

# VisiUMouse: An Ubiquitous Computer Vision Technology for People with Motor Disabilities

Krishna Ferreira Xavier  
Instituto Federal Rio-Grandense -  
WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
tsixav@gmail.com

Vinicius Kruger da Costa  
Universidade Federal de Pelotas  
(UFPel), Instituto Federal  
Rio-Grandense - WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
viniciusdacosta@pelotas.ifsul.edu.br

Rafael Cunha Cardoso  
Universidade Federal de Pelotas  
(UFPel), Instituto Federal  
Rio-Grandense - WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
rafaelcardoso@pelotas.ifsul.edu.br

Jamir Alves Peroba  
Instituto Federal Rio-Grandense -  
WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
perobajamir@gmail.com

Adriano Oliveira Lima Ferreira  
Instituto Federal Rio-Grandense -  
WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
adriano.jaguara@gmail.com

Marcelo Bender Machado  
Universidade Federal de Pelotas  
(UFPel), Instituto Federal  
Rio-Grandense - WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
marcelo@ifsul.edu.br

Tatiana Aires Tavares  
Universidade Federal de Pelotas  
(UFPel) - WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
tatiana@inf.ufpel.edu.br

Andréia Sias Rodrigues  
Universidade Federal de Pelotas  
(UFPel), Instituto Federal  
Rio-Grandense - WeTech (IFSul)  
P.O. Box 15.064-91.501-970  
Pelotas - RS, Brasil  
andreia.sias@inf.ufpel.edu.br

## ABSTRACT

VisiUMouse is a solution that enables accessibility to computer's use by people with upper limbs motor disability. This Assistive Technology uses the concepts of Computational Vision to recognition and face tracking, using the eye as reference point. Through video input, that tracks the eye movement, we control the mouse cursor by webcam, providing an amplification in the functional abilities of these users and, consequently, promoting inclusion. The VisiUMouse was evaluated following the Fitts's law protocol, which involves common pointing, select and click tasks, used as metrics to verify interaction with the computer.

## KEYWORDS

Computer Vision, Assistive Technology, VisiUMouse, Accessibility, Ubiquitous Technology, OpenCV, Face Tracking

## 1 INTRODUÇÃO

O último levantamento feito pelo IBGE (Instituto Brasileiro de Geografia e Estatísticas) indica que existem 45,6 milhões de pessoas com deficiência [8], representando 23,9% da população brasileira. Deste total, 7% apresentam deficiência motora que engloba desde

restrições suaves até outras, mais severas. Diante desse número expressivo existem grandes desafios a serem enfrentados para prover maior autonomia e independência na realização de tarefas cotidianas destas pessoas, como por exemplo, prover acessibilidade ao uso do computador.

O desenvolvimento de novos sistemas, *softwares* e *hardwares* concebem uma evolução nas estruturas sociais, corporativas e acadêmicas. Porém existem áreas ainda com muitas demandas, sendo uma delas a área de Tecnologia Assistiva (TA).

A área da TA pode ser definida como o conjunto de sistemas, tecnologias e inovações que visam aumentar, ou até mesmo permitir, a ampliação das habilidades funcionais de uma pessoa, com algum tipo de deficiência, possibilitando sua inclusão social e familiar [4]. O desafio de formar as condições necessárias à acessibilidade das tecnologias por pessoas com deficiência é uma garantia estabelecida pelo Decreto Brasileiro 7.612 [7], o qual promove o acesso, o desenvolvimento e a inovação em TA.

É evidente a evolução das tecnologias que possibilitam cada vez mais a criação e a melhoria de soluções em vários âmbitos sociais, digitais, acadêmicos e profissionais. Este avanço permite que cada indivíduo tenha acesso a um arsenal de conhecimento, através da internet. O crescimento das possibilidades em todas as áreas de conhecimento representa um novo modelo mundial, mas para aquelas pessoas que possuem plenas condições de utilizar um computador e a internet.

Porém um indivíduo que tenha a mobilidade prejudicada, principalmente nos membros superiores, responsáveis pelos modos de

interação mais comumente utilizados com o computador, como *mouse* ou teclado, não terá a mesma possibilidade. Considerando que os movimentos da cabeça e olhos são os últimos movimentos a serem perdidos, grande parte destes indivíduos pode utilizar estas habilidades para realizar a interação com o computador [12].

Nesse contexto, o objetivo desse artigo é apresentar o VisiU-Mouse, uma TA capaz de permitir o uso do computador por pessoas com algum tipo de deficiência motora, por meio do rastreamento do movimento da cabeça, usando os olhos como ponto de referência. Isto é feito através da entrada de vídeo, via *webcam*, utilizando os conceitos e premissas da Visão Computacional (VC), uma subárea da Inteligência Artificial (IA) que tem como objetivo dar significado às imagens digitais com base em análise para extrair algum conhecimento [14].

O restante do trabalho está organizado da seguinte forma: a seção 2 apresenta um contexto de TA relacionado aos conceitos de VC utilizados neste trabalho. Na seção 3 são descritos os trabalhos relacionados. A seção 4 detalha o protótipo desenvolvido, o VisiU-Mouse; a seção 5 descreve o experimento com o protótipo e a seção 6 apresenta os resultados e discussão, a seção 7 finaliza o trabalho com as conclusões e trabalhos futuros.

## 2 TECNOLOGIA ASSISTIVA E A VISÃO COMPUTACIONAL

Existem TA que permitem acessibilidade para usuários atípicos<sup>1</sup> em tarefas simples como navegar na internet, estudar, interagir em redes sociais, etc. Contudo ainda constitui-se um desafio da área de Interação Homem Computador (IHC) desenvolver soluções mais naturais que não envolvam o uso de tecnologias vestíveis, por exemplo. Nesse sentido, a computação ubíqua proporciona um campo profícuo de pesquisa através de suas aplicações para criação de novas possibilidades [13] de TA.

São ainda de difícil acesso os dispositivos de TA que dão suporte a usuário com apenas o controle do movimento da cabeça. Em muitos destes casos a pessoa com deficiência tem capacidade cognitiva plena, permitindo que a pessoa interaja com o seu ambiente através da movimentação da cabeça e olhos Rodrigues et al. [16].

Considerando as entradas de dados por vídeo, voz ou acionamento mecânico é possível supor que a primeira opção possibilita uma interação mais natural, considerando que o movimento da face e dos olhos pode indicar a direção que o usuário deseja ir. Através da VC é possível dar significado às imagens digitais capturadas desses movimentos, com base em sua análise [14], possibilitando assim sua interpretação.

Um vídeo pode ser considerado um conjunto de imagens digitais, que é uma representação numérica de uma imagem bidimensional, tornando possível o armazenamento, transferência, processamento ou reprodução por meios eletrônicos. A representação e extração dos dados de uma imagem ou vídeo são tarefa da área de VC, que permite extrair informações das imagens através do processamento [14]. Uma das técnicas para detectar objetos em uma imagem digital é o algoritmo de Viola-Jones, implementado pela biblioteca OpenCV, ambos descritos nas subseções seguintes.

### 2.1 Algoritmo de Viola-Jones

O algoritmo de Viola-Jones foi proposto em [19], para o reconhecimento de objetos em imagens digitais, e tem como característica o alto desempenho de processamento e o alto índice de assertividade, podendo chegar até a 99,7% dependendo do número de etapas treinadas, acurácia do classificador e objeto para qual foi treinado. O Viola-Jones trabalha com algoritmo em cascata que se utiliza de arquivos de padrões e características chamados de *haar features*, organizadas em uma estrutura de árvore em XML (*eXtensible Markup Language*).

Os objetos procurados pela estrutura de detecção envolvem as somas de *pixels* de imagem dentro de áreas retangulares. Estes valores são comparados com os valores dos estágios do arquivo XML, para determinar se a área examinada da imagem possui ou não o objeto que pretende detectar, a Figura 1 apresenta esse processo.



Figure 1: Processo de somas de pixels para detecção de objetos do algoritmo de Viola-Jones.

### 2.2 OpenCV

A OpenCV (*Open Source Computer Vision Library*) é uma biblioteca multiplataforma desenvolvida pela Intel Corporation [10], que implementa diversos módulos para VC, os quais são utilizados para o Processamento de Imagens e Vídeo, Estruturas de Dados, Álgebra Linear e mais de 350 algoritmos de VC. OpenCV possui, em seu módulo central a implementação do algoritmo de Viola-Jones para detecção de objetos. A qual esta presente em diversos projetos, a seção de trabalhos relacionados apresenta alguns deles.

## 3 TRABALHOS RELACIONADOS

Existem diversos projetos, produtos e estudos em TA, baseados na VC, que utilizam-se de *softwares* para detecção de objetos através de uma *webcam* [5, 6, 11, 15]. A detecção dos movimentos nesses projetos é feita através de uma entrada de vídeo que rastreia várias partes do corpo humano, sendo que a mais comum é o rastreamento de movimentos da cabeça (*video-based*) [1].

Um dos objetivos de análise desses trabalhos é não depender de dispositivos de alto poder computacional, utilizando tecnologias disponíveis no próprio computador. Das soluções *video-based* a mais adotada é o *software* CameraMouse [6] que frequentemente é usado como base de pesquisas e de comparações, como em Kurauchi [9].

O CameraMouse pode ser configurado para capturar o movimento de qualquer parte móvel do corpo para manipular o ponteiro do computador, através dos recursos de vídeo, pela técnica de espelhamento, ou seja, o movimento da cabeça é refletido no movimento do ponteiro. A aplicação foi desenvolvida para auxiliar pessoas com deficiência motora, e o seu público-alvo são as pessoas que não têm

<sup>1</sup>Usuário atípico: usuários com alguma deficiência física.

controle confiável das mãos, mas que podem executar movimentos com a cabeça [6], por exemplo.

O projeto HeadDev [2] é semelhante ao CameraMouse, porém faz o rastreamento apenas dos movimentos da cabeça, não permitindo configurar o objeto que o *software* deve rastrear. Seu funcionamento também utiliza a técnica de espelhamento.

Já os *softwares* Facial Human-Computer Interface [3], Facial Mouse [5] e BlinkMouse [11] tem como objetivo aplicar o modo de interação feita pelo CameraMouse, refinando a funcionalidade do clique, principalmente nos requisitos de precisão e configuração.

Baseado nos trabalhos relacionados percebe-se a utilização dos conceitos de VC e o uso da OpenCV, com a técnica de espelhamento do movimento e rastreamento da face principalmente. Nesse contexto o VisiUMouse tem como objetivo implementar uma técnica de movimento alternativa, batizada de Técnica de Zona Neutra e de Movimento (TZNM) com o rastreamento dos olhos para o controle do ponteiro, facilitando uma implementação futura de clique pelo piscar dos olhos e *Eye Tracking*.

#### 4 PROTÓTIPO DO VISIUMOUSE

VisiUMouse é um projeto de TA, baseado em VC, com o objetivo de ajudar pessoas com algum tipo de deficiência motora, principalmente nos membros superiores, a utilizar o computador sem necessidade de colocar sensores ou algum tipo de *hardware* no corpo do usuário, permitindo assim a substituição do *mouse* comum por uma nova interface humano computador.

O funcionamento do VisiUMouse é composto pela captura de vídeo de uma *webcam*, do processamento desta imagem com o uso de algoritmos de rastreamento de objetos [19], os quais tornam possível reconhecer a posição atual dos olhos do usuário. Para evitar que outro par de olhos, de um segundo usuário, seja captado e este passe a controlar o cursor, foram desenvolvidas condições para que o VisiUMouse determine se o conjunto de olhos rastreados pertence ao mesmo indivíduo.

Genericamente o processamento da imagem contém 3 estágios: Captura, Análise e Compressão da Imagem. A **Captura** trata da aquisição dos dados de entrada de vídeo; a **Análise** trata do reconhecimento do objeto alvo. Para o reconhecimento é necessário um conjunto de padrões do alvo, esse processo é nomeado de Treinamento e é responsável por gerar um arquivo XML que contém características e padrões dos objetos alvos. A **Classificação** é o processo final que define se um determinado pedaço da imagem contém o objeto que pretende detectar. A **Compressão da Imagem** é o processo feito para reduzir a redundância dos dados, de forma a armazenar ou transmitir esses mesmos dados de forma eficiente. [17].

A primeira versão do VisiUMouse tem seu funcionamento baseado no movimento da cabeça/face, usando os olhos como ponto de referência e a TZNM. A medida do deslocamento dos olhos é feita através do processamento da posição anterior, ou seja, se a medida for superior a configurada é feita a movimentação do ponteiro (zona de movimento da TZNM), se for menor o ponteiro permanece parado (zona neutra da TZNM). Para processar para qual lado o cursor deve ir é verificada a posição dos olhos em relação às linhas (ilusórias) vermelha e verde, como pode ser visto na Figura 2.



Figure 2: Layout do VisiUMouse: telas de rastreamento e processamento do deslocamento dos olhos.

Para determinar a direção do movimento do ponteiro é feita uma comparação com a posição dos olhos em relação às linhas verde e vermelha. Por exemplo, se o olho direito estiver próximo da linha verde e o olho esquerdo próximo da linha vermelha, indica que o usuário está com a cabeça inclinada para esquerda e que o cursor do mouse deve ir para o mesmo lado, como mostrado na Figura 2 (1: para esquerda). Também é possível fazer movimentos para direita, para cima e para baixo (vide Figura 2), além da combinação desses movimentos, tornando possível realizar movimentações em diagonal. No momento em que o movimento do cursor do *mouse* se inicia, o usuário não precisa mais deslocar os olhos para manter o movimento, ele apenas precisa permanecer na mesma posição, podendo tornar o controle do *mouse* menos cansativo, e assim evitando o desgaste físico do usuário <sup>2</sup>.

O protótipo do VisiUMouse utiliza o algoritmo de Viola-Jones e a OpenCV, onde também foram desenvolvidos um sistema calibrador para reiniciar os pontos de referência dos olhos, um sistema capaz de simular as principais funcionalidades do mouse entre outros sistemas.

#### 5 EXPERIMENTO

O experimento foi executado por 9 participantes voluntários e teve como objetivo fazer uma avaliação comparativa do *software* VisiUMouse com o *mouse* convencional para avaliar o seu funcionamento e objetivo. Para tanto foi utilizado um protocolo baseado na lei de Fitts, proposto por Soukoreff and MacKenzie [18], envolvendo tarefas comuns como apontar, selecionar e clicar, que são usadas como métricas para verificar a interação com o computador. O protocolo foi configurado para 6 etapas, com 5 alvos cada, podendo ter o tamanho de 120, 150 ou 180 *pixels*. O tamanho (gerados aleatoriamente) de cada um dos 5 alvos são iguais para cada fase <sup>3</sup>.

O teste foi dividido em 3 etapas: a primeira consistiu na aprendizagem de uso da aplicação, onde o participante era familiarizado com o funcionamento do VisiUMouse durante 5 minutos; a segunda etapa consistiu no experimento em si com o VisiUMouse; na última etapa, o participante utilizou o *mouse* convencional para a execução das mesmas tarefas da etapa anterior. Os resultados são computados pelo *software* do protocolo.

Para as três etapas foi utilizado o clique por tempo, denominado *dwell time* [6], configurado para o tempo 5 segundos para efetuar

<sup>2</sup>Foi disponibilizado um vídeo com o funcionamento básico do VisiUMouse, em: <http://visiумouse.com/func/>

<sup>3</sup>Foi disponibilizado gravações dos testes em: <http://visiумouse.com/videos>

o clique. Cabe ressaltar que em paralelo ao desenvolvimento do VisiUMouse foi criado um sistema de clique por tempo para ser usado nos testes<sup>4</sup>. As demais configurações do VisiUMouse como velocidade foram mantidas iguais em todos os testes.

O experimento foi conduzido com um notebook (LG S460) equipado com Windows 10, processador 2,4GHz (i3), 4 GB RAM, monitor de 14 polegadas, resolução de 1366 x 768 pixels e utilizou-se a webcam do próprio computador.

## 6 RESULTADOS E DISCUSSÃO

A Figura 3 apresenta o gráfico dos resultados da assertividade dos cliques nos alvos. O software VisiUMouse (VUM, em azul) teve assertividade média de 99,63%, representando apenas um erro com o participante 8. O mouse comum (Mouse, em laranja) teve assertividade de 100%, que já era esperado devido a grande familiaridade que os usuários já tinham com este dispositivo.

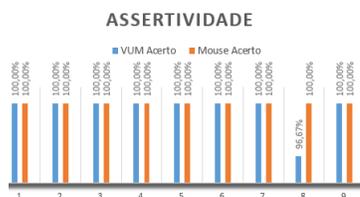


Figure 3: Resultados da assertividade dos cliques nos alvos.

A Figura 4 apresenta o gráfico com os resultados do tempo médio de cada clique (TMC), em milissegundos. O TMC do mouse comum foi de aproximadamente 6372,4 milissegundos. O VisiUMouse (VUM) teve o TMC de aproximadamente 16579,5 milissegundos<sup>5</sup>.

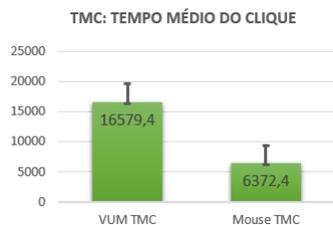


Figure 4: Resultados do teste comparativo.

## 7 CONCLUSÕES E TRABALHOS FUTUROS

O resultado de assertividade de 99,6% do protótipo do VisiUMouse indica seu funcionamento como uma tecnologia que permite o controle do computador apenas com o movimento dos olhos. Além de validar a utilização da TZNM que pode permitir que o usuário com menos mobilidade consiga controlar o ponteiro, tornando possível também um desgaste físico menor, principalmente na região do pescoço. O resultados do TMC apontam uma diferença de tempo que pode ser consequência da inexperiência de uso com o VisiUMouse em comparação com o mouse tradicional, a qual pode ser

amenizada com mais uso e com melhor configuração de velocidade do ponteiro, por exemplo.

O projeto VisiUMouse está na fase inicial de seu desenvolvimento e tem como metas futuras: aperfeiçoar o rastreamento, diminuir o gasto computacional com processamento da imagem, implementar o clique pelo piscar dos olhos, aprimorar a calibração inicial e por fim disponibilização ao público.

## REFERENCES

- [1] AMER Al-Rahayfeh and MIAD Faezipour. 2013. Eye tracking and head movement detection: A state-of-art survey. *IEEE journal of translational engineering in health and medicine* 1 (2013), 2100212–2100212.
- [2] Adelson Rui Alves, Nilson Ribeiro Modro, Alex Luiz de Souza, and Luiz Cláudio Dalmolin. Ferramentas de Tecnologia Assistiva para Auxílio no Processo de Aprendizagem de Pessoas com Paralisia Cerebral: O Software Comunicare. (????).
- [3] Rui Azevedo Antunes, Luís Brito Palma, Fernando V Coito, Hermínio Duarteramos, and Paulo Gil. 2016. Intelligent human-computer interface for improving pointing device usability and performance. In *Control and Automation (ICCA), 2016 12th IEEE International Conference on*. IEEE, 714–719.
- [4] Rita Bersch. 2013. Introdução a Tecnologia Assistiva. (2013). [http://www.assistiva.com.br/Introducao\\_Tecnologia\\_Assistiva.pdf](http://www.assistiva.com.br/Introducao_Tecnologia_Assistiva.pdf)
- [5] Zhen-Peng Bian, Junhui Hou, Lap-Pui Chau, and Nadia Magnenat-Thalmann. 2016. Facial position and expression-based human-computer interface for persons with tetraplegia. *IEEE journal of biomedical and health informatics* 20, 3 (2016), 915–924.
- [6] James Gips, Margrit Betke, and Peter Fleming. 2000. The Camera Mouse: Preliminary investigation of automated visual tracking for computer access. In *In Proc. Conf. on Rehabilitation Engineering and Assistive Technology Society of North America*. 98–100.
- [7] Elaine Maria Bessa Rebelo Guerreiro. 2012. A acessibilidade e a educação: um direito constitucional como base para um direito social da pessoa com deficiência. *Revista Educação Especial* 25, 43 (2012), 217–232.
- [8] IBGE. 2010. Instituto Brasileiro de Geografia e Estatística - Censo Demográfico 2010. <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/>. (11 2010). <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/>
- [9] Andrew Kurauchi, Wenxin Feng, Carlos Morimoto, and Margrit Betke. 2015. HMAGIC: head movement and gaze input cascaded pointing. In *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments*. ACM, 47.
- [10] Robert Laganière. 2017. *OpenCV 3 Computer Vision Application Programming Cookbook*. Packt Publishing Ltd.
- [11] Joanna Marnik. 2014. BlinkMouse-On-Screen Mouse Controlled by Eye Blinks. In *Information Technologies in Biomedicine, Volume 4*. Springer, 237–248.
- [12] Davide Mulfari, Antonio Celesti, Maria Fazio, and Massimo Villari. 2015. Human-computer Interface Based On Iot Embedded Systems For Users With Disabilities. In *Internet of Things. User-Centric IoT*. Springer, 376–383.
- [13] Danielle Teixeira Oliveira, Marla Dore Carvalho, and Tuane Lisboa Silva Paixão. 2016. Ubiquitous Computing: Gestures Interaction Applied the Learning Disabilities in Process Literacy. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*. ACM, 107–110.
- [14] Simon JD Prince. 2012. *Computer vision: models, learning, and inference*. Cambridge University Press.
- [15] Guilherme Matheus Ramos, Victor ON Sales, and Cesar AC Teixeira. 2016. LE-TRAS: Communication with Lines, Triangles and Rectangles. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*. ACM, 215–218.
- [16] Andréia Sias Rodrigues, Vinicius da Costa, Márcio Bender Machado, Angélica Lacerda Rocha, Joana Marini de Oliveira, Marcelo Bender Machado, Rafael Cunha Cardoso, Cleber Quadros, and Tatiana Aires Tavares. 2016. Evaluation of the Use of Eye and Head Movements for Mouse-like Functions by Using IOM Device. In *International Conference on Universal Access in Human-Computer Interaction*. *Human-Computer Interaction International*, 81–91.
- [17] André Lira Rolim and Ed Porto Bezerra. 2008. Um sistema de identificação automática de faces para um ambiente virtual de ensino e aprendizagem. In *Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*. ACM, 129–132.
- [18] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies* 61, 6 (2004), 751–789.
- [19] Paul Viola and Michael Jones. 2001. 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. IEEE, I–I.

<sup>4</sup>Sistema de clique por tempo. Disponível em: <http://visiouse.com/click-time>

<sup>5</sup>Resultados disponíveis em: <http://visiouse.com/WTIC>