

# PROPOSTA DE UMA INTERFACE NÃO CONVENCIONAL PARA REABILITAÇÃO MOTORA SUPOSTADA POR AMBIENTES DE REALIDADE VIRTUAL

Eduardo Souza Santos, Marlus Dias Silva, Webert Vieira Arantes,  
Lázaro Vinícius de Oliveira Lima, Alexandre Cardoso, Edgard Lamounier

Universidade Federal de Uberlândia  
Uberlândia, MG, Brasil

{eduardo, marludias}@mestrado.ufu.br, webert3d@yahoo.com.br,  
lazavini@gmail.com, {alexandre, lamounier}@ufu.br

## Abstract

*This work aims to present a system that allows the creation of non conventional interfaces for navigation on Virtual Reality environments. In this system, images captured by a webcam are processed and converted into patterns that, after being presented to an artificial neural network, become references for computer commands. During the training phase, each pattern is related to an action that will allow the user to act within a VR environment. Since this system has been designed to help handicapped people, during their rehabilitation process, the interaction with the computer is performed without mouse and keyboard. In order to validate the proposed system, two test cases have been developed. The first one asks the user to move an object through a labyrinth. In the second one, by menu navigation, the user is able to construct 3D objects using solid primitives such as spheres, cubes, cylinders and cones.*

## 1 Introdução

Segundo o Censo 2000, aproximadamente 24.6 milhões de pessoas, cerca de 14,4% da população brasileira, apresentam algum tipo de deficiência. [5].

Pode-se classificar as pessoas que perderam algum tipo de função motora como paraplégicas, que perderam os movimentos dos músculos inferiores, tetraplégicas, que não movimentam os músculos superiores e inferiores, e hemiplégicas, que não possuem os movimentos de um lado do corpo. No Brasil, cerca de 937 mil pessoas possuem alguma das deficiências motoras citadas e, embora o último censo não tenha mensurado a quantidade exata, estima-se que cerca de 200 mil dessas pessoas sejam portadores de

tetraplegia [7, 10].

O desenvolvimento de dispositivos e ferramentas que minimizem as limitações desses portadores de deficiência motora e/ou ajudem em sua reabilitação, pode se mostrar importante na melhora da sua qualidade de vida e contribuir para o seu lazer, comunicação e expressão. Para tal, pode-se utilizar ambientes de Realidade Virtual (RV), que segundo [6] é uma “interface avançada do usuário”, propiciando a visualização, movimentação e interação do usuário, em tempo real, em ambientes tridimensionais gerados por computador.

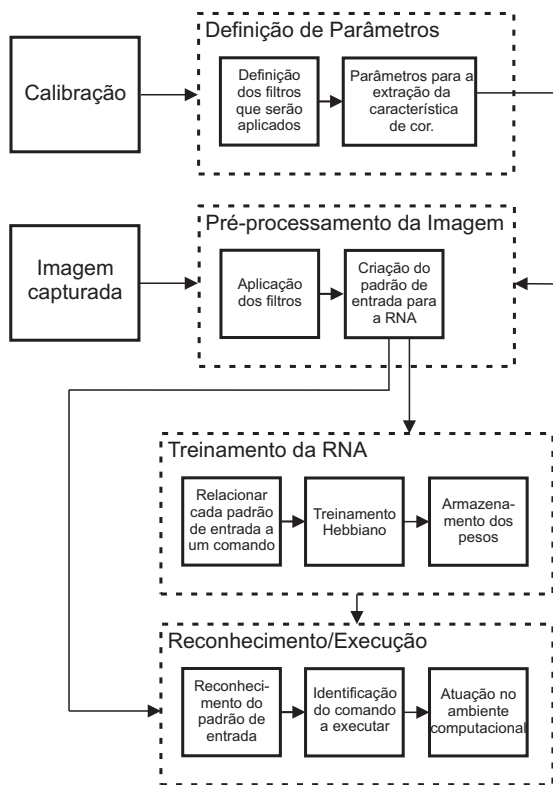
Desta forma, este trabalho propõe desenvolver uma interface homem-máquina não convencional, que possibilitará a interação do usuário com o computador a partir da captação de imagens de uma webcam e com o auxílio de uma Rede Neural Artificial. Pretende-se que se possa utilizar ambientes de Realidade Virtual, de forma simples e intuitiva, sem o uso do mouse e teclado. A execução dos comandos no programa será feita através do reconhecimento, por parte de uma Rede Neural Artificial, dos gestos realizados por uma pessoa ou objetos posicionados à frente da webcam.

Outros dispositivos, os quais utilizam técnicas similares foram pesquisados. A maioria tendo como objetivo o posicionamento do cursor do mouse no monitor através da movimentação da cabeça ou expressões faciais. O *HeadMouse*, da *Origin Instruments Corp.*, e o *SmartNav*, da *Natural-Point Inc.* utilizam sensores de infravermelho e marcadores colocados na cabeça do usuário para reconhecimento dos movimentos. O *HeadMouse 2*, do Grupo de Robótica da *University of Lleida*, aplica algoritmos de visão computacional para reconhecimento de faces e expressões.

Enquanto o sistema proposto utiliza apenas uma webcam, as duas primeiras aplicações, citadas acima, exigem equipamentos adicionais e a terceira, limita-se ao reconhe-

cimento de faces e expressões. O presente trabalho permite a configuração do sistema para uso de qualquer objeto presente na cena, captada pela *webcam*, para reconhecimento dos movimentos e gestos [9, 1].

A Figura 1 mostra o diagrama funcional do software desenvolvido. Pode-se observar que este foi dividido em três etapas. Na primeira, são definidos os filtros que serão aplicados na imagem capturada, facilitando a extração de características. Na segunda, padrões de treinamento da RNA são criados a partir das imagens tratadas, comandos são relacionados a estes padrões e a Rede Neural Artificial é treinada. Na terceira etapa, são aplicados os filtros às imagens capturadas, gerando os padrões de entrada que serão submetidos à RNA, para o reconhecimento. Com o padrão reconhecido executa-se o comando relacionado ao mesmo.



**Figura 1. Diagrama Funcional.**

É importante ressaltar que apesar do sistema desenvolvido ter sido pensado primeiramente para a criação de interfaces com ambientes virtuais que seriam utilizados na reabilitação de portadores de deficiências motoras, ele também pode ser empregado em aplicações educacionais, jogos e softwares. [6].

Este trabalho é descrito da seguinte forma: a Seção 2 aborda o desenvolvimento das etapas do projeto, desde a captura da imagem até o controle do ambiente. A Seção 3 apresenta os resultados obtidos e a Seção 4 discorre acerca

das conclusões e trabalhos futuros.

## 2 Desenvolvimento do Sistema

De forma resumida o que se pretende com o sistema é capturar uma imagem colorida, através de uma *webcam*, converte-la em preto e branco (PB), onde os *pixels* pretos representarão a região de interesse, transformar esta imagem PB em um vetor bipolar (composto por uma cadeia de -1 e 1) e apresentar este vetor à uma Rede Neural Artificial como um padrão a ser treinado e reconhecido. Posteriormente, através do reconhecimento deste padrão o comando associado a ele será executado.

### 2.1 Captura e Processamento das Imagens

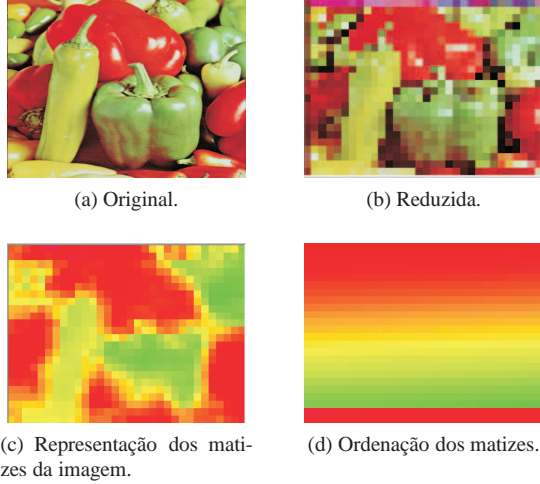
A *webcam* no sistema pode ser configurada para a captação de imagens com tamanho de 320x240 ou 640x480 *pixels*. A manipulação de imagens destes tamanhos exige um esforço computacional muito grande para um sistema em tempo real. Logo, primeiramente é aplicado um filtro de redução a esta imagem, onde o usuário define o tamanho final desta. Nos estudos realizados, optou-se pelo tamanho final de 32x24 *pixels*, que foi a configuração que obteve a melhor relação tempo de processamento/qualidade.

Em seguida seleciona-se as características do(s) objeto(s) que ativar(ão) a execução dos comandos da aplicação. Nos testes realizados separou-se a região que contém a cabeça do usuário, por atender um grande número de portadores de limitações motoras.

Para a segmentação da imagem optou-se pela separação por cores e o sistema oferece duas opções. Na primeira é utilizado o padrão RGB, em que a cor do *pixel* é definida pela intensidade de seus componentes: vermelho (*Red*), verde (*Green*) e azul (*Blue*). Para esta opção é calculado o grau de luminosidade da cor ( $0.3 \cdot R + 0.59 \cdot G + 0.11 \cdot B$ ) de cada *pixel* da imagem e estes são transformados em preto se sua luminosidade for menor que um limiar definido pelo usuário na interface de calibração ou, em branco, se for maior ou igual. Esta opção se mostrou muito limitada para a segmentação da imagem, pois a luminosidade, neste caso, não é suficiente para separar as cores com precisão [8, 3].

A segunda opção foi o padrão de cor HSV, no qual a cor do *pixel* é definida pelos parâmetros matiz (*Hue*), saturação (*Saturation*) e luminância (*Value*) [8]. Nesta opção o usuário define os limites superiores e inferiores de duas faixas de cores, baseadas nos valores dos matizes da imagem, e os *pixels* cujo valor de seu matiz pertencerem a estas faixas serão transformados em preto, caso contrário em branco.

Para facilitar a definição das duas faixas de cores, foi criado um sistema de representação dos matizes da imagem. Como as componentes H, S e V são independentes entre si,



**Figura 2. Representação dos matizes de uma imagem.**

atribuiu-se ao valor da intensidade de S e V dos *pixels* o seu maior patamar, gerando-se a Figura 2c. A Figura 2d é gerada a partir da ordenação dos *pixels* da Figura 2c através dos valores dos matizes. O usuário através de cliques do mouse sobre esta imagem ordenada, seleciona as faixas de cores que serão utilizadas como limiares para a segmentação.

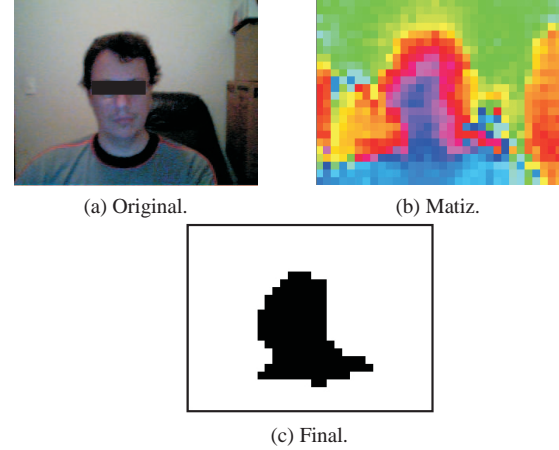
Para facilitar a separação dos objetos de interesse na imagem, outros três filtros podem ser usados pelo usuário: filtro de suavização de cor (*smoothing*) usado logo após o filtro de redução da imagem, o filtro de suavização do matiz aplicado depois do filtro de suavização de cor e o filtro de redução de ruídos aplicado na imagem em preto e branco já segmentada. Ao final desta etapa, através da utilização dos filtros e parâmetros, é gerada uma imagem final com as regiões dos objetos de interesse ficando em evidência, como mostrado na Figura 3.

## 2.2 Rede Neural Artificial (RNA)

A capacidade de generalização, a tolerância a dados ruidosos e a velocidade foram os principais fatores da escolha do uso de uma Rede Neural Artificial como o método de reconhecimento de padrões no sistema proposto [2, 4].

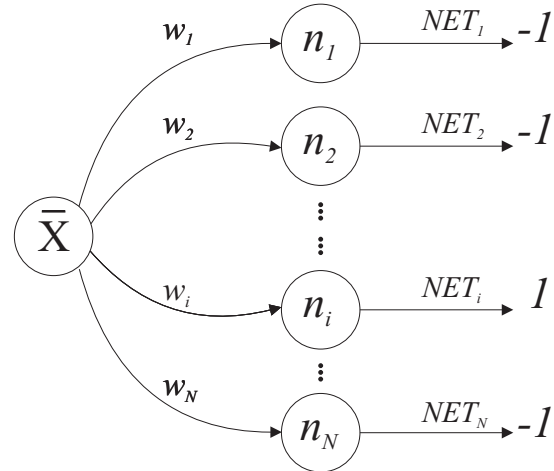
No presente trabalho, devido a simplicidade de implementação e o seu bom desempenho, foi utilizada uma rede Hebb de uma camada, com  $N$  neurônios, Figura 4, onde  $N$  consiste no número de padrões apresentados no treinamento da rede, ou seja, o número de imagens, selecionadas pelo usuário, que definirão os comandos a serem executados.

A utilização de uma rede com treinamento pelo método de Hebb, também se justifica à medida que a calibração do sistema deve ser realizada para cada usuário e ambiente.



**Figura 3. Etapa de definição dos filtros e parâmetros.**

Portanto, um treinamento rápido se mostrou de suma importância, visto que, a demora na inicialização do sistema poderia inviabilizar a utilização do mesmo.

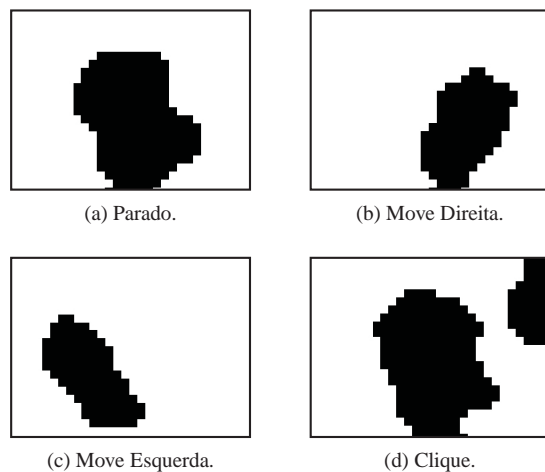


**Figura 4. Arquitetura da RNA.**

## 2.3 Treinamento da Rede Neural Artificial

Nesta etapa, a imagem capturada é filtrada gerando uma imagem final, em preto e branco, que é relacionada a um dos comandos a serem executados. Em seguida, cria-se um vetor a partir desta imagem, em que os *pixels* em branco e preto são representados pelos valores -1 e 1, respectivamente. Ao término da seleção das imagens, os vetores são apresentados à Rede Neural Artificial como padrões de treinamento.

A Figura 5 mostra algumas imagens PB e seus respectivos comandos. Tais imagens foram obtidas a partir de mo-



**Figura 5. Imagens PB treinadas.**

vimentos da cabeça do usuário, sendo que, na Figura 5d também foi utilizada a mão.

## 2.4 Reconhecimento dos padrões e execução dos comandos

Com a Rede Neural Artificial treinada, a cada imagem capturada são realizados os seguintes procedimentos:

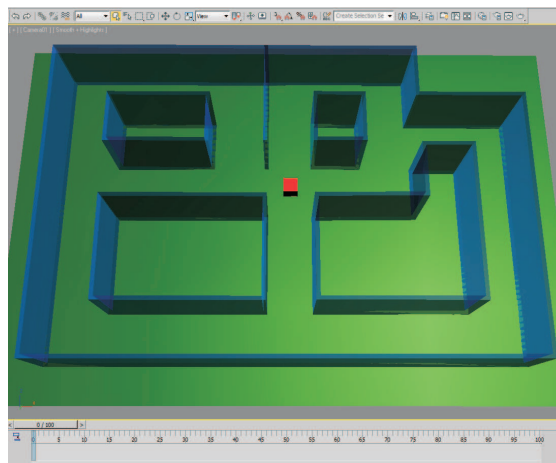
- As características são extraídas da imagem capturada, convertidas para um vetor bipolar que é apresentado a uma RNA. (Ex. uma imagem de 32x24 pixels gera um vetor bipolar com dimensão de 768 posições).
- Calcula-se o valor do somatório dos pesos ( $w$ ) multiplicado pelo padrão de entrada ( $\bar{x}$ ), gerando o valor de NET.
- Por fim, determina-se o comando que será executado de acordo com o neurônio que obteve o maior valor de NET calculado.

## 2.5 Ambientes de Realidade Virtual

Foram desenvolvidos para as experimentações, dois ambientes de Realidade Virtual. A linguagem utilizada foi o MAXScript do ambiente de produção Autodesk 3ds Max®.

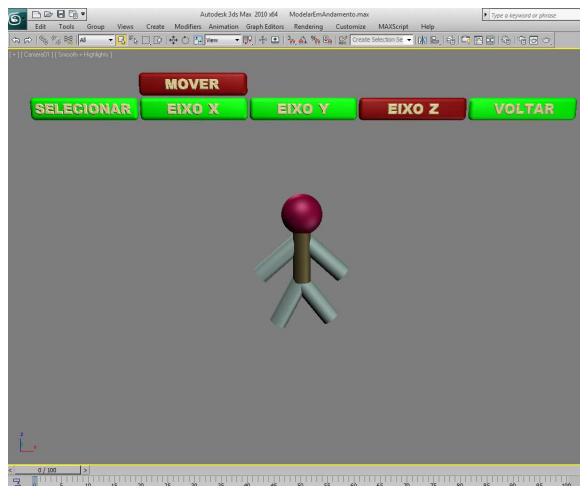
Como mostra a Figura 6, o primeiro protótipo possibilita, em tempo real, a movimentação de um cubo em um labirinto virtual através do reconhecimento dos gestos do usuário. A cada utilização do ambiente, um conjunto de novos gestos pode ser cadastrado, aumentando assim o nível de imersão e interatividade do usuário.

O segundo protótipo desenvolvido é mostrado na Figura 7, e teve como objetivo possibilitar a modelagem de objetos 3D, usando para navegação nos menus somente gestos



**Figura 6. Imagem do Primeiro Protótipo: "Labirinto Virtual".**

do usuário. Nesta aplicação é possível criar formas básicas como esferas, cubos, cilindros, cones, e aplicar a estes transformações de translação, rotação e escala, nos três eixos. Objetos mais complexos podem ser modelados através da união destas formas básicas.



**Figura 7. Imagem do Segundo Protótipo: "Modelagem Virtual".**

Ambientes como estes dois protótipos sugerem a criação de novas possibilidades de auxílio à reabilitação de pacientes com algum tipo de lesão que necessite de exercícios motores.

### 3 Resultados e Limitações

Como mencionado nas seções anteriores, antes do início da utilização dos protótipos é necessário realizar a calibração do sistema segundo o espectro de cor dos objetos de interesse. Esta fase se mostrou trabalhosa em alguns locais, dependendo da quantidade de elementos de cores fortes, principalmente as cores que se aproximam do vermelho. Para um ambiente controlado, onde as cores são mais uniformes, as configurações foram feitas com rapidez.

Em cinco lugares diferentes (duas salas de aula, dois laboratórios e um escritório pessoal) onde os testes foram realizados três deles não apresentaram problemas. Dois ambientes tiveram de ser modificados para se conseguir algum resultado.

A partir da avaliação realizada por pessoas ligadas a área de Engenharia Biomédica, considerou-se os resultados obtidos pelo sistema como sendo satisfatório, ou seja, o seu uso pode sim trazer benefícios a usuários com deficiência motora em fase de reabilitação.

### 4 Conclusões e Trabalhos Futuros

O desenvolvimento deste trabalho mostrou resultados satisfatórios a respeito da utilização de Redes Neurais Artificiais para o reconhecimento de gestos na atuação em ambientes de Realidade Virtual e possível utilização do mesmo como auxiliar em programas de reabilitação motora de pacientes.

O projeto deverá ser desenvolvido, também, em uma linguagem de programação mais robusta como C++ ou C#, para que possa ser realizada uma comparação de desempenho, visto que a linguagem MaxScript do 3ds Max<sup>®</sup> contém limitações ao que se refere a captura de imagens de uma *webcam* e velocidade de execução.

A melhoria nos algoritmos de processamento e segmentação das imagens consiste em uma necessidade inerente à evolução do sistema, logo deseja-se aplicar métodos mais eficientes e precisos nesta parte do sistema.

A criação de novos ambientes de Realidade Virtual possibilitará um comparativo para se definir quais terão maior eficiência no processo de auxílio na reabilitação motora de pacientes e pesquisa de novas aplicabilidades para o sistema proposto.

### Referências

- [1] O. I. Corporation. Headmouse extreme, 2009. Disponível em: <<http://www.orin.com/access/headmouse/>>. Acesso em: 15 de junho de 2009.
- [2] L. Fausett. *Fundamentals of Neural Networks*. 1994.
- [3] R. C. Gonzales and R. E. Woods. *Digital Image Processing*. Pearson Education, 2002.
- [4] S. Haykin. *Redes Neurais Princípios e prática*. Bookman, 2001.
- [5] IBGE. Censo 2000, 2000.
- [6] C. Kirner and R. Siscouto. *Realidade Virtual e Aumentada: Conceitos, Projeto e Aplicações*. Sociedade Brasileira de Computação, 2007.
- [7] MOTRIX. O que é tetraplegia?, 2009. Disponível em: <<http://intervox.nce.ufrj.br/motrix/>>. Acesso em: 15 de junho de 2009.
- [8] H. Pedrini and W. R. Schawartz. *Análise de Imagens Digitais: princípios, algoritmos e aplicações*. thomson, 2008.
- [9] N. Point<sup>®</sup>. What is smartnav?, 2009. Disponível em: <<http://www.naturalpoint.com/smartnav/>>. Acesso em: 15 de junho de 2009.
- [10] L. C. Rodrigues. A inclusão do portador de deficiência no contexto das organizações. Master's thesis, Universidade Federal Rural do Rio de Janeiro Instituto de Ciências Humanas e Sociais, 2007.