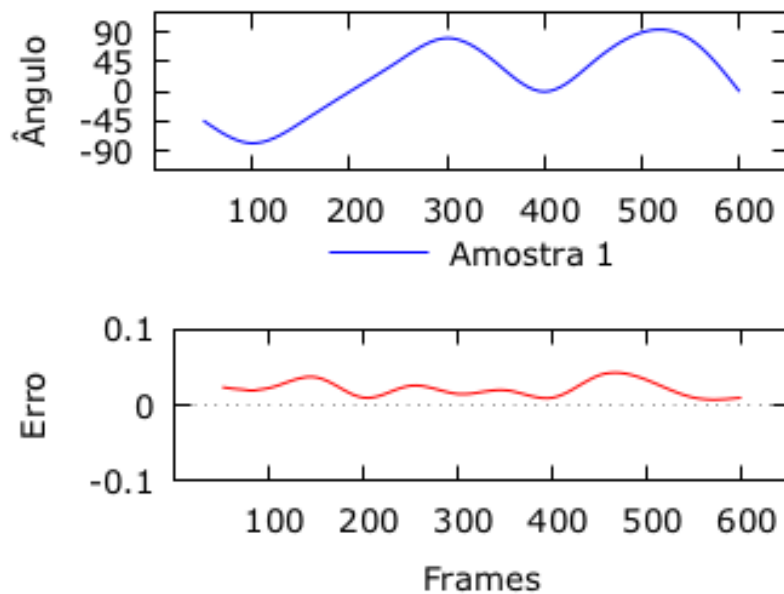


Figura 4.18: Rotação da cabeça amostra 1

Fonte: Elaborada pelo autor

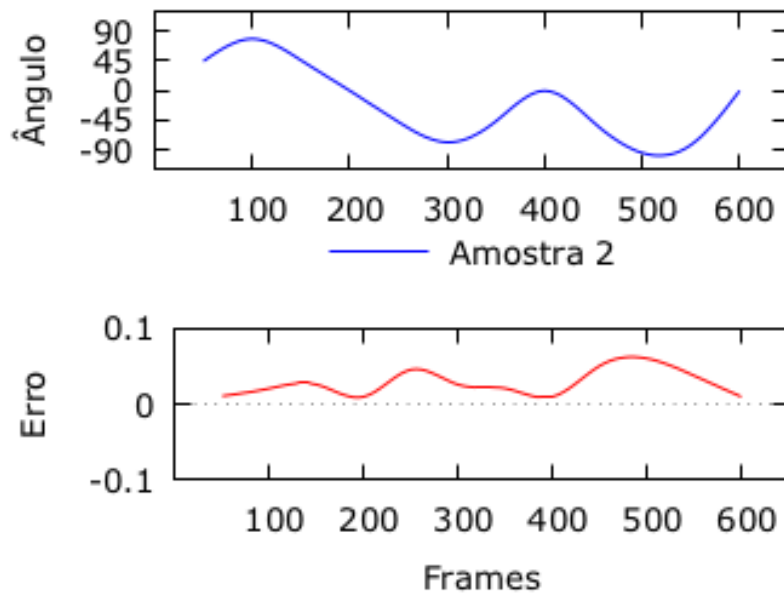


Figura 4.19: Rotação da cabeça amostra 2
 Fonte: Elaborada pelo autor

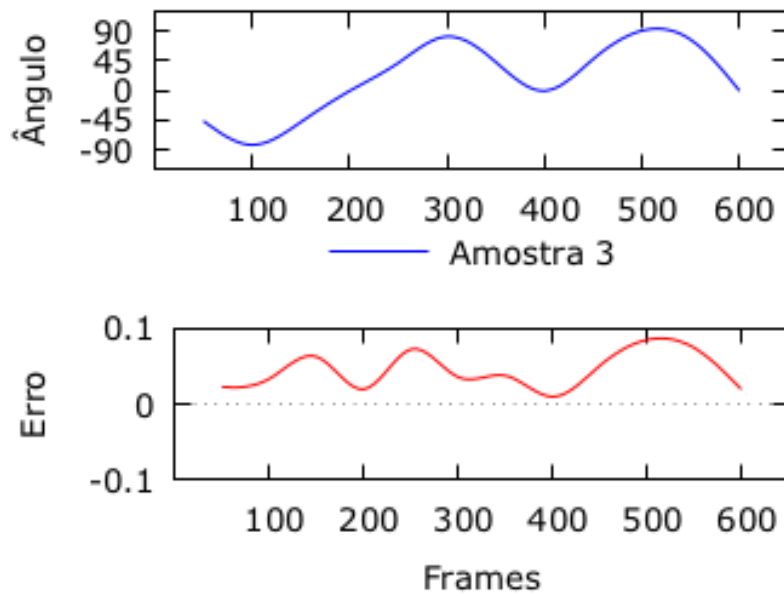


Figura 4.20: Rotação da cabeça amostra 3
 Fonte: Elaborada pelo autor

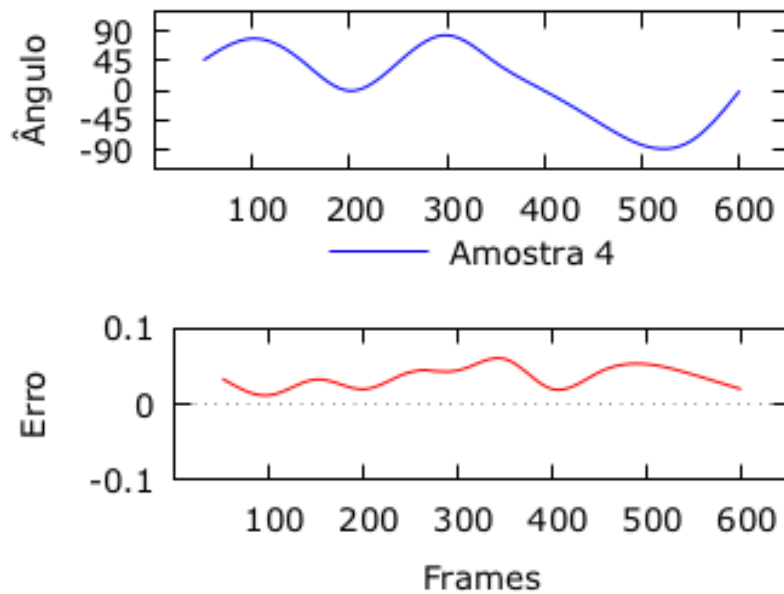


Figura 4.21: Rotação da cabeça amostra 4
Fonte: Elaborada pelo autor

4.4 Resultados da integração das técnicas

Algumas experiências foram realizadas num ambiente simulado, utilizando o MatLab[®]. A cadeira deveria seguir uma trajetória, sem choques com as paredes e sempre para a área definida. A coleta destas informações foi manual e para cada cenário foram realizadas 100 testes, neste experimento testamos:

- Acurácia;
- Precisão;
- Taxa de falso positivos.

Na [Figura 4.22](#), aprestamos o funcionamento do simulador e como ele reage aos comandos, que são gerados pelo mesmo software que usamos para a comunicação com a cadeira física.

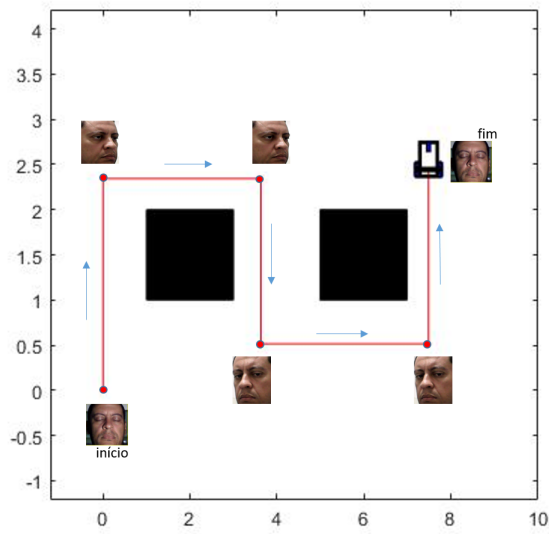


Figura 4.22: Comportamento da navegação no simulador

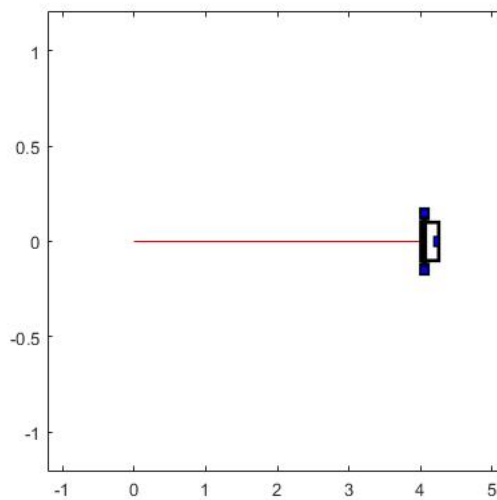


Figura 4.23: Caminho percorrido pela cadeira: cenário 1

No cenário da [Figura 4.23](#) o indivíduo deveria parar a cadeira no intervalo do eixo das abscissas entre $[3, 4]$, e seu início se dá em $x = 0$, a velocidade da cadeira simulada é 1km/h que é reduzida para 0.5km/h quando detectadas rotações.

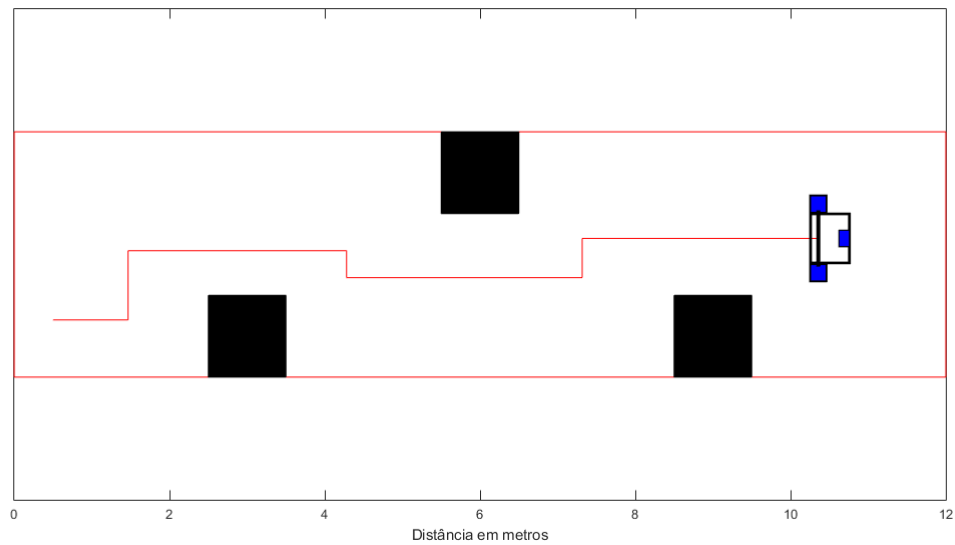


Figura 4.24: Caminho percorrido pela cadeira: cenário 2

No cenário da [Figura 4.24](#) o indivíduo deveria parar a navegação a cadeira até $x = 10$ desviando dos objetos no percurso. O cenário apresenta a possibilidade do paciente navegar em corredores sem a utilização de sensores para desvio de obstáculos.

No cenário 3, representado na [Figura 4.25](#), o paciente realiza a navegação entre dois objetos que estão a 2m de distância e no cenário 4 este desafio é amplificado ao realizar o caminho entre cômodos de uma planta baixa. Neste cenário, representado na [Figura 4.26](#), foram realizados testes nos qual a cadeira desloca-se de diversos pontos de saída e chegada.

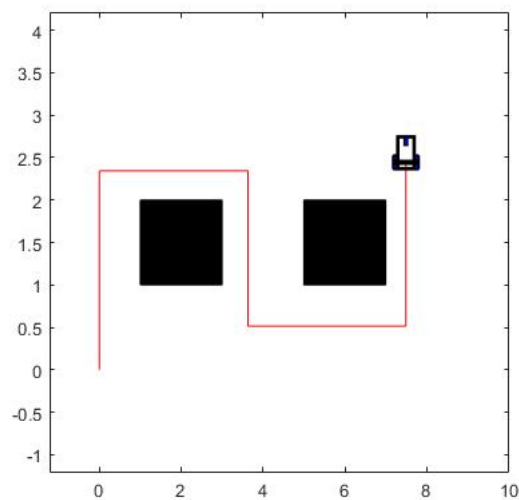


Figura 4.25: Caminho percorrido pela cadeira: cenário 3

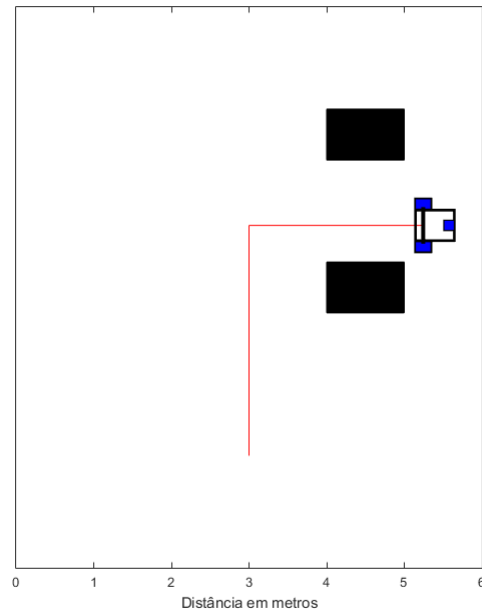


Figura 4.28: Caminho percorrido pela cadeira: cenário 6

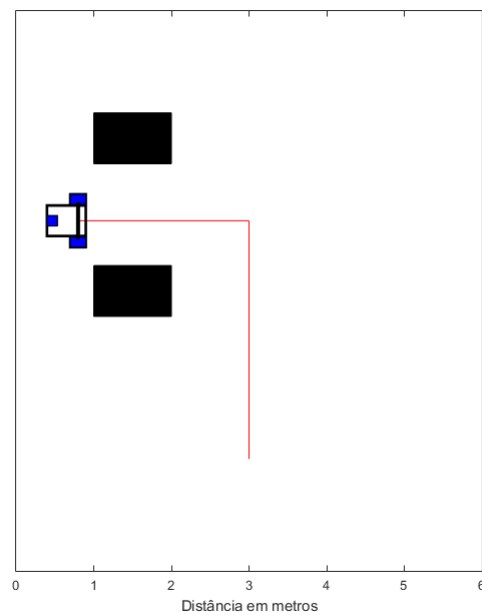


Figura 4.29: Caminho percorrido pela cadeira: cenário 7

Na [Tabela 4.5](#) apresentamos os resultados obtidos, de maneira geral, nas 100 simulações realizadas.

As condições de iluminação foram controladas e para captação dos gestos utilizou-se uma webcam padrão com resolução de 800×600 a taxa de 20 quadros por segundo.

Tabela 4.5: Resultados geral dos testes realizados

Cenário	Falso Positivos	Precisão	Acurácia
1	1.0%	99.01%	97.6%
2	2.0%	98.03%	96.29%
3	1.0%	99.01%	97.6%
4	2.0%	98.03%	96.29%
5	1.0%	99.01%	97.6%
6	1.0%	99.01%	97.6%
7	1.0%	99.01%	97.6%

Os testes com o protótipo foram realizados, nas salas e corredores da Faculdade e as informações da posição da cadeira, ao longo do tempo, foram obtidas por um acelerômetro acoplado a mesma. Os cenários foram organizados sob condições de iluminação constante e realizados sem a presença de indivíduos na cadeira propriamente dita.

Em todos os casos, o caminho percorrido foi comparado ao deslocamento realizado pela cadeira utilizando o padrão de navegação da cadeira, seu *joystick*. A CR não dispõe de hardware ou implementação para o desvio de obstáculo, cabendo ao usuário efetuar a correção da rota durante a navegação com o movimento da cabeça.

No primeiro cenário, foi utilizado um corredor no qual o usuário deveria deslocar a cadeira. Sendo representado este deslocamento na [Figura 4.30](#). Na [Figura 4.31](#) utilizamos uma sala de aula, na qual o indivíduo deveria sair e chegar no mesmo ponto.

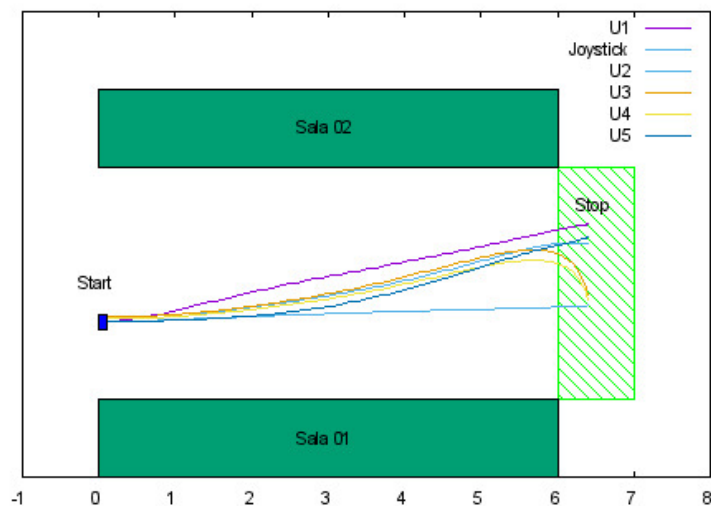


Figura 4.30: Cenário 1: deslocamento horizontal

Utilizando duas mesas, separadas por uma distância de 2m, o usuário é convidado a realizar o deslocamento do início até o ponto de parada, passando entre as duas e os resultados deste deslocamento são descritos na [Figura 4.32](#). Os testes representados na

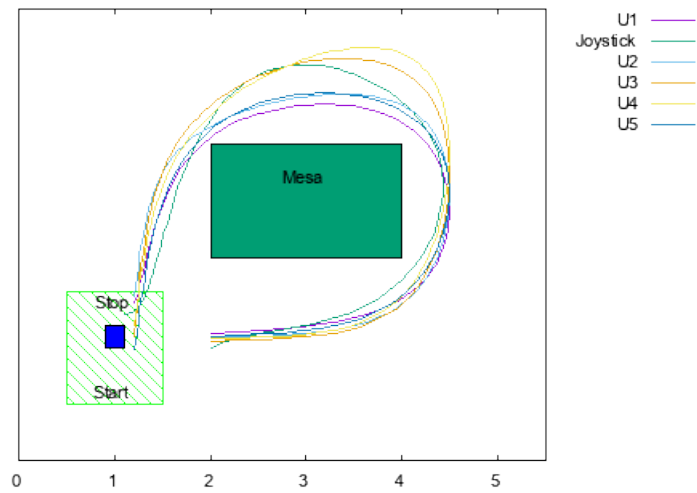


Figura 4.31: Cenário 2: deslocamento ao redor de um obstáculo

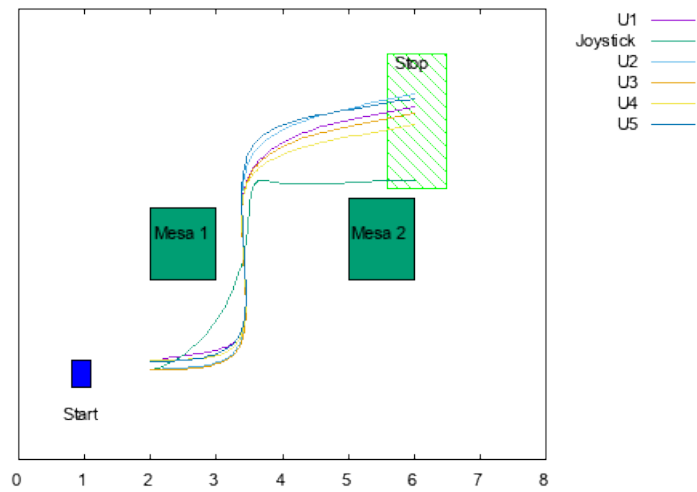


Figura 4.32: Cenário 3: deslocamento com desvio de obstáculo

Figura 4.33 e Figura 4.34 apresentam um contorno em L, muito comum em apartamentos e casas, para o deslocamento entre um cômodo e outro.

No último teste realizado, foram dispostos, ao longo de um corredor, três objetos que simulam obstáculos como vasos, cadeiras, entre outros. Os objetos estão a 2m de distância entre si e o percurso total percorrido foi de 10m, estes testes estão representados na Figura 4.35.

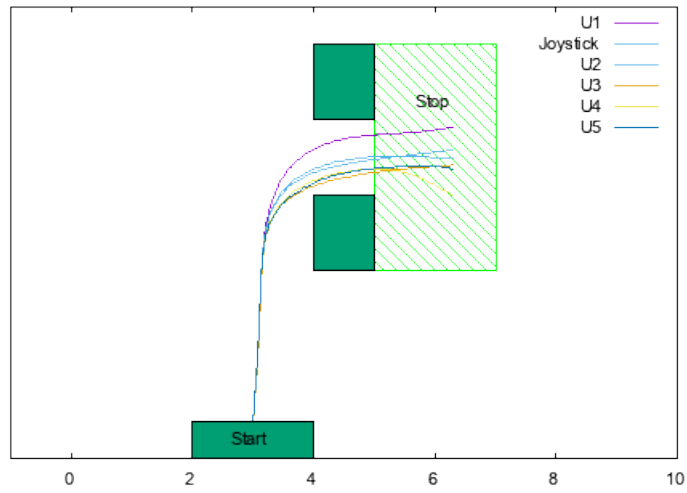


Figura 4.33: Cenário 4: deslocamento em L a esquerda

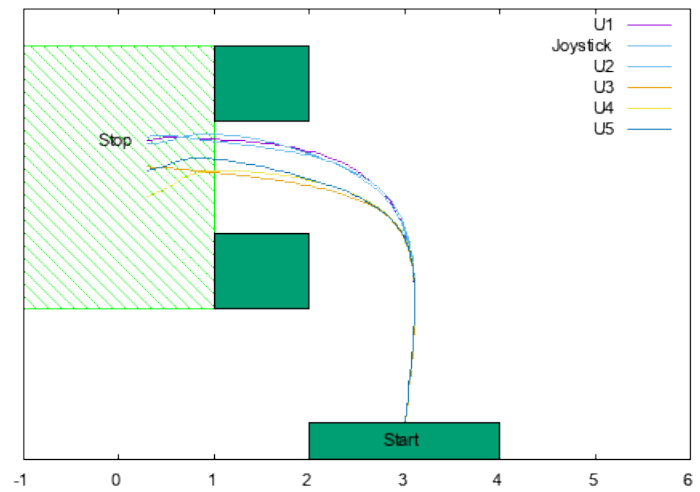


Figura 4.34: Cenário 5: deslocamento em L a direita

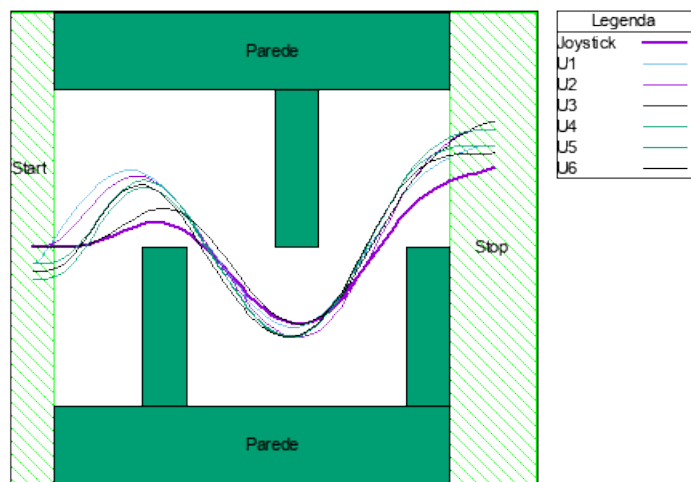


Figura 4.35: Cenário 5: deslocamento em L a direita

Nos testes realizados com a CR física ocorreram deslizamentos ao rotacionar em torno do seu eixo e até mesmo no movimento retilíneo. Esses deslizamentos ocorreram por fatores como a ausência de usuários na cadeira, o que a deixou muito leve, além de irregularidades no piso.

4.5 Análise dos Resultados

O modelo apresentado neste trabalho utiliza uma câmera de resolução 800 × 600 que mostrou-se eficiente na captação da face com a taxa de quadros constante. Comparativamente, Tabela 4.6, o modelo se caracteriza como HMI.

Tabela 4.6: Tabela Comparativa entre nosso projeto e alguns projetos de CR
 Fonte: Elaborada pelo autor

AUTOR	ENTRADAS											MODELOS										
	Sonar	Infravermelho	Scanner laser	Câmera	Sensor contato/colisão	Odometria	Bussola	GPS	Sensor acelerômetro	Giroscópios	Microfone	Joystick/Game Control	Touch screen	EOG	EMG	EEG	Acelerômetros	Movimento da Cabeça	Movimento dos olhos	Bio-Sinal	Controle por voz	
(DYMOND; POTTER, 1992)				✓					✓			✓						✓				
(ADACHI <i>et al.</i> , 1998)	✓	✓		✓					✓			✓						✓	✓			
(KUNO <i>et al.</i> , 2000.)	✓			✓					✓								✓	✓	✓	✓	✓	
(BAREA <i>et al.</i> , 2002)				✓								✓	✓					✓				
(TAYLOR; NGUYEN, 2003)	✓	✓		✓	✓				✓	✓	✓	✓	✓	✓				✓				
(PARIKH <i>et al.</i> , 2007)	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓					✓		✓	✓	
(URBANO <i>et al.</i> , 2008)	✓	✓	✓		✓	✓		✓	✓	✓		✓								✓		
(MANOGNA <i>et al.</i> , 2010)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓			✓				
(LEE; SAMARAS, 2011)	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓					✓	✓		✓		
(BRAGA, 2012)	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓										✓
(HALAWANI <i>et al.</i> , 2012)	✓			✓				✓	✓													✓
(BONARINI <i>et al.</i> , 2012)	✓	✓		✓												✓						
(PEREZ <i>et al.</i> , 2012)	✓	✓		✓				✓	✓				✓			✓		✓				
(CHACKO <i>et al.</i> , 2013)	✓										✓		✓						✓			
(JAIN; JOSHI, 2014)																				✓		
(PORTHIYAS <i>et al.</i> , 2014)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓								✓		
Nossa Abordagem				✓														✓	✓			

A detecção dos olhos fechados e da rotação da cabeça se revelou extremamente rápida, média de 0.02ms e portanto a resposta aos comandos é quase imediata. A avaliação de falsos positivos, que, em média, representam 3% para cada grupo de 100 imagens se revelou persistente.

Os testes com o protótipo revelou que o modelo tem boa adaptabilidade a diferentes condições de fundo, sexo, pigmentação da pele, idade e oclusão.

Observamos que a resposta do modelo à rotação de face foi bem satisfatória com um erro

Tabela 4.7: Comparativo entre nossa abordagem e os métodos de detecção de olhos fechados

Trabalho	Análise					
	Acurácia	Resposta (s)	Varição Escala	Varição Pose	Varição iluminação	Taxa de FP
Abordagem baseada no conhecimento						
Salehi <i>et al.</i> (2012)	97%	0.3	✓	✓	✓	N/A
Sebari & He (2013)	96%	0.1	✓	✓	✓	2%
Rautaray & Agrawal (2015)	97%	0.1	✓	✓	✓	2%
Murawski & Rózanowski (2013)	70%	0.3	✗	✗	✗	N/A
Lethaus <i>et al.</i> (2011)	93%	00.5	✓	✗	✗	3%
Wu & Trivedi (2010)	92.43%	N/A	✓	✓	✓	7.26%
Reinders (1997)	98%	0.2	✓	✓	✓	2%
Isiaka & M (2014)	100%	N/A	✗	✗	✗	N/A
Abordagem baseada características invariantes						
Ghimire & Lee (2013)	98.3%	0.2	✓	✓	✓	2%
Cootes <i>et al.</i> (1995)	92.45%	0.2	✓	✓	✓	2%
Toth & Madabhushi (2012)	91.92%	0.2	✓	✓	✓	2%
Xu & Lee (2013)	98.3%	00.1	✓	✓	✓	3.5%
Krafka <i>et al.</i> (2016)	98.4%	0.005	✗	✗	✗	3.8%
Parada <i>et al.</i> (2015)	98.75%	0.003	✓	✓	✓	0.5%
Wood <i>et al.</i> (2015)	99.23%	0.005	✓	✓	✓	2.3%
Abordagem baseada Templates						
Cootes <i>et al.</i> (2001)	96.4%	✗	✗	✗	✗	N/A
Brunelli & Poggio (1993)	95.32%	✗	✗	✗	✗	N/A
Kossaifi <i>et al.</i> (2014)	92%	✗	✗	✗	✗	N/A
El-Sayed & Khafagy (2014)	100%	✗	✗	✗	✗	N/A
Li <i>et al.</i> (2001)	95%	✗	✗	✗	✗	N/A
Rajpathak <i>et al.</i> (2009)	96%	✗	✗	✗	✗	N/A
Hemlata & Motwani (2012)	90%	✗	✗	✗	✗	N/A
Abordagem baseada na Aparência						
Jia <i>et al.</i> (2012)	95%	0.005	✓	✓	✓	4.1%
Guo <i>et al.</i> (2014)	94%	0.03	✓	✓	✓	2%
Schmidt & KasiĀdski (2007)	98%	0.003	✓	✓	✓	1%
Eriksson & Papanikolopoulos (2001)	96%	0.01	✓	✓	✓	3%
Bacivarov <i>et al.</i> (2008)	95.5%	N/A	✓	✗	✗	N/A
Soetedjo (2012)	91%	N/A	✓	✓	✓	0.5%
Liu <i>et al.</i> (2012)	97.67%	0.289	✗	✗	✗	5.05%
Zhu <i>et al.</i> (2002)	91.5%	✓	✓	✓	✓	N/A
Abordagem baseada métodos híbridos						
Hassaballah & Ido (2009)	97.8%	0.3	✗	✗	✗	0.4%
Majumder <i>et al.</i> (2013)	96%	0.2	✗	✗	✓	N/A
Jin <i>et al.</i> (2006)	90.75%	0.023	✗	✗	✗	N/A
Chennubhotla & Jepson (2004)	95%	0.01	✗	✗	✗	0.1%
Nossa Abordagem	98%	0.2	✓	✓	✓	2%

assinalado não superior a 1% para 68 pontos característicos assinalados. O erro, em geral, ocorre quando o usuário está a uma distância diferente da estabelecida por padrão, que seria de 50cm, isso por que os testes não foram realizados com os voluntários na cadeira. Notamos que o erro médio vai para a casa de 2% quando se estabelece uma distância diferente a qual os templates foram treinados.

Notamos que existe, no detector HCC, uma oscilação enquanto o indivíduo fecha e abre os olhos. Esse conjunto de erros se estabiliza em poucos milissegundos. Esta oscilação, Figura 4.36, causou “trancos” na cadeira e a solução para este problema foi dada usando um tempo de espera para a estabilização da detecção (cerca de 1 segundo) e a partir daí a cadeira teve sua partida e parada realizada suavemente.

Tabela 4.8: Análise geral dos métodos de detecção de movimentos da Cabeça

Trabalho	Acurácia	Análise	
		Resposta (s)	Erro
Abordagem baseada em algoritmos de visão computacional			
Siriteerakul <i>et al.</i> (2011)	95%	0.2	3%
Zhao <i>et al.</i> (2012)	92.86%	0.1	5%
Zhao & Yan (2011)	90%	0.1	4%
Basu & Pentland. (1996)	93%	N/A	0.2%
Liao & Cohen (2005)	96%	N/A	0.1%
Liu <i>et al.</i> (2012)	95%	N/A	3%
Murphy-Chutorian & Trivedi (2010)	94%	N/A	2%
Liu & Zhao (2011)	93%	N/A	4%
Zhao <i>et al.</i> (2011)	90%	N/A	5%
Qin & Shelton (2016)	96%	N/A	3%
Abordagem baseada Acelerômetros e Giroscópios			
Kiruluta <i>et al.</i> (1997)	98%	N/A	1%
Kupetz <i>et al.</i> (2010)	97%	N/A	3%
Song <i>et al.</i> (2011)	96%	0.1	2%
Nguyen <i>et al.</i> (2004)	93.75%	N/A	2%
King <i>et al.</i> (2005)	99.05%	N/A	1%
Nossa Abordagem	97.8%	0.2	2%



Figura 4.36: Análise da resposta do detector

Fonte: Elaborada pelo autor

O software mostrou-se estável, quanto a travamentos, durante a realização dos testes que o modelo se comportando como esperado. Algumas colisões foram observadas durante os testes no protótipo, em todos os cenários apresentados.

Conclusões

O objetivo deste trabalho foi desenvolver um modelo computacional para o gerenciamento dos movimentos básicos de uma CR motorizada utilizando expressões faciais. O modelo foi desenvolvido para ser usado por pessoas que possuem tetraplegia em ambientes domésticos, hospitalares ou similares que possam fornecer mapas e estruturas que se adequem aos critérios mínimos de acessibilidade.

Para atender a estas metas, trilhou-se o seguinte caminho:

- Estudados os conceitos de reconhecimento de padrões, classificadores em cascata, modelos de forma ativa e conceitos de estatística além de metodologias utilizadas em trabalhos relacionados, assim como inúmeros projetos de CR desenvolvidos por outras instituições e laboratórios de investigação;
- Desenvolvido o conceito para a navegação baseada em expressões faciais;
- Construído um protótipo real de CR, a partir de um modelo comercial. Sendo agregados a este uma câmera USB e uma placa arduino conectada ao controle da cadeira;
- Desenvolvido um classificador em cascata, específico para a detecção de olhos fechados;
- Desenvolvido um modelo de forma ativa capaz de detectar o contorno do rosto humano;
- Estudado e desenvolvida a integração entre o ASM e a fórmula de correlação de Pearson;
- Projetado um simulador da CR que permitiu testar o modelo desenvolvido;
- Implementadas as funções de operação da CR no protótipo;
- Implementados, separadamente, os módulos do modelo. Além disso, ambos foram avaliados de forma individual e integrada no conjunto tanto em simulação com no protótipo real da CR;
- Validado o modelo, assim com seus módulos, juntamente com o protótipo real da CR;

Nosso projeto foi baseado em quatro ideias: a primeira que era possível se incorporar tecnologias não intrusivas a uma cadeira de rodas comum que permita a um indivíduo

com tetraplegia assumir o controle do seu equipamento. Nosso modelo dispõe de três equipamentos, uma câmera, uma placa controladora e um computador, sendo este último, atualmente e largamente miniaturizado. Soma-se o fato de que a adaptação da cadeira ao modelo foi simples, necessitando apenas da conexão entre a placa controladora e os controles manuais da cadeira. A utilização de somente uma câmera, sem intrusividade ao usuário, a própria estrutura do protótipo além dos testes mostraram que nosso modelo é suficientemente flexível para permitir facilmente a transformação de CR motorizadas em uma CR com suporte a pessoas com tetraplegia. Esta ideia inicial foi possível graças ao desenvolvimento de uma metodologia para a navegação adequadas ao contexto do público alvo.

A metodologia de navegação mostrou-se bem promissora. A partir do desenvolvimento e análise primeiro de um novo classificador em cascata com o objetivo de estabelecer o movimento de “seguir em frente” e “parar”, e, em seguida, integrar um modelo de distribuição de pontos ao coeficiente de correlação de Pearson com o objetivo de realizar a rotação da cadeira utilizando o movimento da cabeça. Estas duas ideias se estabelecem com a verificação feita da sua eficiência pelos testes realizados, que foram conduzidos comparativamente com o controle próprio da cadeira. Os resultados obtidos permitiram confirmar as capacidades resultantes do modelo, entretanto, testes com o módulo demonstraram um excessivo escorregamento da cadeira como também a necessidade de ajuste da ergonomia da cadeira para conforto do paciente.

Nosso classificador em cascata foi capaz de identificar com precisão de maior que 95% o fechar dos olhos de um paciente e detector criado para indicar a rotação da cabeça obteve acerto igual ou superior a 97%. Em ambos os detectores conseguimos contornar alguns problemas como a variância de escala este de duas formas sendo a primeira o estabelecimento da distância ideal para a câmera e a segunda por meio da redução da dimensionalidade e padronização das imagens objeto da detecção.

Ambos os classificadores se comportaram muito bem com relação a variação de pose e iluminação sendo esta última muito em parte pela utilização de algoritmos de otimização. Apesar do modelo ser ajustado para boas condições de luz os testes realizados permitem dizer que mesmo em condições de pouca luz ($40 < X < 70$ lux) a funcionalidade do modelo continua ativa. Notamos também que barbas, bigodes, bonés e óculos não interferem no desempenho do modelo. Em particular ao integrar um modelo de distribuição de pontos ao coeficiente de correlação de Pearson a detecção do giro da cabeça ficou extremamente precisa, mas em alguns momentos de transição, principalmente entre L3, C e C, R3 ocorreram falsos positivos devido a proximidade da correlação.

A última ideia (o casamento entre os modelo de detecção para olhos fechados e rotação de cabeça) obteve a mesma eficiência que a navegação realizada pelo controlador manual

da cadeira com erro médio de 2%. O sistema desenvolvido é auto-configurável a qualquer usuário, sem a necessidade de treinamento tornando assim a navegação na forma sente-use e totalmente aberto à incorporação de novas interações com o usuário, como uma tela de saída. O modelo também mostrou-se capaz de entender a diferença entre piscar o olho e indicar um comando bem como o girar simples da cabeça e a indicação de giro para a cadeira. A rotação da cadeira pode ainda receber ajustes, pois notamos um “tranco” muito significativo no giro. Isso em parte pela forma de rotação do protótipo que para girar rotaciona as rodas traseiras invertidamente. Neste sentido o conjunto de objetivos para esta tese foi alcançado.

Esta tese apresentou uma síntese, análise e contribuições às perspectivas dos autores mais relevantes nas áreas relacionadas a navegação de CR por meio de visão computacional. Embora existam muitos trabalhos e diferentes abordagens nesta área, pode-se verificar a inexistência de uma CR que apresente as reais capacidades e permita a utilização no dia-a-dia por um indivíduo com tetraplegia. Cada projeto concentra-se em problemas mais ligados ao aspecto teórico que prático. Desta forma, é necessário conhecer os principais requisitos e tecnologias disponíveis. A tese fornece uma visão geral destes requisitos e tecnologias.

A tese, encaminha para o desenvolvimento do conceito de integração total para demais módulos, sendo este aplicado tanto em hardware como software. Os objetivos desta abordagem são: tornar mais perceptíveis os componentes envolvidos, permitir a manutenção e substituição dos módulos sem afetar o resto do sistema; simplificar o desenvolvimento; e incrementar a flexibilidade e fiabilidade.

Embora o módulo de navegação esteja estável outros conceitos necessitam ser propostos através da plataforma SmartWheels (SW). Até o momento, não foi possível colocar em prática ou testar todas elas. O primeiro caso inclui aquelas características que estão diretamente relacionadas com a ideia de um projeto mais amplo. Assim, com o protótipo foi possível verificar a compatibilidade entre o suporte de hardware da CR motorizadas comuns. Adicionalmente, as características visuais e ergonômicas não foram afetadas pela montagem dos dispositivos de hardware, atingindo assim outra meta proposta do projeto.

Além disso, é importante enfatizar que o objetivo de construir protótipos de baixo custo foram também parcialmente alcançados, quando utilizamos equipamentos que possuem um custo abaixo de R\$ 800,00. Todavia, o modelo é integrável e não impede a adição de novos sensores para o suporte de hardware. Pelo contrário, é aberta à adição de outros sensores desde que o impacto do seu custo no custo final do protótipo seja reduzido.

Os resultados obtidos nos testes evidenciaram a eficácia do modelo e portanto mostramos a viabilidade de um modelo computacional para o gerenciamento dos movimentos básicos

de uma CR motorizada utilizando expressões faciais. Outros testes podem ser conduzidos que tornem possível a integração a outros elementos como o desvio de obstáculos assim como a colaboração com outras CRs.

Finalmente, os conceitos propostos que não foram concluídos até a escrita desta tese bem como ideias para melhorias da plataforma desenvolvida, serão discutidos na próxima seção de trabalhos futuros.

5.1 *Trabalhos Futuros*

O projeto SmartWheels teve seu início simultâneo com o início desta tese o que exigiu a formação de um banco de informação para estudo e desenvolvimento. Além disso, a alta complexidade do sistema e os diversos domínios científicos envolvidos no tema desta tese, exigiram um grande esforço inicial para o desenvolvimento do projeto. Inevitavelmente, a implementação do modelo e ajuste no protótipos real acabaram consumindo o tempo disponível e condicionando o desenvolvimento do projeto. Assim, alguns recursos, funções e dispositivos que seriam interessantes acrescentar a este trabalho, não foram implementadas devido aos prazos do projeto. Diante dessa situação, a seguir são listados, a seguir, importantes recursos que poderão servir de ponto de partida para trabalhos de desenvolvimento e melhoramentos futuros da plataforma implementada no âmbito da investigação científica:

- Implementar e desenvolver um módulo de reconhecimento de voz. Esta nova abordagem poderá ser utilizada em conjunto com a detecção facial ou para execução de tarefas específicas como “vá para sala”. Sugerimos a API Google Cloud Speech por ter suporte para mais de 80 idiomas e trabalhar com basicamente qualquer tipo de aplicação. Quando recebe um comando, o sistema se conecta a um servidor online a fim de reconhecê-lo, retornando para o usuário em forma de ação;
- Aumentar o realismo do simulador de cadeira de rodas desenvolvido em MatLab para que os comportamentos da CRI no mundo real e simulado sejam ainda mais semelhantes;
- Realizar experiências que envolvam voluntário reais em testes de condução e avaliação dos algoritmos e que os participantes envolvidos não sejam somente indivíduos com deficiências, mas também pessoas que apresentem diferentes níveis de dificuldades;
- Desenvolver um módulo de controle de decisão inteligente. Tal módulo deverá ser responsável por estabelecer níveis de confiança e gerir entradas de acordo com a sua

percepção, buscando, deste modo, evitar conflitos, comandos com ruídos ou outra situação perigosa;

- Propor melhorias ao módulo de localização, relacionado com novas tecnologias para reduzir a incerteza sobre a posição e orientação da CR. Novas abordagens devem incluir *map matching*, odometria visual e sistemas de localização global de maneira a reduzir os erros de localização sistemáticos e não sistemáticos;
- Desenvolver uma Interface Gráfica de Utilizador intuitiva e amigável tanto para portadores de deficiências com limitações severas ou para pessoas com idades avançadas que já possuem dificuldades motoras. Em conjunto, poderia também ser desenvolvido um conversor de textos em voz e um assistente virtual para melhorar o processo de integração com o utilizador;
- Propor melhorias para aumentar a robustez do reconhecimento de expressões faciais.

O desenvolvimento dos trabalhos futuros sugeridos conduziria à criação de uma CR com alguma inteligência, inovadora e com elevada capacidade no auxílio a pessoas com graves deficiências e com um alto potencial para ser transformada num produto comercial.

Referências Bibliográficas

- ADACHI, Yoshihisa *et al.* Intelligent wheelchair using visual information on human faces. *proceedings of the IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, p. 354–359, 1998. [1.1](#), [2.2.1](#), [2.1](#), [4.6](#)
- ALGHOWINEM, S. *et al.* Head pose and movement analysis as an indicator of depression. In: *Proc. Humaine Association Conf. Affective Computing and Intelligent Interaction (ACII)*. [S.l.: s.n.], 2013. p. 283–288. ISSN 2156-8103. [2.3](#)
- ALSHAQAQI, Belal *et al.* Driver drowsiness detection system. In: *IEEE. Systems, Signal Processing and their Applications (WoSSPA), 2013 8th International Workshop on*. [S.l.], 2013. p. 151–155. [2.3](#), [2.3.1](#), [2.3.1](#), [2.3.1](#)
- ANDRADE, Valéria Sousa de; PEREIRA, Leani Souza Máximo. Influência da tecnologia assistiva no desempenho funcional e na qualidade de vida de idosos comunitários frágeis: uma revisão bibliográfica. *Rev. Bras. Geriatr. Gerontol*, v. 12, n. 1, p. 113–22, 2009. [1.3](#)
- ARAI, Kohei; MARDIYANTO, Ronny. Eyes based eletric wheel chair control system -i (eye) can control electric wheel chair. *International journal of advanced computer science & applications*, v. 2, p. 98, 2011. [1.1](#), [1.1](#), [2.3](#), [2.3.1](#), [2.3.1](#)
- BACIVAROV, Ioana; IONITA, Mircea; CORCORAN, Peter. Statistical models of appearance for eye tracking and eye-blink detection and measurement. *IEEE transactions on consumer electronics*, IEEE, v. 54, n. 3, p. 1312–1320, 2008. [2.3](#), [4.7](#)
- BAILEY, M. *et al.* Development of vision-based navigation for a robotic wheelchair. In: *2007 IEEE 10th International Conference on Rehabilitation Robotics*. [S.l.: s.n.], 2007. p. 951–957. ISSN 1945-7898. [1.1](#)
- BAREA, R. *et al.* Wheelchair guidance strategies using eog. *Journal of Intelligent and Robotic Systems*, v. 34, p. 279–299, 2002. [2.2.1](#), [2.1](#), [2.3](#), [4.6](#)
- BASU, Irfan Essa Sumit; PENTLAND., Alex. Motion regularization for model-based head tracking. In *Proceedings of the IEEE Int'l Conf. on Pattern Recognition (ICPR '96)*. Vienna, Austria., v. 3, p. 611–616, 1996. [1.1](#), [2.3.2](#), [2.4](#), [4.8](#)
- BATUR, A. U.; HAYES, M. H. Adaptive active appearance models. *IEEE Transactions on Image Processing*, v. 14, n. 11, p. 1707–1721, nov. 2005. ISSN 1057-7149. [2.3](#)
- Automatic closed eye correction*. 2433–2436 p. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5414163>. [2.3](#)
- BENOIT, Alexandre *et al.* Using human visual system modeling for bio-inspired low level image processing. *Computer vision and Image understanding*, v. 114(7), p. 758–773, 2010. [3.1.3](#), [3.5](#)
- BERGASA, L.M. *et al.* Commands generation by face movements applied to the guidance of a wheelchair for handicapped people. *Pattern Recognition, . Proceedings. 15th International Conference on*, v. 4, p. 660–663, 2000. [1.1](#), [1.1](#), [2.2.1](#)

- BEYMER, D. J. Face recognition under varying pose. In: *Proc. CVPR '94. IEEE Computer Society Conf Computer Vision and Pattern Recognition*. [S.l.: s.n.], 1994. p. 756–761. ISSN 1063-6919. 2.3, 3.1.4
- BIRKBECK, N.; JAGERSAND, M. Visual tracking using active appearance models. In: *Proc. First Canadian Conf. Computer and Robot Vision*. [S.l.: s.n.], 2004. p. 2–9. 2.3, 3.3
- BLACK, Michael J.; YACOOB, Yaser. Recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision*, v. 25, p. 23–48, 1997. 3.2
- BLEDSON, W. W. The model method in facial recognition. *Panoramic Research Inc., Palo Alto, CA, Rep. PRI:15,* 1966. 3.2.0.4
- BONARINI, Andrea *et al.* Introducing lurch: A shared autonomy robotic wheelchair with multimodal interfaces. In: *Proceedings of IROS 2012 Workshop on Progress, Challenges and Future Perspectives in Navigation and Manipulation Assistance for Robotic Wheelchairs*. [S.l.: s.n.], 2012. 1.3, 1.1, 1.2, 2.1, 2.3, 4.6
- BRADSKI, Gary; KEAHLER, Adrian. *OpenCV - Computer Vision with the OpenCV Library*. [S.l.]: O'Reilly Media Inc., 2008. 1, 3.1.2
- BRAGA, Rodrigo A. M. *Plataforma de desenvolvimento de cadeiras de rodas inteligentes*. Tese (Doutorado) — Faculdade de Engenharia da Universidade do Porto - Programa de Doutoramento em Engenharia Informática, Porto, Portugal, Setembro 2012. 1, 1.4, 1.1, 2.2.1, 2.3, 2.1, 2.3, 4.1.1, 4.6
- BRAGA, Rodrigo A. M. *et al.* Intellwheels - a development platform for intelligent wheelchairs for disabled people. In: *ICINCO-RA (1)*. [S.l.: s.n.], 2008. p. 115–121. 1.1, 2.2.1
- _____. Concept and design of the intellwheels platform for developing intelligent wheelchairs. In: CETTO, Juan Andrade; FERRIER, Jean-Louis; FILIPE, Joaquim (Ed.). *Informatics in Control, Automation and Robotics*. [S.l.]: Springer Berlin Heidelberg, 2009, (Lecture Notes in Electrical Engineering, v. 37). p. 191–203. 1, 2.2.1
- BRUNELLI, Roberto; POGGIO, T. Face recognition: features versus templates. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 15, n. 10, p. 1042–1052, Oct 1993. ISSN 0162-8828. 2.2, 2.3, 4.7
- BUDZAN, Sebastian; WYÅijGOLIK, Roman. Face and eyes localization algorithm in thermal images for temperature measurement of the inner canthus of the eyes. *Infrared Physics & Technology*, v. 60, p. 225 – 234, 2013. ISSN 1350-4495. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1350449513001096>>. 2.3.1, 2.3.1
- CARBON, FREEDOM. *Manual do Proprietário - Cadeiras de rodas motorizadas FREEDOM CARBON*. 2. ed. rua Conde de Porto Alegre, 155, 2016. 1.1
- CHACKO, Jobby K *et al.* Microcontroller based eog guided wheelchair. *World Academy of Science, Engineering and Technology (WASET) International Journal of Medical, Health, Biomedical and Pharmaceutical Engineering*, v. 7, n. 11, 2013. 2.2.1, 2.1, 2.3, 4.6
- CHAU, Michael; BETKE, Margrit. *Real time eye tracking and blink detection with usb cameras*. [S.l.], 2005. 2.3.1

CHELLAPPA, R.; WILSON, C. L.; SIROHEY, S. Human and machine recognition of faces: a survey. *IEEE Transactions on pattern Analysis and Machine Intelligence*, v. 83, n. 5, p. 705–741, maio 1995. 3.2.0.4

CHEN, X. *et al.* Medical image segmentation by combining graph cuts and oriented active appearance models. *IEEE Transactions on Image Processing*, v. 21, n. 4, p. 2035–2046, abr. 2012. ISSN 1057-7149. 3.3

CHENNUBHOTLA, Chakra; JEPSON, Allan. Perceptual distance normalization for appearance detection. In: IEEE. *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. [S.l.], 2004. v. 2, p. 23–27. 2.3, 4.7

COOK, A.M.; HUSSEY, S. M. *Assistive technologies: Principles and Practices*. [S.l.]: Mosby-Year Book, Inc, 1995. 1, 1.3

COOTES, Timothy F; EDWARDS, Gareth J; TAYLOR, Christopher J. Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE Computer Society, v. 23, n. 6, p. 681–685, 2001. 2.2, 2.3, 3, 4.7

COOTES, Timothy F *et al.* Active shape models-their training and application. *Computer vision and image understanding*, Elsevier, v. 61, n. 1, p. 38–59, 1995. 2.2, 2.3, 3.1.6, 3.3, 3.3.1.2, 3.3.1.2, 3.3.1.3, 3.3.2, 4.7

CRUZ, Daniel Marinho Cezar da; IOSHIMOTO, Maria Teresa Augusto. Tecnologia assistiva para as atividades de vida diária na tetraplegia completa c6 pÃ§s-lesÃ£o medular. *Revista TriÃngulo -Revista do Programa de PÃ§s-GraduaÃ§Ã£o em EducaÃ§Ã£o da UFMT*, v. 3, n. 2, p. 177–190, jul/dez 2010. 1, 1.3, 3.5

DUNTEMAN, G.H. *Principal components analysis*. [S.l.]: Sage Publications, Inc, 1989. 3.3.1.3

Robust eye localization in video by combining eye detector and eye tracker. 242–245 p. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6460117>>. 1.1

DYMOND, E.; POTTER, R. Head movements for control of assistive technology. *Engineering in Medicine and Biology Society, 14th Annual International Conference of the IEEE*, v. 4, p. 1527–1528, 1992. 1.1, 1.1, 2.2.1, 2.1, 2.3, 4.6

EKMAN, P.; ROSENBERG, E.L. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press, 2005. (Series in Affective Science). ISBN 9780199792726. Disponível em: <<https://books.google.com.br/books?id=UXapcWqtO-sC>>. 1.1, 2.3.2, 3.1.5, 3.5, 3.5

EL-SAYED, Mohamed A; KHAFAGY, Mohamed A. An identification system using eye detection based on wavelets and neural networks. *International Journal of Computer and Information Technology*, v. 1, n. 2, p. 43–48, nov. 2014. 2.3, 4.7

ERIKSSON, Martin; PAPANIKOLOPOULOS, Nikolaos P. Driver fatigue: a vision-based approach to automatic diagnosis. *Transportation Research Part C: Emerging Technologies*, v. 9, n. 6, p. 399 – 413, 2001. ISSN 0968-090X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0968090X00000450>>. 2.3.1, 2.3.1, 2.3, 4.7

Eye-tracking for detection of driver fatigue. 314–319 p. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=660494>>. 2.3.1, 2.3.1

ETHEVE, Remy; RICCIO, Daniel; DUGELAY, Jean-Luc. State of the art in 3d face recognition. *Eurocom RR 06 160*, p. 14–31, 2006. 3.2.0.4

FAZLI, Saeid; ESFEHANI, Parisa. Tracking eye state for fatigue detection. *International Conference on Advances in Computer and Electrical Engineering (ICACEE'2012)*, p. 17–18, 2012. 2.3.1, 2.3.1

FIGUEREDO, Marcos. *Reconhecimento de faces aplicado ao problema de pessoas desaparecidas - Estudo de caso do Eigenface*. Dissertação (Mestrado) — Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial do SENAI/CIMATEC, 2011. 1.1, 2.3.1, 3.2.0.4

FILHO, João Gomes. *Ergonomia do objeto: sistema técnico de leitura ergonômica*. 2. ed. São Paulo: Escrituras Editora, 2010. 19 p. 3.17

FREUND, Yoav; SCHAPIRE, Robert E. A short introduction to boosting. In: *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*. [S.l.]: Morgan Kaufmann, 1999. p. 1401–1406. 3.2.0.3

FURTADO, Vasco. *Tecnologia e Gestão da Informação na Segurança Pública*. [S.l.]: Garamond, 2002. 4.1.9

GAO, Wei; ZHOU, Zhi-Hua. Approximation stability and boosting. *Algorithmic Learning Theory*, v. 21, p. 59–73, 2010. 3.2.0.3

GHIMIRE, Deepak; LEE, Joonwhoan. A robust face detection method based on skin color and edges. *JIPS*, v. 9, n. 1, p. 141–156, 2013. 2.2, 2.3, 4.7

GONG, Shaigang; XIANG, Tao. Understanding facial expression. In: *Visual Analysis of Behaviour*. [S.l.]: Springer London, 2011. p. 69–93. 2.3.1, 3.2.0.4

GONZALEZ, Rafael C.; WOODS, Richard E.; EDDINS, Steven L. *Digital Image Processing using MATLAB*. New Jersey-USA: [s.n.], 2004. 3.1, 3.1.2, 4.1.9

GRUNDMANN, M.; KWATRA, V.; ESSA, I. Auto-directed video stabilization with robust l1 optimal camera paths. In: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. [s.n.], 2011. p. 225–232. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5995525>>. 4.2

GUO, Lie *et al.* Pedestrian detection for intelligent transportation systems combining adaboost algorithm and support vector machine. *Expert Systems with Applications*, Elsevier, v. 39, n. 4, p. 4274–4286, 2012. 2.2

A new algorithm detects pilot fatigue based on machine vision. 5271–5276 p. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6853121>>. 2.3, 4.7

HALAWANI, Alaa *et al.* Active vision for controlling an electric wheelchair. *Intelligent Service Robotics, Springer Science & Business Media B.V*, Vol.5(2), p. 89–98, 2012. 1, 1.1, 1.2, 2.1, 2.3.1, 4.6

HASSABALLAH, M; IDO, Shun. Eye detection using intensity and appearance information. In: *MVA*. [S.l.: s.n.], 2009. p. 346–349. 2.3, 4.7

- HEMLATA, A; MOTWANI, Mahesh. Fast eye detection using boundary value and template based method. *International Journal of Computer Theory and Engineering*, IACSIT Press, v. 4, n. 6, p. 880, 2012. 2.3, 4.7
- HJELMAS, Erik; Low Boon Keen. Face detection: Asurvey. *Computer Vision and Image Undertanding*, v. 83', n. 83, p. 236–274, 2001. 3.2.0.4
- HOMMEL, S.; HANDMANN, U. Realtime aam based user attention estimation. In: *Proc. IEEE 9th Int. Symp. Intelligent Systems and Informatics*. [S.l.: s.n.], 2011. p. 201–206. ISSN 1949-047X. 3.3
- HORTON, Michael; CAMERON-JONES, Mike; WILLIAMS, Raymound. Multiple classifier object detection with confidence measures. *advances in artificial intelligence*, v. 20, p. 559–568, 2007. 3.2.0.4
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, v. 24, p. 417–441,498–520, 1933. 3.3.1.3
- IBGE, Instituto Brasileiro de Geografi a e EstatÃnstica. CaracterÃnsticas da populaÃgÃo e dos domicÃnlios resultados do universo. *Censo DemogrÃqfico 2010*, v. 1, p. 161, 2011. 1, 1, 1.4
- ISIAKA, Fatima; M, Ibrahim Admu. Pattern recognition in eye movement validation. *International Refereed Journal of Engineering and Science (IRJES)*, v. 3, n. 8, p. 45–55, ago. 2014. 2.3, 4.7
- JAFFE, David L. An ultrasonic head position interface for wheelchair control. *Journal of Medical Systems*, v. 6, n. 4, p. 337–342, 1982. ISSN 1573-689X. Disponível em: <<http://dx.doi.org/10.1007/BF00992877>>. 3.3
- JAIN, Monika; JOSHI, Hitesh. Tongue operated wheelchair for physically disabled people. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, v. 4, n. 1, 2014. 2.2.1, 2.1, 2.3, 4.6
- JENKINS, R. Limitations in faci identification: The evidence. *FEATURE OF THE WEEK*, p. 5–12, 2008. 2.3.1
- JI, Qiang; YANG, Xiaojie. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, v. 8, p. 357–377, 2002. 2.3.1, 2.3.1
- JIA, Haipeng *et al.* Accelerating viola-jones face detection algorithm on gpus. In: IEEE. *High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICISS)*, 2012 IEEE 14th International Conference on. [S.l.], 2012. p. 396–403. 2.2, 2.3, 4.7
- JIN, Lizuo *et al.* A hybrid classifier for precise and robust eye detection. In: IEEE. *18th International Conference on Pattern Recognition (ICPR'06)*. [S.l.], 2006. v. 4, p. 731–735. 2.3, 4.7
- JOLLIFFE, IT. *Principal component analysis*. [S.l.]: Springer verlag, 2002. 2.3.1, 3.3.1.3
- KANADE, T. *Picture processing system by computer complex and recognition of human faces*. Tese (Doutorado) — Dept. of Information Science, Kyoto University,, 1973. 3.2.0.4
- KAPLAN, Robert M *et al.* The quality of well-being scale: applications in aids, cystic fibrosis, and arthritis. *Medical care*, JSTOR, p. S27–S43, 1989. 1

- KHAN, A. S.; ALIZAI, L. K. Introduction to face detection using eigenfaces. In: *Proc. International Conference on Emerging Technologies ICET '06*. [S.l.: s.n.], 2006. p. 128–132. [3.2.0.4](#)
- KING, L.M.; NGUYEN, H.T.; TAYLOR, P.B. Hands-free head-movement gesture recognition using artificial neural networks and the magnified gradient function. *Engineering in Medicine and Biology Society. IEEE-EMBS 2005. 27th Annual International Conference of the*, p. 2063–2066, 2005. [1.1](#), [2.2.1](#), [2.4](#), [4.8](#)
- KIRULUTA, Andrew; EIZENMAN, Moshe; PASUPATHY, Subbarayan. Predictive head movement tracking using a kalman filter. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS*, v. 27, n. 2, p. 326–331, 1997. [2.3.2](#), [2.4](#), [4.8](#)
- KOSSAIFI, J.; TZIMIROPOULOS, G.; PANTIC, M. Fast Newton active appearance models. In: *Proc. IEEE Int. Conf. Image Processing (ICIP)*. [S.l.: s.n.], 2014. p. 1420–1424. ISSN 1522-4880. [2.2](#), [2.3](#), [4.7](#)
- KRAFKA, Kyle *et al.* Eye tracking for everyone. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2016. p. 2176–2184. [2.3](#), [4.7](#)
- KUNO, Yoshinori *et al.* Intelligent wheelchair remotely controlled by interactive gestures. *Pattern Recognition Proceedings. 15th International Conference on*, v. 4, p. 672–675, 2000. [1.2](#), [2.1](#), [4.6](#)
- KUNO, Y. *et al.* Intelligent wheelchair based on the integration of human and environment observations. *Information Intelligence and Systems, Proceedings. 1999 International Conference on*, p. 342–349, 1999. [1.1](#), [2.2.1](#), [2.3](#)
- KUPETZ, D. J.; WENTZELL, S. A.; BUSHA, B. F. Head motion controlled power wheelchair. In: *Proceedings of the 2010 IEEE 36th Annual Northeast Bioengineering Conference (NEBEC)*. [S.l.: s.n.], 2010. p. 1–2. ISSN 2160-6986. [2.3.2](#), [2.4](#), [4.8](#)
- LAKSHMANAN, S. *et al.* Comparison between eigenfaces and fisherfaces for estimating driver pose. In: *Proc. IEEE Intelligent Transportation Systems*. [S.l.: s.n.], 2001. p. 889–894. [3.2](#)
- LEE, Chan-Su; SAMARAS, D. Analysis and synthesis of facial expressions using decomposable nonlinear generative models. In: *Proc. IEEE Int Automatic Face & Gesture Recognition and Workshops (FG 2011) Conf.* [S.l.: s.n.], 2011. p. 847–852. [2.2.1](#), [2.1](#), [4.6](#)
- LETHAUS, Firas *et al.* Using pattern recognition to predict driver intent. In: SPRINGER. *International Conference on Adaptive and Natural Computing Algorithms*. [S.l.], 2011. p. 140–149. [2.3](#), [4.7](#)
- LI, Ying; QI, Xiang-lin; WANG, Yun-jiu. Eye detection by using fuzzy template matching and feature-parameter-based judgement. *Pattern Recognition Letters*, Elsevier, v. 22, n. 10, p. 1111–1124, 2001. [2.3](#), [4.7](#)
- LIAO, Wei-Kai; COHEN, I. Classifying facial gestures in presence of head motion. *Computer Vision and Pattern Recognition - Workshops. CVPR Workshops. IEEE Computer Society Conference on*, p. 77, 2005. [2.3.2](#), [2.4](#), [4.8](#)

- LIN, Chern-Sheng; CHANG, Kai-Chieh; JAIN, Young-Jou. A new data processing and calibration method for an eye-tracking device pronunciation system. *Optics & Laser Technology*, v. 34, p. 405–413, 2002. [2.3.1](#), [2.3.1](#)
- LIU, Jian-zheng; ZHAO, Zheng. Head movement recognition based on lk algorithm and gentleboost. *Networked Computing and Advanced Information Management (NCM), 7th International Conference on*, p. 232–236, 2011. [1](#), [2.2.1](#), [2.3.2](#), [2.4](#), [4.8](#)
- LIU, P.; REALE, M.; YIN, L. 3D head pose estimation based on scene flow and generic head model. In: *Proc. IEEE Int. Conf. Multimedia and Expo*. [S.l.: s.n.], 2012. p. 794–799. ISSN 1945-7871. [2.3.2](#), [2.4](#), [4.8](#)
- LIU, X. *et al.* 3d head pose estimation with convolutional neural network trained on synthetic images. In: *2016 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2016. p. 1289–1293. [2.3.2](#)
- LIU, Xue; TAN, Xiaoyang; CHEN, Songcan. Eyes closeness detection using appearance based methods. In: SPRINGER. *International Conference on Intelligent Information Processing*. [S.l.], 2012. p. 398–408. [2.3](#), [4.7](#)
- MAHALANOBIS, Prasanta Chandra. On the generalized distance in statistics. *Proceedings of the National Institute of Sciences (Calcutta)*, v. 2, p. 49–55, 1936. [3.1.6](#)
- MAJUMDER, Goutam; BHOWMIK, Mrinal Kanti; BHATACHARJEE, Debotosh. Automatic eye detection using fast corner detector of north east indian (nei) face images. *Procedia Technology*, Elsevier, v. 10, p. 646–653, 2013. [2.3](#), [4.7](#)
- MALLIN, Sandra SV. Uma metodologia de design, aplicada ao desenvolvimento de tecnologia assistiva para portadores de paralisia cerebral. *Curitiba: Editora da UFPR*, 2004. [1](#), [1.3](#)
- MANOGNA, S.; VAISHNAVI, S.; GEETHANJALI, B. Head movement based assist system for physically challenged. In: *2010 4th International Conference on Bioinformatics and Biomedical Engineering*. [S.l.: s.n.], 2010. p. 1–4. ISSN 2151-7614. [1.1](#), [2.2.1](#), [2.1](#), [2.3](#), [3.3](#), [4.6](#)
- MARTÍNEZ, Aleix; BENAVENTE, Robert. *The AR Face Database*. Bellatera, 1998. Cites in Scholar Google: <http://scholar.google.com/scholar?hl=en&lr=&client=firefox-a&cites=1504264687621469812>. Disponível em: <http://www.cat.uab.cat/Public/Publications/1998/MaB1998>>. [4.1.8](#)
- MATSUSHITA, Y. *et al.* Full-frame video stabilization with motion inpainting. v. 28, n. 7, p. 1150–1163, 2006. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1634345>>. [4.2](#)
- MELO, Valdenice Rumão de. *Avaliação da qualidade de vida de paciente com lesão medular acompanhados em regime ambulatorial*. Dissertação (Dissertação de Mestrado) — Universidade Federal de Pernambuco CCS. Neuropsiquiatria e Ciências do Comportamento, 2009. [1](#)
- MEYLAN, Laurence; ALLEYSSON, David; SÜSTRUNK, Sabine. Model of retinal local adaptation for the tone mapping of color filter array images. *JOSA A*, Optical Society of America, v. 24, n. 9, p. 2807–2816, 2007. [3.1.3](#)

- MURAWSKI, K; RÓŻANOWSKI, K. Pattern recognition algorithm for eye tracker sensor video data analysis. *Acta Physica Polonica, A.*, v. 124, n. 3, 2013. [2.3](#), [4.7](#)
- MURPHY-CHUTORIAN, E.; TRIVEDI, M. M. Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness. *IEEE Transactions on Intelligent Transportation Systems*, v. 11, n. 2, p. 300–311, June 2010. ISSN 1524-9050. [2.3.2](#), [2.4](#), [4.8](#)
- NGUYEN, H.T.; KING, L.M.; KNIGHT, G. Real-time head movement system and embedded linux implementation for the control of power wheelchairs. *Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, USA*, p. 1–5, 2004. [1.1](#), [2.2.1](#), [2.4](#), [4.8](#)
- NOCK, Richard; NIELSEN, Frank. A real generalization of discrete adaboost. *17th European Conference on Artificial Intelligence*, v. 17, p. 509–515, 2006. [3.2.0.3](#)
- ONER, M.; HU, Gongzhu. Analyzing one-channel EEG signals for detection of close and open eyes activities. *Advanced Applied Informatics (IIAIAAI), 2013 IIAI International Conference on*, p. 318–323, 2013. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6630367>. [2.3](#)
- PARADA, Francisco J *et al.* Experteyes: Open-source, high-definition eyetracking. *Behavior research methods*, Springer, v. 47, n. 1, p. 73–84, 2015. [2.3](#), [4.7](#)
- PARIKH, Sarangi P *et al.* Integrating human inputs with autonomous behaviors on an intelligent wheelchair platform. *IEEE Intelligent Systems*, IEEE, v. 22, n. 2, p. 33–41, 2007. [1.2](#), [1.2](#), [2.1](#), [4.6](#)
- PEREZ, Elisa *et al.* Robotic wheelchair controlled through a vision-based interface. *Robotica*, v. 30, p. 691–708, 9 2012. ISSN 1469-8668. Disponível em: http://journals.cambridge.org/article_S0263574711000919. [1.1](#), [1.2](#), [2.1](#), [2.3](#), [4.6](#)
- PORTHIYAS, Danny *et al.* intellichair: Assistive technology for powered wheelchair users. *IEEE Canada International Humanitarian Technology Conference*, v. 1, 2014. [2.2.1](#), [2.1](#), [4.6](#)
- QIN, Z.; SHELTON, C. R. Social grouping for multi-target tracking and head pose estimation in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 38, n. 10, p. 2082–2095, Oct 2016. ISSN 0162-8828. [2.3.2](#), [2.4](#), [4.8](#)
- RAJPATHAK, Tanmay; KUMAR, Ratnesh; SCHWARTZ, Eric. Eye detection using morphological and color image processing. In: *2009 Florida Conference on Recent Advances in Robotics, FCRAR*. [S.l.: s.n.], 2009. p. 1–6. [2.3](#), [4.7](#)
- RAUTARAY, Siddharth S; AGRAWAL, Anupam. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, Springer, v. 43, n. 1, p. 1–54, 2015. [2.2](#), [2.3](#), [4.7](#)
- REINDERS, MJT. Eye tracking by template matching using an automatic codebook generation scheme. In: *Third Annual Conference of ASCI, ASCI, Delft*. [S.l.: s.n.], 1997. p. 85–91. [2.3](#), [4.7](#)
- REIS, LuÃns Paulo *et al.* Intellwheels mmi: A flexible interface for an intelligent wheelchair. In: BALTES, Jacky *et al.* (Ed.). *RoboCup*. Springer, 2009. (Lecture Notes in Computer Science, v. 5949), p. 296–307. ISBN 978-3-642-11875-3. Disponível em: <http://dblp.uni-trier.de/db/conf/robocup/robocup2009.html#ReisBSM09>. [1.2](#), [4.1.1](#)

ROSOT, Maria Augusta Doetzer; DISPERATI, Attilio Antonio; FILHO, Afonso Figueredo. Medição da área de anéis de crescimento usando imagens digitais. *Revista Árvore/SIF*, v. 1, 2001. 3.1.1

SAINI, Vandna; SAINI, Rekha. Driver drowsiness detection system and techniques: A review. (*IJCSIT*) *International Journal of Computer Science and Information Technologies*, v. 5, p. 4245–4249, 2014. 2.3.1, 2.3.1

SALEHI, Bahram *et al.* Object-based classification of urban areas using vhr imagery and height points ancillary data. *Remote Sensing, Molecular Diversity Preservation International*, v. 4, n. 8, p. 2256–2276, 2012. 2.2, 2.3, 4.7

SCHAPIRE, R. E. The strength of weak learnability. *Machine Learning*, v. 5, p. 197–227, 1990. 3.2.0.3

SCHMIDT, Adam; KASIĄDŃSKI, Andrzej. The performance of the haar cascade classifiers applied to the face and eyes detection. In: KURZYŃSKI, Marek *et al.* (Ed.). *Computer Recognition Systems 2*. Springer Berlin Heidelberg, 2007, (Advances in Soft Computing, v. 45). p. 816–823. ISBN 978-3-540-75174-8. Disponível em: <http://dx.doi.org/10.1007/978-3-540-75175-5_101>. 2.3.1, 2.3, 4.7

SEBARI, Imane; HE, Dong-Chen. Automatic fuzzy object-based analysis of vhsr images for urban objects extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 79, p. 171–184, 2013. 2.2, 2.3, 4.7

SIRITEERAKUL, T.; SATO, Y.; BOONJING, V. Estimating change in head pose from low resolution video using lbp-based tracking. In: *Intelligent Signal Processing and Communications Systems (ISPACS), 2011 International Symposium on*. [S.l.: s.n.], 2011. p. 1–6. 2.3.2, 2.4, 4.8

SOETEDJO, Aryuanto. Eye detection based-on color and shape features. *International Journal of Advanced Computer Science and Applications*, Citeseer, v. 3, n. 5, 2012. 2.3, 4.7

SONG, You; LUO, Yunfeng; LIN, Jun. Detection of movements of head and mouth to provide computer access for disabled. *Technologies and Applications of Artificial Intelligence (TAAI), International Conference on*, p. 223–226, 2011. 1, 2.3.2, 2.4, 4.8

SPIEGEL, M.R.; SCHILLER, J.J.; SRINIVASAN, R.A. *Teoria E Prob. Probabilidade E Estatística*. BOOKMAN COMPANHIA ED, 2004. (Coleção Schaum). ISBN 9788536302973. Disponível em: <https://books.google.com.br/books?id=11_3Qgwt9YoC>. 4.1.3

TAYLOR, P. B.; NGUYEN, H. T. Performance of a head-movement interface for wheelchair control. *Proceedings of the 25th Annual International Conference of the IEEE EMBS*, p. 17–21, 2003. 1.1, 2.2.1, 2.1, 4.6

TOTH, Robert; MADABHUSHI, Anant. Multifeature landmark-free active appearance models: Application to prostate mri segmentation. *Medical Imaging, IEEE Transactions on*, IEEE, v. 31, n. 8, p. 1638–1650, 2012. 2.2, 2.3, 4.7

URBANO, Margarida *et al.* Adaptation of powered wheelchairs for quadriplegic patients with reduced strength. *Emerging Technologies and Factory Automation, 2008. ETFA 2008. IEEE International Conference on*, p. 1104–1107, 2008. 2.2.1, 2.1, 4.6

- VARELA, Renata Cristina Bertolozzi; OLIVER, Fátima Corrêa *et al.* A utilização de tecnologia assistiva na vida cotidiana de crianças com deficiência. *Ciência & Saúde Coletiva*, ABRASCO, v. 18, n. 6, p. 1773–1784, 2013. [1.3](#)
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. [S.l.: s.n.], 2001. v. 1, p. I–511–I–518 vol.1. ISSN 1063-6919. [2.3.1](#), [2.2](#), [3.2.0.1](#), [3.2.0.2](#), [3.2.0.3](#), [3.2.0.4](#)
- VIOLA, Paul; JONES, Michael J. Robust real-time face detection. *International journal of computer vision*, Springer, v. 57, n. 2, p. 137–154, 2004. [3](#), [3.1.4](#), [3.2](#), [3.2.0.3](#), [3.2.0.4](#), [3.5](#)
- WEI, Lai *et al.* Evaluating the performance of a face movement based wheelchair control interface in an indoor environment. *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, p. 387–392, 2010. [2.2.1](#), [2.3.1](#)
- WOOD, Erroll *et al.* Rendering of eyes for eye-shape registration and gaze estimation. In: IEEE. *IEEE International Conference on Computer Vision (ICCV)*. [S.l.], 2015. p. 3756–3764. [2.3](#), [4.7](#)
- WU, Junwen; TRIVEDI, Mohan M. An eye localization, tracking and blink pattern recognition system: Algorithm and evaluation. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, ACM, v. 6, n. 2, p. 8, 2010. [2.3](#), [4.7](#)
- XU, Wenkai; LEE, Eung-Joo. Eye detection and tracking using rectangle features and integrated eye tracker by web camera. *Int. J. Multimedia Ubiquitous Eng*, Citeseer, v. 8, p. 25–34, 2013. [2.3](#), [4.7](#)
- YI, Zhang; XIAOLIN, Feng; YUAN, Luo. Intelligent wheelchair system based on semg and head gesture. *The Journal of China Universities of Posts and Telecommunications*, v. 22, n. 2, p. 74 – 95, 2015. ISSN 1005-8885. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1005888515606422>. [3.3](#)
- ZHAO, X. *et al.* Accurate landmarking of three-dimensional facial data in the presence of facial expressions and occlusions using a three-dimensional statistical facial feature model. v. 41, n. 5, p. 1417–1428, 2011. [2.4](#), [4.8](#)
- ZHAO, Youen; YAN, Hua. Head orientation estimation using neural network. In: *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*. [S.l.: s.n.], 2011. v. 3, p. 2075–2078. [2.3.2](#), [2.4](#), [4.8](#)
- ZHAO, Zheng; WANG, Yuchuan; FU, Shengbo. Head movement recognition based on lucas-kanade algorithm. *Computer Science & Service System (CSSS), International Conference on*, p. 2303–2306, 2012. [1](#), [1.1](#), [2.2.1](#), [2.4](#), [3.3](#), [4.8](#)
- ZHU, Zhiwei; FUJIMURA, Kikuo; JI, Qiang. Real-time eye detection and tracking under various light conditions. In: ACM. *Proceedings of the 2002 symposium on Eye tracking research & applications*. [S.l.], 2002. p. 139–144. [2.3](#), [4.7](#)
- ZHUA, Zhiwei; JIB, Qiang. Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Computer Vision and Image Understanding*, v. 98, p. 124–154, 2005. [2.3.1](#)