UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS

JOSÉ AUGUSTO STUCHI

Registro de Imagens por Correlação de Fase para
Geração de Imagens Coloridas em Retinógrafos
Digitais Utilizando Câmera CCD Monocromática

São Carlos
2013

JOSÉ AUGUSTO STUCHI

Registro de Imagens por Correlação de Fase para Geração de Imagens Coloridas em Retinógrafos Digitais Utilizando Câmera CCD Monocromática

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências, Programa de Engenharia Elétrica.

Área de concentração: Processamento de Sinais e Instrumentação

Orientador: Prof. Dr. Marcelo Andrade da Costa Vieira

São Carlos
2013

Trata-se de versão corrigida da dissertação. A versão original se encontra disponível na EESC/USP que aloja o programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

S932r Stuchi, José Augusto
Registro de imagens por correlação de fase para
geração de imagens coloridas em retinógrafos digitais
utilizando câmera CCD monocromática / José Augusto
Stuchi; orientador Marcelo Andrade da Costa Vieira. São
Carlos, 2013.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2013.

1. Retinógrafo digital. 2. Registro de imagens. 3.
Correlação de fase. 4. Imagem colorida RGB. 5. Câmera
CCD. 6. Iluminação a LED. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **JOSÉ AUGUSTO STUCHI**.

Título da dissertação: "Registro de imagens por correlação de fase para geração de imagens coloridas em retinógrafos digitais utilizando câmera CCD monocromática".

Data da defesa: 10/06/2013

Comissão Julgadora:

Resultado:

Prof. Dr. **Marcelo Andrade da Costa Vieira (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Associado **Evandro Luís Linhari Rodrigues**
(Escola de Engenharia de São Carlos/EESC)

Aprovado

Prof. Dr. **Alexandre Luis Magalhães Levada**
(Universidade Federal de São Carlos/UFSCar)

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e Presidente da
Comissão de Pós-Graduação:

Prof. Titular **Denis Vinicius Coury**

Aos meus pais, José Ramiro e Augusta, pelo incentivo incondicional e educação, além de serem exemplos para mim de caráter, ética e força de vontade.

À minha irmã Nádia, meu cunhado Fernando e minha sobrinha Caroline, que sempre me apoiaram e incentivaram nos estudos, mesmo com os atos mais simples.

À minha noiva Roberta que me encorajou no desenvolvimento desse trabalho e que, além de ser minha companheira querida, é a pessoa com a qual aprendo muito e, na maioria das vezes, tento espelhar meus atos.

AGRADECIMENTOS

Ao Prof. Dr. Marcelo Andrade da Costa Vieira, pelo incentivo, paciência e inúmeros ensinamentos e considerações acadêmicas fornecidos no desenvolvimento desse trabalho.

Ao Prof. Dr. Evandro Luís Linhari Rodrigues pelo apoio nesse trabalho e pelas inúmeras orientações sobre o mercado de trabalho que tem me disponibilizado desde a graduação.

Ao Prof. Dr. Jacob Scharcanski, da Universidade Federal do Rio Grande do Sul, pelos ensinamentos iniciais fornecidos relacionados ao tema de registro de imagens utilizado nesse trabalho.

À Universidade de São Paulo e aos professores que me ajudaram a desenvolver meus conhecimentos desde o início da graduação.

Aos meus colegas da USP, que muitas vezes me ajudaram a suportar momentos de grande *stress* de maneira bem humorada e, além disso, foram responsáveis por grande parte dos conhecimentos que obtive no período.

À sociedade que custeou os meus estudos durante graduação e mestrado. Que eu possa aproveitar bem esse privilégio, que infelizmente muitos não tiveram, e trazer bons retornos à mesma.

Ao Dr. Mario Antonio Stefani pelo apoio e incentivo à pesquisa nas dependências da Opto Eletrônica S.A. e pela motivação em tentar fazer do Brasil um país diferente.

Ao Eng. Flávio Pascoal Vieira e ao Físico Diego Lencione, que tanto me ajudaram no desenvolvimento desse trabalho. Ambos me trazem muito orgulho em saber que temos recursos humanos tão capacitados em nosso país.

Ao Gerente Daniel dos Santos Júnior pela grande paciência, ensinamentos e força de vontade em superar as inúmeras dificuldades enfrentadas nesse projeto.

A todos os companheiros de trabalho da Opto, em especial André Martins, Glauco Zerbini, Eduardo Bandeira, Luís Teles, Elisa Barbarini, Ivan Filgueiras, Arineiza Pinheiro, Giuliano Rossi, Alessandro Mota, Tiago Ortega, Luiz Otoboni, Ailton Vaccare, Renato Romero e Saulo Bombonato que tiveram uma participação importantíssima para o sucesso deste projeto, influenciando grandemente no meu aprendizado.

À Samsung Eletrônica que me concedeu importantíssimas horas para a finalização deste trabalho.

A minha família que sempre me apoiou e investiu nos meus estudos, deixando até mesmo de realizar atividades básicas para o meu sucesso. Muito obrigado pela confiança!

A minha noiva Roberta, pelo amor, carinho, confiança e inúmeros incentivos no desenvolvimento desse trabalho. Por várias vezes ela entendeu com paciência a necessidade do cumprimento de alguma atividade relacionada ao mestrado, mesmo que isso custasse um momento em que poderíamos estar juntos. Obrigado pelo companheirismo!

Por fim, agradeço a Deus, o qual me deu muita força para superar inúmeras dificuldades, iluminação para enfrentar problemas difíceis e discernimento para que pudesse melhorar e me tornar uma pessoa mais madura.

RESUMO

Stuchi, J. A. **Registro de Imagens por Correlação de Fase para Geração de Imagens Coloridas em Retinógrafos Digitais Utilizando Câmera CCD Monocromática.** 2013. Dissertação de Mestrado – Departamento de Engenharia Elétrica e de Computação, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2013.

A análise da retina permite o diagnóstico de muitas patologias relacionadas ao olho humano. A qualidade da imagem é um fator importante já que o médico normalmente examina os pequenos vasos da retina e a sua coloração. O equipamento normalmente utilizado para a visualização da retina é o retinógrafo digital, que utiliza sensor colorido com filtro de Bayer e luz (*flash*) branca. No entanto, esse filtro causa perda na resolução espacial, uma vez que é necessário um processo de interpolação matemática para a formação da imagem. Com o objetivo de melhorar a qualidade da imagem da retina, um retinógrafo com câmera CCD monocromática de alta resolução foi desenvolvido. Nele, as imagens coloridas são geradas pela combinação dos canais monocromáticos R (vermelho), G (verde) e B (azul), adquiridos com o chaveamento da iluminação do olho com LED vermelho, verde e azul, respectivamente. Entretanto, o pequeno período entre os *flashes* pode causar desalinhamento entre os canais devido a pequenos movimentos do olho. Assim, este trabalho apresenta uma técnica de registro de imagens, baseado em correlação de fase no domínio da frequência, para realizar precisamente o alinhamento dos canais RGB no processo de geração de imagens coloridas da retina. A validação do método foi realizada com um olho mecânico (*phantom*) para a geração de 50 imagens desalinhadas que foram corrigidas pelo método proposto e comparadas com as imagens alinhadas obtidas como referência (*ground-truth*). Os resultados mostraram que retinógrafo com câmera monocromática e o método de registro proposto nesse trabalho podem produzir imagens coloridas da retina com alta resolução espacial, sem a perda de qualidade intrínseca às câmeras CCD coloridas que utilizam o filtro de Bayer.

Palavras chave: Retinógrafo Digital, Registro de Imagens, Correlação de Fase, Imagem Colorida RGB, Câmera CCD, Iluminação a LED.

ABSTRACT

Stuchi, J. A. **Image Registration Using Phase Correlation to Generate Color Images in Digital Fundus Cameras Using Monochromatic CCD camera.** 2013. Master Degree Dissertation – Department of Electrical and Computing Engineering, Engineering School of São Carlos, University of São Paulo, São Carlos, 2013.

The analysis of retina allows the diagnostics of several pathologies related to the human eye. Image quality is an important factor since the physician often examines the small vessels of the retina and its color. The device usually used to observe the retina is the fundus camera, which uses color sensor with Bayer filter and white light. However, this filter causes loss of spatial resolution, since it is necessary a mathematical interpolation process to create the final image. Aiming at improving the retina image quality, a fundus camera with monochromatic CCD camera was developed. In this device, color images are generated by combining the monochromatic channels R (red), G (green) and B (blue), which were acquired by switching the eye illumination with red, green and blue light, respectively. However, the short period between the flashes may cause misalignment among the channels because of the small movements of the eye. Thus, this work presents an image registration technique based on phase correlation in the frequency domain, for accurately aligning the RGB channels on the process of generating retina color images. Validation of the method was performed by using a mechanical eye (phantom) for generating 50 misaligned images, which were aligned by the proposed method and compared to the aligned images obtained as references (ground-truth). Results showed that the fundus camera with monochromatic camera and the method proposed in this work can produce high spatial resolution images without the loss of quality intrinsic to color CCD cameras that uses Bayer filter.

Keywords: Digital Fundus Camera, Image Registration, Phase Correlation, RGB Color Image, CCD Camera, LED Illumination.

LISTA DE FIGURAS¹

Figura 1: Esquema simplificado de formação de imagem pelo olho	19
Figura 2: Representação do olho humano.....	20
Figura 3: Exemplo de um retinógrafo digital	21
Figura 4: Esquemático de funcionamento de um retinógrafo digital com iluminação a LED.....	22
Figura 5: Exemplo de uma imagem de fundo de olho obtida com a técnica de retinografia colorida.	24
Figura 6: Exemplo de uma imagem de fundo de olho obtida com a técnica de retinografia anerítra.	25
Figura 7: Exemplo de uma imagem de fundo de olho obtida com a técnica de angiografia fluoresceínica	26
Figura 8: Dinâmica da circulação sanguínea do exame de angiografia fluoresceínica – precoce (esquerda), média (centro) e tardia (direita)	26
Figura 9: Exemplo de uma imagem de fundo de olho obtida com a técnica de angiografia com ICG.	27
Figura 10: Mosaico da retina formada pelo processamento de várias imagens obtidas em diferentes tempos	29
Figura 11: Imagens da retina iluminadas com diferentes comprimentos de onda (λ)	30
Figura 12: Análise da concentração de oxigênio da retina utilizando equipamento <i>Oxymap</i>	31
Figura 13: Arranjo básico do filtro de Bayer (esquerda) e esquema de filtros espectrais utilizado para a formação da imagem na qual a luz incidente é dividida nas componentes RGB (direita)	32
Figura 14: Formação da imagem colorida com filtro de Bayer na qual para cada posição da matriz final tem-se apenas a componente de uma das cores	32
Figura 15: Imagem simulada obtida com câmera monocromática (direita) e colorida (esquerda)	33
Figura 16: Exemplo de <i>color aliasing</i>	34
Figura 17: Câmera multiespectral com roda de filtros que permite a obtenção de diferentes λ	35
Figura 18: Imageamento colorido com o uso de três sensores, na qual filtros ópticos separam as componentes RGB e espelhos as encaminham para o respectivo sensor.....	36
Figura 19: Exemplo da iluminação da retina para a formação de imagem colorida com a arquitetura proposta.....	37
Figura 20: Esquema básico de captura de imagem colorida usando CCD monocromático conforme arquitetura proposta.....	38
Figura 21: CCD monocromático de 11 Megapixels utilizado na arquitetura proposta	39
Figura 22: LEDs utilizados para a obtenção das imagens coloridas na arquitetura proposta	40

¹ As fontes das imagens estão indicadas individualmente junto a sua ocorrência ao longo do trabalho. Para o caso dos créditos serem do próprio autor esta informação é omitida.

Figura 23: Protótipo montado a partir da arquitetura proposta	40
Figura 24: Exame colorido e RF obtidos com apenas um disparo na arquitetura proposta	41
Figura 25: Exemplo de imagem monocromática de ótima qualidade obtida com a arquitetura proposta (exame FA).....	42
Figura 26: <i>Color Checker</i> utilizado para a calibração de potência dos LEDs	42
Figura 27: Imagem colorida obtida antes (esquerda) e depois da calibração de cores (direita)	43
Figura 28: Área do nervo óptico com desalinhamento dos canais devido à movimentação do olho durante o processo de aquisição das imagens dos três canais RGB.....	44
Figura 29: Diferença de intensidade entre o canal vermelho (esquerda), verde (centro) e azul (direita) de exame realizado com olho doente	46
Figura 30: Passos básicos envolvidos no registro entre duas imagens.....	47
Figura 31: Relações geométricas para a correlação cruzada entre duas imagens	49
Figura 32: Relação da região de busca com a região do <i>template</i> no método de correlação cruzada	50
Figura 33: Imagem e <i>template</i> para busca (esquerda) e negativo da imagem da matriz da correlação cruzada obtida (direita) ao realizar o método de correlação cruzada entre o <i>template</i> e a imagem da retina.....	51
Figura 34: Esquemático do resultado da correlação de fase entre dois canais que permite encontrar a transformação de translação entre ambos – Nesse caso, é mostrado o negativo da imagem resultante, no qual o ponto preto representa a máxima correlação entre os dois quadros	54
Figura 35: Resultado para correlação de fase, mostrando o negativo do pico de correlação (parte inferior) entre as duas imagens (para superior) – o ponto preto representa o <i>offset</i> entre as duas imagens, sendo esse deslocamento calculado em relação ao centro do resultado da correlação de fase.....	54
Figura 36: Esquemático do resultado da correlação de fase da magnitude do sinal na forma <i>log-polar</i> de dois canais, o qual permite encontrar a escala e a rotação entre ambos – o ponto preto representa a máxima correlação entre as duas imagens	57
Figura 37: Paciente posicionado para o exame com cabeça fixa (testa e queixo apoiados) o que dificulta o desalinhamento entre os <i>flashes</i> por rotação e escala	59
Figura 38: Imagens correspondentes ao canal G (superior esquerda) e ao canal R (superior direita) e o resultado da correlação de fase (inferior), mostrando a aplicação do método em imagens com diferentes intensidades de informação devido à iluminação diferente de cada canal	60
Figura 39: Canais combinados (50% Canal R + 50% Canal G) antes (esquerda) e depois do registro (direita)	61

Figura 40: Imperfeição no sistema óptico gerando três pontos na imagem devido à movimentação do olho	62
Figura 41: Diagrama simplificado para corrigir a imagem pelo método proposto, no qual a ramificação para calcular os parâmetros mostra a atuação de duas <i>threads</i> para o processamento paralelo da correção do canal vermelho e azul em relação ao verde	63
Figura 42: Diagrama simplificado para calcular os parâmetros de registro entre dois canais, retornando a escala, rotação e <i>offset</i> horizontal e vertical entre eles.....	64
Figura 43: Diagrama simplificado para calcular a correlação de fase entre dois quadros	65
Figura 44: Imagem original com desalinhamento entre os canais antes da aplicação da correção proposta.....	67
Figura 45: Canal G (superior esquerda) e R (superior direita) antes da correção e o produto espectral cruzado entre eles para a correção de escala e rotação (inferior esquerdo) e translação (inferior direito).	68
Figura 46: Canal G (superior esquerda) e B (superior direita) antes da correção e o produto espectral cruzado entre eles para a correção de escala e rotação (inferior esquerdo) e translação (inferior direito).	69
Figura 47: Imagem corrigida com canais alinhados (esquerda) e comparação da região do nervo óptico antes (superior direita) e depois do alinhamento (inferior direita)	70
Figura 48: Imagem do nervo óptico obtida com arquitetura proposta (esquerda) e com retinógrafo Canon CX-1 de 16 Megapixels (direita) e o espectro de Fourier de ambas as imagens (inferior).....	71
Figura 49: Imagem obtida com a arquitetura proposta mostrando boa definição dos vasos da retina, microvasos do nervo óptico e fibras nervosas.....	72
Figura 50: Olho mecânico utilizado para a avaliação do método proposto	73
Figura 51: Imagem sem a máscara na qual aparece a borda da lente captadora (esquerda) e com máscara inserida via software e quadrado usado na comparação (direita).....	74
Figura 52: Conjunto para avaliação 1 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita).....	75
Figura 53: Conjunto para avaliação 2 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita).....	75
Figura 54: Conjunto para avaliação 3 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	75

Figura 55: Conjunto para avaliação 4 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	76
Figura 56: Conjunto para avaliação 5 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	76
Figura 57: Conjunto para avaliação 6 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	76
Figura 58: Conjunto para avaliação 7 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	77
Figura 59: Conjunto para avaliação 8 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita)	77
Figura 60: Conjunto para avaliação 9 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os <i>flashes</i> (centro) e imagem corrigida pelo método proposto (direita).....	77
Figura 61: Diagrama do sistema de medida da média da similaridade estrutural (MSSIM) , no qual exibe de forma básica a análise de luminância, contraste e estrutura entre os dois sinais.....	79
Figura 62: Gráfico comparativo do NMSE entre a correção apenas da translação e a correção completa	81
Figura 63: Gráfico comparativo do MSSIM entre a correção apenas da translação e a correção completa	81

LISTA DE TABELAS

Tabela 1: Cálculo do NMSE e MSSIM para cada conjunto após realizar a correção de translação e a correção completa (translação, rotação e escala)..... 80

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

DMRI – Degeneração Macular Relacionada à Idade

AIDS – *Acquired immune deficiency syndrome* – Síndrome da Imunodeficiência Adquirida

LED – *Light Emission Diode* – Diodo Emissor de Luz

CCD – *Charge Coupling Device* – Dispositivo de Carga Acoplada

RGB – *Red, Green, Blue* – Vermelho, Verde, Azul

CMOS – *Complementary Metal Oxide-Semiconductor* – Semicondutor Metal-Óxido Complementar

CD – *Compact Disk* – Disco Compacto

DVD – *Digital Versatile Disk* – Disco Digital Versátil

DICOM – *Digital Image Communications in Medicine* – Comunicações de Imagem Digital em Medicina

RF – *Red Free* – Retinografia sem a componente verde

FA – *Fluorescein Angiography* – Angiografia com Fluoresceína

ICG – *Indocyanine Green* – Angiografia com Indocianina Verde

pixel – *picture element* – elemento de imagem

fps – *frames per second* – quadros por segundo

DFT – *Discrete Fourier Transform* – Transformada Discreta de Fourier

IDFT – *Inverse Discrete Fourier Transform* – Transformada Discreta de Fourier Inversa

MTF – *Modulation Transfer Function* – Função de Transferência de Modulação

λ – comprimento de onda da luz

SUMÁRIO

1. Introdução.....	15
1.1. Justificativa.....	16
1.2. Objetivos	17
1.3. Organização da dissertação	18
2. Fundamentação Teórica	19
2.1. Anatomia do olho humano	19
2.2. Retinógrafo	21
2.2.1. Princípios de funcionamento	22
2.2.2. Tipos de exames.....	24
2.3. O estado da arte na retinografia.....	28
2.4. Obtenção de imagens coloridas.....	31
2.5. Métodos alternativos de obtenção de imagens coloridas.....	35
2.6. Retinógrafo colorido com uso de sensor monocromático	37
2.7. Registro de imagens.....	44
2.7.1. Passos para registro de imagens.....	46
2.7.2. Principais métodos para registro de imagens.....	48
2.7.2.1. Método para registro usando correlação cruzada.....	48
2.7.2.1. Método para registro usando correlação de fase.....	52
3. Metodologia Proposta	59
4. Resultados e Discussões.....	67
4.1. Exemplo da aplicação do método.....	67
4.2. Avaliação do método proposto.....	73
5. Conclusões	85
Referências.....	89

1.Introdução

A análise da imagem da retina por médicos oftalmologistas permite a detecção de várias doenças de fundo de olho, tais como glaucoma, retinose pigmentar, degeneração macular relacionada à idade (DMRI) e retinopatia diabética. Além das doenças relacionadas ao olho, por ser a única região do corpo humano na qual a condição vascular do paciente pode ser observada ao vivo e de maneira não intrusiva, outras doenças também podem ser reveladas, tais como diabetes, distúrbios na tireoide, câncer, AIDS, leucemia, toxoplasmose e inflamações reumáticas (OLIVEIRA, TREVELIN, *et al.*, 2009).

Em países sul-americanos com médio nível de desenvolvimento, como Brasil, Argentina e Chile, as doenças de retina como a retinopatia diabética e o glaucoma são as principais responsáveis pela perda de visão. Nesses países, doenças como catarata, principal causa de cegueira em países pobres na África, podem ser tratadas de forma rápida e eficiente. Isso ocorre visto que as prioridades dos programas de prevenção e combate à cegueira nos países em desenvolvimento se concentravam nas doenças relacionadas à córnea e ao cristalino, tais como catarata e cicatrizes na córnea, enquanto que os investimentos em tratamento de doenças da retina não tinham prioridade por serem consideradas causas incomuns de cegueira. Além disso, criou-se o falso paradigma de que o investimento necessário para o tratamento de patologias na retina não se justificava pelos resultados clínicos obtidos, devido aos altos custos relacionados à compra de equipamentos para o tratamento do fundo do olho, tais como retinógrafos, campímetros, lasers e topógrafos, geralmente de difícil acesso em países em desenvolvimento, além do tempo e recursos investidos em treinamento (MODUGNO, 2008).

No entanto, conforme os países vão se desenvolvendo e possuem mais tecnologia e recursos para investimento em saúde, o foco para o tratamento das doenças que causam a perda de visão passa a ser o das patologias do segmento posterior do olho (correspondente ao fundo do olho), já que para as doenças relacionadas ao segmento anterior (frente do olho) já existem técnicas mais conhecidas e consolidadas para o tratamento, como cirurgia de catarata e corretiva de miopia, astigmatismo e hipermetropia. (MODUGNO, 2008).

Além disso, a maior causa da cegueira entre pessoas com mais de 50 anos é a degeneração macular relacionada à idade (DMRI) (NEHEMY, 2006). Essa doença atinge em

sua maioria pessoas brancas nas quais, com o passar do tempo, surgem manchas na pele devido à oxidação. As mesmas manchas se formam na retina, enfraquecendo as camadas do fundo do olho e, como consequência, provocam fissuras da camada que impermeabiliza a retina, gerando novos vasos. Quando esses vasos crescem, forma-se um tecido fibrótico cicatricial, que resulta na perda irreversível da visão (BRANDÃO, 2005).

Com o aumento da expectativa de vida da população, cada vez mais pessoas são afetadas pela DMRI. De acordo com Brandão (2005), calcula-se que a cada ano surjam 700 mil casos da doença, sendo 80 mil casos no Brasil.

Dessa forma, um maior investimento na detecção e tratamento das doenças de retina tem-se tornado necessário ao longo dos anos, uma vez que se caracterizam como uma das principais causas de cegueira. Assim, o desenvolvimento de equipamentos para a detecção e tratamento das doenças do fundo do olho é de suma importância, dado que sem eles não se pode prevenir e tratar possíveis patologias, que são, normalmente, descobertas apenas quando parte da visão do paciente já foi perdida e não pode mais ser recuperada.

1.1. Justificativa

O trabalho apresentado nessa dissertação de mestrado está relacionado ao trabalho de Vieira (2013), cujo objetivo foi a construção de um retinógrafo digital com iluminação a LED usando câmera CCD monocromática de alta resolução.

O uso dos LEDs coloridos caracteriza uma inovação do equipamento. Tal fato permite simplificar bastante a eletrônica, além de que os LEDs possuem vida útil mais longa e não dissipam tanto calor quanto as lâmpadas, normalmente utilizadas em projetos de retinógrafos, simplificando também, dessa forma, o sistema de dissipação térmica. Já o uso da câmera CCD monocromática permite o aumento da resolução espacial do sistema, possibilitando gerar imagens com um nível maior de detalhamento, o que permite aumentar a precisão no diagnóstico médico. Nas câmeras CCD coloridas, normalmente estão presentes os filtros de Bayer, técnica na qual é necessário um processo de interpolação matemática para a obtenção da imagem, o que causa uma diminuição na resolução espacial (BAYER, 1976), conforme será explicado em mais detalhes no Capítulo 2. Outro fator a se considerar é que, com o uso de sensores monocromáticos e a eliminação dos filtros de Bayer do

sistema de imagem, pode-se diminuir a quantidade de luz que deve ser utilizada para iluminar a retina do paciente, já que se elimina a atenuação causada pelo filtro. Tal fato permite a diminuição do incômodo ao paciente durante o exame, pois uma menor quantidade de luz chega a seu olho.

No entanto, para exames coloridos, a arquitetura com um sensor monocromático e LEDs cromáticos RGB apresenta alguns problemas como configuração da potência de cada um dos três LEDs (calibração de cores) e desalinhamento entre os canais durante a geração da imagem colorida. O problema da calibração de cores pode ser resolvido utilizando um cartão de cores padrão, na qual se busca aproximar as cores alterando a potência de cada um dos LEDs RGB, até que as cores da imagem obtida estejam próximas das cores especificadas no cartão de cores (VIEIRA, 2013). Já o problema do desalinhamento entre os canais ocorre uma vez que, para produzir uma imagem colorida com a arquitetura proposta, disparam-se três *flashes* no olho do paciente (um vermelho, um verde e um azul), com um pequeno intervalo entre eles. Nesse intervalo, o olho pode movimentar-se, voluntária ou involuntariamente, o que pode causar um desalinhamento entre os canais, gerando um efeito de sombra na imagem resultante, atrapalhando ou até mesmo inviabilizando o diagnóstico médico.

Dessa forma, para garantir o uso clínico desse retinógrafo, existiu a necessidade de se desenvolver um método para corrigir o desalinhamento entre os três canais, alinhando-os (registrando) adequadamente antes da geração da imagem colorida. Além de registrar corretamente, o método deve ser rápido o bastante para que imagens de alta resolução possam ser alinhadas de forma transparente ao operador durante o exame. Dessa forma, o médico não precisa aguardar por um longo tempo desde o momento do disparo dos *flashes* até a imagem capturada ser exibida no monitor.

1.2. Objetivos

Este trabalho tem como objetivo desenvolver um método de registro de imagem para alinhamento dos canais monocromáticos RGB da imagem da retina para corrigir o desalinhamento causado pela movimentação do olho entre os *flashes* coloridos e com um tempo baixo de processamento.

As principais contribuições deste trabalho são:

- Viabilização do uso de LEDs e CCDs monocromáticos em retinógrafos digitais para a obtenção de imagens coloridas e monocromáticas da retina;
- Aumento na resolução espacial das imagens sem aumento de custo, devido à eliminação do filtro de Bayer, normalmente usado em sensores coloridos;
- Possível redução do custo final do equipamento para retinografia, uma vez que utiliza um único sensor monocromático, enquanto que várias arquiteturas utilizam dois sensores, um colorido e outro monocromático, sendo esse último usado para a obtenção de imagens de exames no qual há a administração de contraste, conforme será explicado posteriormente;
- Diminuição do incômodo ao paciente durante o exame, devido à diminuição da potência luminosa a ser aplicada no olho com o uso de câmera CCD sem filtro de Bayer.

1.3. Organização da dissertação

O presente trabalho está dividido em cinco capítulos, sendo este o primeiro, no qual se apresentou uma noção geral do contexto em que se insere retinografia e a importância do presente trabalho.

O segundo capítulo aborda a fundamentação teórica do projeto, como o funcionamento dos retinógrafos digitais, que são os principais equipamentos utilizados para a observação da retina, as técnicas de captura de imagens coloridas (convencionais e alternativas), bem como técnicas para o registro de imagens nas quais este trabalho foi fundamentado.

No terceiro capítulo, a metodologia proposta para o problema do registro das imagens da retina obtidas com diferentes fontes de iluminação é apresentada.

No quarto, os resultados e discussões são apresentados, bem como a validação do método obtida com um olho mecânico.

Por fim, no quinto capítulo é apresentada a conclusão do trabalho e também algumas sugestões de trabalhos futuros.

Vale ressaltar que os resultados aqui obtidos foram desenvolvidos com finalidade comercial e que a Opto Eletrônica é detentora do projeto e direitos sobre o produto.

2. Fundamentação Teórica

A seguir serão discutidos aspectos que fundamentam o presente trabalho, tais como: olho humano, retinografia, influência de sensores coloridos na qualidade da imagem, o uso de sensores monocromáticos para obtenção de imagens coloridas, problemas do uso de sensores monocromáticos e suas possíveis soluções.

2.1. Anatomia do olho humano

O olho é um sistema óptico complexo que funciona como receptor de luz, convertendo sinais luminosos em pulsos elétricos que são levados ao cérebro para processamento. Descrevendo de maneira bastante simplificada a formação de uma imagem no olho, pode-se considerar que a luz refletida do ambiente incide na córnea onde ocorre a refração; após essa refração, a luz passa pela íris que controla a abertura da pupila e o quanto de luz passará para a parte interna do olho; em seguida, a luz incide no cristalino, que focaliza os raios incidentes sobre a retina, formando uma imagem invertida, conforme ilustra a Figura 1 (CARVALHO, 2006).

Figura 1: Esquema simplificado de formação de imagem pelo olho

Fonte: Adaptado de Carvalho (2006)

Quando o olho forma a imagem exatamente sobre a retina, este é chamado de emétrepe. No caso de possuir alguma deformação ou problema de focalização, a imagem pode ser formada fora do plano da retina. Nesse caso, o olho é considerado míope (imagem formada antes da retina) ou hipermetrópe (imagem formada depois da retina) (CARVALHO, 2006).

O olho é formado por três tecidos denominados túnica oculars, mostrados na Figura 2. A mais externa é composta pela córnea e pela esclera, responsáveis, principalmente, pela proteção das demais estruturas. A túnica intermediária é composta pela íris, pelo corpo ciliar e pela coróide. Essa parte é bastante vascularizada e também é responsável pelo ajuste de foco (cristalino) e controle da entrada de luz (íris). Por fim, a túnica interna é composta pela retina, que se caracteriza como um tecido complexo, no qual estão localizadas células fotorreceptoras, que são conectadas ao cérebro através do nervo óptico (SMITH, 2000).

Figura 2: Representação do olho humano

Fonte: Adaptado de Vieira (2013)

A retina contém vasos sanguíneos, fibras nervosas e células conhecidas como cones e bastonetes, sensíveis à luz. O nervo óptico e seu ponto cego associado estão localizados onde as fibras nervosas saem do globo ocular e se dirigem ao cérebro. Ligeiramente deslocado do eixo do globo ocular está a mácula. Na mácula, há uma maior concentração de cones, sendo que na região central, chamada fóvea, encontram-se apenas cones. A partir da fóvea, os bastonetes começam a aparecer até que apenas eles estejam presentes, nos pontos mais distantes da mácula (SMITH, 2000).

Na retina, existem cerca de 7 milhões de cones, 125 milhões de bastonetes e apenas 1 milhão de fibras nervosas. Os cones são responsáveis pela percepção de cores e detalhamento da imagem. Já os bastonetes são responsáveis pela percepção de luz, relacionando-se com a visão periférica de baixa resolução, porém de grande sensibilidade (CARVALHO, 2006).

Ressalta-se ainda que são formadas duas imagens, uma em cada olho, que são processadas pelo cérebro, proporcionando a noção de tridimensionalidade dos objetos. Esse processamento é conhecido como visão estereoscópica (CARVALHO, 2006).

Conforme se pode observar, a estrutura do olho humano é extremamente complexa e com alto grau de organização. Assim, a obtenção de imagens do fundo do olho é um procedimento bastante complicado, uma vez que, para conseguir capturar as imagens, é necessário iluminar a retina, cuja principal função é absorver luz, e captar o sinal refletido e espalhado da mesma. Além disso, há inúmeras outras barreiras para a luz, como a córnea, o cristalino, a íris e o humor vítreo, até atingir a retina, conforme se pode visualizar na Figura 2. Estudos mostram que a energia que pode ser captada da retina está na ordem de 1 a 12% da energia incidente, variando de acordo com o comprimento de onda da luz aplicada (DELORI e PFLIBSEN, 1989).

2.2. Retinógrafo

Os retinógrafos, também conhecidos como câmeras de retina, realizam a captura de imagens da retina e de suas estruturas como o nervo óptico, a mácula e a fóvea. Ao observar essas estruturas, diversas patologias podem ser identificadas. A Figura 3 mostra um retinógrafo produzido pela Opto Eletrônica S.A.

A seguir serão descritos os princípios básicos de funcionamento, os tipos de exame e as inovações da área.

Figura 3: Exemplo de um retinógrafo digital

Fonte: Opto Eletrônica S. A.

2.2.1. Princípios de funcionamento

O retinógrafo é um equipamento constituído de um sistema óptico que permite que a imagem da retina seja magnificada e, com o uso de lentes, projetada em um anteparo. No mesmo caminho óptico são acoplados filtros espectrais e um sistema de captura de imagens que pode ser caracterizado por uma máquina de fotografia analógica ou ainda por um sistema de captura digital (MODUGNO, 2008).

Dessa forma, o equipamento se caracteriza como a integração de vários subsistemas: eletrônico, óptico, mecânico e computacional, sendo este último presente apenas em retinógrafos digitais. Na Figura 4 é exibido o esquemático do funcionamento básico de um retinógrafo digital: o computador realiza o controle do equipamento e dispara comandos para a eletrônica de controle; esta aciona os LEDs de alta potência que emitem a luz que é conduzida pelo sistema de iluminação até o olho do paciente; a luz refletida do olho do paciente retorna pelo sistema de captação e, de acordo com o ajuste do foco, a imagem é formada normalmente no detector CCD de uma câmera colorida, que por sua vez digitaliza a imagem e a encaminha ao computador para armazenamento.

Figura 4: Esquemático de funcionamento de um retinógrafo digital com iluminação a LED

O subsistema eletrônico possui várias funções de controle do equipamento, tais como controle das fontes de iluminação, controle do sensor para captura de imagens e posicionamento dos filtros. A captura das imagens pode ser realizada tanto na forma analógica quanto na digital. Na analógica, uma câmera fotográfica equipada com filme colorido ou preto e branco é utilizada. Após a realização dos exames, o filme precisa ser revelado para possibilitar o diagnóstico. Já nos sistemas digitais, a captura é realizada com sensores fotossensíveis, normalmente do tipo CMOS ou CCD. Os sinais captados por esses sensores podem ser transferidos para um microcomputador e visualizados pelo médico imediatamente após a aquisição da imagem, além de permitir o armazenamento em mídias digitais (CDs, DVDs ou cartões de memória) (MODUGNO, 2008).

O subsistema óptico pode ser dividido em dois módulos: o de captação e iluminação. O módulo de captação tem por objetivo projetar a imagem da retina no sensor ou filme com a maior resolução possível. Já o módulo de iluminação é responsável por projetar homogeneamente a energia de uma fonte de luz na retina do paciente (MODUGNO, 2008). Essa iluminação é projetada na retina usando um anel de iluminação (NOHDA, 1982). O subsistema óptico deve ser bastante eficiente, evitando que os pacientes sejam submetidos a um alto nível de iluminação, uma vez que é necessária uma fonte luminosa de alta intensidade para conseguir boas imagens da retina.

O subsistema mecânico deve suportar o subsistema eletrônico e óptico, além de permitir o posicionamento preciso do subsistema óptico (captura e iluminação) no olho do paciente. Normalmente, os sistemas de retinografia possuem um *joystick* que permite ao operador movimentar o equipamento até a posição correta para obter as imagens da retina.

Por fim, o subsistema computacional, presente apenas em retinógrafos digitais, é responsável pelo armazenamento das imagens do retinógrafo, utilizando para isso um sistema de banco de dados, organizado por pacientes, médicos e clínicas. Além disso, ele permite o controle do equipamento por um painel no software, além de exibir as imagens em tempo real na tela e permitir sua edição (brilho, contraste, gama, etc). Alguns sistemas mais modernos permitem o acesso remoto ao banco de dados de imagem, normalmente realizado no padrão *DICOM* (*Digital Imaging and Communication in Medicine*) (ALLIANCE, MEDICAL IMAGING & TECHNOLOGY, 2012). Dessa forma, o médico consegue realizar o diagnóstico remotamente, agilizando o trâmite do exame e o tratamento do paciente.

2.2.2. Tipos de exames

Para realizar o diagnóstico de doenças do fundo do olho, dois tipos de exames podem ser realizados: angiografias e retinografias. Nas angiografias, existe a necessidade da aplicação de contraste na corrente sanguínea do paciente; no caso das retinografias, não é necessário o uso de drogas. Ambos podem ser feitos conjuntamente, na mesma visita do paciente, o que permite um diagnóstico complementar, uma vez que revelam informações diferentes da condição da retina do paciente (CARVALHO, 2006).

As retinografias são divididas em dois tipos: coloridas e anerítras. As retinografias coloridas objetivam fornecer uma imagem colorida do segmento posterior do olho, conforme exibido na Figura 5. Neste exame, alterações na coloração podem indicar danos no tecido da retina ou outra doença. Normalmente, os médicos comparam imagens atuais com as anteriores para fornecerem diagnósticos mais precisos. Para a obtenção da imagem, normalmente ilumina-se a retina com luz branca e é utilizado um sensor colorido para aquisição. Para evitar radiação excessiva no olho do paciente, devido à alta intensidade da iluminação, comumente incluem-se filtros que barram o infravermelho ($\lambda > 800 \text{ nm}$) e o ultravioleta ($\lambda < 400 \text{ nm}$) da fonte luminosa (VIEIRA, 2013).

Figura 5. Exemplo de uma imagem de fundo de olho obtida com a técnica de retinografia colorida
Fonte: Adaptado de Vieira (2013)

No caso das retinografias anerítras, conhecidas como RF (*red free*), o exame é realizado sem componentes da cor vermelha, sendo que o espectro de iluminação corresponde a uma parcela com $500 \text{ nm} < \lambda < 580 \text{ nm}$ (verde). Como o sangue é

vermelho e os vasos da retina possuem sangue, a consequência é que esses vasos aparecem com nível de cinza escuro na imagem, o que aumenta bastante o contraste em relação o fundo do olho (CARVALHO, 2006), conforme pode ser visualizado na Figura 6.

Figura 6: Exemplo de uma imagem de fundo de olho obtida com a técnica de retinografia anértra
Fonte: Adaptado de Vieira (2013)

Da mesma forma das retinografias, as angiografias também são divididas em dois tipos: fluoresceínica e com indocianina verde. A angiografia fluoresceínica, conhecida como FA (*fluorescein angiography*), é utilizada quando se pretende estudar a dinâmica da circulação sanguínea, possibilitando identificar problemas circulatórios nos vasos retinianos. Nesse exame, utiliza-se uma substância conhecida como fluoresceína sódica, droga que possui a propriedade de absorver energia eletromagnética e reemitir na forma de luz. Isso ocorre devido instabilidade eletrônica da substância quando iluminada, sendo que seu pico de excitação ocorre com $450 \text{ nm} < \lambda < 500 \text{ nm}$ e sua emissão com $510 \text{ nm} < \lambda < 550 \text{ nm}$ (CARVALHO, 2006). A Figura 7 apresenta um exemplo de imagem obtida no exame de angiografia fluoresceínica.

Para realizar um exame de FA, a fluoresceína é injetada na corrente sanguínea do paciente, sendo que em poucos instantes a droga atinge os vasos que irrigam a retina. O olho é então iluminado com luz no comprimento de onda capaz de excitar a droga, que por sua vez reemite a luz. Essa é filtrada no comprimento específico de reemissão (filtro passa faixa com $510 \text{ nm} < \lambda < 550 \text{ nm}$) e a imagem é capturada pelo sensor.

Figura 7: Exemplo de uma imagem de fundo de olho obtida com a técnica de angiografia fluoresceínica
 Fonte: Adaptado de Vieira (2013)

Uma característica interessante do exame FA é permitir ao médico capturar várias imagens sequenciais e estudar a dinâmica da circulação, conforme se pode observar na Figura 8: após injetar o contraste, pode-se visualizar, inicialmente, o momento em que a droga passa pelas artérias (precoce); após alguns segundos, visualiza-se o momento em que passa pelas veias (média); por fim, depois de alguns minutos, pode-se visualizar possíveis hemorragias, indicadas por manchas brancas na retina (tardia) (VIEIRA, 2013).

Figura 8: Dinâmica da circulação sanguínea do exame de angiografia fluoresceínica – precoce (esquerda), média (centro) e tardia (direita)
 Fonte: Adaptado de Vieira (2013)

No caso da angiografia com indocianina verde, conhecida como ICG (*indocyanine green*), a técnica é semelhante a que usa fluoresceína, porém a substância utilizada é a indocianina verde, que possui pico de absorção com $790 \text{ nm} < \lambda < 810 \text{ nm}$ e pico de emissão próximo a 835 nm . O objetivo desse exame é estudar os vasos mais profundos do olho, presentes na coróide (CARVALHO, 2006), estrutura localizada entre a retina e a camada

esclerótica. Isso é possível uma vez que os picos de absorção e emissão estão no infravermelho próximo, que é pouco absorvido nas camadas da retina anteriores à coroide (VIEIRA, 2013). A Figura 9 exibe um exemplo de imagem obtida na angiografia com indocianina verde.

Figura 9: Exemplo de uma imagem de fundo de olho obtida com a técnica de angiografia com ICG

Fonte: Adaptado de Vieira (2013)

Para a obtenção das imagens nos exames de angiografia, é necessário o uso de um sensor de alta sensibilidade uma vez que a potência da luz reemitida pela droga geralmente é baixa. Além disso, no caso do exame ICG, o sensor deve ter a capacidade de capturar no infravermelho próximo, já que o pico de emissão está em torno de 835 nm. Tal fato faz com que muitos fabricantes utilizem dois sensores na arquitetura do equipamento, sendo um colorido para a obtenção das imagens da retinografia colorida e outro monocromático, com maior sensibilidade e resposta espectral mais larga, para a obtenção das imagens monocromáticas e com contraste.

Além dos quatro exames básicos, há outra característica relevante a ser citada quanto à retinografia: a necessidade de dilatar ou não a pupila do paciente com colírio específico. Tal processo de dilatação da pupila é conhecido como midríase (CSILLAG, 2005). Dessa forma, equipamentos em que existe a necessidade da dilatação são conhecidos como midriáticos, enquanto aqueles na qual não existe essa necessidade são conhecidos como não-midriáticos.

Conforme as especificações dos equipamentos não-midriáticos mais modernos, a pupila do paciente deve estar com abertura mínima de 3,3 mm (TOPCON, 2012), razão pela

qual esses exames são realizados em salas escuras. No entanto, o exame mais efetivo é realizado na forma midriática, uma vez que permite que se acople mais luz no olho do paciente, conseguindo, conseqüentemente, imagens de melhor qualidade. Outra desvantagem do exame não-midriático é que, como os equipamentos utilizam iluminação de alta potência, a pupila do paciente demora alguns segundos ou até minutos para retornar ao diâmetro mínimo requerido depois do disparo do *flash* para a captura da imagem, prolongando o tempo do exame.

2.3. O estado da arte na retinografia

Pesquisas atuais em retinografia visam a permitir que o exame seja realizado de uma forma mais rápida, segura e confortável aos pacientes.

Grandes empresas do segmento, principalmente alemãs e japonesas, como Carl Zeiss e Canon, investem em pesquisa para a otimização dos equipamentos, sendo que as primeiras patentes em retinografias datam da década de 60 (VIEIRA, 2013). No Brasil, a Opto Eletrônica, única empresa brasileira a desenvolver retinógrafos, também desenvolve pesquisa com os mesmos fins citados acima.

De modo geral, as pesquisas objetivam melhorar os subsistemas que integrados constituem o retinógrafo digital. Para o subsistema mecânico, a tendência atual é deixá-lo cada vez mais reduzido e leve, permitindo ao operador um maior conforto e precisão ao usar o equipamento. No subsistema óptico, investe-se em lentes e filtros com melhor qualidade óptica para permitir que seja necessária menos radiação para obter imagens da retina do paciente, além de permitir que as imagens tenham maior qualidade e definição. Já no subsistema eletrônico, um dos avanços atuais é o uso de iluminação com LEDs, que possuem tempo de vida maior que as lâmpadas usadas nos equipamentos mais antigos, além do uso de sensores com resoluções cada vez maiores. Por fim, no subsistema de software, os maiores avanços são relacionados a ferramentas de auxílio ao diagnóstico, como, por exemplo, cálculo de área de lesões, mosaico de imagens (foto panorâmica da retina), conforme exibido na Figura 10, e uso de imagens estereoscópicas (3D), além da possibilidade de acesso remoto às imagens usando o padrão DICOM, o que agiliza o exame e diagnóstico.

Figura 10: Mosaico da retina formada pelo processamento de várias imagens obtidas em diferentes tempos

Fonte: <http://www.kowa.eu/medicals/en/nonmydaD.php> Acesso em: 04/05/2012

No âmbito acadêmico, novos estudos estão sendo realizados, principalmente relacionados com a interação da luz com as estruturas do fundo do olho. Nesse contexto, estudos de polarimetria, imagens multiespectrais e de oximetria são os que mais se destacam (VIEIRA, 2013).

Nos estudos de polarimetria, propõe que se polarize a luz no sistema de iluminação e a imagem seja capturada em vários ângulos de polarização. Tal técnica busca identificar retinopatias que modificam as estruturas do fundo do olho e que podem ser mais bem visualizadas em determinadas condições de polarização da luz (MIURA, 2009).

No caso das imagens multiespectrais, busca-se variar o comprimento de onda da fonte de iluminação para se obter diferentes informações do fundo do olho, conforme exibido na Figura 11. A busca, nesse caso, é pelo comprimento de onda que ressalte as principais patologias da retina (EVERDELL, STYLES, *et al.*, 2009).

Figura 11: Imagens da retina iluminadas com diferentes comprimentos de onda (λ)

Fonte: Adaptado de Everdell *et al* (2009)

Outro estudo atual relacionado à retinografia é a oximetria, o qual visa a representar a saturação de oxigênio na hemoglobina que circula pelos vasos retinianos. Nessa técnica, a retina é iluminada com os comprimentos de onda de 522 nm, 548 nm, 569 nm e 586 nm, na qual a hemoglobina com oxigênio possui uma resposta muito próxima à hemoglobina sem oxigênio. A resposta a essas diferentes iluminações é usada para calibração do sistema. Já para os comprimentos de onda de 558 nm e 577 nm, as respostas são diferentes no caso de possuir ou não oxigênio, sendo possível determinar a saturação do mesmo ao longo das veias e artérias da retina (RAMELLA-ROMAN, 2008).

Um equipamento para o estudo da oximetria da retina foi criado em 2008 pela empresa islandesa chamada *Oxymap*. Uma análise realizada pelo equipamento é exibida na Figura 12. No entanto, conforme o próprio *website* da empresa anuncia, atualmente o equipamento está sendo usado apenas para investigação e pesquisa em universidades dos Estados Unidos e Europa (OXYMAP, 2008).

Figura 12: Análise da concentração de oxigênio da retina utilizando equipamento *Oxymap*

Fonte: Adaptado de *Oxymap* (2008)

2.4. Obtenção de imagens coloridas

Grande parte dos retinógrafos digitais presentes atualmente no mercado utilizam sensores coloridos. Tais sensores permitem tanto a realização dos exames coloridos quanto em tons de cinza (RF, FA e ICG) (MODUGNO, 2008). A maioria desses sensores possui apenas sensibilidade à intensidade da luz, não conseguindo distinguir cores. Dessa forma, para se obter imagens coloridas, usa-se um conjunto de filtros espectrais que permitem que apenas uma parcela do espectro visível chegue a um determinado elemento sensível do sensor. Cada parcela espectral é conhecida por canal de cor (GONZALEZ e WOODS, 2009).

O conjunto de filtros espectrais mais utilizado em sensores coloridos é chamado de filtro de Bayer, sendo que o arranjo mais comum consiste de uma matriz 2 x 2 que utiliza 4 filtros para formar um pixel colorido, sendo um filtro para o canal vermelho, dois para o verde e um para o azul (BAYER, 1976), conforme arranjo mostrado na Figura 13.

Figura 13: Arranjo básico do filtro de Bayer (esquerda) e esquema de filtros espectrais utilizado para a formação da imagem na qual a luz incidente é dividida nas componentes RGB (direita)

Existem diversos outros arranjos, porém todos com a mesma característica de que, para cada *pixel*, é permitida a passagem apenas de uma parte do espectro luminoso incidente (VIEIRA, 2013). Assim, para a composição da imagem final, é necessário um processo de interpolação matemática conhecido como *demosaicing* (LI, GUNTURK e ZHANG, 2008). Isso ocorre uma vez que cada *pixel* no sistema de cores RGB deve ser representado por três cores e, na constituição física do sensor, para cada posição tem-se apenas a componente de uma das cores (vermelha, verde ou azul), conforme pode ser visualizado na Figura 14.

Figura 14: Formação da imagem colorida com filtro de Bayer na qual para cada posição da matriz final tem-se apenas a componente de uma das cores

Fonte: Adaptado de Li, Gunturk e Zhang (2008)

Existem diversos algoritmos que realizam o *demosaicing*, os quais objetivam preencher, por um processo matemático, os *pixels* que estão faltando em cada canal de cor. Alguns algoritmos são implementados no domínio da frequência, mas a maioria se concentra no processamento no domínio do espaço (GONZALEZ e WOODS, 2009). Além disso, existem técnicas que mesclam as informações dos três canais de cores para obter o valor do *pixel*

faltante, não utilizando apenas as informações do canal que está sendo interpolado (LI, GUNTURK e ZHANG, 2008).

Dessa forma, um sensor colorido com filtro de Bayer apresenta dois problemas intrínsecos à sua arquitetura: o primeiro está relacionado à perda de resolução devido à interpolação matemática necessária para formar a imagem colorida usando o método *demosaicing* e o segundo está relacionado ao fenômeno conhecido como *color aliasing* (GUTTOSCH, 2005).

A perda de resolução devido ao *demosaicing* pode ser observada ao se fotografar um alvo preto e branco em forma de degrau. Conhecendo o processo de interpolação usado pelo sensor, esse teste também pode ser simulado numericamente. Uma simulação para esse fenômeno pode ser encontrado no trabalho de Vieira (2013). Nesse caso, foi considerado um degrau do preto para o branco fotografado com uma câmera monocromática com óptica perfeita. Depois, considerou-se o mesmo degrau fotografado por uma câmera colorida, também com óptica perfeita. A designação de óptica perfeita foi utilizada apenas para simplificar o problema, sendo que nesse caso, os componentes ópticos não estariam afetando a função de transferência do sistema. Se a interpolação utilizada for a linear, conforme utilizado em Vieira (2013), tem-se a saída apresentada na Figura 15.

Figura 15: Imagem simulada obtida com câmera monocromática (direita) e colorida (esquerda)

Fonte: Adaptado de Vieira (2013)

Conforme se pode observar, na imagem com a câmera colorida, a borda sofreu uma deterioração. Tal fato ocorre, uma vez que os processos de interpolação provocam uma resposta espúria, reduzindo a resolução espacial da imagem, ou seja, piora a taxa de contraste e dificulta a identificação de pequenas estruturas na retina (VIEIRA, 2013).

Ainda, conforme o trabalho de Vieira (2013), um experimento realizado para medida do limite de resolução espacial de um retinógrafo digital com câmera monocromática e com câmera colorida (filtro de Bayer) utilizando o mesmo sensor (tamanho de *pixel* de 9 μm e ampliação óptica de 1,85) mostrou que a menor estrutura possível de ser observada para o modelo colorido foi de 54 μm , e para o modelo

monocromático foi de 30 μm , o que comprova maior resolução espacial dos sistemas monocromáticos.

O segundo problema dos CCDs coloridos com filtro Bayer é o problema conhecido como *color aliasing*. Tal problema causa uma distorção de cores, principalmente em imagens com bordas inclinadas ou com altas frequências (GUTTOSCH, 2005). Um exemplo do problema é exibido na Figura 16, na qual foi feita a simulação de uma borda inclinada sendo capturada com CCD colorido com filtro de Bayer. Esse fenômeno ocorre uma vez que imagem capturada inicialmente possui, para cada posição, a informação de apenas uma componente de cor, conforme se pode visualizar nos canais separados da Figura 16, parte superior direita. Esses canais passam então por um processo de interpolação matemática, formando os canais interpolados, exibidos na parte inferior direita. Nesse momento, para cada posição da imagem, têm-se as três componentes de cor, uma vez que a informação faltante no passo anterior foi gerada pelo processo de interpolação. Ao mesclar os três canais para formar a imagem final no CCD, essa informação adicional criada gera um serrilhamento colorido na borda da imagem resultante, o que caracteriza o problema.

Figura 16: Exemplo de *color aliasing*

Fonte: Adaptado de Guttosch (2005)

No contexto de retinografia, tem-se outro fato bastante interessante em desfavor do uso de sensores coloridos. Dos principais exames realizados pelo médico, apenas um utiliza a característica de cor da imagem. Os outros três – RF, FA e ICG – são exames em tons

de cinza. Dessa forma, o uso do sensor colorido, além de poder aumentar o custo final do equipamento, piora a resolução das imagens devido aos problemas relatados anteriormente e uma abordagem usando sensores monocromáticos para o desenvolvimento de retinógrafos digitais aparece como uma alternativa viável para combater tais problemas.

2.5. Métodos alternativos de obtenção de imagens coloridas

Existem várias formas de se obter uma imagem colorida sem o uso do filtro de Bayer. As primeiras tentativas que se têm notícia na literatura foram obtidas a partir de uma roda de filtros, na qual a partir da obtenção de três imagens monocromáticas filtradas (três canais de cor), formava-se a imagem colorida (OCEAN THIN, 2012). A Figura 17 exibe um exemplo de câmera com roda de filtros.

Figura 17: Câmera multiespectral com roda de filtros que permite a obtenção de diferentes λ

Fonte: Adaptado de Spectrocam (2012)

Esse sistema apresenta algumas vantagens como excelente reprodução de cor, facilidade na troca dos filtros e necessidade de apenas um sensor para a obtenção de imagens coloridas. No entanto, alguns pontos são desfavoráveis, como a demora na

aquisição das imagens, devido aos elementos mecânicos móveis e ao tempo para o posicionamento dos filtros, além dos quadros serem capturados de forma não simultânea.

Uma segunda solução alternativa para a obtenção de imagens coloridas pode ser conseguida com três sensores, um para cada canal de cor, conforme pode ser visualizado na Figura 18. Nesse caso, um conjunto óptico realiza a separação das cores. Esse tipo de arquitetura é encontrado em algumas câmeras de vídeo profissionais (SONY CORPORATION, 2012). A grande vantagem dessa arquitetura em relação à anterior é que os três canais de cor são capturados simultaneamente e não há perda de resolução espacial, como no caso do filtro de Bayer. A desvantagem, nesse caso, está relacionada com o custo. Normalmente, sensores de alta sensibilidade e resolução são bastante caros e o uso de três sensores encarece consideravelmente o equipamento final.

Figura 18: Imageamento colorido com o uso de três sensores, na qual filtros ópticos separam as componentes RGB e espelhos as encaminham para o respectivo sensor

Fonte: Adaptado de Vieira (2013)

No contexto de retinografia, como citado anteriormente, apenas um dentre os quatro principais exames de retina é colorido. Mesmo assim, a maioria dos fabricantes utilizam sensores coloridos que, além de poder encarecer o equipamento, faz com que as imagens percam resolução espacial devido aos processamentos envolvidos com a utilização do filtro de Bayer.

No entanto, os retinógrafos apresentam uma particularidade que pode ser explorada para a construção de um sistema com sensor monocromático que permite a obtenção de imagens coloridas: o fato da iluminação da cena ser controlável. Nesse caso, a iluminação aplicada na retina é controlada e pode-se, ao invés de iluminar com luz branca,

como nos sistemas tradicionais com sensores coloridos, atuar no sistema de iluminação. Assim, ilumina-se com luz vermelha, verde e azul com um pequeno espaço de tempo entre os *flashes*, e se obtêm três imagens sequenciais, conforme a proposta de Vieira (2013). Com as imagens de cada um dos três canais (RGB), pode-se então construir a imagem colorida.

2.6. Retinógrafo colorido com uso de sensor monocromático

O equipamento utilizando a nova arquitetura será descrito sucintamente a seguir, baseado no trabalho de Vieira (2013), o qual foi realizado em conjunto com o trabalho apresentado nessa dissertação.

Nessa proposta, para a obtenção da imagem colorida, não se atua mais no elemento sensor, como os sistemas tradicionais com filtro de Bayer, mas sim no sistema de iluminação. Assim, a cada conjunto de três imagens, consegue-se toda a informação necessária para a obtenção de uma imagem colorida, sem a perda de resolução espacial que ocorre com o uso de sensores com filtro de Bayer. Um exemplo da formação da imagem usando iluminação cromática é exibido na Figura 19.

Figura 19: Exemplo da iluminação da retina para a formação de imagem colorida com a arquitetura proposta

Além dessa vantagem, como não possui filtro sobre os elementos sensíveis, a quantidade de energia que chega ao sensor é maior, dado que, por melhor que seja o filtro espectral, ele nunca apresenta uma taxa de transmitância de 100% por toda a banda de interesse (VIEIRA, 2013). Dessa forma, consegue-se obter imagens de boa qualidade com menor energia aplicada ao olho do paciente, diminuindo o incômodo do exame.

A proposta apresentada é bastante conveniente, uma vez que atende às necessidades dos exames em tons de cinza (RF, FA e ICG) e, ainda assim, permite a obtenção de imagens coloridas como resolução espacial maior, utilizando apenas um único sensor monocromático.

A Figura 20 apresenta com mais detalhes a arquitetura proposta. O equipamento possui três LEDs coloridos, representados na parte superior da figura. Em diferentes e pequenos intervalos de tempo (t_1 , t_2 e t_3) os *flashes* são disparados na retina do paciente. As imagens são então capturadas por um sensor monocromático e os três canais são combinados para a formação da imagem colorida.

Figura 20: Esquema básico de captura de imagem colorida usando CCD monocromático conforme arquitetura proposta

Fonte: Adaptado de Stuchi *et al.* (2012)

Tais características inovadoras no sistema de iluminação e uso de sensores motivaram a empresa brasileira Opto Eletrônica a financiar o projeto de um equipamento de

retinografia completo, objetivando a obtenção de imagens de boa qualidade e resolução e um menor incômodo ao paciente.

Para a captura das imagens, o equipamento possui CCD monocromático Kodak com 4008×2672 *pixels* (aproximadamente 11 Megapixels), cada um com $9 \mu\text{m} \times 9 \mu\text{m}$ de dimensão, o qual é exibido na Figura 21. Além disso, o sensor possui resposta monocromática para o espectro de $400 \text{ nm} < \lambda < 850 \text{ nm}$, alta sensibilidade, taxa de captura de 7 fps em *full resolution* (utilizado para a obtenção das fotografias da retina) e 13 fps para o modo *binning* ativo (imagem de menor resolução utilizada na previsão para o posicionamento do paciente).

Figura 21: CCD monocromático de 11 Megapixels utilizado na arquitetura proposta

Fonte: <http://www.svs-vistek.com> Acesso em: 04/04/2013

No caso do sistema de iluminação, é utilizado sistema a LED sincronizado com o sistema de captura do equipamento. Para a iluminação da retina e obtenção das imagens, utilizam-se três LEDs coloridos de alta potência, tendo os picos de emissão com comprimentos de onda (λ) próximos de 626 nm (vermelho), 528 nm (verde) e 463 nm (azul), exibidos na Figura 22. Já para o caso da previsão e posicionamento do paciente, utiliza-se um LED infravermelho próximo, com pico de emissão em torno de $\lambda = 780 \text{ nm}$ (VIEIRA, 2013). Tal fato faz com que o paciente tenha o incômodo minimizado durante o posicionamento para a obtenção da imagem, dado que enxerga apenas uma pequena parcela da emissão do LED infravermelho. A câmera, por outro lado, possui sensibilidade até aproximadamente $\lambda = 850 \text{ nm}$, o que permite ao operador visualizar as estruturas da retina para o posicionamento e melhor foco.

Figura 22: LEDs utilizados para a obtenção das imagens coloridas na arquitetura proposta

Fonte: Adaptado de Vieira (2013)

A partir dessa inovação arquitetural, construiu-se o subsistema mecânico, óptico, eletrônico e computacional que permite a captura de imagens da retina conforme a proposta de iluminação cromática e sensor monocromático apresentada. Uma imagem do protótipo montado é exibido na Figura 23.

Figura 23: Protótipo montado a partir da arquitetura proposta

Fonte: Adaptado de Vieira (2013)

São vários os benefícios da nova arquitetura. Dentre eles, pode-se citar:

1. possível redução do custo devido ao uso do sensor um único monocromático;
2. imagens com maior resolução espacial, principalmente nos exames monocromáticos, devido ao não uso do filtro de Bayer;
3. mais conforto ao paciente durante o exame, uma vez que se pode reduzir a intensidade da iluminação, devido ao não uso do filtro de Bayer;

4. possibilidade da diminuição do tempo do exame, uma vez que uma imagem no exame colorido já disponibiliza, simultaneamente, a imagem do exame RF.

O benefício descrito no quarto item, no qual o exame colorido pode disponibilizar uma imagem do exame RF ocorre, uma vez que, para obter uma imagem colorida, ilumina-se a retina com luz vermelha, verde e azul, obtendo três quadros distintos. No entanto, no exame RF, ilumina-se a retina com o mesmo LED verde da imagem colorida, sem a adição de filtros adicionais. Assim, para se obter o exame RF de um exame colorido, basta obter o canal verde da imagem colorida, conforme pode ser visualizado na Figura 24. Tal fato agiliza o exame, uma vez que o operador do equipamento, com apenas um posicionamento do paciente e disparo da fotografia colorida, pode obter imagens referente a dois exames.

Figura 24: Exame colorido e RF obtidos com apenas um disparo na arquitetura proposta
Fonte: Opto Eletrônica S.A.

Apesar das grandes vantagens apresentadas para a nova arquitetura, alguns problemas foram verificados. Inicialmente, após os ajustes dos ganhos da câmera, verificou-se o seguinte: as imagens monocromáticas apresentavam excelente qualidade, conforme exibido na Figura 25, porém as coloridas apresentavam variações nas cores, dado que os LEDs possuem diferentes intensidades luminosas, o que influencia diretamente na imagem colorida final.

Figura 25: Exemplo de imagem monocromática de ótima qualidade obtida com a arquitetura proposta (exame FA)

Fonte: Opto Eletrônica S.A.

Após a análise das imagens, verificou-se a necessidade de um sistema de calibração de cores, para que a potência aplicada nos LEDs pudesse reproduzir fielmente as cores da retina. Vieira (2013) descreve a calibração dos LEDs considerando o sistema como uma caixa preta, na qual o mesmo fotografa iterativamente um alvo colorido (conhecido como *Color Checker* – exibido na Figura 26), até que a cor obtida na imagem esteja muito próximo à presente no cartão de cores (VIEIRA, 2013).

Figura 26: *Color Checker* utilizado para a calibração de potência dos LEDs

Fonte: <http://xritephoto.com> Acesso em: 04/04/2013

Para esse ajuste de cor, a corrente elétrica que vai para os LEDs é ajustada através de parâmetros de calibração presentes no *firmware* da placa de controle. Encontrado o valor para o qual o sistema realiza a aquisição da imagem com a cor correta, tais valores são

gravados na memória do microcontrolador presente na placa de controle de corrente dos LEDs. A Figura 27 mostra um exemplo da imagem colorida obtida com o sistema de iluminação antes (esquerda) e após a calibração (direita).

Figura 27: Imagem colorida obtida antes (esquerda) e depois da calibração de cores (direita)

Fonte: Opto Eletrônica S.A.

Corrigido o problema do ajuste das cores, verificou-se então um segundo problema: durante os três *flashes* o olho pode piscar e a imagem sair com as cores distorcidas, uma vez que a informação presente em um dos quadros, por exemplo, não condiz com a informação presente nos outros dois. Nesse caso, a instrução para os operadores do equipamento foi para que a imagem fosse simplesmente excluída, uma vez que, mesmo em exames com sensores coloridos, o paciente pode piscar durante a aquisição da imagem e a imagem tem que ser excluída.

Por fim, notou-se um problema bastante grave relacionado à arquitetura proposta: o olho pode movimentar-se durante os três *flashes*. Como o sensor possui uma taxa de 7 fps, cada quadro é capturado em aproximadamente 143 ms. Nesse caso, para se capturar os três quadros, no mínimo se gasta 429 ms. Conforme em Condez, Macknik e Hubel (2004), o olho humano apresenta um comportamento conhecido como microssacadas, que são pequenos e rápidos movimentos durante uma fixação voluntária com duração aproximada a 25 ms. Assim, em 429 ms o olho pode movimentar, voluntaria ou involuntariamente, o que causa um desalinhamento nos canais e uma sensação de sombra na imagem final, conforme pode ser visto na Figura 28.

Figura 28: Área do nervo óptico com desalinhamento dos canais devido à movimentação do olho durante o processo de aquisição das imagens dos três canais RGB.

Fonte: Opto Eletrônica S.A.

O foco do presente trabalho é, portanto, resolver esse problema do desalinhamento da maneira mais precisa e rápida possível, uma vez que a imagem alinhada deve aparecer ao operador do equipamento logo após a aquisição da imagem, sem que ele espere um longo tempo por isso. No entanto, a imagem possui aproximadamente 11 Megapixels, o que resulta em 33 Megabytes para ser processado, no caso da imagem colorida. Assim, além da eficiência no alinhamento, o método deve ser otimizado para conseguir realizar sua função em um curto período de tempo, o que torna a execução do mesmo uma tarefa bastante desafiadora. A seguir, algumas técnicas de alinhamento de imagens, também conhecido como registro de imagens, serão apresentadas.

2.7. Registro de imagens

Em muitas aplicações de processamento de imagens é necessária a comparação *pixel a pixel* de duas imagens obtidas, por exemplo, em tempos diferentes. Para realizar esse tipo de comparação, é necessário alinhar (registrar) espacialmente as imagens, corrigindo,

dessa forma, diferenças de translação, rotação e escala. Em alguns casos, pode-se eliminar ou minimizar esse problema simplesmente com a calibração dos sensores. No entanto, na maioria dos casos, as imagens apresentam desalinhamento e a correção é necessária (PRATT, 2007).

O registro de imagens é largamente utilizado, principalmente nas áreas de sensoriamento remoto, imagens médicas e visão computacional. No geral, conforme em Zitová e Flusser (2003), essas aplicações podem ser divididas em quatro grupos principais: aquisições em diferentes pontos de vista, aquisições em diferentes tempos, aquisição com diferentes sensores e aquisição da cena e registro com um modelo.

No caso de imagens em diferentes pontos de vista (análise multivisualização), imagens da mesma cena são adquiridas de pontos de vista diferentes, permitindo aumentar o tamanho da visualização 2D ou 3D da cena. Como exemplo, pode-se citar o sensoriamento remoto, na qual se constrói um mosaico das imagens da área explorada (ZITOVÁ e FLUSSER, 2003).

Para o caso de imagens em diferentes tempos (análise multitemporal), imagens da mesma cena são adquiridas em tempos diferentes, frequentemente em uma base regular e possivelmente em diferentes condições. Na maioria dos casos, o objetivo é encontrar e avaliar mudanças na cena que aparecem entre as aquisições consecutivas. Como exemplo, pode-se citar equipamentos médicos, que possuem a função de monitorar a evolução de uma doença ou tumor (ZITOVÁ e FLUSSER, 2003).

Já para a aquisição com diferentes sensores (análise multimodal), imagens da mesma cena são adquiridas com sensores diferentes, com o objetivo de integrar informação obtida por diferentes fontes e ganhar uma representação da cena mais complexa e detalhada. Como exemplo, pode-se citar novamente o sensoriamento remoto, na qual a fusão da informação dos sensores com diferentes características, como imagens pancromáticas, pode oferecer uma melhor resolução espacial, além de poder oferecer imagens coloridas/multiespectrais com melhor resolução espectral (ZITOVÁ e FLUSSER, 2003).

Por fim, para a aquisição da cena e registro com um modelo, o objetivo é registrar a imagem real com uma imagem artificial, normalmente modelando a real, como uma representação computacional. Como exemplo, pode-se citar o sensoriamento remoto, na

qual se pode registrar os dados obtidos por satélite ou por aviões em mapas ou outras camadas de Sistemas de Informações Geográficas (ZITOVÁ e FLUSSER, 2003).

O presente trabalho concentra-se principalmente na análise multitemporal, na qual as imagens da retina são obtidas em instantes distintos. No entanto, outro grupo também deve receber destaque neste trabalho: a análise multimodal, ou seja, a aquisição com diferentes sensores. Embora usando o mesmo sensor, a iluminação aplicada é alterada e a informação da retina muda consideravelmente, o que dificulta ainda mais o registro dessas imagens.

A Figura 29 mostra a diferença de informação entre os três canais de cor obtidos a partir de um olho doente. As imagens foram obtidas com uma diferença aproximada de 150 ms entre os quadros e com iluminação RGB, conforme especificado anteriormente. Nelas, nota-se claramente as diferenças de intensidade das imagens nos diferentes canais, o que dificulta o registro.

Figura 29: Diferença de intensidade entre o canal vermelho (esquerda), verde (centro) e azul (direita) de exame realizado com olho doente

Fonte: Opto Eletrônica S.A.

2.7.1. Passos para registro de imagens

Conforme Zitová e Flusser (2003), a maioria dos métodos de registro de imagens consiste de quatro passos básicos: detecção de características, busca pelas características detectadas, estimativa do modelo de transformação e reamostragem da imagem e transformação, conforme exibido na Figura 30.

Figura 30: Passos básicos envolvidos no registro entre duas imagens

Fonte: Adaptado de Zitová e Flusser (2003)

Na detecção das características, objetos salientes e distintivos, como bordas, contornos, cantos e cruzamentos de linhas, são manualmente ou automaticamente detectados. No caso da busca de características, a correspondência entre as características detectadas na imagem a ser corrigida e na imagem referência é estabelecida. Vários descritores e medidas de similaridade podem ser utilizados para obter as relações espaciais entre as características detectadas. Para o caso da estimativa do modelo de transformação, o tipo e os parâmetros das funções de mapeamento para realizar o registro entre o par de imagens é estimado. Tais parâmetros são calculados conforme as características entre as duas imagens. Por fim, na reamostragem da imagem e transformação, a imagem a ser alinhada é transformada por meio das funções de mapeamento, corrigindo, dessa forma, o deslocamento inicial entre as duas imagens.

2.7.2. Principais métodos para registro de imagens

Os algoritmos de detecção de características são classificados em métodos baseados em características ou métodos baseados em área. Os métodos baseados em características são baseados na extração de estruturas salientes das imagens. Regiões significantes, linhas ou pontos (cantos de regiões e intersecções de linhas) são entendidos como características. Elas precisam ser distintivas e, além disso, deve ser possível de detectar eficientemente nas duas imagens. Esses métodos são recomendados se as imagens possuem objetos facilmente detectáveis. Esse é, frequentemente, o caso de aplicações de sensoriamento remoto, na qual as imagens contêm vários detalhes como cidades, rios, estradas, florestas, etc. Para o caso de imagens médicas, foco deste trabalho, os métodos baseados em características não são recomendados, uma vez que as imagens nem sempre são tão ricas em detalhes. Nesse caso, os métodos baseados em área são mais recomendados (ZITOVÁ e FLUSSER, 2003). Os métodos baseados em área, conhecido também como métodos de correlação ou *template matching*, mesclam o passo de detecção de características com o passo de busca das mesmas. Esses métodos processam as imagens, sem tentar detectar objetos salientes, realizando, na maior parte dos casos, a busca por regiões de tamanho constante, conhecidos como *templates*.

Existem várias funções de correlação, sendo neste trabalho abordados os métodos de *cross correlation* (correlação cruzada) e *phase correlation* (correlação de fase). Outros métodos e mais detalhes podem ser encontrados em Zitová e Flusser (2003) e Pratt (2007), esse último tomado como referência para as descrições abaixo.

2.7.2.1. Método para registro usando correlação cruzada

O método da correlação cruzada realiza o processamento da imagem no domínio do espaço. Considere $F_1(j, k)$ e $F_2(j, k)$, com $1 \leq j \leq J$ e $1 \leq k \leq K$, representando duas imagens discretas para serem registradas. Nesse caso, F_1 é considerada a imagem referência, sendo $F_2(j, k)$ descrita conforme a equação (1):

$$F_2(j, k) = F_1(j - j_0, k - k_0) \quad (1)$$

Para a Eq. (1), $F_2(j, k)$ é a versão transladada de $F_1(j, k)$ em que (j_0, k_0) são os deslocamentos (*offsets*) de translação. A correlação cruzada normalizada entre o par de imagens é definido por

$$R(m, n) = \frac{\sum_j \sum_k F_1(j, k) F_2\left(j - m + \frac{M+1}{2}, k - n + \frac{N+1}{2}\right)}{[\sum_j \sum_k [F_1(j, k)]^2]^{1/2} [\sum_j \sum_k [F_2\left(j - m + \frac{M+1}{2}, k - n + \frac{N+1}{2}\right)]^2]^{1/2}} \quad (2)$$

para $m = 1, 2, \dots, M$ e $n = 1, 2, \dots, N$, em que M e N são inteiros ímpares.

Na Figura 31, o retângulo tracejado especifica as fronteiras da região da função de correlação a partir do qual o canto superior esquerdo de $F_2(j, k)$ move-se no espaço em relação à $F_1(j, k)$. Os limites dos somatórios são dados por

$$\text{MAX} \left\{ 1, m - \frac{M-1}{2} \right\} \leq j \leq \text{MIN} \left\{ J, J + m - \frac{M+1}{2} \right\} \quad (3)$$

$$\text{MAX} \left\{ 1, n - \frac{N-1}{2} \right\} \leq k \leq \text{MIN} \left\{ K, K + n - \frac{N+1}{2} \right\} \quad (4)$$

Figura 31: Relações geométricas para a correlação cruzada entre duas imagens

Fonte: Adaptado de Pratt (2007)

Tais limites são indicados pela região escura na Figura 31 para o *offset* (a, b) . Essa região, conforme pode ser observada, caracteriza a região de intersecção entre as duas imagens e é conhecida como janela da função de correlação (PRATT, 2007).

Para o cálculo da correlação, usa-se uma janela de tamanho constante para diminuir o número de cálculos. Essa janela de tamanho $P \times Q$, também conhecida como *template*, é definida pelos seguintes limites no somatório:

$$m \leq j \leq m + J - M \quad (5)$$

$$n \leq k \leq n + K - N \quad (6)$$

Ainda, na Figura 31, as linhas pontilhadas especificam o máximo tamanho da região do *template*, que está no centro de $F_2(j, k)$. O tamanho do vetor da função de correlação $M \times N$, o da região de busca $J \times K$ e o tamanho do *template* $P \times Q$ são relacionados por:

$$M = J - P + 1 \quad (7)$$

$$N = K - Q + 1 \quad (8)$$

Para o caso especial em que a janela de correlação tem tamanho constante, a função de correlação pode ser reformulada como um processo de busca por *template*, conforme exibido na Figura 32. Considere que $S(u, v)$ denota um área de busca de tamanho $U \times V$ de $F_1(j, k)$, em que a borda superior esquerda tem um *offset* de (j_s, k_s) . Considere que $T(p, q)$ denota um *template* de tamanho $P \times Q$ extraído de $F_2(j, k)$, cuja borda superior esquerda tem um *offset* de (j_T, k_T) . Claramente, $P < U$ e $Q < V$.

Figura 32: Relação da região de busca com a região do *template* no método de correlação cruzada
Fonte: Adaptado de Pratt (2007)

Nesse caso, a função de correlação cruzada normalizada pode ser expressa por

$$R(m, n) = \frac{\sum_u \sum_v S(u, v) T\left(u - m + \frac{M+1}{2}, v - n + \frac{N+1}{2}\right)}{[\sum_u \sum_v [S(u, v)]^2]^{1/2} [\sum_u \sum_v [T\left(u - m + \frac{M+1}{2}, v - n + \frac{N+1}{2}\right)]^2]^{1/2}} \quad (9)$$

para $m = 1, 2, \dots, M$ e $n = 1, 2, \dots, N$, em que

$$M = U - P + 1 \quad (10)$$

$$N = V - Q + 1 \quad (11)$$

Para o cálculo de $R(m, n)$, os limites dos somatórios são:

$$m \leq u \leq m + P - 1 \quad (12)$$

$$n \leq v \leq n + Q - 1 \quad (13)$$

A biblioteca de visão computacional *OpenCV*, utilizada neste trabalho, implementa o método de correlação cruzada de maneira bastante eficiente, possibilitando que o processamento das iterações e somatórios seja realizado de maneira bastante rápida.

O protótipo da função do *OpenCV* é apresentado a seguir:

```
void cvMatchTemplate(const CvArr* image, const CvArr* templ, CvArr* result, int method);
```

Nesse método, passa-se a imagem na qual o *template* será procurado (*image*), o *template* (*templ*), o vetor resultante da correlação (*result*) e o método de correlação a ser utilizado (*method*), dentre os quais se encontra o método de correlação cruzada normalizado demonstrado. Com o resultado retornado em *result*, basta localizar o ponto com o valor máximo dessa matriz, o qual caracteriza o ponto de maior correlação com o *template* (OpenCV , 2012).

A Figura 33 mostra um exemplo do método sendo aplicado para encontrar um *template* na imagem. Nessa imagem, a matriz de correlação de fase é mostrada invertida (negativo) para melhor visualização do ponto máximo. Na imagem original, o fundo é escuro e o ponto de máximo é claro. Tal procedimento será adotado para a exibição de todas as matrizes de correlação presentes neste trabalho.

Figura 33: Imagem e *template* para busca (esquerda) e negativo da imagem da matriz da correlação cruzada obtida (direita) ao realizar o método de correlação cruzada entre o *template* e a imagem da retina

À esquerda, é exibida a imagem e um *template* para ser localizado na mesma; à direita, é exibida a matriz de correlação obtida utilizando o método de correlação cruzada

normalizado. Para a exibição, foi utilizado um *threshold* na imagem e após obtido o negativo para ressaltar o pico de correlação (ponto preto na imagem).

Para o caso de registro de duas imagens com esse método, escolhe-se *templates* da imagem referência, que podem ser caracterizados por bordas e cantos, e realiza-se a busca na imagem desalinhada. No entanto, um dos problemas dessa abordagem é a escolha automática dos *templates* a serem buscados. Além disso, o método de correlação cruzada explora pela busca direta das intensidades, sem nenhuma análise estrutural. Consequentemente, o método é sensível a variações na intensidade introduzidas por ruído, variações na iluminação, uso de diferentes sensores (ZITOVÁ e FLUSSER, 2003) ou ainda diferentes fontes de iluminação.

2.7.2.1. Método para registro usando correlação de fase

Uma alternativa para os problemas citados são os métodos baseados em área, utilizando processamento no domínio da frequência, também conhecidos como métodos de Fourier.

A Transformada Discreta de Fourier (*DFT*) de uma imagem digital $f(x, y)$ de tamanho $M \times N$ é dada por:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (14)$$

Já a Transformada Discreta de Fourier Inversa (*IDFT*), para a mesma imagem é dada por:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (15)$$

Mais detalhes sobre transformadas podem ser encontrados em Gonzalez e Woods (2009).

Dentre os métodos de registro baseados nessas transformadas, destaca-se o conhecido como correlação de fase, que se baseia no teorema de deslocamento de Fourier. Diferente do método de correlação cruzada, esse método é praticamente insensível a ruídos, não uniformidades e variações na iluminação das cenas. Além disso, como o processamento é realizado no domínio da frequência e depende de poucas operações matemáticas, o

processamento é bastante rápido, principalmente para imagens grandes (ZITOVÁ e FLUSSER, 2003).

No método de correlação de fase, calcula-se o produto espectral cruzado entre a imagem de referência e a imagem a ser alinhada, procurando picos na matriz resultante no domínio do espaço (PRATT, 2007). Conforme em Pratt (2007), considere a imagem $F_1(x, y)$ como imagem referência e $F_2(x, y)$ como a imagem a ser alinhada, sendo que $F_2(x, y)$ é a imagem $F_1(x, y)$ translada por um *offset* de (x_0, y_0) em relação a (x, y) :

$$F_2(x, y) = F_1(x - x_0, y - y_0) \quad (16)$$

A propriedade do deslocamento de Fourier afirma que um deslocamento da imagem no espaço resulta em um deslocamento na fase do sinal na frequência:

$$F(x - a, y - b) \rightarrow \mathcal{F}(\omega_x, \omega_y) e^{-i(\omega_x a + \omega_y b)} \quad (17)$$

Assim, no domínio da frequência, tem-se $\mathcal{F}_2(\omega_x, \omega_y)$ como:

$$\mathcal{F}_2(\omega_x, \omega_y) = \mathcal{F}_1(\omega_x, \omega_y) e^{-i(\omega_x x_0 + \omega_y y_0)} \quad (18)$$

A partir dessa propriedade, o fator de deslocamento de fase exponencial pode ser calculado pelo produto espectral cruzado, $G(\omega_x, \omega_y)$, expresso por:

$$G(\omega_x, \omega_y) = \frac{\mathcal{F}_1(\omega_x, \omega_y) \mathcal{F}_2^*(\omega_x, \omega_y)}{|\mathcal{F}_1(\omega_x, \omega_y) \mathcal{F}_2^*(\omega_x, \omega_y)|} = e^{i(\omega_x x_0 + \omega_y y_0)} \quad (19)$$

Possuindo $G(\omega_x, \omega_y)$ no domínio da frequência, pode-se calcular a transformada inversa, obtendo $G(x, y)$, conhecida como matriz de *offset* espacial, dada por:

$$G(x, y) = \delta(x - x_0, y - y_0) \quad (20)$$

Com $G(x, y)$, basta encontrar o ponto máximo da matriz que, assim como na correlação cruzada, caracteriza o pico de correlação. Na Figura 34 é exibido o esquemático do resultado da correlação de fase (matriz de *offset* espacial), conforme explicitado acima, exibindo como os valores do *offset* horizontal e vertical podem ser encontrados. Para a exibição foi obtido o negativo da imagem da matriz de correlação para facilitar a visualização.

Figura 34: Esquemático do resultado da correlação de fase entre dois canais que permite encontrar a transformação de translação entre ambos – Nesse caso, é mostrado o negativo da imagem resultante, no qual o ponto preto representa a máxima correlação entre os dois quadros

Na Figura 35 são exibidas duas imagens e a matriz de *offset* espacial obtida.

Figura 35: Resultado para correlação de fase, mostrando o negativo do pico de correlação (parte inferior) entre as duas imagens (para superior) – o ponto preto representa o *offset* entre as duas imagens, sendo esse deslocamento calculado em relação ao centro do resultado da correlação de fase

O centro dessa matriz (ponto vermelho) representa o ponto em que $x_0 = 0$ e $y_0 = 0$. Para essa matriz, o ponto de máximo está no segundo quadrante, próximo ao centro. Esse ponto representa o *offset* entre as duas imagens em relação ao centro. No caso desse exemplo, o *offset* da segunda imagem em relação à primeira é $x_0 = 10 \text{ pixels}$ e $y_0 = 23 \text{ pixels}$. Para a exibição foi obtido o negativo da imagem da matriz de correlação para facilitar a visualização.

O algoritmo de correlação de fase, conforme já especificado, é robusto em relação a ruídos, variações de iluminação e possui rápido processamento no domínio da frequência. No entanto, ele não consegue, inicialmente, detectar variações de escala e rotação da imagem, apenas de translação. Para resolver esse problema, o algoritmo de correlação de fase pode ser estendido para a correção de escala e rotação, utilizando coordenadas polares, conforme descrito em Pratt (2007).

Considere um par de imagens em que a segunda imagem é translada por um *offset* (x_0, y_0) e rotacionada por um ângulo θ_0 em relação à primeira imagem. Dessa forma, temos:

$$F_2(x, y) = F_1(x \cos \theta_0 + y \sin \theta_0 - x_0, -x \sin \theta_0 + y \cos \theta_0 - y_0) \quad (21)$$

Fazendo a transformada de Fourier dos dois lados, obtêm-se a seguinte relação:

$$\mathcal{F}_2(\omega_x, \omega_y) = \mathcal{F}_1(\omega_x \cos \theta_0 + \omega_y \sin \theta_0, -\omega_x \sin \theta_0 + \omega_y \cos \theta_0) e^{-i(\omega_x x_0 + \omega_y y_0)} \quad (22)$$

A componente da rotação pode ser isolada calculando a magnitude $\mathcal{M}_1(\omega_x, \omega_y)$ e $\mathcal{M}_2(\omega_x, \omega_y)$.

Vale ressaltar que, neste ponto, a influência do *offset* (x_0, y_0) foi removida, uma vez que a informação do mesmo afeta a fase ($e^{-i(\omega_x x_0 + \omega_y y_0)}$) e não a magnitude do sinal, conforme descrito na propriedade do deslocamento de Fourier. Representando as variáveis da magnitude na forma polar, tem-se:

$$\mathcal{M}_2(\rho, \theta) = \mathcal{M}_1(\rho, \theta - \theta_0) \quad (23)$$

A partir da equação (23), pode-se utilizar o método de correlação de fase para se determinar o ângulo de rotação θ_0 , da mesma forma em que foi utilizado para o cálculo da translação.

Caso a segunda imagem também tenha escala aplicada em relação à primeira com os fatores de escala (a, b) , pela propriedade da escala da Transformada de Fourier, tem-se:

$$\mathcal{F}_2(\omega_x, \omega_y) = \frac{1}{|ab|} \mathcal{F}_1\left(\frac{\omega_x}{a}, \frac{\omega_y}{b}\right) \quad (24)$$

Convertendo as variáveis da frequência para a escala logarítmica, a escala pode ser convertida em um movimento translacional. Dessa forma,

$$\mathcal{F}_2(\log \omega_x, \log \omega_y) = \frac{1}{|ab|} \mathcal{F}_1(\log \omega_x - \log a, \log \omega_y - \log b) \quad (25)$$

Assim como na correção da rotação, o método de correlação de fase pode ser aplicado para determinar os fatores de escala (a, b) .

Conforme em Reddy e Chatterji (1996), se os fatores de escala horizontal e vertical forem iguais, ou seja, $a = b$, tem-se na forma polar:

$$F_1(\rho_1, \theta_1), \text{ onde } \rho_1 = (x^2 + y^2)^{1/2} \text{ e } \theta_1 = \operatorname{tg}^{-1}\left(\frac{y}{x}\right) \quad (26)$$

$$F_2(\rho_2, \theta_2), \text{ onde } \rho_2 = \left(\left(\frac{x}{a}\right)^2 + \left(\frac{y}{a}\right)^2\right)^{1/2} \text{ e } \theta_2 = \operatorname{tg}^{-1}\left(\frac{\frac{y}{a}}{\frac{x}{a}}\right) = \operatorname{tg}^{-1}\left(\frac{y}{x}\right) = \theta_1 \quad (27)$$

Dessa forma, se existe uma transformação de translação (x_0, y_0) , rotação (θ_0) e escala (a) entre F_1 e F_2 , a magnitude do espectro de Fourier das duas imagens na forma polar é relacionado por:

$$\mathcal{M}_2(\rho, \theta) = \mathcal{M}_1\left(\frac{\rho}{a}, \theta - \theta_0\right) \quad (28)$$

Vale ressaltar que os fatores de translação (x_0, y_0) não aparecem na equação da magnitude, uma vez que alteram apenas a fase no domínio da frequência. Dessa forma, pode-se calcular as três transformações de maneira independente.

Ao analisar a equação (28), nota-se a razão $\frac{\rho}{a}$. Conforme realizado na equação (25), pode-se aplicar o logaritmo, chegando na seguinte expressão na forma *log-polar* (REDDY e CHATTERJI, 1996):

$$\mathcal{M}_2(\log \rho, \theta) = \mathcal{M}_1\left(\log \frac{\rho}{a}, \theta - \theta_0\right) = \mathcal{M}_1(\log \rho - \log a, \theta - \theta_0) \quad (29)$$

Assim, a partir da transformação para a escala *log-polar* e usando o algoritmo de correlação de fase, pode-se determinar o valor da escala a e o ângulo θ_0 entre as duas

imagens com uma única operação. A correção da translação é posteriormente realizada, conforme especificado nas equações de (16) a (20).

Para a correção de escala e rotação, o resultado da correlação de fase é exibido na Figura 36, do qual podem ser obtidos os valores do logaritmo da escala (deslocamento horizontal em relação ao centro) e a rotação (deslocamento vertical em relação ao centro).

Figura 36: Esquemático do resultado da correlação de fase da magnitude do sinal na forma *log-polar* de dois canais, o qual permite encontrar a escala e a rotação entre ambos – o ponto preto representa a máxima correlação entre as duas imagens

Considerado o processamento no domínio da frequência, iniciou-se nesse trabalho a implementação de algoritmos para realizar o registro das imagens da retina obtidas em tempos diferentes e com diferentes fontes de iluminação.

3. Metodologia Proposta

Ao analisar as imagens da retina com a movimentação do olho entre os *flashes*, notou-se que o maior desalinhamento acontecia devido à translação das imagens, uma vez que o paciente, durante o exame, fica com a cabeça fixa entre a queixeira e testeira do equipamento, como pode ser visto na Figura 37, o que deve tornar o desalinhamento das imagens por rotação e escala menos importante. Mesmo assim, a proposta desse trabalho é realizar o registro das imagens dos canais RGB para a formação da imagem colorida considerando não só a correção da translação, como também da rotação e escala, utilizando para isso a técnica de correlação de fase.

Figura 37: Paciente posicionado para o exame com cabeça fixa (testa e queixo apoiados) o que dificulta o desalinhamento entre os *flashes* por rotação e escala

Fonte: Adaptado de Vieira (2013)

Diferentemente do exemplo apresentado na Figura 35, os pares de imagens a serem registrados são obtidos com diferentes fontes de iluminação e, nesse caso, a diferença de informação entre as duas imagens é significativa, o que deixa o trabalho ainda

mais desafiador. Essa diferença de informação entre os canais pode ser visualizada na Figura 38, na qual também é exibido o negativo da imagem da matriz de correlação de fase entre ambas.

Figura 38: Imagens correspondentes ao canal G (superior esquerda) e ao canal R (superior direita) e o resultado da correlação de fase (inferior), mostrando a aplicação do método em imagens com diferentes intensidades de informação devido à iluminação diferente de cada canal

Como se pode observar, os canais G e R da Figura 38 apresentam informações bastante distintas, como por exemplo, variações na iluminação e mesmo informações diferentes – alguns vasos visualizados no canal verde não podem ser visualizados no canal vermelho. Nesse caso, como se pode visualizar no resultado da correlação de fase, surgem outros pontos na imagem, porém o de maior intensidade, nesse exemplo, é o que está deslocado em relação ao centro 27 *pixels* para direita e 107 para cima, conforme obtido pelo

método de correlação de fase. O resultado da correlação entre as duas imagens pode ser visualizado na Figura 39.

Figura 39: Canais combinados (50% Canal R + 50% Canal G) antes (esquerda) e depois do registro (direita)

Vale ressaltar que o método de correlação de fase já foi utilizado para o processamento e registro de imagens da retina, conforme em Ryan e Heneghan (1999), Schardosim, Scharcanski *et al.* (2011) e Chanwimaluang (2006). No entanto, o registro de imagens da retina obtidas com diferentes fontes de iluminação (vermelho, verde e azul), como é o caso desse trabalho, é uma abordagem nova, uma vez que as imagens a serem alinhadas apresentam diferenças significativas, não só de intensidade de iluminação, mas também das estruturas do olho, as quais podem ser visualizadas no diferentes canais.

Ainda, durante a implementação do método, foi detectado o seguinte problema: se existirem imperfeições no CCD ou no conjunto de lentes, como pequenos pontos de poeira, pode-se gerar componentes de alta frequência no espectro e, como esses pontos não se movimentam durante os três *flashes*, o método pode convergir para a concentração de pontos na origem (centro) do produto espectral calculado. Nesse caso, ao localizar o pico da matriz da correlação de fase resultante, esse estará no centro da imagem e o deslocamento retornado será $x_0 = 0$ e $y_0 = 0$. Para evitar esse problema, antes da realização do registro, um filtro passa-baixa deve ser aplicado na imagem referência e na imagem a ser corrigida para atenuar as altas frequências do espectro. Nesse trabalho, foi utilizado um filtro passa-baixa da média com janela 3 x 3 (GONZALEZ e WOODS, 2009).

Com a aplicação do filtro passa-baixa, a influência desses pequenos pontos de imperfeições diminuíram e o pico do produto espectral cruzado ocorreu na região da maior correlação entre os vasos retinianos, conforme esperado para que o registro fosse realizado corretamente.

Na Figura 40 o problema de imperfeições no CCD ou no conjunto de lentes é exibido. Nesse caso, visualizamos três pontos na imagem resultante. Tal fato ocorre uma vez que imperfeição estava fixa no sistema óptico e o olho moveu-se durante os três *flashes*. Como o alinhamento é baseado nas informações da retina e, além disso, foi aplicado o filtro passa-baixa para remover as componentes das imperfeições do resultado da correlação de fase, é possível identificar a formação de três pontos em cores primárias do sistema subtrativo de cores (magenta, ciano e amarelo), dado que para cada *flash* a imperfeição impediu a chegada de apenas um canal de cor ao sensor.

Figura 40: Imperfeição no sistema óptico gerando três pontos na imagem devido à movimentação do olho
Fonte: Adaptado de Vieira (2013)

Além do filtro passa-baixa, como o que importa é a informação das estruturas da retina para o correto registro das imagens, aplicou-se também um filtro passa-alta Laplaciano (GONZALEZ e WOODS, 2009) para o realce das estruturas do fundo do olho, conforme equação que segue:

$$h = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (30)$$

Na Figura 41 é exibido o diagrama simplificado do algoritmo para realizar a correção da imagem. Na ramificação do algoritmo, depois de dividir os canais, são disparadas duas

threads para processamento paralelo: uma para realizar o registro entre o canal verde e vermelho e outra para realizar o registro entre o canal verde e azul. *Threads* são linhas de controles independentes que podem existir dentro de um mesmo processo computacional, ou seja, o processo pode-se dividir em duas ou mais tarefas que podem ser executadas de maneira concorrente (TANENBAUM e WOODHULL, 2000). Tal fato possibilita que o processamento seja realizado de maneira mais rápida e a imagem corrigida seja exibida em tempo menor ao operador do equipamento.

Figura 41: Diagrama simplificado para corrigir a imagem pelo método proposto, no qual a ramificação para calcular os parâmetros mostra a atuação de duas *threads* para o processamento paralelo da correção do canal vermelho e azul em relação ao verde

Nesse diagrama, inicialmente a imagem original a ser corrigida é adquirida dos canais RGB. Após, aplica-se um filtro passa baixa, que minimiza os efeitos de pequenas imperfeições no sistema óptico e CCD, e um filtro passa alta para destacar as bordas e facilitar o cálculo dos parâmetros de registro. Calcula-se, então, os parâmetros (escala,

rotação e translação) para posterior registro do canal verde em relação ao vermelho (lado esquerdo) e do verde em relação ao direito (lado direito). A função “Calcular Parâmetros” mostrada na Figura 41 será explicada em detalhes no diagrama seguinte (Figura 42). Após serem obtidos os parâmetros da matriz de transformação, os canais originais são então escalados, rotacionados e transladados. Por fim, os canais são novamente mesclados para a formação da imagem colorida, que é retornada pela função. Vale ressaltar que a filtragem passa-baixa e a passa-alta não afetam a qualidade final da imagem, uma vez que são utilizadas apenas para os cálculos dos parâmetros; a correção após esse cálculo é realizada utilizando os canais originais da imagem.

A seguir, é exibido na Figura 42 o diagrama que calcula os parâmetros de registro entre os dois quadros (f_1 e f_2), retornando a escala, ângulo e *offset* de translação entre eles.

Figura 42: Diagrama simplificado para calcular os parâmetros de registro entre dois canais, retornando a escala, rotação e *offset* horizontal e vertical entre eles

O diagrama para calcular os parâmetros para o registro entre dois quadros funciona da seguinte forma: inicialmente um quadrado do centro de lado 1500 pixels é obtido para

remover a região da máscara da imagem, que visa remover as imperfeições da lente captadora ou problemas no sensor. Tal máscara será explicada em detalhes no próximo capítulo. Após, calcula-se a transformada de Fourier dos dois quadros e sua magnitude. Ainda, ambas as magnitudes são convertidas para a forma *log-polar* para ser possível aplicar o algoritmo de correlação de fase e obter a escala e o ângulo entre as duas imagens, conforme explicado das equações (21) a (29). O algoritmo de correlação de fase é então aplicado às duas magnitudes na forma *log-polar* e a escala e o ângulo entre os dois quadros é então retornado. Faz-se então a correção de escala e ângulo para que, posteriormente, se realize a correção da translação. Um quadrado de 1500 pixels do centro da imagem corrigida é novamente recortado e o método de correlação de fase é aplicado mais uma vez para o cálculo dos *offsets* de translação. Os valores encontrados para escala, rotação e translação são então retornados pela função. O método de correlação de fase entre dois quadros (f_1 e f_2) é explicado no diagrama simplificado exibido na Figura 43.

Figura 43: Diagrama simplificado para calcular a correlação de fase entre dois quadros

Nesse diagrama os dois quadros são convertidos para o domínio da frequência e o produto espectral cruzado entre eles é calculado conforme equação (19). Após, faz-se a transformada de Fourier inversa e calcula-se o centro da imagem, no qual o *offset* horizontal e vertical é zero. Localiza-se, então, o pico de correlação, que indica a máxima similaridade

entre os dois quadros. O deslocamento horizontal e vertical desse pico em relação ao centro caracteriza o *offset* horizontal e vertical entre os dois quadros.

Mais detalhes sobre a implementação do algoritmo pode ser encontrado no apêndice desse trabalho, no qual consta o algoritmo em pseudocódigo que foi utilizado para o desenvolvimento das funcionalidades exibidas nos diagramas.

O desenvolvimento das rotinas presentes no método proposto foi realizado utilizando a *IDE C++ Builder 2010* e a biblioteca de processamento de imagens *OpenCV*. Conforme já explicitado, um dos pontos críticos é a velocidade de processamento, uma vez que o operador deseja disparar o *flash* e, em poucos instantes, visualizar a imagem capturada. Dessa forma, optou-se pela utilização da linguagem C++ e da biblioteca *OpenCV* devido ao seu ótimo desempenho, principalmente para o cálculo das transformadas de Fourier presentes no método.

4. Resultados e Discussões

4.1. Exemplo da aplicação do método

Um exemplo da aplicação do método proposto para o registro é apresentado a seguir. Nesse exemplo, a imagem original obtida será exibida, bem como o resultado da correlação de fase entre os canais e a respectiva imagem corrigida. Em seguida, exibe-se a avaliação dos resultados obtidos pelo método proposto para corrigir a movimentação do olho entre os três *flashes*. Na Figura 44 pode-se visualizar a imagem original obtida do olho de um paciente, sem o alinhamento proposto neste trabalho. O desalinhamento entre os canais se torna visível pelo efeito de sombra nos vasos retinianos.

Figura 44: Imagem original com desalinhamento entre os canais antes da aplicação da correção proposta
Fonte: Adaptado de Stuchi, Vieira e Rodrigues (2012)

Para o processamento, selecionam-se os canais vermelho, verde e azul dessa imagem. Como o canal verde apresenta o melhor contraste para as imagens da retina, esse é tomado como referência e os canais vermelho e azul serão alinhados em relação a ele.

Os canais verde e vermelho desse exemplo podem ser visualizados na Figura 45. Nota-se que existe bastante diferença entre os dois canais, o que dificulta o registro das

duas imagens. Pode-se visualizar que a diferença entre os canais não é apenas de intensidade do nível de cinza, mas também estrutural, uma vez que no canal vermelho não é possível visualizar vários vasos presentes no canal verde.

Figura 45: Canal G (superior esquerda) e R (superior direita) antes da correção e o produto espectral cruzado entre eles para a correção de escala e rotação (inferior esquerdo) e translação (inferior direito).

Fonte: Adaptado de Stuchi, Vieira e Rodrigues (2012)

A matriz resultado da correlação de fase para a correção de escala com rotação e translação entre os canais também é exibida na Figura 45. Nessa representação, o ponto vermelho foi introduzido apenas como referência para o centro, onde o deslocamento em x e y é zero. Para essa exibição, o produto espectral foi normalizado em 8 bits e foi aplicado um *threshold* para remoção dos valores inferiores ao nível de cinza 200. Após, foi obtido o

negativo da imagem, objetivando, dessa forma, facilitar a visualização dos pontos que indicam o deslocamento entre os canais. Nesse exemplo, para a correção de escala e rotação (imagem inferior esquerda), o pico ocorreu bem no centro da imagem, o que resultou em uma escala de 1.0 e ângulo de 0° . Já para a correção de translação (imagem inferior direita), o pico à esquerda do centro foi localizado com o método proposto e, a partir dele, obteve-se um *offset* de 95 *pixels* na horizontal e 4 *pixels* na vertical do canal vermelho em relação ao verde.

O mesmo foi realizado para os canais verde e azul, que são exibidos na Figura 46.

Figura 46: Canal G (superior esquerda) e B (superior direita) antes da correção e o produto espectral cruzado entre eles para a correção de escala e rotação (inferior esquerda) e translação (inferior direita).

Fonte: Adaptado de Stuchi, Vieira e Rodrigues (2012)

Diferente do canal vermelho, o azul apresenta informação de intensidade e estrutural bastante parecida com o canal verde, o que facilita a correção. O resultado da correlação de fase entre eles também é exibido na Figura 46. Nesse caso, para a correção de rotação e escala (imagem inferior esquerda), foi encontrado 0 *pixels* na horizontal e 3 *pixels* na vertical, o que resulta em uma escala de 1.0 e um ângulo de 0.72° . No caso da translação (imagem inferior direita), o pico também está à esquerda do centro e, depois do cálculo usando o método proposto, encontrou-se o *offset* de 77 *pixels* na horizontal e 11 *pixels* na vertical do canal azul em relação ao verde. Obtidos os deslocamentos entre os canais, a imagem pode ser corrigida aplicando as correções de escala, rotação e translação medidas com o método proposto. Na Figura 47, a imagem resultante após o alinhamento dos canais é exibida. Nessa imagem, diferentemente da imagem original, nota-se que o efeito de sombra criado pelo desalinhamento não está mais presente, além de apresentar uma maior definição dos vasos da retina.

Figura 47: Imagem corrigida com canais alinhados (esquerda) e comparação da região do nervo óptico antes (superior direita) e depois do alinhamento (inferior direita)

Fonte: Adaptado de Stuchi, Vieira e Rodrigues (2012)

Pode-se notar, pelos valores obtidos, que a variação dos canais devido à escala e rotação é muito pequena, principalmente pelo fato do paciente estar com a cabeça fixa no

equipamento, com seu queixo e testa apoiados sobre a queixeira e testeira, respectivamente, conforme explicado anteriormente e exibido na Figura 37.

A seguir é apresentada na Figura 48 uma comparação entre duas imagens da região do nervo óptico do mesmo olho do paciente, sendo que uma foi obtida utilizando o sistema proposto nesse trabalho e a outra obtida utilizando um retinógrafo comercial Canon CX-1 com câmera CCD colorida de 16 Megapixels. O espectro de Fourier para as duas imagens também é exibido.

Figura 48: Imagem do nervo óptico obtida com arquitetura proposta (esquerda) e com retinógrafo Canon CX-1 de 16 Megapixels (direita) e o espectro de Fourier de ambas as imagens (inferior)

Pode-se notar que a imagem obtida com a câmera CCD monocromática e o método de registro proposto apresenta mais detalhes do que a que foi obtida com uma câmera CCD colorida que utiliza filtro de Bayer. Ao analisar o espectro de Fourier de ambas as imagens, nota-se mais componentes de alta frequência na imagem obtida pela arquitetura proposta

neste trabalho do que no retinógrafo comercial. Pode-se observar uma "nuvem branca" mais visível no espectro da esquerda, à medida que se afasta do centro, mostrando que a imagem gerada pela presente proposta possui componentes de alta frequência com maior amplitude do que na imagem do retinógrafo comercial que usa câmera CCD colorida com filtro de Bayer. No entanto, a figura tem apenas caráter ilustrativo, dado que, apesar das imagens terem sido capturadas no mesmo dia e nas melhores condições de foco obtidas por um operador experiente, algumas variações podem ocorrer como acoplamento de iluminação no olho do paciente e influência do sistema óptico no resultado final da imagem.

A Figura 49 exibe outra imagem obtida com a arquitetura proposta de maneira ampliada, na qual se pode visualizar a boa definição dos vasos da retina e microvasos do nervo óptico.

Figura 49: Imagem obtida com a arquitetura proposta mostrando boa definição dos vasos da retina, microvasos do nervo óptico e fibras nervosas

Fonte: Opto Eletrônica S. A.

Também se pode observar nesta imagem ranhuras saindo do nervo óptico para a periferia, que são fibras nervosas presentes na retina (CSILLAG, 2005) e que são muito difíceis de serem visualizadas em exames de retinografia.

4.2. Avaliação do método proposto

Para a avaliação e cálculo de erro do método proposto não foi possível utilizar o olho humano, uma vez que se torna impossível conseguir duas imagens idênticas da retina na mesma condição, sem movimentação, sendo uma obtida com o sistema proposto e a outra com a arquitetura tradicional com *flash* branco e sensor colorido. Dessa forma, para a avaliação, foi utilizado um olho mecânico da *Ocular Instruments Inc* (OCULAR INC, 2010), exibido na Figura 50, que simula o olho natural contendo superfícies como lentes do cristalino e retina.

Figura 50: Olho mecânico utilizado para a avaliação do método proposto
Fonte: Adaptado de Ocular Instruments (2010)

Como o olho mecânico não apresenta movimentação entre os *flashes*, dado que pode ser fixo no equipamento, pode-se gerar uma imagem referência (*ground truth*), sem o desalinhamento dos canais, e outra com movimentação do olho para a avaliação do desempenho do método. Assim, inicialmente produziram-se duas imagens: uma do olho mecânico parado e outra com movimentação aleatória entre os *flashes*. A imagem obtida com a movimentação foi então corrigida pelo método proposto e pôde ser comparada com a imagem referência.

Para realizar a validação do método proposto, foram obtidos 50 conjuntos de imagens do olho mecânico com diferentes configurações de foco, iluminação e posicionamento. Cada conjunto contém inicialmente duas imagens: uma imagem colorida obtida do olho parado (referência) e outra obtida com a movimentação do olho entre os *flashes* realizada manualmente. No caso da imagem movimentada, a obtenção dos quadros foi realizada a cada 1 s, permitindo assim ao operador humano movimentar o olho mecânico antes da obtenção do próximo canal.

O método de registro de imagens proposto neste trabalho foi, então, aplicado nas imagens desalinhadas de cada conjunto e a imagem resultante também foi adicionada ao mesmo para posterior comparação com a imagem referência. Na comparação, um quadrado de 1500 *pixels* do centro de cada uma das imagens foi tomado para o cálculo da similaridade entre elas. O quadrado é utilizado para remover da comparação a região da máscara circular preta inserida via software na imagem que, depois da correção, é diminuída para remover o efeito do desalinhamento fora da região útil da imagem. A máscara é utilizada para que as bordas da lente captadora não apareçam na imagem que é mostrada para o médico. A Figura 51 apresenta a imagem antes e depois da aplicação da máscara, bem como o quadrado tomado para o cálculo.

Figura 51: Imagem sem a máscara na qual aparece a borda da lente captadora (esquerda) e com máscara inserida via software e quadrado usado na comparação (direita)

Utilizando tais conjuntos, foi realizada a análise qualitativa e quantitativa para a correção em cada um deles. A análise qualitativa foi feita utilizando inspeção visual dos 50 conjunto de imagens criados. Alguns conjuntos são exibidos nas Figuras 52 a 60. Analisando tais conjuntos, pode-se notar que foram obtidos com diferentes condições de iluminação, posição do olho e foco.

Figura 52: Conjunto para avaliação 1 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 53: Conjunto para avaliação 2 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

A região branca presente nas imagens do olho mecânico representa o nervo óptico e as estruturas avermelhadas representam os vasos da retina. No caso da figura 53, existe também uma estrutura de aparência metálica no lado esquerdo da imagem para representar um parasita da doença toxoplasmose. Tal parasita se aloja na retina, podendo causar danos irreversíveis à visão do paciente.

Figura 54: Conjunto para avaliação 3 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 55: Conjunto para avaliação 4 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 56: Conjunto para avaliação 5 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 57: Conjunto para avaliação 6 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Novamente a estrutura metálica que representa o parasita da doença toxoplasmose é exibida na Figura 57. Nesse caso, pode-se visualizar que devido à movimentação entre os três *flashes* e ao consequente desalinhamento entre os três canais, a imagem central apresenta uma coloração vermelha, verde e azul na região que deveria ser metálica. No entanto, após a correção, a imagem da esquerda é exibida com a coloração correta.

Figura 58: Conjunto para avaliação 7 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 59: Conjunto para avaliação 8 - imagem referência do olho mecânico parada (direita), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Figura 60: Conjunto para avaliação 9 - imagem referência do olho mecânico parada (esquerda), imagem desalinhada com movimentação entre os *flashes* (centro) e imagem corrigida pelo método proposto (direita)

Avaliando as imagens exibidas, pode-se notar que, visualmente, o registro funcionou corretamente, uma vez que a imagem central desalinhada foi corrigida, resultando a imagem mostrada à direita. Tal imagem resultante é muito similar à imagem referência (*ground truth*), conforme se pode observar comparando as imagens da direita com a da esquerda.

No entanto, devido ao fato da análise qualitativa ser subjetiva, também foi realizada a análise quantitativa. Para realizar essa avaliação, as imagens coloridas obtidas com a combinação dos canais RGB foram convertidas para imagens que representassem apenas os tons de cinza. Tal conversão foi realizada uma vez que se pretendeu avaliar o algoritmo de registro a partir da informação da posição dos vasos, sem levar em consideração a coloração. Para a conversão, o canal *I* (intensidade) do sistema de cor HSI (*Hue, Saturation e Intensity*) foi extraído, conforme equação (GONZALEZ e WOODS, 2009):

$$Y = \frac{1}{3}(R + G + B) \quad (31)$$

Obtendo as imagens em tons de cinza, métodos objetivos de avaliação de qualidade de imagem foram calculados. O objetivo dessa avaliação é medir o erro e a similaridade entre a imagem referência e a corrigida, visando a demonstrar que a correção é eficiente e precisa, ou seja, possui baixo erro e alta similaridade. Foram considerados os seguintes métodos: NMSE (*Normalized Mean Square Error*) (WANG e BOVIK, 2009) e MSSIM (*Mean Structural Similarity Index*) (WANG, BOVIK, *et al.*, 2004).

O cálculo do NMSE para o par de imagens $I(i, j)$ e $K(i, j)$ de tamanho $m \times n$ é dado pelas seguintes expressões:

$$\bar{I} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(i, j) \quad (32)$$

$$\bar{K} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} K(i, j) \quad (33)$$

$$NMSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \frac{[I(i, j) - K(i, j)]^2}{\bar{I}\bar{K}} \quad (34)$$

Já o cálculo do SSIM busca a comparação não apenas da intensidade ponto a ponto dos *pixels* da imagem (luminância), como no caso do NMSE, mas também há uma comparação estrutural e de contraste. Tal fato permite uma métrica de similaridade mais próxima à percepção humana, retratando, dessa forma, com maior fidelidade a relação entre as imagens. O diagrama básico dos passos para o cálculo do SSIM é exibido na Figura 61.

Figura 61: Diagrama do sistema de medida da média da similaridade estrutural (MSSIM) , no qual exibe de forma básica a análise de luminância, contraste e estrutura entre os dois sinais

Fonte: Adaptado de Wang, Bovik *et al*(2004)

O índice SSIM é calculado usando várias janelas na imagem restaurada e na imagem referência. Tipicamente, usa-se tamanhos de janela de 8×8 pixels. O cálculo para duas janelas x e y de tamanho $W \times W$ é dado por:

$$SSIM(x, y) = l(x, y)c(x, y)s(x, y) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right) \quad (35)$$

onde $l(x, y)$, $c(x, y)$ e $s(x, y)$ são as similaridades da luminância, contraste e estrutura; μ_x e μ_y é a média e σ_x e σ_y a variância de x e y , respectivamente; σ_{xy} é a covariância de x e y e C_1 , C_2 e C_3 são constantes obtidas experimentalmente pelos autores. Na prática, a média dos índices SSIM (MSSIM) calculada para todas as $W \times W$ janelas extraídas da imagem é usada para avaliar a qualidade global da imagem. O resultado do índice MSSIM é um valor decimal entre -1 e 1, sendo que o valor 1 é alcançado apenas no caso das duas imagens serem idênticas (WANG, BOVIK, *et al.*, 2004).

Cálculos de ambos os parâmetros foram realizados para todos os 50 conjuntos de imagens descritos anteriormente, obtendo os resultados exibidos na Tabela 1. Nesse conjunto, calculou-se inicialmente apenas para a correção de translação e, após, calculou-se para a correção completa (escala, rotação e translação) para que se pudesse comparar o ganho entre as duas correções. Vale ressaltar que, quanto maior o erro (NSME), menor a similaridade entre as imagens (MSSIM).

Tabela 1: Cálculo do NMSE e MSSIM para cada conjunto após realizar a correção de translação e a correção completa (translação, rotação e escala)

Conjunto	Correção de translação		Correção completa	
	NMSE ($\times 10^{-4}$)	MSSIM	NMSE ($\times 10^{-4}$)	MSSIM
1	2,83	0,9814	2,69	0,9827
2	9,54	0,9898	8,85	0,9902
3	2,97	0,9778	2,71	0,9806
4	5,81	0,9748	5,88	0,9748
5	4,76	0,9742	4,57	0,9785
6	3,19	0,9825	3,19	0,9825
7	6,34	0,9846	6,34	0,9846
8	2,51	0,9854	2,75	0,9860
9	6,62	0,9757	6,57	0,9760
10	7,43	0,9801	7,83	0,9812
11	3,06	0,9760	2,84	0,9794
12	3,86	0,9773	4,46	0,9805
13	3,76	0,9762	3,47	0,9798
14	38,62	0,9783	39,35	0,9789
15	4,99	0,9755	4,21	0,9772
16	4,64	0,9776	4,34	0,9801
17	3,92	0,9753	3,08	0,9795
18	55,27	0,9823	54,91	0,9825
19	2,95	0,9766	2,65	0,9799
20	2,59	0,9759	2,59	0,9759
21	3,98	0,9825	4,29	0,9840
22	53,43	0,9806	53,43	0,9806
23	4,92	0,9762	4,38	0,9767
24	18,46	0,9867	15,13	0,9874
25	11,61	0,9894	11,00	0,9899
26	4,48	0,9831	4,71	0,9828
27	4,61	0,9867	5,05	0,9792
28	7,71	0,9862	7,58	0,9864
29	11,03	0,9774	11,03	0,9774
30	39,41	0,9644	38,08	0,9650
31	5,30	0,9837	5,50	0,9846
32	5,34	0,9836	5,54	0,9846
33	8,47	0,9847	8,44	0,9848
34	3,46	0,9837	3,88	0,9845
35	30,83	0,9714	30,83	0,9714
36	2,25	0,9774	2,01	0,9810
37	4,15	0,9835	4,22	0,9834
38	4,72	0,9851	4,65	0,9852
39	25,13	0,9858	25,13	0,9858
40	30,35	0,9801	31,37	0,9798
41	5,91	0,9862	6,23	0,9861
42	22,54	0,9860	22,48	0,9860
43	5,19	0,9864	5,19	0,9864
44	17,32	0,9848	19,79	0,9854
45	11,03	0,9776	14,18	0,9775
46	2,90	0,9796	2,68	0,9822
47	4,90	0,9782	5,24	0,9777
48	6,40	0,9760	6,18	0,9788
49	38,04	0,9753	38,59	0,9775
50	65,20	0,9824	65,20	0,9824

A seguir, na Figura 62, seguem os dados da tabela anterior na forma de gráfico que permite comparar o NMSE entre a correção apenas da translação e a correção completa.

Figura 62: Gráfico comparativo do NMSE entre a correção apenas da translação e a correção completa

A Figura 63 apresenta o gráfico para a comparação dos valores do índice MSSIM.

Figura 63: Gráfico comparativo do MSSIM entre a correção apenas da translação e a correção completa

O valor médio ($\overline{NMSE}$ e $\overline{MSSIM}$) e o desvio padrão (σ) encontrados para a correção de translação para cada uma das duas medidas considerando os 50 conjuntos são exibidos a seguir:

$$\overline{NMSE}_{Translação} = 0,00127 \quad \sigma = 0,00153$$

$$\overline{MSSIM}_{Translação} = 0,98045 \quad \sigma = 0,00506$$

Para a correção completa, foram encontrados os seguintes valores:

$$\overline{NMSE}_{Completa} = 0,00127 \quad \sigma = 0,00154$$

$$\overline{MSSIM}_{Completa} = 0,98131 \quad \sigma = 0,00460$$

Analisando os valores encontrados e os gráficos da Figura 62 e 63, pode-se notar que o método foi bastante eficiente para a correção da escala, rotação e do *offset* de translação entre os canais das imagens do olho mecânico. O valor médio do *NSME* mostra que o erro entre as imagens foi bastante baixo, estando na faixa de 10^{-4} . Além disso, a medida do *MSSIM* com valor médio próximo de 1, o qual significa máxima similaridade entre as imagens, ou seja, imagens idênticas, mostra que as imagens corrigidas são bastante similares às imagens referência, comprovando a eficiência do método de alinhamento proposto.

Avaliando a correção completa, pode-se verificar que essa apresentou um ganho bastante pequeno em relação à correção apenas da translação. O *NMSE* médio se manteve praticamente idêntico e o *MSSIM* médio melhorou em 0,086%. Além disso, para verificar se a diferença entre a correção completa e a da translação é estatisticamente significativa, levando em consideração os valores de *NMSE* e *MSSIM* encontrados, o teste *t-student* pareado foi aplicado (DEGROOT e SCHERVISH, 2012). Esse teste pode ser utilizado para determinar se dois conjuntos de dados são significativamente diferentes. O método retorna o fator *p* que corresponde à probabilidade de duas médias pertencerem ao mesmo conjunto. Nesse caso, os valores de *p* obtidos ao realizar o teste entre o conjunto da correção da translação e o conjunto da correção completa foram os seguintes:

- Teste *t-student* pareado para os conjuntos *NMSE*: $p = 0,9300$
- Teste *t-student* pareado para os conjuntos *MSSIM*: $p = 0,0015$

Normalmente, o valor considerado como limite para verificar se a diferença medida entre dois conjuntos é estatisticamente significativa é $p < 0,01$, ou seja, a probabilidade dos

dois valores médios encontrados pertencerem ao mesmo conjunto de dados é menor que 1%.

Ao analisar o valor de p encontrado para os conjuntos do NMSE e MSSIM obtidos para a correção apenas da translação e os obtidos para a correção completa, o valor encontrado foi de 0,93 para o NMSE e 0,0015 para o MSSIM. No caso do NMSE, como $p > 0,01$, a diferença não é estatisticamente significativa, uma vez que a probabilidade das duas médias pertencerem ao mesmo conjunto de dados é de 93%. No entanto, para o MSSIM, a probabilidade das duas médias pertencerem ao mesmo conjunto de dados é bastante pequena, o que mostra que a correção completa é estatisticamente significativa, ou seja, melhor que a correção apenas da translação. No entanto, como o ganho da correção completa é bastante pequeno, menor que 0,1%, essa não apresenta vantagem perceptível se comparada com a correção apenas da translação.

Além disso, o custo de processamento para a correção completa aumentou 4 vezes, principalmente devido à transformação dos canais para a forma *log-polar* e a execução do algoritmo de correlação de fase duas vezes, uma para cálculo da escala com rotação e outra para o cálculo da translação.

Quanto ao desempenho, ao executar o método em uma máquina *Intel Core 2 Duo* de 2,2 Ghz com 2 Gb de memória RAM com sistema operacional Windows 7 rodando seus processos padrões, a média de tempo para a correção da translação de um par de imagens foi de 1,2s. Tal fato faz com que o registro seja um processo transparente ao usuário, devido ao seu baixo tempo de processamento, podendo ser utilizado durante o exame para a correção das imagens capturadas.

5. Conclusões

O método proposto neste trabalho se caracteriza como sendo uma ferramenta automática e bastante robusta para a correção dos canais devido à movimentação do olho durante os três *flashes* cromáticos para a obtenção de imagem colorida com sensor monocromático.

Um fator bastante interessante do método é a sua capacidade de realizar o registro de imagens com diferenças significativas de informação (estruturais e de intensidade), como é o caso do canal vermelho em relação azul e verde. Isso ocorre devido à robustez do método de correlação de fase em relação a diferenças de iluminação nas imagens a serem registradas.

Os resultados obtidos nesse trabalho mostraram que a metodologia de registro proposta pode ser utilizada como ferramenta para propiciar a obtenção de imagens coloridas da retina a partir de uma câmera monocromática. Na avaliação qualitativa com as imagens do olho mecânico, os resultados mostraram que, após a correção, as imagens apresentavam-se perfeitamente registradas, não se notando desalinhamento por inspeção visual mesmo quando as imagens foram analisadas de forma ampliada (*zoom*). Para a avaliação quantitativa, o método resultou em baixíssima taxa de erro (avaliada pela NMSE) e alta taxa de similaridade (avaliada pelo MSSIM), o que o caracteriza como bastante confiável para o alinhamento.

Outro fator bastante importante dessa arquitetura que usa sensor monocromático e LEDs cromáticos é o fato de permitir a obtenção, juntamente com a imagem colorida, dos três canais RGB separadamente. Como o exame RF é realizado com a iluminação verde, ele pode ser extraído diretamente do exame colorido, o que possibilita diminuir o tempo do exame e o desgaste do paciente.

Os resultados mostram que o método proposto viabiliza o projeto de utilizar apenas um CCD monocromático para a obtenção de imagens coloridas da retina, uma vez que consegue corrigir o movimento do olho humano entre os *flashes* em menos de 1,2 segundos para a correção da translação, o que é bastante aceitável durante um exame médico desse tipo. No entanto, devido ao aumento no tempo de processamento da correção completa, optou-se apenas pela correção da translação durante o exame, dado que os efeitos da

rotação e escala são muito pequenos ou inexistentes. Nesse caso, se o médico ou operador detectarem que ocorreu a rotação ou escala da imagem, ele pode aplicar uma ferramenta de pós-processamento que executa o método completo. Essa abordagem foi utilizada para que, durante o exame, a imagem fosse retornada o mais rápido possível ao operador, buscando novamente diminuir o tempo de exame e desgaste do paciente.

A arquitetura proposta traz também benefícios para a qualidade da imagem, uma vez que o uso do sensor monocromático possibilita o aumento da resolução espacial em relação a equipamentos com CCDs coloridos que utilizam filtro de Bayer, tanto para exames que utilizam imagens monocromáticas quanto para os que usam imagens coloridas. Outro fator bastante interessante é o fato que, como não utiliza o filtro de Bayer, menos iluminação pode ser utilizada no olho do paciente, uma vez que não tem a atenuação do sinal nos filtros espectrais. Tal fato permite um maior conforto ao paciente durante o exame.

Por fim, há também a possibilidade de uma redução no custo do equipamento, dado que alguns equipamentos que utilizam sensores coloridos, usam também um segundo sensor para a obtenção das imagens monocromáticas, principalmente para a realização do exame ICG, no qual a captura é realizada no infravermelho próximo. Nesse caso, o uso de um único sensor monocromático diminui o preço do equipamento, já que pode ser utilizado para a realização dos quatro exames. Por ser mais barato, o equipamento se torna, consequentemente, mais acessível à comunidade oftalmológica, o que possibilita uma maior detecção e tratamento de doenças com um equipamento de ótima qualidade de imagem.

Apesar do método presente neste trabalho já estar sendo utilizado em equipamentos comerciais para a detecção de doenças de pacientes com problema na retina, sugere-se uma avaliação mais completa utilizando imagens clínicas e não apenas imagens do olho mecânico. Nesse caso, pode-se pensar em testar com dois equipamentos com ópticas idênticas, sendo um projetado para câmera colorida com *flash* branco e outro com câmera monocromática e *flash* RGB. Obtendo as duas imagens, uma de cada equipamento e do mesmo olho do paciente, sob as mesmas condições, pode-se então avaliar a influência do filtro de Bayer na definição final da imagem. No entanto, esta abordagem possui um fator bastante negativo: o custo do projeto e construção de um equipamento com *flash* branco e câmera colorida é bastante alto. Para contornar o problema do custo, outra abordagem pode ser utilizada, levando em consideração a avaliação de oftalmologistas. Nesse caso, as imagens coloridas obtidas com equipamentos com sensores coloridos ou monocromáticos

poderiam ser exibidas aleatoriamente e os médicos escolheriam qual possui a melhor definição. Ambas as abordagens de avaliação ficam como sugestões para trabalhos futuros.

Referências

ALLIANCE, MEDICAL IMAGING & TECHNOLOGY. DICOM Homepage. **Digital Imaging and Communications in Medicine**, 2012. Disponível em: <<http://medical.nema.org/>>. Acesso em: 03 Maio 2012.

BAYER, B. E. **Color Imaging Array**. 3971065, 20 Jul 1976.

BRANDÃO, A. O universo surpreendente das pesquisas sobre a retina. **Pharmacia Brasileira - Oftalmologia**, p. 21-23, Junho 2005.

CARVALHO, E. G. **Desenvolvimento de um Sistema Óptico para Retinografia e Angiografia Digital (Dissertação de Mestrado)**. IFSC-USP. São Carlos. 2006.

CHANWIMALUANG, T. Hybrid retinal image registration using phase correlation. **IEEE Transactions on Information Technology in Biomedicine**, Janeiro 2006. 129-142.

CONDEZ, S. M.; MACKNIK, S. L.; HUBEL, D. H. The role of fixational eye movements in visual perception. **Nature Reviews - Neuroscience**, Março 2004. 229-240.

CSILLAG, A. **Atlas of the Sensory Organs**. 1ª. ed. New Jersey: Humana Press Inc., 2005.

DEGROOT, M. H.; SCHERVISH, M. J. **Probability and Statistics**. 4ª Edição. ed. Boston: Pearson Education, 2012.

DELORI, F. C.; PFLIBSEN, K. P. Spectral reflectance of the human ocular fundus. **Applied Optics**, Boston, v. 28, n. 6, p. 1065-1077, Mar 1989.

EVERDELL, N. L. et al. Multispectral Imaging of the Ocular Fundus using LED Illumination. **Proceedings of SPIE 7371**, Munich, 09 Julho 2009.

GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital de Imagens**. 3ª Edição. ed. São Paulo: Prentice Hall, 2009.

GUTTOSCH, R. J. Investigation of Color Aliasing of High Spatial Frequencies and Edges for Bayer-Pattern Sensors and Foveon X3 Direct Image Sensors. **Foveon Inc.**, Santa Clara, CA , 2005.

LI, X.; GUNTURK, B.; ZHANG, L. Image Demosaicing: A Systematic Survey. **Proceedings of the SPIE**, San Jose - USA, v. 6822, p. 68221J-68221J-15 , janeiro 2008.

MIURA, M. **Imaging Polarimetry of Human Eye (PhD Thesis)**. Sendai: University of Tsukuba, 2009.

MODUGNO, R. G. **Uma Contribuição ao Projeto de Retinógrafos Digitais (Dissertação de Mestrado)**. EESC-USP. São Carlos. 2008.

NEHEMY, M. B. Degeneração macular relacionada à idade: novas perspectivas. **Arquivo Brasileiro de Oftalmologia**, São Paulo, v. 69, n. 6, p. 955-958, Novembro 2006.

NOHDA, M. **Fundus Observation and Photographing Optical System**. US Patent 4322137, 30 mar. 1982.

OCEAN THIN, F. SpectroCam. **SpectroCam**, 2012. Disponível em: <www.spectrocam.com>. Acesso em: 29 Aug 2012.

OCULAR INC. Ocular Imaging Eye Model. **Ocular**, 2010. Disponível em: <<http://www.ocularinc.com/products/educational-aides/ocular-imaging-eye-model.html>>. Acesso em: 12 maio 2012.

OLIVEIRA, T. B. et al. Desenvolvimento e resultados preliminares de um sistema cromático de iluminação para oftalmoscópios indiretos. **Arquivos Brasileiros de Oftalmologia**, São Paulo, v. 72, n. 2, p. 146-151, Março 2009.

OPENCV. **OpenCV Wiki**, 2012. Disponível em: <<http://opencv.willowgarage.com>>. Acesso em: 09 maio 2012.

OXYMAP. Oxymap. **Oxymap**, 2008. Disponível em: <www.oxymap.com>. Acesso em: 04 Maio 2012.

PRATT, W. K. **Digital Image Processing**. Los Altos - California: Wiley, 2007.

RAMELLA-ROMAN, J. C. E. A. Measurement of oxygen saturation in the retina with a spectroscopic sensitive multi aperture camera. **OSA - Optics Express**, Washington, v. 16, p. 170-182, 28 Abril 2008.

REDDY, B. S.; CHATTERJI, B. N. A FFT Based Technique for Translation, Rotation and Scale Invariant Image Registration. **IEEE Transactions on Image Processing**, Kharagpur, v. 5, p. 1266-1271, 8 August 1996.

RYAN, N.; HENEGHAN, C. Image Registration Techniques for Digital Ophthalmic Images. **Irish Signal and Systems Conference**, Galway, Junho 1999.

SCHARDOSIM, L. et al. Motion Detection in Infrared Retinal Image Sequences. **Instrumentation and Measurement Technology Conference 10**, Hangzhou(China), 12 Maio 2011. 1-5.

SMITH, W. J. **Modern Optical Engineering**. 3ª. ed. New York: McGraw-Hill, 2000.

SONY CORPORATION. **HDR-FX1 Alta Definição**, 2012. Disponível em: <<http://www.sony.pt/product/hdd-hdv/hdr-fx1/>>. Acesso em: 01 Setembro 2012.

STUCHI, J. A.; VIEIRA, M. A. C.; RODRIGUES, E. L. L. Registro de Imagens por Correlação de Fase para Geração de Imagens Coloridas em Retinógrafos Digitais Utilizando Câmera CCD Monocromática de Alta Resolução. **VIII Workshop de Visão Computacional**, Goiânia, 27-30 maio 2012.

TANENBAUM, A. S.; WOODHULL, A. S. **Sistemas Operacionais - Projeto de Implementação**. 2a. ed. Porto Alegre: Bookman, 2000.

TOPCON, C. Topcon Global Gateway. **TRC NW300 non-mydratic fundus camera**, 2012. Disponível em: <<http://www.topcon-medical.eu/eu/products/36-trc-nw300.html#specifications>>. Acesso em: 04 maio 2012.

VIEIRA, F. P. **Proposta de inovação no sistema de aquisição de imagens aplicado a retinógrafos digitais (Dissertação de Mestrado)**. EESC-USP. São Carlos. 2013.

WANG, Z. et al. Image Quality Assessment: From Error Visibility to Structural Similarity. **IEEE Transactions on Image Processing**, v. 13, n. 4, p. 600-612, 2004.

WANG, Z.; BOVIK, A. C. Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. **IEEE Signal Processing Mag**, v. 26, n. 1, p. 98-117, 2009.

ZITOVÁ, B.; FLUSSER, J. Image registration methods: a survey. **Image and Vision Computing** **21**, p. 977-1000, 2003.

Apêndice

A seguir é exibido o algoritmo em pseudocódigo utilizado para o alinhamento dos três canais e posterior explicação do funcionamento do mesmo. Como o canal verde é o que apresenta a melhor definição e contraste dos vasos da retina, o mesmo é utilizado como imagem referência, sendo que os outros canais são alinhados em relação a ele.

```

01: Método CorrigirImagem (imagem) retorna imagemCorrigida {
02:     canalR  $\leftarrow$  ExtrairCanalR (imagem)
03:     canalG  $\leftarrow$  ExtrairCanalG (imagem)
04:     canalB  $\leftarrow$  ExtrairCanalB (imagem)
05:     passaBaixa  $\leftarrow$  CriarPassaBaixa (3)
06:     passaAlta  $\leftarrow$  CriarPassaAlta (3)
07:     canalRf  $\leftarrow$  Filtrar (canalR, passaBaixa)
08:     canalGf  $\leftarrow$  Filtrar (canalG, passaBaixa)
09:     canalBf  $\leftarrow$  Filtrar (canalB, passaBaixa)
10:     canalRf  $\leftarrow$  Filtrar (canalRf, passaAlta)
11:     canalGf  $\leftarrow$  Filtrar (canalGf, passaAlta)
12:     canalBf  $\leftarrow$  Filtrar (canalBf, passaAlta)
13:     [escR angR offsetXR offsetYR]  $\leftarrow$  CalcularParâmetros (canalGf, canalRf)
14:     [escB angB offsetXB offsetYB]  $\leftarrow$  CalcularParâmetros (canalGf, canalBf)
15:     canalR  $\leftarrow$  Escalar (canalR, escR)
16:     canalB  $\leftarrow$  Escalar (canalB, escB)
17:     canalR  $\leftarrow$  Rotacionar (canalR, angR)
18:     canalB  $\leftarrow$  Rotacionar (canalB, angB)
19:     canalR  $\leftarrow$  Transladar (canalR, offsetXR, offsetYR)
20:     canalB  $\leftarrow$  Transladar (canalB, offsetXB, offsetYB)
21:     imagemCorrigida  $\leftarrow$  Mesclar (canalR, canalG, canalB)
22:     retornar imagemCorrigida
23: }

24: Método CalcularParâmetros (f1, f2) retorna [escala angulo offsetX offsetY] {
25:     f1r  $\leftarrow$  RecortarCentro (f1r, 1500, 1500)
26:     f2r  $\leftarrow$  RecortarCentro (f2r, 1500, 1500)
27:     F1  $\leftarrow$  DFT (f1r)
28:     F2  $\leftarrow$  DFT (f2r)
29:     MF1  $\leftarrow$  Magnitude(F1)
30:     MF2  $\leftarrow$  Magnitude(F2)
31:     MF1logpol  $\leftarrow$  ConverterLogPolar (MF1)

```

```

32:      MF2logpol ← ConverterLogPolar (MF2)
33:      [logEscala anguloPixels] ← CorrelaçãoFase (MF1logpol, MF2logpol)
34:      escala ← CalculaEscala (logEscala)
35:      angulo ← CalculaAngulo(anguloPixels)
36:      f2c ← Escalar (f2f, escala)
37:      f2c ← Rotacionar (f2c, angulo)
38:      f2r ← RecortarCentro (f2c, 1500, 1500)
39:      [offsetX offsetY] ← CorrelaçãoFase (f1r, f2r)
40:      retornar [escala angulo offsetX offsetY]
41: }

42: Método CorrelaçãoFase (f1, f2) retorna [offsetX offsetY] {
43:      F1 = DFT (f1)
44:      F2 = DFT (f2)
45:      G = F1 * Conjugado (F2) / módulo (F1 * Conjugado(F2))
46:      g = IDFT (G)
47:      xc = Largura (f1) / 2
48:      yc = Altura (f1) / 2
49:      pico = LocalizarMáximo (g)
50:      pico.x = xc – Pico.x
51:      pico.y = yc – Pico.y
52:      retornar [pico.x pico.y]
53: }

```

A base do alinhamento proposto possui os três métodos apresentados. O primeiro método, que objetiva realizar a correção completa da imagem colorida, funciona da seguinte forma:

- Nas linhas 02 a 04, os canais são extraídos da imagem colorida;
- Nas linhas 05 e 06, a máscara para o filtro passa-baixa e passa-alta são criadas, conforme especificado nas equações (30) e (31);
- Da linha 07 a 12, as imagens são filtradas – inicialmente aplica-se o filtro passa-baixa para remover pequenas imperfeições no sistema óptico ou na câmera (*dead pixels*, por exemplo) e, após, aplica-se o filtro passa-alta para ressaltar as estruturas do fundo do olho, o que facilita o alinhamento;
- Nas linhas 13 e 14, os parâmetros para o registro dos canais são calculados pela chamada do método *CalcularParâmetros*, o qual toma o canal verde

como referência. Tal método retorna a escala, ângulo, e *offset* entre os canais;

- Nas linhas 15 e 16, realiza-se a correção de escala do canal vermelho e azul;
- Nas linhas 17 e 18, realize-se a correção de rotação do canal vermelho e azul;
- Finalizando as correções, nas linhas 19 e 20, é realizada a correção de translação do canal vermelho e azul;
- Corrigido o canal vermelho e azul em relação ao verde, os três canais são então combinados na linha 21 e a imagem corrigida é retornada na linha 22.

Já o segundo método, utilizado para calcular os parâmetros de registro (escala, rotação e translação) entre os canais, tem o seguinte funcionamento:

- Nas linhas 25 e 26, um quadrado do centro da imagem de lado 1500 *pixels* é obtido para remover a região da máscara preta, extraíndo, dessa forma, apenas a informação da retina a ser alinhada;
- Calcula-se então a transformada de Fourier dos dois canais, conforme exibido nas linhas 27 e 28. A magnitude do sinal do domínio da frequência é então obtida para os dois espectros nas linhas 29 e 30;
- Nas linhas 31 e 32, as magnitudes são convertidas para a forma *log-polar*, a qual permite obter a escala e o ângulo diretamente usando algoritmo de correlação de fase;
- O algoritmo de correlação de fase é então aplicado na linha 33, o qual retorna o logaritmo da escala entre as duas imagens e o ângulo entre elas representado em pixels e não em graus;
- Na linha 34, o valor da escala é calculado a partir de seu logaritmo. Se a escala existente entre os dois canais for a , o método de correlação de fase retorna $x = \log_b a$. Dessa forma, o método *CalculaEscala* encontra o valor de a conforme o valor de x retornado pelo método de correlação de fase da seguinte forma: $b^x = a$. O valor da base varia com a largura da imagem, conforme explicitado em (REDDY e CHATTERJI, 1996);

- Na linha 35, ocorre a conversão do ângulo em *pixels* para graus através do método *CalculaAngulo*. Esta conversão é bastante simples, sendo executado o seguinte processamento: $\theta_0 = \frac{360^\circ}{largura} \cdot \theta_p$. Para o algoritmo apresentado, $largura = 1500 \text{ pixels}$, θ_p é o ângulo em pixels retornado pelo método de correlação de fase e θ_0 representa o ângulo entre os dois canais em graus;
- A partir do valor da escala e ângulo calculados, o segundo canal é corrigido nas linhas 36 e 37, dado que para calcular a transformação de translação entre eles, deve-se antes corrigir a escala e rotação;
- Depois da correção, novamente o quadrado do centro do segundo canal é recortado na linha 38;
- O método de correlação de fase é então aplicado novamente na linha 39 para o cálculo da translação entre os dois canais, retornando o *offset* horizontal e vertical entre os dois canais;
- Por fim, na linha 40, a escala, ângulo, e *offset* entre os dois canais são retornados.

O terceiro método é que executa o algoritmo de correlação de fase. Ele funciona da seguinte maneira:

- Nas linhas 43 e 44, a transformada de Fourier deslocada para o centro das imagens a serem alinhadas é calculada;
- Na linha 45, calcula-se o produto espectral cruzado, que consiste na multiplicação da imagem F1, no domínio da frequência, pelo conjugado complexo da imagem F2, dividido pelo módulo dessa multiplicação. Essa operação permite isolar a componente da fase do sinal, da qual será calculado o deslocamento;
- Na linha 46, calcula-se a transformada de Fourier inversa do produto espectral cruzado, que caracteriza a matriz de deslocamento espacial;
- Nas linhas 47 e 48, calcula-se o centro da imagem, que será utilizado para o cálculo do deslocamento entre os canais, dado que as transformadas foram calculadas deslocadas para o centro e, nesse ponto, o *offset* entre os dois canais é zero;

- Por fim, na linha 49, procura-se pela localização do ponto de máximo da matriz de deslocamento espacial e, nas linhas 50 e 51, o *offset* é calculado em relação ao centro da imagem, o qual é retornado na linha 52.