

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

HUMBERTO MOREIRA GONÇALVES

**DETECÇÃO ACURADA DA PUPILA MEDIANTE
OCLUSÕES POR CÍLIOS COM SUPORTE
TEMPORAL ESTENDIDO**

MONOGRAFIA

CAMPO MOURÃO

2017

HUMBERTO MOREIRA GONÇALVES

**DETECÇÃO ACURADA DA PUPILA MEDIANTE
OCLUSÕES POR CÍLIOS COM SUPORTE
TEMPORAL ESTENDIDO**

Trabalho de Conclusão de Curso de Graduação apresentado à disciplina de Trabalho de Conclusão de Curso 2, do Curso de Bacharelado em Ciência da Computação do Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Frank Helbert Borsato

**CAMPO MOURÃO
2017**



ATA DA DEFESA DO PROJETO DE TCC

Às **14:00** do dia **14 de maio de 2017** foi realizada na sala **E103** da UTFPR-CM a sessão pública da defesa do projeto de Trabalho de Conclusão do Curso de Bacharelado em Ciência da Computação do(a) acadêmico(a) **Humberto Moreira Goncalves**. Estavam presentes, além do(a) acadêmico(a), os membros da banca examinadora composta por: **Prof. Dr. Frank Helbert Borsato** (orientador), **Prof. Dr. Diego Bertolini Gonçalves** e **Prof. Dr. Lucio Geronimo Valentin**. Inicialmente, o(a) acadêmico(a) fez a apresentação do seu trabalho, sendo, em seguida, arguido(a) pela banca examinadora. Após as arguições, sem a presença do(a) acadêmico(a), a banca examinadora o(a) considerou _____ na disciplina de Trabalho de Conclusão de Curso **2** e atribuiu, em consenso, a nota _____ (_____). Este resultado foi comunicado ao(à) acadêmico(a) e aos presentes na sessão pública e, posteriormente, deverá ser registrado no sistema acadêmico pelo professor responsável de TCC. Em seguida foi encerrada a sessão e, para constar, foi lavrada a presente Ata que segue assinada pelos membros da banca examinadora, após lida e considerada conforme.

Observações: _____

Campo Mourão, **14 de maio de 2017**

Prof. Dr. Diego Bertolini Gonçalves
Membro 1

Prof. Dr. Lucio Geronimo Valentin
Membro 2

Prof. Dr. Frank Helbert Borsato
Orientador

Resumo

Gonçalves, Humberto Moreira. Detecção acurada da pupila mediante oclusões por cílios com suporte temporal estendido. 2017. 40. f. Monografia (Curso de Bacharelado em Ciência da Computação), Universidade Tecnológica Federal do Paraná. Campo Mourão, 2017.

Contexto: Um rastreador de olhar é um dispositivo que estima a posição do olho, e através de um processo de calibração, é capaz de estimar também o ponto focal. Diversas tecnologias foram propostas para o rastreamento, incluindo lentes especiais e eletrodos acoplados ao redor do olho. Dentre as tecnologias, o uso de vídeo se destaca pela facilidade de uso e conforto. Em um rastreador de olhar baseado em vídeo, características físicas do olho são tipicamente exploradas, como a borda da pupila e reflexos presentes na superfície da córnea. Rastreadores comerciais ainda são pouco acessíveis (valor superior a R\$3000,00), o que motivou diversas iniciativas de baixo custo. Tais iniciativas no entanto, possuem limitações quanto a velocidade e acurácia, geralmente resultado do emprego de webcams e alto processamento demandado. Tipicamente, o processamento é restrito a uma região de interesse (RI) definida ao redor da posição esperada da pupila na imagem. Para otimizar tal região, trabalhos recentes exploram o movimento da pupila no que é chamado de suporte temporal estendido (STE). Os cílios e pálpebras podem afetar a acurácia, visto que podem ocluir parte da pupila. Como esta é geralmente utilizada no mapeamento para o ponto focal, um método robusto a oclusões é altamente desejável.

Objetivo: O trabalho visa estudar uma abordagem de detecção da pupila mediante oclusões por cílios com STE de forma a melhorar a acurácia na detecção da pupila.

Método: Estudo e implementação da nova abordagem de detecção da pupila com STE. O STE proposto se utiliza de um modelo exponencial do movimento do olho. Comparação da pupila estimada no método proposto com as pupilas rotuladas por Świrski et al. (2012). Como métrica na comparação da acurácia se utiliza a distância Hausdorff entre as elipses.

Resultados: Foi obtida uma melhor acurácia na detecção da pupila. Houve uma redução no tempo de processamento com a redução da região processada. A predição dos movimentos dos olhos usando o modelo exponencial mostrou-se viável somente a altas taxas de quadros.

Conclusões: Houve uma melhora na acurácia na presença de oclusão por cílios e redução no esforço computacional necessário. Entretanto, a redução na RI deve vir acompanhada de uma previsão acurada na posição da pupila. Durante sacadas, estimar o movimento do olho

usando um modelo exponencial mostrou-se viável somente a altas taxas de quadros. Obteve-se uma redução no tempo de processamento ocasionado pela redução na área processada, proporcionando também a redução do impacto negativo das oclusões por cílios.

Palavras-chaves: Detecção de pupila. Estimativa de sacadas. Rastreamento de olhar.

Abstract

Gonçalves, Humberto Moreira. . 2017. 40. f. Monograph (Undergraduate Program in Computer Science), Federal University of Technology – Paraná. Campo Mourão, PR, Brazil, 2017.

Context: An eye-tracker is a device that estimates the eye position, and using a calibration process, is also able to estimate the point of gaze. Several techniques have been proposed for tracking de eyes, including special contact lenses and electrodes coupled around the eyes. Among the techniques, the use of video stands out for its ease of use and comfort. In a video-based eye-tracker, biological features of the eye are typically exploited, such as the pupil border and reflexes on the surface of the cornea. Commercial devices are still expensive (worth over R\$ 3000.00), which has motivated several low-cost initiatives. However, such proposals have limited speed and accuracy, usually result of the use of off-the-shelf cameras and the high processing requirements. Typically, the processing is within a region of interest (ROI) defined around the expected pupil position. To improve such region, recent works exploit the pupil movement in what is called extended temporal support (ETS). Eyelashes and eyelids can also affect the accuracy as they may occlude part of the pupil. As the pupil center is generally used in the gaze mapping, a robust occlusion handling method is highly desirable.

Objective: This work aims to study a pupil detection approach with the presence of eyelid occlusions exploiting an ETS to improve the pupil detection accuracy.

Method: Study and implementation of a new approach to pupil detection with ETS. The proposed ETS exploits an exponential eye movement model. Comparison of the accuracy of the proposed method with annotated pupils from Świrski et al. (2012). The metric used in the comparison of accuracy is the Hausdorff distance between ellipses.

Results: The accuracy of the pupil detection was improved. There was a reduction in processing time with the reduction of the processed region. The performance of the prediction of the eye movements using an exponential model showed feasible only at high frame rates.

Conclusions: The accuracy was improved in the presence of eyelash occlusions and the computational effort per frame was reduced. However, the reduction in the ROI must follow an accurate prediction of the pupil position. During saccades, estimating eye movements using an exponential model showed feasible only at high frame rates. There was a reduction

in the processing time caused by the reduction in the processed area, also providing a reduction in the negative impact of eyelash occlusions.

Keywords: Pupil detection. Saccade estimation. Eye tracking.

Lista de figuras

2.1	Estrutura do olho humano simplificado em corte horizontal	13
2.2	Perfil ideal da velocidade em razão do tempo	14
2.3	Tobii Pro Glasses 2	15
2.4	Deteção da pupila com a técnica da pupila clara e escura	16
3.1	Fluxograma de etapas de estimação da pupila.	17
3.2	Representação simplista da implementação da abordagem proposta	19
3.3	Comparação do traçamento de raios em diferentes abordagens	22
3.4	Pupila oclusão por cílios	22
3.5	Exemplo de uso da fórmula da trajetória da sacada	24
4.1	Estimação da borda da pupila	28
5.1	Comportamento do raios com diferentes tamanhos	31
5.2	Desempenho de número de raios em relação ao tempo de processamento	33
5.3	Desempenho da abordagem com 12 raios em relação ao tempo de processamento	33
5.4	Desempenho da abordagem com 36 raios em relação ao tempo de processamento	34
5.5	Desempenho da abordagem com 72 raios em relação ao tempo de processamento	34
5.6	Trajetória da sacada com diferentes quantidades de pontos amostrados	35
5.7	Desempenho do algoritmo <i>levmar</i>	36

Lista de tabelas

5.1	Tamanho do raio e distância Hausdorff variando o número de raios	32
-----	--	----

Sumário

1	Introdução	8
1.1	Objetivos	10
1.2	Estrutura do trabalho	11
2	Referencial Teórico	12
2.1	Estrutura do olho humano	12
2.2	Rastreadores de olhar	14
3	Detector de pupila proposto	17
3.1	Estimativa da região candidata à borda da pupila	18
3.2	Refinamento da borda pupila	19
3.2.1	Traçamento de raios	20
3.2.2	Oclusão	22
3.3	Suporte temporal estendido (STE)	23
3.3.1	Trajectoria da sacada	23
4	Metodologia	26
4.1	Escolha da base de dados de imagens	26
4.2	Estimativa da borda da pupila	27
4.3	Escolha da base de dados para o suporte temporal estendido	29
4.4	Métricas de comparação de desempenho	29
5	Resultados	30
5.1	Avaliação da pupila estimada	30
5.2	Análise do tempo de processamento	32
5.3	Análise da trajetória da sacada	34
6	Conclusão	37
	Referências	39

Introdução

O objetivo de um rastreador de olhar é estimar a direção em que o usuário está olhando, chamado neste trabalho de ponto focal, através de um processo de calibração. Diversas tecnologias foram criadas ao longo dos anos para o rastreamento do olhar, como lentes de contato e eletrodos em contato direto com a pele ao redor dos olhos (HOLMQVIST et al., 2011). No entanto, o desconforto gerado para o usuário foi uma barreira para o uso de tais dispositivos (MORIMOTO; MIMICA, 2005). Dentre as tecnologias utilizadas, o rastreamento de olhar explorando imagens de vídeo se torna de notável interesse devido a sua usabilidade e conforto. Em um rastreador de olhar baseado em vídeo, o uso das características biológicas do olho, como a borda da pupila e reflexos presentes na córnea são comumente utilizados (OHNO et al., 2002; LI; PARKHURST, 2005).

A relação custo \times desempenho dos rastreadores ainda é um desafio. Rastreadores de olhar comerciais ainda são pouco utilizados devido ao seu custo. Apesar dos avanços nos últimos anos que melhoraram aspectos como usabilidade e desempenho, os rastreadores comerciais ainda são caros para alguns segmentos. Neste contexto, foi proposto como alternativa para baratear o custo de rastreadores de olhar a utilização de câmeras de computador (*webcams*). Tais câmeras geralmente possuem baixa taxa de captura de quadros, e as que possuem altas taxas de quadros apresentam baixa resolução de imagem. Uma baixa resolução pode afetar a acurácia de um rastreador. Uma alta taxa de quadros utiliza mais recursos do sistema computacional, que precisa processar mais quadros em um mesmo intervalo de tempo. Para reduzir o custo computacional, define-se uma região de interesse na imagem (BORSATO, 2016). Essa região é definida a partir de uma aproximação da região da pupila estimada no quadro anterior.

Oclusão é também um desafio encontrado em rastreadores de olhar baseados em vídeo. Como geralmente a estimativa da pupila é um ponto chave no mapeamento do ponto focal, oclusões podem afetar a acurácia do rastreador, visto que a pupila pode ser erroneamente

estimada. Visando melhorar a acurácia do rastreamento de olhar, procurar alternativas para minimizar o impacto de oclusões é desejável.

Entre os rastreadores baseados em vídeo, destaca-se o uso de rastreadores de olhar montados na cabeça (do inglês, *head-mounted*). Neles, a câmera é montada em um dispositivo que é fixado a cabeça através de óculos, bandanas e até capacetes. Além da câmera, geralmente é utilizada uma fonte de luz próxima ao olho responsável por iluminá-lo. Tal fonte também produz um reflexo na córnea que pode ser usado para reduzir o impacto de movimentos da câmera em relação ao olho.

Para obter o ponto focal a partir da imagem, é necessário um processo de calibração que é repetido toda vez que o dispositivo é usado. Durante a calibração, diversos alvos são exibidos ao usuário que é instruído a observá-los, o centro da pupila é estimado e relacionado com o ponto sendo observado através do ajuste dos parâmetros de um polinômio. Para estimar o centro da pupila, pode ser utilizado o centro de massa na imagem (AGUSTIN et al., 2010) ou estimar uma elipse que aproxime a borda da pupila, e então obter seu centro (LI; PARKHURST, 2005). A diferença entre o centro da pupila e o reflexo na córnea pode ser utilizado para substituir o centro da pupila nos cálculos a fim de obter uma maior robustez a movimentos entre o olho e a câmera. Posteriormente ao processo de calibração, o reflexo na córnea e o centro da pupila são utilizados no mapeamento para o ponto focal. Portanto, a correta estimativa da pupila é determinante no desempenho do rastreador, sendo necessário realizá-lo em cada quadro do vídeo. Assim surgiram alternativas para minimizar o custo desse processo, um deles é o ajuste da região de interesse na imagem utilizando a informação da pupila de quadros anteriores para obter uma região aproximada, que é então considerada a região candidata para estimar a pupila no quadro atual (LI; PARKHURST, 2005).

Abordagens recentes utilizam a relação de características biológicas do olho, geralmente movimentos dos olhos, obtidas nos quadros anteriores para ajustar a região de interesse. Por exemplo, o método proposto por Li e Parkhurst (2005) (Starburst) utiliza o centro da pupila do quadro anterior como ponto inicial na busca da pupila do quadro atual, assumindo que tal ponto está dentro da pupila atual na imagem. Tal método assume um limite superior para o movimento da pupila entre quadros adjacentes. Isto pode ser chamado de suporte temporal, e depende da taxa de quadros da câmera. Borsato (2016) propõem utilizar informações de quadros mais distantes no tempo para otimizar a região de interesse e assim reduzir o custo computacional na detecção da pupila, no que chama de suporte temporal estendido.

Visando fornecer uma alternativa para reduzir o impacto negativo da oclusão por cílios, foi limitado a detecção da pupila apenas em uma região de interesse. O intuito de usar uma região de interesse é atrativo para uma possível redução em tempo de processamento e reduzir o erro para estimar a pupila. Para estipular uma região de interesse é necessário saber em que local na imagem esta a pupila. É desejado que essa região de interesse seja reduzida

e limitada a região da borda da pupila corretamente. Sendo necessário saber o quão preciso é essa região, ou seja, o quão perto esta da borda da pupila, foi utilizado neste trabalho o suporte temporal estendido, que utiliza informações e características da pupila estimada em quadros anteriores.

A redução da região candidata através do aumento do suporte temporal dispõe uma possibilidade para rastreadores de olhar aumentarem a taxa de quadros, em razão da redução nos dados a serem processados.

1.1. Objetivos

O objetivo do trabalho é propor um método de estimar a borda da pupila, baseado na análise da movimentação do olho e suas características. O método envolve estipular uma região candidata na imagem, que obrigatoriamente deve conter a borda da pupila. Pretende identificar sacadas, estimar seus parâmetros, e assim reduzir essa região. Em vista que a identificação das sacadas tem como propósito aumentar o grau de certeza da posição da pupila na imagem.

Sacadas são movimentos rápidos do olhos, tornando um desafio para rastreadores com câmeras com baixas taxas de quadros lidarem. Porém, rastreadores que utilizam câmeras com altas taxas de quadros tendem a lidar melhor com o desafio de capturarem eventos de movimentação pequenos.

Desta forma para isso ser viável, ou seja, o emprego do suporte temporal estendido é necessário que haja uma taxa de quadros adequada. Não foi estipulada uma taxa de quadros ideal para análise desses movimentos dos olhos. Se essa taxa de quadros for alta, é necessário a redução no tempo de processamento de cada quadro, para que possibilite a sua execução em tempo hábil. A redução de tempo de processamento neste caso pela redução da região candidata a borda da pupila, proporcionando uma redução na região possível para estimação de cada pupila.

Deseja-se que a região esteja mais aproximada da borda da pupila provendo uma minimização de espaço a ser processado. Como a região está mais próxima a borda da pupila tende-se a reduzir a interferência negativa das oclusões por cílios na acurácia da estimação da borda.

Portanto, é possível diminuir a região candidata com um maior grau de certeza, contribuindo para a redução no tempo de processamento de cada quadro. Auxiliando a redução no tempo de processamento da região e possibilitando o uso de câmeras com uma alta taxa de quadros.

Os objetivos deste trabalho são:

1. **Diminuir o esforço computacional necessário para estimar a pupila em cada quadro:** a motivação para isso é que possibilite a identificação de movimentos do olho

e estimule o uso de câmera com maiores taxas de quadro.

2. **Reduzir o efeito negativo de oclusões ocasionadas por cílios sobre a estimativa da pupila:** Oclusão ainda é um desafio enfrentado devido a seu impacto nos algoritmos estudados, apresentar uma alternativa para evitar que sua presença altere negativamente a acurácia de rastreadores

Com a redução no tempo de processamento objetivado pelo trabalho, é possível que utilizar uma câmera com uma maior taxa de quadros por segundo. Atualmente rastreadores de olhar de baixo custo geralmente não operam com uma alta taxa de quadros. Também, os algoritmos empregados são desafiados pelas oclusões e baixa taxa de quadros presentes em câmeras de baixo custo dificultando a estimação da pupila acuradamente, movimentos pequenos nos olhos são difíceis de serem previstos com antecedência ou no momento em que estão ocorrendo. Portanto alguns desafios encontrados são:

1. Estimar a pupila acuradamente com presença de oclusões por cílios;
2. Detectar movimentos dos olhos com baixas taxas de quadros.

Visando fornecer uma alternativa para a taxa de processamento dos rastreadores de baixo custo e conseqüentemente aumentando a taxa de quadros, a região candidata a pupila foi limitada a região próxima a pupila. Foi produzida uma nova abordagem robusta a oclusões por cílios minimizando o impacto negativo. Os resultados foram mensurados através de o quão distante está a pupila estimada na nova abordagem para a pupila apresentada pela base de dados disponibilizado pelo método proposto em Świrski et al. (2012), que será citado como Swirski para facilitar a diferenciação.

1.2. Estrutura do trabalho

O restante do trabalho possui a seguinte estrutura: No Capítulo 2 apresenta-se o referencial teórico, relacionando características do olho e seus movimentos, técnicas utilizadas e medidas de estimação da pupila. No capítulo Capítulo 3 é descrita a proposta do trabalho, apresentando a abordagem que foi desenvolvida e seu funcionamento. O Capítulo 4 descreve a metodologia empregada neste trabalho. O Capítulo 5 apresenta os experimentos e suas análises. E por fim, Capítulo 6 apresenta o fechamento do trabalho e as conclusões obtidas ao longo deste.

Referencial Teórico

Neste capítulo é apresentado o referencial teórico que embasa a pesquisa realizada. Descrevendo sobre a estrutura do olho, suas características biológicas e como o olho se movimenta para focar diferentes objetos. Explicando também um pouco das tecnologias, estruturas e algoritmos empregadas em rastreadores de olhar baseados em imagens de vídeo.

2.1. Estrutura do olho humano

A Figura 2.1 apresenta um corte horizontal da estrutura do olho humano simplificado. O olho humano tem formato aproximadamente esférico e possui um diâmetro de cerca de 20mm. Três membranas cercam o olho: a córnea e a esclerótica; a coroide; e a retina. A córnea é um tecido duro e transparente que cobre a camada anterior do olho, e a esclerótica é uma membrana opaca que envolve o globo ocular. A coroide está abaixo da esclerótica, contém uma rede de vasos sanguíneos responsáveis pela nutrição do olho. Sendo fortemente pigmentado ajudando a reduzir a quantidade de luz que entra no olho. É dividida por corpo ciliar e íris. A íris contrai ou expande para controlar a quantidade de luz que entra no olho. A abertura central, também chamada de pupila, varia de tamanho, entre 2 a 8mm. A retina é a membrana mais interna, situada na parte posterior, composta por células sensíveis a luz (GONZALEZ; WOODS, 2001).

O olho humano é capaz de enxergar um faixa limitada do espectro de ondas eletromagnéticas. A faixa de luz visível humana é em torno de 300 a 700nm. Existem ondas menores que 300nm, denominadas ultravioletas, raios-X e gamas, e maiores que 700 nm, as ondas infravermelhas, de micro-ondas e de radio (GONZALEZ; WOODS, 2001).

Em rastreadores de olhar, geralmente é usada uma fonte de luz infravermelha de baixa intensidade próxima a câmera do rastreador. Quando a luz é direcionada ao olho, parte dela é refletida na córnea, gerando um reflexo que pode ser capturado por câmeras, podendo

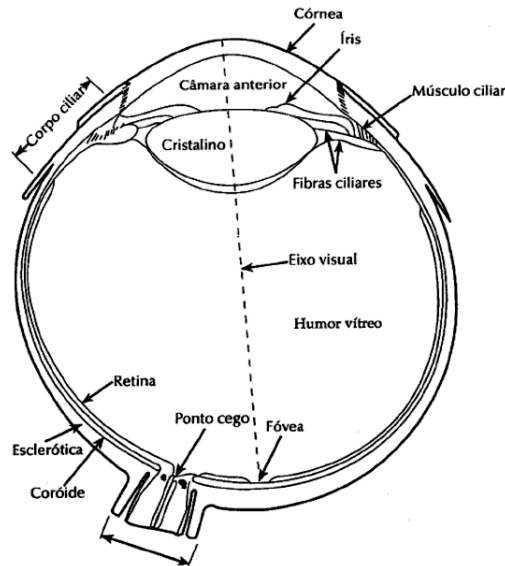


Figura 2.1. Estrutura do olho humano simplificado em corte horizontal.
Fonte: adaptado de Gonzalez e Woods (2001).

ser utilizada na calibração do rastreador, conforme apresentado em Seção 2.2 (MORIMOTO; MIMICA, 2005).

O olho realiza vários tipos de movimentos para focar um ponto de interesse. Durante a leitura, ou quando se procura por um objeto, o olho executa um movimento chamado de sacada. Entre as sacadas, o olho humano permanece relativamente imóvel num período chamado de fixação. Sacadas são movimentos rápidos do olho, chegando a $500^\circ/s$. Além das sacadas, o olho realiza três outros movimentos: perseguição (do inglês, *pursuit*), convergência (do inglês, *vergence*), e vestibular (do inglês, *vestibular*). Os movimentos de perseguição ocorrem quando os olhos seguem um objeto ou alvo. Convergência ocorre quando os olhos se movem para focar objetos a diferentes distâncias. O movimento vestibular ocorre quando os olhos rotacionam para compensar os movimentos do corpo e cabeça e manter a mesma direção de visão (RAYNER, 1998).

O termo sequência principal (do inglês, *main sequence*) é emprestado da astronomia, e explica a relação entre o brilho e a temperatura de uma estrela. No contexto de sacada do olho, sequência principal refere-se à relação entre a duração e magnitude, e entre o pico de velocidade e a sua magnitude. A duração da sacada, a velocidade média, e o pico de velocidade aumentam conforme a magnitude da sacada aumenta. Um pico de velocidade pode ser mensurado, entretanto a duração da sacada apresenta problemas para ser medida. Algumas vezes, há dificuldade de se determinar o exato começo ou término de uma sacada. Uma forma de determinar a duração da sacada é usar a velocidade com que o olho se move em razão do tempo (BUSWELL, 1935).

A velocidade da sacada cresce rapidamente até atingir a velocidade máxima, também conhecido como pico da velocidade, que ocorre antes do ponto médio do movimento, e depois cai a uma taxa ligeiramente mais lenta, até a localização do alvo ser atingida. Exemplos da

função da sacada é apresentado na Figura 2.2. Quanto maior a amplitude da sacada, maior é a velocidade e a duração. A duração da sacada, por exemplo com 5° tem duração de 20ms, já uma de 10° tem duração de 25ms e uma 20° de 30 a 40ms (HARRIS; WOLPERT, 2006).

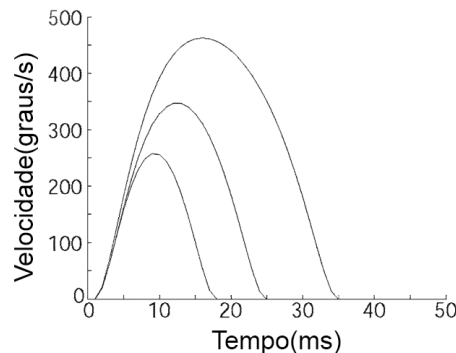


Figura 2.2. Perfil ideal da velocidade (graus por segundo) da sacada em razão do tempo (milissegundos) para 5° , 10° e 20° .

Fonte: Harris e Wolpert (2006).

A pupila realiza movimentos de constrição e dilatação para se ajustar a intensidade de luz presente no ambiente. Ellis (1981) realizou um estudo em um ambiente escuro, com o auxílio de uma fonte de luz infravermelha de baixa intensidade e invisível ao olho humano simulou a mudança de luz presente e observou-se o quanto a pupila ampliava e contraía na presença de luz. Assim, foi estipulado a velocidade máxima de 8mm/s para constrição e 3mm/s para dilatação da pupila.

2.2. Rastreadores de olhar

Os primeiros dispositivos construídos com o intuito de estimar os movimentos dos olhos são datados no século 19, sendo mecânicos e altamente invasivos. Apenas no começo do século 20, o princípio fotográfico foi introduzido, tornando-se a técnica dominante para rastrear olhar (HOLMQVIST et al., 2011). Apesar das descobertas que eles trouxeram, ainda eram grandes e complexos levando o seu uso ficar restrito apenas em laboratório (YARBUS, 1967; WADE, 2010; BUSWELL, 1935).

Nas últimas décadas o cenário mudou consideravelmente. O rastreamento de olhar tonou-se mais acessível, aumentando sua popularidade entre pesquisadores de muitas áreas, como analistas de usabilidade, cientistas de esportes, neurofisiologistas, desenvolvedores de jogos, entre outros (DUCHOWSKI, 2002).

Embora diversas técnicas foram investigadas para estimar os movimentos dos olhos ao longo dos anos, tal como eletro-oculografia e lentes de contato especiais (YOUNG; SHEENA, 1975), a video-oculografia se destaca como a técnica dominante, sendo utilizada em pesquisa e aplicações comerciais. A vídeo-oculografia é uma técnica que permite estimar os movimentos

oculares por meio de vídeo (KIM et al., 2006). Os dispositivos que se utilizam desta técnica são chamados de rastreadores de olhar (MORIMOTO; MIMICA, 2005).

Um rastreador de olhar baseado em vídeo é composto por uma câmera de vídeo que capta as imagens do olho e um computador que processa as imagens e estima os movimentos, armazena e analisa os dados. Rastreadores de olhar baseados em vídeo podem ser usados em uma configuração remota ou montados na cabeça. Em um sistema remoto, a câmera é fixada geralmente abaixo ou perto do monitor do computador a ser utilizado no experimento. Em uma configuração de rastreador montado na cabeça, a câmera é fixada em óculos ou capacetes. O trabalho focou em rastreador de olhar baseado em vídeo montado na cabeça, um exemplo de é apresentado na Figura 2.3, um rastreador de olhar comercial chamado *Tobii Pro Glasses 2* que atinge taxas de 30, 60 e 120 quadros por segundos.



Figura 2.3. Exemplo de rastreador de olhar baseado em vídeo, Tobii Pro Glasses 2

Algoritmos de rastreamento de olhar são divididos em duas abordagens: baseadas em características (do inglês, *feature-based*) e baseadas em modelos (do inglês, *model-based*). As abordagens baseadas em modelos procuram encontrar a melhor modelagem matemática para o problema, por exemplo, um modelo que se ajuste a borda da pupila. Por outro lado, as abordagens baseadas em características detectam e exploram propriedades na imagem do olho, como a pupila parecer com uma região escura e elíptica e a íris parecer mais clara. Nas abordagens baseadas em características, há variações entre os algoritmos estudados, embora sejam diferentes, a utilização de gradientes de intensidade são empregados para localizar partes dos olhos, como reflexão na córnea e pupila (OHNO et al., 2002) e a fronteira entre íris e esclera, conhecida como limbo (ZHU; JI, 2005)

Há três abordagens para iluminação do olho, com a luz ambiente, com uma fonte de luz e com duas fontes de luz num arranjo estruturado. A pupila com a iluminação ambiente deve se apresentar como uma pupila escura e sem muito contraste perante a íris. Com uma fonte de luz é utilizada para a câmera ser capaz de capturar a reflexão causada na córnea, geralmente é utilizada a reflexão na córnea no processo de calibração, apresentada na Figura 2.4a, também chamada de pupila escura. E o arranjo estruturado de duas fontes de luz é utilizado para produzir uma pupila clara (Figura 2.4b) e outra escura, assim através da

operação de subtração das imagens é provável que encontrem uma região que interseccionem ambas Figura 2.4 (EBISAWA; SATOH, 1993).

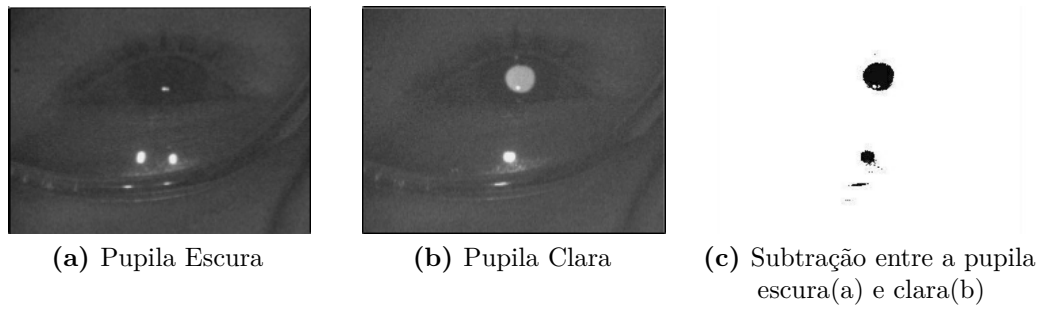


Figura 2.4. Representação de pupila escura, clara e a operação de diferença entre ambas imagens.
Fonte: Morimoto e Mimica (2005)

Detector de pupila proposto

Dado imagens do olho capturas por uma câmera montada na cabeça ou por um rastreador de olhar baseado em vídeo na montado na cabeça, a proposta é detectar a pupila acuradamente mediante oclusões. Para tal, foi estendido o método proposto por Ohno et al. (2002) que é um método baseado em características do olho. Consiste em estimar a região candidata à borda da pupila após um quadro estiver pronto para ser processado e o refinamento da borda. A primeira é iniciada com a aquisição de um novo quadro e sempre que houver uma região candidata é realizado o refinamento, assim que após a estimação da borda da pupila é reiniciado o processo até que não haja mais quadros para serem processados. A Figura 3.1 mostra um diagrama simplificado das etapas.

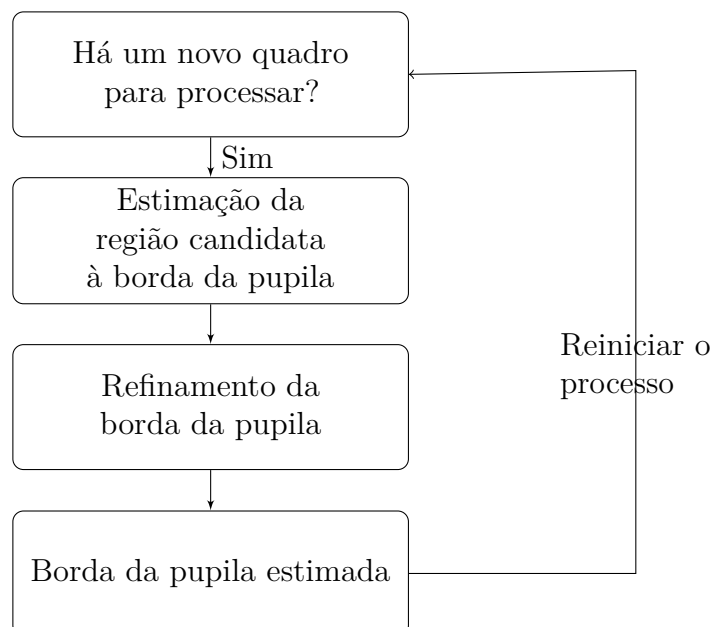


Figura 3.1. Fluxograma de etapas de estimação da pupila.

A estimação da região candidata à borda da pupila é descrita na Seção 3.1. A Seção 3.2 define o método de refinamento da pupila, a Seção 3.3 define a utilização do STE.

3.1. Estimativa da região candidata à borda da pupila

A região candidata à borda da pupila é a região de interesse na imagem do olho. Espera-se que a pupila esteja dentro dessa região em qualquer quadro. Alguns trabalhos não a definem como uma região, apenas um ponto a partir do qual o refinamento começa, não limitando a região de busca (OHNO et al., 2002; LI; PARKHURST, 2005). Trabalhos recentes definem uma região que restringe o processo de refinamento à essa região, que corrobora com a redução no custo computacional (BORSATO, 2016).

Ohno et al. (2002) apresentam um método que procura uma região escura e pequena na imagem, características esperadas da pupila. Primeiramente a imagem é capturada pelo sensor com uma resolução de 640×480 . Para reduzir o ruído e o custo de processamento, a resolução da imagem é reduzida para $1/4$ da imagem original. Então a imagem reduzida é utilizada no passo de limiarização, cujo intuito é detectar pequenas regiões escuras, como a pupila. A limiarização consiste em gerar regiões onde a vizinhança dos pixels são similares. Para cada região o contorno é avaliado para detectar se tal região é uma região candidata à pupila ou não. Para tal propósito, os valores da vizinhança dos pixels de cada região são usados para distinguir entre pupila e outras regiões, como narinas. Assim que a região da pupila é localizada através dos passos de limiarização, uma aproximação inicial do centro da pupila é realizada para então iniciar o refinamento.

Starbust utiliza um ponto como aproximação do centro da pupila para estimação de sua borda. O processo de refinamento começa a partir desse ponto em direção a borda da imagem, com intuito de encontrar a borda da pupila. Não é estipulado uma região máxima de procura da pupila, que pode prosseguir até a borda da imagem caso um determinado limiar não seja atingido durante a busca (LI; PARKHURST, 2005).

Uma abordagem diferente é utilizada em (BORSATO, 2016). Utilizando luz estruturada para gerar tanto pupilas claras quanto escuras, o método explora a diferença entre quadros adjacentes para obter informações sobre a posição atual da pupila. A diferença entre pupila clara e escura definem uma região de interesse proporcional a interseção da pupila em quadros adjacentes. O refinamento começa a partir da região retangular assim definida.

No detector proposto, a região de interesse não será definida como um ponto ou como uma região retangular, mas tomará a forma da pupila. A região será formada por duas elipses com os seguintes requisitos: a elipse interna (e_{in}) deve obrigatoriamente estar contida na pupila do quadro atual, enquanto a elipse externa (e_{out}) deve obrigatoriamente conter toda a pupila do quadro, representado por Figura 3.2. Quanto mais se aproximar a elipse interna

da externa sem comprometer os requisitos, menos pixels serão processados, resultando em menos esforço computacional. Note que nesta abordagem é utilizado apenas um tipo de luz, caracterizando um pupila escura na imagem. A aproximação simples é dada por conhecer a

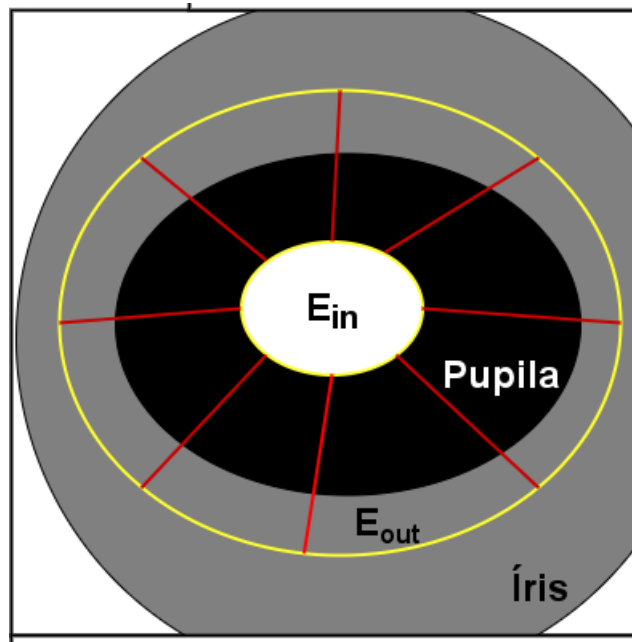


Figura 3.2. Representação simplista da implementação da abordagem proposta, os segmentos de reta em vermelho são chamados de raios para facilitar sua denominação, a pupila é representada por uma elipse preta na imagem, a íris é representada por uma elipse cinza grande e maior que o corte feito na imagem, a elipse interna (e_{in}) e externa (e_{out}) em amarelo. Os raios são limitados e contidos entre e_{in} e e_{out} .

velocidade máxima da pupila é possível otimizar a região candidata, limitando a sua possível região máxima de busca. Também é possível reduzir a região através do monitoramento da posição da pupila ao longo do tempo, como será descrito na Seção 3.3.

3.2. Refinamento da borda pupila

A região candidata a borda da pupila definida pelo processo descrito na Seção 3.1 é ainda uma descrição muito grosseira da borda. Esta região deve ser refinada para que melhor represente a borda da pupila na imagem. Uma das técnicas utilizadas é o traçamento de raios (OHNO et al., 2002; LI; PARKHURST, 2005). O objetivo é traçar raios sobre a região de interesse, e ao longo destes identificar rampas de intensidade dos pixels. Utilizar rampas de intensidade é possível devido ao fato da pupila ser mais escura que a íris quando capturada sob iluminação infravermelho. Candidatos a borda da pupila ao longo dos raios podem ser definidos por diversos métodos, geralmente associados a um limiar pré-estabelecido ou valores extremos na rampa de intensidades (OHNO et al., 2002; LI; PARKHURST, 2005).

Uma vez que todos os candidatos estão definidos, uma elipse é ajustada a eles, definindo o contorno da pupila. Duas alternativas são comumente empregadas para tal, o

método de mínimos quadrados ou *Random Sample Concensus* (RANSAC). No primeiro, todos os pontos são considerados, e pontos que não pertencem de fato a borda da elipse (pupila) tem um efeito negativo bastante pronunciado no resultado. RANSAC é um método estatístico que seleciona pontos aleatoriamente para o ajuste, seguido pelo cálculo do suporte dado pelos pontos restantes, separando os candidatos em *inliers* que é o conjunto de candidatos que pertence a borda da elipse, e *outliers* que é o conjunto de candidatos que possui valores discrepantes da borda da elipse. O método maximiza a fração de *inliers*.

3.2.1. Traçamento de raios

O traçamento de raios começa em uma região ou em um pixel interno a pupila. Essa região ou pixel tipicamente é definido manualmente. Os raios são traçados em múltiplas direções em sentido a borda da pupila. Tais raios são traçados para que ultrapassem o limite entre da borda da pupila e íris. O intuito é analisar as posições onde há rampas de intensidades. As rampas de intensidades identificadas ao longo do raio serão consideradas pontos para compor o conjunto de pontos candidatos à borda da pupila. As rampas de intensidade exemplificam justamente a transição entre a borda da pupila e a íris, geralmente a pupila é escura e íris mais clara principalmente na técnica de pupila escura.

Ohno et al. (2002) introduz a procura da borda da pupila pelo traçamento de raios. O processo é iniciado de um ponto interno na pupila na imagem inicialmente. É feito um processo para eliminar ruídos como cílios, pálpebras e o reflexo na córnea. Posteriormente, é feito o traçamento de raios a partir do centro estipulado, é definido um limiar de detecção de borda. Esse limiar é calculado sobre o brilho médio da área candidata, e os pixels mais brilhantes são adicionados ao conjunto de candidatos à borda da pupila. Se houver uma estimativa de pupila já feita no quadro anterior é utilizado como aproximação do centro do novo quadro, representado por Figura 3.3a.

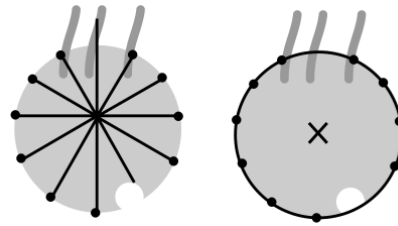
Starbust também utiliza o traçamento de raios para estimar a pupila. Em todo quadro usa-se uma aproximação do centro da pupila para começar a traçar raios. Para o primeiro quadro, é escolhido um ponto interno na pupila e para os quadros seguintes é usado o centro da pupila estimado no quadro anterior. Para escolher pontos candidatos a borda da pupila é utilizado um limiar que pretende localizar a borda ao traçar os raios. Esse limiar é usado para a verificação da diferença entre um pixel e outro adjacente iterativamente ao longo do raio. O pixel que exceder o limiar é considerado como pixel candidato a borda da pupila e assim é feito com todos os raios. Após os candidatos serem selecionados é preciso refinar a borda. É feito um processo de detecção de candidatos iterativamente. Nesse processo é desejável evitar que raios sejam traçados desnecessários e se aproximar o máximo da borda da pupila, representado por Figura 3.3b. Assim, há o processo de seleção de candidatos que é iterado até que a borda estimada da pupila não se altere o suficiente entre uma iteração e outra no mesmo quadro, é esperado encontrar vários pontos candidatos à

borda da pupila e o número de candidatos a borda da pupila pode ser elevado. Portanto é classificado cada candidato em *inliers* e *outliers*. O conjunto de *inliers* são os pontos que se comportam adequadamente como a borda da pupila e o restante são considerados *outliers*. É utilizado RANSAC para avaliar probabilisticamente e aleatoriamente a melhor pupila a cada subconjunto de 5 candidatos (LI; PARKHURST, 2005).

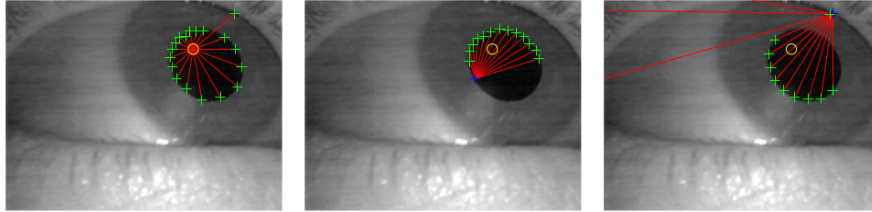
Borsato (2016) apresenta um método de traçamento de raios de uma região interna na pupila. Tal região é obtida através da interseção da pupila em quadros adjacentes, é assumido que sempre há uma região que a pupila de dois quadros adjacentes interseccionam. O refinamento começa da melhor aproximação da borda da pupila. Quanto maior a região interseccionada, maior será a segurança devido a chance da borda da pupila estar próxima da região interseccionada e conseqüentemente menor é o tamanho do raio justamente porque a borda da pupila tende a estar próxima a região. E o inverso também é válido, quanto menor é a região interseccionada, menor é a segurança e maior chance da borda estar distante da região e conseqüentemente maior vai ser o tamanho do raio. Tal processo de ampliar e reduzir o tamanho do raio é feito para diminuir o custo computacional, produzindo raios com tamanhos adequados e buscando diferenças de intensidades ao longo dos raios. Ao longo do raio é feita uma convolução de derivada da gaussiana, para que haja um pico de intensidade correspondente a mudança de intensidade de pixels da pupila e íris. É introduzido um processo para descartar os raios que apresentam vários picos de intensidade, ou seja, aqueles que houveram alguma interferência de oclusão, ruído e entre outros. Assim todos os pontos candidatos são utilizados na estimação da melhor elipse estimada, representado por Figura 3.3c.

Estendendo o método traçamento de raios (OHNO et al., 2002) e limitação do tamanho do raio e da região interna na pupila (BORSATO, 2016) é desejado encontrar um intervalo otimizado entre e_{in} e e_{out} . Então o traçamento de raios estará restrito a região de interesse entre as duas elipses, e_{in} e e_{out} , denominado s , que deve corroborar para reduzir o custo computacional, já que é subconjunto da imagem original e menos pixels para processar. É altamente desejável que s seja reduzido, como descrito na Seção 3.3, mas mantendo o pré-requisito de conter borda da pupila, representado por Figura 3.3d.

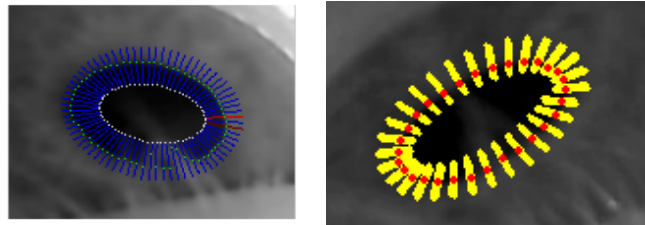
Assim, foi feito o traçamento de raios em s . As intensidades dos pixels ao longo de cada raio formam um vetor que é convolucionado por uma aproximação da derivada da gaussiana. Tal convolução resulta em picos que coincidem com rampas de intensidade nos dados originais. Assume-se que tais rampas coincidem com a região de transição entre a pupila e a íris. Um ponto candidato a borda da pupila é adicionado na resposta mais intensa da convolução, assim como em (BORSATO, 2016). Neste trabalho não haverá uma filtragem de candidatos a borda da pupila, todos os candidatos serão utilizados na região.



(a) Abordagem Ohno et al. (2002)
adaptado



(b) Abordagem Li e Parkhurst (2005) adptado



(c) Abordagem Borsato
(2016) adaptado

(d) Abordagem proposta

Figura 3.3. Comparação do traçamento de raios em diferentes abordagens.

3.2.2. Oclusão

A oclusão por cílios ocorre quando cílios entram na frente da pupila, apresentado na Figura 3.4, alterando a posição ideal de um ponto candidato *inlier* a borda da pupila, geralmente, causando um ou mais pontos *outliers*, valores discrepantes ou distantes dos demais pontos que afetam a acurácia.



Figura 3.4. Pupila ocluída por cílios.
Fonte: Adaptado de Swirski.

Ohno et al. (2002) realiza o refinamento duplo com a estimação da borda e centro da pupila para redução do impacto das oclusões. Li e Parkhurst (2005) classifica *inliers* e *outliers* assumindo que os *inliers* estão posicionados próximos uns aos outros, com um pequeno limiar de erro e qualquer outro ponto candidato não classificado como *inlier* é considerado *outlier*.

Borsato (2016) assume que haverá apenas um pico no raio convolucionado condizente com o efeito da convolução na rampa de intensidade entre pupila e íris, se houver mais de um ponto em qualquer raio, os pontos são classificados como *outliers*.

Na abordagem proposta, a redução do impacto da oclusão por cílios é realizada pela redução do intervalo s entre as duas elipses, apresentado na Seção 3.3. A elipse (borda da pupila) é estimada usando todos os pontos candidatos através do método de quadrados mínimos diretos (FITZGIBBON et al., 1999). Nesse método não há um processo explícito de filtragem de candidatos em *inliers* e *outliers*.

3.3. Suporte temporal estendido (STE)

Para definir e_{in} e e_{out} , pode ser utilizado o Suporte Temporal Estendido (STE) (BORSATO, 2016). Inicialmente deve ser escolhido um ponto interno e uma região retangular externa a pupila semelhante ao STE (BORSATO, 2016), após o início do processo possibilita a sua redução a região de interesse para as duas elipses. Para a otimização e minimização da região de interesse s foi considerado dois fatores biológicos. O reflexo da pupila à luz e os limites superiores na movimentação do olho, que ocorrem durante longas sacadas. A predição da posição da pupila pode ser realizada através da Seção 3.3.1. Entretanto, mesmo com variação das elipses através das taxas limites de reflexo da pupila à luz, as maiores variações de posição da pupila se dá pela ocorrência de sacadas do olho. A relação entre reflexo da pupila à luz e os limites superiores na movimentação do olho em termos de variação da posição da pupila é caracterizada pelo fato de que os limites de movimentação do olho alterarem tão significativamente a posição da pupila na imagem perante as alterações ocasionadas na posição da pupila referente aos reflexos à luz, que os reflexos podem ser desprezados. Portanto, o foco será dado aos limites superiores na movimentação do olho, mais precisamente aos movimentos de sacada do olho.

3.3.1. Trajetória da sacada

O olho realiza sacadas em um curto espaço de tempo, o comportamento da sacada pode ser mensurado através de limiares de velocidade e tempo. A Equação (3.1) demonstra a equação da trajetória de sacadas em relação ao tempo. Onde $p1$, $p2$, $p3$ são três parâmetros que precisam ser determinados pelo ajuste dos dados. Sendo t o tempo discretizado e o $f(t)$ a função da exponencial no instante t (HAN et al., 2013).

$$f(t) = p1 \left[1 - \exp \left[- \left(\frac{t}{p2} \right)^{p3} \right] \right] \quad (3.1)$$

Os três parâmetros podem ser ajustados no modelo associando com diferentes atributos das sacadas. O parâmetro $p1$ corresponde a amplitude da sacada, $p2$ pode ser usado

para reescalar a duração da sacada e p3 é mencionado seu uso na tendência de queda da função. Uma sacada foi considerada com um limiar de velocidade maior que $20^\circ/s$ e duração maior que 20ms.

A sacada tem duração pequena e a fórmula da trajetória de sacada requer pelo menos 3 quadros adjacentes, nomeados Q_n , Q_{n+1} e Q_{n+2} . Um exemplo é a Figura 3.5, que demonstra a amplitude(graus) em razão do tempo(ms).

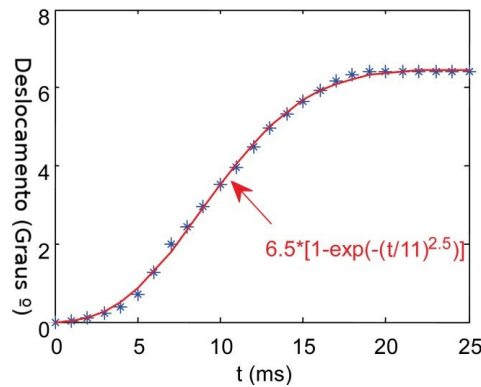


Figura 3.5. Um exemplo do movimento da sacada. Em azul a trajetória da sacada. Em vermelho uma aproximação usando a fórmula da trajetória da sacada. Fonte: Han et al. (2013).

Para a identificação do começo da sacada é preciso de um limiar para a movimentação do olho. É estabelecido o limiar utilizado $20^\circ/s$ e maior que 20ms de duração, o mesmo utilizado no estudo Han et al. (2013). Quando um novo quadro Q_n é processado, é armazenado seu centro após a estimação da borda. O segundo quadro Q_{n+1} utiliza o centro da pupila estimado no Q_n como aproximação para seu centro. Os centros dos dois primeiros quadros são utilizados para formar um vetor de distância entre ambos, fornecendo uma possível posição de centro do Q_{n+2} . Após o terceiro quadro ser estimado, é avaliado se os limiares de velocidade e tempo entre os quadros são excedidos, para iniciar o processo de predição do movimento da sacada. E mesmo que as condições para início do processamento de uma sacada não sejam atendidas, mantêm a análise sobre a movimentação do olho nos próximos quadros até que uma sacada seja identificada.

Para testar a predição da sacada em relação ao tempo foi utilizado dados brutos do *SMI Hi-Speed Eye Tracker*, um rastreador de olhar que opera a 500hz. Utilizando de dados como posição do olho na tela para medir o quanto se moveu, assim foi utilizado o pacote do R *saccades*¹ que estima em qual momento houve uma fixação, ou seja, o momento em que o olho se manteve fixo em um ponto ou região da tela.

Portanto o pacote *saccades* estima as fixações e neste trabalho o interessante são os momentos em que houve movimento dos olhos. Selecionando os intervalos de tempo que não houve fixações é obtido intervalos que o olho estava se movendo. Com os dados extraídos esse momentos e salvo em arquivos para o emprego da fórmula da trajetória da sacada.

¹ <<https://cran.r-project.org/web/packages/saccades/index.html>>

Como os dados foram obtidos em pixels na tela foi realizado uma conversão para posição do olho em ângulo através das Equações (3.2) e (3.3). Sendo *distância* o coeficiente que representa a distância euclidiana em pixels da posição central da tela naquele instante e *distancia monitor* é a distância da pessoa em relação ao monitor, e *ângulo* é o angulo da posição do olho em relação ao centro da tela.

$$distância = \sqrt{(x - centroX)^2 + (y - centroY)^2} \quad (3.2)$$

$$ângulo = \arctan(distância/(distancia monitor)) \quad (3.3)$$

Para estimar a trajetória das sacadas é utilizado o método de quadrados mínimos não-linear de Levenberg-Marquardt, que está disponível na *API* implementada em linguagem de programação C++ conhecida como *levmar* (LOURAKIS, Julho 2004). Utilizou-se o *levmar* para estimar a exponencial que representa cada sacada extraída do *ground-truth* da sacada, usando todos os pontos disponíveis a 500Hz. Entretanto como houve um número excessivo de sacadas para a comparação e além disso, utilizou-se o processo humano de verificação para identificar sacadas adequadas ao contexto. Portanto, foi selecionado 20 sacadas para a comparação, foram utilizadas essas sacadas para estimar as fórmulas exponenciais de trajetória de cada uma, denominadas de f_1 até f_k , sendo que 1 e k o número de cada sacada, e o intervalo de 20 sacadas.

Para simular as mesmas 20 sacadas selecionadas em diferentes taxa de amostragens de quadros, como: 30, 60, 100, 120, 150, 187, 250, 500hz, empregou-se a interpolação entre os pontos obtidos de cada sacada a 500hz, para suprir a necessidade de um ponto em um instante de tempo diferente aqueles já existentes. As taxas de amostragens foram selecionadas em vista que há câmeras disponíveis no mercado com essas taxas de amostragens de quadros.

Metodologia

Neste capítulo é apresentado a metodologia empregada na realização do trabalho. Foram escolhidas duas bases de dados, uma para verificação da acurácia das estimativas da pupila mediante oclusão por cílios da nova abordagem, assim foi utilizada a disponibilizada por Swirski em (ŚWIRSKI et al., 2012), descrito em Seção 4.1. E para compor a segunda base de dados, foram escolhidas alguns dados das bases de dados constituídas por dados brutos computados pelo *SMI Hi-Speed Eye Tracker*, para testar o STE (predição da sacada) descrito na Seção 4.3.

4.1. Escolha da base de dados de imagens

É necessário pontuar quesitos para avaliar o algoritmo na base de dados, assim a mesma precisa atender alguns quesitos, como:

1. **Ground-Truth:** é um considerado um gabarito, ou seja, há resposta que são assumidas como corretas e são aceitas. Nesse contexto, é uma base de dados que contém os coeficientes da elipse ou círculo que forma a borda da pupila, no máximo de quadros possíveis;
2. **Oclusões da pupila por cílios:** para avaliar a abordagem desenvolvida é necessário testar a sua performance perante as oclusões por cílios.

Uma base de dados que atende aos requisitos necessários foi disponibilizada em Świrski et al. (2012), contendo imagens dos dois olhos em uma resolução de 640×480 , a uma taxa de 30Hz, possuindo uma rotulação em algumas imagens da melhor pupila estimada, aproximada em 25% das imagens, possuindo apenas pupila escuras. Entretanto a taxa de quadros não é ideal para o emprego do STE, nessa taxa de quadros dificulta a identificação de movimentos dos olhos.

Após a base ser escolhida, foi realizado iterativamente diferentes configurações de tamanhos de raios variando entre um valor mínimo e máximo. O valor mínimo foi de 15 pixels particularmente escolhido devido o tamanho do vetor de convolução implementado e o valor máximo foi de 100 pixels escolhido por ser aproximadamente um quarto do menor lado da imagem. E também foi configurado o número de raios, variando entre 6 a 72 raios, escolhidas pelo resultado da divisão entre o número de raios e 360° resultar em um número inteiro. O programa desenvolvido foi feito com o auxílio das linguagens de programação C e C++ e com a biblioteca de visão computacional de código aberto *Open Source Computer Vision Library* (OpenCV) que esta disponível em C++ e Python, entretanto para o contexto do problema e do trabalho foi utilizado a versão implementada em C++.

Foi estabelecido no primeiro quadro um ponto dentro da pupila, semelhante a inicialização de Starbust e Ohno et al. (2002), a quantidade de raios e o tamanho foram definidos previamente. Assumindo como um ponto a elipse e_{in} , traçando os raios nas diferentes direções e grosseiramente estabelecendo a elipse e_{out} . Na Figura 4.1 demonstra o impacto precisão da estimação da pupila com o diminuição da região entre as elipses.

4.2. Estimativa da borda da pupila

Será apresentado alguns exemplos do processo de traçamento, seleção dos candidatos a borda e estimação da pupila. Todas as sub-imagens na Figura 4.1 foram cortadas a partir da imagens originais para melhor exibição. Na Figura 4.1a é dada como entrada para a abordagem processar, há uma oclusão da pupila por cílios. Como é a primeira imagem do vídeo é selecionado um ponto interno na pupila para começar o processo de traçamento de raios. Assim os raios são traçados como é mostrado na Figura 4.1b, representados em vermelhos, nesta primeira imagem foi utilizado um círculo devido a sua fácil parametrização em comparação a uma elipse. Posteriormente, em Figura 4.1c é realizado a convolução e selecionado os pontos candidatos, representado em amarelo sobrepondo os raios vermelhos. O local em que se encontram no raio, demonstra o maior pico de intensidade na convolução aproximada da derivada da gaussiana. Na Figura 4.1d ajusta-se a elipse com a utilização de todos os pontos candidatos borda da pupila e forma-se a elipse que representa a borda da pupila (em amarelo). Na Figura 4.1e é demonstrado em verde a elipse estimada e em azul a elipse rotulada. Observa-se que a elipse ajustada da Figura 4.1e é semelhante a da elipse rotulada pelo Swirski.

Após os raios traçados, foi convolucionado a derivada da gaussiana ao longo do mesmo para obter e identificar as rampas de intensidade que correspondem a borda da pupila. Selecionou-se o maior pico de intensidade de cada raio convolucionado para ser um ponto candidato que compõe o conjunto com todos os pontos candidatos a borda da pupila. Para ajustar a elipse que estima a borda da pupila foi utilizado o ajustamento da disponível no

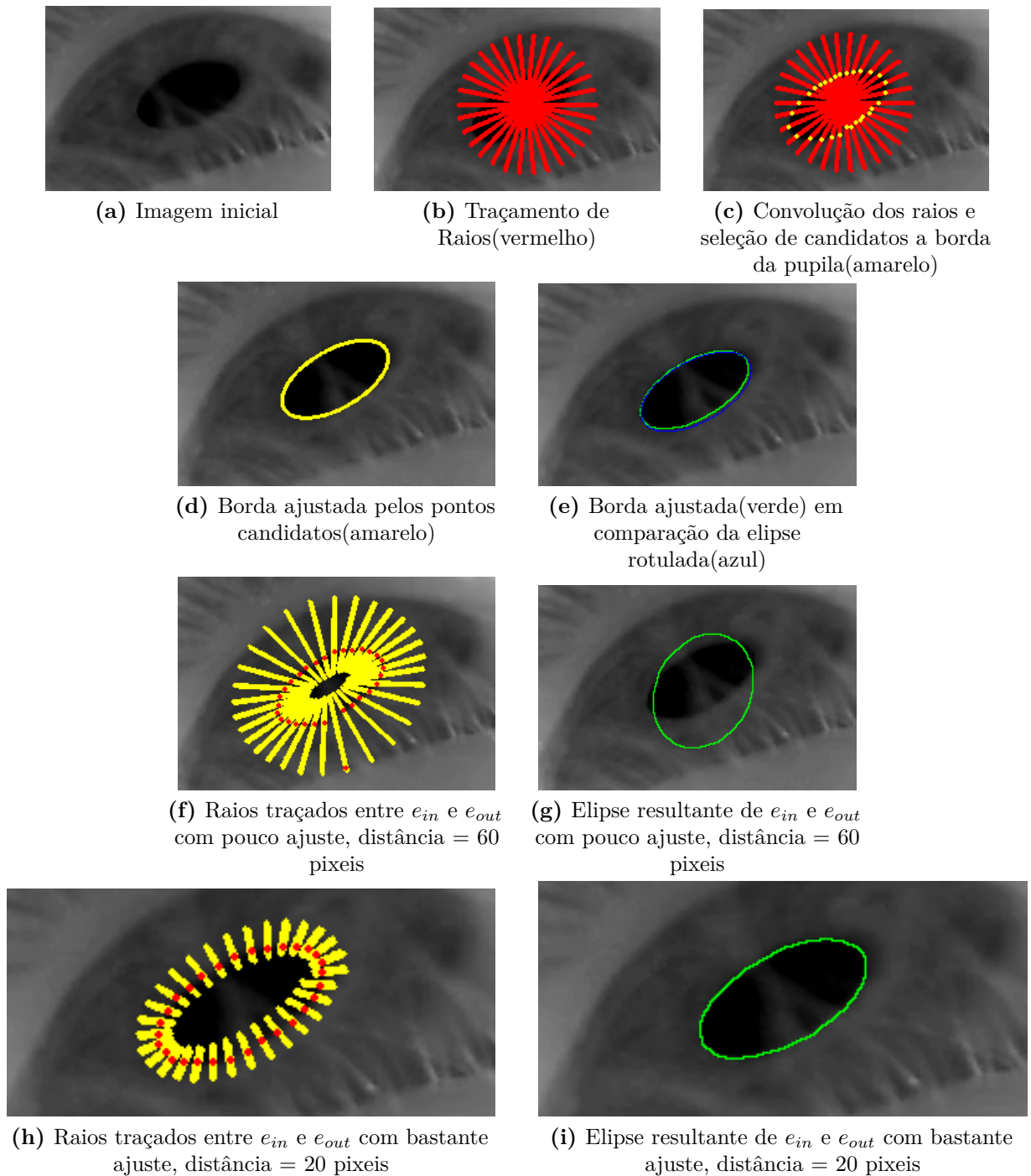


Figura 4.1. Estimação da borda da pupila através do conjunto de todos os pontos candidatos a pupila com diferentes distâncias entre e_{in} e e_{out} com 30 raios.

OpenCV através da função *fitellipse()*. A função utilizada necessita de 5 ou mais pontos para ajustar a elipse. Nesse caso foi utilizado todos os pontos candidatos da borda da pupila, não descartando nenhum candidato.

Para calcular a distância Hausdorff entre a pupila estimada e a pupila do *ground-truth*, foi utilizado todos os quadros que possuem uma elipse rotulada previamente na base de dados para fazer a comparação. Foi comparado a borda da pupila estimada com o a elipse disponível

no gabarito através da distância Hausdorff. Portanto, foi persistido a distância em um arquivo para a análise em cada configuração. Também foi anotado o tempo de processamento de cada quadro em cada configuração para realizar a comparação.

4.3. Escolha da base de dados para o suporte temporal estendido

Para compor a base de dados para testar o suporte temporal estendido foi utilizado dados de 3 bases de dados: ETHyma¹, 3D_Gaze² e Eyetracker_Images_LIVE³. São bases de dados que foram utilizadas em projetos de rastreamento de olhar. Foram encontrados arquivos com dados de vários indivíduos que estimou-se o olhar para diferentes objetos que foram dispostos na tela. Os dados brutos disponíveis tem coeficientes extraídos tanto do olho esquerdo como do olho direito.

Logo foi feito o mesmo processo, a estimação da trajetória da sacada nas diferentes taxas de amostragens de quadros, para a comparação de distância do ponto calculado naquele instante da exponencial gerada entre 500hz e as outras taxas de amostragens inferiores. Com o intuito de saber se há uma discrepância suficientemente grande com a utilização de menos pontos amostrados, no mesmo intervalo de tempo. Também foi utilizado a métrica de tempo de processamento de cada exponencial em cada novo ponto inserido como entrada para o algoritmo do *levmar*. Então, foi armazenado uma lista de cada sacada e sua exponencial e suas variações em cada taxa de amostragem diferente, além disso, foi anotado quanto tempo precisou para processar cada ponto novo dado ao algoritmo para re-estimar a exponencial da trajetória.

4.4. Métricas de comparação de desempenho

Para comparar as elipses estimadas para cada configuração de tamanho de raio e número de raios foi armazenado o tempo de processamento necessário para computá-lo e a sua distância Hausdorff para o *ground-truth*. A distância Hausdorff entre duas elipses é definida pela distância Euclidiana máxima entre qualquer ponto na primeira elipse para o ponto mais próximo na segunda elipse. A acurácia é definida pela distância Hausdorff e a precisão pelo desvio padrão da mesma. A avaliação da predição da sacada foi feita utilizando tempo de processamento, o número de iterações do algoritmo *levmar* e a distância entre a predição com $n - 1$ pontos do próximo ponto e a posição real do próximo ponto.

¹ <<http://ivc.univ-nantes.fr/en/databases/ETHyma/>>

² <http://ivc.univ-nantes.fr/en/databases/3D_Gaze/>

³ <http://ivc.univ-nantes.fr/en/databases/Eyetracker_Images_LIVE/>

Resultados

Neste capítulo é relatado os resultados obtidos pela pesquisa. O objetivo desse trabalho é avaliar o impacto da redução do tamanho dos raios entre e_{in} e e_{out} . Então minimizando a interferência negativa da oclusão por cílios. E assim corroborando para redução no tempo de processamento de cada quadro do vídeo.

5.1. Avaliação da pupila estimada

Utilizou duas formas estimar a pupila, variando a quantidade de raios e o seu tamanho. Obteve-se o tempo de processamento de cada quadro e a distância Hausdorff de cada elipse ajustada para cada configuração de tamanho e quantidade de raios.

Analisou-se o impacto na alteração das configurações dos raios, tanto na quantidade quanto no seu tamanho, para estipular uma faixa onde a redução do tamanho não diminua a distância e que o tempo computacional seja adequado ao contexto.

É demonstrado na Figura 5.1 um exemplo de redução e diminuição do impacto negativo da oclusão na estimação da pupila. Para exemplificar a configuração do raio e o que interfere a oclusão na mesma Figura 5.1, é mostrado três casos de tamanhos variando de 20, 40 e 60 pixels, representando a distância entre um ponto de e_{in} e e_{out} perpendicularmente. Foram obtidos os valores ao longo dos raios e estão representados nas Figuras 5.1a, 5.1d e 5.1g e convolucionados correspondendo nas Figuras 5.1b, 5.1e e 5.1h e Figuras 5.1c, 5.1f e 5.1i são as elipses para 20, 40 e 60 pixels entre as elipses respectivamente da mesma imagem. As diferenças entre os níveis de pixels e o local onde foram encontrados os picos de intensidades estão diferentes devido o formato da disposição dos raios em torno da pupila. Entretanto é visto na Figura 5.1h que o impacto de um candidato *outlier* na estimação da pupila é negativo, resultante do algoritmo de ajustamento da elipse que utiliza todos os pontos candidatos. Nessa abordagem não há um método de filtragem de candidatos entre candidatos *inliers* e

outliers. A filtragem é feita pela diminuição do tamanho do raio, correspondendo a segurança que a diminuição do tamanho entre as duas elipses resulta.

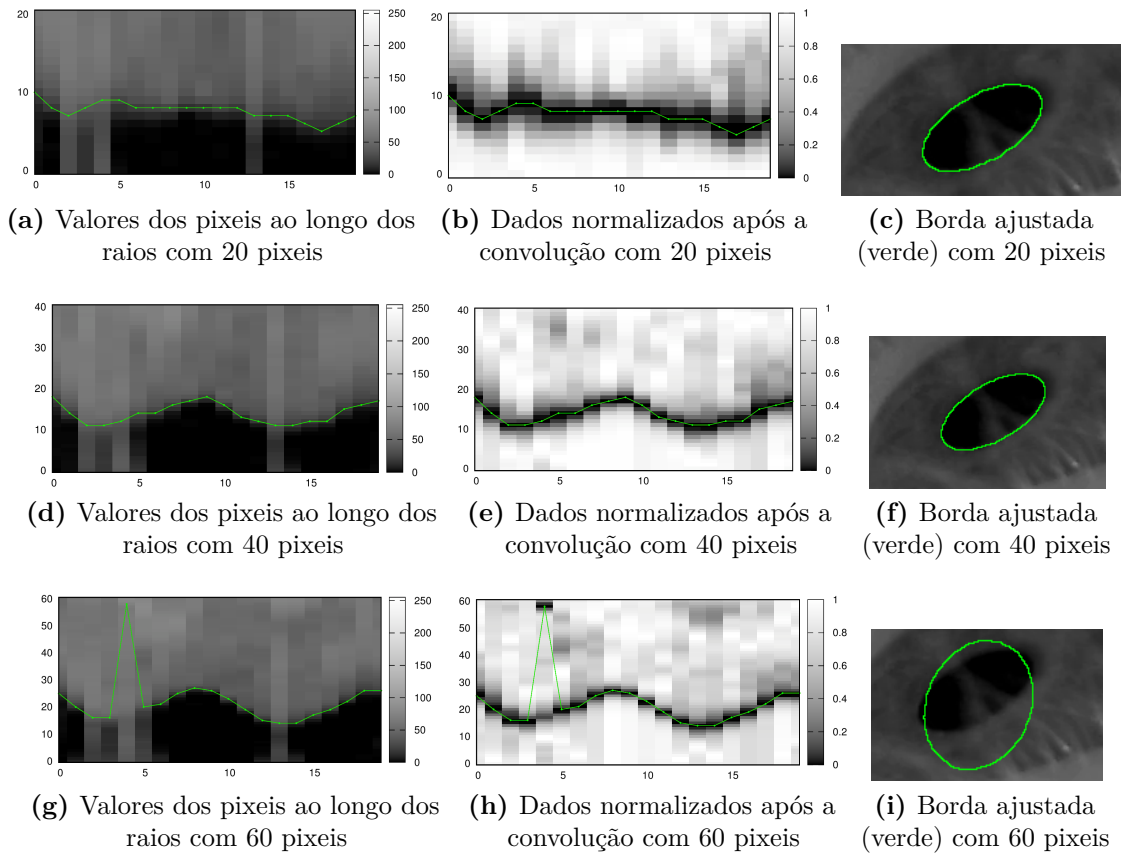


Figura 5.1. Comportamento do raios com diferentes tamanhos

Como a Figura 5.1 foi só um exemplo de desempenho do detector de pupila proposto em relação as oclusões. Um método de comparação para saber se houve uma redução na acurácia, foi proposto a medição da distância Hausdorff entre duas pupilas, a estimada pelo método proposto e a disponível na base de dados disponibilizada no trabalho do Świrski et al. (2012). Portanto, houve uma comparação entre as duas pupila em todas as imagens que possuem pupila rotuladas como elipses. Os dados foram dispostos no formato da Tabela 5.1.

A Tabela 5.1 mostra a média da distância e o desvio padrão para diferentes configurações tanto no tamanho quanto na quantidade de raios. Houve uma redução na distância média ao se empregar raios menores, apresentado por Figura 5.3, Figura 5.4 e por Figura 5.5. Com o aumento em número de quantidade de raios e com a utilização de raios menores houve uma redução da distância causado pelo aumento de raios que são estimados corretamente e a diminuição do tamanho dos raios que possibilita a redução da possível oclusão da pupila. Isso se deve a quantidade de raios que foram sendo computados na região que não está sendo oclusa. Ou seja, há mais chances de que quantos mais raios e com raios menores aumente a quantidade de pontos candidatos a borda da pupila serão estimados

adequadamente, isto é, em um ponto candidato a pupila *inlier*. Portanto, diminuir o raio é um processo que possibilita a redução do impacto negativo da oclusão da pupila na sua estimativa. Um método para diminuir o raio é através da predição da posição da pupila, dado que a predição fornece uma grau de certeza maior, justamente necessário para reduzir o raio sem que não haja o traçamento incorreto de raios em torno da borda da pupila. O algoritmo utilizado para ajustar a borda necessita de todos os pontos candidatos para ajustar a elipse, então quanto mais pontos estimados corretamente melhor.

Tabela 5.1. Tamanho do raio e distância Hausdorff variando o número de raios

Números de Raios	Tamanho do Raio					
	15	36	72	15	36	72
	Distância Hausdorff (pixeis)			Desvio Padrão (pixeis)		
6	1,68	1,77	2,10	1,13	0,97	1,66
12	1,26	1,67	2,78	0,93	0,94	2,59
24	1,08	1,61	1,84	0,75	0,92	1,70
36	1,05	1,48	2,05	0,80	0,90	2,46
50	1,01	1,45	1,89	0,73	0,86	2,07
60	0,99	1,50	2,00	0,66	0,91	2,48
72	0,90	1,47	2,30	0,55	0,88	2,50

O número de raios deve influenciar o tempo de processamento, mostrado na Figura 5.2. Pode-se notar o impacto da utilização de mais raios e a sua performance em relação ao tempo de processamento. Uma maior quantidade de raios, como 72 raios, é perceptível que há um desvio padrão e uma média maior apropriado em relação 60 raios. Entretanto há execuções que tiveram o mesmo tempo de processamento. O número de raios influencia positivamente no comportamento do impacto da oclusão quando o tamanho do raio é menor. Justamente devido a tendência de ter mais raios distribuídos igualmente em torno da pupila, contribuindo também a diminuição do raio que limita o tamanho do erro.

5.2. Análise do tempo de processamento

Analisou o impacto das configurações dos raios em relação de tempo gasto para realizar o processamento de cada quadro. Para saber se houve uma redução no tempo de processamento com utilização de menos raios e raios menores. Com menos raios houve menos raios para serem convolucionados, implicando em menos pixeis a serem processados. Com raios menores houve menos pixeis a serem processados na convolução de cada raio, assim limitando o esforço computacional. Observou-se que o tempo médio gasto e desvio padrão é menor em raios menores, há ressalva para alguns raios que apresentam um acréscimo no tempo de execução. Uma tendência desejada devida a menor quantidade de pixeis a serem processados ao longo do raio, já que seu tamanho delimita o processamento. Além disso, o tempo de processamento foi mais influenciado pela quantidade de raios. Houve

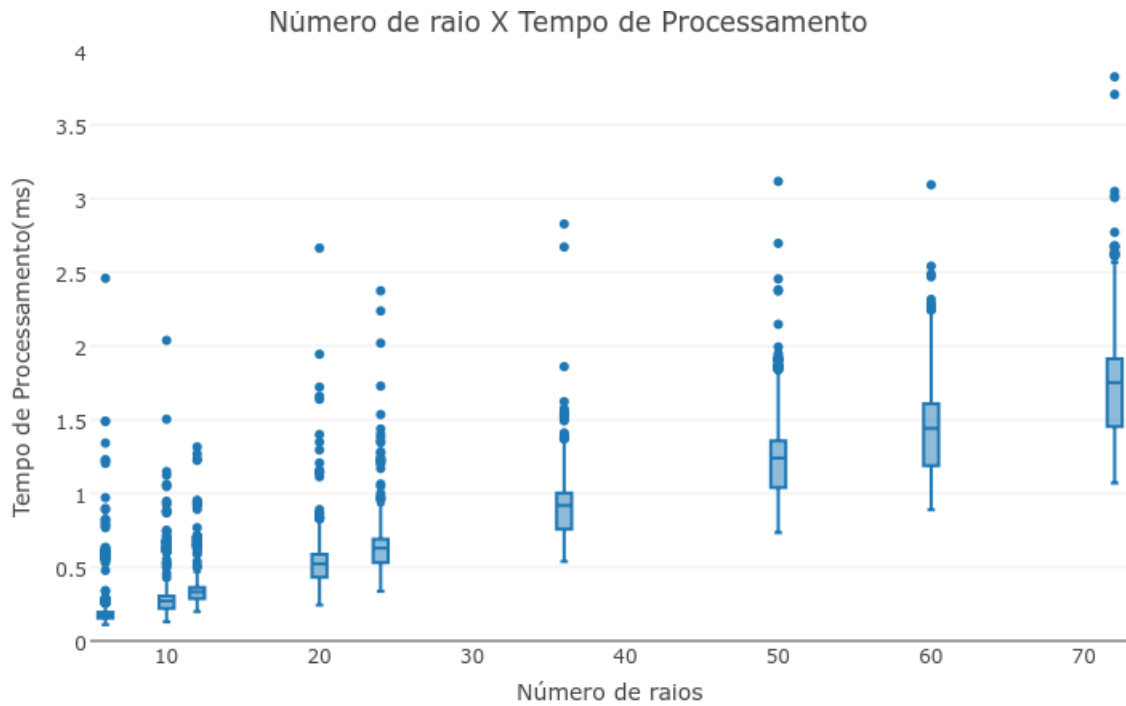


Figura 5.2. Desempenho de número de raios em relação ao tempo de processamento

um aumento no tempo de processamento geral no aumento de quantidade de raios. Isso é apresentado em Figuras 5.3 e 5.5, foi notado que houve uma variação no eixo y, no tempo de processamento a medida que mais raios foram adicionados para serem processados. Essa variação é caracterizada pelo aumento de pixels a serem processados de forma geral no processamento.

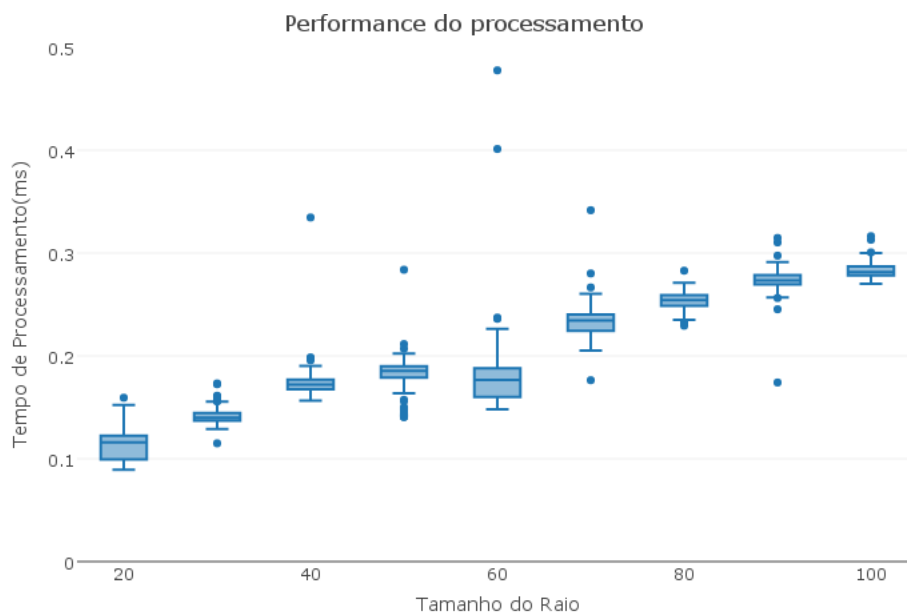


Figura 5.3. Desempenho da abordagem com 12 raios em relação ao tempo de processamento

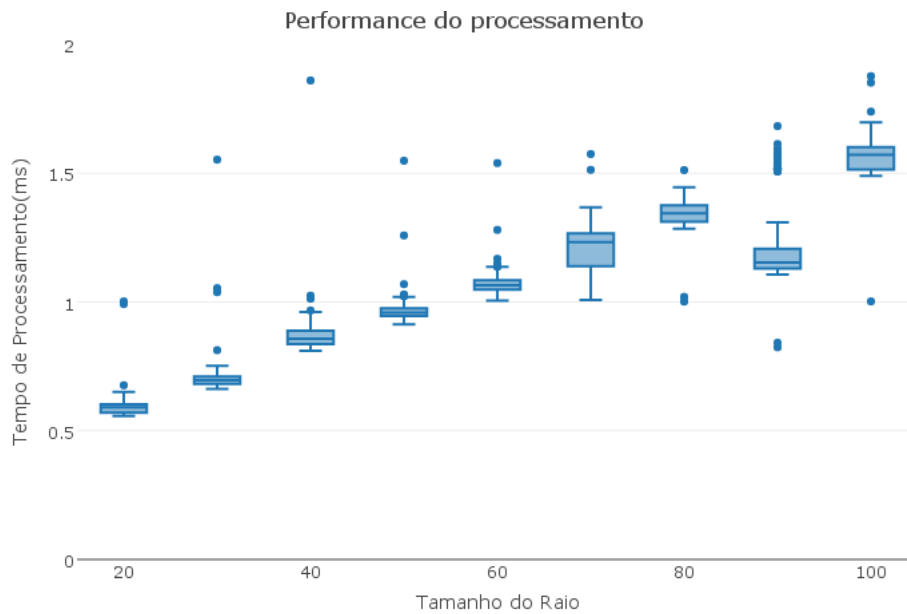


Figura 5.4. Desempenho da abordagem com 36 raios em relação ao tempo de processamento

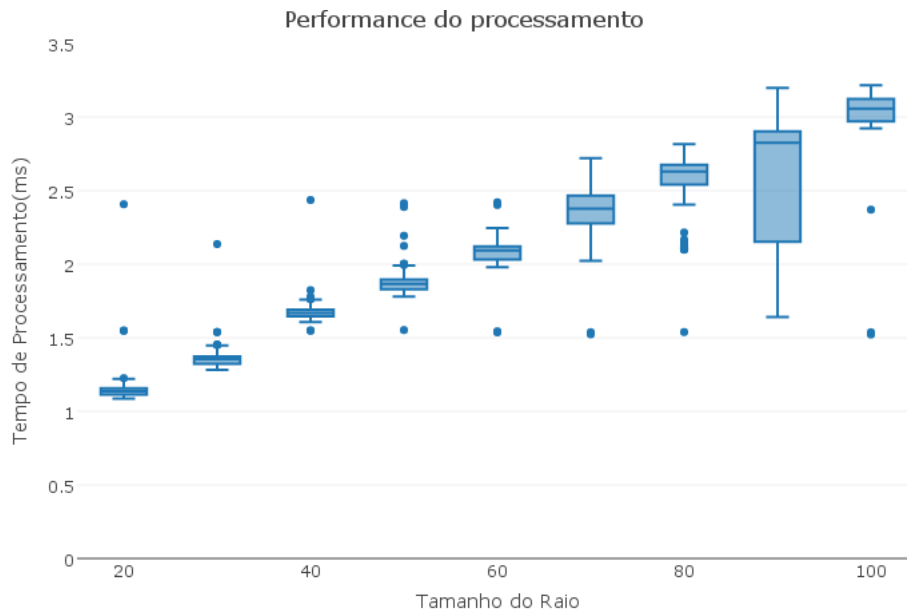


Figura 5.5. Desempenho da abordagem com 72 raios em relação ao tempo de processamento

5.3. Análise da trajetória da sacada

O intuito de analisar a trajetória da sacada é devido a necessidade de saber uma posição mais precisa do olho quando esta se movendo. Świrski et al. (2012) utilizam o quadro anterior como uma aproximação do centro da pupila para um próximo quadro, entretanto isso não apresenta uma precisão em movimentos grandes do olho, como a sacada.

Inicialmente foi usado o pacote *saccades* do R para a identificação das sacadas,

entretanto, a implementação se mostrou bastante ruidosa quando executando sobre as bases testadas, selecionando erroneamente períodos de fixação e pequenos deslocamentos como sacadas. Dentre as muitas sacadas obtidas, foram selecionadas 20 exemplares.

A Figura 5.6 é um exemplo de predição a trajetória da sacada. A fórmula da trajetória da sacada sofreu um ajuste do modelo exponencial usando *levmar* a um conjunto de pontos. O conjunto de pontos variou entre 3, 4, 5, até o total de pontos presente na sacada, para que gerasse uma fórmula para sacada. Do terceiro até décimo a distância entre o comportamento da sacada real foi grande, porém acima do décimo ponto houve um bom acerto, ou seja, o comportamento foi ajustado e a sua distância entre a predição e trajetória real da sacada foi menor que 1 grau.

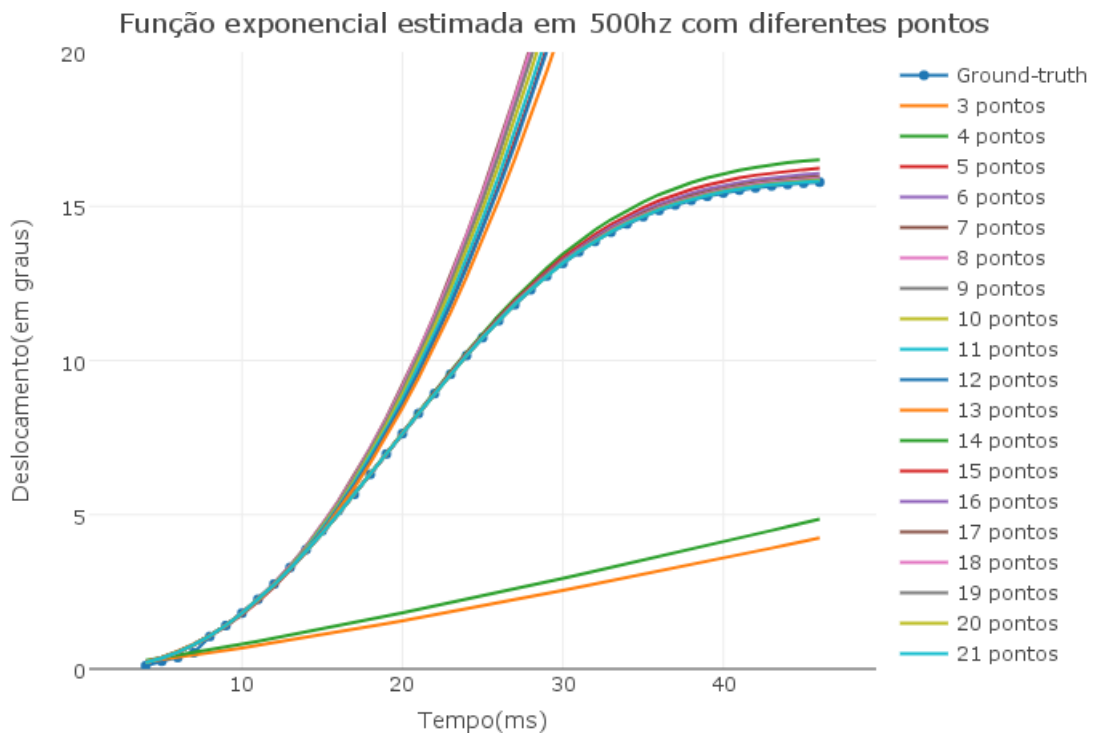


Figura 5.6. Trajetória da sacada com diferentes quantidades de pontos amostrados

O *ground-truth* nesse contexto são os pontos originais obtidos em 500hz extraídos pelo *saccades*. Após a estimativa das trajetórias das sacadas com a fórmula e em diferentes taxa de amostragens. Foi obtido a distância da previsão em diferentes taxas de amostragens e do *ground-truth*. Observando o gráfico pode se concluir que há alguns picos acima de 5 graus de diferença entre o *ground-truth* e as diversas taxas de amostragens. Porém, isso pode ser reavaliado utilizando limiares de distância máxima percorrida devido ao conhecimento de que o olho não move tanto em tão pouco tempo.

O tempo de processamento é menor que 1ms, a função estimada pelo *levmar* minimiza os erros com o passar do tempo. Inicialmente é usado um palpite para a função, parâmetro difícil de estimar devido a variação da amplitude da sacada e duração em tempo que esta ocorrendo, assim no início geralmente é alta o número de iterações, causando um maior tempo

de processamento. Porém, ao passar o tempo com a estimação recebe mais pontos, a função tende a minimizar os erros utilizando menos iterações para estimar as próximas funções, reduzindo assim o esforço computacional. No começo do processamento pelo algoritmo de utilizado pelo *levmar*, o número de iteração é alto no começo em comparação ao resto, demonstrado na Figura 5.7.

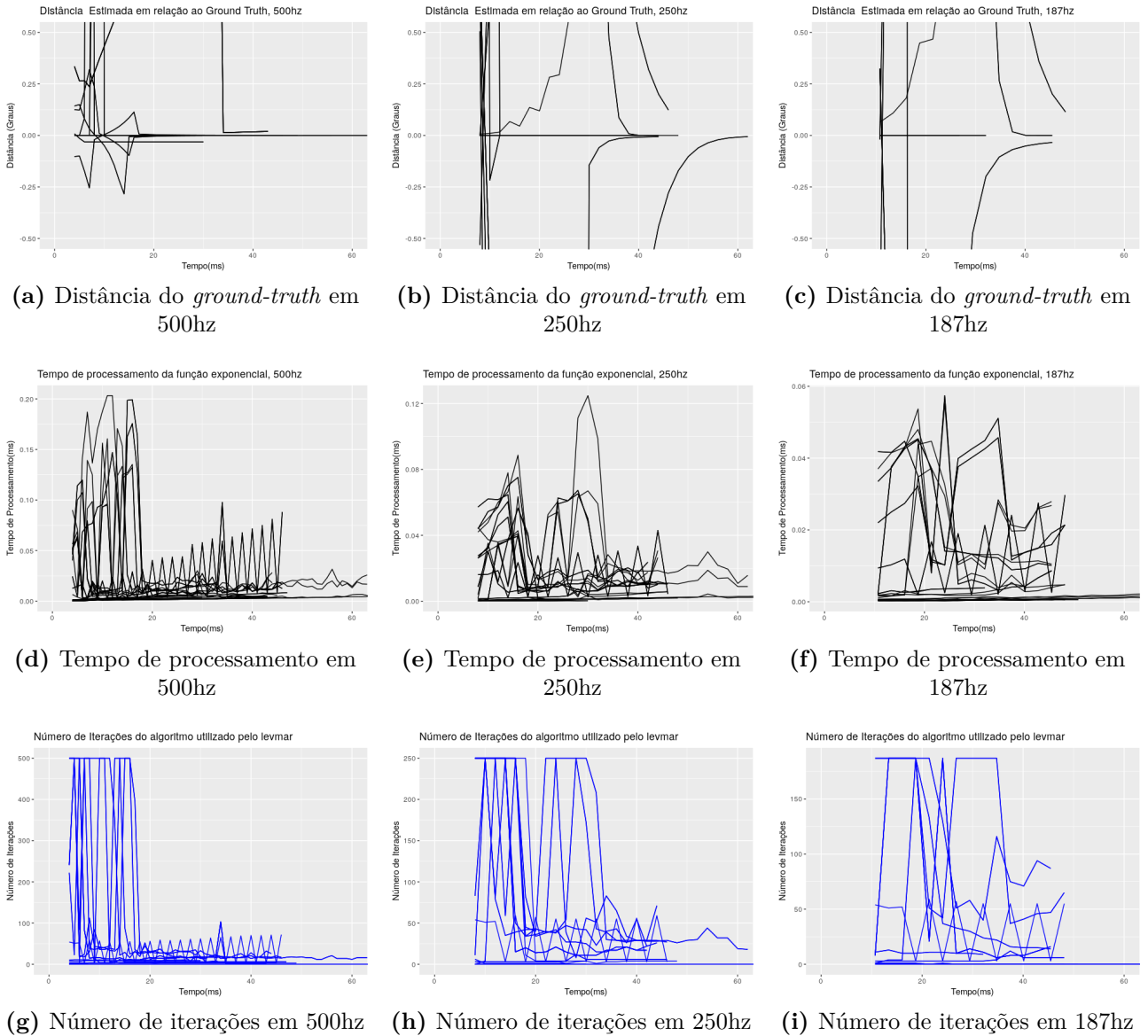


Figura 5.7. Desempenho do algoritmo *levmar* em relação a distância do *ground-truth*, tempo de processamento e número de iterações

Após feito a comparação com vários tempo de amostragens, foi constatado que para taxa de amostragens igual ou menor 120hz não permite uma adequada estimativa da trajetória da sacada, devido a diminuir a quantidade de pontos disponibilizado por cada taxa.

Conclusão

O intuito de rastreadores de olhar é fornecer a posição onde o usuário está olhando. São utilizados em pesquisas e estudos de usabilidade de interfaces, diagnósticos de doenças, entre outros. Um dos desafios enfrentado pelos rastreadores de baixo custo é usar uma câmera com uma alta de quadros, e conseguir processar todo os quadros em tempo hábil devido a capacidade de sua câmera e também de seus algoritmos que não se adaptam em altas taxas. Porém existem rastreadores comerciais que atingem taxas de 1Khz ou superiores, mas eles são rastreadores destinados a pesquisas, caros, grandes e intrusivos. Estes necessitam de algoritmos que precisam lidar com essas altas taxas de quadros. Porém tempo de processamento não é o único desafio, já que a pupila oclusa dificulta a sua estimativa afetando a acurácia do rastreador.

O objetivo do trabalho é reduzir o tempo de processamento para a estimativa da borda da pupila e reduzir o impacto da oclusão por cílios, ocasionando no geral um melhora na acurácia. Um dos métodos possíveis é a redução da região a ser processada, para que além de impactar positivamente no tempo de processamento, contornar o desafio da presença da oclusão por cílios. Foi estendido o método Ohno et al. (2002) empregando o traçamento de raios. Foi empregado a convolução ao longo do raio com intuito de identificar rampas de intensidade, que após convolucionados apresentam pico de intensidade na mesma posição das rampas. Isso pode ser possível devido a diferença entre pixels na pupila que se apresenta escura e a íris que se apresenta mais clara. Assim com raios que transpassam a região da borda da pupila, isso torná-se viável e um cenário ideal. As posições onde se encontram a rampa de intensidade é considerada uma candidato a borda da pupila, que forma com um conjunto de pontos utilizando vários raios. Esse conjunto de pontos é utilizado no processo de refinamento para o ajuste da elipse que melhor se enquadra. Porém, desafios como a oclusão dificultam a estimativa da pupila, produzindo candidatos inadequados para o processo de refinamento, e como nesse trabalho não há um método de filtragem de candidatos adequados

e inadequados para refinamento. Um meio de reduzir o impacto da oclusão foi necessário.

Neste trabalho foi empregado uma região reduzida, implicando em diminuir o tamanho dos raios traçados. Acarretando em um processo de saber precisamente em qual lugar na imagem a pupila se localiza, uma alternativa para saber é utilizar informações dos quadros anteriores, como em Starbust utilizando a posição do centro da pupila estimada em um quadro e utilizado como palpite para o próximo. Porém, não provê uma região tão precisa necessária nesse trabalho, assim surge a carência de algo que consiga predizer onde esta a borda da pupila. Aqui nesse trabalho, foi estudado a predição de movimentação do olho. Forneceu uma região precisa em uma alta taxa de quadros, como em 500, 250 ou 187hz, mas em taxas inferiores não foi obtido bons resultados devido sua dependência a uma alta taxa de quadros.

Reduzindo o tamanho dos raios pode-se perceber que a distância Hausdorff diminuía entretanto também variava conforme a quantidade de raios aumentava. Isso deve-se ao fato de que se há uma oclusão, raios maiores podem adquirir na convolução vários picos de intensidade, podendo ocasionar uma estimativa da borda da pupila inadequada. Para contornar esse desafio é empregado raios menores, que contribuem para redução do impacto da oclusão ao longo do raio, e após a convolução, tendem a estar mais comportados com a presença de apenas um pico de intensidade correspondendo ao desejo inicial da pesquisa.

Então analisando os resultados obtidos, pode-se concluir que o tempo de processamento pode ser reduzido utilizando uma região reduzida para o processamento. Entretanto, tem que se atentar a quantidade de raios necessários traçados na imagem, porque esse é um fator determinante para o tempo de processamento. Em vista que a acurácia aumenta com a uma quantidade de raios maiores, correspondendo a diluição da possível oclusão em mais raios, e portanto diminuindo no seu impacto negativo no trabalho. Contudo, a utilização de uma quantidade maior de raios impacta negativamente no tempo de processamento, mas afeta positivamente a acurácia de modo geral. Outro fator que altera a acurácia é o tamanho do raio, que pode afetar positivamente ou negativamente dependendo do seu tamanho. No trabalho foi aplicado raios variados e chegou a conclusão que raios acima de 40-50 pixels, começam a dificultar a estimação da pupila devido a inserção de partes da imagem desnecessária para a identificação do que é pupila ou não.

Esses resultados podem ajudar pesquisas futuras a desenvolver um rastreador de olhar com uma câmera de alta taxas de quadros, que consiga processar todos os quadros em tempo hábil utilizando as redução da região candidata e a predição da posição da pupila. E como trabalhos futuros, planeja-se usar algoritmo de contornos para saber onde está a pálpebra e os cílios para assim, não traçar raios desnecessários. Otimizando e limitando a região de procura a apenas parte visível do olho.

Referências

- AGUSTIN, Javier San et al. Evaluation of a low-cost open-source gaze tracker. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. New York, NY, USA: ACM, 2010. (ETRA '10), p. 77–80. ISBN 978-1-60558-994-7. Disponível em: <<http://doi.acm.org/10.1145/1743666.1743685>>.
- BORSATO, F. H. *Rastreador de olhar de alto desempenho e baixo custo baseado em iluminação estroboscópica diferencial*. 144 p. Tese (Doutorado) — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2016.
- BUSWELL, G T. How people look at pictures: a study of the psychology and perception in art. *Chicago University of Chicago Press*, p. 198, 1935. ISSN 0033-2909.
- DUCHOWSKI, Andrew T. A breadth-first survey of eye-tracking applications. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, v. 34, n. 4, p. 455–470, 2002. ISSN 0743-3808.
- EBISAWA, Yoshinobu; SATOH, S-I. Effectiveness of pupil area detection technique using two light sources and image difference method. In: IEEE. *Engineering in Medicine and Biology Society, 1993. Proceedings of the 15th Annual International Conference of the IEEE*. [S.l.], 1993. p. 1268–1269.
- ELLIS, CJ. The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, BMJ Publishing Group Ltd., v. 65, n. 11, p. 754–759, 1981.
- FITZGIBBON, Andrew; PILU, Maurizio; FISHER, Robert B. Direct least square fitting of ellipses. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 21, n. 5, p. 476–480, 1999.
- GONZALEZ, Rafael C.; WOODS, Richard E. Digital image processing. In: . 2nd. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001. cap. 2, p. 135–103. ISBN 0201180758.
- HAN, Peng et al. Trajectory prediction of saccadic eye movements using a compressed exponential model. *Journal of vision*, The Association for Research in Vision and Ophthalmology, v. 13, n. 8, p. 27–27, 2013.
- HARRIS, Christopher M; WOLPERT, Daniel M. The main sequence of saccades optimizes speed-accuracy trade-off. *Biol Cybern*, v. 95, p. 21–29, 2006.
- HOLMQVIST, Kenneth et al. Eye Tracking : A Comprehensive Guide to Methods and Measures. Oxford University Press, 2011. Disponível em: <<http://ukcatalogue.oup.com/product/9780199697083.do>>.

- KIM, Soo Chan et al. A new method for accurate and fast measurement of 3d eye movements. *Medical Engineering & Physics*, v. 28, n. 1, p. 82–89, 2006. ISSN 1350-4533. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1350453305000779>>.
- LI, Dongheng; PARKHURST, Derrick J. Starburst: A robust algorithm for video-based eye tracking. *Elsevier Science, Citeseer*, v. 6, 2005.
- LOURAKIS, M.I.A. *levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++*. Julho 2004. <http://www.ics.forth.gr/~lourakis/levmar/>+. Acessado em 15 Novembro de 2016.
- MORIMOTO, Carlos H.; MIMICA, Marcio R M. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*, v. 98, n. 1, p. 4–24, 2005. ISSN 10773142.
- OHNO, Takehiko; MUKAWA, Naoki; YOSHIKAWA, Atsushi. Freegaze: A gaze tracking system for everyday gaze interaction. In: *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*. New York, NY, USA: ACM, 2002. (ETRA '02), p. 125–132. ISBN 1-58113-467-3. Disponível em: <<http://doi.acm.org/10.1145/507072.507098>>.
- RAYNER, Keith. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, American Psychological Association, v. 124, n. 3, p. 372, 1998.
- ŚWIRSKI, Lech; BULLING, Andreas; DODGSON, Neil. Robust real-time pupil tracking in highly off-axis images. In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. New York, NY, USA: ACM, 2012. (ETRA '12), p. 173–176. ISBN 978-1-4503-1221-9. Disponível em: <<http://doi.acm.org/10.1145/2168556.2168585>>.
- WADE, Nicholas J. Pioneers of eye movement research. *i-Perception*, SAGE Publications, v. 1, n. 2, p. 33–68, 2010.
- YARBUS, Alfred L. Eye movements during perception of moving objects. In: *Eye Movements and Vision*. Boston, MA: Springer US, 1967. p. 159–170. ISBN 978-1-4899-5379-7. Disponível em: <http://dx.doi.org/10.1007/978-1-4899-5379-7_7>.
- YOUNG, Laurence R; SHEENA, David. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, Springer, v. 7, n. 5, p. 397–429, 1975.
- ZHU, Zhiwei; JI, Qiang. Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Computer Vision and Image Understanding*, v. 98, n. 1, p. 124–154, 2005. ISSN 10773142.