

ESTUDO DA NAVEGAÇÃO DE UMA CADEIRA DE RODAS UTILIZANDO CONCEITOS DE VISÃO COMPUTACIONAL

STUDY OF THE NAVIGATION OF A WHEELCHAIR USING COMPUTATIONAL VISION CONCEPTS

Marcos Batista Figueredo¹, Roberto Luiz Souza Monteiro^{1,2,3}

Universidade do Estado da Bahia – UNEB¹; Faculdade de Tecnologia SENAI/CIMANTEC – FTSC²; Programa de Modelagem Computacional – FTSC³

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. For people who have the so called paraplegias or tetraplegias, the wheelchair is an important means of exercising their citizenship. Several researches seek to make navigation simple and efficient, but, in general, the presented solutions have a great amount of sensing, intrusiveness and high cost. We propose a computational model that allows the navigation of a wheelchair using facial expressions. Unlike the works studied, we suggest a model that is based on two facial expressions: the pose of the head and the closing of the eyes, and only an input sensor, a USB camera. The model converts facial expressions into commands for navigating the chair, and two techniques make interpretation: Cascade Classifiers and Active Shape Models (ASM). In the first, it uses a classifier capable of detecting the closure of the eyes and in the second the marriage between the ASM response and the Pearson correlation coefficient. The tests show that the model has excellent accuracy and precision and a robust performance in the detection of closed eyes 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%.

Key words: Wheelchair, Assistive Technology, Cascade Classifiers, Active Shape Models

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. Para pessoas que possuem as chamadas paraplegias ou tetraplegias a cadeira de rodas é um importante meio para exercer sua cidadania. Diversas pesquisas buscam tornar a navegação simples e eficiente, mas, em geral, as soluções apresentadas possuem uma grande quantidade de sensoriamento, intrusividade e alto custo. Propomos um modelo computacional que permita a navegação de uma cadeira de rodas utilizando expressões faciais. Ao contrário dos trabalhos estudados, sugerimos um modelo que se baseia em duas expressões faciais: a pose da cabeça e o fechamento dos olhos, e apenas um sensor de entrada, uma câmera USB. O modelo converte as expressões faciais em comandos para a navegação da cadeira e duas técnicas fazem a interpretação: Classificadores em cascata e Modelos de Forma Ativa(ASM). Na primeira, utiliza um classificador capaz de detectar o fechamento dos olhos e na segunda o casamento entre a resposta do ASM e o coeficiente de correlação de Pearson. Os testes 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: Cadeira de Rodas, Tecnologia Assistiva, Classificadores em Cascata, Active Shape Models.

Introdução

Diversas doenças ou acidentes podem levar um paciente a graves danos na medula espinhal ocasionando a perda de movimento nas partes inferiores e superiores. 14% da população brasileira possui algum tipo de deficiência, seja ela visual, motora, auditiva, entre outras. Dentre esse grupo, cerca de 4% não possuem qualquer tipo de movimento nos membros inferiores e superiores¹.

Os trabalhos de Braga², Halawani³, Song⁴ e Zhao⁵ indicam ações tecnológicas voltadas ao controle da cadeira que visem maior autonomia e independência de pacientes que perderam o movimento dos membros inferiores e superiores, associados ao baixo custo do equipamento, e em geral estas ações prolongam e melhoram sua qualidade de vida,⁶⁻⁷.

Motivado por esta realidade iniciou-se um projeto de cunho interdisciplinar que visa dotar uma cadeira de rodas com elementos que tornem tanto possível como viável a mobilidade, mesmo que parcial, de indivíduos com lesões na medula que utilizem como controle apenas expressões faciais.

Diversos trabalhos correlatos apresentam modelos que buscam tratar essa questão com abordagens variadas, dos quais alguns merecem destaque.

O trabalho de Dymond⁸ propõe uma cadeira que utiliza os movimentos da cabeça para realizar a navegação utilizando um sistema de vídeo. SumitBasu⁹ 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. Kiruluta¹⁰ procura diminuir a ação do movimento da cabeça supondo que com poucos movimentos a cadeira já pode seguir uma rota estimada por esta posição diminuindo assim o esforço do usuário. Adachi¹¹ tenta também prever o desejo do usuário utilizando apenas as informações da face por meio de detecção de 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 permite indicar a localização desses pontos e assim determinar a direção que o usuário deseja guiar a cadeira.

Kuno¹² utiliza os gestos faciais integrados a uma certa inteligência da cadeira para gerar o movimento, tentando assim compreender as intenções do usuário. Nesta perspectiva o

usuário indica a direção e a cadeira é responsável por todo o processo de navegação. O sistema, que gera os comandos, é baseado em movimentos faciais e rastreadores 2D, além de um detector de cor e um difuso. Já Bergassa¹³ trabalha com uma arquitetura específica para tentar definir comandos que gerem o movimento de guiagem. Não leva em consideração os aspectos tratados em Kuno¹², mas destaca que o modelo se revela 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.

Em Taylor¹⁴ e Nguyen¹⁵ retoma-se as ideias de sensores colocados na cabeça do usuário para detecção do movimento num sistema desenvolvido para plataforma livre. Este mesmo conceito foi revisto em King¹⁶ e obtendo índices de acerto entre 98% e 99% também passou a utilizar um acelerômetro.

Liao¹⁷ 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 utilizando o Active Shape Models para extrair características da face para, em seguida, montar um modelo em 3D em consonância com a Facial Actinos Coding System FACS.

Manogna¹⁸ cria um sistema que controla diretamente a rotação do motor com base no movimento da cabeça. O dispositivo é fixado na cabeça do paciente e com base em sensores produz movimento. O mesmo tipo de estudo foi realizado por Wei¹⁹ que avaliou um equipamento similar em um ambiente fechado. Lee²⁰ detecta o abrir e fechar da boca para movimentos extras e estuda o problema da ergonomia da cabeça. Define ainda, angulações mínimas e máximas para o indivíduo e o tempo total de uso da cadeira sem cansar o paciente. Destaca-se também os trabalhos de Braga²¹, Halawani³, Perez²², Ribeiro²³, Bailey²⁴.

Ao longo dos trabalhos analisados percebe-se que aspectos da visão computacional sempre estão associadas a algum tipo de hardware que dá suporte para movimentação de cadeira tornando-a quase sempre em um equipamento estranho, constrangedor e que depende de treinamento específico do paciente o que pode vir a causar uma difícil adaptação ao equipamento. Neste trabalho adota-se uma abordagem diferente, baseada em dois paradigmas da visão computacional, a técnica de classificadores em cascata e modelos

deformáveis para obter um modelo de guiagem da cadeira por um paciente que não dependa de treinamento prévio e sem a utilização de sensores ou quaisquer acoplamentos à cabeça do usuário.

O modelo proposto detecta e rastreia os movimentos da cabeça e algumas expressões faciais que são traduzidas em comandos para a cadeira. O modelo foi testado num protótipo, mas sem a presença de humanos, se revelando bastante estável. Obteve-se uma taxa de acerto na navegação da ordem de 97%.

Metodologia

O modelo

O modelo computacional proposto busca utilizar duas ações do usuário. A primeira que, a partir do estado do olho (aberto, fechado), informa à cadeira duas possíveis ações: a) seguir em frente; b) parar.

Cada comando é acionado quando o outro está inativo, ou seja, se o estado atual for “seguir”, o seguinte será “parar” e vice-versa. A caracterização dessa ação se dá num movimento de fechamento dos olhos de maneira não

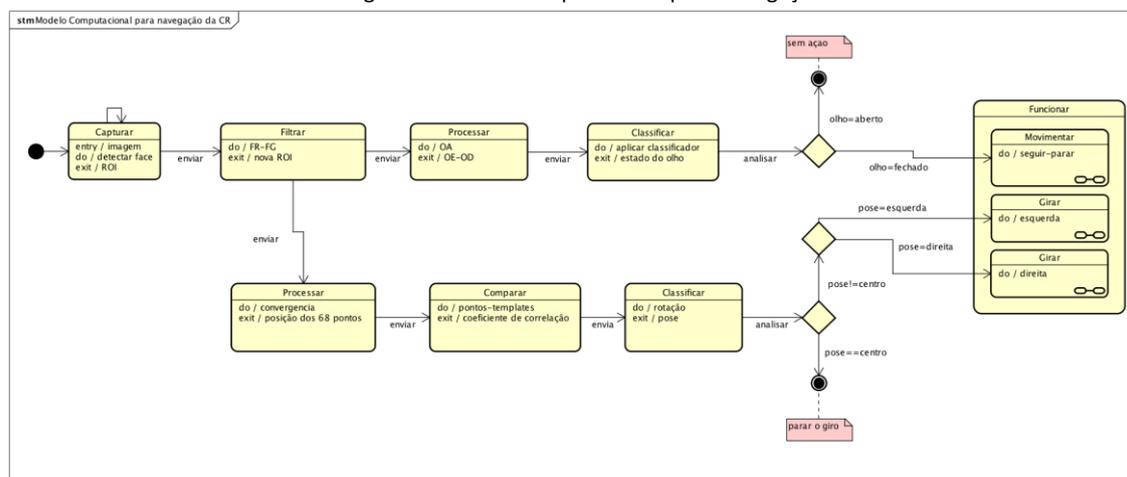
natural, ou seja, permanecer com o olho fechado, por um período de tempo superior a 2000 milissegundos. Segundo Ekman²⁵ o olho humano demora 280 milissegundos para piscar e, portanto existe uma diferenciação razoável entre o ato natural de piscar e a ação de controle do modelo.

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 de “girar” a CR para esquerda ou direita e “parar o giro” quando a cabeça retornar à posição centralizada.

O modelo apresentado na Figura 1, 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°.

Figura 1: Modelo computacional para navegação da CR



A operação de detectar face visa à redução do espaço de busca e é realizada utilizando a técnica de Viola²⁶ ao qual aplicou-se 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 menos 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 (<200lux) utilizando o filtro de retina²⁷ (FR-Filtro de Retina).

A ROI é enviada, simultaneamente, para extração da região do olho por meio de operações algébricas - OA²⁵ 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 completo. Nesse momento, o modelo de forma ativa (ASM) se revela extremamente eficiente na variação de escala.

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

A cadeira

O modelo foi implementado para um protótipo fabricado pela freedom carbon, produzido pela Freedom Veículos Elétricos Ltda, que possui duas rodas traseiras motorizadas com dois movimentos frente e ré, que utilizam baterias de 12V/45Ah, um joystick tradicional e um módulo de controle (arduino¹).

Utilizou-se um notebook com processador de 1.6GHz dual-core Intel Core i5 (Turbo Boost up to 2.7GHz) com 3MB e 4GB de memória. Além disso, dispomos para conexão entre o computador e a cadeira um placa do microcontrolador baseado no ATmega328 (arduino UNO), que dispõe de 14 pinos digitais de entrada/saída, 6 entradas analógicas e conexão USB.

Sistema de navegação para ambientes internos

Duas expressões foram escolhidas para realizar as ações sobre a cadeira: o fechar dos olhos e o movimento lateral da cabeça. A escolha destes se dá pelo fato de serem movimentos naturais aos comandos.

1. Fechar os olhos - Esta expressão facial não é afetada pela patologia e como o

¹ **Arduino**, é uma plataforma de prototipagem eletrônica de hardware livre e de placa única, projetada com um microcontrolador Atmel AVR com suporte de entrada/saída embutido, uma linguagem de programação padrão, a qual tem origem em Wiring, e é essencialmente C/C++. O objetivo do projeto é criar ferramentas que são acessíveis, com baixo custo, flexíveis e fáceis de se usar por artistas e amadores. Principalmente para aqueles que não teriam alcance aos controladores mais sofisticados e de ferramentas mais complicadas.

movimento natural de abrir e fechar os olhos dura milissegundos, seu contraponto torna-se ideal para iniciar ou encerrar o movimento da cadeira.

2. Movimento lateral da cabeça - O paciente tem um grau de liberdade para movimentação da cabeça o que permite tanto um grau satisfatório de acerto na identificação da posição como no conforto na navegação.

Fechar os olhos - Comando Avançar/Parar

Devido à natureza específica desta implementação, a imagem de entrada sempre conterá uma face. A câmera fica na cadeira a uma distância média de 1m do indivíduo com um ângulo de inclinação mínimo 30° e máximo de 60° além disso a imagem terá sempre uma boa condição de luz (>100lux). Assim, para detecção da face utiliza-se o modelo Viola²⁸ amplamente utilizado e com reconhecida eficiência²⁹. A detecção do estado dos olhos utiliza a premissa que se um detector facial consegue detectar com sucesso a face humana então ele deverá reconhecer também um olho fechado. Para realização e verificação dessa hipótese deve-se treinar um classificador que identifique este objeto na imagem de entrada.

Este método que basicamente combina quatro conceitos-chave, é composto por: 1. Características retangulares, chamadas de recursos de Haar; 2. Imagem Integral; 3. Algoritmo de aprendizagem (AdaBoost); 4. Um classificador em cascata.

As características de Haar codificam a existência de contrastes entre as regiões orientadas da imagem. Pode-se definir qualquer característica de Haar em uma janela de detecção com W × H pixel conforme a formula:

$$C = \sum_{i=1}^N \omega_i \text{RecSum}(r_i)$$

Em que ω_1 é arbitrariamente escolhido como fator de peso e $\text{RecSum}(r_i)$ é a soma das intensidades dos pixels. O retângulo r_i é descrito como uma função com cinco parâmetros: $r = (x, y, w, h, \theta)$, onde x, y são as coordenadas da posição superior do conjunto de pixels, w e h define as dimensões do retângulo, e $\theta = 0^\circ, 45^\circ$ representa o grau de rotação. A presença de uma característica de Haar numa imagem de treinamento é determinada pela subtração do valor do pixel médio da região pelo valor médio do pixel da região clara. Se a diferença for acima

de um limiar (definido durante a aprendizagem), dizemos que a característica está presente.

A imagem integral cria, a partir da entrada matricialmente pensada, uma nova representação de saída, somando-se os valores de cada pixel à esquerda e acima, inclusive. Essa representação aumenta a velocidade na extração de características, visto que passa a depender apenas de uma operação de subtração simples entre dois pixels.

De posse do conjunto de dados a próxima etapa consiste no treinamento utilizando o algoritmo de aprendizagem, nele utilizamos imagens de 24x24 com exatamente 10000 imagens de olhos fechados coletados em Salvador com iluminação controlada.

A idéia chave do algoritmo de aprendizagem é transformar todos os dados em um conjunto de classificadores fracos que são então selecionados e combinados para formarem um classificador forte. De posse do limiar, os dados são submetidos a uma árvore de decisão.

Resultados do Classificador

Para obter uma detecção precisa do classificador, vários parâmetros podem ser modificados durante o processo de treinamento. A influência desses parâmetros altera a complexidade dos classificadores fracos e conseqüentemente aspectos como falsos positivos e positivos são influenciados. Na Tabela 1 apresentam-se os classificadores criados e seus respectivos parâmetros.

Na Tabela 1 temos que MinHitRate é 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; Wcout - contagem máxima de árvores falsas para todas as etapas da cascata de classificadores²⁸.

Estimação da pose da cabeça - Movimento lateral da cabeça

O controle da direção em uma cadeira de rodas tem sido usado de diversas formas como destacado em Yi³⁰, Jaffe³¹, Manogna¹⁸ e Zhao⁵. Frequentemente a solução proposta envolve o uso de algum hardware específico para captar a rotação da cabeça. Nesta proposta verifica-se a

viabilidade do uso da rotação da cabeça, mas sem o uso de hardware e sim utilizando marcos facial que compõe um modelo deformável ou flexível. Os modelos de forma ativa (ASM- Active Shape Model) foi apresentado por Cootes³², e trata-se de um modelo estatístico que é frequentemente utilizado em diversas aplicações Chen³³, Birk³⁴, Hommel³⁵. No modelo, 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. Cootes³² ainda utiliza a distância de Mahalanobis² entre os marcos e da 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 num objeto em outra fonte. De posse destes 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, estes com um número médio de 2% de falsos positivos. Por simplificação tomaremos o modelo em duas etapas iniciais:

1. Construção do PDM- Modelo de Distribuição de Pontos;
2. Detecção dos níveis de cinza (LGL) em cada ponto;

PDM- Modelo de Distribuição de Pontos

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

Etiquetagem do conjunto de treino

Nesta etapa determina-se a localização dos pontos mais adequados nas imagens do conjunto de treino, sendo esta realizada

² A **distância de Mahalanobis** é uma medida de [distância](#) introduzida pelo matemático [indiano Prasanta Chandra Mahalanobis](#) em [1936](#). Formalmente, a distância de Mahalanobis entre um grupo de valores com média. É baseada nas [correlações](#) entre variáveis com as quais distintos padrões podem ser identificados e analisados.

manualmente e considerando locais que melhor descrevem o objeto que neste estudo será uma face. Este conjunto é formado por pontos-chave da face que são exatamente o redor da face, dos

olhos, nariz e boca que são pontos de elevado valor de curvatura, vértice e bifurcação.

Tabela 1: Parâmetros utilizados para a realização do estudo

Parâmetro	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

Essa etapa é realizada em todo o conjunto de treino o que equivale a $N = 120$ imagens e $n = 68$ pontos marcados em cada, onde o ponto (x_{ij}, y_{ij}) pertence a j -ésima coordenada da forma da i -ésima imagem onde o tamanho de x_i é $2n$ e $1 \leq i \leq N$.

Alinhamento das formas

Por meio de rotações, translações e redimensionamento (escala), passa-se a alinhar os pontos com o objetivo de minimizar a soma quadrática das distâncias entre pontos equivalentes. Iterativamente, utilizam-se a forma:

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

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 redimensiona-lo de forma que $|\bar{x}| = 1$;
3. Fazer \bar{x}_0 como a 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 redimensiona-la de forma que $|\bar{x}| = 1$;
7. Repetir os passos anteriores até que o valor da forma média convirja.

Aparência dos níveis de cinza

Em cada ponto da forma os níveis de cinza descrevem a aparência típica da vizinhança que está localizada o contorno do ponto. Isto é, examinamos as regiões de níveis de cinza próximas.

Localização dos pontos característicos em uma nova imagem

De posse de uma nova imagem contendo o elemento treinado podemos utilizar o algoritmo para localizar a posição dos pontos característicos numa nova imagem seguindo o algoritmo 1 (Figura 2).

Treinou-se esse modelo com um número de 30 imagens de 5 pessoas diferentes com uma rotação da cabeça entre $0 \leq \theta \leq 80$, tomando 68 pontos característicos. Incluindo amostras de pessoas de pele clara e escura, com algum tipo de oclusão, cabelos curtos e longos, com e sem óculos e com iluminação controlada e constante. A resolução de vídeo original é de 800 por 600 pixels. As imagens da cabeça são cortadas e alinhadas usando um algoritmo de detecção de Viola-Jones²⁸ (Figura 2). Estas imagens foram utilizadas tanto para o treinamento do ASM como templates $T(x,y)$, que vão definir a posição da cabeça. Assim dada uma imagem $I(x,y)$ e um template $T(x,y)$ como as mesmas dimensões, a correlação normalizada entre eles é definida como:

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

Onde $\bar{I} = \frac{1}{n} \sum_{x,y} I$ e $\bar{T} = \frac{1}{n} \sum_{x,y} T$. 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. Note-se que a resposta dos

modelos é simétrica, evidenciando a simetria vertical visual da cabeça. Note também que a queda do vector de resposta de seu pico é monótona, gradual e tem um baixo desvio padrão (desvio padrão médio é de 0,0431).

Figura 2: Algoritmo utilizado para localização dos pontos característicos em uma nova imagem

Algorithm 1: algoritmo para localização dos pontos característicos numa nova imagem

Input : $P, \bar{x}, \bar{y}_j, C_{y_j}$
output: Localização dos pontos característico numa nova imagem

- 1 Localizar a face da imagem de entrada usando o Algoritmo de Viola-jones;
- 2 Localizar a forma média \bar{x} , no centro da imagem facial;
- 3 $b = 0$;
- 4 **repeat**
- 5 **for** $j = 1$ to n **do**
- 6 Calcular os níveis de cinza na nova imagem
- 7 $f_j = (f_{j0}, f_{j1}, \dots, f_{jn_x-1})^T$
- 8 Calcular as derivadas de f_j :
- 9 Normalizar as derivadas de f_{y_j} $f_{y_j} = \frac{df_j}{\sum_{k=0}^{n_x-2} |df_{jk}|}$
- 10 Minimizar a distância entre df_j e y_j usando a distância de Mahalanobis.
- 11 **end**
- 12 Obter o vetor de movimento e calcular o menor deles para deslocar os pontos
- 13 Movimentar os pontos até a posição correta
- 14 Determinar $b = P^T(xf - \bar{x})$
- 15 Limitar b baseado nos autovalores obtidos no conjunto de treinamento e atualizar b
- 16 Estimar $xf = \bar{x} + Pb$
- 17 **until** distância seja mínima;

Resultados e Discussões

Nas experiências que foram realizadas num ambiente controlado, a cadeira deveria seguir uma trajetória sem choques com as paredes e sempre para a área definida. A coleta destas informações foi manual, para cada cenário, que represente a movimentação básica de um paciente e buscou-se obter: Acurácia; Precisão; Taxa de falso positivos.

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.

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, comparou-se o deslocamento realizado pelo modelo

(U1,U2,U3,U4,U5 e U6) comparado ao deslocamento realizado pelo joystick, que é a forma de navegação padrão da cadeira. 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, descrito na Figura 3, utilizou-se uma sala de aula na qual o indivíduo deveria sair e chegar no mesmo ponto dando uma volta completa ao redor de uma mesa de 1.8m de comprimento por 1.5m de largura. Nesta rotina observou-se o comportamento da cadeira quanto à rotação em si, nível de deslizamento ao realizar o movimento e também os elementos de estudo principais do modelo com acurácia, resposta e erro.

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 apresentados na Figura 4.

Os testes representados nas Figuras 5 e 6 apresentam um contorno em L, muito comum em apartamentos e casas, para o deslocamento entre um cômodo e outro, respectivamente 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.

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. Os resultados numéricos dos dados estão descritos na Tabela 2.

Tabela 2: Resultado geral dos testes realizados

Cenário	Falso Positivo	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%

Figura 3: Primeiro Cenário - uma sala de aula na qual o indivíduo deveria sair e chegar no mesmo ponto dando uma volta completa ao redor de uma mesa.

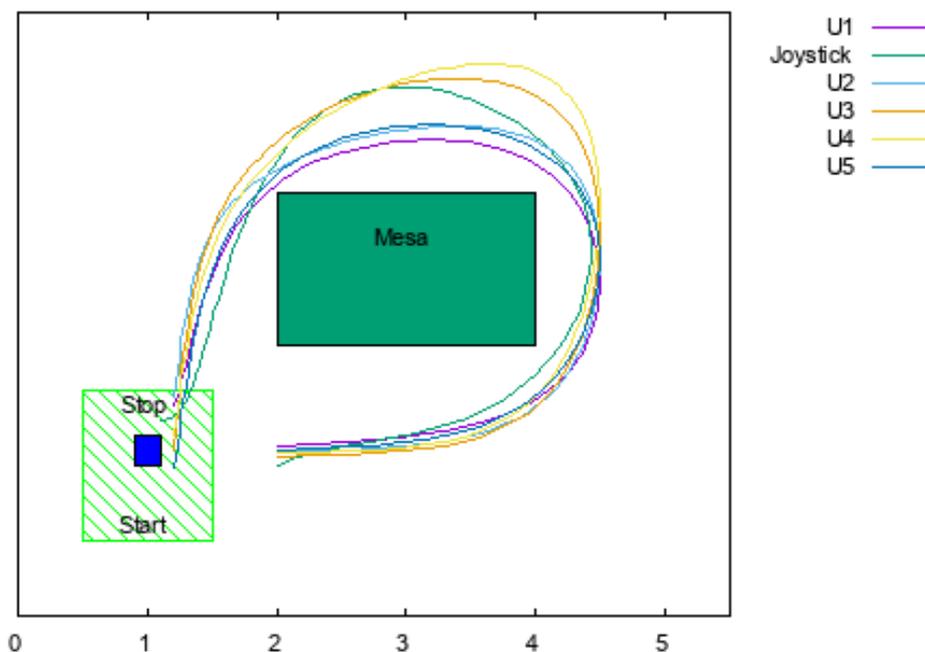


Figura 4: Cenário onde, utilizando duas mesas separadas o usuário é convidado a realizar o deslocamento do início até o ponto de parada, passando entre as duas.

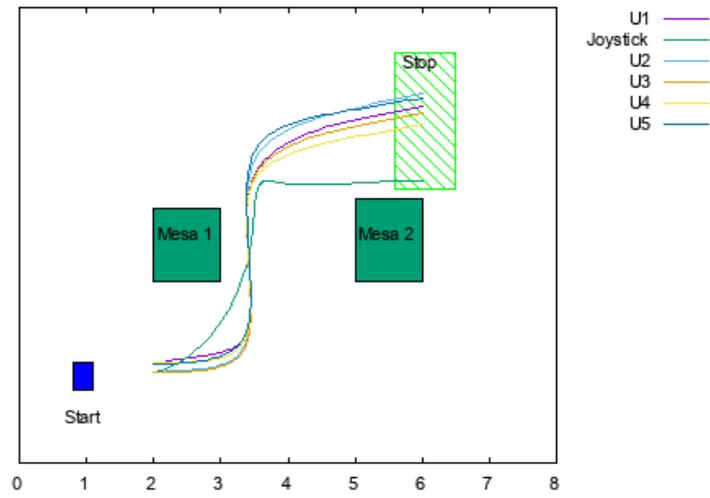


Figura 5: Cenário utilizando um contorno em L, com dois obstáculos.

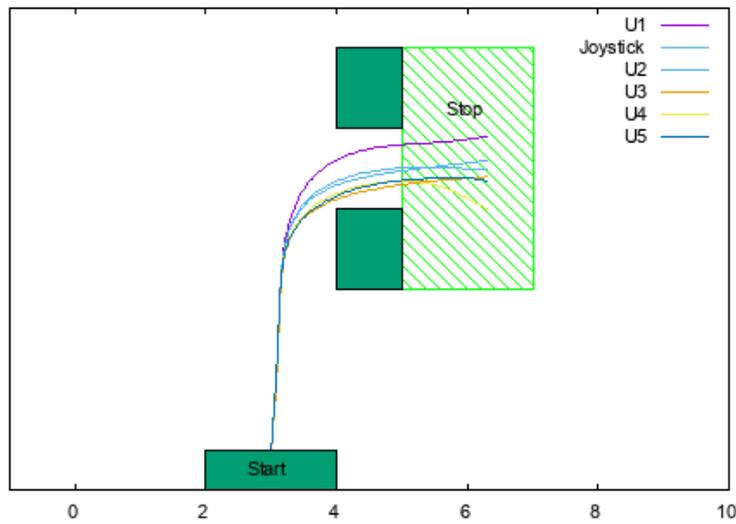
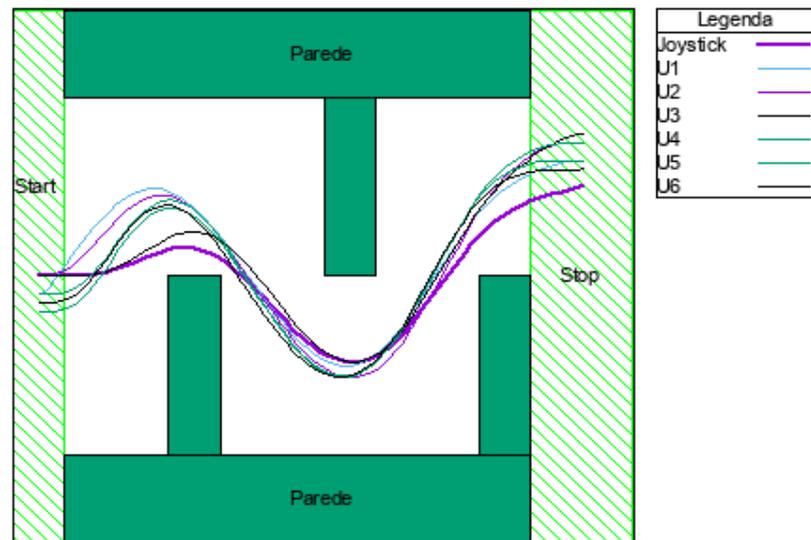


Figura 6: Cenário utilizando um contorno em L, com tres obstáculos.



Discussão

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.

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 da imagem, gênero, pigmentação da pele, idade e oclusão. Observou-se que a resposta do modelo à rotação de face foi bem satisfatória com um erro 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 porque os testes não foram realizados com os voluntários na cadeira. Notou-se que o erro médio vai para a casa de 2% quando se estabelece uma distância diferente da qual os templates foram treinados.

Notou-se 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 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.

O software mostrou-se estável quanto a travamentos e durante a realização dos testes se comportou como esperado em relação ao projetado. Algumas colisões foram observadas durante os testes no protótipo em todos os cenários apresentados.

Conclusões

Neste trabalho foi proposto controlar a navegação de uma cadeira de rodas baseada em expressões faciais e no movimento da cabeça sendo que o primeiro baseado em características de Haar e o segundo em modelos de forma ativa. A detecção do fechamento dos olhos indica as ações de seguir e parar enquanto o giro da cabeça indica a rotação da cadeira em torno do

seu eixo. Estes elementos mostraram-se bastante simples e de fácil assimilação por parte dos indivíduos. Os resultados experimentais mostraram que essa abordagem é promissora. Embora os testes do sistema atual tenha ainda um erro acentuado na detecção da rotação do rosto, utilizar-se-á um número maior de templates para correção da falha. O número de falsos positivos encontrado nos testes não foi significativa, cerca de 0.01% de todas as detecções analisadas e, portanto, um fator de pouca relevância mais que, quando da utilização de usuários com a real deficiência, deve ser tratado.

Para muitas pessoas uma cadeira de rodas (CR) representa a principal conexão com o mundo e com ela se deslocando, trabalhando ou realizando sua socialização diária. Para muitos, se não todos, a forma como a cadeira o 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. A preocupação deste trabalho foi tornar o paciente com tetraplegia, capaz de deslocar sua CR num ambiente interno de forma suave e sem esforço com domínio total do movimento e independência.

Centrado na perspectiva das Tecnologias Assistivas (TA) os resultados levam a crer que com alguns ajustes relativos a questões como desvio de obstáculos e deslizamento das rodas, é possível apresentar um protótipo para pacientes reais e validar completamente esta forma de navegação promovendo assim aspectos já discutidos e que serão propostos pelo usuário final.

Nos trabalhos futuros será feito a incorporação de um algoritmo de desvio de obstáculos como também um para controlar melhor a cadeira em corredores e portas bem como acrescentar módulos para controle por voz e deslizamento das rodas.

Referências

1. de Geografia e Estatística IBGE IB. Características da população e dos domicílios Resultados do universo. Censo Demográfico 2010. 2011;1:161.
2. Braga RAM. Plataforma de desenvolvimento de cadeiras de rodas inteligentes. Faculdade de Engenharia da Universidade do Porto - Programa de Doutoramento em Engenharia Informática. Porto, Portugal; 2012.

3. Halawani A, Ur Réhman S, Li H, Anani A. Active vision for controlling an electric wheelchair. *Intelligent Service Robotics*, Springer Science & Business Media BV. 2012;Vol.5(2):89–98.
4. Song Y, Luo Y, Lin J. Detection of Movements of Head and Mouth to Provide Computer Access for Disabled. *Technologies and Applications of Artificial Intelligence (TAAI)*, International Conference on. 2011;p. 223–226.
5. Zhao Z, Wang Y, Fu S. Head Movement Recognition Based on Lucas- Kanade Algorithm. *Computer Science & Service System (CSSS)*, International Conference on. 2012;p. 2303–2306.
6. de Melo VR. Avaliação da qualidade de vida de paciente com lesão medular acompanhados em regime ambulatorial [Dissertação de Mestrado]. Universidade Federal de Pernambuco CCS. *Neuropsiquiatria e Ciências do Comportamento*; 2009.
7. da Cruz DMC, Ioshimoto MTA. TECNOLOGIA ASSISTIVA PARA AS ATIVIDADES DE VIDA DIÁRIA NA TETRAPLEGIA COMPLETA C6 PÓS-LESÃO MEDULAR. *Revista Triângulo -Revista do Programa de Pós-Graduação em Educação da UFMT*. 2010 jul/dez;3(2):177–190.
8. Dymond E, Potter R. Head movements for control of assistive technology. *Engineering in Medicine and Biology Society, 14th Annual International Conference of the IEEE*. 1992;4:1527–1528.
9. Sumit Basu IE, Pentland A. Motion Regularization for Model-Based Head Tracking. In *Proceedings of the IEEE Int'l Conf on Pattern Recognition (ICPR '96) Vienna, Austria*. 1996;3:611–616.
10. Kiruluta A, Eizenman M, Pasupathy S. Predictive Head Movement Tracking Using a Kalman Filter. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS?PART B: CYBERNETICS*. 1997;27(2):326–331.
11. Adachi Y, Kuno Y, Shimada N, Shirai Y. Intelligent Wheelchair Using Visual Information on Human Faces. *proceedings of the IEEE/RSJ Intl Conference on Intelligent Robots and Systems*. 1998;p. 354–359.
12. Kuno Y, Nakanishi S, Murashima T, Shimada N, Shirai Y. Intelligent wheelchair based on the integration of human and environment observations. *Information Intelligence and Systems, Proceedings 1999 International Conference on*. 1999;p. 342–349.
13. Bergasa LM, Mazo M, Gardel A, Barea R, Boquete L. Commands generation by face movements applied to the guidance of a wheelchair for handicapped people. *Pattern Recognition, Proceedings 15th International Conference on*. 2000;4:660–663.
14. Taylor PB, Nguyen HT. Performance of a Head-Movement Interface for Wheelchair Control. *Proceedings of the 25th Annual International Conference of the IEEE EMBS*. 2003;p. 17–21.
15. Nguyen HT, King LM, 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*. 2004;p. 1–5.
16. King LM, Nguyen HT, Taylor PB. 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*. 2005;p. 2063–2066.
17. Liao WK, Cohen I. Classifying Facial Gestures in Presence of Head Motion. *Computer Vision and Pattern Recognition - Workshops CVPR Workshops IEEE Computer Society Conference on*. 2005;p. 77.
18. Manogna S, Vaishnavi S, Geethanjali B. Head Movement Based Assist System for Physically Challenged. In: *2010 4th International Conference on Bioinformatics and Biomedical Engineering*; 2010. p. 1–4.
19. Wei L, Hu H, Lu T, Yuan K. 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*. 2010;p. 387–392.
20. Lee CS, 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*; 2011. p. 847–852.
21. Braga RAM, Petry M, Moreira AP, Reis LP. Concept and Design of the Intellwheels Platform for Developing Intelligent Wheelchairs. In: *Cetto JA, Ferrier JL, Filipe J, editors. Informatics in Control, Automation and Robotics*. vol. 37 of *Lecture Notes in Electrical Engineering*. Springer Berlin Heidelberg; 2009. p. 191–203.
22. Perez E, Soria C, Nasisi O, Bastos TF, Mut V. Robotic wheelchair controlled through a vision-based interface. *Robotica*. 2012 9;30:691–

708. Available from: http://journals.cambridge.org/article_S0263574711000919.
23. Ribeiro MP. CAMIN - CADEIRA MÓVEL INTELIGENTE: UM SISTEMA DE NAVEGAÇÃO UTILIZANDO UMA ADAPTAÇÃO DO MÉTODO DOS CAMPOS POTENCIAIS ARTIFICIAIS [Dissertação de Mestrado]. Curso de Mestrado em Sistema e Computação do Instituto Militar de Engenharia,. INSTITUTO MILITAR DE ENGENHARIA Praça General Tibúrcio, 80-Praia Vermelha Rio de Janeiro-RJ CEP 22290-270; 2007. Available from: http://www.comp.ime.br/pos/images/repositorio-dissertacoes/2007-Monael_Ribeiro.pdf.
24. Bailey M, Chanler A, Maxwell B, Micire M, Tsui K, Yanco H. Development of Vision-Based Navigation for a Robotic Wheelchair. In: 2007 IEEE 10th International Conference on Rehabilitation Robotics; 2007. p. 951–957.
25. Ekman P, Rosenberg EL. What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). Series in Affective Science. Oxford University Press; 2005. Available from: <https://books.google.com.br/books?id=UXapcWqtO-sC>.
26. Viola P, Jones MJ. Robust real-time face detection. International journal of computer vision. 2004;57(2):137–154.
27. Benoit A, Caplier A, Durette B, Héroult J. Using human visual system modeling for bio-inspired low level image processing. Computer vision and Image understanding. 2010;114(7):758–773.
28. 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. vol. 1; 2001. p. 511–518 vol.1.
29. Turk M, Pentland A. Eigenface for recognition. J Cognitive Neuroscience 3(1). 1991;p. 71–96.
30. Yi Z, Xiaolin F, Yuan L. Intelligent wheelchair system based on sEMG and head gesture. The Journal of China Universities of Posts and Telecommunications. 2015;22(2):74 – 95. Available from: <http://www.sciencedirect.com/science/article/pii/S1005888515606422>.
31. Jaffe DL. An ultrasonic head position interface for wheelchair control. Journal of Medical Systems. 1982;6(4):337–342. Available from: <http://dx.doi.org/10.1007/BF00992877>.
32. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models- their training and application. Computer vision and image understanding. 1995;61(1):38–59.
33. Chen X, Udupa JK, Bagci U, Zhuge Y, Yao J. Medical Image Segmentation by Combining Graph Cuts and Oriented Active Appearance Models. IEEE Transactions on Image Processing. 2012 Apr;21(4):2035–2046.
34. Birkbeck N, Jagersand M. Visual tracking using active appearance models. In: Proc. First Canadian Conf. Computer and Robot Vision; 2004. p. 2–9.
35. Hommel S, Handmann U. Realtime AAM based user attention estimation. In: Proc. IEEE 9th Int. Symp. Intelligent Systems and Informatics; 2011. p. 201–206.

Endereço para Correspondência

Universidade do Estado da Bahia - UNEB
Rua Silveira Martins, 2555, Cabula.
Salvador - Ba
CEP.: 41150-000

e-mail: marcos.b.figueredo@gmail.com

Recebido em 26/01/2017
Aprovado em 27/03/2017
Publicado em 29/03/2017