

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**LUCAS ROYES SCHARDOSIM**

**DETECÇÃO E ESTIMAÇÃO DE  
MOVIMENTO ROBUSTAS A ROTAÇÃO,  
ESCALA E TRANSLAÇÃO EM  
SEQUÊNCIAS DE IMAGENS EM  
INFRAVERMELHO DA RETINA**

Porto Alegre  
2014

**LUCAS ROYES SCHARDOSIM**

**DETECÇÃO E ESTIMAÇÃO DE  
MOVIMENTO ROBUSTAS A ROTAÇÃO,  
ESCALA E TRANSLAÇÃO EM  
SEQUÊNCIAS DE IMAGENS EM  
INFRAVERMELHO DA RETINA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Engenharia de Computação, Processamento de Sinais

**ORIENTADOR: Prof. Dr. Jacob Scharcanski**

Porto Alegre  
2014

**LUCAS ROYES SCHARDOSIM**

**DETECÇÃO E ESTIMAÇÃO DE  
MOVIMENTO ROBUSTAS A ROTAÇÃO,  
ESCALA E TRANSLAÇÃO EM  
SEQUÊNCIAS DE IMAGENS EM  
INFRAVERMELHO DA RETINA**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: \_\_\_\_\_

Prof. Dr. Jacob Scharcanski, UFRGS

Doutor pela (University of Waterloo – Waterloo, Canadá)

Banca Examinadora:

Prof. Dr. Eduardo Barros da Silva, COPPE-UFRJ  
Doutor pela University of Essex – Colchester, Inglaterra

Prof. Dr. Luciano Silva da Silva, Engenharia Agroindustrial Agroquímica-  
FURG  
Doutor pela UFRGS – Porto Alegre, Brasil

Prof. Dr. Valner João Brusamarello, PPGEE-UFRGS  
Doutor pela UFSC – Florianópolis, Brasil

Coordenador do PPGEE: \_\_\_\_\_

Prof. Dr. Alexandre Sanfelice Bazanella

Porto Alegre, agosto de 2014.

## **DEDICATÓRIA**

Dedico este trabalho aos meus pais Caetano e Maria, irmã Chris, esposa Kellen, primos, amigos e tias, em especial, pela dedicação e apoio em todos os momentos.

## **AGRADECIMENTOS**

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Aos colegas do PPGEE e PPGC pelo seu auxílio nas tarefas desenvolvidas durante o curso.

Ao Prof. Dr. Jacob Scharcanski.

Ao pessoal da secretaria do PPGEE, Miriam, Alessandra e Roberto.

À Opto Eletrônica pela base de dados.

Ao CNPq pela provisão da bolsa de mestrado.

## RESUMO

Esta dissertação de mestrado versa sobre um algoritmo de detecção e de estimação de movimento robustas a rotação, escala e translação em sequências de imagens da retina em infravermelho. O algoritmo desenvolvido detecta movimentos da retina com base em medidas de distribuição estatísticas, como a entropia conjunta e a informação mútua entre os quadros adjacentes de um vídeo. A técnica de estimação de movimento inspira-se nos métodos de deslocamento de fase e na transformação log-polar para estimar os parâmetros de movimento. Esta abordagem produz uma solução robusta para este tipo de imagens intrinsecamente ruidosas e comumente utilizadas no diagnóstico e tratamento a laser do edema macular diabético.

**Palavras-chave:** Engenharia Elétrica, Processamento de Sinais, Estimação, Infravermelho, Registro, Imagem, Rastreamento, Log-Polar.

## **ABSTRACT**

This thesis it is about an algorithm of retinal motion detection and estimation robust to rotation, scaling and translation in infrared images sequences. This algorithm detects retinal movements based on measures of statistical distribution, such as joint entropy and mutual information between adjacent frames of a video. The technique of motion estimation is inspired by the methods of phase shift and log-polar transformation to estimate the parameters of motion. This approach yields a robust solution for this type intrinsically noisy images and commonly used in the diagnosis and laser treatment of diabetic macular edema.

**Keywords: Electrical Engineering, Signal Processing, Estimation, Infrared, Register, Image, Tracking and Log-polar.**

## SUMÁRIO

<b>LISTA DE ILUSTRAÇÕES</b> . . . . .	8
<b>LISTA DE TABELAS</b> . . . . .	12
<b>LISTA DE ABREVIATURAS</b> . . . . .	13
<b>LISTA DE SÍMBOLOS</b> . . . . .	14
<b>1 INTRODUÇÃO</b> . . . . .	17
<b>2 REVISÃO DO ESTADO DA ARTE</b> . . . . .	24
<b>3 DETECÇÃO E ESTIMAÇÃO DOS MOVIMENTOS DA RETINA EM SEQUÊNCIAS DE IMAGENS EM INFRAVERMELHO</b> . . . . .	26
3.1 Detecção dos Movimentos da Retina . . . . .	26
3.2 Identificação de Inconsistências nos Movimentos da Retina . . . . .	35
3.3 Estimação dos Movimentos da Retina . . . . .	38
3.4 Sumário do Método Proposto . . . . .	54
3.5 Classificação dos Movimentos da Retina . . . . .	55
3.6 Identificação e Ajuste dos Parâmetros Críticos para a Detecção e Estimação dos Movimentos da Retina . . . . .	56
3.7 Método Comparativo para Detecção e Estimação do Movimento da Retina . . . . .	57
<b>4 RESULTADOS EXPERIMENTAIS</b> . . . . .	61
4.1 Ajustes de Parâmetros . . . . .	62
4.2 Resultados Obtidos . . . . .	64
4.3 Discussão . . . . .	75
<b>5 CONCLUSÃO</b> . . . . .	78
5.1 Trabalhos Futuros . . . . .	79
<b>APÊNDICE A CONTRIBUIÇÕES DESTE TRABALHO</b> . . . . .	80
<b>REFERÊNCIAS</b> . . . . .	82



## LISTA DE ILUSTRAÇÕES

Figura 1:	(a) Imagem em seção da área macular da retina obtida por Tomografia de Coerência Óptica(OCT). Área macular sem alterações, com depressão normal na zona central da fóvea e sem espessamento da retina; (b) Imagem também obtida por OCT da área macular de um doente com <i>EMD</i> . Nota-se o espessamento da área foveal, com perda da depressão foveolar e com áreas de menor refletância (mais escuras) que correspondem a líquido intra retiniano (NOGUEIRA et al., 2007). . . . .	18
Figura 2:	(a) Tomografia de coerência óptica mostrando o EMD antes da injeção intravítrea; (b) com 1 semana; (c) 1 mês após a injeção (REIS et al., 2010). . . . .	20
Figura 3:	Imagem de fundo de olho e retinógrafo digital de tratamento a laser (KERNT et al., 2011). . . . .	21
Figura 4:	Ilustração de quadros consecutivos do vídeo #3 do conjunto de dados.	23
Figura 5:	Ilustração de um par de quadros adjacentes e as respectivas magnitudes do gradiente: (a) quadro $I(t)$ ; (b) quadro $I(t + 1)$ ; (c) magnitude do gradiente de $I(t)$ , $ \overrightarrow{\nabla I(t)} $ ; (d) magnitude do gradiente de $I(t + 1)$ , $ \overrightarrow{\nabla I(t + 1)} $ . . . . .	28
Figura 6:	Amostra dos primeiros 5 quadros do vídeo#1 do conjunto de dados.	31
Figura 7:	Amostra dos primeiros 5 quadros do vídeo#2 do conjunto de dados.	31
Figura 8:	Amostra dos primeiros 5 quadros do vídeo#3 do conjunto de dados.	31
Figura 9:	Amostra dos primeiros 5 quadros do vídeo#4 do conjunto de dados.	31
Figura 10:	Amostra dos primeiros 5 quadros do vídeo#5 do conjunto de dados.	31
Figura 11:	Amostra dos primeiros 5 quadros do vídeo#6 do conjunto de dados.	31
Figura 12:	Amostra dos primeiros 5 quadros do vídeo#7 do conjunto de dados.	32
Figura 13:	Ilustração de uma piscada de olhos. Os quadros 149 – 154 de 6 quadros em infravermelho da retina são descritos acima como (a) - (f), respectivamente; movimento da retina é detectado como inconsistente para os quadros adjacentes 149 – 153, ou seja, (b)-(d). O piscar de olhos ocorre nos quadros 150 – 152, onde a entropia conjunta calculado é $H(X, Y) \leq \eta$ . . . . .	37
Figura 14:	Ilustração dos valores da entropia conjunta $H(X, Y)$ , $GT$ , $H(X, Y) > \eta$ , $D(X, Y) \leq \tau$ para uma sequência em infravermelho na proximidade de uma piscada ilustrada na Figura 13. A piscada de olhos começa no quadro 149 e termina no quadro 153. . . . .	38

Figura 15:	Ilustração do método de deslocamento de fase para estimar os parâmetros de translação. (a) $f(x, y) \equiv I(x, y, t)$ ; (b) $g(x, y) \equiv I(x, y, t+1)$ ; (c) Transformada de Fourier de $f(x, y)$ , $F(u, v)$ ; (d) Transformada de Fourier de $g(x, y)$ , $G(u, v)$ ; (e) Função impulsiva correlação-cruzada $I_{\Delta}(x, y)$ ; (f) Vista superior da função impulsiva correlação-cruzada $I_{\Delta}(x, y)$ . . . . .	41
Figura 16:	Ilustra a transformação polar e log-polar. (a) Qualquer ponto no espaço $(x, y)$ pode ser representado no espaço polar $(r, a)$ ; (b) Qualquer ponto no espaço $(r, a)$ pode ser representado no espaço log-polar $(\rho, a)$ . . . . .	43
Figura 17:	Exibe a transformação log-polar para uma figura que representa um rosto onde é possível verificar a transposição dos elementos de uma face: a boca em vermelho, o nariz em amarelo, os olhos em azul e o contorno em verde. (a) é o rosto no domínio cartesiano $(x, y)$ ; (b) é o rosto no domínio log-polar $(\log r, a)$ . . . . .	44
Figura 18:	Exibe a transformação log-polar para a imagem de um babuíno. (a) é a imagem do babuíno no domínio cartesiano $(x, y)$ ; (b) é a imagem do babuíno no domínio log-polar $(\log r, a)$ . . . . .	45
Figura 19:	Ilustra o efeito de uma rotação entre dois quadros no domínio log-polar. (a) representa o quadro $I$ no plano cartesiano; (b) é o quadro $I$ depois da transformação log-polar, $I_p$ ; (c) é $I$ rotada por $45^\circ$ ; (d) é a sua contraparte no espaço log-polar. . . . .	47
Figura 20:	Exibe o efeito de uma rotação e ampliação entre dois quadros consecutivos da retina. (a) representa um quadro da retina em infravermelha no tempo $t$ ; (b) é a sua contraparte no espaço log-polar; (c) é o quadro consecutivo em $t+1$ com uma ampliação de 2 vezes, uma rotação de $30^\circ$ e uma translação nula; (d) é o quadro $I(t+1)$ no domínio log-polar. . . . .	48
Figura 21:	Ilustra passos do uso da transformação log-polar para estimar os parâmetros; (a) quadro da retina $I(t)$ ; (b) quadro da retina $I(t+1)$ ; (c) $I'_1$ é o recorte da região central de $I(t)$ ; (d) $I'_2$ é o recorte da região central de $I(t+1)$ ; (e) $I'_{1p}$ é a transformação log-polar de $I'_1$ ; (f) $I'_{2p}$ é a transformação log-polar de $I'_2$ . . . . .	50
Figura 22:	Exibe os mapas de máxima correlação-cruzada, $p_{map}$ construídos pelo algoritmo de estimação dos parâmetros de movimento; (a) $p_{map}$ resultante da busca esparsa pelo máximo pico com a posição estimada $(90, 141)$ ; (b) $p_{map}$ resultante após a busca completa na região estimada do pico. O máximo pico de correlação-cruzada acontece na posição $(92, 142)$ . Os parâmetros de translação $(\Delta x, \Delta y)$ são dados pela posição do pico em relação ao centroide. . . . .	51
Figura 23:	Ilustra a correlação-cruzada $p(x'' + mm, y'' + nm)$ realizada na posição $(x'' + mm, y'' + nm)$ tal que a posição do pico traduz nos parâmetros de rotação e escala são dados pelos deslocamentos $da$ e $d(\log_b r)$ . . . . .	52

Figura 24:	Ilustra a comparação entre os métodos. (a) e (b), ou seja, $I(t)$ e $I(t + 1)$ , são quadros consecutivos com grande movimento entre si. (c) $I(t + 1)$ corrigido pelo método proposto; (d) $I(t + 1)$ corrigido pelo método em (REDDY; CHATTERJI, 1996); (e) resultado da correlação-cruzada com o método proposto contendo apenas um pico bem definido; (f) resultado da correlação-cruzada com o método em (REDDY; CHATTERJI, 1996) contendo vários picos; (g) e (h) são as visões de cima das correlações-cruzadas para ambos os métodos. Neste exemplo, o método proposto alcança maior $PSNR$ entre $I(t)$ e $I(t + 1)$ corrigido por parâmetros estimados. . . . .	60
Figura 25:	Valores da métrica universal $D(X, Y)$ (curva azul) para 200 quadros do conjunto teste das imagens da retina em infravermelho. Os segmentos de linhas tracejadas em vermelho indicam as possíveis regiões de movimento pequeno ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge [  \theta   \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$ ), os segmentos de linha na cor preta mostram o conjunto verdade $GT$ . . . . .	64
Figura 26:	Valores da entropia conjunta $H(X, Y)$ (curva azul)para para 200 quadros do conjunto teste das imagens da retina em infravermelho. Os segmentos de linhas tracejadas em vermelho indicam as possíveis regiões de movimento pequeno ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge [  \theta   \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$ ), os segmentos de linha na cor preta mostram o conjunto verdade $GT$ . . . . .	65
Figura 27:	O ângulo $\theta$ para cada par de quadros adjacentes estimado pelo método proposto para os 400 primeiros quadros do conjunto teste. . . . .	65
Figura 28:	O fator de escala $S$ para cada par de quadros adjacentes estimado pelo método proposto para os 400 primeiros quadros do conjunto teste. . . . .	66
Figura 29:	A translação $\sqrt{\Delta x^2 + \Delta y^2}$ para cada par de quadros adjacentes estimada pelo método proposto para os 400 primeiros quadros do conjunto teste. . . . .	66
Figura 30:	Histograma da $PSNR$ entre cada par de quadros para o método proposto e o método em(REDDY; CHATTERJI, 1996) nos conjuntos de treinamento e teste, respectivamente: (a) método proposto em (REDDY; CHATTERJI, 1996) no conjunto de treinamento; (b) método proposto em (REDDY; CHATTERJI, 1996) no conjunto de teste; (c) método proposto no conjunto de treinamento; (d) método proposto no conjunto de teste. . . . .	68
Figura 31:	Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #1,(a) $I_{222}(t)$ ; (b) $I_{223}(t + 1)$ ; (c) $I_{223}(t + 1)$ corrigido por $-\theta, 1/S, -\Delta x$ e $-\Delta y$ . Os parâmetros estimados são $\theta = -15.58^\circ, S = 1.36, \Delta x = -23$ e $\Delta y = -21$ , atingindo um $PSNR = 42.76dB$ . Métrica $D(X, Y) = 0.9815$ e entropia conjunta $H(X, Y) = 2.05$ . . . . .	69
Figura 32:	Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #2,(a) $I_{625}(t)$ ; (b) $I_{626}(t + 1)$ ; (c) $I_{625}(t + 1)$ corrigido por $-\theta, 1/S, -\Delta x$ e $-\Delta y$ . Os parâmetros estimados são $\theta = 10.39^\circ, S = 0.86, \Delta x = 9$ e $\Delta y = 10$ , atingindo um $PSNR = 40.10dB$ . Métrica $D(X, Y) = 0.9823$ e entropia conjunta $H(X, Y) = 1.9222$ . . . . .	70

- Figura 33: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #3,(a)  $I_{61}(t)$ ; (b)  $I_{62}(t + 1)$ ; (c)  $I_{61}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 0^\circ$ ,  $S = 1$ ,  $\Delta x = -4$  e  $\Delta y = 1$ , atingindo um  $PSNR = 40.87dB$ . Métrica  $D(X, Y) = 0.9742$  e entropia conjunta  $H(X, Y) = 2.2725$  . 71
- Figura 34: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #4,(a)  $I_{95}(t)$ ; (b)  $I_{96}(t + 1)$ ; (c)  $I_{95}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 0.86^\circ$ ,  $S = 1.29$ ,  $\Delta x = -20$  e  $\Delta y = 2$ , atingindo um  $PSNR = 38.30dB$ . Métrica  $D(X, Y) = 0.9794$  e entropia conjunta  $H(X, Y) = 1.7502$  . 72
- Figura 35: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #5,(a)  $I_{37}(t)$ ; (b)  $I_{38}(t + 1)$ ; (c)  $I_{37}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = -21.65^\circ$ ,  $S = 1$ ,  $\Delta x = 3$  e  $\Delta y = -24$ , atingindo um  $PSNR = 40.33dB$ . Métrica  $D(X, Y) = 0.9815$  e entropia conjunta  $H(X, Y) = 1.9912$  . . . . . 73
- Figura 36: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #6,(a)  $I_{191}(t)$ ; (b)  $I_{192}(t + 1)$ ; (c)  $I_{191}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = -22.51^\circ$ ,  $S = 1.36$ ,  $\Delta x = -13$  e  $\Delta y = -31$ , atingindo um  $PSNR = 36.59dB$ . Métrica  $D(X, Y) = 0.9848$  e entropia conjunta  $H(X, Y) = 2.4488$  . . . . . 74
- Figura 37: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #7,(a)  $I_{204}(t)$ ; (b)  $I_{205}(t + 1)$ ; (c)  $I_{204}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 3.46^\circ$ ,  $S = 1.18$ ,  $\Delta x = -11$  e  $\Delta y = 3$ , atingindo um  $PSNR = 38.23dB$ . Métrica  $D(X, Y) = 0.9742$  e entropia conjunta  $H(X, Y) = 2.0064$  . 75

## LISTA DE TABELAS

Tabela 1:	Tabela verdade para detecção do movimento da retina. . . . .	33
Tabela 2:	Resultados para detecção do movimento com as 4 medidas de informações testadas para o conjunto treino. . . . .	34
Tabela 3:	Tipos de movimentos detectados entre quadros e a condição de detecção. . . . .	56
Tabela 4:	Tipos de movimentos detectados entre quadros e a condição de detecção para o método comparativo. . . . .	57
Tabela 5:	Resultados experimentais da ANOVA dos parâmetros de detecção do movimento da retina. . . . .	63
Tabela 6:	Resultados experimentais da ANOVA dos parâmetros de estimação do movimento da retina. . . . .	63
Tabela 7:	Resultados comparativos para diferentes figuras de mérito. . . . .	67
Tabela 8:	Resultados comparativos em termos da média do PSNR. . . . .	67

## LISTA DE ABREVIATURAS

ACC	Acurácia
ANOVA	Análise de variância
AMIR	Anormalidades microvasculares intra-retinianas
CCD	<i>Charge-coupled device</i> , dispositivo de carga acoplada
EMD	Edema Macular Diabético
FDM	Figura de mérito
FP	Falso positivo
FN	Falso negativo
GT	<i>Ground truth</i> , conjunto verdade
M	Número de linhas de um quadro da retina
N	Número de colunas de um quadro da retina
<i>pixels</i>	<i>Picture elements</i> , elementos da imagem
PSNR	<i>Peak signal-to-noise rate</i> , taxa de pico sinal-ruído
VEGF	<i>Vascular endothelial growth factor</i> , fator de crescimento do endotélio vascular
VN	Verdadeiro negativo
VP	Verdadeiro positivo
VPP	Valor preditivo positivo

## LISTA DE SÍMBOLOS

$\Sigma$	Somatório
$\tau$	Limiar de detecção de movimento relacionado a métrica de distância
$\eta$	Limiar de detecção de movimento relacionado a entropia conjunta
$\gamma$	Limiar de estimação de movimento relacionado aos parâmetros de translação
$\alpha$	Limiar de estimação de movimento relacionado ângulo de rotação
$\sigma$	Limiar de estimação de movimento relacionado fator de escala
$\theta$	Ângulo de rotação entre dois quadros consecutivos
$S$	Fator de escala entre dois quadros consecutivos
$\Delta x$	Translação horizontal entre dois quadros consecutivos
$\Delta y$	Translação vertical entre dois quadros consecutivos
$\nabla(A)$	Gradiente de uma matriz qualquer $A$ .
$\partial(A)$	Derivada parcial de uma matriz qualquer $A$
$\wedge$	Operador lógico E
$\vee$	Operador lógico OU
$I(t)$	Quadro no instante atual $t$ de um vídeo
$I(t)$	Quadro no instante seguinte $t + 1$ de um vídeo
$X$	Magnitude do gradiente de um quadro no instante atual $t$ de um vídeo
$Y$	Magnitude do gradiente de um quadro no instante seguinte $t + 1$ de um vídeo
$d(X, Y)$	Métrica de informação teórica da distância entre $X$ e $Y$
$D(X, Y)$	Métrica de informação teórica da distância entre $X$ e $Y$ limitada entre os valores $[0, 1]$
$H(X, Y)$	Métrica de informação teórica da entropia conjunta entre $X$ e $Y$
$MI(X, Y)$	Métrica de informação teórica da informação mútua entre $X$ e $Y$
$HC(X, Y)$	Métrica de informação teórica da entropia condicional entre $X$ e $Y$
$[A]^T$	Transposta da matriz $A$

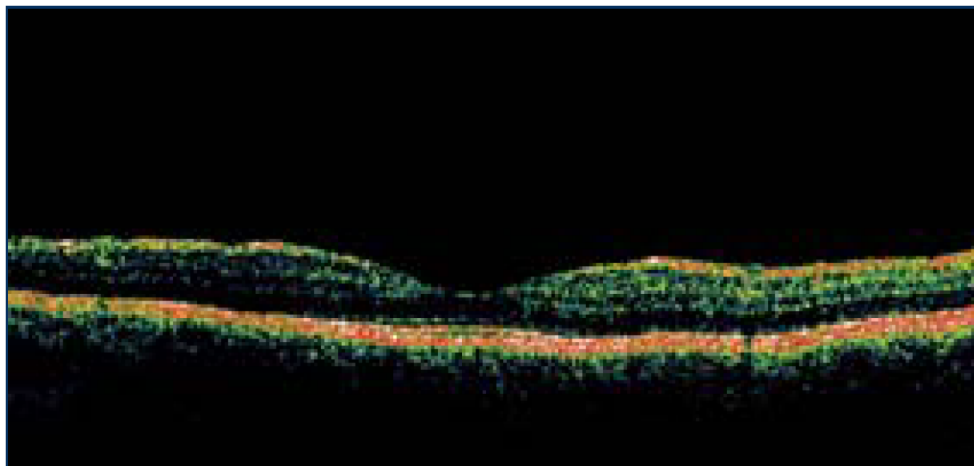
$ A $	Magnitude da matriz A
$prob_t(m)$	Probabilidades de que as magnitudes do gradiente $m$ ocorra no quadro $I(t)$
$prob_{t+1}(n)$	Probabilidades de que as magnitudes do gradiente $n$ ocorra no quadro $I(t + 1)$
$prob_{t,t+1}(m, n)$	Probabilidade conjunta das magnitudes dos gradientes $m$ e $n$ ocorram na posição $(x, y)$ dos quadros $I(t)$ e $I(t + 1)$
$p(x, y)$	Correlação-cruzada
$p_{map}$	Mapa de correlação-cruzada
$f(x, y)$	Quadro no instante atual $t$ de um vídeo no domínio da imagem
$g(x, y)$	Quadro no instante atual $t + 1$ de um vídeo no domínio da imagem
$F(u, v)$	Quadro no instante atual $t$ de um vídeo no domínio de frequência, mesmo que a transformada de Fourier de $f(x, y)$
$G(u, v)$	Quadro no instante atual $t + 1$ de um vídeo no domínio de frequência, mesmo que a transformada de Fourier de $g(x, y)$
$MC(x, y)$	Máscara
$\mathcal{F}(A)$	Transformada de Fourier da matriz A
$\mathcal{F}^{-1}(A)$	Transformada inversa de Fourier da matriz A
$x$	Coordenada horizontal do plano cartesiano
$y$	Coordenada vertical do plano cartesiano
$r$	Coordenada horizontal no espaço polar
$a$	Coordenada vertical no espaço log-polar
$b$	Base exponencial relacionada ao tamanho do quadro de um vídeo da retina
$I$	Imagem de um quadro da retina no plano cartesiano
$I_p$	Imagem de um quadro da retina no espaço log-polar
$M$	Número de linhas de uma imagem
$N$	Número de colunas de uma imagem
$w$	Tamanho da janela de busca
$da$	Distância, ao longo do eixo $a$ , entre o centroide da imagem e o pico da correlação-cruzada
$d(\log_b r)$	Distância, ao longo do eixo $\log_b r$ , entre o centroide da imagem e o pico da correlação-cruzada
$x_c$	Coordenada horizontal do centroide da imagem
$y_c$	Coordenada vertical do centroide da imagem
$x_{pico}$	Coordenada horizontal do pico correlação-cruzada



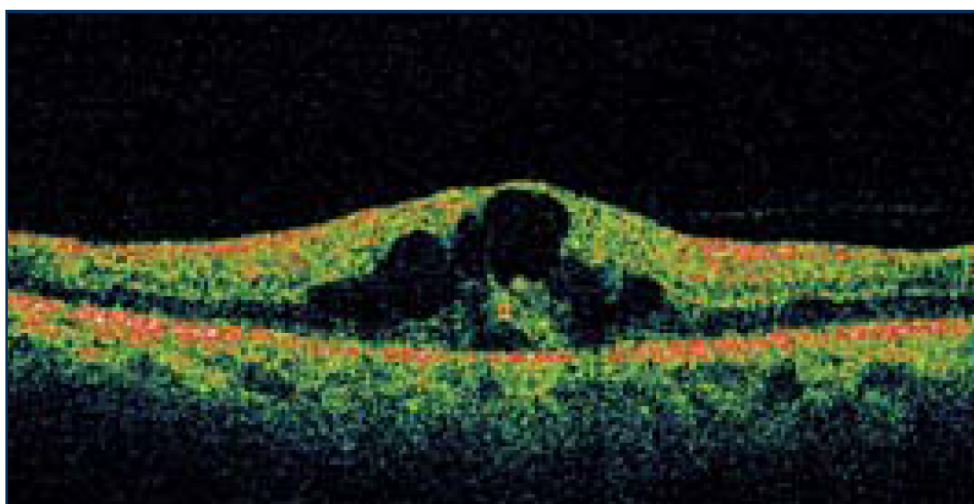
$y_{pico}$	Coordenada vertical do pico correlação-cruzada
$I_{\Delta}(x, y)$	Função impulsiva da correlação-cruzada.
$M1$	Magnitude da transformada de Fourier de $F$
$M2$	Magnitude da transformada de Fourier de $G$
max	Máximo valor de uma função
<i>hipass</i>	Filtro passa alta

# 1 INTRODUÇÃO

Edema macular diabético (EMD) é uma das mais graves consequências da retinopatia diabética e a complicação visual mais comum em pessoas com esta doença. Estudos histopatológicos dos olhos humanos indicam que há uma deposição de material nas camadas internas da retina (KOZAK et al., 2011; KOZAK, 2011; BRESSLER et al., 1994). Além disso, há uma redistribuição de melanina no epitélio pigmentar da retina: as células doentes lançam sua melanina e as células vizinhas recebem (BOULTON et al., 1990). Novos vasos coroidais podem se formar, o que representa um risco grave na perda da visão. Dessa forma, a realização de imagens em infravermelho da retina é necessária na avaliação de tais casos. A Figura 1 ilustra a diferença entre uma retina saudável e uma com *EMD*.



(a)



(b)

Figura 1: (a) Imagem em seção da área macular da retina obtida por Tomografia de Coerência Óptica(OCT). Área macular sem alterações, com depressão normal na zona central da fóvea e sem espessamento da retina; (b) Imagem também obtida por OCT da área macular de um doente com *EMD*. Nota-se o espessamento da área foveal, com perda da depressão foveolar e com áreas de menor refletância (mais escuras) que correspondem a líquido intra retiniano (NOGUEIRA et al., 2007).

No entanto, nem os pequenos depósitos nem as alterações precoces são observáveis com métodos clínicos padrão. Imagens de infravermelho próximo (com comprimento de onda entre  $800nm$  a  $2500nm$ ) são adequadas para a investigação de estruturas sub-retinianas, pois podem revelar a maioria dos detalhes do fundo de olho (ELSNER et al., 1990). Além disso, imagem infravermelha é preferida para a detecção da patologia, visto que é menos agressiva para a retina do paciente. O sistema de imagem em infravermelho também detecta a patologia, apesar da presença de hemorragia ou catarata, o que pode não ser detectado em outros estudos *in vivo*, como angiografia com contraste (ELSNER

et al., 1996). Uma dificuldade com as imagens de fundo de olho pode ser atribuída à incapacidade de separar a luz refletida e a luz espalhada, uma vez que a luz infravermelha é menos absorvida do que a luz visível e, assim, espalha-se por longas distâncias. Consequentemente, as imagens em infravermelho do fundo de olho muitas vezes são ruidosas e apresentam pouca definição de detalhe. Portanto, imagens em infravermelho do fundo de olho de cenas estáticas tiradas em instantes diferentes (ou seja, sequências de imagens de infravermelho) tendem a apresentar flutuações de intensidade dos *pixels* no tempo. Também, tais imagens podem apresentar mudanças substanciais na intensidade da imagem de fundo - *background* - pois por menor que seja o movimento da retina, isso pode causar o espalhamento de luz infravermelha.

Os sinais clínicos observáveis do EMD incluem microaneurismas, hemorragias tipo ponto ou borrão, exsudato e anormalidades microvasculares intra-retinianas (AMIR), e esses sinais podem evoluir com o tempo. É bem conhecido que a detecção precoce pode permitir oportuno tratamento e minimização na deterioração da acuidade visual do paciente (NARASIMHA-IYER et al., 2006). Infelizmente, em alguns casos, a EMD tende a evoluir; e, em tais casos, muitas vezes são usados tratamentos a base de laser para evitar uma maior deterioração na acuidade visual (SCHARDOSIM et al., 2011).

Até o presente momento, a tecnologia de imagem em infravermelho da retina e os tratamentos a laser ainda estão evoluindo. Particularmente, a técnica de laser amarelos têm encontrado importância na fotocoagulação focal (comprimento de onda de aproximadamente  $586nm$ ), por ser uma técnica usual para o tratamento de EMD, além de ser utilizada como um complemento à administração de drogas anti-VEGF (Fator de Crescimento do Endotélio Vascular) (KOZAK, 2011). No tratamento a laser, um perito escolhe as áreas de lesão a serem tratadas interativamente, tentando evitar que a EMD se espalhe nestas áreas. Entretanto, o movimento da retina voluntário (ou mesmo involuntário) pode perturbar o tratamento e as áreas saudáveis também podem ser atingidas, degradando ainda mais a acuidade visual do paciente (SCHARDOSIM et al., 2011). Estimar os movimentos dos olhos durante o tratamento a laser minimiza a possibilidade de queimadura acidental, o que é particularmente importante quando o tratamento está perto da fóvea (KOZAK et al.,

2011). Portanto, a detecção e estimação dos movimentos da retina são importantes etapas nos tratamentos a laser para EMD.

O tratamento com fotocoagulação a laser é o método convencional para tais casos, reduzindo o risco de perda visual moderada em aproximadamente 50% dos olhos tratados (REIS et al., 2010). O uso de fotocoagulação de micropulso tem sido uma terapia a laser alternativa menos destrutiva, com a sua relação risco-benefício mais favorável. O tratamento pode ser justificado precocemente, permitindo a melhoria ou a estabilização de uma função visual menos comprometida (REIS et al., 2010). A Figura 2 mostra a evolução da EMD antes e após o tratamento com injeção intravítrea, onde é possível visualizar a diminuição do espessamento da área foveal.

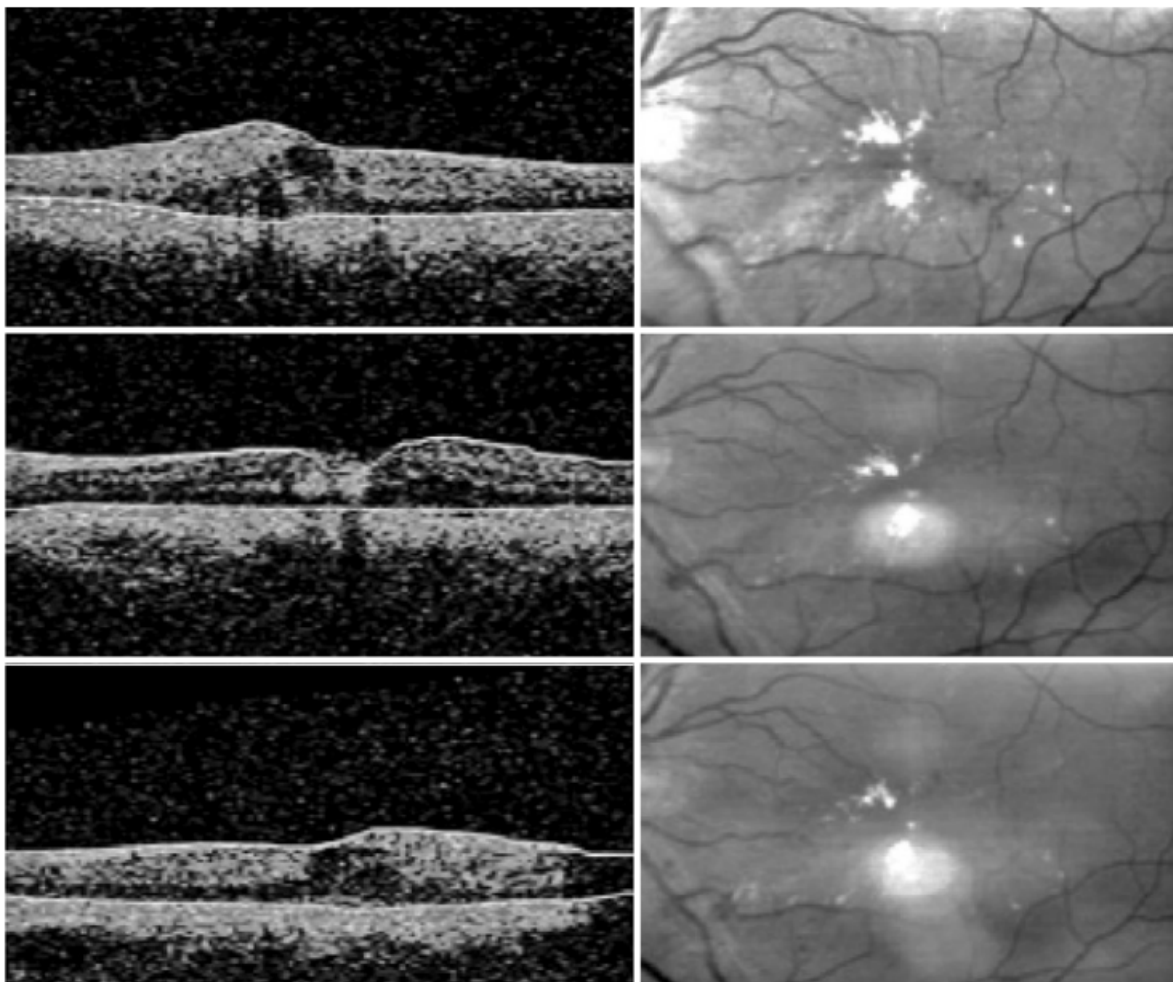


Figura 2: (a) Tomografia de coerência óptica mostrando o EMD antes da injeção intravítrea; (b) com 1 semana; (c) 1 mês após a injeção (REIS et al., 2010).

Uma breve visão geral de um sistema de tratamento a laser comumente utilizado é

apresentado aqui para contextualizar o trabalho. Alguns dos componentes desse equipamento são: sistema de imagem, visor óptico, controle do feixe do laser, unidade do laser e controlador central. A Figura 3 ilustra um médico posicionando o feixe utilizando o visor óptico e o sistema de imagem. O programa permite encontrar a localização das lesões que foram anteriormente detectadas usando uma câmara de fundo ou outro método de diagnóstico. Uma mira é exibida sobre todos os pontos de tratamento, a mudança ocorre de um ponto para outro. Quando o médico tem o feixe na correta posição, o tratamento a laser começa. Os pontos são tratados um por um, até terminar ou até que o processo seja interrompido pelo médico ou automaticamente. Um dos problemas desse tipo de equipamento é que não há a detecção e estimação do movimento da retina, isto é, caso o paciente se mova, pode haver danos em áreas saudáveis.

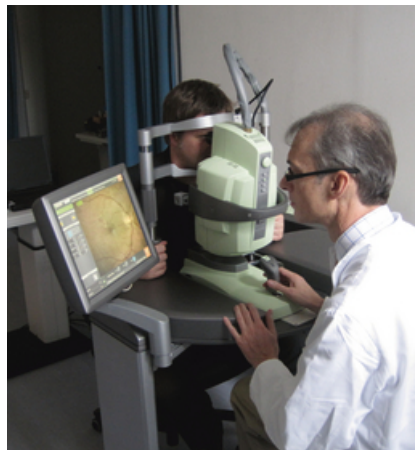


Figura 3: Imagem de fundo de olho e retinógrafo digital de tratamento a laser (KERNT et al., 2011).

Nesta dissertação, apresentamos um método de detecção e estimação de movimento robustas a rotação, escala e translação em sequências de imagens em infravermelho que é adaptável ao ruído de fundo e mudanças de iluminação. Os algoritmos desenvolvidos e apresentados permitem que haja um reposicionamento da mira. Esse sistema de rastreamento ajusta a posição do laser para estimar, isto é, corrigir pequenos movimentos da retina, correspondendo a menos de  $50\mu m$ . Se um movimento maior da retina é detectado, o laser é imediatamente interrompido, como um sistema de segurança. O processo é reiniciado e a tarefa de posicionamento do feixe e o tratamento são retomados. Assim, o algoritmo desenvolvido aumentará a segurança no tratamento do EMD.

Essa abordagem se baseia em informações estatísticas globais em vez de confiar em características locais (como vasos ou estruturas salientes), já que em sequências de imagens em infravermelho estas saliências consomem tempo computacional e podem ser enganosas devido ao ruído e artefatos, não contendo informações de cor e têm menor resolução do que outras modalidades de imagens da retina (exemplo: imagens de fundo de olho a cores). Portanto, vasos, intersecções e outras características importantes não são facilmente rastreadas, uma vez que elas tendem a desaparecer ou serem confundidas com artefatos ao longo dos quadros do vídeo.

Os resultados experimentais indicam que o método proposto potencialmente pode detectar movimentos da retina e estimá-los em sequências de imagens de infravermelho, ajudando a melhorar a confiabilidade dos sistemas de tratamento a laser para EMD. A Figura 4 mostra uma sequência de quadros da retina de um dos vídeos do conjunto de dados utilizados nesta dissertação. Nas 10 imagens são verificados movimentos de rotação, escala e translação entre quadros vizinhos. Entre (a) e (b) há uma translação; entre (c) e (d) há uma translação, escala e rotação; entre (e) e (f) há uma translação, escala e rotação; entre (g) e (h) há uma pequena translação e entre (i) e (j) não há movimento.

Conforme será detalhado nas seções posteriores, as principais contribuições desta dissertação estão relacionadas ao desenvolvimento de algoritmos para detectar e estimar os movimentos da retina durante o tratamento a laser do EMD. A criação destes algoritmos geraram dois artigos. O primeiro é uma publicação na revista internacional *Computerized Medical Imaging and Graphics* (CMIG), (SCHARCANSKI et al., 2013). E a segunda contribuição é um trabalho apresentado e publicado nos anais da conferência realizada em Montevideo (Uruguai) *International Instrumentation and Measurement Technology Conference* (I2MTC 2014), (SCHARDOSIM; SCHARCANSKI, 2014).

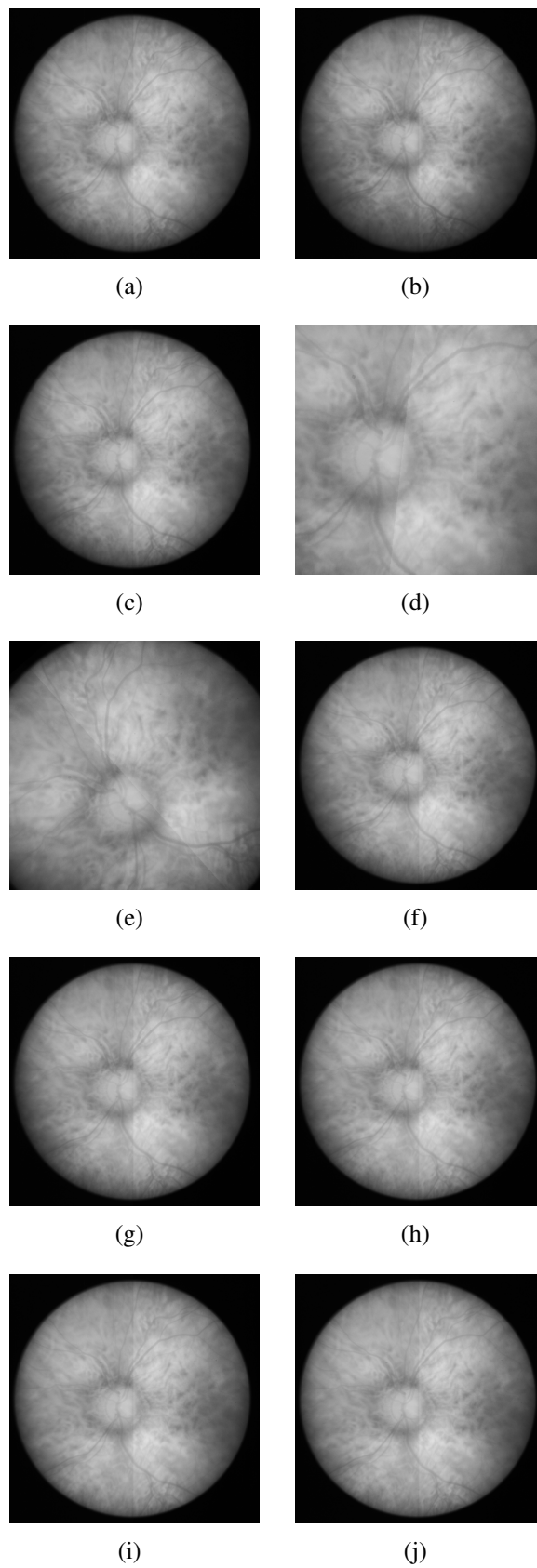


Figura 4: Ilustração de quadros consecutivos do vídeo #3 do conjunto de dados.



## 2 REVISÃO DO ESTADO DA ARTE

Registro (alinhamento) de imagem da retina é um problema conhecido; poucos métodos foram desenvolvidos e alguns deles são baseados na detecção de características locais, tais como vasos ou outras características salientes. No entanto, registro de imagem da retina rápido e preciso ainda é um problema desafiador (CHANWIMALUANG; FAN; FRANSEN, 2006; ZHENG et al., 2009). Além disso, métodos de registro de imagem da retina são usados principalmente para a comparação de imagens, a fim de encontrar alterações locais ou acompanhamento evolutivo de doenças, que estão fora do escopo deste trabalho.

Da mesma forma, o registro de imagem robusto a rotação, escala e translação é um problema relatado e tem sido proposto um grande número de métodos, e muitas vezes estes métodos, os quais muitas vezes são fundamentos no domínio de frequência. Reddy e Chatterji (1996) utilizam uma abordagem no domínio de Fourier para coincidir imagens que são transladadas, rotadas e escaladas em relação uma à outra. Métodos de Fourier diferem de outras estratégias de registro, pois estes pesquisam pela imagem ótima no domínio da frequência. Esse método não é o mais adequado para imagens da retina em infravermelho, pois ele tende falhar na estimativa de grandes movimentos entre quadros, como explicado na Seção 3.7. Outra abordagem para registrar imagens de forma robusta à rotação, escala e translação baseia-se no registro utilizando a transformação log-polar. Wolberg e Zokai (2000) propõem realizar um registro da imagem hierárquica, primeiro usando uma transformação log-polar para trazer as imagens em alinhamento aproximado, mesmo na presença de ângulos arbitrários de rotação e uma vasta gama de mudanças de

escala. Seu propósito é fornecer uma boa estimativa inicial para um módulo de registro com transformação afim, com base na otimização não linear de mínimos quadrados, embora estime os parâmetros de movimento corretamente, o módulo de registro afim não é aplicável em sequências de imagens da retina.

### **3 DETECÇÃO E ESTIMAÇÃO DOS MOVIMENTOS DA RETINA EM SEQUÊNCIAS DE IMAGENS EM INFRAVERMELHO**

Nesta dissertação, o movimento da retina é detectado medindo dissimilaridades entre quadros adjacentes utilizando técnicas fundamentadas em estatísticas e teoria da informação. Diferenças substanciais entre quadros adjacentes são atribuídas ao movimento da retina, uma vez que as estruturas nessas sequências de imagens em infravermelho são supostamente estáticas, conforme detalhado na Seção 3.1. Na Seção 3.2 são analisados os movimentos atípicos na retina. Em seguida, se for detectado movimento, estima-se os parâmetros de deslocamento em termos de rotação, escala e translação entre dois quadros consecutivos, utilizando uma abordagem inspirada nos métodos de deslocamento de fase e transformação log-polar que serão detalhados na Seção 3.3. Um resumo do método é proposto na Seção 3.4. Uma vez que o deslocamento é detectado e os parâmetros estimados, uma classificação do movimento da retina é necessário para determinar se o laser será desligado ou não durante o tratamento de EMD como será detalhado na Seção 3.5. Depois, na Seção 3.6 serão analisados os limiares de detecção e estimação do movimento da retina. Por fim, na Seção 3.7 será apresentado um método comparativo.

#### **3.1 Detecção dos Movimentos da Retina**

Sequências de imagens em infravermelho do fundo de olho são idealmente estáticas e mudanças quadro a quadro não são muito esperadas durante um vídeo. Neste contexto,

grandes dissimilaridades proveem indicações de movimento da retina. Porém, simples métodos de diferenças quadro a quadro, por exemplo subtração entre *pixels* não deve funcionar bem nesta aplicação devido ao ruído e à variação da iluminação de fundo. A abordagem utilizada detecta o movimento da retina em sequências de imagens em infravermelho, com base na comparação de pares de quadros adjacentes usando métricas da teoria de informação, como explicado a seguir.

A iluminação de fundo pode variar substancialmente em sequências de imagens em infravermelho, mas as magnitudes dos gradientes locais são menos suscetíveis a tais variações do que os tons de cinza. Em outras palavras, seja  $I(t)$  um quadro no tempo  $t$  de um vídeo da retina em infravermelho, e a magnitude do gradiente local seja  $|\overrightarrow{\nabla I(t)}|$ . Assumindo que  $I(t)$  e  $I(t + \Delta t)$  são capturados a partir da retina estática nos tempos  $t$  e  $t + \Delta t$ . O gradiente local é:

$$\overrightarrow{\nabla I(t)} = \left[ \frac{\partial I(t)}{\partial x}, \frac{\partial I(t)}{\partial y} \right]^T, \quad (1)$$

e desde que  $\frac{\partial I(t)}{\partial x} = \frac{\partial(I(t)+rd)}{\partial x}$  e  $\frac{\partial I(t)}{\partial y} = \frac{\partial(I(t)+rd)}{\partial y}$ , se a iluminação de fundo variar apenas por uma constante, tal que  $I(t + \Delta t) = I(t) + rd$  (onde  $rd$  é uma constante), é obtido:

$$|\overrightarrow{\nabla I(t)}| = |\overrightarrow{\nabla(I(t) + rd)}|. \quad (2)$$

Portanto, as magnitudes de gradientes locais no instante  $t$  não são afetadas por constantes nos tons de cinza, ou seja, são invariantes para pequenas mudanças de iluminação de fundo, as quais podem ser aproximadas por escalas lineares. A Figura 5 ilustra um par de quadros adjacentes e as respectivas magnitudes do gradiente.

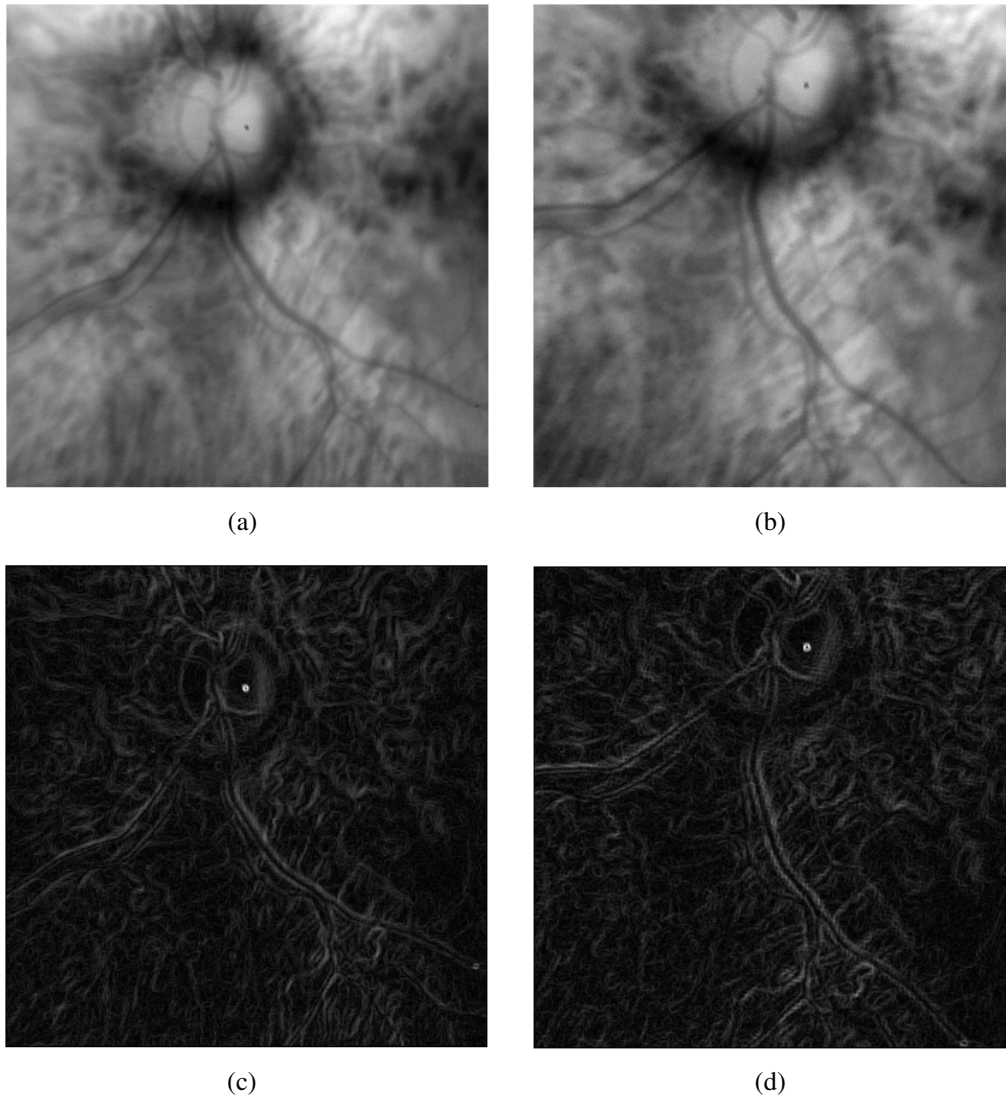


Figura 5: Ilustração de um par de quadros adjacentes e as respectivas magnitudes do gradiente: (a) quadro  $I(t)$ ; (b) quadro  $I(t + 1)$ ; (c) magnitude do gradiente de  $I(t)$ ,  $|\overrightarrow{\nabla I(t)}|$ ; (d) magnitude do gradiente de  $I(t + 1)$ ,  $|\overrightarrow{\nabla I(t + 1)}|$ .

Lembrando que as imagens da retina em infravermelho são ruidosas, as magnitudes do gradiente local  $|\overrightarrow{\nabla I}|$  também serão contaminadas por ruído e este pode ser não-estacionário nestas imagens. Portanto, é necessário um método robusto para medir dissimilaridades em quadros adjacentes. Neste trabalho usa-se uma abordagem de teoria da informação, uma vez que, neste caso, esta teoria não precisa de estimativas do ruído ou das variações de iluminação. Além disso, essa abordagem pode acomodar praticamente qualquer distribuição de probabilidade dos artefatos de imagem. Neste caso, as estatísticas  $X = |\overrightarrow{\nabla I(t)}|$  e  $Y = |\overrightarrow{\nabla I(t + 1)}|$  podem ser estimadas diretamente a partir dos tons de cinza. Para medir as dissimilaridades entre quadros são utilizadas medidas de distri-

buição de estatísticas genéricas. Neste trabalho foram consideradas: a informação mútua  $MI(X, Y)$ , a entropia conjunta  $H(X, Y)$ , a métrica universal  $D(X, Y)$  e a entropia condicional  $HC(X, Y)$ . Para explicar melhor as medidas de dissimilaridades entre quadros, é definida uma notação. Seja  $MI(X, Y)$  a informação mútua entre  $X$  e  $Y$  (KRASKOV et al., 2005):

$$MI(X, Y) = \sum_{m \in X} \sum_{n \in Y} prob_{t,t+1}(m, n) \log \frac{prob_{t,t+1}(m, n)}{prob_t(m)prob_{t+1}(n)}, \quad (3)$$

onde,  $prob_t(m)$  e  $prob_{t+1}(n)$  são as probabilidades de que as magnitudes dos gradientes  $m$  e  $n$  ocorram nos quadros  $I(t)$  e  $I(t + 1)$ , respectivamente, e  $prob_{t,t+1}(m, n)$  é a probabilidade conjunta de que as magnitudes dos gradientes  $m$  e  $n$  ocorram na posição  $(x, y)$  dos quadros  $I(t)$  e  $I(t + 1)$ . Em outras palavras,  $prob_t(m)$  e  $prob_{t+1}(n)$  são vetores com 256 posições cada e representam os histograma de ocorrências das magnitudes  $m$  e  $n$  dos gradientes respectivamente para os quadros  $I(t)$  e  $I(t + 1)$ . Já  $prob_{t,t+1}(m, n)$  é uma matriz de tamanho  $256 \times 256$ . A entropia conjunta  $H(X, Y)$  entre  $X$  e  $Y$  é:

$$H(X, Y) = - \sum_{m \in X} \sum_{n \in Y} prob_{t,t+1}(m, n) \log(prob_{t,t+1}(m, n)), \quad (4)$$

e uma distância entre  $X$  e  $Y$  é (veja os detalhes em KRASKOV et al., 2005):

$$d(X, Y) = H(X, Y) - MI(X, Y). \quad (5)$$

Uma vez que  $H(X, Y) \geq 0$  e  $MI(X, Y) \geq 0$ , então  $d(X, Y) \leq H(X, Y)$ , e uma versão com limites da distância  $d(X, Y)$  na Eq. 5, (KRASKOV et al., 2005):

$$D(X, Y) = \frac{d(X, Y)}{H(X, Y)} \leq 1, \quad (6)$$

e a entropia condicional  $HC(X|Y)$  entre  $X$  e  $Y$  é definida por:

$$HC(X|Y) = H(X, Y) - H(Y), \quad (7)$$

onde,

$$H(Y) = - \sum_{n \in Y} prob_{t+1}(n) \log(prob_{t+1}). \quad (8)$$

A escolha de uma medida de informação teórica adequada para a detecção de movimento da retina em infravermelho foi fundamentada na avaliação experimental de dados reais. A empresa Opto Eletrônica disponibilizou 7 vídeos da retina em infravermelho e esse é o conjunto de dados analisados nesta pesquisa. Os vídeos são compostos de 2710 quadros obtidos por um retinógrafo não midriático digital. Conforme as Figuras 6 - 12 a seguir, que ilustram parte do conjunto de dados, onde é possível verificar as variações de iluminação e movimento entre os quadros. Esse tipo de equipamento é frequentemente utilizado no diagnóstico e tratamento de EMD. Ruído e artefatos da iluminação são comumente encontrados em imagens da retina em infravermelho. Além disso, foi adicionado aleatoriamente em 50% dos quadros um fator de rotação e/ou escala. A adição destes fatores gerou um conjunto de dados onde há rotação, escala e translação nos quadros. A rotação adicionada foi aleatória dentro da faixa de  $[-45^\circ, 45^\circ]$ . Da mesma forma, o fator de escala adicionado aos quadros também foi aleatório dentro do intervalo  $[0.5, 2]$ . Além das translações que ocorrem naturalmente com frequência entre os quadros. Para a determinação dos parâmetros críticos, ruído gaussiano com média zero e variância 0.0001 foram adicionados a cada quadro para aumentar o erro aleatório no conjunto de dados de treinamento.

O conjunto de dados contém 7 sequências de vídeo. A parte inicial de cada um dos 7 vídeos são concatenados para formar o conjunto de treinamento, e os quadros restantes são concatenados para formar o conjunto de teste. Assim, os primeiros 65% dos quadros de cada vídeo foram concatenados e fazem parte dos dados de treinamento, e os 35% finais de cada sequência de vídeo também foram concatenados para serem incluídos no conjunto de dados de teste. O número total de quadros utilizados foram 2710, com 1762 quadros no conjunto de treinamento e 948 quadros no conjunto de teste. Portanto, foram estimados os parâmetros críticos do algoritmo proposto para a detecção e estimação do movimento da retina usando o conjunto de dados de treinamento, mas testado o desempenho do método em um conjunto de dados distintos, como sugerido em (BISHOP, 2006).

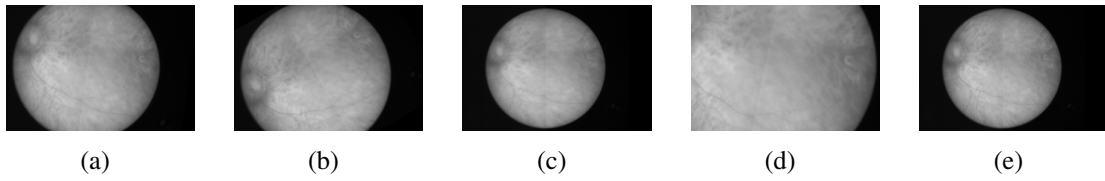


Figura 6: Amostra dos primeiros 5 quadros do vídeo#1 do conjunto de dados.

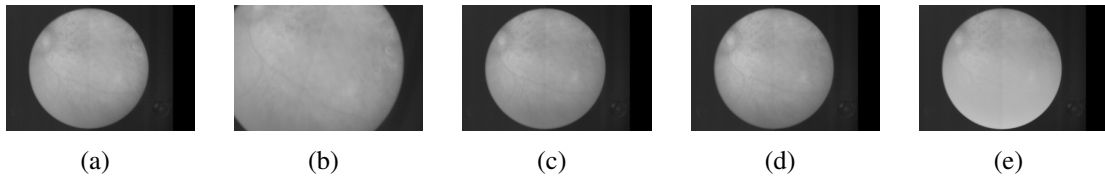


Figura 7: Amostra dos primeiros 5 quadros do vídeo#2 do conjunto de dados.

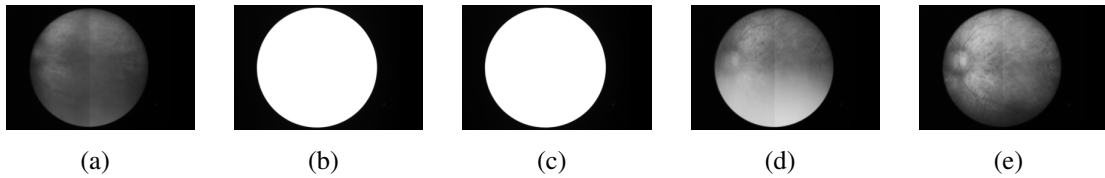


Figura 8: Amostra dos primeiros 5 quadros do vídeo#3 do conjunto de dados.

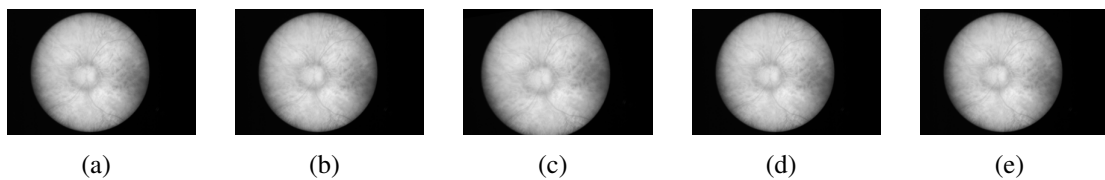


Figura 9: Amostra dos primeiros 5 quadros do vídeo#4 do conjunto de dados.

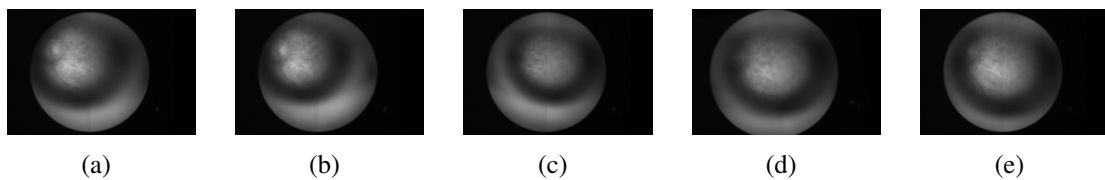


Figura 10: Amostra dos primeiros 5 quadros do vídeo#5 do conjunto de dados.

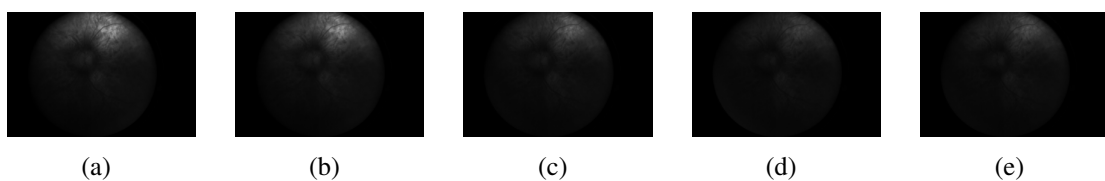


Figura 11: Amostra dos primeiros 5 quadros do vídeo#6 do conjunto de dados.



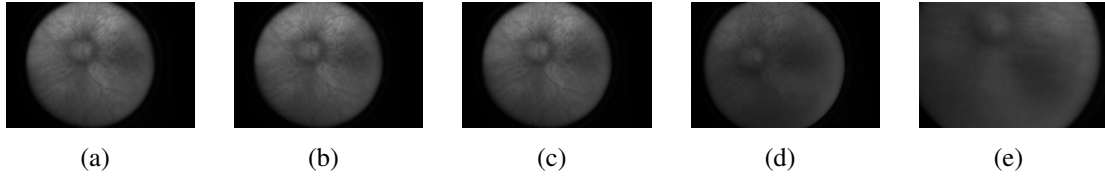


Figura 12: Amostra dos primeiros 5 quadros do vídeo#7 do conjunto de dados.

Foram calculadas as classificações binárias valor preditivo positivo ( $VPP$ ), acurácia ( $ACC$ )<sup>1</sup> e a figura de mérito  $FDM$  definida como a média entre  $VPP$  e  $ACC$ :

$$VPP = \frac{VP}{VP + FP} \quad (9)$$

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (10)$$

$$FDM = \frac{VPP + ACC}{2}. \quad (11)$$

onde  $VP \rightarrow$  Verdadeiro Positivo: se nenhum movimento é detectado (concorda com o *Ground truth*( $GT$ ), conjunto verdade),  $VN \rightarrow$  Verdadeiro Negativo: se o movimento é detectado (concorda com o  $GT$ ),  $FN \rightarrow$  Falso Negativo: se nenhum movimento é detectado (discorda com o  $GT$ ) e  $FP \rightarrow$  Falso Positivo: se o movimento é detectado (discorda com  $GT$ ). Concordar com o  $GT$  implica uma das seguintes situações: (a) dado o  $GT(t, t + 1) = 1$  para um dado par de quadros adjacentes, o algoritmo também não detecta movimento entre os quadros nos tempos  $t$  e  $t + 1$ ; ou (b) o  $GT(t', t' + 1) = 0$  para um outro par adjacente de quadros, e o algoritmo detecta movimento entre quadros nos tempos  $t'$  e  $t' + 1$ .

Na prática, levando em consideração apenas a detecção do movimento da retina, isto é, sem considerar a estimação, as condições acima podem ser resumidas em:  $VP$ , não há movimento da retina e o laser é ligado;  $VN$ , há movimento da retina e o laser não é

<sup>1</sup>Para evitar um erro de terminologia da palavra acurácia no âmbito da instrumentação são utilizados as classificações binárias  $VPP$  e  $ACC$  nas formas comumente utilizadas e definidas em revistas internacionais na área de processamento de imagens (ZHAO et al., 2014; LIU et al., 2012) já que essas classificações não são definidas como medidas quantitativas pelo VIM( Vocabulário Internacional de Metrologia), (JCGM, 2008).

ligado;  $FN$ , não há movimento da retina e o laser não é ligado e  $FP$ , há movimento da retina e o laser é ligado. Das 4 possibilidades, a única que é prejudicial ao paciente é a situação de  $FP$ , ver Tabela 1. Conforme está detalhado nas Seções 3.3 e 3.5, é possível detectar e estimar o movimento da retina de forma que o laser seja reposicionado para o local correto.

Tabela 1: Tabela verdade para detecção do movimento da retina.

	Há movimento	Não há movimento
Laser ligado	FP	VP
Laser desligado	VN	FN

As classificações binárias ou taxas de acerto  $VPP$  e  $ACC$  foram obtidas com todas as 4 medidas de informações (Eqs. 3, 4, 6 e 7) e comparando os resultados de detecção de movimento da retina com o conjunto verdade ( $GT$ ), onde  $GT$  é uma sequência binária construída por inspeção visual dos vídeos (quadro a quadro). Para cada quadro do vídeo o  $GT$  apenas informa se há movimento perceptível ('0') ou não ('1'), isto é, se não houver movimento perceptível e a cena local do vídeo for estática. Um programa simples de visualização de imagem foi utilizado para observar as sequências de imagens, quadro a quadro, para determinar se o movimento ocorreu entre os quadros atual e o quadro anterior. As posições dos pontos de fácil identificação nas imagens em infravermelho foram observadas para verificar a existência de movimento. O processo de obtenção do  $GT$  apenas considera se houve mudanças posicionais ou não da retina entre os quadros.

Neste trabalho, o movimento da retina é detectado quando o conjunto de *pixels* correspondente às estruturas da retina muda entre os tempos  $t$  e  $t + \Delta t$ . Nestas sequências de imagens da retina em infravermelho a iluminação varia de um quadro para outro. Como os gradientes locais são menos afetados pelas mudanças da iluminação do que tons de cinza, são utilizadas as magnitudes dos gradientes (em vez de intensidades dos tons de *pixels*) para representar a informação de *pixel* local quando compara-se quadros adjacentes para detecção do movimento da retina. Dessa forma, a abordagem proposta detecta o movimento da retina na sequência de imagens em infravermelho, medindo a dissimilaridade entre quadros ruidosos adjacentes utilizando técnicas fundamentadas em estatísticas e teoria da informação. Neste caso, diferenças substanciais entre quadros adjacentes são

atribuídos ao movimento da retina, uma vez que as estruturas nestas imagens da retina são supostamente estáticas. Em seguida, como será detalhado na Seção 3.3, se o movimento for detectado, são estimados os parâmetros de deslocamento de movimento com base em uma abordagem que utiliza uma busca iterativa de transformação log-polar + deslocamento de fase.

A Tabela 2 resume os resultados obtidos para detecção do movimento com as 4 medidas de informações testadas para o conjunto treino.

Tabela 2: Resultados para detecção do movimento com as 4 medidas de informações testadas para o conjunto treino.

Medida de informação	VPP(%)	ACC(%)	Limiar otimizado
$H(X, Y)$	70.89	61.57	1.56
$MI(X, Y)$	100	39.68	0.26
$D(X, Y)$	94.79	68.31	0.89
$HC(X Y)$	67.87	50.73	1.56

Comparando todas as medidas de informação testadas,  $MI(X, Y)$  apresenta o maior  $VPP$  e a menor  $ACC$  (ver Tabela 2). Isso significa que  $MI(X, Y)$  é muito sensível, mas não é prático para detecção de movimento, uma vez que raramente detecta cenas estáticas. Por outro lado, a medida  $D(X, Y)$  oferece o melhor compromisso entre o  $VPP$  e a  $ACC$ , e foi escolhido para método de detecção de movimento. A medida de entropia conjunta  $H(X, Y)$  responde bruscamente para movimento inconsistente, conforme será visto na Seção 3.2. Os limites de detecção foram otimizados para cada medida de informação testada para obter o valor máximo da  $FDM$ .

A quantidade  $D(X, Y)$  satisfaz as propriedades básicas de uma métrica, isto é, a desigualdade do triângulo, a não-negatividade, a indiscernibilidade e a simetria (KRASKOV et al., 2005). Portanto, se  $D(X, Y) \rightarrow 1$  então  $X$  e  $Y$  são maximamente distintos. Ou seja, se a métrica de distância  $D(X, Y) \rightarrow 1$ ,  $X = |\overrightarrow{\nabla I(t)}|$  e  $Y = |\overrightarrow{\nabla I(t+1)}|$  são significativamente diferentes. Nesse caso, as diferenças significativas são associadas ao movimento da retina entre os quadros  $t$  e  $t + 1$ . Caso contrário, se  $D(X, Y) \rightarrow 0$ ,  $X = |\overrightarrow{\nabla I(t)}|$  e  $Y = |\overrightarrow{\nabla I(t+1)}|$  são muito semelhantes e nenhum movimento da retina é estimado entre quadros  $t$  e  $t + 1$ . Há exceções para o pressuposto que as diferenças entre quadros são de-

vidas ao movimento da retina, como a piscada de olhos, mas estes são eventos ocasionais e apresentam características específicas. Por exemplo, o piscar de olhos aparece como uma transição brusca entre os quadros, e o quadro em que o olho está fechado aparece brilhante e homogêneo. Assim, tais exceções podem ser também tratadas.

A entropia conjunta  $H(X, Y)$ , Eq. 4, aplicada a dois quadros consecutivos é máxima se estes dois quadros são significativamente diferentes um do outro. A informação mútua é uma medida de quanta informação é compartilhada entre dois quadros, e tem um valor máximo quando os dois quadros analisados são muito semelhantes. A métrica universal definida pela Eq. 6 é limitada em  $[0, 1]$ , e fornece uma estimativa da distância entre os dois quadros analisados. Logo, a métrica  $D(X, Y)$  tem valores próximos de 1 para quadros muito distintos e tende a 0 para os quadros semelhantes.

### 3.2 Identificação de Inconsistências nos Movimentos da Retina

Situações inconsistentes (isto é, atípicas) geram dissimilaridades entre quadros que podem originar situações diferentes de movimento da retina. Tais situações especiais normalmente apresentam características, permitindo identificar o movimento da retina como inconsistente. Por exemplo, piscada de olhos é caracterizada por um quadro subitamente brilhante na sequência de imagem da retina. Imagens fora de foco, causadas por movimento da cabeça, também mostram características particulares como grandes valores de dissimilaridades repentinas entre quadros adjacentes.

Nesses casos, as medidas de informações  $H(X, Y)$  e  $D(X, Y)$  possuem comportamentos distintos. É observado experimentalmente que: (a)  $H(X, Y)$  é menos sensível a pequenos movimentos do que  $D(X, Y)$ , mas a métrica universal é menos sensível a grandes dissimilaridades causadas, por exemplo, por piscada de olhos ou outra situação atípica; (b)  $D(X, Y)$  responde de forma igual para pequenos e grandes movimentos, gerando muitos alarmes falsos em situações de movimentos atípicos; e (c)  $D(X, Y)$  aumenta enquanto  $H(X, Y)$  diminui quando um movimento local ocorre. A abordagem para detectar as dissimilaridades locais causadas por fatores além dos movimentos triviais da retina (isto é, quando imagens capturadas bruscamente tornam-se fora de foco, piscadas de olhos

ou imagens borradas) consiste em observar  $H(X, Y)$  e  $D(X, Y)$ , simultaneamente. Tais situações podem ser discriminadas como segue: (a)  $H(X, Y) \leq \eta$  e  $D(X, Y) > \tau$  podem indicar inconsistência no movimento da retina, onde  $\eta$  é uma constante experimental (veja a Seção 4 adiante); (b)  $H(X, Y) > \eta$  e  $D(X, Y) > \tau$  indica possivelmente um movimento ordinário da retina; e (c)  $H(X, Y) > \eta$  e  $D(X, Y) \leq \tau$  indica provavelmente um movimento pequeno ou não notável da retina ( $H(X, Y)$  diminui em tais situações, mas ainda,  $H(X, Y) > \eta$ ). Na Seção 3.5, a seguir, é desenvolvida uma classificação absoluta dos movimentos da retina utilizando os parâmetros estimados. Um exemplo de movimento atípico está ilustrado na Figura 13 adiante, cujos valores da entropia conjunta  $H(X, Y)$  na proximidade de uma piscada de olhos estão retratados na Figura 14 na sequência. Para o quadro anterior a piscada de olhos (quadro 149),  $D(X, Y) \leq \tau$  indicando que não há movimento da retina nesta parte da sequência. Ainda assim,  $H(X, Y) \leq \eta$  para a mesma região devido às intensidades dos *pixels* dos quadros adjacentes serem muito semelhantes (isto é, as intensidades de todos os *pixels* tendem a aumentar durante a piscada de olhos), e esta inconsistência é detectada.

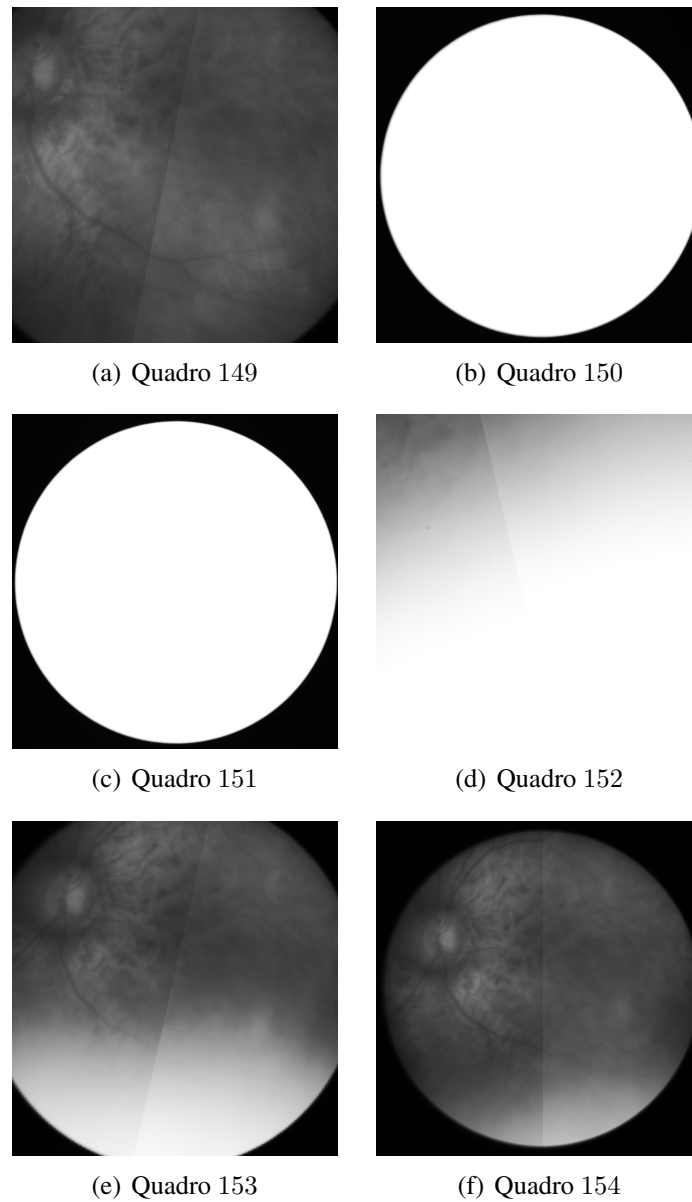


Figura 13: Ilustração de uma piscada de olhos. Os quadros 149 – 154 de 6 quadros em infravermelho da retina são descritos acima como (a) - (f), respectivamente; movimento da retina é detectado como inconsistente para os quadros adjacentes 149 – 153, ou seja, (b)-(d). O piscar de olhos ocorre nos quadros 150 – 152, onde a entropia conjunta calculado é  $H(X, Y) \leq \eta$ .

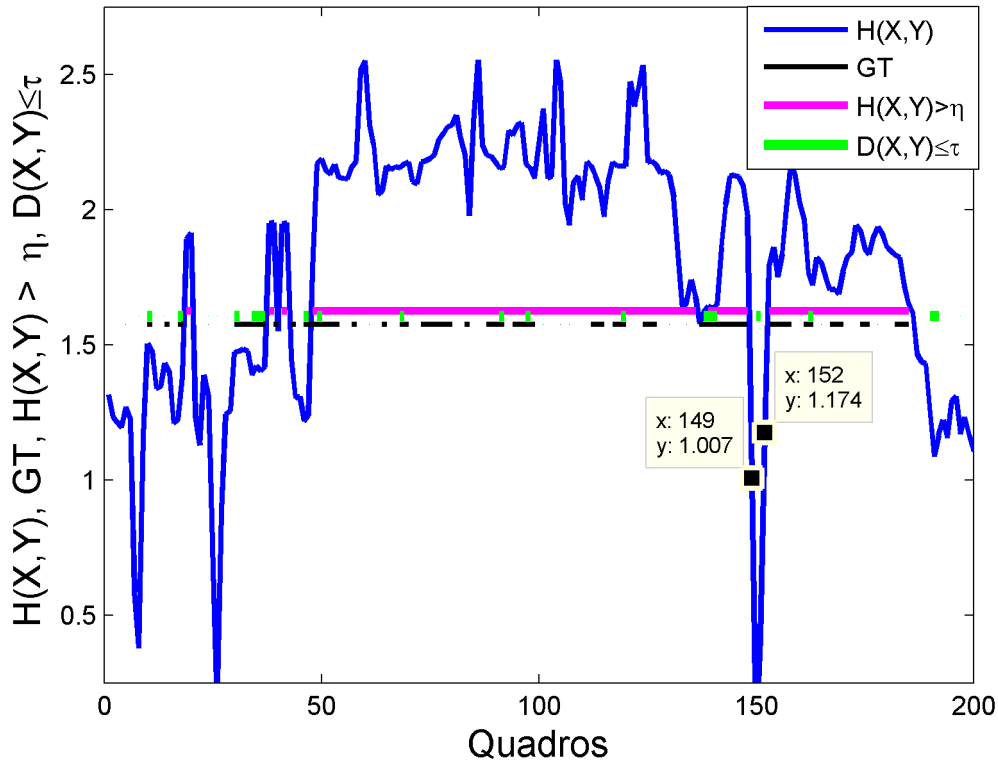


Figura 14: Ilustração dos valores da entropia conjunta  $H(X,Y)$ ,  $GT$ ,  $H(X,Y) > \eta$ ,  $D(X,Y) \leq \tau$  para uma sequência em infravermelho na proximidade de uma piscada ilustrada na Figura 13. A piscada de olhos começa no quadro 149 e termina no quadro 153.

### 3.3 Estimação dos Movimentos da Retina

Uma vez que o movimento da retina é detectado utilizando as medidas de informações mencionadas, Eqs. 3, 4 e 6, é necessário estimar o deslocamento da retina entre quadros consecutivos. Supondo que o movimento da retina seja detectado no quadro  $I(t+1)$  com respeito ao quadro  $I(t)$ , para estimar este movimento da retina em  $I(t+1)$  é necessário estimar o movimento em termos dos parâmetros de rotação ( $\theta$ ), escala ( $S$ ) e translação ( $\Delta x, \Delta y$ ). Obtidos os parâmetros de movimento, a imagem  $I(t+1)$  pode ser corrigida por  $-\theta, 1/S, -\Delta x$  e  $-\Delta y$ , então o quadro  $I(t)$  e o quadro  $I(t+1)$  corrigido estarão alinhados. Para estimar os parâmetros de movimento  $\theta, S, \Delta x$  e  $\Delta y$ , a combinação de dois métodos resumidos nos Algoritmos 1 e 2 é utilizada (p. 42 e 53). Translações simples da retina em um par consecutivo de quadros são estimados com o método de deslocamento de fase. A propriedade de deslocamento de fase da transformada de Fourier diz que um

deslocamento espacial entre duas funções implica um deslocamento de fase no domínio da frequência (REN; JIANG; VLACHOS, 2010). Seja  $f(x, y) \equiv I(x, y, t)$ , Figura 15(a), e  $g(x, y) \equiv I(x, y, t + 1)$ , Figura 15(b), sendo  $M \times N$  o tamanho dos quadros de um vídeo em infravermelho da retina ( $M$  é o número de linhas e  $N$  é o número de colunas), onde  $g(x, y) = f(x + \Delta x, y + \Delta y)$ , e  $F(u, v)$ , Figura 15(c), e  $G(u, v)$ , Figura 15(d), são as suas transformadas de Fourier discretas, respectivamente. Neste caso, no domínio da frequência é obtido:

$$F(u, v) = G(u, v) \exp\left[-2\pi i \left(u \frac{\Delta x}{M} + v \frac{\Delta y}{N}\right)\right], \quad (12)$$

que implica:

$$P(u, v) = \frac{F(u, v)G^*(u, v)}{|F(u, v)G^*(u, v)|} = \exp\left[-2\pi i \left(u \frac{\Delta x}{M} + v \frac{\Delta y}{N}\right)\right], \quad (13)$$

onde  $'*$ ' denota o complexo conjugado, e  $P(u, v)$  é o espectro de potência cruzado dos dois quadros. Se os quadros  $g(x, y)$  e  $f(x, y)$  possuem informações idênticas, mas visualmente deslocadas, a transformada inversa de Fourier de  $P(u, v)$  dará a correlação-cruzada normalizada:

$$p(x, y) = \mathcal{F}^{-1}(P). \quad (14)$$

Idealmente,  $p(x, y)$  contém uma função delta, um único pico em  $(\Delta x, \Delta y)$  e a posição deste pico representa o deslocamento translacional entre  $g(x, y)$  e  $f(x, y)$  no domínio espacial. Foi verificado experimentalmente que o ruído tende a aparecer na transformada inversa de Fourier de  $P(u, v)$ , portanto  $p(x, y)$ , ao longo do eixo horizontal ( $x$ ) e vertical ( $y$ ), mas perto do centro da imagem. Para reduzir o efeito do ruído na estimativa do  $(\Delta x, \Delta y)$ , antes de computar  $I_{\Delta}(x, y)$  com a transformada inversa de Fourier  $p(x, y)$  é aplicada uma máscara  $MC(x, y)$  em  $p(x, y) = \mathcal{F}^{-1}(P(u, v))$ , onde  $\mathcal{F}^{-1}$  denota a transformada inversa de Fourier e  $MC(x, y)$  possui as mesmas dimensões de  $p(x, y)$ :



$$MC(x, y) = \begin{cases} 0 & , \text{ Se } x_c - \kappa_x \leq x \leq x_c + \kappa_x \text{ e } y = 0; \\ 0 & , \text{ Se } y_c - \kappa_y \leq y \leq y_c + \kappa_y \text{ e } x = 0; \\ 1 & , \text{ Caso contrário;} \end{cases} \quad (15)$$

onde,  $\kappa_x = 3$  e  $\kappa_y = 3$  nos experimentos realizados e  $(x_c, y_c)$  são as coordenadas do centroide de  $MC$ .

Onde a função impulsiva da correlação-cruzada é:

$$I_{\Delta}(x, y) = MC(x, y) \cdot p(x, y). \quad (16)$$

Contudo, na prática, os quadros  $g(x, y)$  e  $f(x, y)$  são ruidosos, e  $I_{\Delta}(x, y)$  pode possuir vários picos além do pico dominante na posição  $(\Delta x, \Delta y)$ . Uma vez que  $I_{\Delta}(x, y)$  usualmente contém vários picos, entre os picos mais altos do conjunto  $\{\max_{I_{\Delta}}\} = \{k \cdot \max_{\forall x, y} |I_{\Delta}(x, y)|, \max_{\forall x, y} |I_{\Delta}(x, y)|\}$  é selecionado o de menor norma Euclidiana (correspondendo à estimativa mais conservadora dos parâmetros de translação  $(\Delta x, \Delta y)$ ). O valor da constante nos experimentos realizados é  $k = 0.95$ . Em outras palavras, são selecionados 5% dos picos mais altos e para cada um desses é calculada a norma Euclidiana da sua distância em relação ao centro da imagem e é escolhido o de menor norma.

Então, o deslocamento espacial translacional  $(\Delta x, \Delta y)$  é estimado como segue:

$$(\Delta x, \Delta y) = \operatorname{argmin}_{\forall x, y \in \{\max_{I_{\Delta}(x, y)}\}} \|I_{\Delta}(x, y)\|, \quad (17)$$

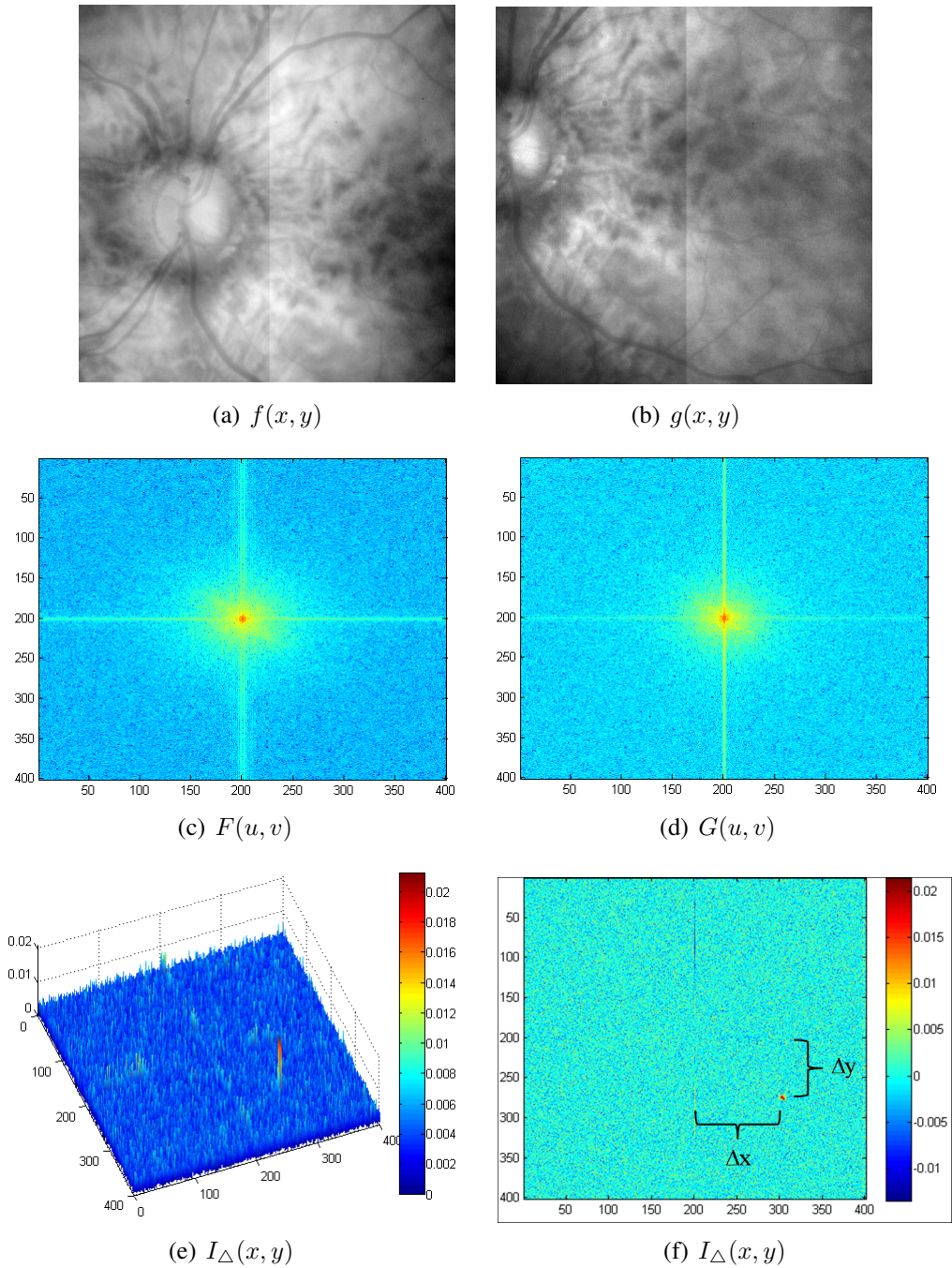


Figura 15: Ilustração do método de deslocamento de fase para estimar os parâmetros de translação. (a)  $f(x, y) \equiv I(x, y, t)$ ; (b)  $g(x, y) \equiv I(x, y, t + 1)$ ; (c) Transformada de Fourier de  $f(x, y)$ ,  $F(u, v)$ ; (d) Transformada de Fourier de  $g(x, y)$ ,  $G(u, v)$ ; (e) Função impulsiva correlação-cruzada  $I_{\Delta}(x, y)$ ; (f) Vista superior da função impulsiva correlação-cruzada  $I_{\Delta}(x, y)$ .

---

**Algoritmo 1:** Método de deslocamento de fase para estimação dos parâmetros de translação.

---

**Entrada:** Um par de quadros consecutivos

**Saída:** Parâmetros de translação:  $\Delta x$ ,  $\Delta y$  e função impulsiva da correlação-cruzada  $I_{\Delta}(x, y)$ .

- 1 Sejam  $f \leftarrow I(t)$  e  $g \leftarrow I(t + 1)$  um par de quadros consecutivos;
  - 2 Suas correspondentes transformadas de Fourier  $F \leftarrow \mathcal{F}(f)$  e  $G \leftarrow \mathcal{F}(g)$ ;
  - 3 Calcula o espectro de potência cruzado usando a Eq. 13;
  - 4 Calcula a transformada inversa de Fourier  $p \leftarrow \mathcal{F}^{-1}(P)$ , a correlação-cruzada;
  - 5 Criação da máscara  $MC(x, y)$ ;
  - 6 Calcula a função impulsiva  $I_{\Delta} \leftarrow MC(x, y) \cdot p(x, y)$ ;
  - 7 Classifica os picos mais altos no conjunto:  
 $\{\max_{I_{\Delta}}\} \leftarrow \{0.95 \cdot \max_{\forall x, y} |I_{\Delta}(x, y)|, \max_{\forall x, y} |I_{\Delta}(x, y)|\}$ ;
  - 8 Seleciona os picos com a menor norma euclidiana do deslocamento;
  - 9 Parâmetros de translação são  $(\Delta x, \Delta y) \leftarrow \operatorname{argmin}_{\forall x, y \in \{\max_{I_{\Delta}}\}} \|I_{\Delta}(x, y)\|$ ;
  - 10 **retorna**  $\Delta x, \Delta y$  e  $I_{\Delta}(x, y)$ .
- 

Se não houver indicação clara de simples translação da retina, ou seja, se  $\max(I_{\Delta}(x, y)) \leq \lambda$ , os parâmetros de rotação, escala e translação são estimados utilizando a transformação log-polar (WOLBERG; ZOKAI, 2000). Em seguida, é descrita a transformação de coordenadas log-polar e o uso dos parâmetros estimados de movimento da retina. Considere o sistema de coordenadas log-polar  $(\rho, a)$ , onde  $\rho$  denota logaritmo da distância radial  $r$  a partir do centroide da imagem  $(x_c, y_c)$ ,  $a$  denota o ângulo e  $b$  é a base do logaritmo que está relacionada com o tamanho da imagem, veja a Figura 16. Quaisquer pontos  $(x, y)$  podem ser representados em coordenadas log-polares:

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2}, \quad (18)$$

$$x = r \cdot \cos(a), \quad (19)$$

$$y = r \cdot \sin(a), \quad (20)$$

$$\rho = \log_b r. \quad (21)$$

$$a = \tan^{-1}\left(\frac{y - y_c}{x - x_c}\right). \quad (22)$$

$$b = \frac{\exp(x_c)}{x_c}. \quad (23)$$

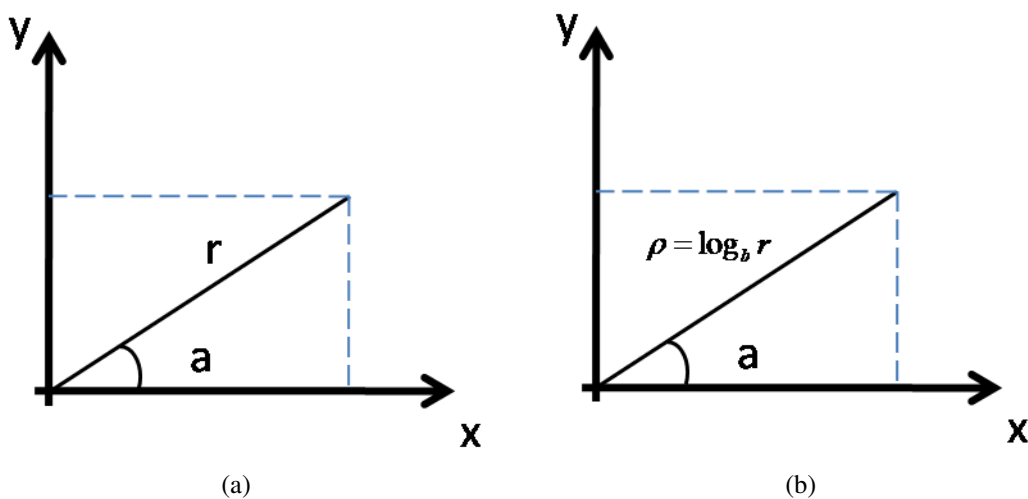
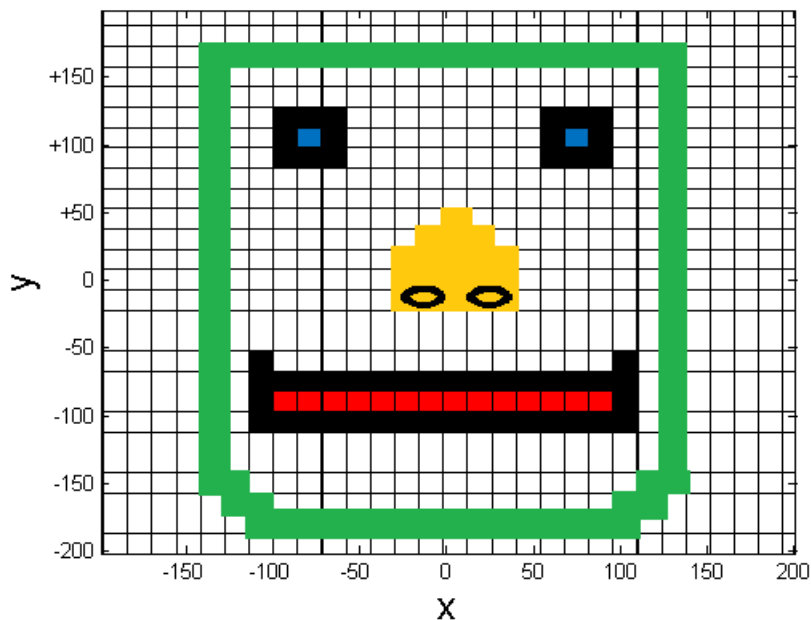


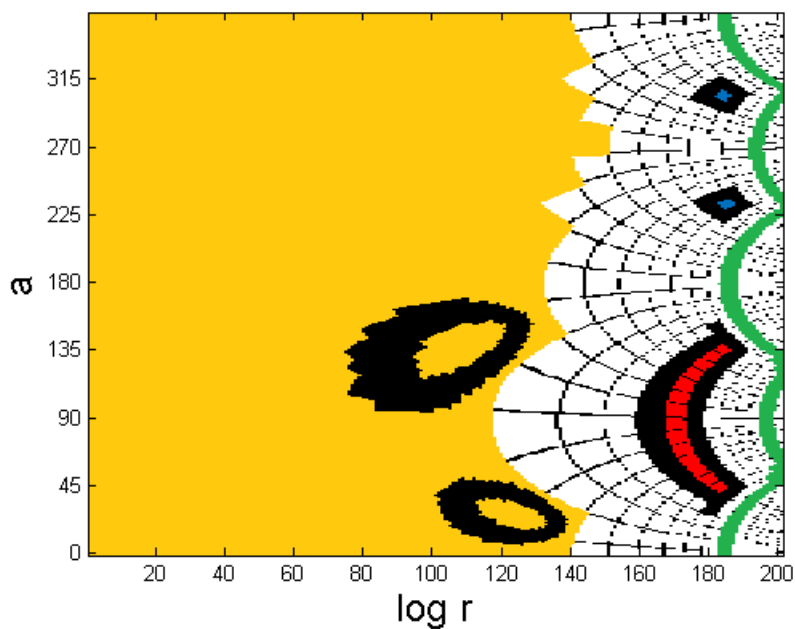
Figura 16: Ilustra a transformação polar e log-polar. (a) Qualquer ponto no espaço  $(x, y)$  pode ser representado no espaço polar  $(r, a)$ ; (b) Qualquer ponto no espaço  $(r, a)$  pode ser representado no espaço log-polar  $(\rho, a)$ .

Aplicando uma transformação de coordenadas log-polar (WOLBERG; ZOKAI, 2000) a uma imagem  $I$ , linhas radiais no espaço cartesiano são mapeadas para linhas horizontais no espaço coordenado log-polar, essa imagem transformada é chamada de  $I_p$ . Supondo que  $\log_b r$  e  $a$  são os eixos horizontal e vertical, respectivamente. A vantagem deste novo espaço de coordenadas é que as mudanças de escalas e de rotações podem ser induzidas por simples modificações dos dados em  $(\log_b r, a)$ . Para exemplificar a transformação log-polar é exibida na Figura 17(a) uma ilustração simples representando um rosto no domínio cartesiano  $(x, y)$ . Já a Figura 17(b) a seguir ilustra a mesma no espaço log-

polar. De forma semelhante, a imagem do babuíno, representada pela Figura 18(a) e a sua respectiva contraparte no domínio log-polar, Figura 18(b), na sequência.



(a)

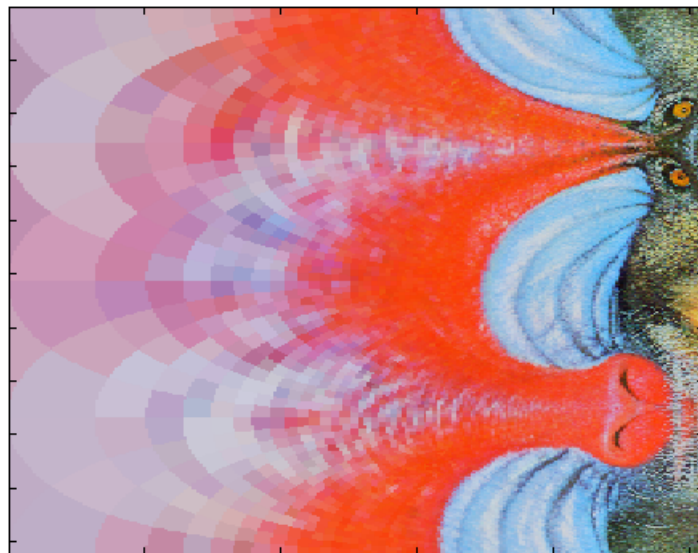


(b)

Figura 17: Exibe a transformação log-polar para uma figura que representa um rosto onde é possível verificar a transposição dos elementos de uma face: a boca em vermelho, o nariz em amarelo, os olhos em azul e o contorno em verde. (a) é o rosto no domínio cartesiano  $(x, y)$ ; (b) é o rosto no domínio log-polar  $(\log r, a)$ .



(a)



(b)

Figura 18: Exibe a transformação log-polar para a imagem de um babuíno. (a) é a imagem do babuíno no domínio cartesiano  $(x, y)$ ; (b) é a imagem do babuíno no domínio log-polar  $(\log r, a)$ .

Seja uma imagem da retina  $I$  representada pela Figura 19(a), e a sua versão rotada, Figura 19(c) - ambas no espaço cartesiano - e suas respectivas contrapartes no domínio log-polar, Figuras 19(b) e 19(d), pode-se verificar que há apenas uma translação ao longo do eixo  $a$  entre as imagens transformadas. Em outras palavras, dados armazenados ao longo das linhas no espaço  $(\log_b r, a)$  representam dados radiais em  $(x, y)$ . Rotar a imagem  $I$ , Figura 19(a), é equivalente a mover as linhas acima ou abaixo na imagem transformada  $I_p$ , Figura 19(b). A Figura 19(c) mostra a imagem  $I$  rotada e o resultado no

domínio log-polar é a Figura 19(d). Note que 50 linhas foram deslocadas para baixo na Figura 19(b) (em uma imagem de 401 linhas), e a imagem na Figura 19(a) foi rotada de  $45^\circ$  (isto é,  $\theta = \frac{50}{401} \cdot 360^\circ \sim 45^\circ$ ). Dessa forma, é notável que as Figuras 19 (b) e (d) diferem apenas por uma translação na direção do eixo vertical.

Considere duas imagens  $I_1$  e  $I_2$  e as suas contrapartes no espaço log-polar  $I_{1p}$  e  $I_{2p}$ . Se  $I_2$  é uma versão girada  $I_1$ , então  $I_{2p}$  é uma versão deslocada de  $I_{1p}$  ao longo do eixo  $a$ . Através da aplicação de correlação-cruzada entre  $I_{1p}$  e  $I_{2p}$  pode ser encontrado o melhor deslocamento  $da$  que ajuste as duas imagens. O deslocamento  $da$  é a distância em *pixels* entre o centroide da imagem e a posição do pico  $(x_{pico}, y_{pico})$  na função impulsivo de correlação-cruzada na equação Eq. 16,  $da = y_c - y_{pico}$ . A correlação-cruzada é comumente usada para encontrar os movimentos translacionais pelo método de deslocamento de fase entre duas imagens, mas este método não funciona bem na presença de escala ou rotação. No entanto, no espaço de coordenadas log-polar, encontrar as componentes de translação  $da$  e  $d(\log_b r)$  entre  $I_{1p}$  e  $I_{2p}$  corresponde a encontrar os parâmetros de rotação entre  $I_1$  e  $I_2$  no espaço cartesiano.

Assim, é possível encontrar a mudança de escala entre dois quadros em sequências de imagens da retina em infravermelho. Considere uma imagem  $I_2$  sendo uma ampliação de duas vezes da imagem  $I_1$ . Para cada ponto  $(x, y)$  em  $I_1$  é mapeado para  $(2x, 2y)$  em  $I_2$ . Para determinar o fator de escala é utilizado o logaritmo. No espaço log-polar  $(x, y) \rightarrow (2x, 2x)$ , e  $(2x, 2y) \rightarrow (\log 2x, \log 2y) \rightarrow (\log x + \log 2, \log y + \log 2)$ , agora se torna evidente que no domínio log-polar a introdução de um fator de escala manifesta-se como um deslocamento de fase em uma imagem transformada no log-polar. Então, em vez de mapear a imagem no espaço de coordenadas  $(x, y)$ , é útil mapear isso no sistema de coordenadas  $(\log_b r, a)$  usando a transformação log-polar que agora pode determinar os melhores deslocamentos  $d(\log_b r)$  e  $da$  que se traduzem em escala e em rotação no espaço cartesiano  $(x, y)$ , onde o deslocamento  $d(\log_b r) = b^{(x_c - x_{pico})}$ .

Dado um par de quadros consecutivos  $I(t)$  na Figura 20(a) e  $I(t + 1)$  na Figura 20(c), onde  $I(t + 1)$  é uma ampliação de 2 vezes e uma rotação de  $30^\circ$  e não há translação, suas contrapartes no espaço log-polar são as Figuras 20(b) e 20(d), respectivamente. No

espaço log-polar elas diferem apenas de uma translação simples ao longo dos eixos  $\log_b r$  e  $a$ . Isso significa que no domínio log-polar elas diferem por uma translação e quando uma correlação-cruzada é aplicada a essas imagens, o deslocamento de fase calculado reflete com precisão o fator de escala de 2 vezes e a rotação de  $30^\circ$  que representam a transformação entre os dois quadros da retina em infravermelho.

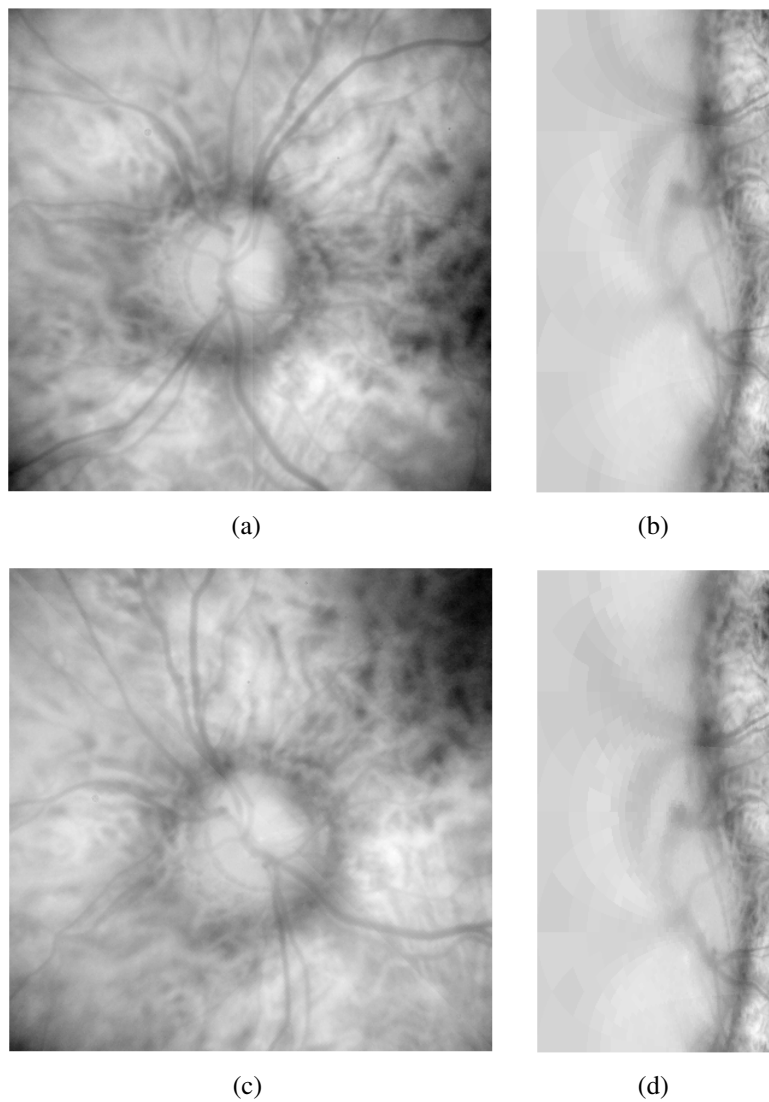


Figura 19: Ilustra o efeito de uma rotação entre dois quadros no domínio log-polar. (a) representa o quadro  $I$  no plano cartesiano; (b) é o quadro  $I$  depois da transformação log-polar,  $I_p$ ; (c) é  $I$  rotada por  $45^\circ$ ; (d) é a sua contraparte no espaço log-polar.



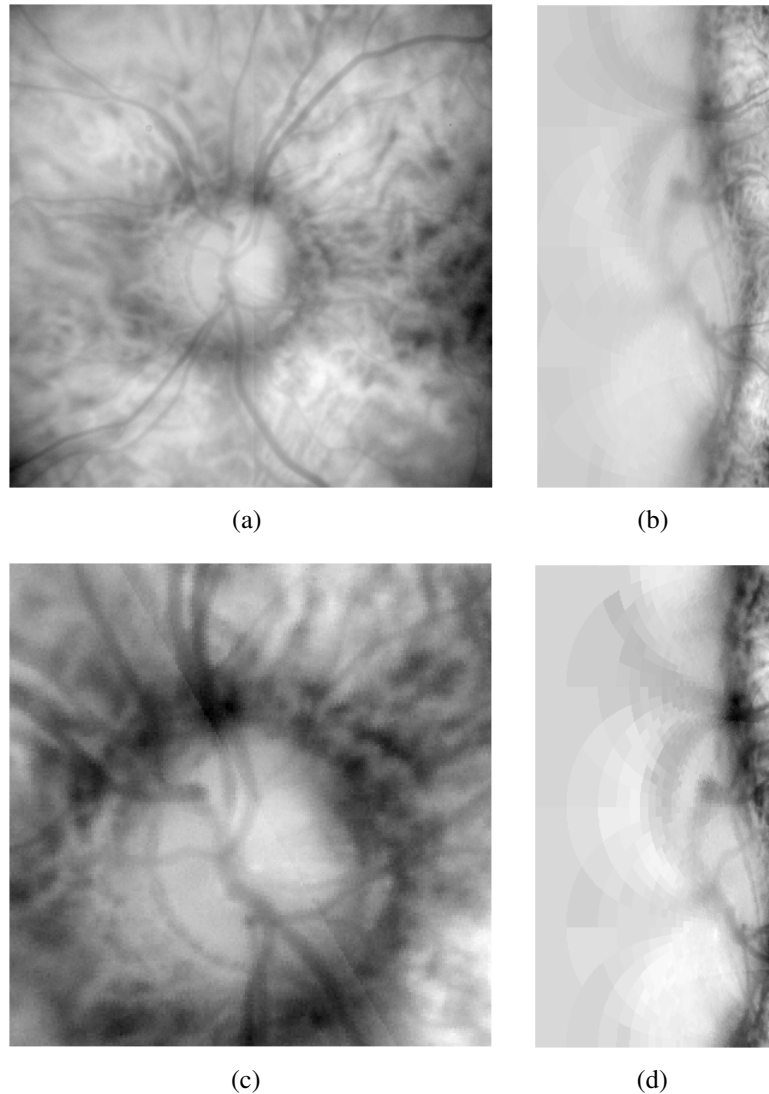


Figura 20: Exibe o efeito de uma rotação e ampliação entre dois quadros consecutivos da retina. (a) representa um quadro da retina em infravermelha no tempo  $t$ ; (b) é a sua contraparte no espaço log-polar; (c) é o quadro consecutivo em  $t + 1$  com uma ampliação de 2 vezes, uma rotação de  $30^\circ$  e uma translação nula; (d) é o quadro  $I(t + 1)$  no domínio log-polar.

A fim de estimar parâmetros de movimento,  $\theta$ ,  $S$ ,  $\Delta x$  e  $\Delta y$ , e então estimar o movimento detectado, o método proposto com base na transformação log-polar requer que o deslocamento de translação seja zero entre  $I(t)$  e  $I(t + 1)$ . Em outras palavras, é necessário obter os parâmetros de translação  $\Delta x$  e  $\Delta y$  antes da realização de correlação-cruzada para obter os parâmetros de rotação e escala. Para isso, é introduzido um algoritmo de otimização de busca esparsa com o uso de transformação log-polar + deslocamento de fase.

Para um par de quadros consecutivos  $I(t)$  e  $I(t + 1)$ , representados nas Figuras 21(a)

e 21(b), a região central de  $I(t)$  é recortada, ou seja, é utilizada uma região de  $I(t)$  localizada no centroide da imagem  $(x_c, y_c)$ , para formar o quadro  $I'_1$ , Figura 21(c). Depois,  $I'_1$  é transformada para o espaço log-polar  $I'_{1p}$  usando as Eqs. 18 - 23 representada na Figura 21(e). Para cada posição  $(x', y')$ ,  $x' = 1 + w/2 : 10 : M - w/2$  e  $y' = 1 + w/2 : 5 : N - w/2$ , a imagem  $I(t + 1)$ , Figura 21(b), é recortada em uma região centrada em  $(x', y')$  para formar a imagem  $I'_2$ , e sequencialmente  $I'_2$  é transformada para log-polar formando a imagem  $I'_{2p}$  exibidas respectivamente nas Figuras 21(d) e 21(f), e para cada posição  $(x', y')$  a correlação-cruzada  $p(x', y')$  entre as imagens  $I'_{1p}(x_c, y_c)$  e  $I'_{2p}(x', y')$  é calculada usando as Eqs. 12, 13 e 14 onde  $f = I'_{1p}(x_c, y_c)$  e  $g = I'_{2p}(x', y')$ . O parâmetro  $w$  refere-se ao tamanho da janela de busca e  $M \times N$  é o tamanho da imagem  $I(t)$ . Neste ponto, a função mapa da máxima correlação-cruzada  $p_{map}(x', y')$  é construída e possui apenas os valores máximos de cada correlação-cruzada  $p(x', y')$  e está ilustrada na Figura 22(a). A procura esparsa otimiza o tempo e a complexidade do algoritmo, mas indica a posição provável do pico mais alto.

Dado o conjunto de todas as correlações-cruzadas calculadas, o detalhamento do pico é realizado calculando a correlação-cruzada com uma pesquisa completa com uma janela de  $5 \times 5$  em torno da posição em que ocorreu o pico mais alto em  $p_{map}(x'', y'')$ , onde  $p_{map}(x'' + mm, y'' + nn)$  é calculado para  $mm = -2 : 1 : 2$  e  $nn = -2 : 1 : 2$ , veja a Figura 22(b). Se  $p_{map}(x'' + mm, y'' + nn)$  é o pico mais alto, os parâmetros de translação são estimados por  $(\Delta x = x'' + mm$  e  $\Delta y = y'' + nn)$ , e a rotação e a escala são estimadas pela posição do pico  $(x_{pico}, y_{pico})$  em  $p(x'' + mm, y'' + nn)$  tal que o ângulo é obtido através do deslocamento  $da$ ,  $\theta = da = x_c - x_{pico}$  e o fator de escala é obtido mediante o deslocamento  $d(\log_b r)$ ,  $S = d(\log_b r) = y_c - y_{pico}$ , conforme Figura 23. Assim, essa metodologia de pesquisa iterativa para encontrar a translação nula entre o par de quadros (posição onde ocorre o pico em  $p_{map}$ ) e as correlações-cruzadas realizadas no domínio log-polar permite encontrar os parâmetros de rotação, escala e translação em imagens da retina em infravermelho. Esta explicação está resumida no Algoritmo 2 (p. 55).

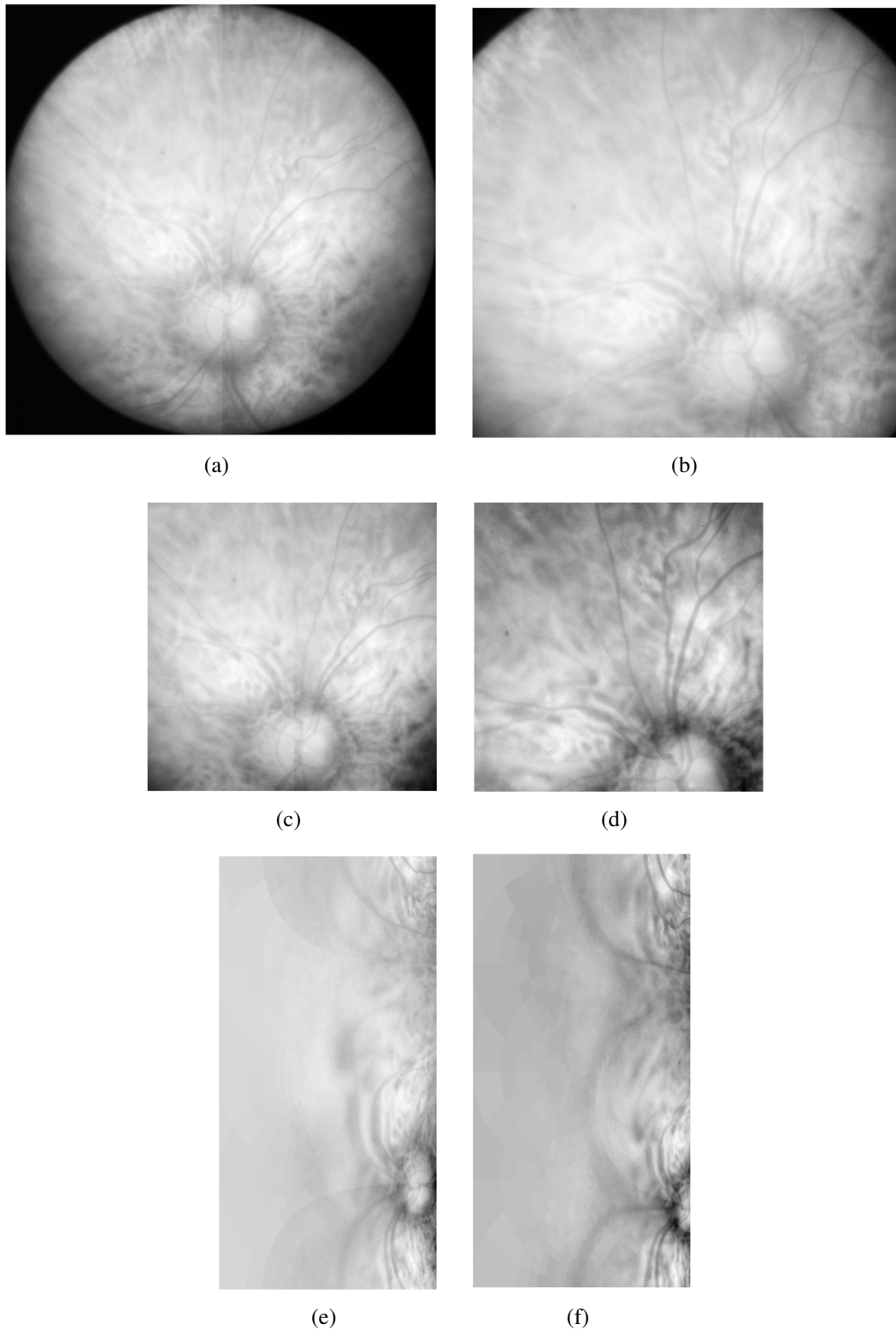
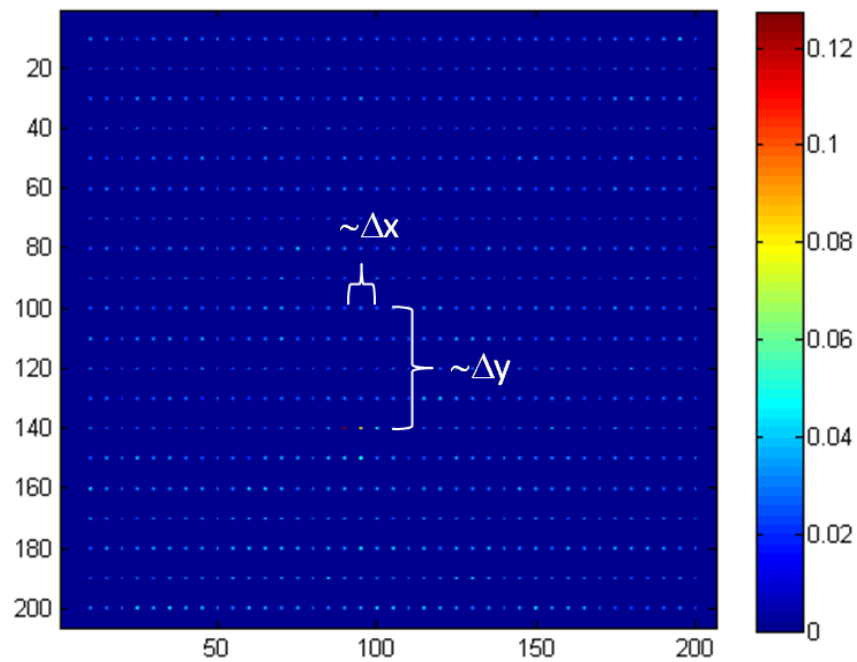
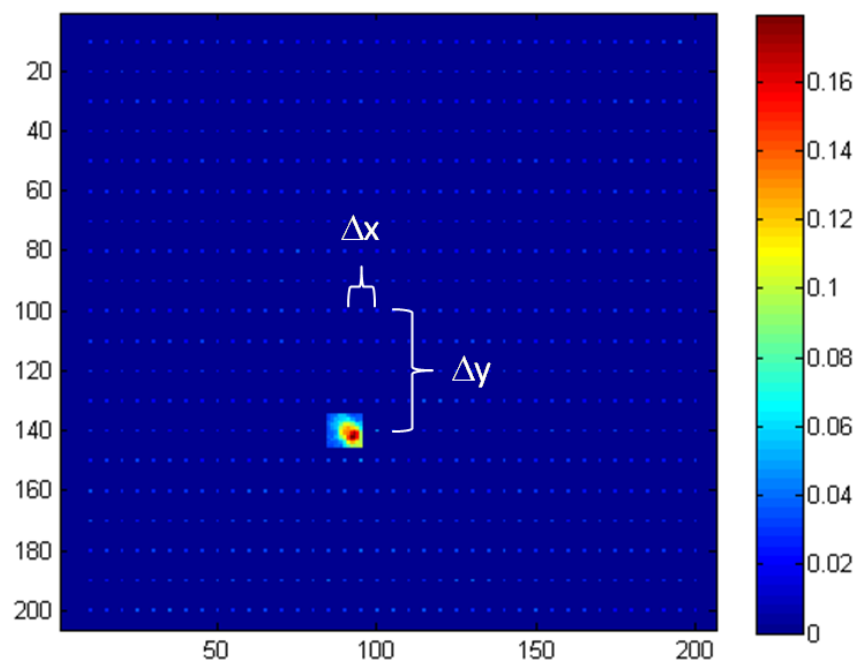


Figura 21: Ilustra passos do uso da transformação log-polar para estimar os parâmetros; (a) quadro da retina  $I(t)$ ; (b) quadro da retina  $I(t + 1)$ ; (c)  $I'_1$  é o recorte da região central de  $I(t)$ ; (d)  $I'_2$  é o recorte da região central de  $I(t + 1)$ ; (e)  $I'_{1p}$  é a transformação log-polar de  $I'_1$ ; (f)  $I'_{2p}$  é a transformação log-polar de  $I'_2$ .

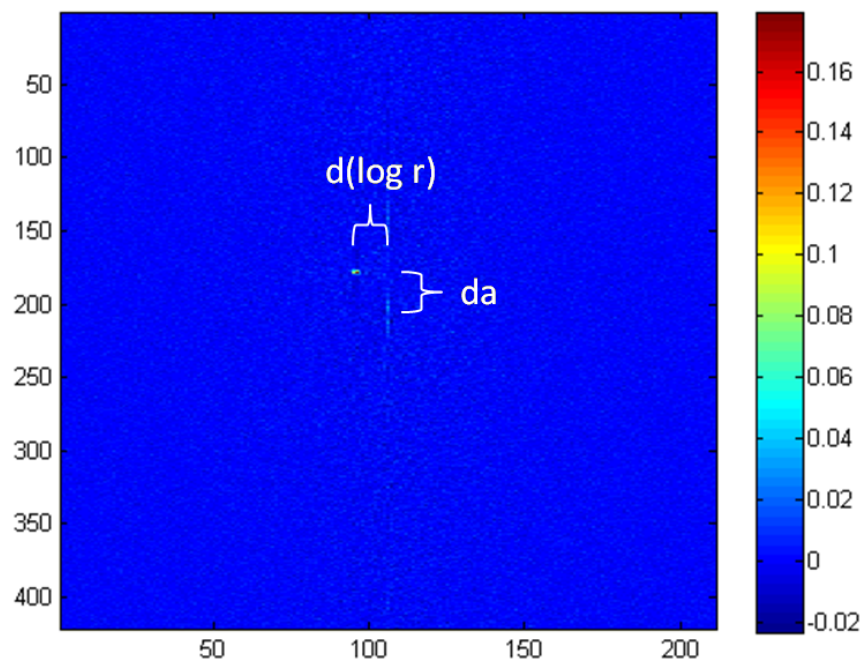


(a)



(b)

Figura 22: Exibe os mapas de máxima correlação-cruzada,  $p_{map}$  construídos pelo algoritmo de estimação dos parâmetros de movimento; (a)  $p_{map}$  resultante da busca esparsa pelo máximo pico com a posição estimada (90, 141); (b)  $p_{map}$  resultante após a busca completa na região estimada do pico. O máximo pico de correlação-cruzada acontece na posição (92, 142). Os parâmetros de translação ( $\Delta x$ ,  $\Delta y$ ) são dados pela posição do pico em relação ao centroide.



(a)

Figura 23: Ilustra a correlação-cruzada  $p(x'' + mm, y'' + nm)$  realizada na posição  $(x'' + mm, y'' + nm)$  tal que a posição do pico traduz nos parâmetros de rotação e escala são dados pelos deslocamentos  $da$  e  $d(\log_b r)$ .

---

**Algoritmo 2:** Método da transformação log-polar para estimar os parâmetros de rotação, escala e translação.

---

**Entrada:** Um par de quadros consecutivos

**Saída:** Parâmetros de estimação para rotação:  $\theta$ , escala:  $S$  e translação:  $\Delta x, \Delta y$

```

1  Sejam  $I(t)$  e  $I(t + 1)$ , um par de quadros consecutivos;
2  Recorta a região central  $I'_1$  de  $I(t)$ ;
3   $(M, N) \leftarrow \text{tamanho}(I(t))$  e  $(x_c, y_c) \leftarrow \text{centroide}(I'_1)$ ;
4  Computa  $I'_{1p}$ , a transformação log-polar de  $I'_1$ ;
5  para  $x' \leftarrow 1 + w/2$  até  $M - w/2$  passo 10 faça
6  |   para  $y' \leftarrow 1 + w/2$  até  $N - w/2$  passo 5 faça
7  |   |   Recorta a região centrada em  $(x', y')$ ,  $I'_2$  de  $I(t+1)$ ;
8  |   |   Computa  $I'_{2p}$ , a transformação log-polar de  $I'_2$ ;
9  |   |    $P \leftarrow$  Espectro de potência cruzada entre  $I'_{1p}$  e  $I'_{2p}$ ;
10 |   |    $p_{map}(x', y') \leftarrow \max(\mathcal{F}^{-1}(P))$ ;
11 |   Encontra o pico máximo  $(x'', y'') \leftarrow \max_{x,y}\{p_{map}\}$ ;
12 |   Detalhamento do pico máximo:
13 |   para  $mm \leftarrow -2$  até 2 passo 1 faça
14 |   |   para  $nn \leftarrow -2$  até 2 passo 1 faça
15 |   |   |   Recorta a região centrada em  $(x'' + mm, y'' + nn)$   $I'_2$  de  $I(t+1)$ ;
16 |   |   |   Computa  $I'_{2p}$ , a transformação log-polar de  $I'_2$ ;
17 |   |   |    $P \leftarrow$  Espectro de potência cruzada entre  $I'_{1p}$  e  $I'_{2p}$ ;
18 |   |   |    $p(x'' + mm, y'' + nn) \leftarrow \mathcal{F}^{-1}(P)$ , correlação-cruzada;
19 |   |   |    $p_{map}(x'' + mm, y'' + nn) \leftarrow \max(p(x'' + mm, y'' + nn))$ , mapa de
20 |   |   |   máxima correlação-cruzada;
21 |   |   |   se  $p_{map}(x'' + mm, y'' + nn)$  é máximo então
22 |   |   |   |    $(x_{pico}, y_{pico}) \leftarrow \max_{x,y}\{p(x'' + mm, y'' + nn)\}$ ;
23 |   |   |   |   Rotação  $\theta \leftarrow x_c - x_{pico}$ ;
24 |   |   |   |   Escala  $S \leftarrow y_c - y_{pico}$ ;
25 |   |   |   |   Translação  $(\Delta x, \Delta y) \leftarrow (x'' + mm, y'' + nn)$ ;
26 retorna  $\theta, S, \Delta x$  e  $\Delta y$ .

```

---

### 3.4 Sumário do Método Proposto

As métricas  $D(X, Y)$  e  $H(X, Y)$  são usadas para detectar o movimento da retina entre dois quadros consecutivos e decidir se os movimentos da retina serão estimados ou não. Normalmente, translações simples - onde não há rotação nem escala - entre um par de quadros consecutivos  $(I(t), I(t + 1))$  apresentam um forte pico na função impulsiva da correlação-cruzada  $I_{\Delta}(x, y)$  gerada a partir do algoritmo de deslocamento de fase. Nestes casos,  $\max(I_{\Delta}(x, y)) > \lambda$  caracteriza um verdadeiro pico, como na linha 4 do Algoritmo 3. Assim, os parâmetros de translação são estimados usando apenas o método de deslocamento de fase. Entretanto, se há rotação e/ou escala, o método da transformação log-polar é utilizado para estimar os parâmetros de movimento. Como sugerido em Reddy e Chatterji (1996),  $\lambda = 0.03$  nos experimentos realizados.

A inovação do método proposto para calcular o deslocamento da retina utiliza duas técnicas diferentes para estimar os parâmetros de movimento da retina: deslocamento de fase e transformação log-polar. Além disso, se faz uso de pesquisa esparsa e do detalhamento do pico para criar uma estimativa confiável dos parâmetros de movimento. Resumindo o método proposto, após o movimento ser detectado, para os 2 quadros consecutivos é aplicado o método de deslocamento de fase de acordo com as Eqs. 12 - 16. Dependendo do valor máximo da função impulsiva da correlação-cruzada, Eq. 16, os parâmetros de translação são estimados corretamente para os casos de translação simples e  $\max(I_{\Delta}(x, y)) > \lambda$ . Nestes casos,  $\Delta x$  e  $\Delta y$  são estimados, o fator de escala é definido para a unidade 1 e a rotação é ajustada para  $0^{circ}$ . Esta abordagem melhora o consumo de tempo e a complexidade do algoritmo. Para os outros casos em que  $\max(I_{\Delta}(x, y)) \leq \lambda$ , presume-se que exista uma rotação e/ou um fator de escala entre os dois quadros consecutivos. Nesse caso, o método da transformação log-polar é utilizado associado a uma pesquisa iterativa na qual os parâmetros de translação, rotação e escala são estimados.

---

**Algoritmo 3:** Método proposto para detecção e estimação do movimento da retina.

---

```

1  Calcula  $D(X, Y)$  e  $H(X, Y)$ , onde  $X = |\overrightarrow{\nabla I(t)}|$  e  $Y = |\overrightarrow{\nabla I(t+1)}|$ ;
2  se  $D(X, Y) \leq \tau \wedge H(X, Y) > \eta$  então
3  |   Calcula o deslocamento de fase entre  $I(t)$  e  $I(t+1)$ ;
4  |   se  $\max(I_{\Delta}(x, y)) > \lambda$  então
5  |   |    $\Delta x$  e  $\Delta y$  são estimados pelo algoritmo de deslocamento de fase;
6  |   |    $S = 1$ ;
7  |   |    $\theta = 0$ ;
8  |   senão
9  |   |    $\theta, S, \Delta x$  e  $\Delta y$  são estimados pelo algoritmo log-polar;
10 senão
11 |   O movimento da retina detectado é grande ou inconsistente;
12 retorna  $\theta, S, \Delta x$  e  $\Delta y$ .
```

---

### 3.5 Classificação dos Movimentos da Retina

Obtidos os valores de  $MI(X, Y)$ ,  $H(X, Y)$  e  $D(X, Y)$  bem como os parâmetros de estimação de movimento  $\theta$ ,  $S$ ,  $\Delta x$  e  $\Delta y$ , é necessário verificar se o movimento detectado é consistente e se pode ser estimado com sucesso. Isso é realizado através da classificação de movimentos da retina. Para este fim são aplicados os seguintes 5 limiares:  $\tau$ ,  $\eta$ ,  $\alpha$ ,  $\sigma$  e  $\gamma$ . Onde o limiar  $\tau$  está associado à métrica  $D(X, Y)$ ,  $\eta$  está associado à entropia conjunta  $H(X, Y)$ ,  $\alpha$  está associado ao ângulo de rotação  $\theta$ ,  $\sigma$  está associado ao fator de escala  $S$  e  $\gamma$  está associado aos parâmetros de translação  $\Delta x$  e  $\Delta y$ . Situações inconsistentes são movimentos da retina atípicos, como piscar de olhos, o que é visto por quadros saturados (ou seja, onde a maioria das intensidades dos *pixels* estão próximas de 255), borrão de movimento ou quadros fora de foco causados por movimentos súbitos de cabeça. Situações consistentes são movimentos típicos da retina que ocorrem quando existe um pequeno deslocamento entre imagens adjacentes, ou seja, pode haver uma pequena rotação e/ou escala e/ou translação da qual esse movimento pode ser estimado corretamente. Já as si-



tuações nas quais ocorrem grande movimento, as medidas teóricas  $D(X, Y)$  e  $H(X, Y)$  bem como os parâmetros estimados de movimento  $\theta$ ,  $S$ ,  $\Delta x$  e  $\Delta y$  indicam situações em que não há necessidade de estimação de movimento, pois o tratamento a laser seria interrompido.

Tabela 3 abaixo mostra a situação especial de pequeno movimento ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge \|\theta\| \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma$ ). Esta situação é comparada com o conjunto verdade ( $GT$ ). Estes limiares são estabelecidos com base na figura de mérito ( $FDM$ ) definida na Eq 11, e é a média entre ( $VPP$ ) e a ( $ACC$ ).

Tabela 3: Tipos de movimentos detectados entre quadros e a condição de detecção.

Tipos de movimentos detectados	Condição para detecção
Pequeno	$D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge [\ \theta\  \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$
Pequeno, mas inconsistente	$D(X, Y) \leq \tau \wedge H(X, Y) \leq \eta \wedge [\ \theta\  \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$
Grande	$D(X, Y) > \tau \wedge H(X, Y) > \eta \wedge [\ \theta\  > \alpha \vee (S \leq 1 - \sigma \vee S \geq 1 + \sigma) \vee \sqrt{\Delta x^2 + \Delta y^2} > \gamma]$
Grande, mas inconsistente	$D(X, Y) > \tau \wedge H(X, Y) \leq \eta \wedge [\ \theta\  > \alpha \vee (S \leq 1 - \sigma \vee S \geq 1 + \sigma) \vee \sqrt{\Delta x^2 + \Delta y^2} > \gamma]$

### 3.6 Identificação e Ajuste dos Parâmetros Críticos para a Detecção e Estimação dos Movimentos da Retina

A detecção e a estimação de movimentos da retina podem ser alteradas através da variação dos limiares  $\tau$ ,  $\eta$ ,  $\alpha$ ,  $\sigma$  e  $\gamma$ , que por sua vez influenciam a figura de mérito  $FDM$  composta pelo  $VPP$  e a  $ACC$ . Idealmente,  $VPP$  e  $ACC$  tendem a 100%, o que significa que os movimentos da retina são detectados corretamente entre todos os pares de quadros consecutivos de acordo com o  $GT$ . Determinou-se os parâmetros críticos usando duas análises de variâncias independentes. O primeiro teste é um projeto de experimento de fatorial completo com 2 fatores controláveis  $\tau$  e  $\eta$  a 5 níveis cada, e a variável de resposta é figura de mérito  $FDM$ . O segundo teste é um projeto de experimento de fatorial completo com 3 fatores controláveis  $\alpha$ ,  $\sigma$  e  $\gamma$  a 5 níveis cada, e a  $FDM$  é variável de resposta, o que será detalhado na Seção 4.1.

### 3.7 Método Comparativo para Detecção e Estimação do Movimento da Retina

O método desenvolvido por Xu et al (2005) utiliza um escâner laser oftalmológico para obter as imagens da retina. O *block matching*, um dos métodos padrão para rastreamento, foi empregado nesse estudo. Em um algoritmo genérico de *block matching* um bloco de *pixels*  $M \times N$  é definido no tempo  $t$ . Este bloco é utilizado para determinar o movimento entre os quadros  $t$  e  $t + 1$  através do encontro do melhor casamento entre este bloco e um outro bloco em uma área de busca no quadro  $t + 1$ .

Já o método proposto por Li et al (2010) usa um escâner laser oftalmológico com óptica adaptativa confocal. Nesse tipo de equipamento o movimento do olho severamente afeta a estabilização das imagens de vídeo e resultam em distorções das imagens da retina. Nesse trabalho, *Scale-Invariant Feature Transform (SIFT)* é utilizado para abstrair pontos salientes estáveis das imagens da retina. O algoritmo de Kanade-Lucas-Tomas (KLT) é aplicado para rastrear esses pontos.

Ambos os métodos (XU et al., 2005) e (LI et al., 2010) abordam a estimação dos parâmetros de movimento da retina em relação à translação. Para validar e comparar os resultados obtidos para a detecção e estimação de movimento da retina pelos algoritmos propostos é utilizado o método comparativo desenvolvido em Reddy e Chatterji (1996), pois trata-se de um método de registro de imagens capaz de estimar rotações, escalas e translações. Desta forma, é o método mais adequado para o tipo de imagens contidas no conjunto de dados desse trabalho.

No método comparativo (REDDY; CHATTERJI, 1996) foram utilizados os mesmos limiares otimizados. A detecção e estimação de movimentos da retina é feita de acordo com critérios da Tabela 4.

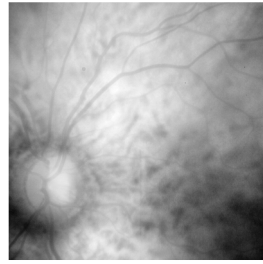
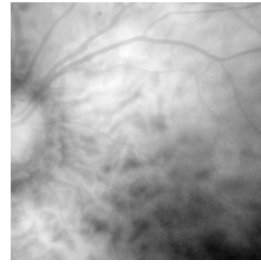
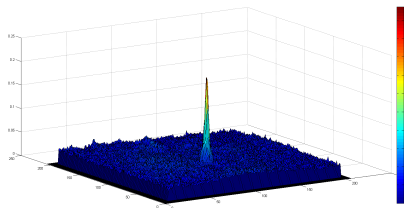
Tabela 4: Tipos de movimentos detectados entre quadros e a condição de detecção para o método comparativo.

Tipos de movimentos detectados	Condição de detecção
Movimento pequeno	$\ \theta\  \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma$
Movimento grande	$\ \theta\  > \alpha \vee (S \leq 1 - \sigma \vee S \geq 1 + \sigma) \vee \sqrt{\Delta x^2 + \Delta y^2} > \gamma$

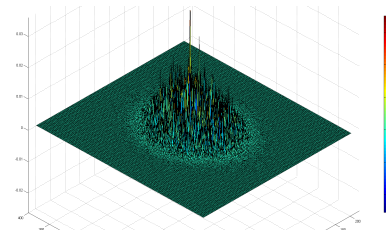
O método proposto aqui fundamenta-se na abordagem que utiliza o deslocamento de fase e a transformação log-polar, sendo diferente do método utilizado em Reddy e Chatterji (1996), no entanto os dois métodos usam o espectro de potência cruzada. Uma vantagem do método proposto nesta dissertação é usar a transformação log-polar onde a informação da imagem está no domínio real, assim as informações utilizadas no espectro de potência cruzada também são reais. Sejam  $f(x, y) = I(t)$  e  $g(x, y) = I(t + 1)$  os quadros consecutivos de imagens da retina que sofrem uma transformação de coordenadas log-polar nomeados  $L1(\log_b r, a)$  e  $L2(\log_b r, a)$ , sendo que  $F(u, v) = \mathcal{F}(L1(\log_b r, a))$  e  $G(u, v) = \mathcal{F}(L2(\log_b r, a))$  são as suas respectivas transformadas de Fourier. O espectro de potência cruzada é  $P(u, v) = \frac{F(u, v)G^*(u, v)}{|F(u, v)G^*(u, v)|}$ , onde  $p(x, y) = \mathcal{F}^{-1}(P)$  é a correlação-cruzada normalizada. No método proposto, isso produz uma função com um pico bem definido em  $p(x, y)$ , veja a Figura 24(e).

Porém, o método de Reddy e Chatterji (1996) mapeia a informação da imagem para o domínio de frequência utilizando apenas as suas magnitudes, conforme foi detalhado a seguir. Como consequência, o espectro de potência cruzada entre duas matrizes é realizada no domínio da frequência. Portanto, assumindo que o espectro de potência cruzada é executado no domínio de frequência para um par de quadros consecutivos  $f(x, y) = I(t)$  e  $g(x, y) = I(t + 1)$ , tais que  $F(u, v)$  e  $G(u, v)$  são as suas respectivas transformadas de Fourier. As magnitudes são definidas por  $M1(u, v) = |F(u, v)|$  e  $M2(u, v) = |G(u, v)|$ . Após a aplicação de um filtro passa-alta sobre  $M1(u, v)$  e  $M2(u, v)$ ,  $M1'(u, v) = \text{hipass}(M1(u, v))$  e  $M2'(u, v) = \text{hipass}(M2(u, v))$ , ou seja, os coeficientes de baixa frequência são atenuados. Essas matrizes  $M1'(u, v)$  e  $M2'(u, v)$  são mapeadas para o sistema coordenado log-polar,  $L1(u', v')$  e  $L2(u', v')$ , respectivamente. Por fim, o espectro de potência cruzado  $P$  é calculado usando  $F = \mathcal{F}(L1(u', v'))$  e  $G = \mathcal{F}(L2(u', v'))$ . A utilização deste método provoca o aparecimento de picos secundários na correlação-cruzada  $p(x, y)$ . Nos casos em que ocorrem rotações e escalas significativas torna-se difícil a detecção do pico correto, ver Figura 24(f). Esse exemplo é ilustrado na Figura 24, onde 24(a) e 24(b) são um par de quadros consecutivos com rotação, escala e translação entre si, ou seja,  $I(t)$  e  $I(t + 1)$ . A Figura 24(c) mostra a imagem

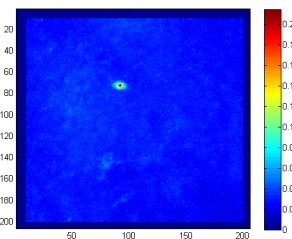
$I(t + 1)$  corrigida pelo método proposto e a Figura 24(d) mostra  $I(t + 1)$  corrigida pelo método Reddy e Chatterji (1996). As Figuras 24(g) e (h) mostram a vista superior das correlações-cruzadas para ambos os métodos.

(a)  $I(t)$ ;(b)  $I(t+1)$ ;(c)  $I(t+1)$  corrigida pelo método proposto. Parâmetros estimados:  $\theta = -22.52^\circ$ ,  $S = 1.29$ ,  $\Delta x = -11$  e  $\Delta y = -31$ , alcançando  $PSNR = 40.01dB$ ;(d)  $I(t+1)$  corrigida pelo método em (REDDY; CHATTERJI, 1996). Parâmetros estimados:  $\theta = -21.65^\circ$ ,  $S = 1.31$ ,  $\Delta x = -14$  e  $\Delta y = -18$ , alcançando  $PSNR = 37.34dB$ ;

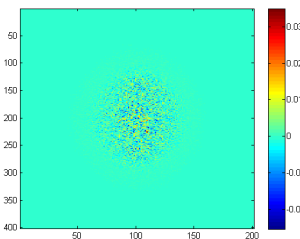
(e) Resultado da correlação-cruzada pelo método proposto;



(f) Resultado da correlação-cruzada pelo método em (REDDY; CHATTERJI, 1996);



(g) Visão superior da correlação-cruzada pelo método proposto;



(h) Visão superior da correlação-cruzada pelo método em (REDDY; CHATTERJI, 1996).

Figura 24: Ilustra a comparação entre os métodos. (a) e (b), ou seja,  $I(t)$  e  $I(t+1)$ , são quadros consecutivos com grande movimento entre si. (c)  $I(t+1)$  corrigido pelo método proposto; (d)  $I(t+1)$  corrigido pelo método em (REDDY; CHATTERJI, 1996); (e) resultado da correlação-cruzada com o método proposto contendo apenas um pico bem definido; (f) resultado da correlação-cruzada com o método em (REDDY; CHATTERJI, 1996) contendo vários picos; (g) e (h) são as visões de cima das correlações-cruzadas para ambos os métodos. Neste exemplo, o método proposto alcança maior  $PSNR$  entre  $I(t)$  e  $I(t+1)$  corrigido por parâmetros estimados.

## 4 RESULTADOS EXPERIMENTAIS

Nesta dissertação são utilizados dados de imagens em infravermelho captadas por um retinógrafo digital não midriático, ou seja, que tem iluminação infravermelha permitindo assim que os estudos da retina podem ser realizados sem a necessidade da dilatação de pupila. A luz infravermelha é utilizada ajustada à faixa de  $940nm$ . O sensor óptico é uma matriz CCD (*Charge-coupled device*, dispositivo de carga acoplada) monocromática com cerca de 11 *MPixels* (isto é,  $4008 \times 2672$  sensores com uma área de  $9.0 \times 9.0 \mu m$  cada). Apenas uma fração da resolução das imagens é utilizada devido ao uso de um *binning* (agrupamento)  $2 \times 2$  e  $\frac{1}{4}$  da resolução total da imagem é disponível, tornando-se mais fácil de trabalhar com 13 quadros/s. Imagens estáticas são tiradas com uma ampliação de 1.85 e cada pixel corresponde a uma área da retina de  $4.86 \times 4.86 \mu m$ , mas em sequências de vídeo cada pixel corresponde a uma área da retina de  $19.46 \mu m^2$ . Para um quadro de retina em infravermelho de  $627 \times 627 pixels$  e uma janela de pesquisa de  $420 \times 420 pixels$  são necessários apenas 805 iterações contra 42436, caso uma busca exaustiva fosse utilizada para estimar os parâmetros de movimento da retina.

Conforme detalhado na Seção 3.1, dado um par de quadros adjacentes nos tempos  $t$  e  $t + \Delta t$  - respectivamente representados por suas magnitudes dos gradientes das imagens,  $X$  e  $Y$  - apenas os casos de pequeno movimento da retina são estimados após serem detectados se forem satisfeitas as seguintes condições  $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge \|\theta\| \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma$ . Isso significa que o laser é reposicionado entre os tempos  $t$  e  $t + \Delta t$  em tais casos. A constante utilizada nos experimentos foi  $\Delta t = 1$ , com foco nos quadros imediatamente adjacentes nas sequências de vídeo

da retina em infravermelho, o que torna a resposta do método mais rápida até mesmo para pequenos movimentos da retina. Os valores de  $\tau$ ,  $\eta$ ,  $\alpha$ ,  $\sigma$  e  $\gamma$  foram otimizados após a determinação dos parâmetros críticos, utilizando ANOVA, conforme detalhado na Seção 4.1.

## 4.1 Ajustes de Parâmetros

Para determinar os parâmetros críticos foram realizados 2 projetos de experimentos de fatorial completo. O primeiro é um projeto de experimento de fatorial completo com 2 fatores controláveis  $\tau$  e  $\eta$  a 5 níveis cada com 3 repetições. Então  $\tau$  e  $\eta$  podem assumir 5 valores:  $\tau \in [0.93 : 0.01 : 0.97]$  e  $\eta \in [1.4 : 0.1 : 1.8]$ . O segundo é um projeto de experimento de fatorial completo com 3 fatores controláveis  $\alpha$ ,  $\sigma$  e  $\gamma$  a 5 níveis cada e 3 repetições. Estes fatores controláveis podem assumir os valores:  $\alpha \in [0 : 15 : 60]$ ,  $\sigma \in [0 : 0.2 : 0.8]$  e  $\gamma \in [0 : 20 : 80]$ . A variável de resposta de ambos os experimentos é a *FDM*. Esses experimentos foram conduzidos de maneira aleatorizada e foi utilizado o conjunto treino corrompido com ruído gaussiano de média zero e variância 0.0001.

Dados estes experimentos, a análise de variância foi utilizada para verificar a hipótese  $H_0$  (MONTGOMERY, 2001), ou seja, analisar se as diferenças entre os tratamentos foram significativos assumindo a distribuição Fisher-Snedecor. Neste caso,  $F_{calculado}$  é comparada com os valores da distribuição Fisher-Snedecor  $F_{tabela}$ , e se  $F_{calculado} > F_{tabela}$  a fonte de variação, ou seja, o limiar é significativo. A Tabela 5 mostra que  $\tau$ ,  $\eta$  e sua correlação  $\tau * \eta$  têm variâncias significativas. Do mesmo modo, a Tabela 6 mostra que  $\alpha$ ,  $\sigma$ ,  $\gamma$  têm variações significativas também, porém suas correlações não. A significância da correlação  $\tau * \eta$  exige que  $\tau$  e  $\eta$  sejam otimizados de forma correlacionada. Enquanto que a não significância das correlações entre  $\alpha$ ,  $\sigma$  e  $\gamma$  possibilita a otimização desses limiares de forma independente.

Tabela 5: Resultados experimentais da ANOVA dos parâmetros de detecção do movimento da retina.

Limiar	Soma Qua.	G.L	Média Qua.	$F_{calc.}$	$F_{tab.}$	Conclusão
$\tau$	275.351	4	68.848	32899	2.557	Significativo
$\eta$	63.400	4	15.850	7575	2.557	Significativo
$\tau * \eta$	2.526	16	0.158	75	1.850	Significativo
Erro	0.105	50	0.002	-	-	-
TOTAL	341.382	74	-	-	-	-

Tabela 6: Resultados experimentais da ANOVA dos parâmetros de estimação do movimento da retina.

Limiar	Soma Qua.	G.L	Média Qua.	$F_{calc.}$	$F_{tab.}$	Conclusão
$\alpha$	227.4	4	56.84	13.54	2.41	Significativo
$\sigma$	116	4	28.99	6.9	2.41	Significativo
$\gamma$	17523.3	4	4380.8	1043.18	2.41	Significativo
$\alpha * \sigma$	40.2	16	2.52	0.6	1.68	Não Significativo
$\alpha * \gamma$	111	16	6.94	1.65	1.68	Não Significativo
$\sigma * \gamma$	36.9	16	2.3	0.55	1.68	Não Significativo
$\alpha * \sigma * \gamma$	18.1	64	0.28	0.07	1.36	Não Significativo
Erro	1049.9	250	4.2	-	-	-
TOTAL	19122.6	374	-	-	-	-

Desta forma, para otimizar os parâmetros de detecção e estimação do movimento da retina, é necessário avaliar os 5 limiares:  $\tau$  e  $\eta$  utilizados na detecção do movimento da retina, e os limiares  $\alpha$ ,  $\sigma$  e  $\gamma$  usados na estimação do movimento. Lembrando que, nesta pesquisa, a detecção do movimento não está amarrada com a estimação. Então é possível otimizar os 5 limiares em 2 conjuntos distintos de acordo com as Eqs. 24 e 25:

$$(\tau, \eta) = \max_{\tau', \eta' \in \mathbb{R}} \{FDM(\tau', \eta')\}, \quad (24)$$

$$(\alpha, \sigma, \gamma) = \max_{\alpha', \sigma', \gamma' \in \mathbb{R}} \{FDM(\alpha', \sigma', \gamma')\}. \quad (25)$$

Isso significa que todos esses 5 limiares são parâmetros críticos e devem ser ajustados para alcançar altos valores de  $FDM$ . Portanto, nos experimentos de detecção de movimento e estimação da retina, os valores dos parâmetros críticos otimizados são  $\tau = 0.9527$ ,  $\eta = 1.6258$ ,  $\alpha = 38.97$ ,  $\sigma = 0.41$  e  $\gamma = 50.15$ .



## 4.2 Resultados Obtidos

A abordagem proposta está resumida nos Algoritmos 1, 2 e 3 (p. 42, 53 e 55), que descrevem como o método proposto detecta o movimento da retina estimando os parâmetros de movimento em termos de rotação, escala e translação. A Figura 25 ilustra os valores da métrica universal  $D(X, Y)$  obtidos numa situação de detecção do movimento da retina, enquanto a Figura 26, na sequência, mostra os valores correspondente da entropia conjunta  $H(X, Y)$  usada para detectar situações inconsistentes. A linha preta tracejada mostra o  $GT$ , e a linha vermelha tracejada indica as regiões que possivelmente têm pequenos movimentos válidos ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge [||\theta|| \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$ ). Para uma melhor visualização, apenas são exibidos 200 quadros do conjunto de teste. As Figuras 27, 28 e 29 mostram os parâmetros de movimento estimados para os primeiros 400 quadros do conjunto teste das imagens da retina em infravermelho. Os resultados apresentados nesta seção serão discutidos na Seção 4.3.

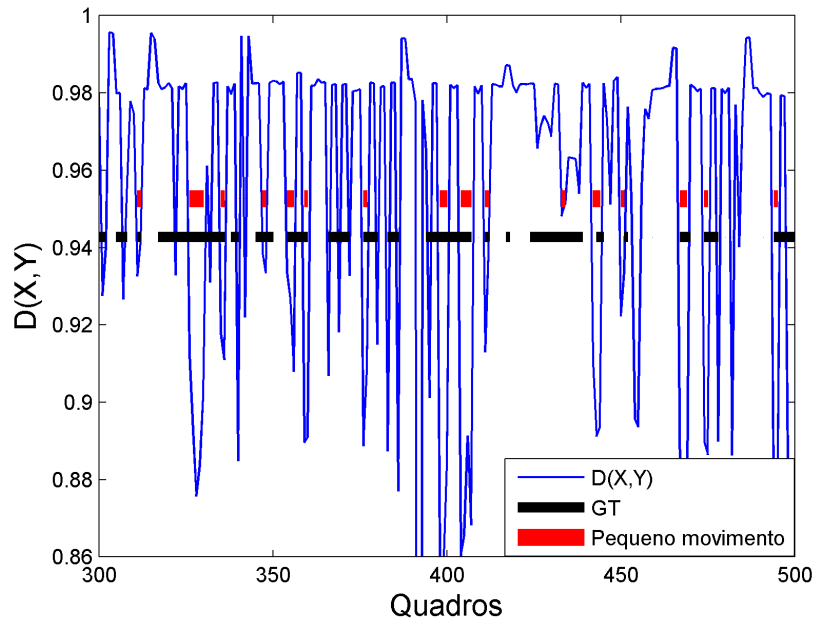


Figura 25: Valores da métrica universal  $D(X, Y)$  (curva azul) para 200 quadros do conjunto teste das imagens da retina em infravermelho. Os segmentos de linhas tracejadas em vermelho indicam as possíveis regiões de movimento pequeno ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge [||\theta|| \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma]$ ), os segmentos de linha na cor preta mostram o conjunto verdade  $GT$ .

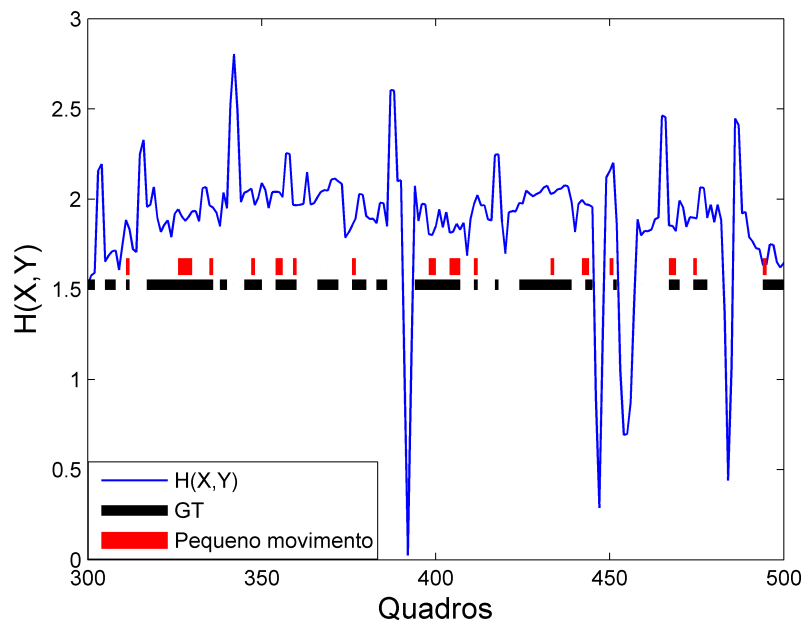


Figura 26: Valores da entropia conjunta  $H(X, Y)$  (curva azul) para 200 quadros do conjunto teste das imagens da retina em infravermelho. Os segmentos de linhas tracejadas em vermelho indicam as possíveis regiões de movimento pequeno ( $D(X, Y) \leq \tau \wedge H(X, Y) > \eta \wedge \|\theta\| \leq \alpha \wedge (1 - \sigma \leq S \leq 1 + \sigma) \wedge \sqrt{\Delta x^2 + \Delta y^2} \leq \gamma$ ), os segmentos de linha na cor preta mostram o conjunto verdade  $GT$ .

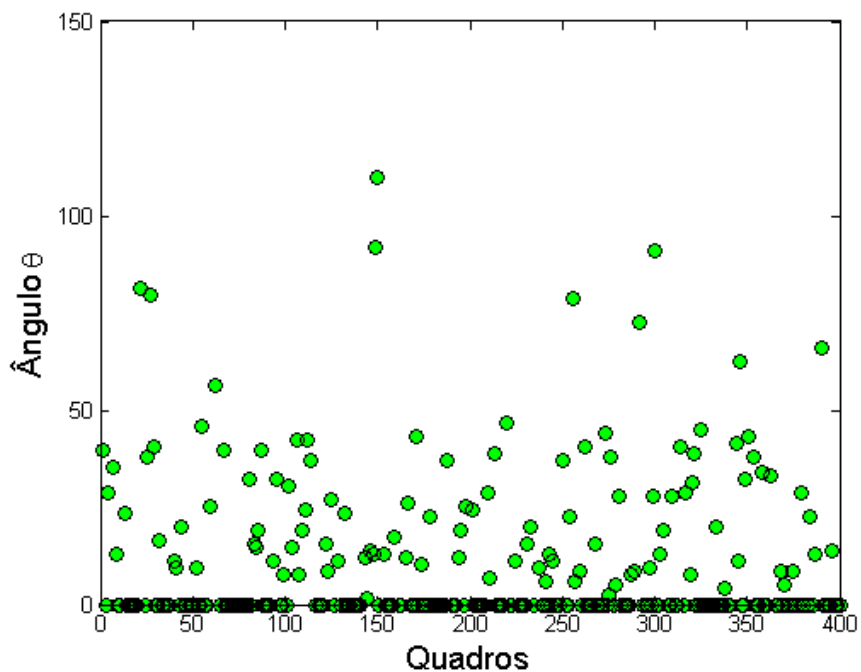


Figura 27: O ângulo  $\theta$  para cada par de quadros adjacentes estimado pelo método proposto para os 400 primeiros quadros do conjunto teste.

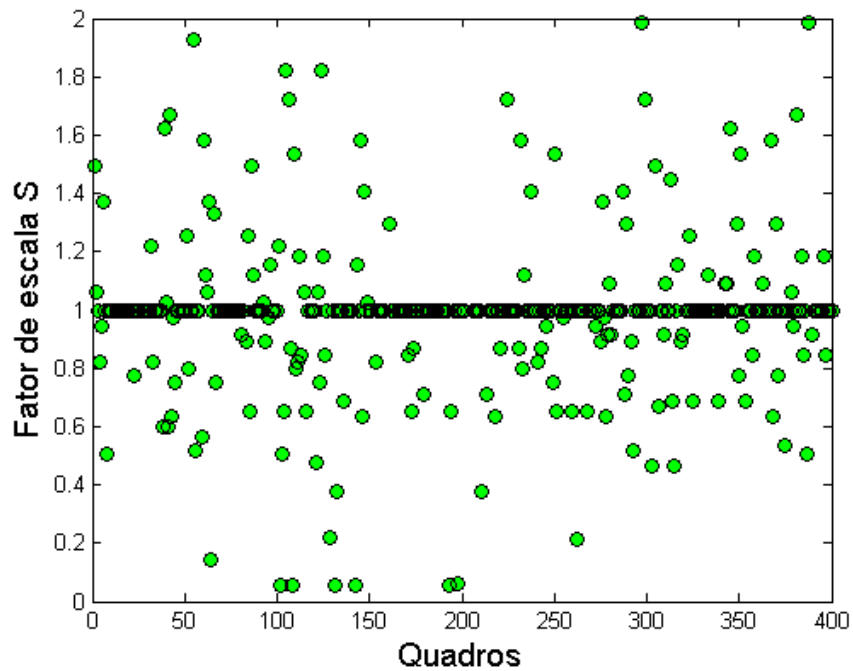


Figura 28: O fator de escala  $S$  para cada par de quadros adjacentes estimado pelo método proposto para os 400 primeiros quadros do conjunto teste.

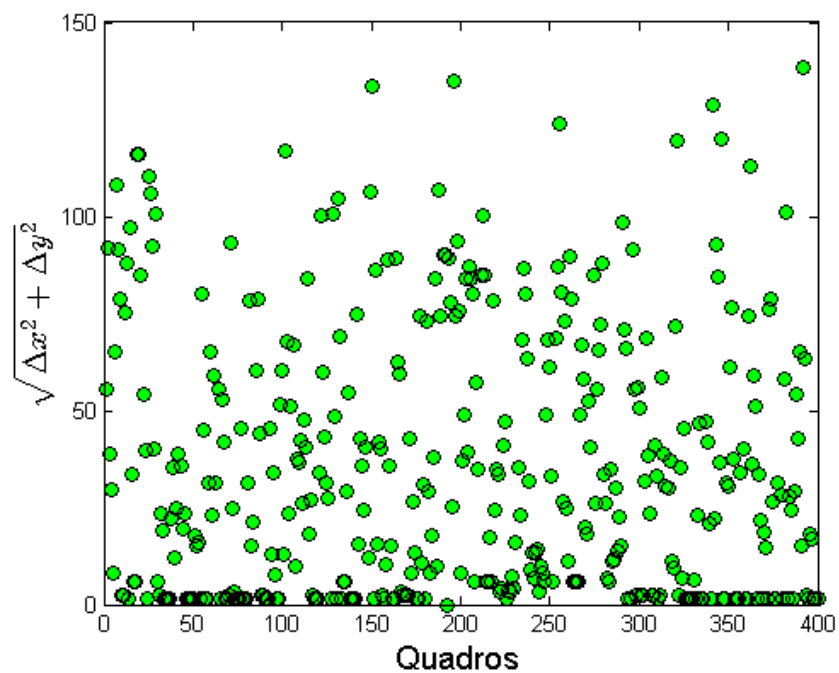


Figura 29: A translação  $\sqrt{\Delta x^2 + \Delta y^2}$  para cada par de quadros adjacentes estimada pelo método proposto para os 400 primeiros quadros do conjunto teste.

A Tabela 7 mostra os resultados comparativos entre o método proposto de detecção e estimação da retina e o método em Reddy e Chatterji (1996). O método elaborado obtém resultados melhores em termos de  $VPP$ ,  $ACC$  e  $FDM$  para os conjuntos de

treinamento e teste. Também foram comparados os parâmetros de movimento utilizando a *Peak Signal-to-Noise Rate* (PSNR - taxa de pico sinal-ruído) entre cada quadro  $I(t)$  e o quadro  $I(t + 1)$  corrigidos pelos parâmetros de movimento estimados por ambos os métodos. A PSNR média para os conjuntos de treinamento e teste estão na Tabela 8.

Tabela 7: Resultados comparativos para diferentes figuras de mérito.

	Método proposto(%)			Método em (REDDY; CHATTERJI, 1996)(%)		
	VPP	ACC	FDM	VPP	ACC	FDM
Conjunto treino	95.51	75.47	85.48	73.96	73.76	73.86
Conjunto teste	95.79	74.65	85.22	69.51	68.95	69.23

Tabela 8: Resultados comparativos em termos da média do PSNR.

	$\overline{PSNR}$	
	Método proposto(dB)	(REDDY; CHATTERJI, 1996)(dB)
Conjunto treino	35.58	28.48
Conjunto teste	32.17	28.25

A Figura 30 mostra os histogramas da PSNR entre cada par de quadros compensados nos conjuntos de treinamento e teste para o método proposto e para o método comparativo (REDDY; CHATTERJI, 1996). Nessa figura percebe-se que os histogramas em (a) e (b) estão deslocados para valores mais baixos de PSNR em relação aos histogramas em (c) e (d). Isso reflete em uma estimativa pior dos parâmetros de movimento da retina para o método (REDDY; CHATTERJI, 1996) em relação aos parâmetros estimados pelos algoritmos propostos nesta dissertação.

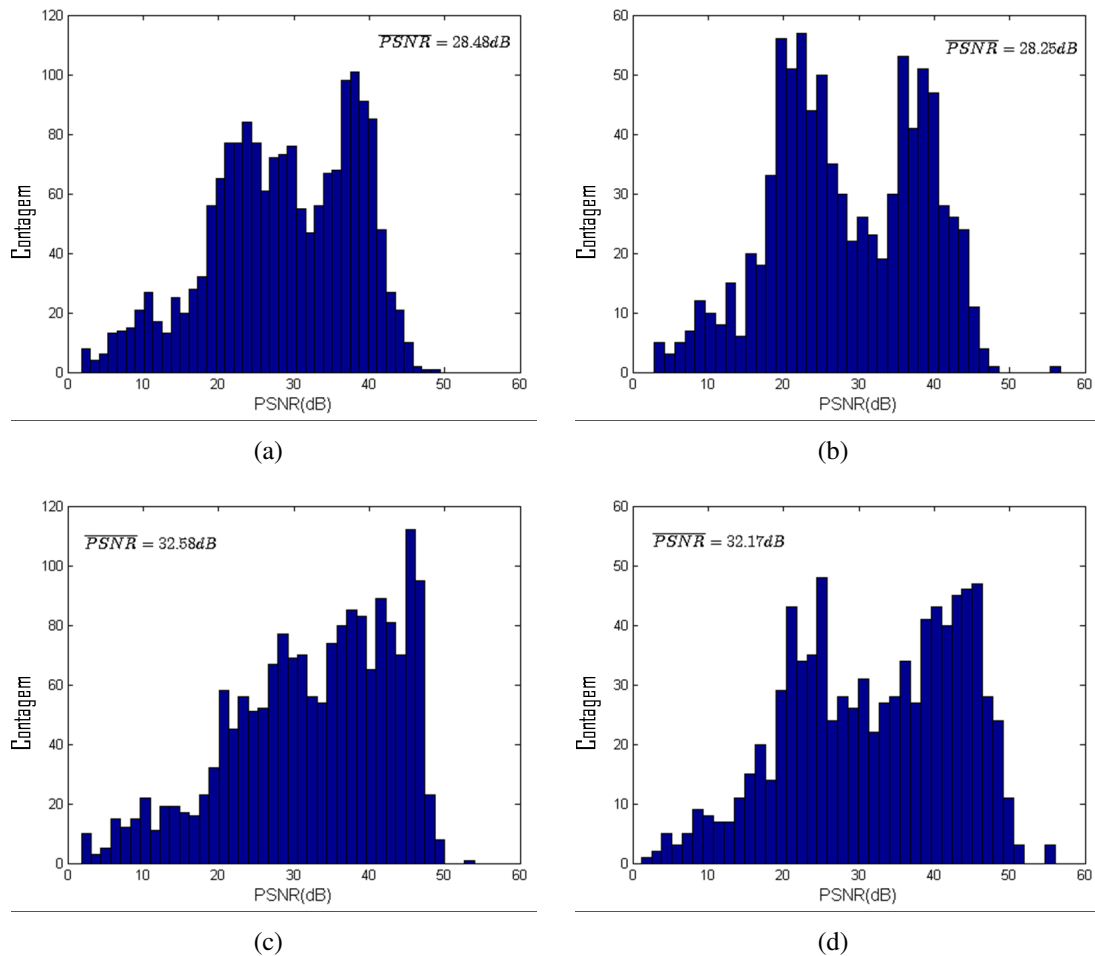
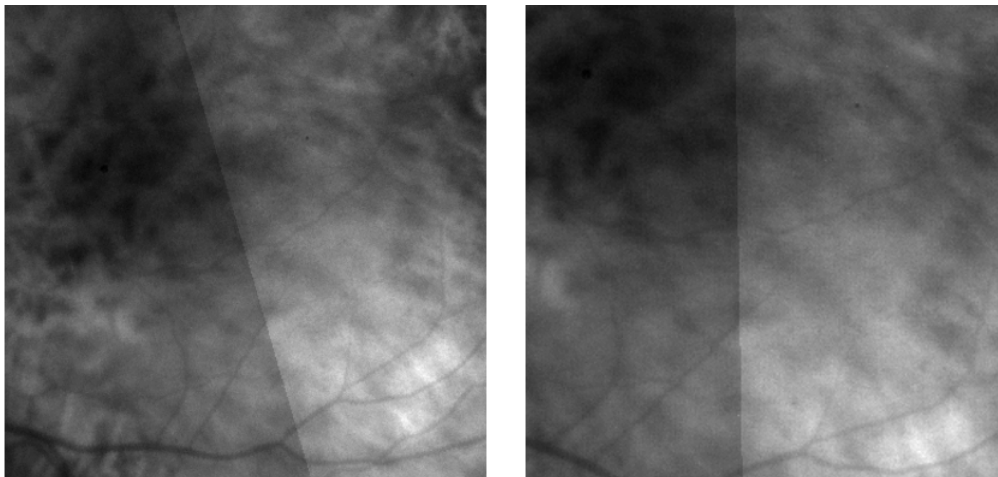


Figura 30: Histograma da PSNR entre cada par de quadros para o método proposto e o método em (REDDY; CHATTERJI, 1996) nos conjuntos de treinamento e teste, respectivamente: (a) método proposto em (REDDY; CHATTERJI, 1996) no conjunto de treinamento; (b) método proposto em (REDDY; CHATTERJI, 1996) no conjunto de teste; (c) método proposto no conjunto de treinamento; (d) método proposto no conjunto de teste.

Para ilustrar os resultados de estimação de movimento visualmente, um exemplo para cada um dos 7 vídeos da retina em infravermelho é apresentado nas Figuras 31 - 37, onde  $I_n(t)$  e  $I_n(t+1)$  são os quadros adjacentes e  $n$  indica o número do quadro dentro do respectivo vídeo. Os quadros  $I_n(t+1)$  corrigidos possuem uma moldura vermelha na borda entre os quadros  $I_n(t)$  e  $I_n(t+1)$  que os compõem. E a taxa PSNR é calculada entre as imagens  $I_n(t)$  e  $I_n(t+1)$  corrigida.



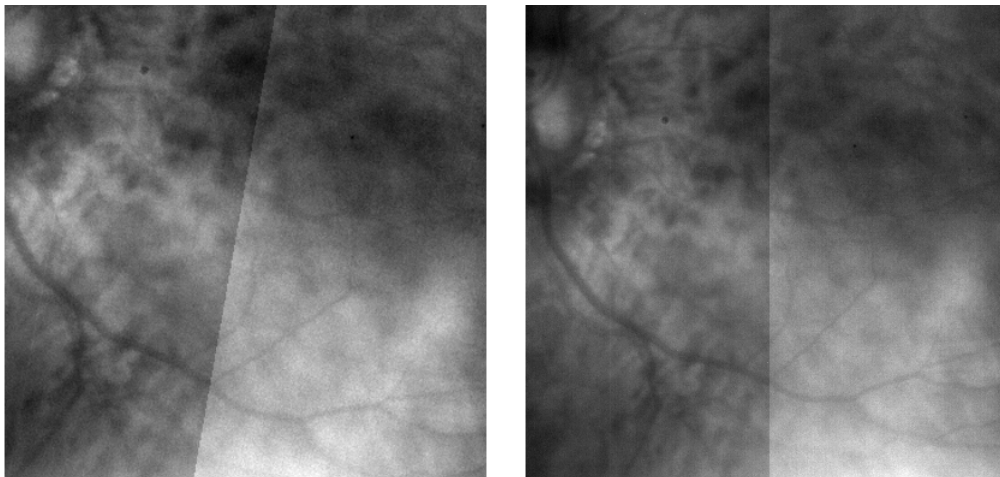
(a)

(b)



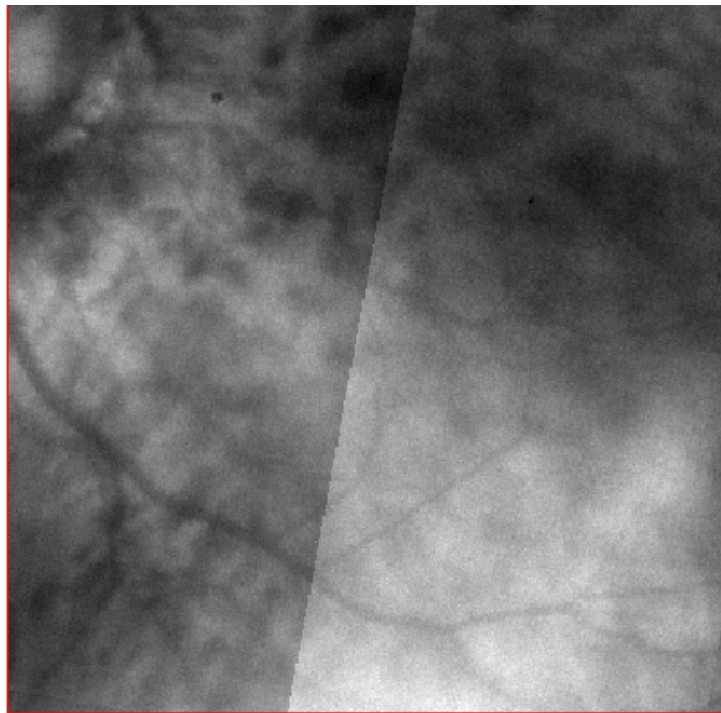
(c)

Figura 31: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #1, (a)  $I_{222}(t)$ ; (b)  $I_{223}(t + 1)$ ; (c)  $I_{223}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = -15.58^\circ$ ,  $S = 1.36$ ,  $\Delta x = -23$  e  $\Delta y = -21$ , atingindo um  $PSNR = 42.76dB$ . Métrica  $D(X, Y) = 0.9815$  e entropia conjunta  $H(X, Y) = 2.05$



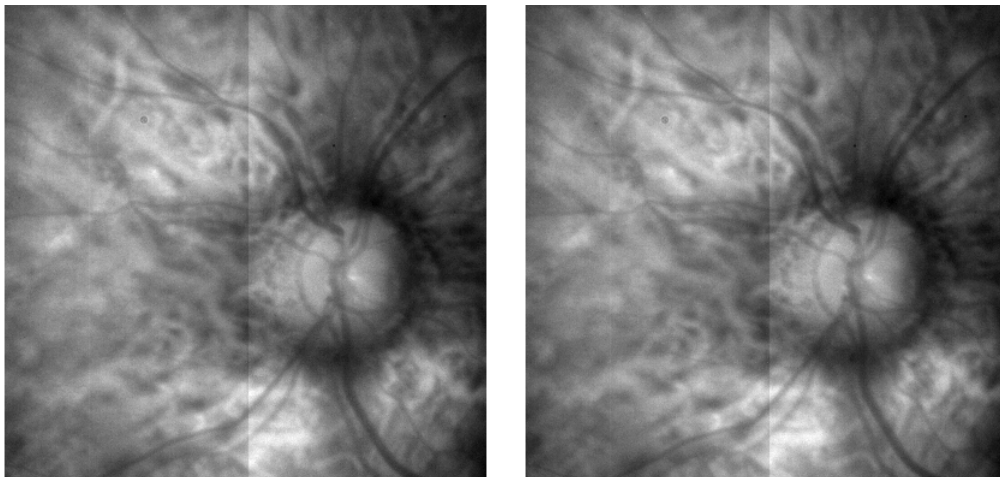
(a)

(b)



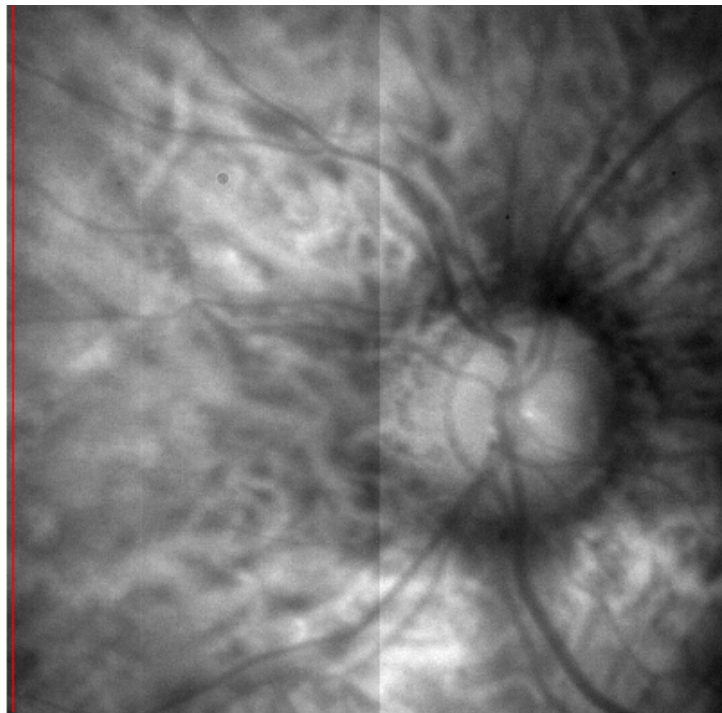
(c)

Figura 32: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #2, (a)  $I_{625}(t)$ ; (b)  $I_{626}(t+1)$ ; (c)  $I_{625}(t+1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 10.39^\circ$ ,  $S = 0.86$ ,  $\Delta x = 9$  e  $\Delta y = 10$ , atingindo um  $PSNR = 40.10dB$ . Métrica  $D(X, Y) = 0.9823$  e entropia conjunta  $H(X, Y) = 1.9222$



(a)

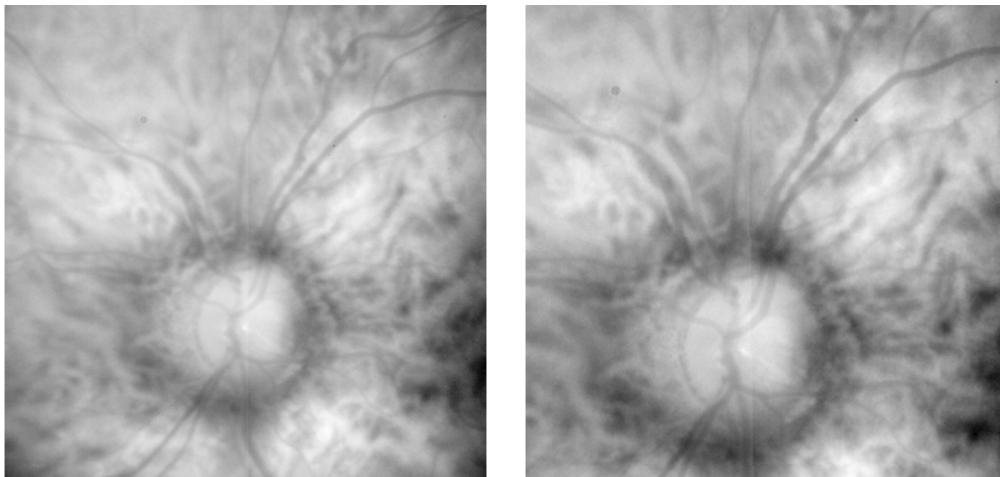
(b)



(c)

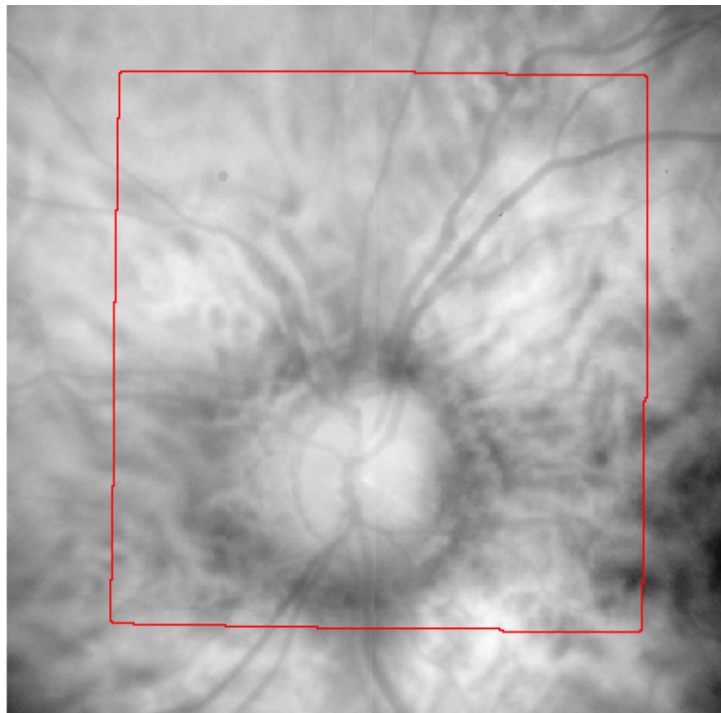
Figura 33: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #3, (a)  $I_{61}(t)$ ; (b)  $I_{62}(t+1)$ ; (c)  $I_{61}(t+1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 0^\circ$ ,  $S = 1$ ,  $\Delta x = -4$  e  $\Delta y = 1$ , atingindo um  $PSNR = 40.87dB$ . Métrica  $D(X, Y) = 0.9742$  e entropia conjunta  $H(X, Y) = 2.2725$





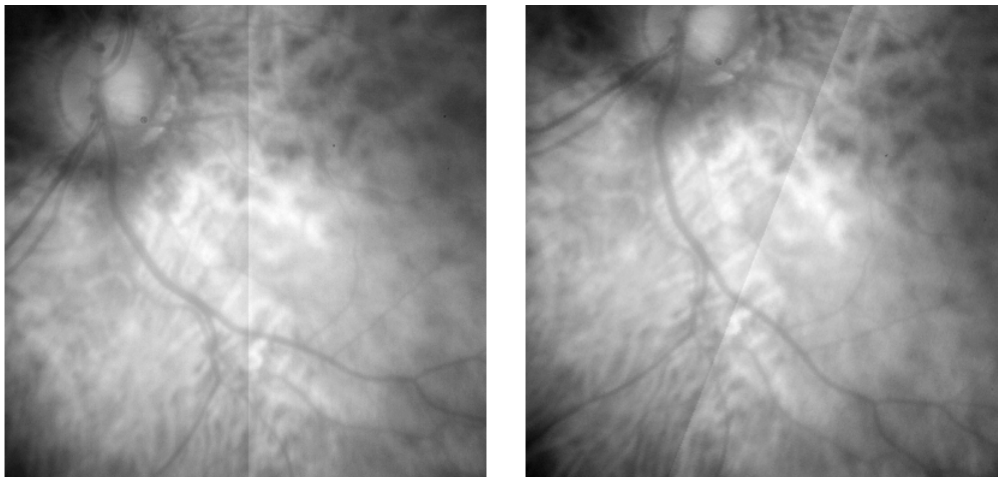
(a)

(b)



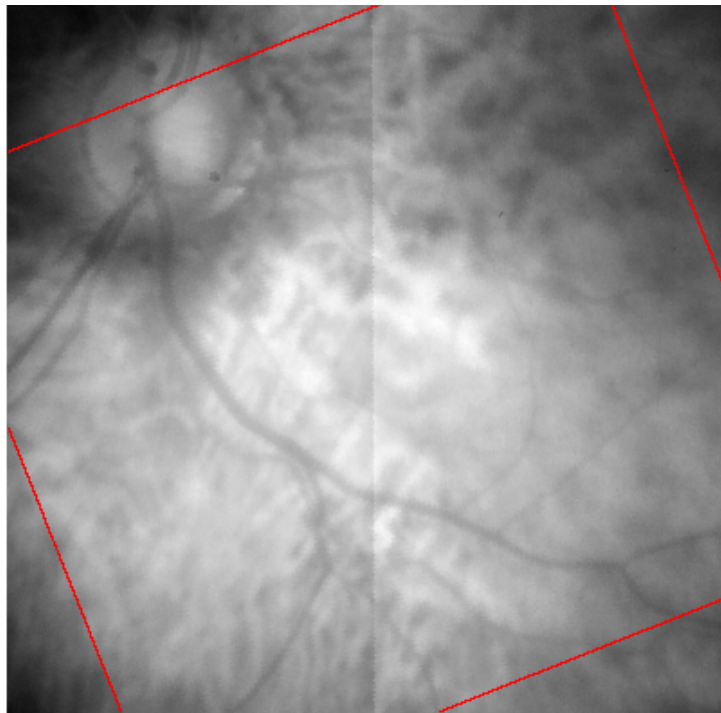
(c)

Figura 34: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #4, (a)  $I_{95}(t)$ ; (b)  $I_{96}(t + 1)$ ; (c)  $I_{95}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 0.86^\circ$ ,  $S = 1.29$ ,  $\Delta x = -20$  e  $\Delta y = 2$ , atingindo um  $PSNR = 38.30dB$ . Métrica  $D(X, Y) = 0.9794$  e entropia conjunta  $H(X, Y) = 1.7502$



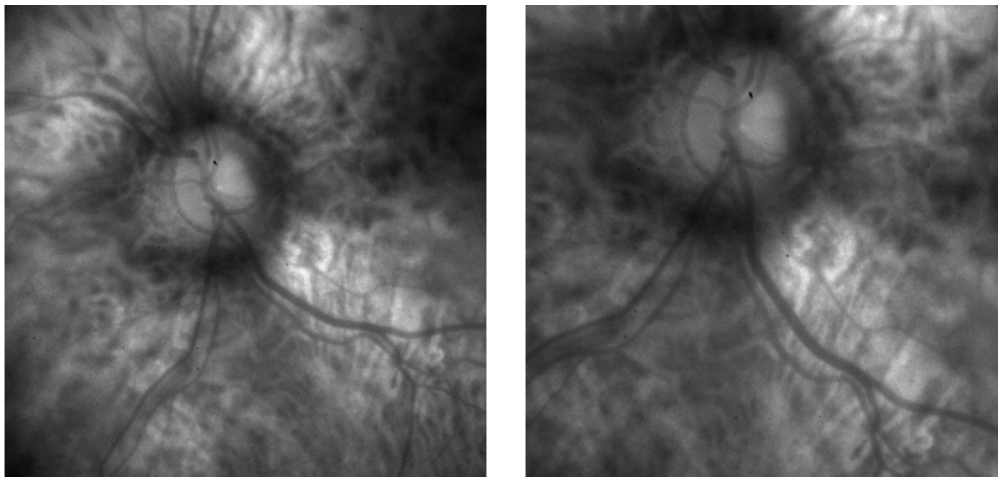
(a)

(b)



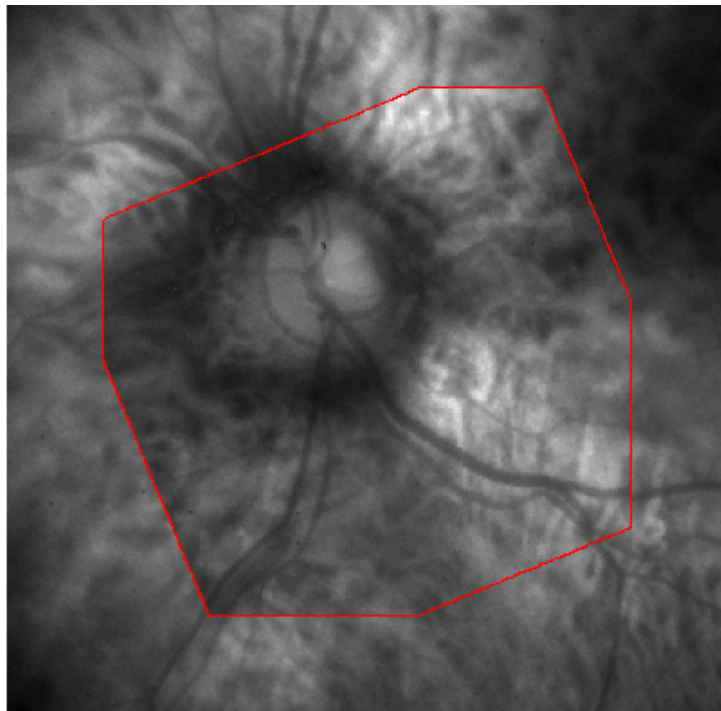
(c)

Figura 35: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #5, (a)  $I_{37}(t)$ ; (b)  $I_{38}(t+1)$ ; (c)  $I_{37}(t+1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = -21.65^\circ$ ,  $S = 1$ ,  $\Delta x = 3$  e  $\Delta y = -24$ , atingindo um  $PSNR = 40.33dB$ . Métrica  $D(X, Y) = 0.9815$  e entropia conjunta  $H(X, Y) = 1.9912$



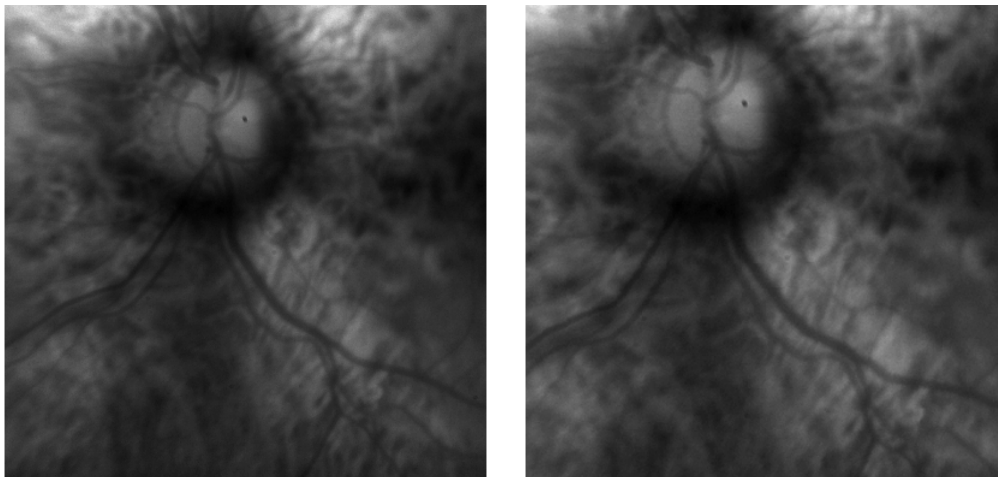
(a)

(b)



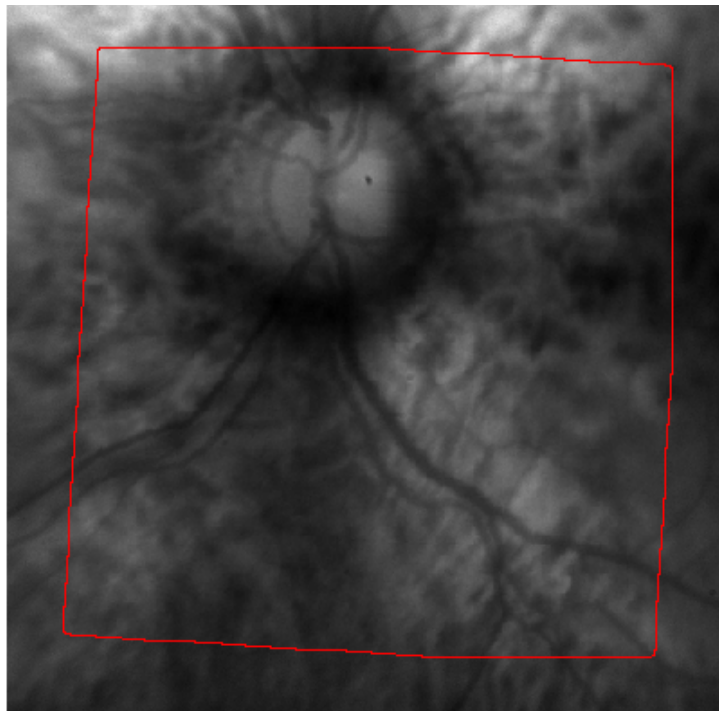
(c)

Figura 36: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #6, (a)  $I_{191}(t)$ ; (b)  $I_{192}(t + 1)$ ; (c)  $I_{191}(t + 1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = -22.51^\circ$ ,  $S = 1.36$ ,  $\Delta x = -13$  e  $\Delta y = -31$ , atingindo um  $PSNR = 36.59dB$ . Métrica  $D(X, Y) = 0.9848$  e entropia conjunta  $H(X, Y) = 2.4488$



(a)

(b)



(c)

Figura 37: Ilustração de estimação de movimento para um par de quadros adjacentes do vídeo #7, (a)  $I_{204}(t)$ ; (b)  $I_{205}(t+1)$ ; (c)  $I_{204}(t+1)$  corrigido por  $-\theta$ ,  $1/S$ ,  $-\Delta x$  e  $-\Delta y$ . Os parâmetros estimados são  $\theta = 3.46^\circ$ ,  $S = 1.18$ ,  $\Delta x = -11$  e  $\Delta y = 3$ , atingindo um  $PSNR = 38.23dB$ . Métrica  $D(X, Y) = 0.9742$  e entropia conjunta  $H(X, Y) = 2.0064$

### 4.3 Discussão

Esta seção analisa os resultados experimentais obtidos para as 2 análises de variâncias realizadas e para os resultados de detecção e estimação do movimento da retina. A seguir,

a discussão de cada uma das tabelas e figuras apresentadas anteriormente.

A Tabela 5, apresentada na Seção 4.1, exhibe os resultados da ANOVA para os limiares de detecção  $\tau$ ,  $\eta$  e a correlação  $\tau * \eta$ . A análise mostra que as variâncias destes limiares e a sua correlação são significativos para a variável de resposta  $FDM$ . Ou seja, estes limiares são críticos para a detecção de movimento da retina e foram otimizados de acordo com a Eq. 24.

A Tabela 6, apresentada na Seção 4.1, exhibe os resultados da ANOVA para os limiares de estimação  $\alpha$ ,  $\sigma$ ,  $\gamma$  e suas correlações  $\alpha * \sigma$ ,  $\alpha * \gamma$  e  $\sigma * \gamma$ . A análise mostra que as variâncias destes limiares são significativas para a variável de resposta  $FDM$ , mas as suas correlações não. Ou seja, estes limiares são críticos independentemente entre si para a estimação do movimento da retina e foram otimizados de acordo com a Eq. 25.

A Figura 25 mostra os valores da métrica de distância  $D(X, Y)$  para cada par de quadros ao longo do vídeo do conjunto teste. Lembrando que  $X = |\overrightarrow{\nabla I(t)}|$  e  $Y = |\overrightarrow{\nabla I(t+1)}|$  são as magnitudes dos gradientes das imagens. É notável que  $D(X, Y)$  é uma métrica confinada no intervalo  $[0, 1]$  e os valores de  $D(X, Y)$  menores que  $\tau = 0.9527$  indicam possivelmente uma situação de pequeno movimento.

A Figura 26 ilustra os valores da entropia conjunta  $H(X, Y)$  para cada par de quadros ao longo do vídeo do conjunto teste. É notável que  $H(X, Y)$  é uma métrica que responde rapidamente para a detecção de movimentos inconsistentes (piscadas, borrões, cena fora de foco). Dessa forma, sempre que os valores de  $H(X, Y)$  são menores que  $\eta = 1.6258$  indicam uma situação atípica.

A Figura 27 evidencia os ângulos  $\theta$  entre cada par de quadros, estimados pelo método proposto. Movimentos da retina com valores de rotação maiores que o limiar otimizado  $\alpha = 38.97^\circ$  não serão estimados. É possível verificar a capacidade do método em estimar desde rotações perto do ângulo zero até rotações acima de  $100^\circ$ .

A Figura 28 exhibe os fatores de escala  $S$  entre cada par de quadros, estimados pelo método proposto. Movimentos da retina com fatores de escala maiores que o limiar otimizado  $\sigma = 0.41$ , ou seja, uma ampliação maior que 1.41 ou uma redução menor que 0.59 não serão estimados também. É possível verificar a capacidade do método em estimar

escalas para reduções menores que 0.2 e ampliações maiores que 2 vezes.

A Figura 29 exhibe as translações  $\Delta x$  e  $\Delta y$  na forma do fator de translação  $\sqrt{\Delta x^2 + \Delta y^2}$  entre cada par de quadros, estimados pelo método proposto. Movimentos da retina com fator de translação maior que o limiar otimizado  $\gamma = 50.15$  não serão estimados. O método proposto consegue estimar translações entre quadros da retina em imagens em infravermelho de até 140 *pixels*.

A Tabela 7 mostra os resultados comparativos para as diferentes figuras de méritos para o conjunto treino e conjunto teste entre o método proposto e o método em Reddy e Chatterji (1996). O método proposto nesta dissertação alcança maiores valores em todas as figuras de mérito testadas em relação ao método comparativo.

A Tabela 8 resume os resultados comparativos em termos da média da taxa PSNR ( $\overline{PSNR}$ ) calculada entre os quadros corrigidos para ambos os métodos. A informação  $\overline{PSNR}$  fornece um valor quantitativo para a qualidade dos parâmetros de deslocamentos estimados pelo método proposto e o método comparativo (REDDY; CHATTERJI, 1996). Novamente o método desenvolvido nesta dissertação atinge valores melhores tanto no conjunto treino quanto no conjunto teste.

Já a Figura 30 ilustra os resultados para os valores da PSNR entre todos os quadros estimados para os dois conjuntos treino e teste. Os resultados estão exibidos em histogramas que contam o número de ocorrência dos valores de PSNR. Nota-se que há um deslocamento para a direita das distribuições no histograma da Figura 30(c) e (d) para o método proposto em relação às distribuições (a) e (b) do método (REDDY; CHATTERJI, 1996). Este resultado mostra que os parâmetros estimados pelo método proposto são melhores do que o método comparativo.

Evidencia-se que o método proposto possui um desempenho melhor nos resultados experimentais obtidos em comparação ao método de Reddy e Chatterji (1996) e tão importante quanto obter resultados melhores é poder validar o método proposto com resultados já estabelecidos na literatura.

## 5 CONCLUSÃO

Esta dissertação foca na detecção e estimação do movimento da retina e a sua aplicação em sistemas de tratamentos para EMD. Qualquer movimento da retina pode fazer com que o laser atinja uma área saudável, o que degradaria ainda mais a acuidade visual do paciente. O método proposto utiliza a informação mútua  $MI(X, Y)$ , a entropia conjunta  $H(X, Y)$  e a métrica de distância  $D(X, Y)$  para detectar o movimento entre quadros. Além disso, é desenvolvida uma abordagem com base nos métodos de deslocamento de fase e na transformação log-polar para estimar os deslocamentos espaciais frente a rotações, escalas e translações que ocorrem em sequências de imagens da retina em infravermelho. Desta forma, foram utilizados 2 limiares para detecção do movimento com base nas métricas  $D(X, Y)$  e  $H(X, Y)$  e mais 3 limiares relacionados aos parâmetros de deslocamentos  $\theta$ ,  $S$ ,  $\sqrt{\Delta x^2 + \Delta y^2}$ . Para analisar a significância desses limiares foram projetados 2 experimentos de fatorial completo a fim de verificar os parâmetros críticos dos algoritmos de detecção e estimação de movimento invariante à rotação, escala e translação em sequências de imagens em infravermelho da retina. Desta forma, todos os 5 limiares foram avaliados criticamente pelas análises de variâncias realizadas e foram otimizados em 2 conjuntos distintos.

Sendo assim, o método proposto alcança uma figura de mérito  $FDM = 85.22\%$  mesmo nas imagens em infravermelho tipicamente ruidosas e com mudanças constantes de iluminação. Os resultados experimentais indicam que o método proposto pode potencialmente detectar e estimar movimentos da retina, com resultados melhores em comparação com outras abordagens descritas na literatura, melhorando a confiabilidade

dos sistemas de tratamento a laser EMD.

## 5.1 Trabalhos Futuros

Durante o mestrado foram desenvolvidos métodos de detecção e estimação de movimento da retina que são adaptáveis ao ruído de fundo e a mudanças na iluminação (SCHARCANSKI et al., 2013; SCHARDOSIM; SCHARCANSKI, 2014). Os resultados experimentais indicam que estes métodos são adequados para a detecção e estimação robustas a rotação, escala e translação de movimento da retina em sequências de imagens de infravermelho, o que são passos importantes em sistemas de tratamento a laser na retina para EMD. Dessa forma, no doutorado, pretende-se expandir estas técnicas de detecção e estimação de movimentos para rastrear em vídeos diferentes tipos de objetos.

Rastreamento visual de objetos é uma tarefa significativa da visão computacional (LI et al., 2013), que pode ser aplicado a muitos domínios, como a vigilância visual, interação homem-computador e compressão de vídeo. Apesar de uma extensa pesquisa sobre o tema, que ainda sofre com dificuldades em lidar com as mudanças complexas de aparência de objetos causadas por fatores como a variação de iluminação, oclusão parcial, deformação e movimento de câmera. Portanto, a modelagem eficaz da aparência 2D de objetos rastreados é uma questão-chave para o sucesso do rastreamento de objetos.



## APÊNDICE A CONTRIBUIÇÕES DESTE TRABALHO

O trabalho desenvolvido durante o mestrado gerou três submissões. A primeira contribuição foi a submissão e publicação na revista internacional *Computerized Medical Imaging and Graphics* (CMIG), (SCHARCANSKI et al., 2013), sob o título "*Motion detection and compensation in infrared retinal image sequences*" onde os principais assuntos abordados foram o método proposto para a detecção do movimento da retina, (ver Seção 3.1), a detecção de inconsistências, (ver Seção 3.2) e a estimação de movimentos da retina frente a translações utilizando a técnica de deslocamento de fase, (ver Seção 3.3).

A segunda contribuição foi a submissão e apresentação de artigo na conferência em Montevideo (Uruguai) *International Instrumentation and Measurement Technology Conference* (I2MTC 2014), (SCHARDOSIM; SCHARCANSKI, 2014), sob o título "*Phase-Shift and Information Measures for Retinal Motion Detection and Compensation with Optimized Parameter Selection*" onde as principais colaborações estão relacionadas à classificação do tipo de movimento da retina, (ver Seção 3.5), e ao projeto de experimento de fatorial completo para a seleção dos parâmetros críticos do algoritmo de detecção e compensação da retina, (ver Seções 3.6) e 4.1.

A terceira contribuição foi o artigo submetido em junho de 2014 para a edição especial da conferência I2MTC 2014 que poderá ser aceito na revista internacional *IEEE Transactions on Instrumentation and Measurement*, sob o título "*Motion Detection and Compensation Invariant to Rotation, Scale and Translation for Infrared Retinal Image Sequences*". As principais contribuições deste artigo estão relacionadas à detecção e compensação de movimento invariante à rotação, escala e translação em sequências de

imagens em infravermelho da retina, (ver Seção 3.3).

## REFERÊNCIAS

- BISHOP, C. M. **Pattern Recognition and Machine Learning (Information Science and Statistics)**. Secaucus, NJ, USA: Springer-Verlag New York, 2006.
- BOULTON, M. et al. Age-related changes in the morphology, absorption and fluorescence of melanosomes and lipofuscin granules of the retinal pigment epithelium. **Vision research**, Oxford, v.30, n.9, p.1291–1303, 1990.
- BRESSLER, N. M. et al. Clinicopathologic correlation of drusen and retinal pigment epithelial abnormalities in age-related macular degeneration. **Retina**, Hagerstown, v.14, n.2, p.130–142, 1994.
- CHANWIMALUANG, T.; FAN, G.; FRANSEN, S. Hybrid retinal image registration. **IEEE Transactions on Information Technology in Biomedicine**, New York, v.10, n.1, p.129–142, 2006.
- ELSNER, A. E. et al. Evaluating the photoreceptor/rpe complex with and SLO, Non-invasive Assessment of the Visual System. **Technical Digest, Optical Society of America**, Washington, D.C, v.3, p.40 – 43, 1990.
- ELSNER, A. E. et al. Infrared imaging of sub-retinal structures in the human ocular fundus. **Vision Research**, Oxford, v.36, n.1, p.191 – 205, 1996.
- JCGM. **International vocabulary of metrology - Basic and general concepts and associated terms (VIM)**, Sevres Cedex, France, 2008.
- KERNT, M. et al. Focal and panretinal photocoagulation with a navigated laser (NAVILAS®). **Acta Ophthalmologica**, Oxford, v.89, n.8, p.662–664, 2011.
- KOZAK, I. Determining clinical performance and treatment accuracy of navigated retinal photocoagulation. **Retina Today**, Wayne, PA, v.118, n.2, p.72–73, 2011.
- KOZAK, I. et al. Clinical Evaluation and Treatment Accuracy in Diabetic Macular Edema Using Navigated Laser Photocoagulator NAVILAS. **Ophthalmology**, Rochester, v.118, n.6, p.1119 – 1124, 2011.
- KRASKOV, A. et al. Hierarchical clustering using mutual information. **EPL (Europhysics Letters)**, Paris, v.70, n.2, p.278, 2005.
- LI, H. et al. Tracking features in retinal images of adaptive optics confocal scanning laser ophthalmoscope using KLT-SIFT algorithm. **Biomedical Optics Express**, [S.l.], v.1, n.1, p.31–40, Aug 2010.

LI, X. et al. A Survey of Appearance Models in Visual Object Tracking. **ACM Transactions on Intelligent Systems and Technology**, New York, NY, USA, v.4, n.4, p.1–48, Oct. 2013.

LIU, N. et al. An Intelligent Scoring System and Its Application to Cardiac Arrest Prediction. **Information Technology in Biomedicine, IEEE Transactions on**, [S.l.], v.16, n.6, p.1324–1331, Nov 2012.

MONTGOMERY, D. C. **Design and analysis of experiments**. 5th.ed. New York / Chichester: John Wiley & Sons, 2001.

NARASIMHA-IYER, H. et al. Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy. **IEEE Transactions on Biomedical Engineering**, [S.l.], v.53, n.6, p.1084–1098, Jun 2006.

NOGUEIRA, V. et al. Diabetic retinopathy - The role of Family Medicine. **Revista Portuguesa de Medicina Geral e Familiar**, Lisboa, v.23, n.5, 2007.

REDDY, B. S.; CHATTERJI, B. N. An FFT-based Technique for Translation, Rotation, and Scale-invariant Image Registration. **IEEE Transactions on Image Processing**, Piscataway, NJ, USA, v.5, n.8, p.1266–1271, Aug. 1996.

REIS, A. d. C. A. et al. Injeção intravítrea de cetorolaco de trometamina em pacientes com edema macular diabético refratário à fotocoagulação retiniana. **Arquivos Brasileiros de Oftalmologia**, Sao Paulo, v.73, p.338 – 342, Aug. 2010.

REN, J.; JIANG, J.; VLACHOS, T. High-Accuracy Sub-Pixel Motion Estimation From Noisy Images in Fourier Domain. **IEEE Transactions on Image Processing**, New York, v.19, n.5, p.1379–1384, 2010.

SCHARCANSKI, J. et al. Motion detection and compensation in infrared retinal image sequences. **Computerized Medical Imaging and Graphics**, [S.l.], v.37, n.5, p.377–385, 2013.

SCHARDOSIM, L. et al. Motion detection in infrared retinal image sequences. In: **IEEE INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC 2011)**, 2011., Binjiang. **Proceedings...** New York: IEEE, 2011. p.1–5.

SCHARDOSIM, L.; SCHARCANSKI, J. Phase-Shift and Information Measures for Retinal Motion Detection and Compensation with Optimized Parameter Selection. In: **IEEE INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC 2014)**, 2014., Montevideo. **Proceedings...** New York: IEEE, 2014. p.124–129.

WOLBERG, G.; ZOKAI, S. Robust image registration using log-polar transform. In: **IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING**, 2000., Vancouver. **Proceedings...** New York: IEEE, 2000. v.1, p.493–496.

XU, Z. et al. Tracking retinal motion with a scanning laser ophthalmoscope. **Journal of Trhabilitation Research and Development**, [S.l.], v.42, n.3, p.373–380, 2005.

ZHAO, Q. et al. Digital facial dysmorphology for genetic screening: hierarchical constrained local model using {ICA}. **Medical Image Analysis**, [S.l.], v.18, n.5, p.699 – 710, 2014.

ZHENG, J. et al. Retinal image registration based on salient feature regions. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC 2009), 2009., Minneapolis. **Proceedings...** New York: IEEE, 2009. v.707, p.102–105.