



ATENÇÃO VISUAL HUMANA E DIREÇÃO AUTÔNOMA: UM EXPERIMENTO

Claudio de Castro Coutinho Filho¹ - Unifesspa
claudio.coutinho@unifesspa.edu.br

Agência Financiadora: FAPESPA

Eixo Temático/Área de Conhecimento: Engenharia Elétrica / Processamento Digital de Sinais

1. INTRODUÇÃO

As Tecnologias da Informação sempre atuaram como um meio facilitador de tarefas humanas. Nos dias atuais, confrontamo-nos com uma série de ferramentas que buscam nos entregar agilidade, acessibilidade, encurtamento de distâncias, entre outras vantagens. Conseguimos interagir com os sistemas computacionais a partir de diferentes atuadores, como as mãos, a fala. Entretanto, a tecnologia percebeu ser possível investigar o que se passa na mente humana através da forma como os olhos da pessoa se comportam.

E os interesses por trás dessa investigação são diversos. Eles vão desde o desenvolvimento de ferramentas que visem reproduzir o comportamento humano – como a condução autônoma de automóveis –, até métodos para otimizar a transmissão de informação – como o reforço de regiões mais “atrativas” de um filme ou imagem em detrimento de zonas que chamem menos a atenção (LE CALLET; NIEBUR, 2013).

Com essa ideia em mente, este trabalho foi concebido no âmbito da exploração dos chamados Modelos de Atenção Visual, que são modelos computacionais desenvolvidos para prever o comportamento da visão humana frente a um estímulo específico (BORJI; ITTI, 2013), para que se propusesse a modelagem da atenção humana durante o ato de dirigir. A pesquisa e o desenvolvimento dos chamados *Autonomous Driving System* (ADS) têm sido de grande interesse da indústria automobilística, pois visa, entre outras coisas, a diminuição de custos de logística (YURTSEVER et al., 2020), melhoria na qualidade de tráfego nas autovias (considerando-se fatores como a diminuição de imprudências e consequentes acidentes) e a ampliação da disponibilidade de serviços.

A ideia inicial deste projeto seria construir um experimento para coleta de dados visuais de observadores comuns. Conforme mostrado nas seções seguintes o experimento seria realizado utilizando a ferramenta PsychoPy, com o auxílio de um rastreador ocular. Considerando que rastreadores oculares de alto desempenho possuem um custo elevado, também surgiu a ideia de se construir um rastreador utilizando equipamentos mais acessíveis como uma *webcam*. Os dados coletados serviriam como ponto de partida para validar um possível modelo de atenção visual que viesse a ser proposto. Esse tipo de dado também é conhecido como *ground truth*.

O andamento do projeto sofreu impacto direto das restrições de distanciamento social causadas pela pandemia de COVID-19, visto que não haveria como coletar dados de pessoas. Mesmo que a problemática do custo de um rastreador ocular de alto desempenho tenha sido parcialmente sanada (o coordenador do projeto teve acesso a um equipamento desse tipo durante seu estágio de Doutorado na Université de Nantes, França), não houve como passar para a fase seguinte do projeto.

Desta feita, o projeto descrito por este artigo aguarda a retomada da possibilidade de procedimentos que envolvam participantes humanos. A continuação do mesmo será possibilitada devido à recente aquisição, por um membro do grupo de pesquisa, de um rastreador ocular comercial.

¹Mestre em Engenharia Elétrica - Professor Assistente da Universidade Federal do Sul e Sudeste do Pará (FACEEL/IGE/Unifesspa).

2. MATERIAIS E MÉTODOS

Resumidamente, o projeto visava o desenvolvimento de um experimento em *software* para coleta de dados visuais. Esses dados seriam, então, utilizados para comparação e proposição de um modelo de atenção visual. Logo, o desenvolvimento do projeto foi dividido em três etapas.

2.1 Ferramentas utilizadas

Abaixo são listadas as ferramentas utilizadas para a construção do experimento.

- PsychoPy: *software* livre de código aberto desenvolvido na linguagem Python. Ele possui uma interface amigável e fornece um *framework* de construção de experimentos que utilizem interação humana como interações visuais táteis e até auditivas. O programa apresenta a facilidade de manuseio e fornece fácil comunicação com periféricos como microfone Mouse e câmera.
- Octave: pacote gratuito de funções Matemáticas que utilizam uma linguagem específica para manipulação de vetores, matrizes e operações lineares e não lineares em geral. Seu principal uso foi para manipulação de imagens que seriam utilizados no experimento.
- Sublime text: editor de texto, com funções voltadas para programação. Muito prático para o desenvolvimento dos códigos do experimento feitos em linguagem Python.

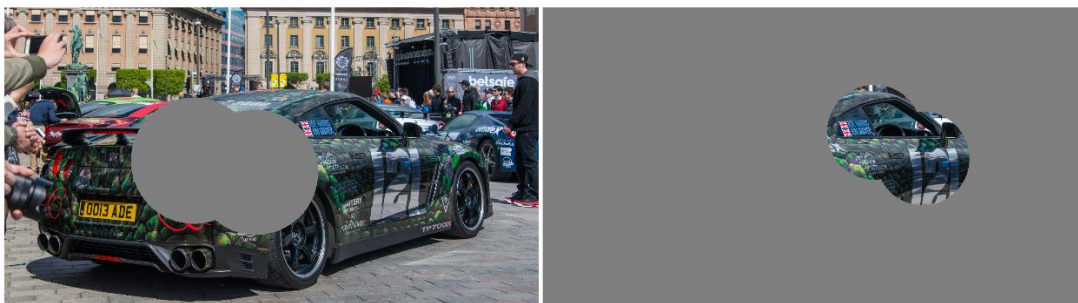
2.2 Etapas

As etapas nas quais o projeto foi dividido serão detalhadas nas subseções a seguir.

2.2.1 Experimento

Para a coleta de dados, seria necessário o desenvolvimento de um *software* capaz de conduzir o experimento. Esse experimento foi construído utilizando-se o *framework* gratuito PsychoPy. Para tanto, seria necessário que o bolsista tivesse conhecimento de manipulação de imagens e da linguagem de programação Python. O experimento a ser desenvolvido consistiria na exibição rápida (em torno de 5 segundos) e progressiva de diversas imagens bidimensionais ao observador, com o intuito de coletar e armazenar as posições das imagens para as quais o observador focou seu olhar. Esses pontos representam as regiões de interesse, ou seja, aquelas que mais chamaram a atenção de um observador em particular.

Figura 1 - Exemplos de exibição dos estímulos com as máscaras central e periférica



Fonte: autoria própria

No mesmo experimento, em algumas imagens, certas regiões da mesma foram mascaradas para que se investigasse o impacto da chamada **atenção periférica**, isto é, o comportamento dos olhos quando sua visão central é obstruída. Essa investigação é necessária para que se entenda como uma pessoa reage a estímulos periféricos e como isso afeta sua visão central, da mesma forma como funciona a atenção de um motorista real.

Essa abordagem de experimento já é bastante difundida na área de rastreamento ocular (DAVID, 2019). A Figura 1 demonstra dois tipos de máscaras utilizadas no experimento.

2.2.2 Coleta de dados

Quando concluído, o programa desenvolvido pelo bolsista deveria coletar dados visuais do participante, ou seja, converter a localização para onde os olhos focassem em coordenadas dentro do plano da tela do computador (tais pontos são chamados de *Point of Regard* ou POR). Ao observador seria apresentado algum estímulo visual que seria uma imagem 2D com conteúdo aleatório esses estímulos seriam apresentados durante poucos segundos. Cada imagem seria apresentada durante o intervalo de 5 segundos ao observador para que não houvesse tempo que ele guardasse informações sobre o que já visualizou. O rastreador ocular então converteria a posição do olhar em coordenadas bidimensionais para uma posterior análise indicar quais porções da imagem seriam mais salientes para o observador. Após a coleta, todos os dados seriam compilados e seria gerado um mapa de calor, indicando as regiões mais salientes para os observadores, como mostra a Figura 2.

É importante salientar que as imagens são apresentadas de maneira aleatória ao observador, de forma que cada participante experencia uma sequência diferente e única dos outros. O conteúdo das imagens é variável, entretanto, ao selecionar as imagens, demos ênfase para aquelas que representassem o ambiente das ruas, como carros, vias, pedestres etc.

Figura 2 - Mapa de calor após coleta de dados, indicando regiões mais salientes da imagem



Fonte: autoria própria

2.2.3 Validação do modelo

Por fim, os dados coletados através do experimento funcionariam como forma de validação do modelo matemático. Através de uma Rede Neural Convolutiva (CNN), o modelo poderia ser “treinado” de forma supervisionada e, após o treinamento, gerar como resultado a região da imagem (o que, na prática, representaria a “visão” do carro autônomo) cuja atenção seria maior, baseando-se em como a atenção periférica humana atuaria em uma situação semelhante. Esta etapa não foi realizada por falta de dados.



3. RESULTADOS E DISCUSSÃO

Após realizar o levantamento bibliográfico, foi necessário fazer o estudo das ferramentas utilizadas para manipulação de imagens e construção do experimento. Este foi concluído com sucesso, bastando conectá-lo a um rastreador ocular capaz de coletar os dados visuais dos participantes. Enquanto não se conseguisse realizar a construção de um rastreador ocular de baixo custo, o experimento foi testado em um rastreador ocular de alto desempenho (EyeLink 1000 Plus) pelo coordenador do projeto durante seu estágio de doutorado em Nantes na França. Os participantes foram colegas de pesquisa do laboratório. O experimento conseguiu com sucesso coletar os dados visuais dos observadores durante alguns testes de funcionamento, entretanto, no momento em que seria iniciada a fase de coleta de dados, instaurou-se a pandemia e o consequente *lockdown* na França, impedindo que houvesse qualquer procedimento com pessoas.

Restou ao bolsista realizar levantamento dos componentes necessários para a construção de um rastreador ocular de baixo custo que seria utilizado após o retorno das atividades na Universidade.

4. CONSIDERAÇÕES FINAIS

Conforme o que foi apresentado acima, concluímos que não foi possível conduzir os experimentos da forma adequada por motivos externos à vontade dos membros da pesquisa. Em outras palavras, como o projeto dependia diretamente dos dados coletados para que se pudesse fazer uma comparação mais adequada, não houve a possibilidade de realizar experimentos com pessoas presencialmente, por restrições impostas pelo distanciamento social em período de pandemia. Os participantes do projeto continuarão engajados em realizar a coleta de dados assim que for possível para que esses dados possam ser utilizados como *ground truth* na comparação com um possível modelo de atenção visual que visse representar o comportamento visual humano ao longo da atividade de condução de um veículo automotor. Isso será possível assim que os testes foram liberados, pois, recentemente, um membro do grupo de pesquisa adquiriu um rastreador ocular de excelente desempenho, que ficará à disposição da pesquisa por parte da equipe. Ao final da condução dos experimentos, o modelo proposto, com toda certeza, representará uma contribuição bastante relevante para a melhoria das condições de um veículo autônomo de se locomover com mais segurança.

REFERÊNCIAS

BORJI, Ali; ITTI, Laurent. **State-of-the-Art in Visual Attention Modeling**. IEEE Transactions On Pattern Analysis And Machine Intelligence, [S.L.], v. 35, n. 1, p. 185-207, jan. 2013. Institute of Electrical and Electronics Engineers (IEEE).

DAVID, Erwan. **L'impact des troubles du champ visuel sur les dynamiques spatio-temporelles de l'observation de scènes naturelles: analyses et modélisation**. 2019. 243 f. Tese (Doutorado) - Curso de Ciência da Computação, Université de Nantes, Nantes, 2019.

LE CALLET, Patrick; NIEBUR, Ernst. **Visual Attention and Applications in Multimedia Technologies**. Proceedings Of The IEEE, [S.L.], v. 101, n. 9, p. 2058-2067, set. 2013. Institute of Electrical and Electronics Engineers (IEEE).

YURTSEVER, Ekim; LAMBERT, Jacob; CARBALLO, Alexander; TAKEDA, Kazuya. **A Survey of Autonomous Driving: common practices and emerging technologies**. IEEE Access, [S.L.], v. 8, p. 58443-58469, 2020. Institute of Electrical and Electronics Engineers (IEEE).