

# Protótipo de um Simulador para Cadeiras de Rodas Guiadas por Expressões Faciais: Estudos Preliminares

Breno de Paula Fernandes<sup>1</sup>, Vinicius A. Saueia da Silva<sup>1</sup>, Hemerson Pistori<sup>1</sup>

<sup>1</sup>Universidade Católica Dom Bosco – UCDB

Grupo de Pesquisa em Engenharia e Computação – GPEC

Av. Tamandaré, 6000 – Jd. Seminário – CEP: 79117-900 – Campo Grande – MS

bpferrandes@acad.ucdb.br, vsaueia@acad.ucdb.br, pistori@ec.ucdb.br

**Abstract.** *This paper describes a computer vision system targeted at the digital inclusion for people with special needs. A system prototype that can produce controlling commands for an electric wheelchair, from face images captured by a webcam, will be presented, together with the skin segmentation technique used for face tracking. The prototype is a module of the SIGUS platform, a programming environment for computer vision based interfaces.*

**Keywords:** *Computer Vision, Digital Inclusion*

**Resumo.** *Este artigo descreve uma aplicação de visão computacional para a inclusão digital de pessoas com necessidades especiais. Uma versão preliminar de um protótipo de um simulador capaz de gerar comandos para uma cadeira de rodas motorizada, a partir de imagens capturadas por meio de uma câmera Web, será apresentado, juntamente com a técnica de segmentação de pele utilizada para rastrear a face humana nas imagens capturadas. O protótipo é parte de um projeto maior, chamado SIGUS, que tem como objetivo a criação de um ambiente de programação para interfaces baseadas em visão computacional.*

**Palavras-chave:** *Visão Computacional, Inclusão Digital.*

## 1. Introdução

De acordo com a Organização Mundial de Saúde (OMS), existem mais de 600 milhões de pessoas com necessidades especiais no mundo. No Brasil, foi realizada uma pesquisa no ano de 2000, pelo Instituto Brasileiro de Geografia e Estatística (IBGE), que detectou a existência de 25 milhões de pessoas com algum tipo de necessidade especial, o que corresponde a 15% da população total [Passafaro 2002].

Necessidades especiais podem surgir por meio de anomalias congênitas ou adquiridas, parciais ou totais, de estrutura ou função fisiológica ou anatômica, podendo ocasionar restrições da capacidade orgânica e da habilidade funcional. O foco deste trabalho, que se insere na área da inclusão digital de pessoas com necessidades especiais, são as pessoas cujos membros superiores e inferiores estão impossibilitados de realizar movimentos e que dependem de cadeiras de rodas para a sua locomoção. O censo brasileiro de 2000 observou que existem 955.287 pessoas tetraplégicas, paraplégicas ou hemiplégicas permanentes.

O desenvolvimento de áreas como robótica, controle automático e visão computacional, entre outras, tem proporcionado o surgimento de diversos equipamentos capazes de tornar mais simples e autônoma a vida de pessoas que dependem de algum tipo de dispositivo auxiliar para locomoção, como paraplégicos e tetraplégicos. No caso dos tetraplégicos, a cadeira de rodas convencional coloca o usuário na dependência de uma outra pessoa, para conduzi-la. Para contornar esse problema, pesquisas estão sendo realizadas visando a construção de dispositivos que permitam que uma cadeira de rodas motorizada possa ser controlada por outros meios de comunicação homem-máquina, como, por exemplo, comandos de voz [Katevas et al. 1997, Simpson and Levine 1997].

Outros mecanismos de controle utilizando, por exemplo, movimentos da boca [Neveryd et al. 1999], língua e cabeça [Yanco and Gips 1997, Carlson et al. 1994], também têm sido experimentados. Um recente estudo indica, no entanto, que a utilização desses mecanismos em situações reais, pode ser muito complicada e frustrante para os usuários [Fehr et al. 2000].

A plataforma SIGUS<sup>1</sup> é um ambiente computacional de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais, que tem como objetivo aumentar o número de programas computacionais destinados a essas pessoas [Pistori and Pereira 2004, Pistori and Neto 2004]. Dois protótipos já foram desenvolvidos utilizando essa plataforma: um jogo da velha que pode ser acionado por meio de movimentos da face e um editor de símbolos alfabéticos da Língua Brasileira de Sinais. Neste artigo será descrito um terceiro protótipo construído por meio da plataforma SIGUS. Esse protótipo simula um ambiente para controle de cadeiras de rodas motorizadas utilizando expressões faciais capturadas por uma webcam.

A Figura 1 mostra exemplos de expressões que um usuário poderia fazer para gerar comandos para guiar uma cadeira de rodas. O usuário poderia olhar para cima para a cadeira seguir em frente, para baixo para parar e para esquerda ou direita para fazer curvas.



**Figura 1. Exemplo de expressões capazes de gerar comandos para uma cadeira de rodas**

Para o desenvolvimento do protótipo aqui apresentado utilizou-se, em grande parte, de classes já implementadas na plataforma SIGUS, como um módulo de extração de características baseado em momentos estatísticos e o módulo de aprendizagem automática supervisionada. A principal contribuição do trabalho aqui apresentado foi o aprimoramento do módulo de rastreamento baseado na detecção da pele humana, cujos principais conceitos serão descritos na seção 3. A seção 4 trata do protótipo do simulador

<sup>1</sup>Disponível gratuitamente, e na forma de software livre, sob os termos da *General Public License (GPL)*, em <http://www.gpec.ucdb.br/sigus>

de cadeira de rodas guiada por movimentos da face, apresentando alguns detalhes sobre as funcionalidades e a interface implementada. Na última seção têm-se a conclusão e sugestões para trabalhos futuros.

## 2. Trabalhos Correlatos

Existem várias pesquisas sobre cadeira de rodas inteligentes guiadas com auxílio de dispositivos especiais, como sensores ultra-sônicos e câmeras web. Uma cadeira de rodas inteligente proposta por [Bergasa et al. 2000] utiliza um rastreador de face que consegue captar comandos para movimentar uma cadeira de rodas. O rastreador de faces é capaz de capturar imagens da face e um interpretador analisa as imagens capturadas, definindo, assim, os comandos que serão responsáveis pela movimentação da cadeira de rodas.

Um outro tipo de cadeira de rodas utilizando sensores combinados com dispositivos de captura de imagens [Kuno et al. 2000, Schlling et al. 1998] é capaz de se movimentar de maneira autônoma, em linha reta, desde que não haja objetos no caminho. Nessa cadeira de rodas os sensores são utilizados para localizar a posição do usuário em um ambiente e para detectar obstáculos. Foi utilizado por [Kuno et al. 2000] uma cadeira de rodas composta por 16 sensores ultra-sônicos e duas câmeras. A primeira câmera calcula a posição da cadeira de rodas no ambiente e captura imagens que estão à frente da cadeira e a segunda captura imagens da face do usuário.

É possível se construir uma cadeira de rodas controlada pela direção do olhar utilizando apenas dispositivos de captura de imagens. Foi desenvolvido por [Matsumoto et al. 2001] uma cadeira de rodas com interface simplificada para o usuário, utilizando apenas comandos gerados pela cabeça e olhar. Para isto, desenvolveu-se um sistema capaz de rastrear a face e a direção do olhar simultaneamente. Assim, quando a cabeça do usuário da cadeira de rodas tem a mesma direção que o olhar, gera-se comandos de movimentação para a cadeira de rodas e a cadeira de rodas segue na direção escolhida.

## 3. Segmentação de Pele Humana

Através da segmentação de pele humana em uma imagem é possível se determinar quais os *pixels* estão relacionados com pele e quais não estão [Hsu et al. 2002]. A detecção da pele é uma fase importante para o rastreamento de seres humanos, pois é nessa fase que se pode distinguir quais os *pixels* pertencem a seres humanos. Foi desenvolvida uma técnica por [Albiol et al. 2000] para localizar regiões que contenham face humana em uma imagem, nas quais a iluminação, forma, cor e pose da face variam muito. Para isso foi necessário integrar algumas informações importantes como textura, cor da pele, formas e posições. Em [Pistori 2003] foram utilizadas técnicas para segmentação de pele baseadas em aprendizagem supervisionada a partir de amostras de pele capturadas em uma fase de calibração do sistema. Estudos comparativos sobre diferentes técnicas para detecção de pele podem ser encontrados em [Terrillon et al. 2000, Vezhnevets et al. 2003, Gomez et al. 2002].

No presente trabalho foi implementado um segmentador de pele, *pixel a pixel*, baseado em aprendizagem estatística. O *pixel* é um elemento importante e representa numericamente cada ponto de uma imagem. Para distinguir *pixels* pele de não pele podem ser utilizados diferentes espaços de cores, obtendo-se assim diferentes características.

Espaços de cores são formas de se representar cores como vetores de um espaço vetorial, normalmente, tridimensional. Operações de mudança de base nesses espaços vetoriais geram diferentes representações para as cores de uma imagem, permitindo assim que determinadas características interessantes presentes em uma imagem sejam destacadas. Dois dos espaços de cores mais citados na literatura são o RGB (*Red, Green, Blue*) e o HSV (*Hue, Saturation, Value*).

No espaço HSV, utilizado neste trabalho, cada cor possui três componentes: matiz, saturação e luminância. A matiz está relacionada com o comprimento de onda espectral que produz a percepção da cor. A saturação está relacionada com a pureza da cor. Uma cor pura é aquela que tem em sua composição poucas combinações de diferentes cores, são as chamadas cores vivas. Cores esbranquiçadas e acizentadas são cores pouco saturadas. As cores vermelho e rosa, por exemplo, possuem a mesma matiz, mas o rosa é menos saturado que o vermelho, pois combina o vermelho com o branco. A luminância está relacionada com o brilho da cor, ou seja, com o quanto ela é mais clara ou mais escura.

Neste trabalho utilizaram-se as informações de matiz (*Hue*) e saturação (*Saturation*), extraídas de uma imagem convertida para o formato HSV (*Hue, Saturation and Value*), e ignorando-se a componente relacionada com a luminância (*Value*). Ignora-se a componente referente a luminância do *pixel*, pois em sistemas de reconhecimento de pele humana não é interessante que o sistema deixe de classificar pele por que a luminância foi alterada. Essa componente pode ser alterada facilmente, por exemplo, devido à variação de iluminação do ambiente.

A Figura 2 mostra exemplos de resultados do segmentador de pele humana em imagens de pessoas com diferentes tons de pele. Uma vez que o sistema pode ser calibrado para cada diferente usuário, com a aprendizagem supervisionada sendo realizada por usuário, é possível atingir um bom desempenho quanto à taxa de acerto, na segmentação da pele. Essa calibração consiste, basicamente, em ajustar, em um pequeno retângulo na tela, uma região da imagem que contenha *pixels* relacionados à pele humana (ver região superior, central, na altura da testa, nas imagens à esquerda da Figura 2).

#### 4. Protótipo Implementado

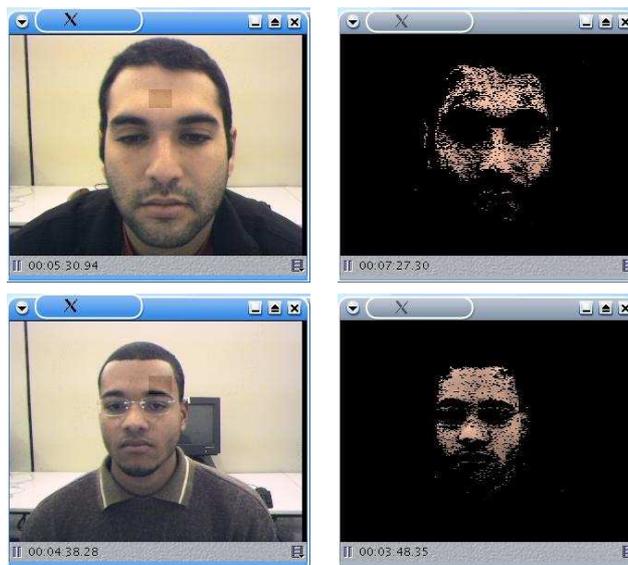
O protótipo do simulador foi desenvolvido no Laboratório de Pesquisa e Desenvolvimento (LP&D) do Grupo de Pesquisa em Engenharia e Computação da Universidade Católica Dom Bosco (UCDB). Foi utilizado durante a pesquisa um computador Pentium IV 1.5 GHZ, 40 GB HD, 128 MB RAM e uma *webcam* Creative GoPlus III. Utilizou-se também a plataforma de desenvolvimento SIGUS, que integra os programas livres e gratuitos ImageJ e Weka, e o programa gratuito, mas não livre, JMF (*Java Media Framework*).

O ImageJ<sup>2</sup> é um *software* de processamento de imagens digitais, desenvolvido em Java, que disponibiliza implementações de vários algoritmos de processamento digital de imagens. O Weka<sup>3</sup> (*Waikato Environment for Knowledge Analysis*), também desenvolvido em Java, facilita a implementação e teste de técnicas de aprendizagem de máquina. O Java Media Framework<sup>4</sup> é um pacote utilizado para capturar imagens em tempo real

<sup>2</sup>Disponível em <http://rsb.info.nih.gov/ij/>

<sup>3</sup>Disponível em <http://www.cs.waikato.ac.nz/~ml/weka/>

<sup>4</sup>Disponível em: <http://java.sun.com/products/java-media/jmf/>



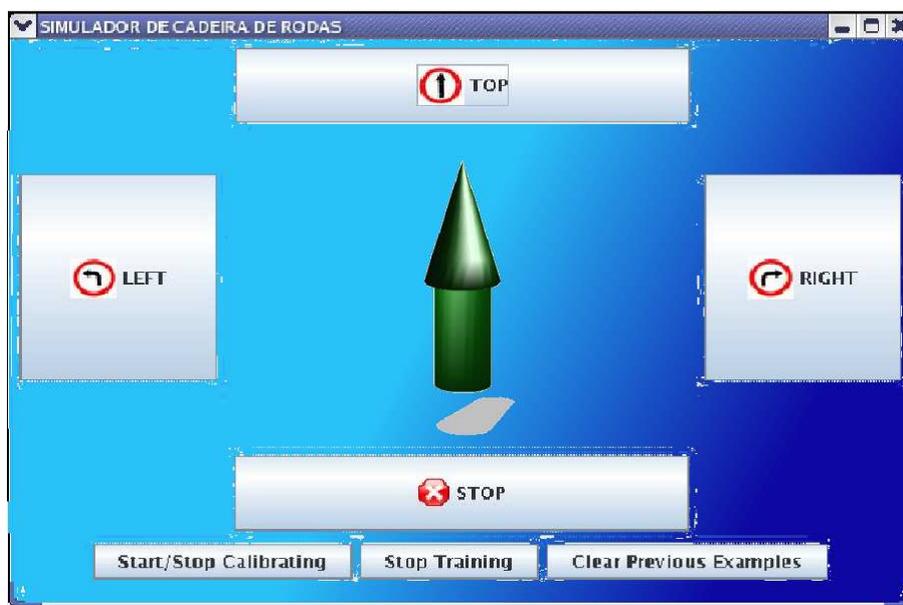
**Figura 2. Resultado da segmentação em pessoas com diferentes tons de pele**

por meio de uma *webcam* acoplada a um computador. A grande vantagem desse pacote é a portabilidade dos sistemas que utilizam a *webcam* para diferentes plataformas de *hardware* e *software*. Com a plataforma SIGUS é possível, de maneira relativamente simples, integrar os módulos de processamento de imagens com o módulo de aprendizagem de máquina, que podem funcionar em tempo real na interpretação de imagens em movimento.

O protótipo que simula a cadeira de rodas foi desenvolvido em Java e é composto de sete botões e uma seta. A seta representa a cadeira de rodas e os botões são utilizados para movimentar a seta por meio de diferentes configurações da face. Os botões para esses movimentos são: RIGHT (virar o eixo da cadeira para a direita), LEFT (virar o eixo para a esquerda), INIT (iniciar movimento da cadeira de rodas) e STOP (parar a cadeira de rodas). Os outros três botões estão relacionados com as fases de calibração do segmentador de pele humana (*Start/Stop Calibration*) e de aprendizagem supervisionada (*Stop Training* e *Clear Previous Exemples*). A Figura 3 mostra a tela principal do protótipo.

O funcionamento do protótipo consiste em três fases: calibração, aprendizagem e execução. Na fase de calibração o usuário deve, utilizando a resposta em tempo real apresentada na tela do computador, preencher um pequeno retângulo, no campo de visão da *webcam*, com uma amostra de *pixels* correspondentes à pele. Uma maneira simples de fazer isso é posicionando a testa ou a mão (se possível) na frente da *webcam*. Essa amostra de pele é utilizada pelo algoritmo de segmentação, que funciona em tempo real, apresentando para o usuário o resultado da segmentação (uma imagem binarizada) na tela do computador (um exemplo dessa imagem foi apresentado na seção 3.). A qualquer momento, o usuário (ou uma pessoa auxiliando o usuário), pode pressionar o botão que finaliza a fase de calibração e iniciar a fase de aprendizagem. A calibração pode ser refeita posteriormente, caso os resultados não estejam sendo satisfatórios, bastando para isso pressionar o mesmo botão.

Na fase de aprendizagem, o usuário deve associar, a cada um dos quatro botões



**Figura 3. Painel principal do protótipo**

de controle da cadeira, uma determinada configuração da face, realizando a configuração e pressionando o botão correspondente. A configuração da face deve ser mantida até que o botão seja pressionado novamente e o sistema pare de coletar exemplos daquela associação. O usuário pode repetir várias vezes esse procedimento, aumentando assim a quantidade de exemplos a serem utilizados pelo algoritmo de aprendizagem supervisionada. Assim que uma quantidade suficiente de exemplos (determinada empiricamente) é capturada, o usuário deve pressionar o botão de término da fase de treinamento (*Stop Training*). A partir desse momento, o sistema passa para a fase de execução, podendo ser agora controlado por meio da realização das configurações de face previamente informadas e associadas aos botões. O terceiro botão, *Clear Previous Examples*, é utilizado para reiniciar a fase de treinamento, apagando-se todos os exemplos obtidos anteriormente. Esse recurso é muito útil quando um exemplo incorreto é inserido no conjunto de treinamento. Isso pode acontecer, por exemplo, quando o botão que termina a fase de associação de uma configuração da face não é pressionado antes da mudança para uma outra configuração, e o sistema passa a capturar uma associação incorreta.

## 5. Conclusão

Dispositivos computacionais e robóticos para facilitar a inclusão digital de pessoas com necessidades especiais são cada vez mais comuns. Diferentes alternativas para controle e comunicação com esses dispositivos estão sendo testadas, como voz, gestos, movimentos da boca entre outras. Neste artigo, foi descrito uma versão preliminar do protótipo de um simulador para uma cadeira de rodas guiada por expressões faciais. Essa versão está disponível na forma de software livre e integra o ambiente de programação SIGUS, que também está em fase de desenvolvimento. A futura integração do protótipo na plataforma SIGUS facilitará bastante a utilização do sistema utilizando diferentes algoritmos de processamento de imagens e aprendizagem automática. Hoje já é possível se escolher entre dezenas de algoritmos clássicos de aprendizagem supervisionada disponíveis por meio do software WEKA, integrado à plataforma SIGUS.

O protótipo ainda precisa ser intensivamente testado e aprimorado, antes de poder ser utilizado por pessoas com necessidades especiais. Melhorias na interface como, por exemplo, a inclusão de um ambiente 3D em que os movimentos da cadeira de rodas possam ser simulados de maneira mais realística, foram planejadas e deverão ser realizadas em breve. As imagens mostradas nesse texto foram obtidas com a *webcam* posicionada acima do monitor capturando quadros frontais da face do usuário. Outra melhoria seria posicionar a *webcam* em uma posição semelhante a qual esta ficaria em uma cadeira de rodas (no braço da cadeira capturando quadros da face do usuário).

## Agradecimentos

Este trabalho recebeu apoio financeiro da Universidade Católica Dom Bosco, UCDB, da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, FUNDECT, e da Agência Financiadora de Estudos e Projetos, FINEP.

## Referências

- Albiol, A., Torres, L., Bouman, C. A., and Delp, E. J. (2000). A simple and efficient face detection algorithm for video database applications. *Proceedings of the IEEE International Conference on Image Processing*, 2:239–242.
- Bergasa, L., Mazo, M., Gardel, A., Barea, R., and Boquete, L. (2000). Commands generation by face movements applied to the guidance of a wheelchair for handicapped people. *15th International Conference on Pattern Recognition (ICPR'00)*, 4.
- Carlson, W., Swan II, J. E., Stredney, D., and Blostein, B. (1994). The application of virtual wheelchair simulation to the determination of environmental accessibility and standards compliance in architectural design. In *Proceedings of the Symposium on Computers and Innovative Architectural Design*, pages 49–54, Baden-Baden, Germany.
- Fehr, L., Langbein, W. E., and Skaar, S. B. (2000). Adequacy of power wheelchair control interfaces for persons with severe disabilities: A clinical survey. *Journal of Rehabilitation Research and Development*, 37(3):353–360.
- Gomez, G., Sanchez, M., and Sucar, L. E. (2002). On selecting an appropriate colour space for skin detection. *MICAI 2002: Advances in Artificial Intelligence, Second Mexican International Conference on Artificial Intelligence. Lecture Notes in Artificial Intelligence*, pages 69–78.
- Hsu, R. L., Mottaleb, M. A., and Jain., A. K. (2002). Face detection in color images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(1):696–706.
- Katevas, N. L., Sgouros, N. M., Tzafestas, S. G., Papakonstantinou, G., Beattie, P., Bishop, J. M., Tsanakas, P., and Koutsouris, D. (1997). The autonomous mobile robot scenario: A sensor-aided intelligent navigation system for powered wheelchairs. *IEEE Robotics and Automation*, 4:60–70.
- Kuno, Y., Murashima, T., Shimada, N., and Shirai., Y. (2000). Intelligent wheelchair remotely controlled by interactive gestures. *Proceedings of the 15th International Conference on Pattern Recognition, ICPR 2000*, 4:672–675.

- Matsumoto, Y., Ino, T., and Ogasawara, T. (2001). Development of intelligent wheelchair system with face and gaze based interface. *Proceedings of 10th IEEE Int. Workshop on Robot and Human Communication (ROMAN 2001)*, pages pp.262–267.
- Neveryd, H., Efring, H., and Bolmsjo, G. (1999). The swedish experience of rehabilitation robotics. In *Proceedings of Workshop Rehabilitation Robotics*, pages 49–54, Hoensbroek.
- Passafaro, E. L. (2002). *Guia de Acessibilidade em Edificações*. CPA - Comissão Permanente de Acessibilidade da SEHAB, 2 edition.
- Pistori, H. (2003). *Tecnologia adaptativa em engenharia de computação: estado da arte e aplicações*. PhD thesis, Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo, SP.
- Pistori, H. and Neto, J. J. (2004). An experiment on handshape sign recognition using adaptive technology: Preliminary results. *Lecture Notes in Artificial Intelligence. XVII Brazilian Symposium on Artificial Intelligence - SBIA'04*, 3171.
- Pistori, H. and Pereira, M. C. (2004). Integração dos ambientes livres weka e imagej na construção de interfaces guiadas por sinais visuais. In *Anais do V Workshop de Software Livre - WSL*, Porto Alegre, RS.
- Schlling, K., Roth, H., Lieb, R., and Stutzle., H. (1998). Sensors to improve the safety for wheelchair users. *Improving the Quality of Life for the European Citizen*, 4:331–335.
- Simpson, R. C. and Levine, S. P. (1997). Adaptive shared control of a smart wheelchair operated by voice control. *International Conference on Intelligent Robots and Systems*, 2:622–626.
- Terrillon, J.-C., Fukamachi, H., Akamatsu, S., and Shirazi, M. N. (2000). Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 54–63.
- Vezhnevets, V., Sazonov, V., and Andreeva, A. (2003). A survey on pixel-based skin color detection techniques. In *Proceedings of Graphicon-2003*, pages 85–92, Moscou, Rússia.
- Yanco, H. A. and Gips, J. (1997). Preliminary investigation of a semi-autonomous robotic wheelchair directed through electrodes. In *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America Annual Conference (RESNA '97)*, pages 414–416. RESNA Press.