

Criação de uma Base de Dados de Rastreamento Ocular no Espectro Infravermelho e de Luz Visível

Guilherme Gomes Luccas Rodrigues¹, Fabricio Batista Narcizo^{2,3}, Mario Tadashi Shimanuki⁴

¹ Cursando Tecnologia em Análise e Desenvolvimento de Sistemas, Bolsista PIBIFSP 2023, IFSP, Campus Caraguatatuba, rodrigues.luccas@aluno.ifsp.edu.br

² Part-time Lecturer, Computer Science Department, IT University of Copenhagen, ITU, fabn@itu.dk

³ Senior Software Engineer, The Office of CTO (OCTO), GN Audio A/S, Jabra, fbncarizo@jabra.com

⁴ Professor do curso de Análise e Desenvolvimento de Sistemas, Instituto Federal de São Paulo, IFSP, Campus Caraguatatuba, mario@ifspcaragua.net

Área de conhecimento (Tabela CNPq): Sistemas de Computação - 1.03.04.00-2

RESUMO: O rastreamento ocular é uma técnica de monitoramento das atividades oculares que tem por objetivo estimar o ponto de maior interesse de um indivíduo no seu campo visual. Atualmente, o rastreamento ocular pode ser encontrado em diversos sistemas computacionais, como dispositivos de controle a base dos movimentos oculares, sistemas de detecção de doenças correlatas ao sistema visual humano, treinamento de modelos de inteligência artificial focados nos movimentos oculares, dentre outros. Este projeto de pesquisa tem como objetivo a criação de uma base de dados de imagens capturadas dos olhos de voluntários nas frequências infravermelho e de luz visível do espectro eletromagnético. Essa base de dados tem como principais vantagens a ampla diversidade dos participantes envolvidos na coleta dos dados, as diferentes angulações dos dispositivos de captura de vídeo, os amplos movimentos oculares com base nos alvos observados, e o controle dos movimentos da cabeça por meio de um apoiador de queixo. Os resultados obtidos com essa base de dados poderão ser utilizados para o desenvolvimento de projetos envolvendo o rastreamento ocular via processamento de imagens em ambos espectros de luz. A base de dados contém imagens dos olhos observando um alvo em movimento por 35 posições pré-estabelecidas (matriz de 7x5) e mais 65 posições aleatórias. Para cada uma das 100 posições observadas, foram capturadas três imagens em câmeras distintas, com o voluntário a 62 centímetros do monitor, utilizando uma câmera posicionada no centro da tela e duas 15° abaixo do centro da tela, sendo duas de luz visível e uma infravermelho.

PALAVRAS-CHAVE: rastreamento ocular; estimativa do olhar; correção ocular.

1 INTRODUÇÃO

O rastreamento ocular (*eye tracking*) é uma tecnologia versátil por conta da sua gama de possibilidades de uso em aplicações computacionais [1], sendo utilizada em diversas áreas além de resultar em muitos projetos nos quais o uso dessas tecnologias é relevante. A técnica de rastreamento ocular consiste, em termos gerais, no monitoramento da posição relativa dos olhos durante tarefas que envolvem a visualização de estímulos visuais, sejam eles imagens, objetos ou textos escritos [2], por exemplo, a Figura 1 ilustra um projeto realizado pelos engenheiros da *LC Technologies*,

no qual desenvolveram um sistema de rastreamento ocular para que pessoas com deficiências físicas possam continuar conectadas ao mundo. Nesse sistema, usuários podem controlar o cursor de um computador apenas pelos movimentos oculares, pois o sistema detecta o ponto de interesse do usuário por meio de cálculos envolvendo a posição da pupila e as reflexões na córnea.



Figura 1. Sistema de reconhecimento ocular para pessoas com deficiências físicas.

Fonte: https://img.vision-systems.com/files/base/ebm/vsd/image/2015/12/snap2_tracker_1305vsd.png?auto=format,compress&w=1050&h=590&fit=max

2 TEORIA

As pesquisas de rastreamento ocular se baseiam na obtenção e análise de dados de diversas pessoas em um ambiente controlado [3], e com base nesses dados é possível realizar estudos para determinar vários fatores, por exemplo, o tempo em que a pessoa olha para determinado alvo, os principais focos de atenção visual, o trajeto dos movimentos oculares, a dilatação da pupila, dentre outros. Pelo fato das técnicas de rastreamento ocular utilizarem cálculos envolvendo o olho humano [4], os resultados podem diferir por conta de fatores como modelos matemáticos, ângulo, formato, dispositivos, e tipo de iluminação, pois mesmos em condições idênticas pode haver diferenças entre indivíduos. Desse modo, para estudos ou projetos feitos a partir do rastreamento ocular, o uso de uma ampla base de dados ocular é de extrema importância.

Uma base de dados de rastreamento ocular, pode ser utilizada para treinar detectores visuais [5], modelos de redirecionamento dos olhos [6][7], detectores biométricos de análise da íris [8], detectores do piscar dos olhos [9], dentre outros. Por exemplo, a base de dados *Columbia Gaze* [10] é amplamente utilizada em estudos de rastreamento ocular, pois contém dados de 56 pessoas somando no total 5,880 imagens, sendo essas capturadas de 5 angulações diferentes em um intervalo de 30° no campo de visão. A base de dados *Columbia Gaze*, também apresenta diversidade em seus dados onde, em suas estatísticas, para cada voluntário foram catalogados, sexo, etnia e o uso de lentes oftálmicas de correção ocular.

Utilizando como base o *Columbia Gaze*, esse projeto fez adaptações a requisitos pré-estabelecidos, como a adição de uma câmera infravermelho para a captação dos dados. Por conta disso, foi iniciado o processo de elaboração de um ambiente de interação do usuário adequado para a coleta desses dados e na elaboração da lógica e cálculos usados para a captura dos dados.

3 MATERIAL E MÉTODOS

A metodologia utilizada definiu um ambiente sistematizado para a aquisição das imagens, a criação de um processo padronizado para a captura, e a elaboração de uma base de dados com todos os arquivos obtidos pelo sistema com a ajuda de voluntários.

Os testes pilotos consistiram em uma breve explicação de como o sistema funciona seguido de uma confirmação de consentimento na coleta de dados e de utilização da imagem em outros projetos ou testes. Para os procedimentos oficiais, além dos mesmos procedimentos feitos nos testes pilotos, foi adicionado um suporte para facilitar a coleta de dados e evitar que alguns dados sejam perdidos por conta de movimentos entre as capturas. Para o procedimento oficial é esperado uma média entre 100 a 110 pessoas para formar uma base de dados com mais de 30.000 imagens oculares capturadas.

Para o procedimento de coleta foram utilizados os seguintes equipamentos: (i) monitor de 24 polegadas, com 1980×1080 de resolução, 54.3 cm de largura × 32.2 cm de altura; (ii) 2 câmeras conectadas via USB 3.0 com 1980×1080 em 30 FPS (*frames per second*); (iii) 1 câmera infravermelho conectada na porta USB 2.0 com resolução de 1980×1080 e a taxa de captura de vídeo em 30 FPS; (iv) 1 suporte de apoio para a cabeça.

Todos os voluntários foram posicionados a uma distância de 62 centímetros do monitor. As câmeras foram posicionadas em 0° grau em relação ao centro do monitor, sendo 1 câmera RGB na parte superior do monitor, 1 câmera RGB e 1 câmera infravermelho na parte inferior. Desse modo, foi possível coletar dados de rastreamento ocular de diferentes ângulos simultaneamente. A Figura 2 ilustra a posição das câmeras e do usuário no ambiente.

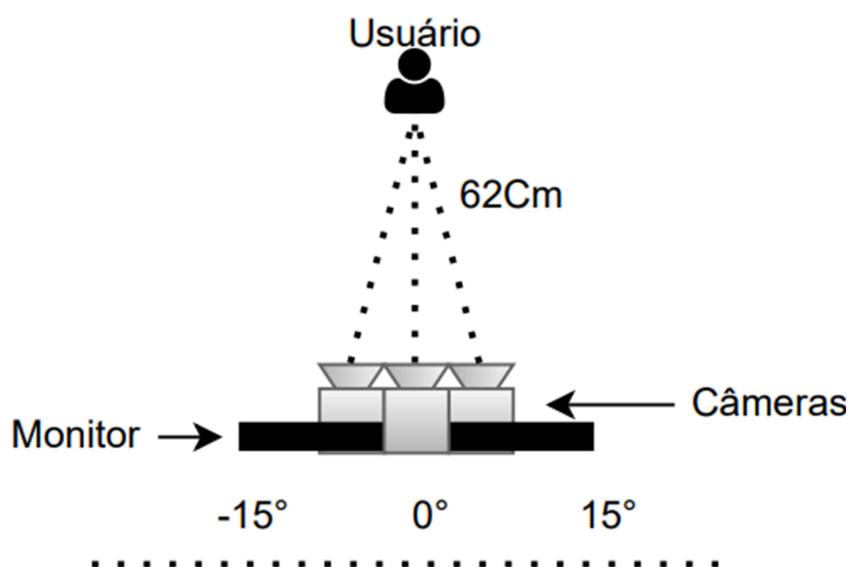


Figura 2. Demonstração gráfica do ambiente.

A Figura 3 ilustra o alvo em sua posição inicial. A captura é finalizada após capturadas 100 imagens, sendo elas de 35 pontos em uma matriz de 7 x 5 apresentadas em ordem randomicamente. Logo após, são coletadas outras 65 imagens em pontos aleatórios no monitor não determinados previamente.

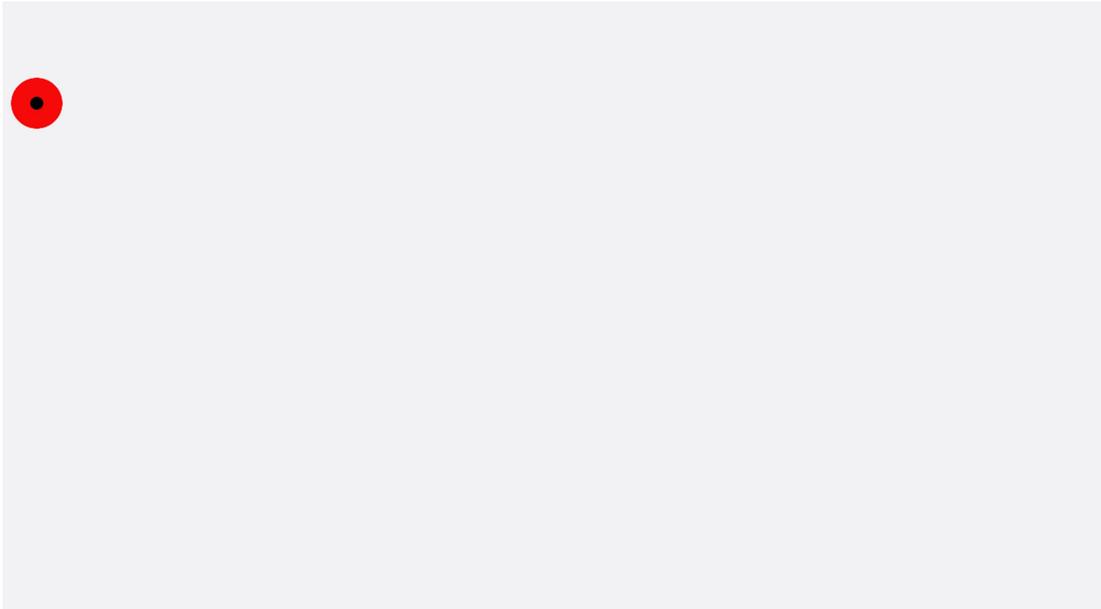


Figura 3. Demonstração do alvo inicial.

4 RESULTADOS E DISCUSSÃO

A Figura 4, ilustra a junção de imagens capturadas de um voluntário observando as duas câmeras visíveis (uma em cima e uma em baixo) e a câmera infravermelho, além do primeiro e último ponto.



Figura 4. Compilado de dados de um voluntário.

Como ilustrado na Figura 5, para evitar a perda de dados por conta de bruscas mudanças na posição do usuário entre as capturas, foi utilizado um apoio para a cabeça, deixando o ambiente de coleta de dados utilizado similar a Figura 5.

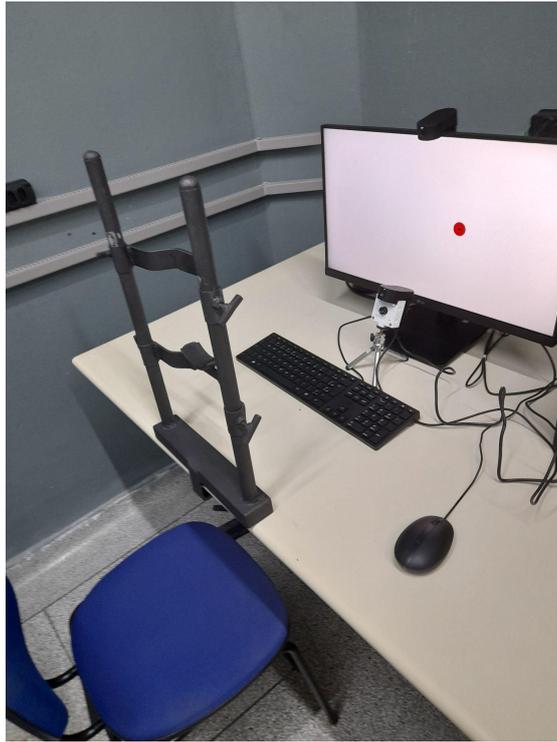


Figura 5. Imagem do setup utilizado.

5 CONSIDERAÇÕES FINAIS

Após a criação da base de dados, o projeto terá atingido o objetivo de criar uma base de dados composta por imagens em câmeras visíveis e infravermelhos, tendo uma diferença entre a diversidade das câmeras e a quantidade de capturas por usuário e capturas totais superior em comparação a base de dados do Columbia Gaze. A criação dessa base de dados beneficia projetos em âmbito acadêmico e científico e poderá auxiliar em futuros projetos de rastreamento ocular que necessitem de uma base de dados similar.

Uma base de dados com o principal foco em reconhecimento ocular auxilia em futuros desenvolvimentos de projetos que necessitem de um treinamento, como projetos que ajudem pessoas que possuem deficiências físicas e/ou cognitivas ou para o treinamento de inteligências artificiais que necessitem de um treinamento para o seu aprendizado, projetos de pesquisas de interesse ativo entre pesquisadores na área de rastreamento ocular.

Agradecimentos:

¹ Os autores desejam agradecer ao IFSP-CAR e ao Programa de Bolsas de Iniciação Científica do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (PIBIFSP 2023) pela concessão da bolsa de Iniciação Científica, o que tornou possível a realização deste projeto.

² Os autores agradecem ao CNPq – *Conselho Nacional de Desenvolvimento Científico e Tecnológico* – pelo apoio financeiro ao autor *Fabricio Batista Narcizo* durante os estudos do seu doutoramento [Processo nº: 229760/2013-9].

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] NARCIZO, Fabricio Batista; DE QUEIROZ, José Eustáquio Rangel; GOMES, Herman Martins. Remote Eye Tracking Systems: Technologies and Applications. 2013 26th Conference on Graphics, Patterns and Images Tutorials, p. 15-22, 2013.
- [2] FORSTER, Renê. Aspectos da Utilização do Rastreamento Ocular na Pesquisa Psicolinguística. Revista de Documentação de Estudos em Lingüística Teórica e Aplicada, 33.2 (2017): 609-44. Web.
- [3] HANSEN, Dan, Witzner; JI, Qiang. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. In IEEE Transactions on Pattern Analysis and Machine Intelligence, [S. l.], v. 32, n. 3, p. 478-500, mar. 2010.
- [4] NARCIZO, Fabricio Batista; DOS SANTOS, Fernando Eustáquio Dantas; HANSEN, Dan Witzner. High-Accuracy Gaze Estimation for Interpolation-Based Eye-Tracking Methods. Vision, 2021, 5, 41. <https://doi.org/10.3390/vision5030041>
- [5] BISSOLI, Alexandre, Daniel Lavino-Junior, Mariana Sime, Lucas Encarnação, and Teodiano Bastos-Filho. "A Human-Machine Interface Based on Eye Tracking for Controlling and Monitoring a Smart Home Using the Internet of Things." Sensors (Basel, Switzerland) 19.4 (2019): 859. Web.
- [6] JIN, Shiwei; WANG, Zhen; WANG, Lei; BI, Ning; NGUYEN, Truong. ReDirTrans: Latent-to-Latent Translation for Gaze and Head Redirection. 2023 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 5547-5556, 2023.
- [7] RUZZI, Alessandro; SHI, Xiangwei; WANG, Xi; LI, Gengyan; DE MELLO, Shalini; CHANG, Hyung Jin; ZHANG, Xucong; HILLIGES, GazeNeRF: 3D-Aware Gaze Redirection With Neural Radiance Fields. 2023 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 9676-9685, 2023.
- [8] NGUYEN, Kien; PROENÇA, Hugo; ALONSO-FERNANDEZ, Fernando. Deep Learning for Iris Recognition: A Survey. arXiv 2210.05866, 2022.
- [9] ZENG, Wenzheng; XIAO, Yang; WEI, Sicheng; GAN, Jinfang; ZHANG, Xintao; CAO, Zhiguo; FANG, Zhiwen; ZHOU, Joey Tianyi. Real-Time Multi-Person Eyeblink Detection in the Wild for Untrimmed Video. 2023 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 13854-13863, 2023.
- [10] SMITH, B. A.; YIN, Q.; FEINER, S. K.; NAYAR, S. K. Gaze Locking: Passive Eye Contact Detection for Human? Object Interaction. ACM Symposium on User Interface Software and Technology (UIST), p 271-280, 2013.