

ESTIMATIVA DO OLHAR UTILIZANDO O RASTREAMENTO DA ÍRIS

BRUNO ADRIANO ALCANTARA FERREIRA¹, FABRICIO B. NARCIZO²,
MARIO T. SHIMANUKI³

¹ Graduando em Tecnologia de Análise e Desenvolvimento de Sistemas, Bolsista PIBIFSP, IFSP, Câmpus Caraguatatuba, brunoa.f@hotmail.com

² Professor no IT University of Copenhagen, ITU, Copenhagen, narcizo@itu.dk

³ Professor no Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, IFSP, Câmpus Caraguatatuba, mario.shimanuki@gmail.com

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

RESUMO: O rastreamento ocular é aplicado no estudo de diversas áreas, como na análise de atletas de alta performance, comportamento de motoristas, jogos online e ferramentas para interação humano-computador, no qual pode-se determinar comportamentos e encontrar padrões. O objetivo deste trabalho é implementar uma aplicação para estimar a posição onde um usuário está focado durante o uso de um computador através de coordenadas capturadas durante o rastreamento da íris, utilizando a câmera de vídeo de um computador. Realizada a estimativa, é possível utilizar os dados capturados para o desenvolvimento de outras implementações como analisar o comportamento humano ou desenvolver aplicações para interagir com o movimento ocular.

PALAVRAS-CHAVE: rastreamento ocular; detecção da íris; visão computacional;

1 INTRODUÇÃO

O rastreamento ocular é um tema que aborda diferentes tecnologias e métodos, podendo ser intrusiva como no implante escleral e na eletro-oculografia ou não intrusiva através da captura de vídeo, o objetivo comum é registrar e analisar os movimentos oculares de um indivíduo, buscando determinar a área de sua atenção e possivelmente detectar padrões de comportamento visual (Barreto, p.168).

O tema é abordado em diferentes áreas, dentre elas analisar atletas de alta performance (Klatt, Noël, Memmert, 2018), no setor comercial empresas investem para analisar a perspectiva do consumidor e melhor posicionar seus produtos (Duarte, Rodrigues, Paiva, 2015), no setor de jogos online (Sundstedt, 2010), ferramentas para auxiliar pessoas com mobilidade reduzida, interação humano-computador (Poole, Ball, 2006) e comportamento de motoristas (Campos *et al.*, 2020).

Visto a ampla área onde o tema é abordado o objetivo é implementar um sistema para estimar o olhar do usuário durante o uso de um computador utilizando apenas um dispositivo de vídeo comum, como a *webcam* de um computador.

2 TEORIA

As tecnologias de rastreamento ocular possuem variadas aplicações, sendo utilizadas para análise de informações ou interação, mostrando-se versátil em contextos na internet, jogos, decoração em pontos de venda e eventos. Na aplicação comercial é eficiente para analisar o comportamento do consumidor e assim melhorar a organização em campanhas e divulgação de produtos por exemplo (Barreto, p. 169), no contexto dos

esportes podemos analisar o comportamento de atletas e também pode-se desenvolver soluções para auxiliar pessoas com mobilidade reduzida a interagir com sistemas.

Existem métodos intrusivos e não intrusivos, o implante escleral, realizado através da inserção de uma lente na esclera para capturar variações no campo eletromagnético, e eletro-oculografia, utilizando eletrodos posicionados próximos aos olhos, são exemplos de métodos intrusivos (Forster, p.612, 2017).

Com o avanço das tecnologias, o rastreamento por vídeo tornou-se largamente utilizado pelo fato de não ser intrusivo, o método mais popular é através de uma câmera de vídeo ou um dispositivo equivalente que projete uma luz infravermelha nos olhos, o reflexo da pupila é capturado e utilizado para estimar o olhar (Barreto, p.172).

Outro método de rastreamento por vídeo é através da íris (Park *et al.*, 2018), sua vantagem é dispensar um aparelho com projeção do infravermelho mas necessita uma iluminação no ambiente.

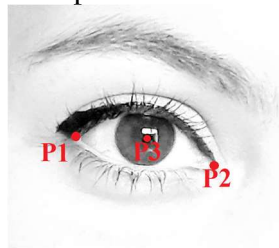
3 MATERIAL E MÉTODOS

Para o desenvolvimento da pesquisa a linguagem Python foi utilizada em conjunto com a biblioteca OpenCV para captura e manipulação de *frames*.

O método de detecção utilizado foi através da detecção da íris, utilizando uma *webcam* Logitech c270 em sessenta *frames* por segundo, posicionada em frente ao usuário na parte superior do monitor.

A implementação do projeto GazeML foi utilizada para o rastreamento da íris e retorno das coordenadas de interesse, os cantos dos olhos (P1 e P2) e o centro da íris (P3) exemplificado na figura 1.

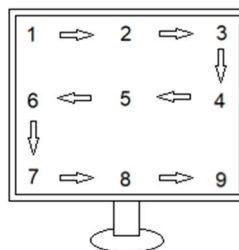
FIGURA 1. Coordenadas de interesse para estimativa do olhar.



Fonte: Galeria do autor

Para iniciar o processo de estimativa é necessário realizar uma calibração, o processo de calibração consiste na exibição de nove pontos na tela para o usuário observar enquanto captura-se as coordenadas de interesse, a figura 2 exemplifica a ordem e sequência em que cada ponto é exibido, um de cada vez.

FIGURA 2. Ordem e sequência dos pontos durante a animação.



Fonte: Galeria do autor

Sessenta *frames* são capturados por ponto onde a animação é exibida, para cada *frame* é aplicada a equação 1 para obter a diferença entre o centro do olho e a posição atual do centro da íris.

$$R = \begin{bmatrix} \frac{P1x + P2x}{2} & \frac{P1y + P2y}{2} \end{bmatrix} - \begin{bmatrix} P3x & P3y \end{bmatrix} \quad (1)$$

em que,

R - diferença entre o centro do olho e a posição atual do centro da íris, pixels;

P1 - coordenadas x e y de P1, pixel;

P2 - coordenadas x e y de P2, pixel;

P3 - coordenadas x e y de P3, pixel,

Para cada posição da animação, calcula-se a média dos seus sessenta *frames*, obtendo uma matriz de 9x2.

$$M = \begin{bmatrix} \frac{\sum_{i=1}^{60} Rx_i}{60} & \frac{\sum_{i=1}^{60} Ry_i}{60} \end{bmatrix} \quad (2)$$

em que,

M - média dos sessenta *frames*, pixel;

R - coordenada x e y, resultante de cada frame processado, pixel.

Após processada a média para cada posição, calcula-se a matriz de calibração, através dela pode-se estimar a direção do olhar do usuário para novas entradas.

$$E_{(R)} = \begin{bmatrix} M^+ \cdot T_x & M^+ \cdot T_y \end{bmatrix} \cdot R \quad (3)$$

em que,

E - coordenadas estimadas, pixel;

M⁺ - matriz resultante da equação 2, pixel;

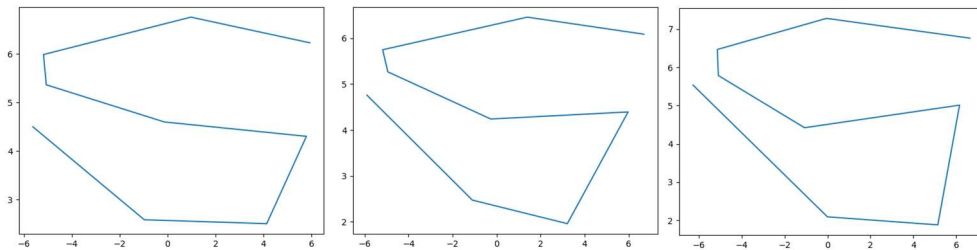
T - vetor com as coordenadas x e y de exibição da animação, pixel.

R - nova entrada para estimativa, pixel;

4 RESULTADOS E DISCUSSÃO

A média dos pontos processados durante cada coordenada da calibração, apresentado na figura 3, gera um resultado semelhante ao caminho percorrido pelo olhar de forma espelhada.

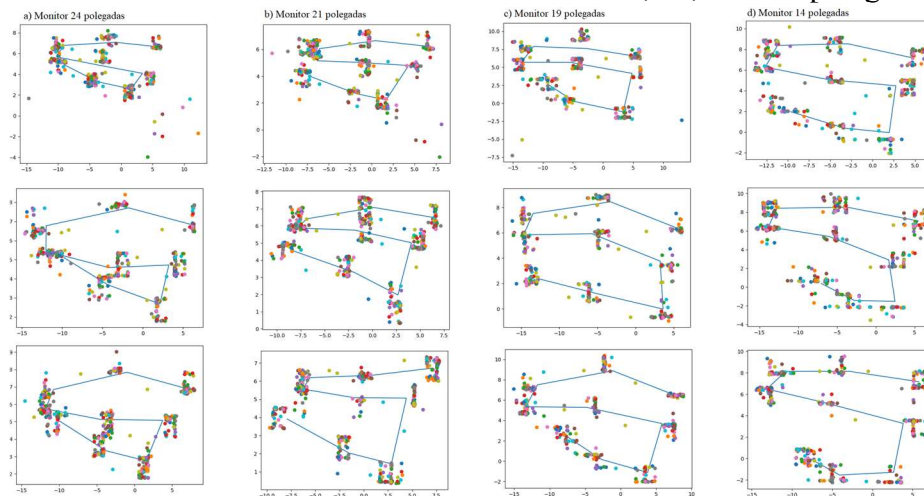
FIGURA 3. Ponto médio durante a calibração.



Fonte: Galeria do autor

A figura 4 apresenta os testes realizados em diferentes tamanhos de monitores, exibindo com linha azul a média dos *frames* processados por ponto na calibração, e nos pontos coloridos o resultado para cada *frame* processado.

FIGURA 4. Testes realizados em monitores de 24, 21, 19 e 14 polegadas.



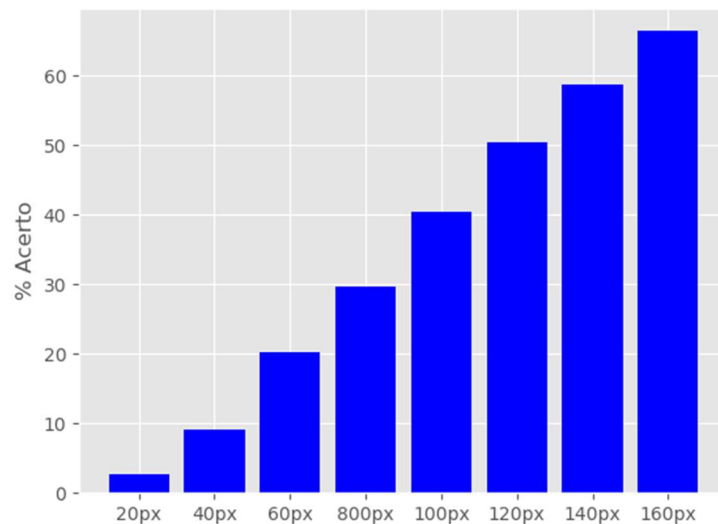
Fonte: Galeria do autor

Cada calibração apresenta um resultado diferente, a distância do usuário para o dispositivo de vídeo, posição e ângulo do rosto influenciam nos valores gerados, portanto a calibração é impactada sempre que houver alterações nessas condições, necessitando uma nova calibração.

Com a calibração executada, é possível processar os novos *frames* e estimar a coordenada para onde o usuário está olhando.

A figura 5 ilustra em uma amostra de cinco mil e quatrocentos frames processados, a porcentagem de acertos. Para o cálculo foi definido uma área em volta do ponto previsto como região de foco, as amostras foram testadas com variação de dez até cem pixels obtendo uma precisão aproximada de 66% na variação de até cento e sessenta pixels e 2.59% em uma variação de até vinte pixels.

FIGURA 5. Porcentagem de acertos da estimativa utilizando monitor de 14 polegadas.



Fonte: Galeria do autor

5 CONSIDERAÇÕES FINAIS

A precisão durante a estimativa do olhar apresenta oscilações quando o usuário altera sua posição, tornando necessário a realização de uma nova calibração e impactando

o resultado da estimativa. A próxima etapa é aplicar a estimativa em um projeto de interação humano-computador, servindo de base para o desenvolvimento de aplicações que auxiliem pessoas com mobilidade reduzida a interagir com sistemas.

AGRADECIMENTOS: Os autores gostariam de agradecer ao Programa Institucional de Bolsas de Iniciação Científica do IFSP (PIBIFSP) – Edição 2020, edital nº 41/2019, pelo apoio financeiro.

REFERÊNCIAS

BARRETO, Ana. Eye tracking como método de investigação aplicado às ciências da comunicação. *Revista Comunicando*, v1, n1, p. 168 - 186, dez. 2012.

FORSTER, Renê. Aspectos da utilização do rastreamento ocular na pesquisa psicolinguística. *DELTA*, São Paulo, v. 33, n. 2, p. 609-644, jun. 2017.

KLATT, Stefanie; NOËL, Benjamin; MEMMERT, Daniel. Eye tracking in high-performance sports: Evaluation of its application in expert athletes. *International Journal of Computer Science in Sport*, dezembro, 2018.

CAMPOS, Cintia; MARCOMINI, Leandro; PANICE, Natália; PIVA, Fernando; LARocca, Ana. Perception analysis of highway quality of service using a driving simulator and eye tracking system. *TRANSPORTES*, v. 28, p. 165-179, ago. 2020.

POOLE, A; BALL, Linden. Eye tracking in human-computer interaction and usability research: Current status and future prospects. *Encyclopedia of Human Computer Interaction*, p. 211-219, jan. 2006.

SUNDSTEDT, Veronica. Gazing at Games: Using eye tracking to control virtual characters. *ACM SIGGRAPH 2010*, jul. 2010.

DUARTE, Paulo; RODRIGUES, Ricardo; PAIVA, Daniel. O impacto das cores, do tamanho e do posicionamento na captação da atenção dos consumidores: uma análise com eye-tracking. *Revista portuguesa de marketing*. v. 18, p. 31-51, 2015.

PARK, Seonwook; ZHANG, Xucong; BULLING, Andreas; HILLIGIES, Otmar. Learning to Find Eye Region Landmarks for Remote Gaze Estimation in Unconstrained Settings. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications (ETRA '18)*. Association for Computing Machinery, NY, USA, 21, 1–10.

GazeML: Gaze estimation using deep learning. Disponível em: <https://github.com/swook/GazeML>. Acesso em: 28 set. 2020.