

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

Projeto para a disciplina MAC0499

**Detecção de padrões de leitura por rastreadores de
baixo desempenho**

Candidato: Carlos Eduardo Leão Elmadjian

Orientador: Prof. Dr. Carlos Hitoshi Morimoto

1 Resumo do Projeto

O uso de rastreadores do olhar em pesquisas científicas se estende hoje a diversas áreas, tais como neurociência, psicologia e ciência da computação [5]. Em particular, tem-se observado um crescente número de estudos sobre novas formas de interação com sistemas digitais por meio do olhar desde o final da década de 1990 [10].

Com o avanço da tecnologia móvel e vestível, rastreadores de olhar podem se revelar um novo e apropriado mecanismo de interação em diversas situações, sobretudo naquelas em que há restrições para o uso das mãos. No entanto, há pelo menos dois grandes impedimentos [10] para isso: o custo proibitivo de rastreadores com grande acurácia e a mobilidade desses aparelhos para experiências ubíquas e prolongadas.

Tais restrições impõem a necessidade de se utilizar equipamentos de desempenho técnico limitado, isto é, câmeras para detecção da pupila e reflexão da córnea com taxa de atualização e consumo de energia baixos, um cenário ainda pouco explorado na literatura.

Logo, pretendemos estudar neste projeto a viabilidade de se empregar, sob condições críticas de rastreamento, algoritmos já conhecidos e bem-sucedidos para detecção de padrões oculares. Em especial, dedicaremos nossa atenção para padrões relacionados à leitura, uma vez que se trata de uma das formas mais recorrentes e significativas de interação entre o usuário e sistemas baseados em janelas [3].

Para tanto, buscaremos não somente adaptar os algoritmos conhecidos, mas também, se necessário, propor soluções inteiramente novas para esse cenário. Nosso algoritmo será submetido a testes para aferir robustez a ruídos e ambientes não controlados, bem como para estabelecer limiares das restrições que garantam um patamar mínimo de acurácia desejável. Em caso de sucesso, criaremos um aplicativo como prova de conceito.

2 Introdução

De acordo com Duchowski [5], os estudos com rastreamento do olhar tiveram até agora três grandes eras: um primeiro período, entre o final do século XIX até 1920, marcado pela descoberta e classificação dos movimentos oculares, seguido, em meados do século XX, pela pesquisa behaviorista e psicologia experimental, chegando, até os dias de hoje, a uma era de estudos difusos e abrangentes, possibilitados por grandes avanços tecnológicos.

Na ciência da computação, há um particular interesse por rastreadores como uma interface para novas formas de interação entre usuário e computadores. Um dos possíveis benefícios do emprego de movimentos oculares está no fato de que, ao contrário das mãos, os olhos podem indicar mais do que ações de controle: eles revelam o interesse de um indivíduo por um objeto, dificuldades de compreensão ou ainda traços de personalidade [3]. Além disso, os olhos podem servir de parâmetro para tarefas não seletivas, isto é, atuar de modo complementar às interações convencionais [5], ou ainda prover meios de interação para pessoas que não dispõem do movimento das mãos.

Dentre os cinco padrões típicos de movimentos oculares num sistema de janelas (inspecionar, buscar, explorar, monitorar e ler), o da leitura é comumente o mais recorrente e o que provê mais dados significativos sobre a relação homem-máquina. Contudo, detectar a leitura — que não necessariamente é sequencial e compenetrada — implica diferenciar esse padrão dos demais. Mais do que isso, um sistema capaz de identificá-lo deve ser robusto a fatores de personalidade, complexidade do texto, tamanho e cor das fontes e distância entre usuário e texto. [3]

Uma das maneiras mais empregadas para a detecção do comportamento de leitura é medir o tempo de fixação em palavras em termos de unidades de caracteres [6, 2], o que

pode ser conseguido com rastreadores de alta acurácia e boa resolução. Entretanto, Biedert et al. [1] evidenciaram a falta de robustez dos métodos convencionais em ambientes não controlados, nos quais não é possível corrigir erros de calibração ou estimar o olhar com a ajuda de pontos de controle.

Ainda assim, existem trabalhos que mostram a viabilidade do rastreamento da leitura em condições pouco ou não controladas, sem que haja uma perda significativa de acurácia [1] e sem que as configurações experimentais interfiram de alguma forma nas atividades monitoradas de leitura [9], possibilitando, inclusive, inferir sobre o conteúdo lido por um usuário [8].

Já em relação a câmeras de baixo custo, os poucos estudos existentes se restringem a interações bastante limitadas, como mecanismos de controle com o piscar dos olhos [4] ou estimativas do centro do olhar com o intuito de determinar gestos [12].

Com o avanço da computação vestível, o desafio agora é verificar se os recentes bons resultados com rastreadores em ambientes não controlados resistem às restrições impostas por dispositivos de baixo desempenho, justificadas por configurações de equipamentos que possam ser produzidos em larga escala e supram a necessidade de uso prolongado.

3 Objetivos

Pretendemos estudar neste projeto o problema da detecção de padrões de leitura (*reading, skimming, scanning*) por rastreadores de olhar funcionando em condições críticas de desempenho, como baixa resolução de imagem e baixa taxa de atualização. Para isso, subdividiremos nossa pesquisa em objetivos secundários.

Em um primeiro momento, iremos estudar os algoritmos já existentes para detecção de leitura em situações controladas ou não para, em seguida, aferir o desempenho desses

algoritmos sob as condições restritivas já descritas, verificando se é possível adaptá-los para obter resultados satisfatórios ou que tipo de solução seria necessária para atacar o problema.

Desse modo, um segundo objetivo seria definir o comportamento do algoritmo proposto, mensurando quais os patamares críticos de funcionamento do *hardware* para a manutenção de uma acurácia desejável. Por exemplo, com o equipamento funcionando a 5Hz, nosso algoritmo deverá ser robusto à perda de movimentos sacádicos, cuja preparação leva em torno de 200 ms e execução de 20 a 40 ms [11].

Por fim, o último objetivo será desenvolver um aplicativo, como prova de conceito, que demonstre o algoritmo em execução, como, por exemplo, um leitor de documentos com possibilidade de auto-rolagem a partir da detecção de leitura pelo rastreador, ou ainda um sistema capaz de exibir a tradução de palavras desconhecidas com base no comportamento do olhar, como sugerido por Hyrskykari [6].

4 Justificativas

Dentre as aplicações do uso de rastreadores de olhar na ciência da computação, nosso projeto visa estudar novas formas de interação, sobretudo as não seletivas, sob condições não controladas ou com restrições técnicas. Primeiro, porque cenários como esse representam a maioria das situações a que estariam submetidos rastreadores de uso massivo. Segundo, porque a despeito da expectativa do uso massivo de rastreadores em algum momento no futuro, não há uma literatura estabelecida a respeito.

Quanto à escolha pela pesquisa de padrões de leitura, há tanto uma razão acadêmica quanto técnica. Do ponto de vista acadêmico, a leitura é o tipo de interação pela qual se manifesta da maneira mais clara o interesse do usuário, o que pode ser determinado

por tempo de fixação ou padrão de varredura sobre o texto (e.g. leitura compenetrada ou dispersa) [3]. Do ponto de vista técnico, a leitura é um tipo de interação não seletiva, isto é, não utiliza o olhar como elemento de controle da visualização, o que, nas condições propostas de restrição deste projeto, poderia acarretar em entraves como o *Toque de Midas*, descrito por Jacob et al [7].

Ademais, a tecnologia disponível atualmente para rastreamento ocular foi projetada para atender nichos como pesquisas em neurociência ou oferecer qualidade de vida a pessoas com deficiência físico-motora. Desse modo, os equipamentos associados possuem configurações de *hardware* bastante específicas (de alto desempenho) e não são produzidos em larga escala, o que implica um custo geral proibitivo.

Entretanto, para que rastreadores de olhar possam se tornar dispositivos de propósito geral e massivo, não basta que haja um barateamento. É preciso ainda que o *hardware* disponível seja produzido em larga escala, possa ser amalgamado a dispositivos com alta mobilidade (e.g. vestíveis) e tenha, portanto, um baixo consumo de bateria.

Como um rastreador é composto basicamente por câmeras e unidades de processamento em tempo real, essas restrições estão voltadas primariamente à câmera da pupila. Para que um rastreador portátil e barato possa consumir pouca energia e funcione em qualquer ambiente, sua câmera deverá ter resolução e frequência baixas — gerando uma matriz de *pixels* pequena e em intervalos espaçados para processamento — além de estar suscetível tanto à radiação visível quanto infravermelha.

O gargalo para viabilizar esse novo paradigma em rastreamento, todavia, não está no *hardware* — já que há diversos modelos de câmeras acessíveis que atendem os requisitos — e sim nos algoritmos, que são voltados para os equipamentos tradicionalmente desenvolvidos para fins de uso restrito.

5 Materiais e Métodos

Neste projeto, basearemos nossa pesquisa em soluções para detecção de padrões de leitura já propostas tanto por experimentos realizados em ambientes controlados quanto não controlados. Portanto, os primeiros meses de projeto estarão voltados para um levantamento completo da bibliografia na área. Entre os trabalhos pré-selecionados, citamos os de Campbell et al., Hyrskykari e Buscher et al. (controlados) [3, 6, 2], bem como os de Biedert et al. e Mazzei et al. (não controlados) [1, 9]

Boa parte do projeto também estará voltado para a implementação dos algoritmos e o desenvolvimento de um aplicativo que utilize nossa solução, caso se revele viável. Para tarefas típicas de processamento de imagens (detecção, extração de *features* e rastreamento), utilizaremos a biblioteca OpenCV. Faremos uso de *bindings* em Python para prototipação rápida e de C++ quando o desempenho for um requisito crucial.

Alguns métodos que estimam a posição do olho utilizam modelos tridimensionais do globo ocular para melhorar a precisão do rastreamento [9]. Nesse caso, faremos uso da biblioteca OpenGL, diretamente em C++, ou então em Python, por meio do PyOpenGL.

Para aferir a qualidade do nosso algoritmo, utilizaremos métricas conhecidas [10] e aplicaremos testes de desempenho e robustez, estabelecendo comparações não apenas entre diferentes algoritmos, porém visando determinar diversos limiares de configuração de *hardware* e seu desempenho correspondente.

6 Plano de Trabalho

O cronograma deste projeto envolverá tarefas que abarcam leitura, estudo e aplicação da bibliografia básica a respeito de detecção de padrões de leitura, estudo e desenvolvi-

mento de algoritmos para condições limitantes, testes de desempenho e implementação dos resultados numa possível prova de conceito. Abaixo, segue uma lista detalhada das atividades a serem cumpridas. Relatórios parciais sobre a conclusão de tarefas relevantes serão postados na página do projeto:

<http://www.linux.ime.usp.br/~elmadjian/mac0499>.

6.1 Lista de Atividades

- (a) **Pesquisa bibliográfica:** Iremos conduzir uma revisão bibliográfica com o intuito de estabelecer quais os algoritmos disponíveis para detecção de leitura por rastreadores de olhar. Essa pesquisa também servirá de base para o aprendizado de técnicas de rastreamento.
- (b) **Ferramentas de desenvolvimento:** Parte do projeto envolve o estudo de ferramentas e bibliotecas de desenvolvimento para que os algoritmos estudados possam ser testados. Como o aluno já possui alguma familiaridade com OpenCV e OpenGL, faremos um estudo mais aprofundado dessas ferramentas, bem como um aprendizado de como utilizá-las em plataformas móveis, como o Android.
- (c) **Implementação dos algoritmos:** Nesta parte, selecionaremos alguns dos algoritmos estudados para implementá-los por meio das ferramentas já citadas, ainda sem impor nenhuma restrição de desempenho aos experimentos.
- (d) **Estudo das condições restritivas:** Será realizada uma investigação mais detalhada a respeito das restrições impostas por equipamentos de baixo desempenho. Em particular, iremos avaliar o comportamento dos algoritmos implementados na última etapa sob condições críticas.

- (e) **Relatório parcial:** Redação do relatório parcial do projeto, conforme exigido pela disciplina.
- (f) **Implementação de novos algoritmos:** Após o estudo das restrições e do desempenho dos algoritmos conhecidos, realizaremos um esforço para adaptá-los às novas condições restritivas com o intuito de garantir um patamar mínimo de acurácia para detectar padrões de leitura. Caso as restrições impeçam todos os algoritmos de serem adaptados, proporemos uma solução única e independente para o nosso cenário.
- (g) **Testes e medições:** Utilizaremos testes convencionados na literatura para mensurar o desempenho do nosso algoritmo, mas também proporemos outras formas de medição. Nesta etapa, também serão estudados os limiares para o rastreamento, como, por exemplo, com que resolução ou com que taxa de atualização o algoritmo se torna ineficaz.
- (h) **Desenvolvimento de um aplicativo:** Caso nossa solução apresente um desempenho satisfatório e verificuemos a viabilidade do emprego de rastreadores em condições críticas, desenvolveremos um pequeno aplicativo apenas como prova de conceito da nossa pesquisa.
- (i) **Pôster e monografia:** Confecção do pôster, apresentação e redação da monografia, conforme exigido pela disciplina.

6.2 Cronograma de Execução do Projeto

Item	03	04	05	06	07	08	09	10	11
a	X	X	X						
b		X	X						
c			X	X					
d				X	X				
e					X				
f					X	X	X		
g							X		
h								X	X
i						X	X	X	X

Tabela 1: Cronograma de atividades a serem desenvolvidas no projeto.

6.3 Atividades realizadas

Leitura da bibliografia (item **a**, ainda em progresso) e familiarização com as ferramentas de desenvolvimento (item **b**, também em progresso).

Referências

- [1] Ralf Biedert, Jörn Hees, Andreas Dengel, and Georg Buscher. A robust realtime reading-skimming classifier. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 123–130, New York, NY, USA, 2012. ACM.
- [2] Georg Buscher, Andreas Dengel, and Ludger van Elst. Eye movements as implicit relevance feedback. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, pages 2991–2996, New York, NY, USA, 2008. ACM.
- [3] Christopher S. Campbell and Paul P. Maglio. A robust algorithm for reading detection. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces*, PUI '01, pages 1–7, New York, NY, USA, 2001. ACM.
- [4] Michael Chau and Margrit Betke. Real time eye tracking and blink detection with usb cameras. Technical report, 2005.
- [5] Andrew T. Duchowski. A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments and Computers*, 34(4):455–470, 2002.
- [6] Aulikki Hyrskykari. Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading. *Computers in Human Behavior*, 22(4):657–671, July 2006.
- [7] Robert J. K. Jacob. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Inf. Syst.*, 9(2):152–169, April 1991.

- [8] Kai Kunze, Yuzuko Utsumi, Yuki Shiga, Koichi Kise, and Andreas Bulling. I know what you are reading: Recognition of document types using mobile eye tracking. In *Proceedings of the 2013 International Symposium on Wearable Computers, ISWC '13*, pages 113–116, New York, NY, USA, 2013. ACM.
- [9] Andrea Mazzei, Shahram Eivazi, Youri Marko, Frederic Kaplan, and Pierre Dillenbourg. 3d model-based gaze estimation in natural reading: A systematic error correction procedure based on annotated texts. In *Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA '14*, pages 87–90, New York, NY, USA, 2014. ACM.
- [10] Alex Poole and Linden J. Ball. Eye tracking in human-computer interaction and usability research: Current status and future. In *Prospects”, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc, 2005.*
- [11] Keith Rayner, Barbara R. Foorman, Charles A. Perfetti, David Pesetsky, and Mark S. Seidenberg. How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2):31–74, 2001-11-01T00:00:00.
- [12] Fabian Timm and Erhardt Barth. Accurate eye centre localisation by means of gradients. In Leonid Mestetskiy and José Braz, editors, *VISAPP*, pages 125–130. SciTePress, 2011.