



V Encontro de Iniciação Científica e
Tecnológica
V EnICT
ISSN: 2526-6772
IFSP – Câmpus Araraquara
22 e 23 de outubro de 2020



PROPOSTA DE APLICAÇÃO PARA AUXILIAR A COMUNICAÇÃO POR MEIO DO MOVIMENTO DE PÁLPEBRAS

YURI HENRY TAKASHI QUEIROZ KANEGAE¹, LUIZ HENRIQUE NUNES²

¹Aluno do Curso Técnico em Informática Integrado, IFSP Campus Araraquara, y.kanegae@aluno.ifsp.edu.br

²Docente no IFSP Campus Araraquara, lhenriquenunes@ifsp.edu.br

Área de conhecimento (Tabela CNPq): Sistemas de Informação - 1.03.03.04-9

RESUMO: No mundo, mais de um bilhão de pessoas apresentam algum tipo de deficiência intelectual ou motora. Dependendo do grau da deficiência, essas pessoas têm problemas em executar tarefas simples ou se comunicar com os demais. Neste sentido, várias áreas da tecnologia vêm se desenvolvendo com o intuito de melhorar a qualidade de vida dessas pessoas e possibilitar a sua comunicação. O objetivo desse trabalho é descrever um sistema que possibilita a pessoas com limitações motoras e comunicacionais uma forma alternativa de se expressar. Essa forma de comunicação é baseada no reconhecimento e transformação de uma determinada sequência de movimentos nas pálpebras em ações.

PALAVRAS CHAVE: tecnologias assistivas, OpenCV, comunicação alternativa.

INTRODUÇÃO

O contingente de pessoas com algum grau de deficiência é enorme. Muitas vezes, as deficiências dessas pessoas acabam por interferir no seu dia-a-dia, impossibilitando desde a execução de tarefas simples até a comunicação e expressão de sentimentos (HUSSAIN, JAGOE, *et al.*, 2018).

O sistema descrito neste trabalho, tem como intuito fornecer uma ferramenta que possibilite a comunicação dos usuários por meio do movimento das pálpebras. Tal reconhecimento, é feito através de um *software*, o qual ao receber uma sequência de imagens por meio de uma *webcam* utiliza um mecanismo similar ao proposto pelo código *morse* (ITU-R, 2009), para converter uma sequência de movimentos em ações. Depois de interpretadas, essas ações possibilitarão a execução de ações pré-definidas como pedidos e frases por parte do seu usuário.

FUNDAMENTAÇÃO TEÓRICA

De acordo com um levantamento da Organização Mundial da Saúde (OMS), uma em cada sete pessoas possuem alguma deficiência, resultando em mais de um bilhão de pessoas (HUSSAIN, JAGOE, *et al.*, 2018). No Brasil, de acordo com um levantamento feito em 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), 23,9% da população tem pelo menos algum tipo de deficiência.

Tecnologias assistivas têm sido desenvolvidas para melhorar a qualidade de vida de pessoas portadoras de alguma deficiência, expandindo a sua capacidade de interação com o ambiente por meio da utilização de computadores, sensores e atuadores (NUNES, MACHADO e NIRAES, 2014). Os trabalhos desenvolvidos apresentam uma grande variedade de soluções para cada situação, como o de Silva, Reidrich e Bassani (2007), que apresenta um sintetizador de voz para leitura de telas.

Diversos trabalhos exploram o conceito de visão computacional e processamento de imagens para suprir as necessidades de comunicação. O sistema proposto em Aksu e Aydin (2020) propõem o uso do algoritmo de Viola Jones para a detecção da face e, considerando qual olho foi fechado, é percorrido uma árvore binária contendo os caracteres. Já o sistema proposto por Królak e Pawel (2011) passa por mais etapas, utilizando o “*Gentle Adaptive Boosting*” para detectar o rosto, definindo a área dos olhos através de

propriedades anatômicas, reconhecendo as piscadas pela aplicação do “*template matching*” e percorrendo uma tabela de caracteres com as piscadas.

Os autores Pandey et al. (2019) também utilizam uma *webcam* para fazer o reconhecimento dos olhos, mas, nesse caso, para reconhecer a face é utilizado um *shape predictor* que coloca 68 marcas em regiões específicas do rosto. Após isso, é aplicada uma equação do *eye aspect ratio*, a qual retorna à relação entre o olho aberto e o olho fechado. Para ser considerada uma piscada é necessário que esta relação se mantenha abaixo de um limiar por um período específico de tempo.

Em geral, os trabalhos citados têm um foco na detecção das piscadas, centrando-se nos estudos e testes de métodos de visão computacional e deixando um de lado a transformação e *interface* do usuário, mesmo que a *interface* do usuário tem grande importância no sistema. Uma má usabilidade de *interface* pode causar desde um desconforto do usuário até a inutilização total do sistema. Levando isso em conta, este trabalho visa não somente no método de visão computacional, mas também como esses dados são traduzidos, processados e mostrados ao usuário.

METODOLOGIA

Desenvolveu-se uma pesquisa exploratória com a análise de trabalhos acadêmicos relacionados as tecnologias assistivas e a detecção de piscadas. Essas pesquisas tinham como intuito conhecer os métodos já estabelecidos na literatura, além de definir os requisitos de funcionamento, os recursos necessários.

A partir deste levantamento, foi possível identificar as principais lacunas e tecnologias utilizadas para este tipo de pesquisa. Neste sentido, foi desenvolvido um protótipo funcional de baixa fidelidade representando o funcionamento do sistema. Uma vez que tal protótipo seja validado, o seu desenvolvimento será composto por uma *interface* com o usuário a ser desenvolvida nas linguagens HTML, JAVASCRIPT e CSS. Para realizar a integração entre a *interface* com o usuário e o processamento das imagens será o utilizado o framework FLASK (FLASK, 2020) juntamente com biblioteca OpenCV (OPENCV, 2020) na linguagem python, a qual será responsável por converter as piscadas dos usuários em ações.

1° Passo

Imagem retirada da webcam do usuário. Não é necessário nenhuma fonte de luz adicional ou equipamento específico.



3° Passo

Escalonamento da imagem a fim de diminuir a sua resolução e o trabalho necessário para processá-la.



2° Passo

Transformação das cores da imagem em escalas de cinza, facilitando a aplicação dos modelos e removendo possíveis ruídos.



4° Passo

Detecção do rosto, aplicação das marcas faciais e cálculo do *eye aspect ratio*, para assim saber se o usuário está piscando.



Figura 1. Passos do processamento da imagem e aplicação das marcas faciais

Fonte: próprio autor

Para realizar o processamento e o reconhecimento das piscadas, este trabalho se baseará na metodologia proposta pelos autores Pandey et al. (2019), a qual possui uma menor taxa de erro quanto ao reconhecimento dos movimentos das pálpebras e uma facilidade na otimização dos cálculos. Para realizar o reconhecimento de uma piscada será necessário realizar a transformação das cores do vídeo em uma escala de cinza, diminuir a resolução da imagem, aplicar o *shape predictor*, calcular *eye aspect ratio* e comparar com os limiares, tal como representado na Figura 1.

Uma vez desenvolvido, este sistema será avaliado de forma quantitativa e qualitativa. A avaliação quantitativa ocorrerá através de uma avaliação de desempenho em que serão quantificados aspectos como taxa de processamento e consumo de memória. Já a avaliação qualitativa, ocorrerá por meio da aplicação de questionários de avaliação aos usuários do sistema.

RESULTADOS E DISCUSSÃO

Como resultado deste projeto até o momento, pode-se destacar o desenvolvimento de um protótipo de baixa fidelidade do sistema. Na Figura 2, é possível observar a página principal do sistema, a qual está dividida em três áreas. A primeira área, destacada no canto superior esquerdo contém a imagem captada pela *webcam*, na qual será realizado o processamento descrito na Figura 1.

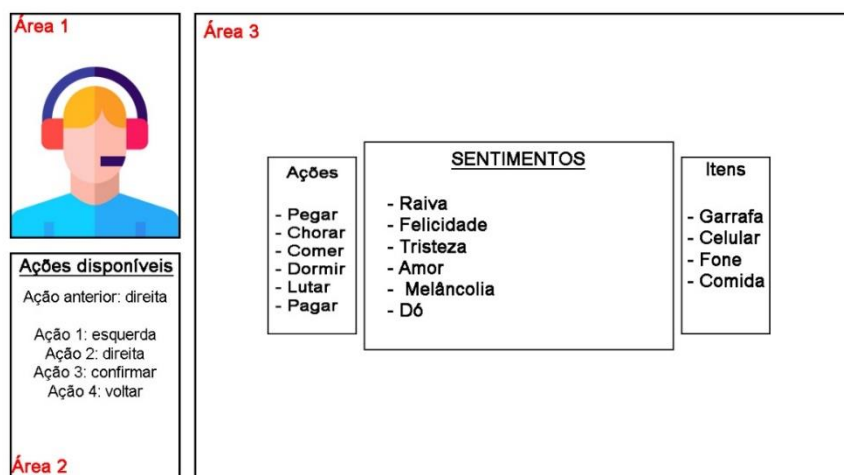


Figura 2. Página principal

Fonte: próprio autor

A segunda área, representa as ações disponíveis ao usuário. A princípio serão utilizadas apenas quatro ações, as quais são utilizadas por meio de uma sequência específica de piscadas, tais como descritas na Tabela 1. A função de cada ação é definida dinamicamente no sistema e por isso são descritas em uma área separada. É esperado que com o número baixo de ações disponíveis, o usuário do sistema passe a domina-lo mais rapidamente e não tenha ações detectadas indevidamente.

Tabela 1. Relação da sequência de piscadas e a ações executadas

Fonte: próprio autor.

Sequência de piscadas	Ação Executada
Fechado, aberto, fechado, fechado e fechado	Esquerda
Fechado, aberto, fechado, aberto e aberto	Direita
Fechado, aberto, fechado, aberto e fechado	Ação 1
Fechado, aberto, fechado, fechado e aberto	Ação 2

A terceira área, corresponde a navegação do usuário no sistema a qual representará as possíveis formas de comunicação do usuário, representando desejos ou necessidades. A interação do usuário com esta parte do sistema ocorre por meio das ações descritas na segunda área que devem ser transmitidas por meio das piscadas.

A princípio, a representação dos desejos e necessidades do usuário devem ser cadastradas por pessoas que não possuem nenhum tipo de limitação. Na Figura 3 é representada a página de cadastro dos desejos e necessidades do usuário, na qual basicamente devem ser incluídos o nome da ação, uma imagem e a categoria a qual pertence. O agrupamento de ações em categorias visa diminuir a quantidade de ações a serem enviadas pelo usuário ao sistema.

<p><u>Cadastrar Ações</u></p> <p><u>Cadastrar Categorias</u></p> <p><u>Editar Ações</u></p> <p><u>Editar Categorias</u></p>	<p style="text-align: center;">CADASTRAR AÇÕES</p> <p>Nome: <input type="text"/></p> <p>Categoria: <input type="text" value="Sentimentos"/> <input type="text" value="Ações"/> <input type="text" value="Itens"/></p> <p>Imagem: <input type="button" value="Carregar Imagem"/></p>
---	---

Figura 3. Cadastro e edição de ações e categorias
Fonte: próprio autor

CONCLUSÕES

Este trabalho apresentou um protótipo de baixa fidelidade de um *software* para permitir a comunicação de pessoas com deficiência por meio dos olhos. Para isso, o protótipo se baseará nas piscadas dos usuários as quais serão convertidas em quatro ações. As finalidades de cada ação são dinâmicas e dependem da página que o usuário estiver.

Os próximos passos deste trabalho serão compostos pelo desenvolvimento do protótipo funcional e da sua avaliação qualitativa e quantitativa.

REFERÊNCIAS

- AKSU, D.; AYDIN, M. A. Human Computer Interaction by Eye Blinking on Real Time. **IEEE**, Girne, Cyprus, 2018 Março 2020. 135-138.
- FLASK. Flask. **Flask**, 2020. Disponível em: <<https://flask.palletsprojects.com/en/1.1.x/>>. Acesso em: 21 set. 2020.
- HUSSAIN, N. et al. **The Importance of Speech, Language and Communication to the United Nations Sustainable Development Goals: A Summary of Evidence**. 1. ed. [S.l.]: [s.n.], v. 1, 2018.
- ITU-R. International Morse Code. **international Telecommunication Union.**, Genebra, Novembro 2009.
- KRÓLAK, A.; PAWEL, S. Eye-blink detection system for human-computer interaction. **Springer**, 2 Outubro 2011. 409-419.
- NUNES, F. L. S.; MACHADO, L. S.; NIRAES, R. M. Evolution of Virtual and Augmented Reality in Health: A Reflection from 15 Years of SVR. **IEEE**, Piata Salvador, Brazil, p. 220-228, Maio 2014. ISSN 978-1-4799-4261-9.
- OPENCV. OpenCV. **OpenCV**, 2020. Disponível em: <<https://opencv.org/>>. Acesso em: 21 set. 2020.
- PANDEY, M. et al. Assistance for Paralyzed Patient Using Eye Motion Detection. **IEEE**, Pune, India, Índia, 25 Abril 2019.
- SILVEIRA, C.; REIDRICH, R. D. O.; BASSANI, P. B. S. Avaliação das tecnologias de softwares existentes para a Inclusão Digital de dEficientes visuais através da utilização de Requisitos de qualidade. **RENOTE - Revista Novas Tecnologias na Educação**, Porto Alegre, v. 5, n. 1, Julho 2007. ISSN 1679-1916.