

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

JOSEMAR SIMÃO

**MÉTODO ITERATIVO PARA CONSTÂNCIA DE COR EM
VÍDEOS UTILIZANDO CORES IDENTIFICADAS NAS
CENAS**

**VITÓRIA
2016**

JOSEMAR SIMÃO

**MÉTODO ITERATIVO PARA CONSTÂNCIA DE COR EM
VÍDEOS UTILIZANDO CORES IDENTIFICADAS NAS
CENAS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Doutor em Engenharia Elétrica.

Orientadora: Prof.^a Dra. Raquel Frizera Vassallo.

Coorientador: Prof. Dr. rer. nat. Hans Jörg Andreas Schneebeli.

VITÓRIA
2016

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica,
Universidade Federal do Espírito Santo, ES, Brasil)

S588m Simão, Josemar, 1972-
Método iterativo para constância de cor em vídeos utilizando
cores identificadas nas cenas / Josemar Simão. – 2016.
152 f. : il.

Orientador: Raquel Frizera Vassallo.

Coorientador: Hans Jörg Andreas Schneebeili.

Tese (Doutorado em Engenharia Elétrica) – Universidade
Federal do Espírito Santo, Centro Tecnológico.

1. Visão por computador. 2. Processamento de imagens -
Técnicas digitais. 3. Robótica. 4. Filtro temporal. 5. Constância de
cor. I. Vassallo, Raquel Frizera. II. Schneebeili, Hans Jörg
Andreas. III. Universidade Federal do Espírito Santo. Centro
Tecnológico. IV. Título.

CDU: 621.3

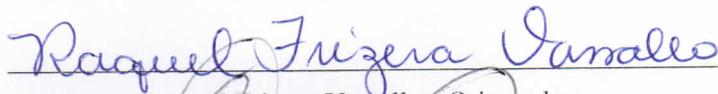
JOSEMAR SIMÃO

**MÉTODO ITERATIVO PARA CONSTÂNCIA DE COR EM
VÍDEOS UTILIZANDO CORES IDENTIFICADAS NAS CENAS**

Tese submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Doutor em Engenharia Elétrica.

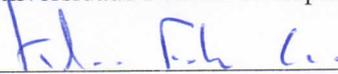
Aprovada em 15 de Agosto de 2016.

COMISSÃO EXAMINADORA

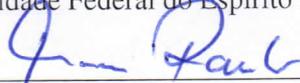


Prof.^a Dra. Raquel Frizera Vassallo - Orientadora
Universidade Federal do Espírito Santo

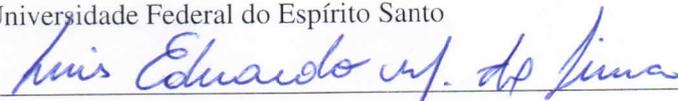
Prof. Dr. rer. nat. Hans Jörg Andreas Schneebeli - Coorientador
Universidade Federal do Espírito Santo



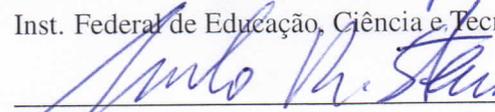
Prof. Dr. Klaus Fabian Côco
Universidade Federal do Espírito Santo



Prof. Dr. Thomas Walter Rauber
Universidade Federal do Espírito Santo



Prof. Dr. Luis Eduardo Martins de Lima
Inst. Federal de Educação, Ciência e Tecnologia do Espírito Santo



Prof. Dr.-Ing. Marcelo Ricardo Stemmer
Universidade Federal de Santa Catarina

*À minha esposa Flávia, ao meu filho André, ao meu pai José Carlos (in memoriam), à
minha mãe Creomar e aos meus irmãos Marcela e José Carlos.*

Agradecimentos

Agradeço a Deus pela paciência, paz e equilíbrio a mim concedidos, sem os quais eu não chegaria ao final dessa jornada.

Agradeço ao Prof. Hans e a Prof.^a Raquel por compartilharem seus conhecimentos e pelas orientações e sugestões que contribuíram para a conclusão deste trabalho.

Também agradeço ao Instituto Federal do Espírito Santo pela licença concedida para dedicação ao curso.

Sumário

1	Introdução	18
1.1	Motivação	18
1.1.1	Aplicações sujeitas à correção	23
1.1.2	Importância da constância de cor	24
1.2	Definição do problema	26
1.3	Metodologia proposta	27
1.4	Estrutura desta tese	29
2	Constância de Cor	30
2.1	Modelagem do sistema de visão	30
2.2	A constância de cor na visão humana	32
2.3	A constância de cor na visão computacional	35
2.3.1	Constância de cor em imagem única	36
2.3.2	Constância de cor em sequências de imagens	40
2.3.3	Abordagem proposta	41
2.3.4	Trabalhos desenvolvidos nesta tese	44
2.3.5	Taxonomia das abordagens para constância de cor	46
2.4	Conclusão	47

3	Métodos Básicos Desenvolvidos para Sequências de Imagens	50
3.1	Introdução	50
3.1.1	A transformação de cores	51
3.1.2	O processo iterativo	53
3.1.3	Sequências de imagens para testes	56
3.2	Método usando Amostras de referência da Imagem Anterior (MAIA)	59
3.3	Método do Filtro Temporal para Amostras de referência (MFTA)	67
3.4	Aplicação do MFTA em sequência de imagens sem referências fixas de cores	73
3.5	Conclusão	77
4	Método do Filtro Temporal para Transformações (MFTT)	78
4.1	Introdução	78
4.2	Melhoria da qualidade das cores de referência	79
4.2.1	Avaliação da filtragem com mediana	80
4.2.2	Avaliação da máscara binária	86
4.3	Uma abordagem baseada em filtragem temporal	92
4.4	O MFTT	93
4.5	Conclusão	96
5	Experimentos Realizados	98
5.1	Ferramentas de programação	98
5.2	Experimentos de segmentação da imagem	99
5.2.1	Segmentação utilizando crescimento de região	99
5.2.2	Segmentação utilizando <i>graph-based image segmentation algorithm</i>	103
5.2.3	Segmentação utilizando quadrados de tamanho fixo	105
5.3	Experimentos de comparação do MFTA	106

5.3.1	Comparação qualitativa	108
5.3.2	Comparação quantitativa	115
5.4	Experimentos com o MFTT	117
5.5	Conclusão	123
6	Conclusões e Trabalhos Futuros	125
A	Espaços de Cores	140
B	Modelo dos Iluminantes	143
C	Modelo das Refletâncias	147
D	Modelo dos Sensores	150

Lista de Tabelas

2.1	Informações de entrada e informações adicionais em algoritmos de constância de cor.	48
2.2	Restrições e a forma de correção de cores nos algoritmos de constância de cor.	49
3.1	Valores RGB médios da placa 19 (branco) nos quatro ambientes.	58
3.2	Erros normalizados para a sequência sem movimentos relativos utilizando áreas de 3x3, 7x7 e 13x13 <i>pixels</i>	71
4.1	Exemplo do matiz ordenado para definição da cor mediana	81
4.2	Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 700 para os critérios (a), (b), (c) e (d)	84
4.3	Valores do módulo do erro entre a cor média de cada região do quadro 700 e a cor média do primeiro quadro	85
4.4	Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 800 para os critérios (a), (b), (c) e (d)	85
4.5	Valores do módulo do erro entre a cor média de cada região do quadro 800 e a cor média do primeiro quadro	85
4.6	Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 900 para os critérios (a), (b), (c) e (d)	86
4.7	Valores do módulo do erro entre a cor média de cada região do quadro 900 e a cor média do primeiro quadro	86
4.8	Valores RGB e HSV médios da região 5x10 <i>pixels</i> da imagem corrigidas do quadro 800 para os dois critérios utilizados	88

4.9	Valores RGB e HSV médios da região 5x10 <i>pixels</i> da imagem corrigidas do quadro 900 para os dois critérios utilizados	89
4.10	Frequência dos desvios entre os pontos localizados pelo <i>tracking</i> e os pontos localizados pelo SURF	91
4.11	Valores RGB e HSV médios da região 5x10 <i>pixels</i> do quadro um capturado e do quadro 700 corrigido	96
4.12	Valores RGB e HSV médios da região 5x10 <i>pixels</i> do quadro 800 e 900 corrigidos	96
5.1	Erro normalizado para comparação dos algoritmos	116

Lista de Figuras

1.1	Curva de energia das radiações eletromagnéticas representando uma cor. . .	18
1.2	Formação da cor para diferentes fontes de iluminação.	19
1.3	Curvas de sensibilidades da câmera Canon 400D para os canais R, G e B. .	20
1.4	Dois imagens RGB, I_a e I'_a , sob iluminações distintas.	23
1.5	Caminho de um robô móvel.	26
1.6	Dois quadro consecutivos e as regiões coloridas identificadas.	28
2.1	Curva espectral de absorção dos três tipos de cones (azul, verde e vermelho) e dos bastonetes.	31
2.2	Curva espectral dos filtros colocados em frente aos projetores para iluminar as superfícies coloridas.	33
2.3	Cada conjunto de cores é iluminado por um grupo de três projetores. . . .	34
2.4	Um longo tubo com abertura ajustável.	35
2.5	Cartão do tipo <i>ColorChecker</i> com 24 placas coloridas.	36
2.6	Curvas de sensibilidade dos sensores das câmeras Sony DXC-930 e Kodak DCS-460.	37
2.7	Variação do matiz das cores nas imagens capturadas.	44
2.8	Variação do matiz das cores nas imagens corrigidas por cada algoritmo. . .	45
3.1	Primeiro quadro da primeira sequência de imagens.	57
3.2	<i>ColorChecker</i> com as regiões quadradas para cálculo da cor média.	57

3.3	Primeiro quadro da segunda sequência de imagens.	58
3.4	Esquema do processo iterativo presente no MAIA.	59
3.5	Triângulo de cromaticidade no plano unitário.	62
3.6	Cores normalizadas plotadas no triângulo de cromaticidade.	62
3.7	Matiz das cores 01, 03, 06, 08 e 12 provenientes das imagens capturadas.	63
3.8	Matiz corrigido com referências fixas do primeiro quadro (<i>ground truth</i>).	64
3.9	Matiz corrigido pelo MAIA para área 3x3 <i>pixels</i>	65
3.10	Matiz corrigido pelo MAIA para áreas (a) 7x7 e (b) 13x13 <i>pixels</i>	66
3.11	Esquema do processo iterativo presente no MFTA.	68
3.12	Resultado do Método Simples (a) (<i>ground truth</i>) e do MFTA (b) para uma área de 3x3 <i>pixels</i>	69
3.13	Resultado do Método Simples (a) (<i>ground truth</i>) e do MFTA (b) para uma área de 7x7 <i>pixels</i>	69
3.14	Resultado do Método Simples (a) (<i>ground truth</i>) e do MFTA (b) para uma área de 13x13 <i>pixels</i>	70
3.15	Resultado do MFTA para uma região de 1 <i>pixel</i> em 10 repetições.	71
3.16	Resultado do matiz das cores corrigidas do quadro 1517 em 100 repetições encadeadas para uma região de 1 <i>pixel</i>	72
3.17	Variação (%) do matiz da cor 08 entre duas repetições consecutivas do quadro 1517.	72
3.18	Primeiro quadro da sequência de imagens e a imagem segmentada com quadrados de 5x5 <i>pixels</i>	74
3.19	Quadro inicial e quadro 700 da sequência corrigidos pelo MFTA.	76
4.1	Região colorida com uma área de 5x5 <i>pixels</i>	80
4.2	Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA com mediana ordenada pelo matiz.	82

4.3	Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA com mediana ordenada pelos canais R, G e B.	82
4.4	Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA em que a segmentação utiliza a mediana ao invés da média aritmética na composição do critério de uniformidade.	83
4.5	Região de 5x10 <i>pixels</i> separada do primeiro quadro.	83
4.6	Região de 5x10 <i>pixels</i> para comparação dos critérios: (a) segmentação e valor da amostra pela média aritmética; (b) segmentação pela média e valor pela mediana do matiz; (c) segmentação pela média e valor pela mediana dos canais R, G e B; (d) segmentação e valor pela mediana dos canais R, G e B.	84
4.7	Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA utilizando uma máscara binária definida pelo brilho dos <i>pixels</i> de cada região colorida.	87
4.8	Quadros 1 e 700 da segunda sequência corrigidos pelo MFTA utilizando uma máscara binária definida pelos três canais R, G e B.	88
4.9	Localização de pontos característicos em duas imagens consecutivas da sequência.	90
4.10	Primeiro e último quadros da sequência corrigidos pelo MFTT para w=30.	96
5.1	Resultado da segmentação do primeiro quadro por crescimento de região (utiliza a implementação em C++ chamada <i>Fast 3D/2D Region Growing (MEX)</i>).	101
5.2	Resultado da segmentação do segundo quadro por crescimento de região (utiliza a implementação em C++ chamada <i>Fast 3D/2D Region Growing (MEX)</i>).	101
5.3	Resultado da segmentação do terceiro quadro por crescimento de região (utiliza a implementação em C++ chamada <i>Fast 3D/2D Region Growing (MEX)</i>).	102
5.4	Resultado da segmentação por <i>graph-based</i> do primeiro quadro.	104
5.5	Resultado da segmentação por <i>graph-based</i> do segundo quadro.	104
5.6	Resultado da segmentação por <i>graph-based</i> do terceiro quadro.	105

5.7	Seis sequências (1-6) para análise. As marcas em cruz vermelha, verde e azul representam os pontos fixos escolhidos para análise do matiz.	107
5.8	Variação do matiz das imagens capturadas. (a)-(f) Sequência de 1 a 6. . . .	109
5.9	Variação do matiz das imagens corrigidas da Sequência 1 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	110
5.10	Variação do matiz das imagens corrigidas da Sequência 2 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	111
5.11	Variação do matiz das imagens corrigidas da Sequência 3 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	112
5.12	Variação do matiz das imagens corrigidas da Sequência 4 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	113
5.13	Variação do matiz das imagens corrigidas da Sequência 5 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	114
5.14	Variação do matiz das imagens corrigidas da Sequência 6 pelo (a) MFTA, (b) <i>Gray-World</i> , (c) <i>Max-RGB</i> e (d) <i>Gray-Edge</i>	115
5.15	Primeiro e último quadros da sequência com movimento e variação de iluminação.	118
5.16	Pontos escolhidos para avaliação das cores.	119
5.17	Variação do matiz original proveniente das imagens capturadas para os três pontos escolhidos.	120
5.18	Variação do matiz das imagens corrigidas para (a) $w = 15$, (b) $w = 30$ quadros. 120	
5.19	Variação do matiz das imagens corrigidas para (a) $w = 150$, (b) $w = 450$ quadros.	121
5.20	Variação do matiz das imagens corrigidas para (a) $w = 900$ e (b) $w = 1800$ quadros.	121
5.21	Variação da intensidade das cores proveniente das imagens capturadas para os três pontos escolhidos.	122
5.22	(a) Correção pelo método <i>Max-RGB</i> e (b) correção pelo método <i>Gray-Edge</i> . 123	
5.23	Correção pelo método <i>Gray-World</i>	123

A.1	Curvas de correspondência de cor para o sistema XYZ.	140
B.1	Curvas espectrais da irradiância solar na faixa visível obtidos por Neckel & Labs e por SOLSPEC/ATLAS 1.	143
B.2	Distribuição espectral relativa de alguns iluminantes padronizados.	144
B.3	Funções base para caracterização dos iluminantes da série D.	145
C.1	Funções espectrais de refletância de superfícies do <i>Colorchecker</i>	148
D.1	Funções de sensibilidade espectral das câmeras CANON 5D e CANON 10D.	151
D.2	Funções de sensibilidade espectral das câmeras KODAK DCS 420 e KODAK DCS 460.	151
D.3	Funções de sensibilidade espectral das câmeras NIKON D1X e NIKON D70.	152
D.4	Funções de sensibilidade espectral das câmeras NIKON D1X e NIKON D70.	152

Resumo

Os algoritmos de constância de cor buscam corrigir as cores das imagens capturadas sob iluminação desconhecida para apresentá-las como se tivessem sido capturadas sob uma iluminação conhecida. Quando se trabalha com sequências de imagens, há variação de iluminação entre elas, o que dificulta a aplicação de diversos algoritmos de visão computacional. Corrigir as imagens e apresentá-las como se estivessem sob a mesma iluminação, simplifica consideravelmente a aplicação de diversos destes algoritmos.

Neste trabalho, as cores de uma sequência de imagens são corrigidas individualmente para que estas tenham a iluminação de um quadro de referência, em geral, o primeiro quadro. Para isto, é utilizada uma transformação linear geral com nove parâmetros, chamada aqui de matriz de mapeamento de cores. Um processo de otimização, por exemplo, mínimos quadrados, é necessário para a obtenção da matriz de mapeamento de cores quando o número de cores de referência for maior que o número de canais de cor. Isto é alcançado pelo cálculo da matriz pseudo-inversa.

Um processo iterativo pode usar as cores de uma imagem para corrigir as cores da imagem seguinte, utilizando regiões comuns a ambas imagens. Para isso, um conjunto de cores da imagem anterior compõe as amostras de referência, e as cores da imagem atual, provenientes de regiões correspondentes às da imagem anterior, compõem as amostras capturadas. Estas amostras são utilizadas para obtenção da matriz de mapeamento de cores. Como o campo visual varia ao longo da sequência, as regiões utilizadas no cálculo da correção devem ser ajustadas a cada iteração.

Este processo iterativo está presente em três métodos de correção de cores desenvolvidos nessa tese. O primeiro método, usado como base, e denominado **Método usando Amostras de Referência da Imagem Anterior (MAIA)**, considera apenas sequências de imagens com variação de iluminação e sem movimento relativo entre a câmera e a cena. Um filtro temporal que utiliza um conjunto de imagens anteriores para produzir amostras de referência mais estáveis é aplicado para atenuar os efeitos do ruído nas imagens, dando origem ao **Método do Filtro Temporal para Amostras de Referência (MFTA)**. Este método, acoplado a um esquema de *tracking* de regiões coloridas, permite a correção de cores em sequências de imagens com movimento relativo entre a câmera e a cena. No entanto, o método ainda apresenta uma sensibilidade grande aos ruídos levando a uma degradação da qualidade das imagens corrigidas. Para contornar este problema, as amostras de referência e as amostras capturadas de um conjunto de iterações anteriores são utilizadas em um filtro temporal caracterizando o terceiro método chamado **Método do Filtro Temporal para Transformações (MFTT)**, que apresenta resultados que permitem o seu uso em diversas áreas, incluindo Robótica Móvel.

Abstract

Color constancy algorithms try to correct the colors of images captured under unknown lighting and present them as if they had been captured under a known lighting. When working with images sequences, there are lighting variations between the images, which hinders the application of various computer vision algorithms. Correcting the images and presenting them as if they were taken under the same lighting enables the application of various of these algorithms.

In this work, the colors of images in a sequence are corrected individually so that they have the same lighting of a reference frame, in general, the first frame. To do that, general linear transformation with nine parameters is used, called here color mapping matrix. An optimization process, for example, least squares method, is necessary to obtain the color mapping matrix when the number of reference colors is higher than the number of color channels. This is achieved through the calculation of the pseudo-inverse matrix.

An iterative process can use the colors of an image to correct the color of the next image, using regions common to both images. A set of colors from the former image composes the reference sample, and the colors of the next image, from regions corresponding to the those of the reference image, compose the captured sample. These samples are used to obtain the color mapping matrix. As the visual field varies throughout the sequence, the regions common to images in the sequence must be adjusted for each iteration.

Three color correction methods based on this approach are shown in this thesis. The first method, the reference method, is called **M**ethod using **R**eference **S**amples from the **P**revious **I**mage (MSPI), and can only be applied to image sequences with lighting variation and no relative movement between the camera and the scene, i.e., no need to adjust the regions. A temporal filter that uses a set of previous images to produce more stable reference samples is applied to attenuate the noise effects on images, giving origin to the **M**ethod of the **T**emporal **F**ilter for **R**eference **S**amples (MTFS). This method, coupled with a scheme for tracking the regions with identified colors, allows color correction in sequences of images with relative movement between the camera and the scene. However, the method still presents a high sensibility to the image noise leading to a degradation of the corrected images. Using reference samples and the captured samples of a set of previous iterations and a temporal filter, the third method called **M**ethod of the **T**emporal **F**ilter for **T**ransformations (MTFT), presents results that allow its use in various areas, including Mobile Robotics.

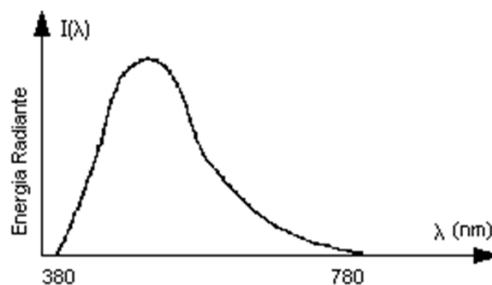
Capítulo 1

Introdução

1.1 Motivação

A cor pode ser definida de forma objetiva como sendo um conjunto de radiações eletromagnéticas representado por uma curva de distribuição espectral contínua (Figura 1.1). Essa forma de representação está associada a um espaço de representação cor de dimensão infinita chamado espaço espectral de cor.

Figura 1.1: Curva de energia das radiações eletromagnéticas representando uma cor.

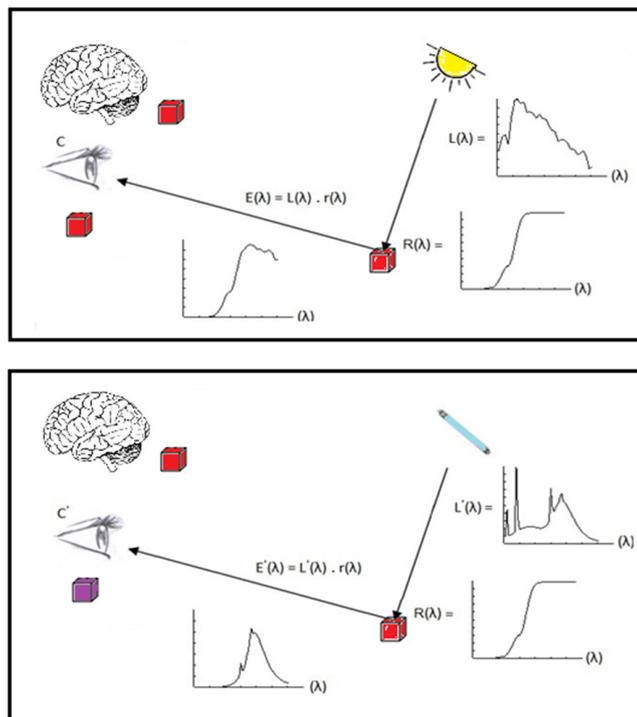


Fonte: Elaborada pelo autor

Quando a luz refletida pela superfície de um objeto alcança os olhos, estímulos são gerados e enviados ao cérebro. A interpretação desses estímulos pelo cérebro define a cor percebida pela visão humana. Estes estímulos são resultante da interação entre refletância dos objetos, o sensor (retina dos olhos) e a fonte de iluminação (também conhecida como o iluminante da cena). Sob variação da composição espectral da iluminação que incidente sobre um objeto, a luz refletida por ele também se altera (Kawakami et al., 2005) provocando variação dos estímulos que chegam ao cérebro, mas a cor percebida ainda é a mesma.

Na Figura 1.2, um objeto de refletância constante, $R(\lambda)$, é iluminado por duas fontes de iluminação diferentes: o Sol e uma lâmpada fluorescente. Estes dois iluminantes são representados pelas distribuições espectrais de radiação eletromagnética, $L(\lambda)$ e $L'(\lambda)$, respectivamente, onde λ é o comprimento de onda. As radiações refletidas por este objeto e que alcançam os olhos, $E(\lambda)$ e $E'(\lambda)$, são diferentes. No entanto, no cérebro a sensação de cor é a mesma, ou seja, a cor C é percebida como sendo idêntica à cor C' .

Figura 1.2: Formação da cor para diferentes fontes de iluminação.



Fonte: Elaborada pelo autor

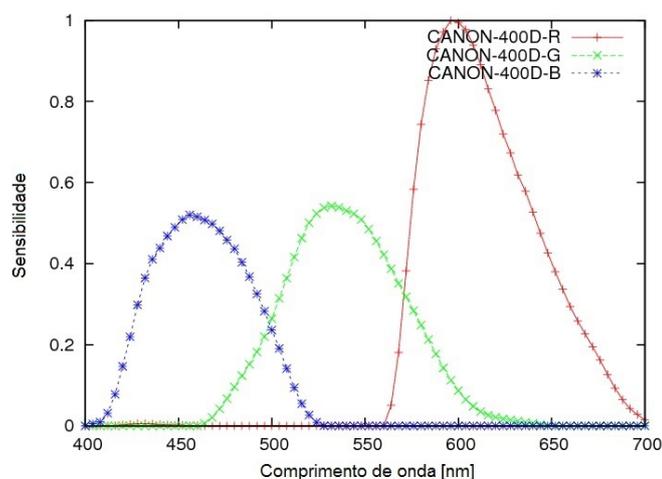
Dentro de certos limites, o sistema visual humano possui uma extraordinária habilidade para identificar as cores como sendo constantes mesmo quando há variação na composição espectral da fonte de iluminação. Esta habilidade é chamada de constância de cor (Olkkonen et al., 2010). A percepção de constância nas cores dos objetos permite que o cérebro humano faça uma associação entre as cores observadas e os objetos que lhe deram origem. Esta associação é tão significativa que, por vezes, a cor observada é atribuída ao objeto como se fosse uma propriedade intrínseca deste. Dessa forma, a percepção da cor torna-se uma interpretação psicofísica da refletância da superfície dos objetos. A expressão 'cor do objeto' será utilizada nesta tese com este sentido. Todo este processamento envolve a retina e o córtex (Land e Mccann, 1971). No entanto, a forma como a mente humana implementa a constância de cor ainda não é completamente entendida (Gijsenij et al., 2011).

A percepção das cores pelo olho humano se dá por uma amostragem em três regiões distintas na faixa visível do espectro eletromagnético. Esta amostragem é feita por três tipos de receptores, sensíveis à luz, existentes na retina (Ebner, 2007). Estes receptores absorvem a luz nas bandas curta, média e longa de comprimentos de onda da luz visível e a resposta destes varia com a alteração do espectro da fonte de luz (Bianco e Schettini, 2011; Gu et al., 2013). O espaço de representação da cor torna-se, então, um espaço tridimensional também chamado de espaço de cor tricromático.

Com o objetivo de representar quantitativamente as cores identificadas por humanos, a *Commission Internationale de l'Éclairage* (CIE) desenvolveu um espaço de cores que pudesse conter toda diversidade de cores observáveis pela visão humana. Este espaço, chamado de espaço XYZ, foi definido a partir das funções de correspondência de cores obtidas para um grupo de indivíduos humanos e representa a resposta de um observador padrão. No entanto, este espaço de cor não é o mesmo implementado nos dispositivos de captura de imagens.

Os dispositivos de captura de imagens coloridas, tais como câmeras digitais e *scanners*, possuem sensores baseados na fisiologia da visão humana e as cores capturadas por eles normalmente são representadas utilizando-se o modelo tricromático, cujos canais são chamados *Red*, *Green* e *Blue*, RGB. Estes canais representam a região de sensibilidade espectral dos sensores utilizados nestes dispositivos. A Figura 1.3 mostra as curvas¹ de sensibilidade espectral dos canais R, G e B da câmera Canon 400D (Kawakami et al., 2013).

Figura 1.3: Curvas de sensibilidades da câmera Canon 400D para os canais R, G e B.



Fonte: (Kawakami et al., 2013)

¹A base de dados numérica e as imagens das curvas podem ser obtidas no endereço eletrônico <http://www.cvl.iis.u-tokyo.ac.jp/~rei/research/cs/zao/database.html>

Para compatibilizar o espaço de cores dos dispositivos de captura de imagem com o espaço XYZ foi definido um espaço chamado *standard RGB* (*sRGB*). Este espaço é um subconjunto do espaço XYZ. Isto significa que todas as cores representadas no espaço *sRGB* possuem representação também no espaço XYZ, mas a recíproca não é verdadeira. As Equações 1.1 e 1.2 permitem converter os valores do espaço *sRGB* para o espaço XYZ e vice-versa.

$$\begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} = \begin{bmatrix} 3,2406 & -1,5372 & -0,4986 \\ -0,9689 & 1,8758 & 0,0415 \\ 0,0557 & -0,2040 & 1,0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1.1)$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,4124 & 0,3576 & 0,1805 \\ 0,2126 & 0,7152 & 0,0722 \\ 0,0193 & 0,1192 & 0,9505 \end{bmatrix} \begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} \quad (1.2)$$

No entanto, nem todos os dispositivos são fabricados com o espaço *sRGB*. Estes dispositivos de captura de imagens possuem um espaço que é dependente do dispositivo e a conversão dos valores de cor gerados por eles para o espaço XYZ depende da determinação de uma matriz de conversão específica para cada dispositivo. A obtenção dessa matriz é um processo de calibragem que deve ser executado para cada dispositivo de captura de imagens a ser utilizado em um sistema de visão computacional.

Espaços de cores que separam as informações de luminância e cromaticidade apresentam uma representação mais intuitiva e adequada à interpretação humana. Os espaços HSI e HSV são dois exemplos bem comuns na literatura. Nestes espaços, o canal H, chamado de matiz, está associado à componente espectral predominante e o canal S, chamado de saturação, indica a proporção de branco associado à cor. Os canais I e V, referentes aos espaços HSI e HSV, respectivamente, estão associados à representação da luminância. Os valores destes canais podem ser obtidos a partir dos valores RGB. As funções de correspondência de cores do espaço XYZ e as equações que convertem valores RGB em HSI e HSV são mostradas no Apêndice A.

As respostas dos sensores dos dispositivos de captura de imagens não caracterizam diretamente a refletância de uma superfície, mas elas são o resultado da medição da luz refletida por ela. Esta luz refletida, como dito anteriormente, é produto da refletância do objeto e a luz que incide sobre a superfície. Assim, os valores medidos pelos sensores estão sujeitos às variações da iluminação e uma mesma superfície iluminada por diferentes iluminantes terá diferentes representações de cor no espaço RGB. Estes valores são definidos pela Equação 1.3.

$$c_m = \int_{\lambda_1}^{\lambda_2} S_m(\lambda)L(\lambda)R(\lambda)d\lambda \quad (1.3)$$

Nesta equação, $[\lambda_1, \lambda_2]$ é a faixa de comprimento de onda do espectro visível, $L(\lambda)$ é o espectro de radiação do iluminante da cena, $R(\lambda)$ é a curva espectral de refletância da superfície observada e $S_m(\lambda)$ é a curva de sensibilidade do sensor para $m \in \{r, g, b\}$.

A variação das cores sob condições variáveis de iluminação é um efeito indesejável que prejudica o desempenho dos algoritmos de visão computacional tais como segmentação, rastreamento e reconhecimento de objetos quando aplicados sobre imagens estáticas (Kamijo et al., 2002; Tu, 2009). Em sequências de imagens, onde há variação temporal da iluminação os resultados são ainda piores.

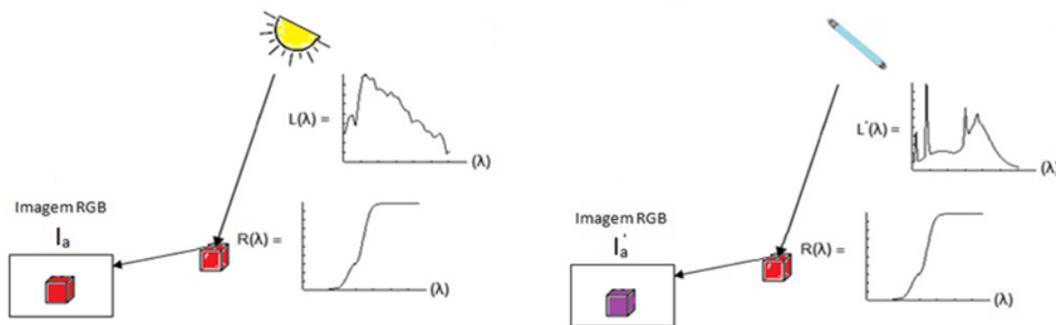
Em geral, os sensores das câmeras digitais não implementam nenhum processo de compensação da iluminação. Por conta disso, sistemas de visão computacional que utilizam cor normalmente impõem restrições quanto à iluminação da cena que está sendo avaliada. Estas restrições têm o objetivo de minimizar a variação das cores em tais sistemas. Sistemas que não podem impor estas exigências em relação à fonte de luz, como aqueles que operam em ambientes externos, devem implementar uma compensação para a variação da fonte de luz. Este é o caso das câmeras fotográficas digitais que aplicam um pré-processamento nas imagens capturadas antes de armazená-las, buscando a constância de cor.

Para se alcançar a constância de cor nas imagens RGB duas alternativas são comumente utilizadas. Uma é a determinação da curva espectral da refletância das superfícies $R(\lambda)$ como função dos valores RGB, das curvas espectrais dos sensores e da distribuição espectral dos iluminantes, ou seja $R(\lambda) = \Phi(c_m, S_m(\lambda), L(\lambda))$. Uma vez que as refletâncias tenham sido determinadas uma imagem corrigida pode ser gerada com valores RGB calculados pela Equação 1.3 para um dado iluminante escolhido como referência. A dificuldade dessa estratégia está justamente na obtenção das curvas de sensibilidade dos sensores e da distribuição espectral do iluminante, que normalmente são desconhecidas.

A segunda alternativa está baseada na definição de uma transformação que leva os valores RGB de uma imagem capturada sob uma iluminação desconhecida para valores RGB de uma imagem como tendo sido capturada sob uma iluminação padrão. Assim, para as imagens I_a e I'_a capturadas sob iluminação desconhecida (Figura 1.4), duas imagens corrigidas seriam geradas $I_c = T(I_a)$ e $I'_c = T'(I'_a)$ por duas transformações distintas T e T' . As imagens corrigidas I_c e I'_c seriam, então, idênticas. Esta estratégia utiliza os valores RGB da imagem capturada como fonte de informação para, juntamente com algumas suposições a respeito das refletâncias, dos iluminantes e dos sensores, obter as transformações de correção. Nesta

estratégia, a dificuldade aparece quando as suposições são violadas restringindo, assim, o escopo de aplicação.

Figura 1.4: Duas imagens RGB, I_a e I'_a , sob iluminações distintas.



Fonte: Elaborada pelo autor

1.1.1 Aplicações sujeitas à correção

A facilidade de aquisição de equipamentos a preços relativamente baixos e o grande volume de informação útil presente em imagens têm despertado o interesse na utilização de sistemas de visão computacional para a realização de diversas tarefas (Gonzalez e Woods, 2008). Informações sobre formas, tamanhos e texturas de objetos e o uso de métodos de visão computacional têm permitido o desenvolvimento de técnicas para aplicação em tarefas de controle, de inspeção, de monitoramento, de navegação e outras. Estas técnicas proporcionam um aumento da precisão, da produtividade e da segurança na realização de diversos tipos de tarefas (Humod et al., 2015). Um exemplo é um sistema de classificação de mamões (Simão, 2003) que avalia o grau de maturação dos frutos pela cor da casca.

Sistemas de visão computacional que tiram proveito da informação de cor podem aprimorar os métodos de segmentação, rastreamento e reconhecimento de objetos desde que encontrem condições de iluminação favoráveis. Muitas tarefas podem se tornar bastantes simples se a cor dos objetos for conhecida com boa exatidão. Reconhecer objetos e destacá-los de uma cena pode resumir-se a uma simples tarefa de comparação com base somente em suas cores (Ebner, 2007).

Todavia, a maioria dessas aplicações alcançam bons resultados quando a iluminação do ambiente é controlada, uma vez que qualquer variação na iluminação interfere diretamente na qualidade das cores utilizadas no processamento (Demaagd et al., 2012). Aplicações em

que o controle da iluminação não é possível normalmente apresentam uma abordagem simplificada, fazendo uso de imagens em tons de cinza (Kunimitsu et al., 2003) ou implementam métodos que sejam invariantes com a iluminação (Upcroft et al., 2014; Kamijo et al., 2002). Aplicações para ambientes externos ou aquelas que se desenvolvem em diferentes ambientes com diferentes fontes de iluminação têm esta variação de iluminação como o maior obstáculo. Assim, métodos que atenuem os efeitos da variação da iluminação produzindo cores mais estáveis são necessários para estas aplicações.

Muitos métodos de constância de cor têm sido propostos na literatura (Gijssenij et al., 2011) para serem utilizados em aplicações que envolvam a captura de imagens estáticas, como, por exemplo, a fotografia. Para a captura de fotos, algumas câmeras permitem um ajuste conhecido na fotografia como balanço de branco. Este ajuste aplica um ganho diferente para cada canal de cor, R, G e B, da imagem capturada. Os valores dos ganhos são pré-definidos conforme o tipo de iluminação do ambiente (luz do dia, nublado, pôr do sol, fluorescente e incandescente) para um ajuste manual. O recurso *Automatic White Balance* (AWB) busca estimar, por alguma heurística, o tipo de iluminação da cena para automaticamente escolher os ganhos a serem aplicados.

Embora o sistema visual humano não garanta uma constância de cor perfeita (Brainard e Wandell, 1986), é possível que ele seja bom o suficiente para que a divergência de cor existente em uma foto, após a correção por algum método de constância de cor, não seja percebida quando se a vê. No entanto, é nos sistemas de visão computacional que a efetividade dos métodos de constância de cor é colocada à prova. Em uma tarefa de reconhecimento de objetos com base em histogramas de cor, o uso dos algoritmos de constância de cor bem conhecidos, tais como: *Gray-World* (Buchsbbaum, 1980), *White-Patch Retinex* ou *Max-RGB* (Land, 1977), *Gamut-Mapping* (Forsyth, 1990) e aqueles baseados em redes neurais, apresentou uma melhora significativa nos resultados em relação aos testes em que eles não foram utilizados. No entanto, estes bons resultados não foram precisos o suficiente para que o desempenho de indexação de cor fosse próximo daquele obtido quando não há alteração na iluminação do ambiente (Funt et al., 1998). Além disso, um outro estudo mostra que a dependência destes algoritmos em relação às suposições feitas têm restringido o uso em aplicações práticas de rastreamento em vídeo (Agarwal et al., 2006a).

1.1.2 Importância da constância de cor

A constância de cor é um problema mal posto (Vazquez-Corral et al., 2012), uma vez que tanto a distribuição espectral do iluminante como as refletâncias da cena são desconhecidas. Nesta situação é impossível determinar, por exemplo, se a diferença de cor entre duas regiões

de uma imagem é devido à mudança das superfícies ou da luz incidente. Este problema não pode ser resolvido sem suposições adicionais (Li et al., 2013) e as mais comuns são:

- As refletâncias são foscas e a radiância que emerge das superfícies não possui componentes especular.
- A iluminação possui uma diversidade espectral limitada.
- A iluminação é uniforme sobre toda a imagem.
- Os ângulos de incidência e reflexão da luz não alteram sua intensidade nem sua composição espectral.
- As curvas de sensibilidade dos sensores são banda estreita e sem sobreposição.

Entretanto, estas suposições restringem as possibilidades de aplicações dos algoritmos de constância de cor, pois o desempenho deles é significativamente deteriorado quando estas suposições são violadas (Celebi e Bogdan, 2014).

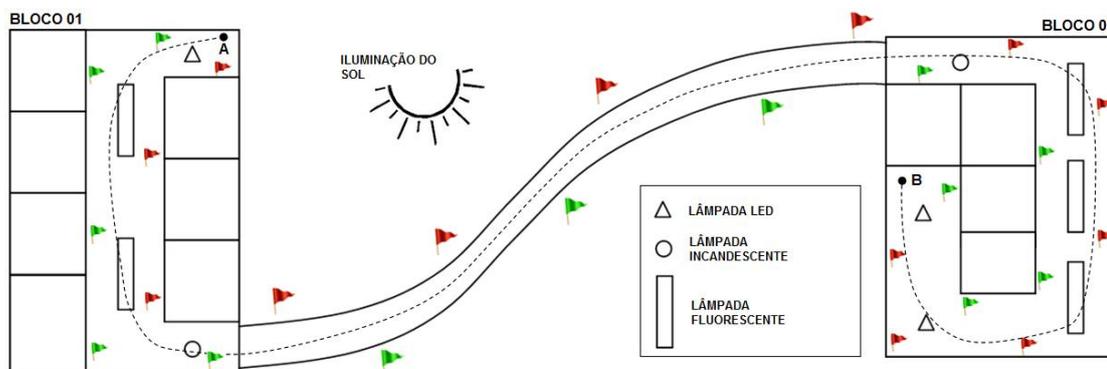
Isto tem motivado diversas pesquisas em busca de uma solução para o problema da constância de cor aplicada em sistemas de visão computacional. Muitas propostas têm sido apresentadas e uma revisão das principais abordagens presentes na literatura é apresentada por Gijsenij et al. (2011). Contudo, a constância de cor ainda é um problema sem solução geral, e obter imagens com cores invariantes sob mudanças de iluminação ainda é um problema aberto em visão computacional (Vazquez-Corral et al., 2012).

Os algoritmos comumente apresentados na literatura buscam alcançar a constância de cor para imagens estáticas. Quando sequências de imagens ou vídeos devem ser corrigidos, estes algoritmos são aplicados individualmente em cada quadro da sequência (Sridharan e Stone, 2004; Renno et al., 2005). No entanto, uma abordagem pouco explorada é aquela na qual se tira proveito da informação correlata existente em quadros consecutivos de um vídeo.

Uma aplicação que utiliza uma sequência de imagens sujeitas à variação de iluminação é a navegação de um robô móvel através de um caminho definido por marcas coloridas. Esta navegação é análoga àquela utilizada por embarcações durante a entrada e a saída dos portos, canais e rios. Nestes casos, os sinais laterais do sistema de balizamento marítimo, conhecido como *International Association of Lighthouse Authorities Lateral System* (IALA), são utilizados para orientação. Este sinais normalmente são boias ou balizas nas cores verde e vermelho.

Na Figura 1.5 é mostrado um esboço contendo duas edificações, BLOCO 01 e BLOCO 02, separadas por uma certa distância. Entre estes dois blocos há uma caminho (linha pontilhada) com origem no ponto A, BLOCO 01, e chegada no ponto B, BLOCO 02. Um robô móvel, que se oriente pela identificação de marcas coloridas para seguir o caminho entre os pontos A e B, está sujeito à mudança de iluminação na medida em que atravessa ambientes com diferentes fontes de iluminação. A identificação das marcas coloridas será prejudicada pela variação das cores capturadas em cada ambiente. Assim, é necessário que as imagens capturadas sejam corrigidas de forma a compensar a variação da iluminação e obter cores mais estáveis e precisas.

Figura 1.5: Caminho de um robô móvel.



Fonte: Elaborada pelo autor

1.2 Definição do problema

Nesta tese, o problema da constância de cor é abordado no contexto de sequências de imagens. As cores correlatas existentes entre os quadros consecutivos destas sequências compõem a principal fonte de informação e possibilitam um relaxamento das restrições relativas aos sensores e aos iluminantes. Os cenários presentes nestas sequências apresentam elementos naturais contendo árvores, plantas, calçadas, edificações e outros elementos que compõem os ambientes internos e externos. O problema abordado pode ser delimitado pelas seguintes características:

- **Imagens sequenciais:** Serão utilizadas imagens do mundo real e de ambientes comuns sujeitos a variação de iluminação para que se tenha pouca ou nenhuma restrição de efeitos indesejáveis como sombras, reflexos, *flashes* e outros. Estas imagens serão extraídas de vídeos que serão produzidos e outros disponíveis em banco de dados

existentes na internet em que seja possível obter informações correlatas entre quadros consecutivos.

- **Transformação:** Busca-se uma transformação T capaz de gerar uma imagem corrigida I_c a partir de uma imagem capturada I_a , tal que $I_c = T(I_a)$. Para isso, dado uma sequência de imagens capturadas $I_0, I_1, I_2 \dots I_n$, a n -ésima transformação T_n é função destas imagens, ou seja, $T_n = \Psi(I_n, I_{n-1}, I_{n-2} \dots I_1, I_0)$.
- **Refletâncias:** As superfícies são consideradas foscas e a componente especular da luz refletida são ignoradas.
- **Iluminante:** A iluminação será considerada uniforme sobre toda a imagem.
- **Sensores:** Nenhuma restrição será considerada em relação aos sensores, apenas que seguem o modelo tricromático.
- **Geometrias:** Os efeitos da geometria e dos ângulos de incidência e reflexão em um ponto qualquer, x_O , na superfície dos objetos serão considerados mínimos, ou seja, o fator de escala $K(x_O)$, que relaciona a radiância que emerge de uma superfície no ponto x_O e a irradiância que atinge este ponto, será considerado unitário.

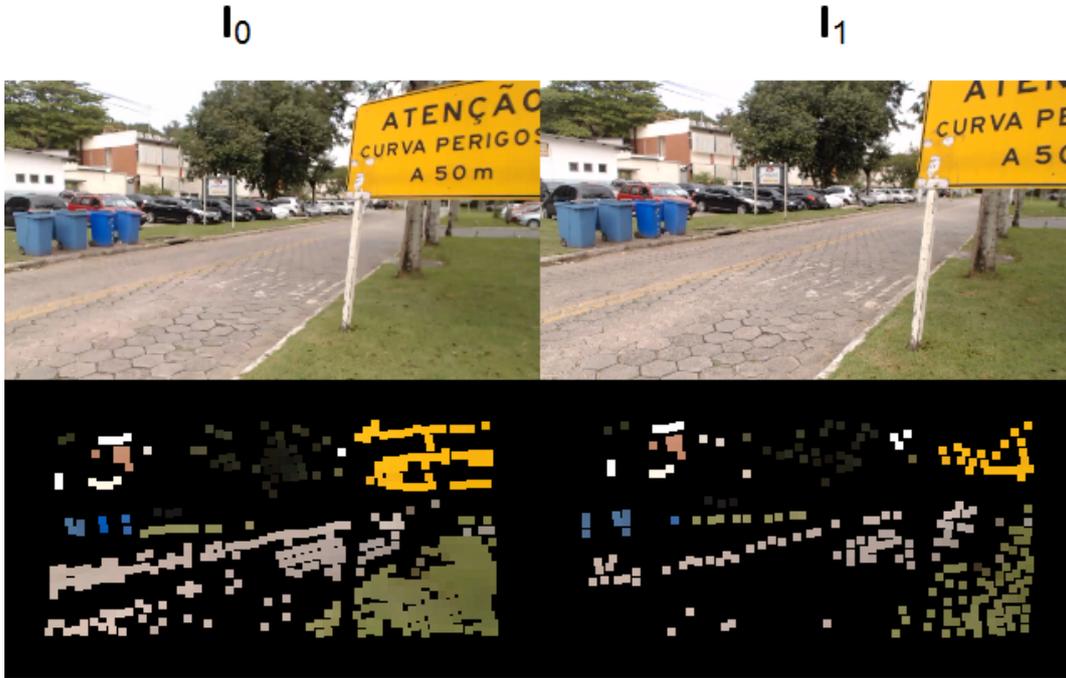
1.3 Metodologia proposta

Em uma sequência de imagens, as informações correlatas existentes entre quadros consecutivos podem servir como meio de propagação de valores de cores de referência úteis na correção dos quadros subsequentes. Esta estrutura fundamenta a hipótese principal dessa tese de que um processo iterativo de correção de cores para imagens subsequentes permite alcançar bons resultados na constância de cor.

Para alcançar a constância de cor em sequências de imagens, a metodologia apresentada nesta tese toma como ponto de partida a existência de regiões coloridas presentes em duas imagens consecutivas, \mathbf{I}_0 e \mathbf{I}_1 , de um vídeo (Figura 1.6). Estas regiões são representadas por suas cores médias na forma de dois conjuntos de cores, \mathbf{C}_0 e \mathbf{C}_1 , gerados para cada imagem, respectivamente. Devido ao movimento relativo entre câmera e cena, algumas regiões que aparecem em uma imagem podem não aparecer na outra e vice-versa. Isso faz com que os conjuntos tenham tamanhos diferentes.

Dados estes dois conjuntos, \mathbf{C}_0 e \mathbf{C}_1 , uma transformação T_1 , que mapeia as cores das duas imagens \mathbf{I}_0 e \mathbf{I}_1 , pode ser obtida utilizando-se os valores de cor correspondentes existentes nos dois conjuntos. A partir de uma estratégia de *tracking* é possível identificar as regiões

Figura 1.6: Dois quadro consecutivos e as regiões coloridas identificadas.



Fonte: Elaborada pelo autor

que são correlatas entre as duas imagens consecutivas. Então, dois novos conjuntos \hat{C}_0 e \hat{C}_1 , são gerados contendo apenas as cores identificadas nas duas imagens. A transformação T_1 é definida como uma função destes dois novos conjuntos como mostrado na Equação 1.4.

$$T_1 = \Psi(\hat{C}_1, \hat{C}_0) \quad (1.4)$$

Considerando que as imagens I_0 e I_1 tenham sido capturadas sob duas fontes de iluminação distintas, $L(\lambda)$ e $L'(\lambda)$, respectivamente, a transformação T_1 é capaz de gerar, a partir das cores da imagem I_1 , uma nova imagem I'_1 cujas cores seriam como se tivessem sido capturadas sob a iluminação $L(\lambda)$. Neste contexto, o conjunto \hat{C}_0 e a fonte de iluminação $L(\lambda)$ são tomados como referências para a correção da imagem I_1 .

Esta estratégia de correção de cores entre duas imagens consecutivas é a base de um processo iterativo em que as cores corrigidas de um quadro são utilizadas como referência para obtenção da transformação que corrigirá o quadro seguinte da sequência de imagens. Este processo se repete para cada iteração enquanto os conjuntos de cores se modificam dinamicamente com a inclusão de novas regiões e a exclusão de outras. Nesta abordagem, a fonte de luz que ilumina a cena do primeiro quadro é tomada como o iluminante de referência para a correção das cores. Caso seja necessário uma correção de cor em relação a um iluminante

padrão, um conjunto de cores calibradas pode ser utilizado para correção do primeiro quadro fazendo com que os quadros seguintes sejam corrigidos em relação a este iluminante padrão.

Para avaliar o método proposto, os resultados alcançados por ele serão comparados com aqueles provenientes dos métodos *Gray-World*, *Max-RGB* e *Gray-Edge* (van de Weijer et al., 2007a). O parâmetro de comparação será um fator de correção calculado pela razão entre a variação da cor corrigida e a variação da cor capturada de um mesmo objeto que apareça, por exemplo, no início e no final do vídeo.

1.4 Estrutura desta tese

Esta tese está organizada em cinco capítulos sendo o primeiro esta introdução. No Capítulo 2, uma revisão da literatura é apresentada, evidenciando os principais trabalhos realizados em constância de cor. Estes trabalhos são organizados em duas tabelas permitindo uma clara distinção entre eles e a identificação de onde o método aqui proposto se localiza no contexto da constância de cor.

No Capítulo 3, são apresentados dois métodos básicos desenvolvidos nesta tese para constância de cor em sequências de imagens. Estes métodos utilizam um processo iterativo para correção das cores dos quadros consecutivos. Um filtro temporal é utilizado para atenuar os efeitos dos ruídos.

No Capítulo 4 é apresentada uma série de testes que buscam a melhoria da qualidade das amostras de cor utilizadas na correção. No entanto, os efeitos dos ruídos presentes nas imagens continuam a degradar a qualidade da correção. Então, um outro método é apresentado.

No Capítulo 5 são detalhados os experimentos realizados para escolha da técnica de segmentação e aqueles relacionados com os métodos de correção de cor desenvolvidos. Os resultados são apresentados e analisados. Por fim, é apresentado um capítulo contendo as conclusões e as sugestões de trabalhos futuros.

Capítulo 2

Constância de Cor

2.1 Modelagem do sistema de visão

A visão humana se manifesta quando a imagem de uma cena é projetada sobre a retina. A retina é uma membrana que está na superfície interna do olho e sobre ela estão distribuídos dois tipos de células fotorreceptoras, os cones e os bastonetes. Existem cerca de 75 a 150 milhões de bastonetes distribuídos sobre a retina e vários deles são conectados a uma única ramificação nervosa reduzindo a quantidade de detalhes percebidos por estes sensores. Os bastonetes são sensíveis a baixos níveis de iluminação e não influenciam a visão colorida. Isso faz com que as cores dos objetos, quando vistos sob baixa iluminação, não sejam percebidas, caracterizando a visão escotópica¹. Os cones são células fotorreceptoras responsáveis pela principal característica da visão humana que é a capacidade de percepção das cores dos objetos. Eles estão localizados principalmente na região central da retina sendo que cada olho possui entre 6 e 7 milhões deles, o que representa de 4% a 10% da quantidade de bastonetes. Cada um destes cones está conectado a uma ramificação nervosa permitindo que pequenos detalhes possam ser observados. A visão gerada pelos cones é chamada visão fotópica² e se manifesta em condições de luz brilhante (Gonzalez e Woods, 2008).

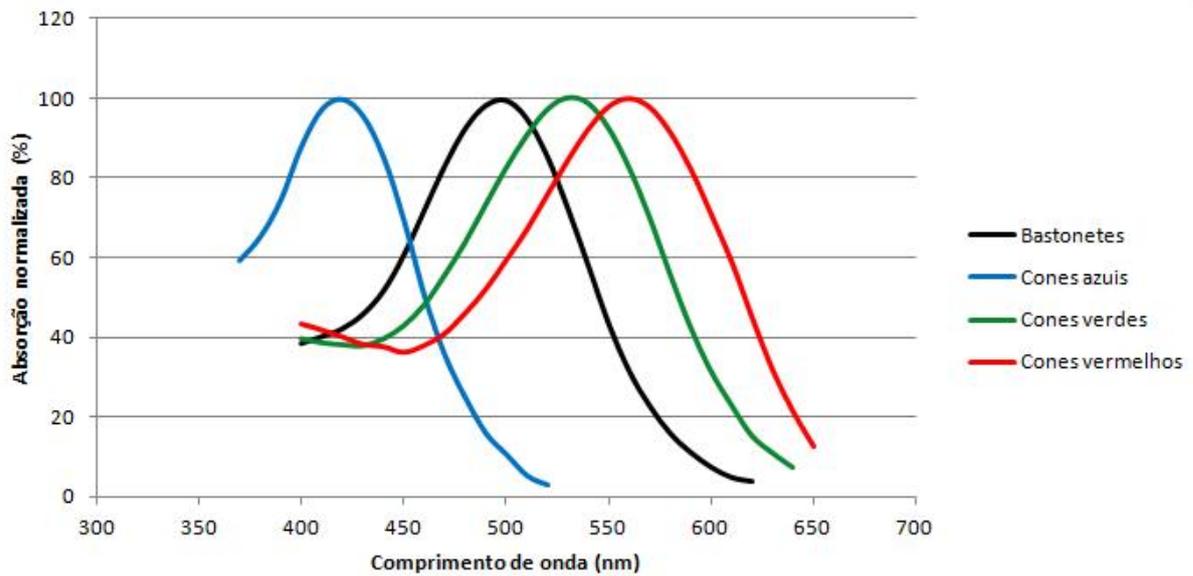
Os cones podem ser de três tipos dependendo do foto-pigmento que possuem. Eles formam três sistemas fisiológicos que são sensíveis a três faixas distintas L(*long*), M(*middle*) e S(*short*) de comprimento de onda da radiação eletromagnética (Figura 2.1). Os cones das três faixas L, M e S possuem a máxima absorção nos comprimentos de onda 559 nm, 531 nm e 419 nm. Estes comprimentos de onda caracterizam três cores distintas, o vermelho,

¹Esta é a visão produzida pelo olho em condições de baixa luminosidade.

²Visão produzida pelo olho em condições de intensidade luminosa que permitam a percepção das cores

o verde e o azul, respectivamente. Os bastonetes têm absorção máxima em 496nm (Dartnall et al., 1983; Ebner, 2007). Existem casos em que a retina possui mais que três tipos de células cones. Esta situação é chamada de tetracromatismo e permite que o observação perceba um número maior de cores distintas. Nos humanos, a incidência de tetracromatismo nas mulheres é maior que nos homens (Jameson et al., 2001).

Figura 2.1: Curva espectral de absorção dos três tipos de cones (azul, verde e vermelho) e dos bastonetes.



Fonte: Elaborada pelo autor a partir dos dados disponibilizados por Dartnall et al. (1983)

Este modelo da visão humana está em concordância com a teoria de Young-Helmholtz (Wyszecki e Stiles, 2000) sendo utilizado na construção dos sensores de imagem. Em tais sensores, o modelo tricromático é denominado de modelo RGB para representação de cores cujos canais são chamados *Red*, *Green* e *Blue*. Cada canal $c_m(x_I)$ medido em uma posição x_I da imagem depende da irradiância $E(\lambda, x_I)$ que alcança o sensor no ponto x_I e da curva de sensibilidade do canal $S_m(\lambda)$ para $m \in \{r, g, b\}$, onde λ é o comprimento de onda da luz no espectro visível. Esta irradiância é proporcional à radiância $L(\lambda, x_O)$ que emerge da superfície do objeto no ponto x_O . Considerando um fator de escala unitário pode-se escrever $E(\lambda, x_I) = L(\lambda, x_O)$.

Para superfícies lambertianas, ou seja, aquelas que possuem um aspecto fosco sem reflexão especular, a refletância no ponto x_O , $R(\lambda, x_O)$, é constante em todas as direções e a radiância que emerge dela é equivalente ao produto entre a radiância emitida pela fonte de luz, $L(\lambda)$, sua refletância e um fator de escala, $K(x_O)$. Este fator depende do ângulo entre o

vetor normal à superfície e o vetor na direção da luz incidente. Admitindo que a luz refletida seja independente do ângulo de incidência, ou seja, $K(x_O) = 1$, a expressão da energia medida em cada canal será dada pela Equação 2.1.

$$c_m(x_I) = \int_{\lambda_1}^{\lambda_2} S_m(\lambda)L(\lambda)R(\lambda, x_O)d\lambda \quad (2.1)$$

Nesta equação, $c_m(x_I)$ é o valor medido na posição x_I da imagem para $m \in \{r, g, b\}$ e $[\lambda_1, \lambda_2]$ é o intervalo de comprimento de onda de resposta do sensor. A Equação 2.1 caracteriza a relação existente entre o iluminante da cena, o sensor e a refletância de uma superfície na formação da cor. Qualquer alteração em uma dessas três componentes altera o valor da cor observada. A cor medida não é uma característica da superfície, mas uma representação da interação entre os componentes que a formam. A única componente que caracteriza intrinsecamente uma superfície é sua refletância. Para que sistemas computacionais possam identificar objetos corretamente utilizando suas cores é necessário que possuam um esquema de correção de cores que opere sob diferentes fontes de iluminação (Novak e Shafer, 1991).

A medição de cada canal de cor é uma amostragem do espectro da radiação que alcança os sensores. Esta amostragem resulta em uma perda de informação capaz de gerar uma ambiguidade na representação da cor no espaço RGB. Por exemplo, duas superfícies com distribuição espectral de refletância distintas geram duas representações RGB distintas quando iluminadas por um dado iluminante. No entanto, sob a iluminação de um outro iluminante, estas duas superfícies podem gerar uma mesma representação RGB. Neste caso elas são chamadas superfícies metaméricas (Wyszecki e Stiles, 2000). Esta ambiguidade na representação pode ocorrer em qualquer espaço de cor tricromático.

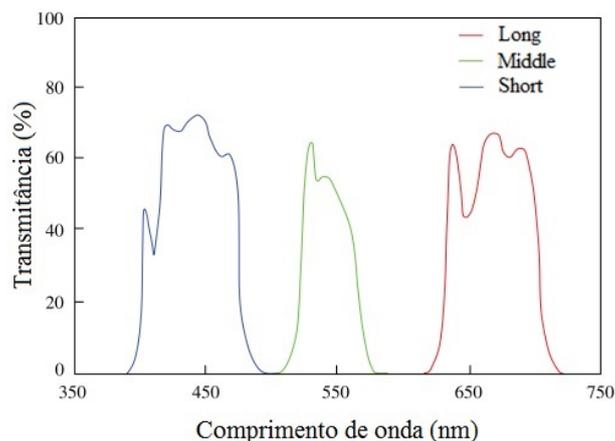
2.2 A constância de cor na visão humana

A constância de cor presente no sistema visual humano é objeto de análise desde o século XVIII quando Gaspard Monge, um matemático francês, criador da Geometria Descritiva (O'Connor e Robertson, 2014), registrou suas observações. Ele considerou de forma categórica que a percepção das cores não depende do valor absoluto de uma grandeza física, mas é influenciada pelo contexto a que está sujeita e pela estimativa do iluminante (Mollon, 2006). Outros nomes como Thomas Young, Hermann von Helmholtz, Ewald Hering, Johannes von Kries, Henry Helson, Deane Brewster Judd e Edwin Herbert Land também fizeram contribuições sobre este assunto desde então (Foster, 2011).

Muitos trabalhos envolvendo as áreas da biologia e psicologia buscam desvendar os processos e estruturas envolvidos no fenômeno da constância de cor na visão humana (Hurlbert, 1999; Kraft e Brainard, 1999). As observações feitas nestes experimentos permitem a formulação de modelos biológicos de constância de cor a serem utilizados em aplicações práticas na área da fotografia e visão computacional (Semo e Spitzer, 2000; Vazquez-Corral et al., 2012). Outros trabalhos que não se baseiam no modelo da visão humana propõem algoritmos frequentemente chamados de constância de cor computacional. Estes modelos computacionais não explicam completamente a constância de cor observada em humanos (Gijsenij et al., 2011).

Um dos trabalhos mais conhecidos é o clássico experimento realizado por Edwin H. Land (Land e McCann, 1971) que verifica a afirmação de Gaspard Monge. Land identificou que as cores observadas não dependem do valor absoluto do espectro de radiação que alcança os olhos. Para isso, ele utilizou três projetores de luz com filtros em bandas distintas de comprimento de onda para iluminar um conjunto de papéis retangulares coloridos dispostos de forma semelhante a uma pintura do artista Piet Mondrian. As bandas foram chamadas de *long-wave*, *middle-wave* e *short-wave*. A razão entre a intensidade da luz que atravessa o filtro e a intensidade que chega nele define a transmitância deste filtro para cada comprimento de onda (Figura 2.2).

Figura 2.2: Curva espectral dos filtros colocados em frente aos projetores para iluminar as superfícies coloridas.



Fonte: (Land e McCann, 1971)

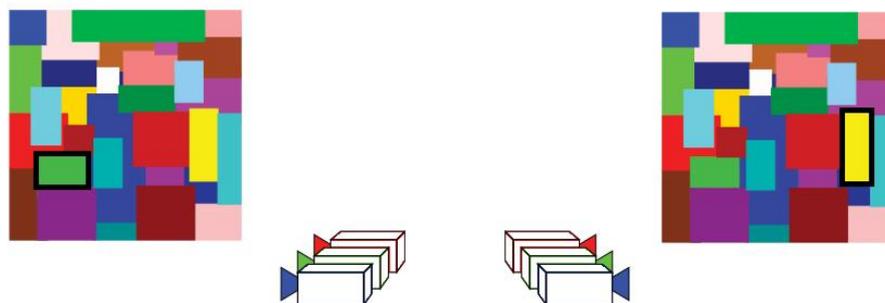
Inicialmente, os três projetores foram ligados simultaneamente e suas intensidades ajustadas para que o conjunto das superfícies aparecesse fortemente colorido e com um branco evidente. Em seguida, ligando um projetor por vez, a luminância que atinge o olho proveniente de um retângulo branco foi medida com um fotômetro. Assim, foram obtidos três

números, um para cada banda do espectro de radiação, que são proporcionais à luminância. Estes valores, para a superfície branca, foram 6, 35 e 60 respectivamente para as bandas *long*, *middle* e *short*. Em seguida, Land iluminou uma superfície de cor marrom escuro e ajustou a intensidade dos projetores de luz para que fossem medidos os mesmos valores provenientes da superfície branca, qual seja, 6, 35 e 60. Ele observou que a sensação de cor da nova superfície permaneceu inalterada, ou seja, marrom escuro.

Land repetiu este experimento para outras superfícies coloridas como amarelo claro, azul, cinza, verde limão e vermelho. Para cada uma destas superfícies ele ajustou os projetores de forma que se medisse sempre 6, 35 e 60 nas bandas *long*, *middle* e *short*. As cores das superfícies apareceram inalteradas e os observadores reportaram que a sensação de cor da série foi amarelo, azul, cinza, verde e vermelho. A conclusão de Land foi que a sensação de cor não depende da luz refletida pelo objeto, mas da refletância. Assim, um observador humano é capaz de estimar as refletâncias dos objetos independentemente do iluminante utilizado. Por outro lado, as câmeras digitais ou analógicas podem fazer, no máximo, o que um fotômetro faz, qual seja, medir a luz refletida por uma superfície (Ebner, 2007).

Em outro experimento, Land (1977) utilizou dois grupos de três projetores para iluminar dois conjuntos distintos de superfícies coloridas (Figura 2.3). Os três primeiros projetores, pertencentes ao primeiro grupo, foram ligados para iluminar uma superfície branca do primeiro conjunto de superfícies. A luz refletida dessa superfície branca foi medida para cada uma das três bandas de comprimento de onda. Em seguida, uma superfície verde pertencente ao primeiro conjunto de superfícies e uma superfície amarela pertencente ao segundo conjunto foram selecionadas. Os dois grupos de projetores, cada um iluminando um conjunto de superfícies, foram ajustados de forma que cada superfície colorida, verde e amarela, refletisse a mesma quantidade de luz que a superfície branca nas três bandas de comprimento de onda.

Figura 2.3: Cada conjunto de cores é iluminado por um grupo de três projetores.



Fonte: Elaborada pelo autor

Quando os seis projetores foram ligados os observadores identificaram as duas superfícies coloridas como sendo, respectivamente, verde e amarela. No entanto, quando os observadores olharam as mesmas superfícies através de tubos (Figura 2.4), identificaram tanto a superfície verde quanto a amarela como sendo de cor branca acinzentada. Os tubos foram posicionados de tal forma que o centro da superfície verde era visto pelo olho esquerdo e o centro da superfície amarela pelo olho direito. O formato dos tubos assegurava também que somente a luz proveniente dos centros das superfícies e nenhuma luz da vizinhança alcançavam os olhos dos observadores.

Figura 2.4: Um longo tubo com abertura ajustável.



Fonte: Elaborada pelo autor

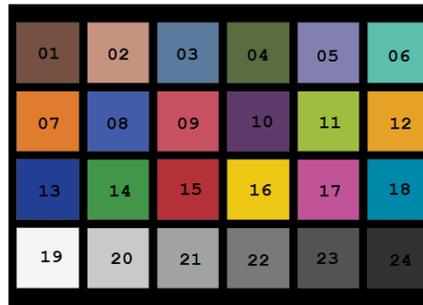
A partir disso, ficou evidente que o tipo do ambiente em torno de um objeto influencia na percepção de sua cor (Ebner, 2007). Land concluiu que o processamento da visão colorida poderia ocorrer na retina, no córtex cerebral ou parcialmente em ambos. Uma vez que a localização exata era desconhecida ele criou o termo *Retinex*, uma combinação de retina e córtex, para descrever o conjunto de mecanismos biológicos que processam a visão colorida humana (Land, 1977).

2.3 A constância de cor na visão computacional

Uma abordagem bastante explorada por diversos autores para se alcançar a constância de cor em sistemas de visão computacional é aquela baseada no modelo de dimensão finita (Brill, 1978; D’Zmura e Lennie, 1986; Finlayson et al., 1994a; Funt et al., 1991; Healey e Slater, 1994). Neste modelo, tanto a refletância das superfícies quanto o iluminante da cena são caracterizados por uma soma ponderada de funções base. O objetivo é estimar os coeficientes que representam as refletâncias a partir dos valores obtidos dos sensores em cada posição da imagem. Para isso, devem ser conhecidas as curvas espectrais dos sensores e as funções base dos iluminantes e das refletâncias. Mesmo com estas informações, quando os coeficientes que representam a iluminação e as refletâncias são desconhecidos têm-se mais variáveis que equações e o problema não pode ser resolvido. Nos Apêndices B, C e D estão descritas as modelagem para iluminantes, refletâncias e sensores.

Por conta disso, para se alcançar uma solução, algumas considerações devem ser feitas. Maloney e Wandell (1986) apresentam um método para recuperar com exatidão os coeficientes das refletâncias em uma imagem, mas eles consideram os iluminantes representados por três funções base e as refletâncias por duas. Buchsbaum (1980) considera como cinza a refletância média de uma imagem. A partir dessa informação ele estima os coeficientes do iluminante para, então, estimar as refletâncias. Esta abordagem deu origem ao algoritmo *Gray-World*. O uso de um cartão de cores do tipo *ColorChecker* (Figura 2.5) com características espectrais conhecidas é sugerido por Novak e Shafer (1991) para estimar a curva de distribuição espectral de um iluminante desconhecido. Com este iluminante determinado, as refletâncias podem ser estimadas.

Figura 2.5: Cartão do tipo *ColorChecker* com 24 placas coloridas.



Fonte: Elaborada pelo autor

2.3.1 Constância de cor em imagem única

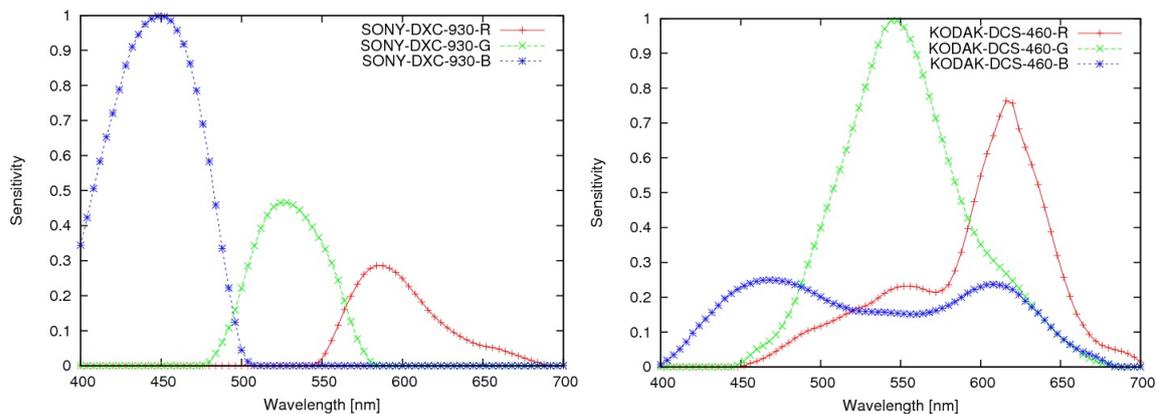
Quando as considerações são referentes às curvas espectrais dos sensores, uma outra abordagem para os algoritmos de constância de cor é definida. A principal consideração é de que as curvas espectrais dos sensores são banda estreita e que não há sobreposição entre elas. No limite deste conceito, estas curvas se aproximam de uma função delta, ou seja, $S_m(\lambda) = \delta(\lambda - \lambda_m)$, onde $S_m(\lambda)$ é a curva de sensibilidade espectral do sensor para os canais $m \in \{r, g, b\}$ e λ_m é a frequência espectral em que a resposta é diferente de zero. As componentes de cor, c_m , de uma superfície com refletância $R(\lambda)$ iluminada por um iluminante $L(\lambda)$, obtidas com sensores banda estreita a partir da Equação 2.1, são dadas pela Equação 2.2.

$$\begin{bmatrix} c_r \\ c_g \\ c_b \end{bmatrix} = \begin{bmatrix} R(\lambda_r)L(\lambda_r) \\ R(\lambda_g)L(\lambda_g) \\ R(\lambda_b)L(\lambda_b) \end{bmatrix} \quad (2.2)$$

Isto conduz a uma simplificação permitindo que as imagens tenham suas cores corrigidas apenas multiplicando os canais de cor por ganhos individuais. Estes ganhos são obtidos pela razão entre as componentes de cor L_m e L'_m do iluminante padrão, $L(\lambda)$, e do iluminante da cena, $L'(\lambda)$, respectivamente. Assim, uma cor presente na cena, C'_m , é transformada para uma cor c_m . Este modelo é comumente chamado na literatura de constância de cor computacional (Gijsenij et al., 2012b).

A consideração de que as curvas espectrais dos sensores se aproximem de funções delta impõe uma restrição muito severa aos sensores das câmeras comerciais uma vez que estes são fabricados sob uma vasta variedade de funções de sensibilidade. A Figura 2.6 mostra as curvas espectrais dos sensores presentes nas câmeras Sony DXC-930 e Kodak DCS-460³ (Kawakami et al., 2013). A câmera Sony DXC-930 é muito utilizada em experimentos realizados para validar os algoritmos apresentados na literatura. As curvas espectrais dos sensores dessa câmera são relativamente estreitas e com pouca sobreposição quando comparados, por exemplo, com aqueles da câmera Kodak DCS-460. Esta última apresenta sensores banda larga e com uma significativa sobreposição. Esta enorme diversidade na sensibilidade dos sensores faz com que os algoritmos baseados na suposição de banda estreita apresentem resultados diferentes para diferentes câmeras.

Figura 2.6: Curvas de sensibilidade dos sensores das câmeras Sony DXC-930 e Kodak DCS-460.



Fonte: (Kawakami et al., 2013)

Os algoritmos de constância de cor computacional baseados na suposição de que os sensores são banda estreita buscam, em primeiro lugar, estimar a cor do iluminante da cena tomando como base a imagem. Em seguida, calculam os coeficientes que escalonarão os canais de cor da imagem, dado que o iluminante padrão seja conhecido, e, por fim, geram

³A base de dados numérica e as imagens das curvas podem ser obtidas no endereço eletrônico <http://www.cvl.iis.u-tokyo.ac.jp/~rei/research/cs/zhao/database.html>

uma nova imagem com cores corrigidas. Estes coeficientes podem ser arranjados na forma matricial, como na Equação 2.3, gerando uma matriz diagonal (Gijssenij et al., 2012b), onde L_r , L_g e L_b são as componentes da cor do iluminante padrão, $L(\lambda)$, enquanto L'_r , L'_g e L'_b são as componentes de cor do iluminante da cena, $L'(\lambda)$. A maioria das abordagens que utilizam este modelo ainda consideram que as superfícies são foscas não possuindo componente especular, que o espectro de frequências da iluminação é uniforme em toda imagem e que os efeitos da geometria dos objetos e dos ângulos de incidência e reflexão são mínimos (Ebner, 2007).

$$\begin{bmatrix} c_r \\ c_g \\ c_b \end{bmatrix} = \begin{bmatrix} \frac{L_r}{L'_r} & 0 & 0 \\ 0 & \frac{L_g}{L'_g} & 0 \\ 0 & 0 & \frac{L_b}{L'_b} \end{bmatrix} \begin{bmatrix} c'_r \\ c'_g \\ c'_b \end{bmatrix} \quad (2.3)$$

Diversos algoritmos que estimam a cor do iluminante a partir de uma única imagem digital têm sido propostos. Gijssenij et al. (2011) propõem uma classificação para estes algoritmos dividindo-os em três categorias:

1. os estáticos;
2. os baseados em escala de cores;
3. os baseados em aprendizagem;

Na primeira categoria, estão os algoritmos cujos parâmetros de ajustes são mantidos fixos ou estáticos. O mais conhecido é o algoritmo *Gray-World* (Buchsbaum, 1980) que considera acromática a média das refletâncias em uma imagem sob iluminação neutra e qualquer desvio nesta média é devido à cor do iluminante. Melhorias neste algoritmo são alcançadas segmentando-se a imagem para computar a média das cores dos segmentos (Gershon et al., 1987; Ebner, 2009; Gijssenij e Gevers, 2007). O algoritmo *Max-RGB*, uma versão simplificada do *Retinex*, considera que a máxima resposta dos canais R, G e B é proveniente de uma superfície branca perfeita de onde pode-se obter a cor do iluminante (Gijssenij et al., 2011; Land, 1977). Este algoritmo possui uma implementação muito simples, no entanto, ele é sensível aos ruídos presentes nas imagens. As abordagens que suavizam as imagens, filtrando os ruídos, antes de se estimar o iluminante geram resultados melhores (Gijssenij e Gevers, 2007; Ebner, 2009). O algoritmo *Gray-Edge* está baseado na hipótese de que a média da variação da cor nas bordas é acromática (van de Weijer et al., 2007a). Uma versão que contempla diferentes tipos de bordas de forma ponderada é apresentada por Gijssenij et al. (2009). O modelo dicromático de reflexão é adotado por alguns métodos que exploram a presença de

superfícies especulares e destaques nas imagens (Lee, 1986; Tominaga e Wandell, 1989; Tan et al., 2004).

Na segunda categoria, estão os algoritmos baseados em *Gamut-mapping* onde supõe-se que o número de cores observadas nas imagens do mundo real, para um dado iluminante, é limitada, ou seja, pertencem a um *Gamut* padrão. Este conjunto de cores é determinado durante uma fase de treinamento pela observação de várias superfícies iluminadas por uma fonte de luz conhecida. Qualquer variação na imagem que seja proveniente de cores que estão fora deste *Gamut* padrão é consequência de um desvio na cor do iluminante (Forsyth, 1990; Barnard, 2000). Aplicando um estimador, o mapeamento mais adequado é selecionado dentre um conjunto de mapeamentos possíveis. Quando aplicado no espaço de cor definido pela razão dos canais R e G pela canal B, (R/B, G/B) (Finlayson, 1996) este algoritmo apresenta um desempenho ligeiramente pior que na abordagem (R, G, B). Melhorias nesta abordagem podem ser alcançadas mapeando o espaço de cromaticidade de volta para o espaço RGB antes da seleção do mapeamento mais adequado (Finlayson e Hordley, 1999, 2000).

A terceira categoria reúne os algoritmos que utilizam técnicas de aprendizagem para estimar o iluminante. Eles são baseados em redes neurais (Cardei et al., 2002), em suporte vetorial para regressão (Funt e Xiong, 2004) e em técnicas de regressão linear (Agarwal et al., 2006b, 2007). Existem abordagens que utilizam estatísticas de baixo nível (Finlayson et al., 2001) e formulação Bayesiana (Brainard e Freeman, 1997; Sapiro, 1999). Uma abordagem que combina a saída de múltiplos algoritmos para estimar um iluminante médio é apresentada por Schaefer et al. (2005). Em Gijsenij e Gevers (2011) a avaliação das propriedades intrínsecas das imagens permite selecionar o algoritmo mais adequado para a correção. Além desses, uma abordagem baseada no conteúdo semântico da cena, como céu e grama, é proposto por van de Weijer et al. (2007b).

Para os casos em que as curvas espectrais dos sensores não são banda estreita, Finlayson et al. (1994b) sugerem uma transformação chamada *spectral sharpening*. Esta transformação tem o objetivo de gerar curvas espectrais mais estreitas utilizando uma combinação linear das curvas existentes dos sensores. Na prática, este efeito é conseguido aplicando-se a transformação sobre os valores RGB da imagem e mapeado-os para outro espaço de cor. A nova imagem produzida é, então, corrigida por algum algoritmo de constância de cor. Por fim, uma transformação inversa deve ser aplicada para se voltar ao espaço original. No entanto, esta transformação não tem o desempenho assegurado para qualquer que seja o algoritmo de constância de cor aplicado (Barnard e Funt, 1998). Além disso, a dependência do conhecimento do iluminante da cena e outras restrições não deixa claro a praticidade da técnica (Barnard et al., 2001).

Abordagens menos comuns na literatura são aquelas que utilizam imagens adicionais (Finlayson et al., 2005; Fredembach e Finlayson, 2008; Tsukada e Ohta, 1990; Xiong e Funt, 2009) ou algum dispositivo diferente (Zaraga e Langfelder, 2010; Nieves et al., 2008) para alcançar a constância de cor. O uso de três cores conhecidas e presentes em uma imagem permite a formulação de uma base para a representação de qualquer refletância. A partir dos valores RGB, os coeficientes representando as refletâncias são determinados sem a necessidade de se conhecer o iluminante. Esta abordagem é chamada de modelo volumétrico (Bril, 1980; Kang, 2006). Outras abordagens buscam obter descritores de cor que sejam invariantes em relação à iluminação sem gerar, necessariamente, uma imagem corrigida que tenha correspondência com a imagem de entrada (Finlayson et al., 1996; Geusebroek et al., 2003; Funt e Finlayson, 1995; Gevers e Smeulders, 1999; Healey e Slater, 1994).

2.3.2 Constância de cor em sequências de imagens

Os algoritmos citados anteriormente são concebidos no contexto de imagens fixas, mas poderiam ser utilizados para corrigir, quadro a quadro, as imagens provenientes de um vídeo (Wang et al., 2011). Este é o caso em que os algoritmos *Gray-World* e *Gamut-Mapping* são avaliados por Renno et al. (2005) quando aplicados sobre imagens individuais provenientes de vídeos gerados em sistemas de vigilância. Talvez, por esta facilidade de aplicação, pouquíssimos trabalhos tratam da constância de cor aplicada a sequências de imagens e mais raros ainda são aqueles onde há mudanças de ambientes constantemente ou que tiram proveito da informação existente em quadros adjacentes. Em um deles, Almonfrey et al. (2011) empregam um controlador servo-visual na navegação de um robô móvel em ambiente externo sob grande variação da iluminação natural. Ele utiliza um conjunto de cores conhecidas nas cenas para obter uma matriz de transformação (Konzen e Schneebeli, 2007). Estas cores são obtidas pela presença permanente de um cartão do tipo *ColorChecker* da X-Rite (X-Rite, 2009) nas imagens capturadas. Não há uso de informação correlata entre quadros consecutivos, apenas a multiplicação da matriz pela imagem capturada para se obter a imagem corrigida.

Wang et al. (2011) obtêm uma melhor estimativa do iluminante utilizando estimativas da cor do iluminante para um grupo de quadros adjacentes de uma sequência de imagens. Recortes de vídeo, chamados de cenas, são separados quando ocorre uma mudança significativa do iluminante. O iluminante é estimado para cada quadro do recorte e um valor médio é calculado a partir destas estimativas. Este iluminante médio é considerado como sendo o iluminante daquela cena e seu valor é utilizado na correção de cada um dos quadros daquele recorte de vídeo. Os métodos de constância de cor utilizados consideram os sensores como banda estreita e são aplicados nos quadros individualmente para estimar o iluminante. Prinnet

et al. (2013) utilizam o modelo dicromático de reflexão considerando que o iluminante da cena não muda com o tempo e a componente de cor difusa é a mesma entre dois quadros consecutivos. A cor do iluminante é obtida por diferenciação das componentes RGB de qualquer ponto com uma componente especular identificado entre dois quadros consecutivos. Uma vez que a cor do iluminante foi identificada, a correção é aplicada por uma matriz diagonal.

Garud et al. (2014) utilizam um esquema de balanço de branco para corrigir quadros intercalados de uma sequência de imagens. Os parâmetros do ajuste são definidos a cada cinco quadros sendo atualizados apenas quando há uma variação significativa do iluminante estimado. A eficiência é avaliada quantitativamente para os quatro retângulos mais brilhantes do cartão do tipo *ColorChecker* presentes na escala de cinza. Os resultados apresentados são, então, apenas relativos ao ajuste de branco e uma avaliação da correção nas demais cores não é apresentada. Além disso, não há comparação com outros métodos.

2.3.3 Abordagem proposta

A abordagem apresentada nesta tese possui características que a distinguem das demais abordagens e algumas características presentes nos métodos conhecidos. Dentre as características principais da abordagem apresentada aqui, tem-se:

- **Uso de sequências de imagens:** A abordagem apresentada nesta tese se aplica a sequências de imagens diferentemente daquelas que tratam de imagens estáticas. Esta característica permite que informações correlatas, existentes entre quadros, sejam utilizadas. Isto contribui para que algumas restrições, que normalmente são feitas pelos algoritmos de constância de cor, sejam relaxadas. Entre estas restrições estão a consideração de que os sensores das câmeras são banda estreita e a necessidade de se conhecer as curvas espectrais das fontes de luz das cenas.
- **Uso de cores RGB:** O uso do espaço RGB de cores se deve, primeiro, por ser o espaço de cores que as câmeras comerciais utilizam para representação das imagens geradas e, segundo, a conversão do espaço RGB para o espaço XYZ é uma tarefa difícil de ser implementada uma vez que se caracteriza como uma etapa de calibração da câmera e deveria ser executada para cada câmera em que o espaço de cores é dependente do dispositivo. Além disso, espaços de cores como o HSI e HSV não seriam adequados devido às indeterminações do matiz (H) quando a saturação é zero. Neste caso, o ruído pode fazer o matiz variar bruscamente. Há também de se considerar a descontinuidade na representação do matiz que pode ser 0 e 1 para a mesma cor vermelha. Outra característica importante da abordagem apresentada é a utilização de valores RGB das cores

presentes nas imagens como fonte de informação para promover a correção de toda a imagem. Muitos trabalhos, principalmente aqueles que consideram o modelo linear de dimensão finita, utilizam as curvas espectrais dos sensores e conjuntos de funções base para refletância e iluminantes como fonte de informação. Câmeras convencionais não fornecem as curvas espectrais, mas sim uma integração dessas curvas sob a forma de três valores escalares (RGB). As tentativas de se determinar as curva a partir desses parâmetros se defronta com o problema de metamerismo, valores RGB idênticos para curvas espectrais diferentes.

- **Uso de várias cores:** As abordagens de constância de cor que consideram a suposição da banda estreita dos sensores necessitam de apenas uma cor característica e seu respectivo valor de referência para determinarem a relação de correção. Em princípio, qualquer cor presente na imagem poderia ser utilizada com esta finalidade. No entanto, o desconhecimento do valor de referência para estas cores dificulta sua escolha. Quando a cor do iluminante é tomada como cor característica, tem-se maior liberdade para a escolha do valor de referência. Todavia, a determinação da cor do iluminante não é uma tarefa fácil. Para isto, as abordagens de constância de cor fazem suposições adicionais e utilizam as diversas cores presentes na imagem para estimar a cor do iluminante. Uma vez que a cor do iluminante tenha sido estimada, ela é utilizada na determinação da relação de correção. Neste processo, muita informação é descartada e quando a suposição da banda estreita dos sensores falha, erros mais expressivos são observados. Por outro lado, a abordagem apresentada nesta tese utiliza diretamente as diversas cores existentes na imagem para determinar a relação de correção. Isto permite que toda informação disponível na imagem seja aproveitada.
- **Uso de um processo iterativo:** A abordagem apresentada nesta tese utiliza um processo iterativo que toma o primeiro quadro como referência e estima, a cada iteração, novas referências para a correção contínua de cores. Estas referências representam as condições de iluminação do quadro de referência e à medida que a sequência de imagens evolui e as referências vão sendo reconstruídas, a informação do iluminante do primeiro quadro se propaga junto com a evolução das sequências de imagens. Este procedimento difere das abordagens apresentadas por Wang et al. (2011) e Prinet et al. (2013) que utilizam a correlação entre quadros adjacentes para correção dos quadros correntes sem transmitirem nenhuma informação para a correção dos quadros seguintes.
- **Sem restrições aos sensores:** Outra característica que restringe a diversidade de câmeras a serem utilizadas é aquela referente à largura de banda de sensibilidade dos sensores. A maioria dos algoritmos considera sensores de banda estreita e sem sobreposição. Em contrapartida, na abordagem apresentada não há nenhuma restrição aos

sensores, apenas que seguem o modelo tricromático de cor. Isto permite que câmeras de baixo custo que possuam sensores banda larga e com sobreposição possam se beneficiar da abordagem de constância de cor apresentada nesta tese.

- Sem restrições aos iluminantes: Algumas abordagens tiram proveito do conhecimento prévio do iluminante da cena, como é o caso daqueles métodos que restringem sua aplicação a ambientes externos iluminados com a luz do Sol. Outros buscam estimar os iluminantes das cenas utilizando curvas espectrais de cores conhecidas (Novak e Shafer, 1991). Na abordagem proposta não há restrição aos iluminantes, pois eles estão implícitos na transformação T .
- Mapeamento por uma transformação linear 3x3: Diferente dos métodos que utilizam uma matriz diagonal suportada pela suposição de que as curvas espectrais dos sensores são banda estreita, nesta tese, a abordagem apresentada utiliza uma matriz de correção com maior grau de liberdade dado por seus nove parâmetros. Isto permite uma diversidade maior de iluminantes e refletâncias que podem ser ajustados nas correções das imagens capturadas, sendo útil, por exemplo, quando se desloca entre ambientes internos com diferentes fontes de iluminação.

Por questões de simplificação, algumas restrições que comumente são feitas na maioria dos trabalhos apresentados na literatura também são consideradas na abordagem apresentada nesta tese. Dentre estas restrições, têm-se:

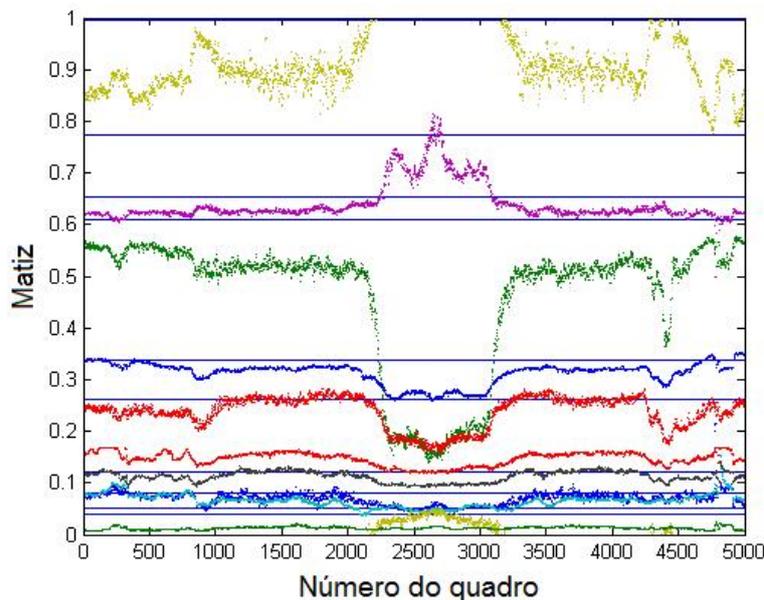
- Iluminação uniforme em toda imagem: Uma das restrições mais frequentes nas abordagens de constância de cor é a consideração de que a iluminação é uniforme em toda imagem. Esta consideração deve falhar em casos bem específicos como próximo a uma parede de cor muito saturada, uma considerável região de sombra e nos casos de cenas com mais de um iluminantes.
- As refletâncias são difusas: A abordagem apresentada nesta tese considera que as superfícies dos objetos são difusas, ou seja, não possuem a componente especular. Supõe-se também que os efeitos das geometrias dos objetos sejam mínimos e não há influência dos ângulos de incidência de reflexão da luz sobre as superfícies dos objetos.
- Refletâncias modeladas por três funções base: Também é considerado que as refletâncias são bem modeladas por uma base tridimensional. Isto quer dizer que o espaço de refletâncias possíveis é formado por todas as combinações lineares de três funções base.

Além de todas as considerações apresentadas, a abordagem descrita nesta tese não considera em seu modelo os efeitos de fatores, tais como: não linearidade dos sensores, distorções das lentes, ruído de corrente escura (*dark current*), ruído devido à natureza quântica da luz (*shot noise*) e ruído de quantização. Uma análise destes ruídos e seu modelamento para sensores CCD pode ser encontrado em Withagen et al. (2005) e Healey e Kondepudy (1994). Em comparação com CCD, os sensores do tipo CMOS são mais ruidosos (Brouk et al., 2008), mas apresentam facilidades de integração e permitem dispositivos mais compactos e de menor consumo de energia (Tian, 2000).

2.3.4 Trabalhos desenvolvidos nesta tese

Simão et al. (2013) comparam a estabilidade do matiz de um grupo de cores ao longo de um percurso que atravessa ambientes não estruturados sujeitos à forte variação de iluminação. As cores são obtidas de um *ColorChecker* presente nas imagens capturadas. Os valores dos matizes das cores capturadas são plotados juntamente com os matizes de referência representados por linhas base (Figura 2.7).

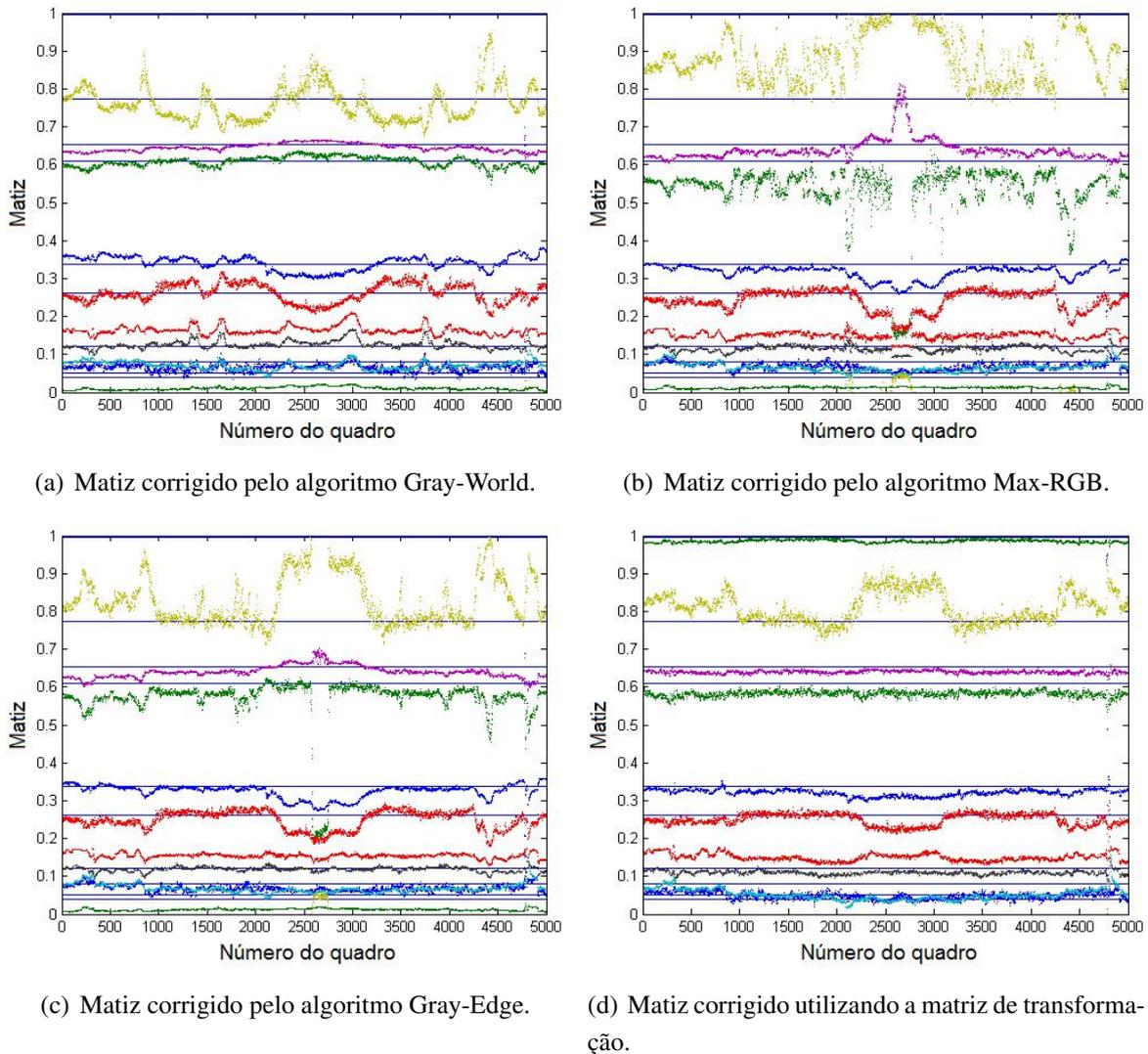
Figura 2.7: Variação do matiz das cores nas imagens capturadas.



Fonte: Elaborada pelo autor

Os resultados dos algoritmos *Gray-World*, *Max-RGB*, *Gray-Edge* são calculados, quadro a quadro, para comparação com o resultado da correção por uma matriz de transformação (Figura 2.8). A matriz é obtida com os valores de cor das imagens capturadas e os respectivos

Figura 2.8: Variação do matiz das cores nas imagens corrigidas por cada algoritmo.



Fonte: Elaborada pelo autor

valores padronizados do *ColorChecker*. Os resultados mostraram que os algoritmos *Max-*RGB**, *Gray-*Edge** produziram resultados mais ruidosos enquanto que a correção utilizando a matriz de transformação e o *Gray-*World** apresentaram os menores erros acumulados. A correção com a matriz 3x3 de transformação apresentou um resultado cerca de 20% melhor que o método *Gray-*World**. A correção é feita de forma direta, ou seja, não é um processo iterativo e não há uso de informações correlatas entre quadros consecutivos.

O uso da matriz de mapeamento de cores aplicada em um processo de correção iterativo é apresentado por Simão et al. (2014). Neste processo, os valores padronizados das cores de um *ColorChecker* presente nas imagens são utilizados apenas para correção do primeiro quadro. Os valores das cores corrigidas são, então, utilizados como referência para correção

do quadro seguinte. Nesta abordagem, os valores das cores presentes no *ColorChecker*, tanto das imagens capturadas quanto das imagens corrigidas estabelecem a correlação entre dois quadros consecutivos. O processo iterativo mostrou-se sensível aos ruídos presentes nas cores e um efeito indesejável de convergência, que gera imagens somente com cores cinza, foi observado. Este efeito foi minimizado aumentando-se a área uniforme contendo as regiões coloridas do *ColorChecker*. A cor utilizada no cálculo da matriz de mapeamento de cores passou a ser a cor média dessas áreas. Com uma área de 17×17 pixels o resultado da correção foi muito próximo da correção direta, mas uma pequena convergência ainda podia se observada.

Simão et al. (2015) apresentaram um método iterativo com filtragem temporal para alcançar a correção em vídeo de ambiente externo sob variação da iluminação, mas sem movimento relativo entre a câmera e a cena. Neste trabalho não há a presença do *ColorChecker* e as cores utilizadas para o cálculo da matriz são os valores médios de regiões quadradas e uniformes localizadas no primeiro quadro. Os valores de cor obtidos do primeiro quadro são utilizados como referência inicial. Essa referência é reconstruída a cada iteração por meio de um filtro temporal. Este filtro calcula os novos valores de referência pela média temporal dos valores corrigidos nas iterações anteriores.

2.3.5 Taxonomia das abordagens para constância de cor

As Tabelas 2.1 e 2.2 organizam os algoritmos separando as publicações por similaridades. Na Tabela 2.1 estes algoritmos são classificados pela origem da informação (uma imagem, duas imagens, sequências de imagens e vídeos) e pela informação adicional necessária para se alcançar a constância de cor. Esta informação adicional pode ser estimada (cor do iluminante da cena) ou conhecida de antemão (cores presentes na cena, funções base para iluminantes e refletâncias, curvas espectrais dos sensores e curvas espectrais de refletâncias).

A constância de cor em sequências de imagens está separada em três categorias: (i) direta, quando os métodos de constância de cor são aplicados individualmente em cada quadro. (ii) com correlação, quando informações correlatas entre dois ou mais quadros são utilizadas para se alcançar a constância de cor e (iii) iterativo, quando a correção de um quadro anterior produz informação para correção do quadro seguinte.

As sequências de imagens não são classificadas em relação à taxa de captura. No entanto, deve-se ter em mente que sequências adquiridas com baixa taxa de captura, dependendo da velocidade dos movimentos, não terão informações que coexistam em quadros consecutivos impedido o uso de técnicas que tiram proveito de informações correlatas. Por outro lado, dependendo das considerações feitas é possível obter informações correlatas mesmo em baixas

taxas de captura. Este é o caso do método apresentado por Wang et al. (2011) em sequência de imagens capturadas com uma taxa de 3 quadros por segundo. Ele considera que a iluminação varia muito lentamente e utiliza a informação correlata de cor dos iluminantes de quadros consecutivos.

Na Tabela 2.2, estes algoritmos são agrupados, no eixo vertical, segundo o mapeamento adotado que pode ser uma matriz diagonal ou uma transformação linear 3x3. No eixo horizontal eles estão organizados quanto às restrições consideradas relativas aos sensores, às refletâncias, aos iluminantes, às geometrias dos objetos e à presença de cores padronizadas na cena. A ausência de sensores banda estreita considera que os sensores são banda larga. Esta regra se aplica também para reflexão difusa/especular, iluminação uniforme/não uniforme, com restrições geométricas/sem restrições geométricas e cores conhecidas presentes/cores conhecidas ausentes.

2.4 Conclusão

Neste capítulo foi apresentada uma descrição básica da fisiologia do sistema visual humano e como ela serviu de inspiração para o modelo dos sensores que equipam as câmeras. O fenômeno da constância de cor no sistema visual humano foi identificado através dos experimentos de Land motivando, assim, a busca por uma solução para os sistemas de visão computacional. Também foram apresentadas as principais abordagens para o problema da constância de cor baseado em uma única imagem e os poucos trabalhos que utilizam informações correlatas em sequências de imagens. Os trabalhos desenvolvidos aqui introduzem a abordagem apresentada nessa tese no contexto de sequências de imagens. Por fim, uma taxonomia para abordagens para constância de cor é apresentada e estruturada na forma de tabelas.

Tabela 2.1: Informações de entrada e informações adicionais em algoritmos de constância de cor.

adicionais nº imagens		Uma cor estimada	Três cores conhecidas	Mais que 3 cores conhecidas	Funções base (Iluminante e Refletância), Curvas dos sensores e curva espectral das refletâncias
Uma imagem		(Land e McCann, 1971) (Land, 1977) (Gijssenij e Gevers, 2007) (Ebner, 2009) (van de Weijer et al., 2007a) (Gijssenij et al., 2009) (Cardei et al., 2002) (Lee, 1986) (Toro e Funt, 2007) (Forsyth, 1990) (Finlayson, 1996) (Finlayson e Hordley, 1999) (Funt e Xiong, 2004) (Agarwal et al., 2007) (Finlayson et al., 2001) (Sapiro, 1999) (Gijssenij et al., 2012a) (Vazquez-Corral et al., 2012) (Bianco e Schettini, 2011) (Gijssenij e Gevers, 2011) (van de Weijer et al., 2007b) (Gijssenij et al., 2012b)	(Bril, 1980)		(Funt et al., 1991) (Brill, 1978) (Buchsbaum, 1980) (Maloney e Wandell, 1986) (D’Zmura e Lennie, 1986) (D’Zmura e Iverson, 1993a) (D’Zmura e Iverson, 1993b) (D’Zmura e Iverson, 1994) (Gershon et al., 1987) (Forsyth, 1990) (Bajcsy et al., 1989)
Duas imagens		(Fischer e Sajjaa, 2008)			(Tsukada e Ohta, 1990) (Fredembach e Finlayson, 2008) (Finlayson et al., 2005)
Seq. imagens	Direto	(Renno et al., 2005)	(Almonfrey et al., 2011) (Simão et al., 2013)		
	Correlação	(Wang et al., 2011) Prinet et al. (2013) Garud et al. (2014)	(Simão et al., 2014) (Simão et al., 2015)		
	Iterativo		ESTA PROPOSTA (sem ColorChecker)		

Tabela 2.2: Restrições e a forma de correção de cores nos algoritmos de constância de cor.

Modelo de transformação	Restrições consideradas	Sensores banda estreita	Iluminante base 3	Refletância base 2	Refletância base 3	Gray-World	White Pach ou Max RGB	Gray-Edge	Reflexão difusa	Iluminação uniforme	Com restrições geométricas	Cores conhecidas presentes na cena
	Trabalho											
Matriz diagonal	(Land e Mccann, 1971)	X					X		X	X	X	
	(Land, 1977)	X					X		X	X	X	
	(Ebner, 2009)	X					X		X	X	X	
	(Lee, 1986)	X								X		
	(Forsyth, 1990)	X							X	X	X	
	(Cardei et al., 2002)	X								X	X	
	(Ebner, 2004)	X				X			X		X	
	(van de Weijer et al., 2007a)	X						X	X	X	X	
	Gijsenij et al. (2009)	X						X	X	X	X	
	(Renno et al., 2005)	X				X			X	X	X	
	(Wang et al., 2011)	X							X	X	X	
	(Fischer e Sajjaa, 2008)	X								X	X	
	(Gijsenij e Gevers, 2007)	X				X			X	X	X	
	(Gijsenij et al., 2009)	X						X	X	X	X	
	(Toro e Funt, 2007)	X								X	X	
	(Finlayson, 1996)	X							X	X	X	
	(Finlayson e Hordley, 1999)	X							X	X	X	
	(Funt e Xiong, 2004)	X							X	X	X	
	(Agarwal et al., 2007)	X							X	X	X	
	(Finlayson et al., 2001)	X							X	X	X	
	(Sapiro, 1999)	X							X	X	X	
	(Gijsenij et al., 2012a)	X						X	X	X	X	
	(Vazquez-Corral et al., 2012)	X							X	X	X	
	(Gijsenij e Gevers, 2011)	X							X	X	X	
	(van de Weijer et al., 2007b)	X							X	X	X	
	(Gijsenij et al., 2012b)	X							X		X	
Transformação linear 3x3	(Buchsbaum, 1980)		X		X	X			X	X	X	
	Maloney e Wandell (1986)		X	X		X			X	X	X	
	(Gershon et al., 1987)		X		X	X			X	X	X	
	(Bril, 1980)				X				X	X	X	X
	(Almonfrey et al., 2011)				X				X	X	X	X
	(Tsukada e Ohta, 1990)		X		X				X	X	X	X
	(Fredembach e Finlayson, 2008)	X	X		X				X	X	X	
	(Novak e Shafer, 1991)		X		X				X	X	X	X
	(Bajcsy et al., 1989)		X		X				X	X	X	
	(D'Zmura e Lennie, 1986)		X		X				X	X	X	X
	(Brill, 1978)		X		X				X	X	X	X
	(Funt et al., 1991)		X		X				X	X	X	
	(Finlayson et al., 2005)		X		X				X	X	X	
ESTA PROPOSTA				X				X	X	X		

Capítulo 3

Métodos Básicos Desenvolvidos para Sequências de Imagens

3.1 Introdução

Na Seção 1.3, Equação 1.3, a transformação T_i , que relaciona as cores entre duas imagens consecutivas, foi caracterizada como a estrutura básica responsável por gerar as imagens corrigidas em uma sequência. A obtenção desta transformação é feita a partir de amostras de cores escolhidas nas duas imagens. Estas amostras, provenientes de cada uma das imagens, são obtidas a partir das mesmas superfícies observadas que localizadas por um esquema de *tracking*, sendo portanto correlatas. O espaço de cores de cada imagem pode se dividido entre as cores utilizadas para obtenção da transformação e as demais cores restantes. Assim, a transformação que se busca deve ser capaz de relacionar, com precisão suficiente, tanto as cores que foram utilizadas na sua obtenção como também as demais cores das imagens.

Uma vez que a transformação tenha sido definida, ela fará parte de um processo iterativo para correção de cores. Neste processo, as cores corrigidas em uma iteração juntamente com as cores provenientes da nova imagem capturadas são utilizadas para obtenção da transformação da iteração seguinte. Esta nova transformação gera uma nova imagem corrigida de onde um novo conjunto de cores é tomado repetindo todo o processo. Portanto, a transformação T_i poder escrita como na Equação 3.1.

$$T_i = \Psi(\hat{\mathbf{C}}_i, \hat{\mathbf{C}}_{i-1}) \quad (3.1)$$

A implementação deste processo iterativo depende de aspectos relacionados com os conteúdos das sequências de imagens e com a forma como as informações de cores são utiliza-

das. Isto permite o desenvolvimento de diversos métodos de constância de cor. Dentre eles, dois métodos básicos são apresentados e avaliados neste capítulo.

Para a avaliação destes métodos, duas sequências de imagens são utilizadas para estruturar os testes a serem executados. A primeira sequência possui regiões coloridas que estão em posições fixas em relação à câmera. As imagens são capturadas sob uma grande variação da iluminação. O objetivo é verificar a qualidade das correções das cores desassociadas do problema de localização. Por outro lado, a segunda sequência é obtida sem variações significativas da iluminação ambiente, mas com a câmera em movimento. Neste caso, o objetivo é verificar a estabilidade das correções enquanto o *tracking* das regiões é executado. A definição da transformação T_i , o processo iterativo e a definição das sequências de imagens para testes estão detalhadas nas seções seguintes.

3.1.1 A transformação de cores

Utilizando o modelo linear de dimensão finita, a refletância das superfícies observadas nas imagens pode ser representada por um pequeno número de funções base, normalmente entre 3 e 5, fornecendo uma boa aproximação espectral para a refletância da maioria das superfícies (Ebner, 2007; Buchsbaum, 1980; Maloney e Wandell, 1986; D’Zmura e Lennie, 1986). Assim, a refletância $R(\lambda)$ de cada superfície pode ser escrita como na Equação 3.2.

$$R(\lambda) = \sum_{n=1}^k r_n \hat{\mathbf{R}}_n(\lambda) \quad (3.2)$$

Nesta equação, $\hat{\mathbf{R}}_n(\lambda)$ é o vetor representando as funções base, k é a quantidade de funções e r_n são os seus coeficientes. Substituindo a Equação 3.2 na Equação 2.1 e trazendo o somatório para fora da integral obtém-se a Equação 3.3.

$$c_m(x_I) = \sum_{n=1}^k \int_{\lambda_1}^{\lambda_2} S_m(\lambda) L(\lambda) r_n(x_O) \hat{\mathbf{R}}_n(\lambda) d\lambda, \quad (3.3)$$

Nesta equação, $c_m(x_I)$ é o valor medido na posição x_I da imagem para $m \in \{r, g, b\}$ e $r_n(x_O)$ são os coeficientes que representam a refletância da superfície no ponto x_O do objeto.

Para que se tenha um sistema de equações determinado, o número de funções base de refletância, k , deve ser igual ao número de canais de cor, m . Assumindo o modelo tricromático de sensores, em que $m \in \{r, g, b\}$, o modelo de refletâncias deve possuir três funções base,

$k = 3$. Assim, a Equação 3.3 passa a ser escrita na forma matricial conforme apresenta a Equação 3.4.

$$\mathbf{C}_{3 \times 1} = \mathbf{A}_{3 \times 3} \mathbf{R}_{3 \times 1} \quad (3.4)$$

Nesta equação, $\mathbf{C}_{3 \times 1}$ é o vetor com os valores medidos das componentes de uma dada cor, $\mathbf{R}_{3 \times 1}$ é o vetor de coeficientes das funções base da refletância correspondente e $\mathbf{A}_{3 \times 3}$ é a matriz que relaciona linearmente a refletância à cor (Ebner, 2007). Os elementos a_{mn} da matriz \mathbf{A} são dependentes do iluminante, do sensor e das funções base das refletâncias, podendo ser escritos como mostrado na Equação 3.5.

$$a_{mn} = \int_{\lambda_1}^{\lambda_2} S_m(\lambda) L(\lambda) \hat{\mathbf{R}}_n(\lambda) d\lambda \quad (3.5)$$

Utilizando os elementos da Equação 3.5 na Equação 3.4 pode ser obter uma versão expandida como mostrado na Equação 3.6.

$$\begin{bmatrix} c_r \\ c_g \\ c_b \end{bmatrix} = \begin{bmatrix} \int_{\lambda_1}^{\lambda_2} S_r(\lambda) L(\lambda) \hat{\mathbf{R}}_1(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_r(\lambda) L(\lambda) \hat{\mathbf{R}}_2(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_r(\lambda) L(\lambda) \hat{\mathbf{R}}_3(\lambda) d\lambda \\ \int_{\lambda_1}^{\lambda_2} S_g(\lambda) L(\lambda) \hat{\mathbf{R}}_1(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_g(\lambda) L(\lambda) \hat{\mathbf{R}}_2(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_g(\lambda) L(\lambda) \hat{\mathbf{R}}_3(\lambda) d\lambda \\ \int_{\lambda_1}^{\lambda_2} S_b(\lambda) L(\lambda) \hat{\mathbf{R}}_1(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_b(\lambda) L(\lambda) \hat{\mathbf{R}}_2(\lambda) d\lambda & \int_{\lambda_1}^{\lambda_2} S_b(\lambda) L(\lambda) \hat{\mathbf{R}}_3(\lambda) d\lambda \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \quad (3.6)$$

Se a curva espectral do iluminante $L(\lambda)$ e as curvas de sensibilidade dos sensores $S_m(\lambda)$ forem conhecidas, a refletância de uma superfície pode ser estimada a partir dos coeficientes r_n da matriz \mathbf{R} da Equação 3.7.

$$\mathbf{R} = \mathbf{A}^{-1} \mathbf{C} \quad (3.7)$$

Dadas duas imagens, \mathbf{I} e \mathbf{I}' , de uma mesma cena, capturadas sob iluminação de dois iluminantes distintos, $L(\lambda)$ e $L'(\lambda)$, pode-se definir dois conjuntos, \mathbf{C} e \mathbf{C}' , com todas as cores presentes nas imagens. Estes dois conjuntos de cores podem ser expressos na forma da Equação 3.4, ou seja, $\mathbf{C} = \mathbf{A}(L(\lambda)) \mathbf{R}$ e $\mathbf{C}' = \mathbf{A}(L'(\lambda)) \mathbf{R}$. Uma vez que estes dois conjuntos de cores são gerados a partir de um mesmo conjunto de refletâncias \mathbf{R} , suas cores se correlacionam pela Equação 3.8.

$$\mathbf{C} = \mathbf{A}(L(\lambda)) \mathbf{A}^{-1}(L'(\lambda)) \mathbf{C}' \quad (3.8)$$

Assim, a transformação que correlaciona as cores de duas imagens é uma matriz de transformação $\mathbf{T} = \mathbf{A}(L(\lambda)) \mathbf{A}^{-1}(L'(\lambda))$. A correlação entre os dois conjuntos de cores pode

ser expressa na forma matricial $\mathbf{C} = \mathbf{T}\mathbf{C}'$. Esta matriz \mathbf{T} de transformação é chamada aqui de matriz de mapeamento de cores.

A matriz de mapeamento de cores \mathbf{T} relaciona todas as cores presentes nas duas imagens, mas pode ser determinada por um número menor de amostras. Para o caso particular de três amostras de cores, esta matriz pode ser obtida a partir dos conjuntos $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3]$ e $\mathbf{C}' = [\mathbf{c}'_1, \mathbf{c}'_2, \mathbf{c}'_3]$. Estes conjuntos são tomados sob uma iluminação padrão, $L(\lambda)$, e uma iluminação desconhecida, $L'(\lambda)$, respectivamente. A matriz de mapeamento de cores é dada pela Equação 3.9.

$$\mathbf{T} = [\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3][\mathbf{c}'_1, \mathbf{c}'_2, \mathbf{c}'_3]^{-1} \quad (3.9)$$

Para minimizar os efeitos dos ruídos, das distorções das lentes, da aproximação no modelo das refletâncias e etc, um número maior, p , de amostras pode ser utilizado. Neste caso, a matriz \mathbf{T} deverá ser obtida por um processo de otimização, por exemplo, utilizando mínimos quadrados. Este processo pode ser alcançado calculando-se a pseudo-inversa (Ben-Israel e Greville, 2003) dada pela Equação 3.10 quando a matriz \mathbf{W} possui linhas linearmente independentes e pela Equação 3.11 quando a matriz possui colunas linearmente independentes.

$$\mathbf{W}^+ = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \quad (3.10)$$

$$\mathbf{W}^+ = \mathbf{W}^T (\mathbf{W} \mathbf{W}^T)^{-1} \quad (3.11)$$

Nesta equação, \mathbf{W}^+ é a pseudo-inversa da matriz \mathbf{W} . Assim, a matriz \mathbf{T} , com dimensão 3×3 , pode ser obtida pela Equação 3.12 para p de amostras.

$$\mathbf{T} = [\mathbf{c}_1, \dots, \mathbf{c}_p][\mathbf{c}'_1, \dots, \mathbf{c}'_p]^+ \quad (3.12)$$

Multiplicando a imagem \mathbf{I}' , adquirida sob uma iluminação desconhecida, $L'(\lambda)$, pela matriz de mapeamento de cores, \mathbf{T} , obtém-se uma imagem, $\hat{\mathbf{I}}$, com cores corrigidas como se tivessem sido capturadas sob a mesma iluminação, $L(\lambda)$, da imagem \mathbf{I} . Isto pode ser escrito na forma $\hat{\mathbf{I}} = \mathbf{T}\mathbf{I}'$.

3.1.2 O processo iterativo

O processo iterativo utiliza as cores corrigidas em uma iteração para determinar a transformação a ser utilizada na iteração seguinte. Os valores das cores são obtidos de regiões

identificadas nas imagens da sequência. Estas regiões são relativas aos objetos presentes na cena, que podem estar parados ou em movimento em relação aos outros objetos e à câmera. Assim, considerando o movimento relativo entre a câmera, a cena e os objetos da cena, cinco situações principais podem ser caracterizadas:

- Estático entre câmera, cena e objetos: Nesta situação não há movimento relativo entre a câmera, a cena e os objetos da cena. Este é o caso, por exemplo, de sequências de imagens de um pátio de cargas durante um intervalo de tempo em que não há operações de carga e descarga. Há apenas a variação da iluminação que incide sobre a cena.
- Movimento da câmera e estático entre cena e objetos: No caso de objetos estáticos em relação à cena, todo o contexto das imagens capturadas apresenta um movimento relativo à câmera quando esta se movimenta. Por exemplo, uma sequência de imagens de um corredor obtida com a câmera presa a um robô móvel enquanto ele se movimenta.
- Movimento da câmera e estático entre câmera e objetos: Quando alguns objetos aparecem fixos em relação à câmera enquanto a cena se modifica ao fundo. É o caso de algumas imagens geradas para avaliação de algoritmos de constância de cor, nas quais uma referência¹ de cor é fixada à câmera. Esta referência permite avaliar a correção de cor alcançada.
- Movimento de objetos e estático entre câmera e cena: Neste caso não há movimento relativo entre câmera e cena, mas um objeto distinto se movimenta em alguma direção. Um exemplo para este caso é um pátio de estacionamento no momento em que um carro esteja entrando. Nele toda a cena é estática e apenas o carro se movimenta.
- Movimento entre câmera, cena e objetos: Neste último caso há movimento relativo entre a câmera, a cena e os objetos. É o caso de sequências de imagens geradas de dentro de um veículo em movimento enquanto um pedestre atravessa a avenida.

As duas primeiras situações são de interesse particular nesta tese. A primeira, por não apresentar movimento relativo entre a câmera, a cena e os objetos da cena, constitui uma situação de menor complexidade. Isto permite que os efeitos da variação da iluminação e do ruído na imagem possam ser isolados e estudados mais facilmente. A segunda, por se enquadrar no contexto de um robô móvel que se desloca por um caminho orientado por marcas coloridas sujeito à variação de iluminação como visto na Figura ???. Assim, estas duas situações são avaliadas nesta tese. Os demais casos, pela complexidade que apresentam, estão fora da proposta dessa tese.

¹Normalmente essa referência é uma bola cinza pequena que aparece em algum canto da imagem

Outro aspecto importante do processo iterativo é quanto ao uso das cores corrigidas. Os valores dessas cores podem ser utilizados diretamente na obtenção da transformação T ou, ainda, serem modificados em alguma etapa de adequação antes da utilização. Estas duas opções caracterizam dois métodos básicos distintos desenvolvidos nesta tese.

Além destes dois métodos, um método de correção que não se caracteriza como processo iterativo também é utilizado para efeito de comparação. Ele é utilizado apenas na situação em que não há movimento relativo entre a câmera, a cena e os objetos da cena, pois depende de que sempre as mesmas regiões sejam identificadas. Neste método, a correção de cores utiliza sempre as mesmas cores de referência como, por exemplo, as cores do primeiro quadro. Simão et al. (2013) identificaram que este método apresenta um desempenho melhor na estabilidade do matiz de um grupo de cores quando comparado com os métodos *Gray-World*, *Max-RGB* e *Gray-Edge*. Este método é referenciado, aqui, como Método Simples e serve como *ground truth*. Apesar de apresentar bons resultados, este método não resolve o problema da constância de cor nas situações em que há movimento relativo entre a câmera e a cena, pois, neste caso, não há referências fixas.

O algoritmo do Método Simples está descrito abaixo. Nele, o cálculo da matriz \mathbf{T} depende da obtenção da pseudo inversa \mathbf{C}'^+ que pode ser feita, por exemplo, pela decomposição de valores singulares. Uma alternativa é a obtenção da matriz \mathbf{T} pela solução de mínimos quadrados para o sistema linear de equações $\mathbf{C}' * \mathbf{T} = \mathbf{C}$.

Algoritmo *Método Simples*

Entrada: Uma lista com as posições fixas das regiões coloridas L_{xy}

Entrada: Uma sequência de imagens V_e

Saída: Uma sequência de imagens V_s

1. Gere uma lista C contendo as cores de referência em formato RGB usando as posições fixas L_{xy} das regiões coloridas
2. $V_s \leftarrow \emptyset$
3. **para cada** $I' \in V_e$
4. **faça** C' recebe os valores médios RGB das regiões coloridas definidas em posições fixas, L_{xy} , da imagem I'
5. $T \leftarrow C * C'^+$; calcula a matriz de transformação usando pseudo-inversa
6. $I \leftarrow T * I'$; aplica a matriz de transformação a toda a imagem
7. insere o quadro corrigido I na sequência V_s
8. avança para a próximo quadro da sequência V_e
9. **retorna** V_s

3.1.3 Sequências de imagens para testes

As sequências de imagens para avaliação dos métodos básicos foram geradas por uma câmera fotográfica digital SONY DSC-H9 e por uma *webcam* Logitech C920 Pro. Estas câmeras possuem ajustes automáticos de balanço de branco e controle de exposição. Estes dois ajustes foram desabilitados durante a captura das imagens.

O balanço de branco automático foi desabilitado para se evitar a interferência com os métodos de constância de cor estudados. Por outro lado, quando uma sequência de imagens é gerada com a câmera em movimento, regiões muito claras ou muito escuras podem aparecer na cena e alterar o brilho médio da imagem. Nesta situação, o controle automático de exposição, buscando manter o brilho médio, altera o brilho de todo o conteúdo da cena. Assim, este controle também foi desativado.

Outro ajuste removido foi a correção gama. Esta correção é uma função exponencial que ajusta a intensidade do sinal de vídeo para exibição em tubos de raios catódicos². Estes tubos possuem uma relação não linear entre o sinal de entrada e o brilho gerado na tela. Para compensar esta não linearidade, uma função inversa (correção gama) é aplicada durante a geração das sequências de imagens.

As câmeras utilizadas na geração das sequências não possuem um recurso para desabilitar a aplicação da correção gama durante a captura. A remoção da correção gama é executada em uma etapa de pré-processamento das imagens antes de serem utilizadas nos métodos de correção de cores.

A primeira sequência de imagens foi gerada com um cartão com placas coloridas, do tipo *ColorChecker*, mantido fixo na posição central das imagens capturadas (Figura 3.1). Nesta sequência o cartão permanece estático em relação à câmera enquanto, ao fundo, a cena do ambiente se movimenta. Este contexto é análogo à terceira situação, relativa ao movimento entre a câmera e a cena, descrita na Seção 3.1.2. Eventualmente, esta sequência pode ser útil para avaliar o efeito da correção nas cores de fundo.

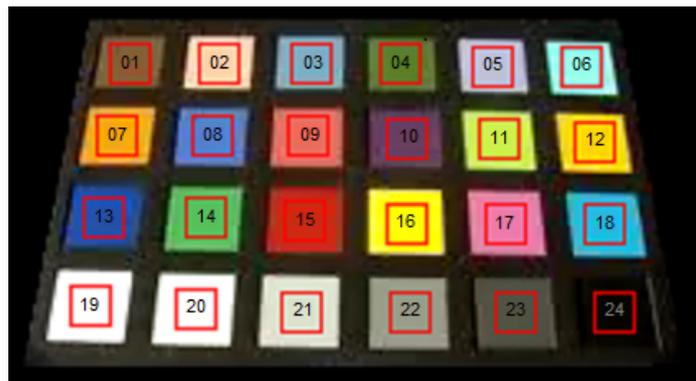
Para adequar esta sequência à primeira situação relativa ao movimento entre a câmera e a cena, uma máscara foi utilizada para isolar somente a imagem do *Colorchecker* como pode ser visto na Figura 3.2. As cores utilizadas para obtenção da matriz de mapeamento de cores são provenientes destas placas coloridas. Elas são caracterizadas pelo valor médio dos canais R, G e B de uma região quadrada centrada em cada placa do *ColorChecker*.

²Os modernos monitores, do tipo LCD, não apresentam a necessidade deste ajuste. Todavia, como a geração das sequências de imagens, por padrão, incluem a correção gama, estes monitores aplicam um ajuste para desfazer esta correção.

Figura 3.1: Primeiro quadro da primeira sequência de imagens.



Fonte: Elaborada pelo autor

Figura 3.2: *ColorChecker* com as regiões quadradas para cálculo da cor média.

Fonte: Elaborada pelo autor

O percurso para geração da sequência de imagens foi escolhido de forma a produzir uma grande variação nas cores das placas do *ColorChecker*. Esta variação deve ser consequência da variação espectral dos iluminantes. Assim, o trajeto percorrido atravessa três zonas com diferentes fontes de iluminação e, em seguida, retorna pelo mesmo caminho. A primeira zona é um ambiente externo, sujeito à iluminação natural, mas sem incidência direta da luz do Sol sobre as cores observadas. A zona seguinte abrange dois ambientes cujas iluminações são produzidas por lâmpadas fluorescentes levemente esverdeadas. A terceira zona possui iluminação fluorescente de tom avermelhado.

Além disso, optou-se para que a sequência de imagens gravadas começasse e terminasse sob iluminação do ambiente externo, ou seja, o primeiro quadro e o último estão sob iluminação do mesmo iluminante. Isto permite encadear sucessivas repetições deste vídeo simulando uma longa sequência de imagens.

A variação da iluminação que alcança o *ColorChecker* pode ser observada através da análise do comportamento da cor medida sobre a placa 19 que é uma superfície branca. Esta superfície reflete todo o espectro de cores visíveis fazendo com que a cor observada sobre ela se aproxime da cor do iluminante da cena. Assim, durante o deslocamento da câmera amostras de cor da superfície branca foram tomadas em quatro instantes distintos sob iluminação direta das fontes de iluminação. As amostras da placa branca e os respectivos valores RGB para cada um destes instantes podem ser vistos na Tabela 3.1.

Tabela 3.1: Valores RGB médios da placa 19 (branco) nos quatro ambientes.

Ambiente	R	G	B	Placa 19
Externo	255	255	255	
1º Lâmp. Fluor.	238	252	181	
2º Lâmp. Fluor.	164	175	116	
3º Lâmp. Fluor.	200	162	61	

A segunda sequência de imagens foi gerada com a câmera em movimento enquanto os objetos das cenas estão estáticos entre si. Nesta sequência foi escolhido uma iluminação uniforme e sem variações significativas. Para isso, as imagens foram capturadas sob iluminação externa em um dia nublado (Figura 3.3). Isto evita a variação da iluminação devido à alternância de nuvens e à incidência direta da luz do Sol.

Figura 3.3: Primeiro quadro da segunda sequência de imagens.



Fonte: Elaborada pelo autor

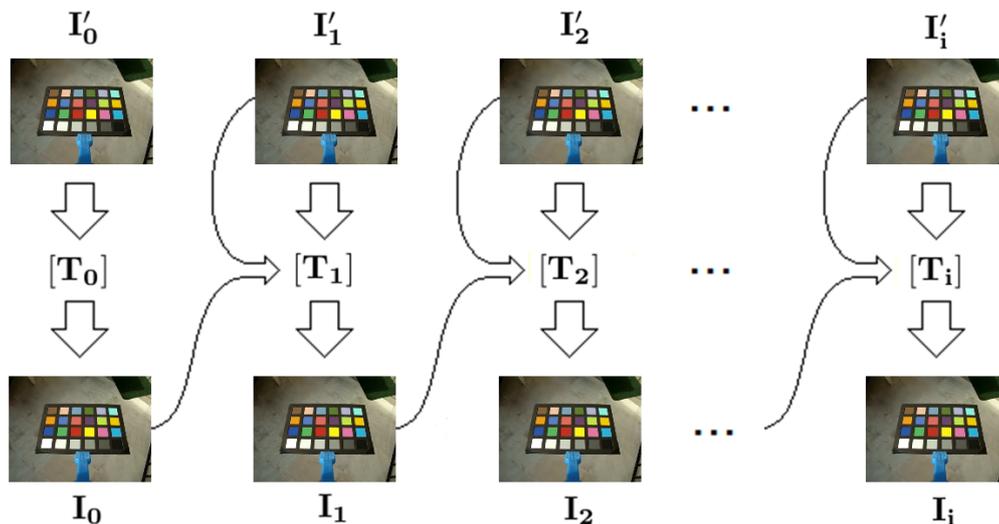
A escolha de uma sequência de imagens com iluminação constante permite avaliar os efeitos do movimento da câmera nas correções das imagens. Este movimento se traduz em

mudanças de ângulos em relação às superfícies. Este comportamento não foi considerado na determinação da transformação e por isso é importante avaliar sua influência. A primeira e a segunda sequências de imagens geradas são referenciadas nesta tese como sequência sem movimento relativo e sequência com movimento relativo, respectivamente.

3.2 Método usando Amostras de referência da Imagem Anterior (MAIA)

Introduzido por Simão et al. (2014), este método possui um processo básico de correção das cores que consiste na multiplicação de cada imagem capturada, I'_i , pela matriz de mapeamento de cores, T_i , para gerar uma imagem corrigida I_i (Figura 3.4). As imagens capturadas são consideradas como tendo sido iluminadas por um iluminante desconhecido, $L'(\lambda)$, e as imagens corrigidas são consideradas como tendo cores relativas a um iluminante de referência $L(\lambda)$. Neste contexto, a imagem corrigida pode fornecer as cores de referência enquanto a imagem capturada fornece as cores sob iluminação desconhecida. Os valores dessas cores são utilizados diretamente no cálculo da matriz de mapeamento de cores.

Figura 3.4: Esquema do processo iterativo presente no MAIA.



Fonte: Elaborada pelo autor

Diferente do Método Simples, em que as cores de referência são sempre as mesmas, no MAIA estas cores são atualizadas a cada iteração. Esta estratégia, apesar de introduzir a propagação do erro, contribui para uma independência das cores do primeiro quadro. Esta independência é essencial para correção de cores em sequências com movimento relativos

entre a câmera e a cena, pois, nestes casos, as referências originais são perdidas com o passar do tempo. Este método, torna-se, então, um caso controlado para facilitação dos estudos.

Este método considera que não há movimento relativo entre a câmera, a cena e os objetos da cena. Desta forma, não é necessário fazer o *tracking* das regiões coloridas, pois elas ocupam posições fixas. Considerando p amostras obtidas de cada imagem, a cada iteração i são geradas duas matrizes de cores. A primeira matriz, $[\mathbf{c}_1, \dots, \mathbf{c}_p]_{i-1}$, contém os valores das cores de referência obtidos da imagem corrigida na iteração anterior, \mathbf{I}_{i-1} . A segunda matriz é formada com os valores das cores sob iluminação desconhecida $[\mathbf{c}'_1, \dots, \mathbf{c}'_p]_i$, obtidos da imagem capturada \mathbf{I}'_i . A matriz de mapeamento de cores \mathbf{T}_i pode ser calculada pela multiplicação da primeira matriz pela pseudo-inversa da segunda matriz conforme apresenta a Equação 3.13.

$$\mathbf{T}_i = [\mathbf{c}_1, \dots, \mathbf{c}_p]_{i-1} [\mathbf{c}'_1, \dots, \mathbf{c}'_p]_i^+ \quad (3.13)$$

O processo iterativo utiliza, então, as cores corrigidas em uma iteração como referências para correção das cores da iteração seguinte. Este processo se repete indefinidamente gerando, a cada iteração, uma nova matriz que mapeia as cores da imagem capturada nas cores da imagem de referência da iteração anterior.

A correção do primeiro quadro é feita pela matriz \mathbf{T}_0 . Esta matriz pode ser a matriz identidade ou uma matriz determinada a partir de cores padronizadas presentes na cena. No primeiro caso, o iluminante e as cores do primeiro quadro são tomados como referência para a correção dos quadros seguintes. No segundo, o iluminante de referência para os quadros seguintes é o próprio iluminante padronizado para o conjunto de cores calibradas escolhidas. Eventualmente, uma imagem com melhor qualidade pode aparecer durante a sequência de imagens e ser utilizada como referência. Para os casos de correção *offline* o método pode ser aplicado para frente e para trás.

O algoritmo do MAIA é descrito abaixo. Nele, as cores de referência iniciais podem ser os valores RGB obtidos sob uma iluminação padronizada ou os valores obtidos do primeiro quadro. O passo 6 é a operação que diferencia o MAIA do Método Simples. No MAIA, as cores de referência não permanecem as mesmas, mas são atualizadas a cada iteração pelas cores corrigidas.

Algoritmo *Método usando Amostras de referência da Imagem Anterior (MAIA)*

Entrada: Uma lista com as posições fixas das regiões coloridas L_{xy}

Entrada: Uma sequência de imagens V_e

Saída: Uma sequência de imagens V_s

1. Gere uma lista C contendo as cores de referência em formato RGB usando as posições fixas L_{xy} das regiões coloridas
2. $V_s \leftarrow \emptyset$
3. **para cada** $I' \in V_e$
4. **faça** C' recebe os valores médios RGB das regiões coloridas definidas em posições fixas, L_{xy} , da imagem I'
5. $T \leftarrow C * C'^+$; calcula a matriz de transformação usando pseudo-inversa
6. $C \leftarrow T * C'$; atualiza as cores de referência
7. $I \leftarrow T * I'$; aplica a matriz de transformação a toda a imagem
8. insere o quadro corrigido I na sequência V_s
9. avança para a próximo quadro da sequência V_e
10. **retorna** V_s

Os experimentos realizados com o MAIA utilizaram as imagens da sequência sem movimento relativo descrita na Seção 3.1.3. Neles, as amostras de cores foram tomadas das placas 01 a 18 do *ColorChecker* (Figura 3.2). As placas 19 a 24, pertencentes à escala de tons de cinza, não foram utilizadas. Suas cores não possuem os valores do matiz definidos no modelo HSV. Além disso, para valores de saturação próximos de zero, o matiz é muito sensível ao ruído. Os valores das amostras foram calculados pela média de uma região quadrada, inicialmente, de 3×3 *pixels*.

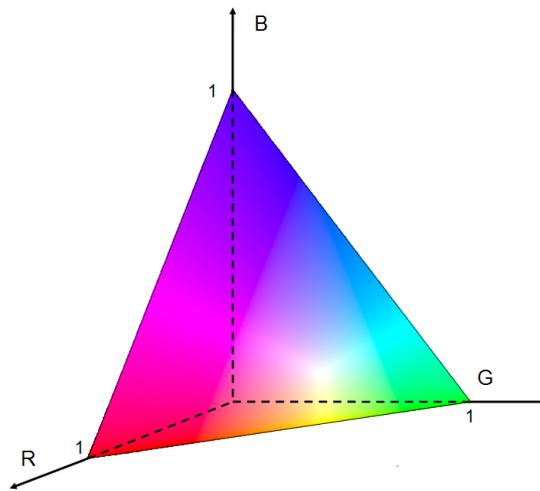
Dentre as 18 cores, algumas foram escolhidas para o cálculo da matriz de mapeamento de cores e outras para a avaliação dos resultados. Esta separação permite avaliar a qualidade da correção alcançada por uma cor que não integrou o cálculo da matriz de mapeamento de cores, T . Para a obtenção da matriz de mapeamento de cores foram escolhidas as cores mais distantes do branco. A determinação dessas cores foi feita a partir da projeção dos pontos do espaço tridimensional RGB sobre o plano unitário. Este plano intercepta os eixos cartesianos nos valores iguais a um, definindo o triângulo de cromaticidades (Figura 3.5).

A projeção é obtida pelo cálculo das componentes normalizadas de cor, dado pelas expressões

$$r = \frac{R}{R+G+B}; \quad g = \frac{G}{R+G+B}; \quad b = \frac{B}{R+G+B}.$$

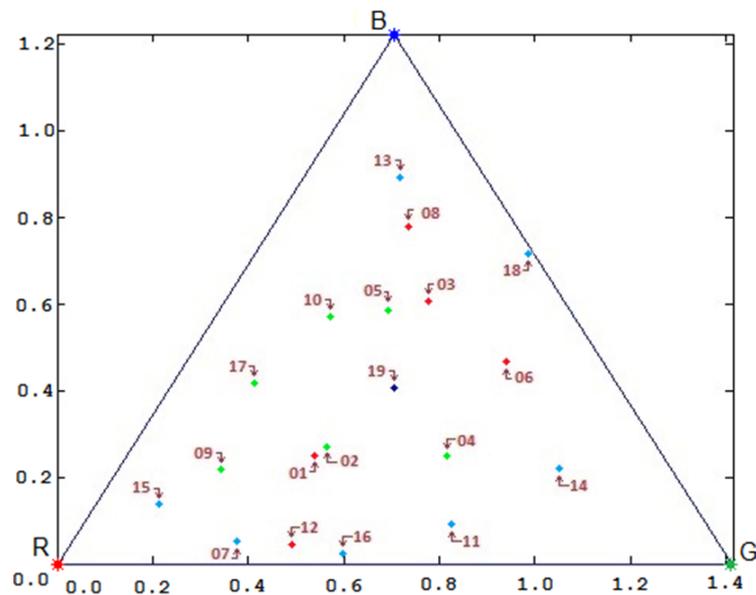
Esta normalização elimina a informação de intensidade e preserva as informações de matiz e saturação. Os valores RGB padronizados das cores 01 a 18, descontado a correção gama, foram utilizados na projeção desses pontos no triângulo de cromaticidades (Figura 3.6). Estes valores estão no sistema de cores sRGB e são referentes ao iluminante D65 padronizado pelo CIE.

Figura 3.5: Triângulo de cromaticidade no plano unitário.



Fonte: Elaborada pelo autor

Figura 3.6: Cores normalizadas plotadas no triângulo de cromaticidade.



Fonte: Elaborada pelo autor

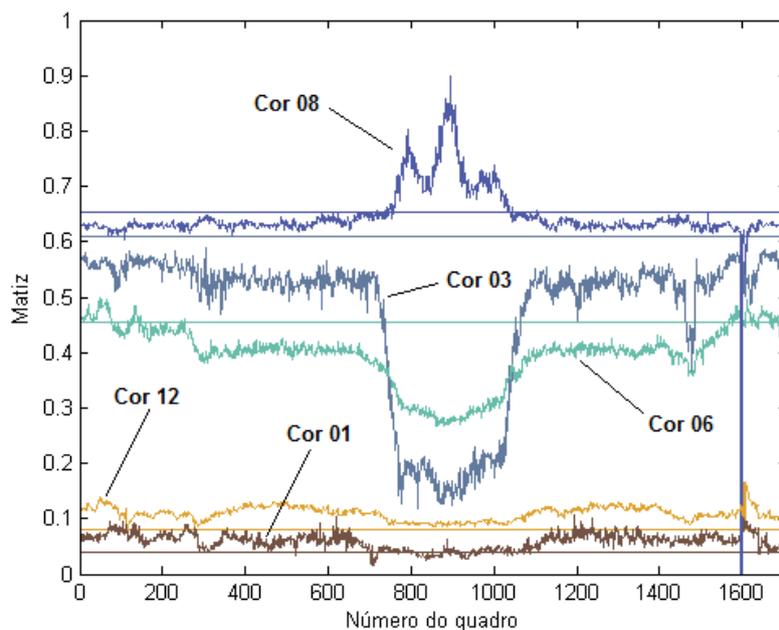
Assim, as cores escolhidas para o cálculo da matriz \mathbf{T} são 07, 11, 13, 14, 15, 16 e 18. Estas são as cores mais saturadas, pois possuem uma menor quantidade de cor branca em suas composições. A disposição dos pontos projetados, correspondentes a essas cores, forma uma envoltória convexa que cobre boa parte da área do triângulo de cromaticidades. O grupo de cores para avaliação é formado pelas cores 01, 03, 06, 08, 12 do *ColorChecker*. As cores

03, 06 e 08 apresentam grande variação do matiz em função da variação da iluminação. Por outro lado, as cores 01 e 12 apresentam pouca variação nas mesmas condições. Isto permite verificar se a correção de algumas cores piora a estabilidade de outras. A escolha dessas cores também permite um gráfico mais organizado sem grandes sobreposições das curvas. Por fim, as cores 02, 04, 05, 09, 10 e 17 não foram utilizadas nos testes.

A análise deste método é feita, inicialmente, de forma qualitativa pela observação da variação do matiz do grupo de cores de avaliação. Esta variação pode ser plotada juntamente com linhas base do matiz das cores de referência provenientes do primeiro quadro. Isto permite avaliar o comportamento do matiz das cores capturadas e das cores corrigidas.

O primeiro gráfico (Figura 3.7) contém a curva do matiz das cores provenientes das imagens capturadas.

Figura 3.7: Matiz das cores 01, 03, 06, 08 e 12 provenientes das imagens capturadas.



Fonte: Elaborada pelo autor

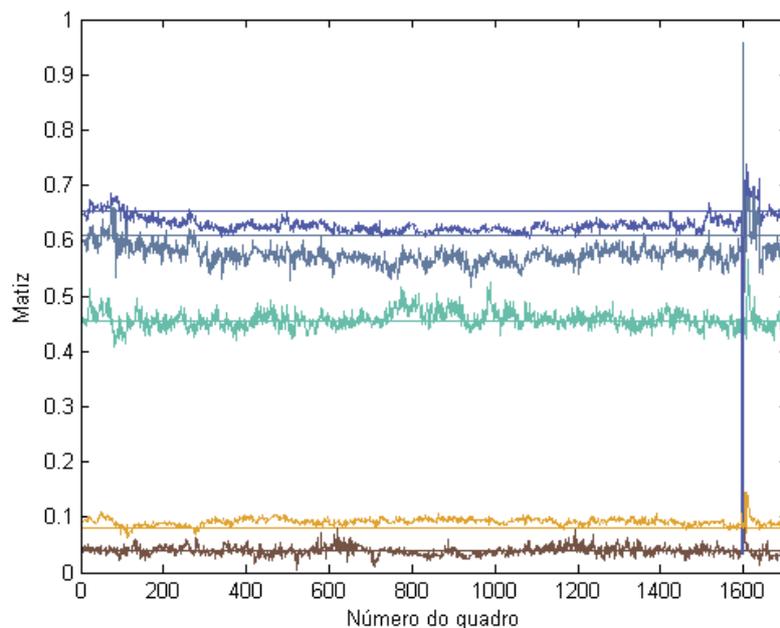
Em torno do quadro 900 há um ponto de simetria das curvas. Este é o ponto de retorno no trajeto executado que atravessa os ambientes com diferentes fontes de iluminação. A simetria existe, pois o caminho de retorno é igual ao caminho de ida mudando apenas o sentido. Entre os quadros 800 e 1000 ocorre um grande distúrbio na cor 03 causado por uma lâmpada fluorescente de tom mais avermelhado. Outro distúrbio ocorre em torno do quadro 1600 causado por uma grande variação da iluminação devido à incidência direta da luz do Sol sobre as placas do *ColorChecker*.

Neste gráfico é possível observar como a variação da iluminação altera o matiz de todas as cores observadas. As cores 03, 06 e 08 sofreram maior variação enquanto as cores 01 e 12 são mais estáveis. As primeiras, apesar de serem azuladas, refletem uma porção do vermelho e sob uma iluminação mais avermelhada, a componente R da cor se torna mais expressiva mudando o matiz observado. Por outro lado, as cores 01 e 12 são mais avermelhadas tendo pouca reflexão na região do espectro referente ao verde e ao azul e sob uma iluminação avermelhada, a componente R aumenta sua intensidade enquanto as componentes G e B têm pouca variação. Isto provoca um aumento da saturação e pouca variação no matiz.

Resultado qualitativo do Método Simples

O Método Simples é importante, pois as referências são fixas e a matriz de mapeamento de cores sempre mapeia as cores capturadas para as mesmas cores de referência. Esta correção permite verificar o quanto uma transformação linear com nove parâmetros consegue aproximar as cores capturadas das cores de referência. Vale ressaltar que este método não é iterativo e que cada quadro é corrigido individualmente sem utilizar nenhuma informação de quadros anteriores. O resultado dessa correção pode ser visto na Figura 3.8.

Figura 3.8: Matiz corrigido com referências fixas do primeiro quadro (*ground truth*).



Fonte: Elaborada pelo autor

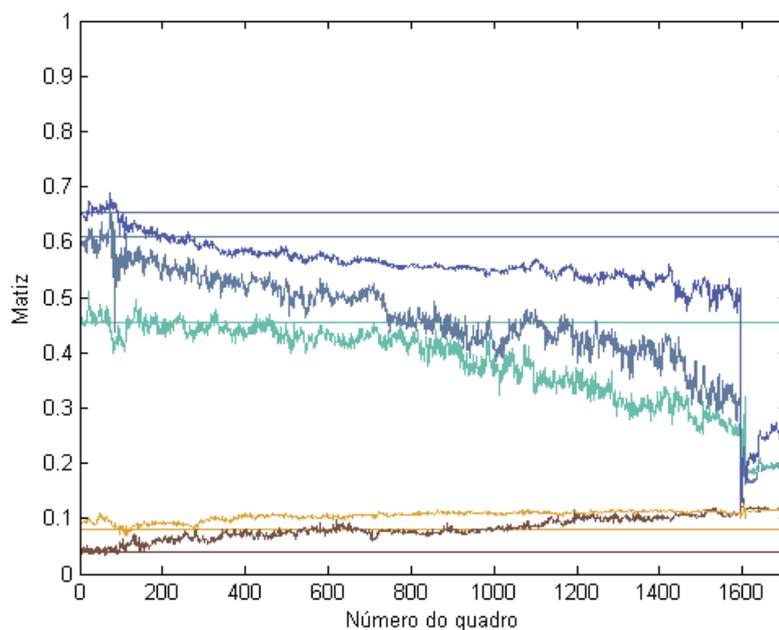
É possível observar que as cores mais estáveis, 01 e 12, tiveram uma melhora na estabili-

dade e uma redução no deslocamento em relação às linhas base de referência. Entretanto, as correções mais expressivas são das cores 03, 06 e 08. Os efeitos da variação da iluminação foram significativamente atenuados tornando as cores mais estáveis ao longo do trajeto. A cor 06 foi a que mais se aproximou da linha de referência enquanto que as cores 03 e 08 mantiveram um pequeno desvio. Assim, este resultado é tomado como base (*ground truth*) para efeitos de comparação com o MAIA.

Resultado qualitativo do MAIA

O MAIA é caracterizado como um processo iterativo que, diferentemente do Método Simples, utiliza as cores corrigidas em uma iteração como referência para a correção na iteração seguinte. O resultado deste método pode ser visto na Figura 3.9. Neste gráfico é possível observar uma indesejável convergência nas curvas do matiz das cores corrigidas. Esta convergência leva a correção para uma imagem com somente valores de cinza.

Figura 3.9: Matiz corrigido pelo MAIA para área 3x3 *pixels*.

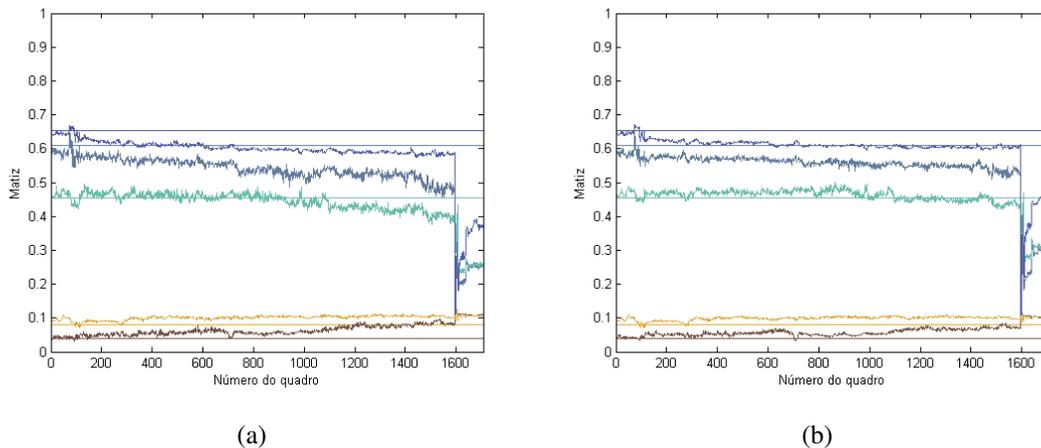


Fonte: Elaborada pelo autor

No entanto, apesar da convergência indesejável existir, há uma tendência das curvas corrigidas se aproximarem de uma reta. Isto indica que os efeitos da variação da iluminação foram suavizados e o efeito da convergência deve possuir uma outra origem. Assim, para

verificar a origem desse efeito, foi alterado o tamanho da região quadrada utilizada no cálculo da cor média. O gráfico da Figura 3.10 mostra o resultado da correção pelo MAIA para áreas de 7×7 e 13×13 *pixels*.

Figura 3.10: Matiz corrigido pelo MAIA para áreas (a) 7×7 e (b) 13×13 *pixels*.



Fonte: Elaborada pelo autor

A média espacial, utilizada no cálculo das amostras de cor, comporta-se como um filtro passa-baixa que atenua os ruídos presentes nas imagens. O aumento no tamanho das áreas aumenta a ordem do filtro, reduzindo de forma mais expressiva os ruídos. Quando a área utilizada no MAIA foi de 7×7 *pixels* houve uma redução do efeito da convergência como pode ser visto na Figura 3.10(a). Similarmente, quando a área aumentou ainda mais, passando para 13×13 *pixels*, a redução no efeito da convergência foi mais significativa (Figura 3.10(b)).

Este resultado permite inferir que a existência do efeito da convergência está associada à presença de ruído nas imagens e que o MAIA apresenta uma significativa sensibilidade a eles. Este ruído está presente nos valores dos *pixels* da imagem. Quando um destes valores se desvia do valor de referência, ele superestima a correção e, conseqüentemente, a nova referência também é superestimada desviando-se do valor médio. Assim, a próxima medida estará mais longe da referência propagando este desvio para as iterações posteriores.

A presença de ruídos nas imagens está associada à qualidade dos sensores das câmeras. Esta qualidade é dependente dos tipos de materiais e tecnologias empregados na construção dos diferentes dispositivos de captura. As câmeras destinadas às aplicações científicas são menos ruidosas, mas possuem um alto custo de aquisição. Por outro lado, as câmeras comuns existentes no mercado possuem um preço mais acessível, mas apresentam uma quantidade maior de ruído. Isto não é um grande problema quando a finalidade das imagens geradas

é a visualização pelo sistema visual humano, mas pode impactar significativamente nos resultados dos sistemas de visão computacional. Para que as câmeras de baixo custo possam ser utilizadas em aplicações como a navegação de robôs móveis, as abordagens para os problemas de visão computacional devem ser mais robustas compensando a menor qualidade desses dispositivos.

Os resultados desta análise qualitativa evidenciam a necessidade de um aperfeiçoamento do MAIA capaz de minimizar os efeitos dos ruídos. Portanto, uma análise quantitativa não será apresentada para este método uma vez que ainda precisa de melhorias.

3.3 Método do Filtro Temporal para Amostras de referência (MFTA)

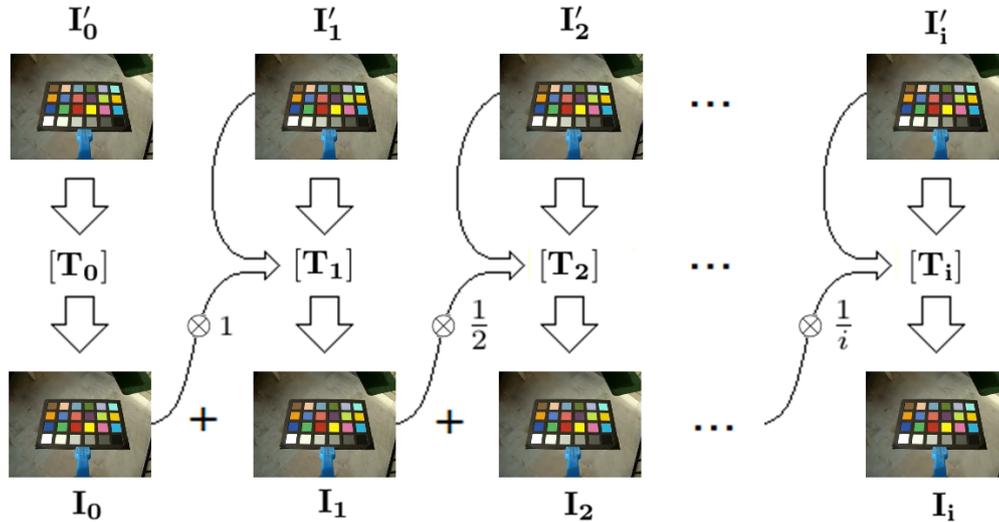
O MAIA, visto anteriormente, utiliza os valores das cores corrigidas na iteração $i - 1$ como cores de referência na iteração i . Devido aos ruídos nas imagens capturadas, as cores de referência apresentam um desvio cada vez que são geradas. Isto ocorre mesmo quando essas cores são selecionadas a partir de regiões fixas. Acontece que a filtragem espacial, aplicada quando se calcula uma cor média de uma região, não é suficiente para eliminar os ruídos presentes nas cores de referência. Assim, buscando alcançar cores de referência mais estáveis, um aprimoramento na filtragem dos ruídos deve ser implementado. Para isso, um estágio com filtragem temporal é adicionado ao processo iterativo caracterizando um novo método chamado aqui de Método do Filtro Temporal para Amostras de referência (MFTA).

A filtragem temporal é alcançada por meio de um filtro passa-baixa calculado pela média aritmética temporal das cores de referência (Simão et al., 2015) como mostrado na Figura 3.11. A cada iteração i , a imagem capturada \mathbf{I}'_i e a matriz de mapeamento de cores \mathbf{T}_i são obtidas da mesma forma como no MAIA. No entanto, as cores de referência $[\mathbf{c}_1, \dots, \mathbf{c}_k]$ são obtidas de uma forma diferente. Elas são estimadas pela média dos valores das cores de referência tomadas das imagens $\mathbf{I}_0 \dots \mathbf{I}_{i-1}$ previamente corrigidas conforme apresenta a Equação 3.14.

$$[\mathbf{c}_1 \dots \mathbf{c}_k]_i = \frac{1}{i} \sum_{n=0}^{i-1} [\mathbf{c}_1 \dots \mathbf{c}_k]_n. \quad (3.14)$$

O cálculo da média das cores de referência corresponde ao novo estágio adicionado ao processo iterativo. Ele é utilizado para alcançar uma estabilidade maior nessas cores e, conseqüentemente, obter a anulação do efeito da convergência para uma imagem com somente valores de cinza. Este cálculo está no passo 7 do algoritmo do MFTA descrito abaixo.

Figura 3.11: Esquema do processo iterativo presente no MFTA.



Fonte: Elaborada pelo autor

Algoritmo Método do Filtro Temporal para Amostras de referência (MFTA)

Entrada: Uma lista com as posições fixas das regiões coloridas L_{xy}

Entrada: Uma sequência de imagens V_e

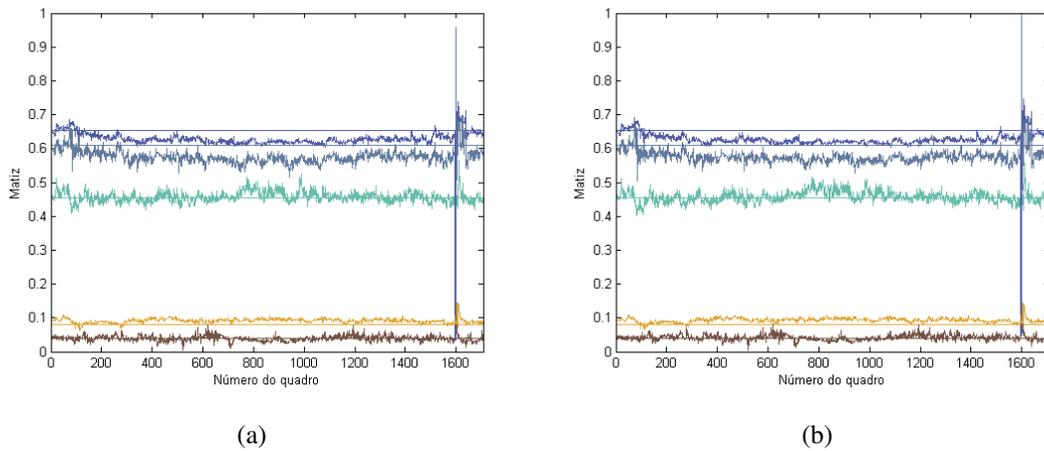
Saída: Uma sequência de imagens V_s

1. Gere uma lista C contendo as cores de referência em formato RGB usando as posições fixas L_{xy} das regiões coloridas
2. $i \leftarrow 0$
3. $V_s \leftarrow \emptyset$
4. **para cada** $I' \in V_e$
5. **faça** C' recebe os valores médios RGB das regiões coloridas definidas em posições fixas, L_{xy} , da imagem I'
6. $T \leftarrow C * C'^+$; calcula a matriz de transformação usando pseudo-inversa
7. $C \leftarrow ((C * i) + T * C') / (i + 1)$; atualiza as cores de referência utilizando média ponderada
8. $i \leftarrow i + 1$
9. $I \leftarrow T * I'$; aplica a matriz de transformação a toda a imagem
10. insere o quadro corrigido I na sequência V_s
11. avança para a próximo quadro da sequência V_e
12. **retorna** V_s

Nas Figuras 3.12, 3.13 e 3.14 é possível ver o resultado do Método Simples (*ground truth*) para áreas 3×3 , 7×7 e 13×13 *pixels* e também as correções feitas pelo MFTA, para as

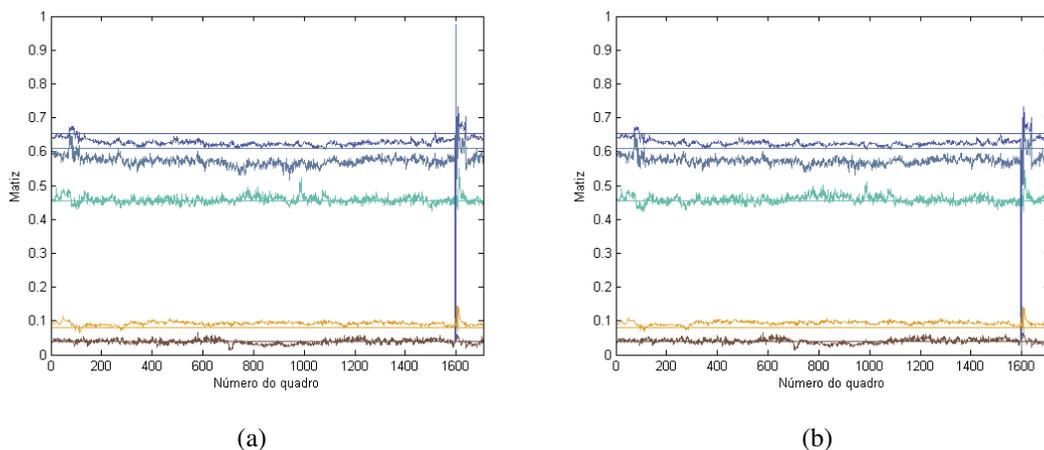
mesmas áreas, respectivamente. Estes resultados foram obtidos utilizando a sequência sem movimentos relativos definida na Seção 3.1.3. As correções feitas pelo MFTA alcançaram resultados muito próximos daquelas obtidas pelo Método Simples (*ground truth*).

Figura 3.12: Resultado do Método Simples (a) (*ground truth*) e do MFTA (b) para uma área de 3x3 *pixels*.



Fonte: Elaborada pelo autor

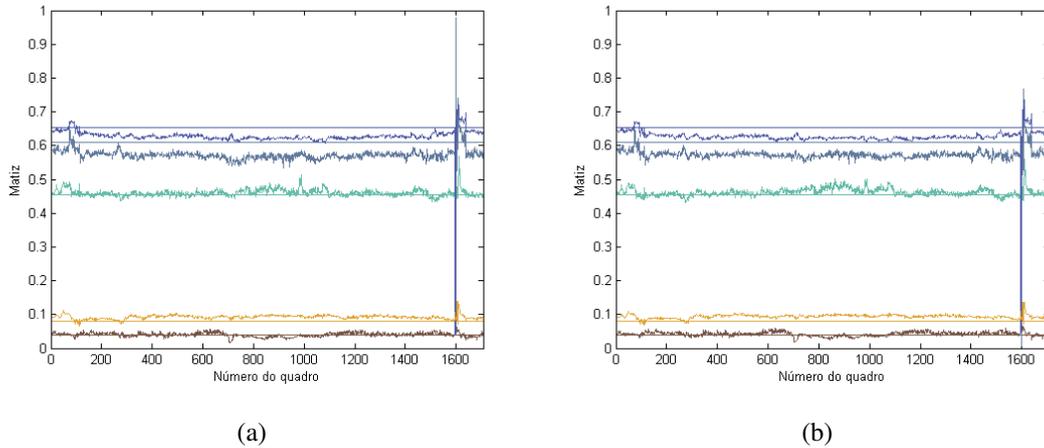
Figura 3.13: Resultado do Método Simples (a) (*ground truth*) e do MFTA (b) para uma área de 7x7 *pixels*.



Fonte: Elaborada pelo autor

A avaliação dos gráficos das Figuras 3.12, 3.13 e 3.14 permitem uma análise qualitativa dos resultados. Uma comparação mais precisa entre o MFTA e o Método Simples pode ser alcançada através de uma avaliação quantitativa usando o erro normalizado do matiz das cores corrigidas. Este erro é calculado pelo somatório da diferença absoluta entre o matiz

Figura 3.14: Resultado do Método Simples (a) (*ground truth*) e do MFTA (b) para uma área de 13x13 *pixels*.



Fonte: Elaborada pelo autor

das cores corrigidas e o matiz das linhas base em todos os quadros da sequência de imagens, dividido pelo número de quadros da sequência. O erro normalizado e_n é dado pela Equação 3.15.

$$e_n = \frac{1}{N} \sum_{i=1}^N \sum_j |\hat{h}_j - h_j| \quad (3.15)$$

Nesta equação, N é o número de quadros da sequência, i é o índice dos quadros, \hat{h}_j e h_j são o matiz das cores corrigidas e o matiz das linhas base para cada cor j , respectivamente.

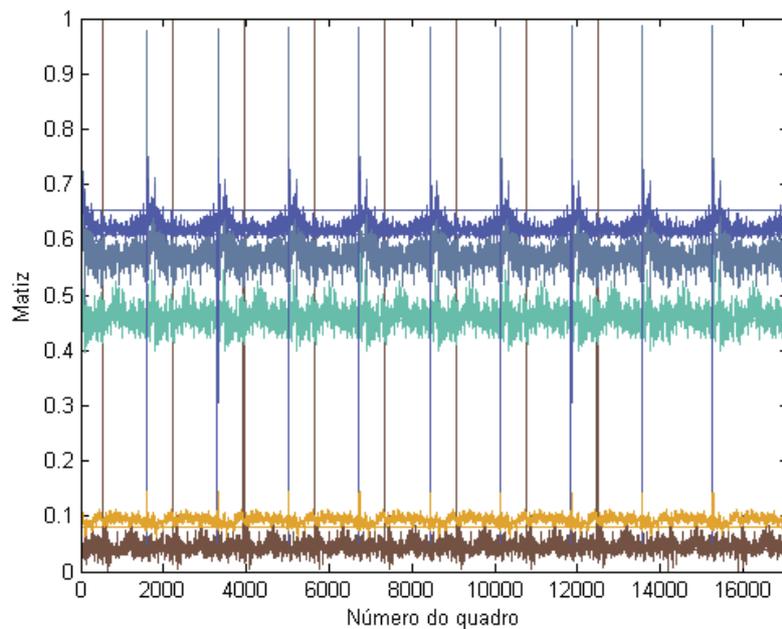
A Tabela 3.2 mostra os resultados do erro normalizado para $N = 1708$. Nesta tabela estão relacionados os valores dos erros normalizados obtidos para as imagens originais capturadas e para aquelas corrigidas pelo Método Simples e pelo MFTA. Os valores revelam que o MFTA apresenta uma correção muito próxima do Método Simples. A maior diferença entre eles é para uma área de 13x13 *pixels*. Neste caso, o MFTA apresenta um erro normalizado que é apenas 3,8% maior que o erro apresentado pelo Método Simples (*ground truth*).

Apesar da diferença entre os erros normalizados ser relativamente pequena, não fica evidente quanto o efeito da convergência foi atenuado. No estudo do MAIA foi possível observar que o efeito da convergência era mais evidente quanto menor eram as áreas das regiões coloridas e quanto maior era o número de quadros transcorridos da sequência de imagens. A Figura 3.15 mostra o resultado da correção das cores pelo MFTA utilizando uma região de apenas 1 *pixel* e sendo executado durante 10 repetições encadeadas da sequência sem movimentos relativos, totalizando 17080 quadros.

Tabela 3.2: Erros normalizados para a sequência sem movimentos relativos utilizando áreas de 3x3, 7x7 e 13x13 *pixels*

Área	Erro normalizado		
	Imagens originais	Método Simples	MFTA
3x3	0,2732	0,0903	0,0937
7x7	0,2752	0,0887	0,0913
13x13	0,2838	0,0881	0,0914

Figura 3.15: Resultado do MFTA para uma região de 1 *pixel* em 10 repetições.

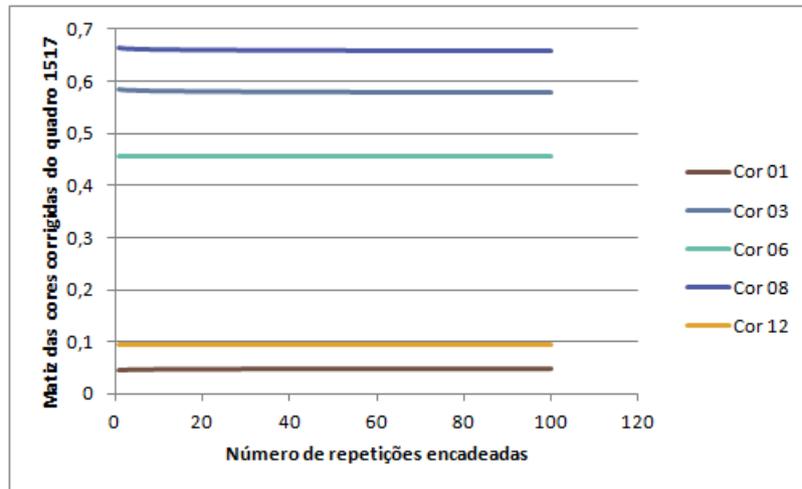


Fonte: Elaborada pelo autor

A análise qualitativa do gráfico revela que o efeito de convergência foi significativamente atenuado mesmo após 10 repetições. Isto reforça a hipótese de que o efeito da convergência é causado pelo ruído e sugere que a filtragem temporal pode ser capaz de anulá-lo. Para melhorar esta avaliação, as amostras do matiz das cores corrigidas foram registradas para o quadro de número 1517, escolhido aleatoriamente, em 100 repetições encadeadas e os seus valores plotados em um gráfico (Figura 3.16). Neste caso, as amostras corrigidas nas 100 repetições sofreram uma pequena variação sendo a cor 08 a que mais variou. O matiz dessa cor passou de 0,66542 para 0,65815 o que corresponde a uma variação de 0,728%. Também é possível observar que esta variação é mais acentuada durante as primeiras repetições e fica

menor a cada nova repetição.

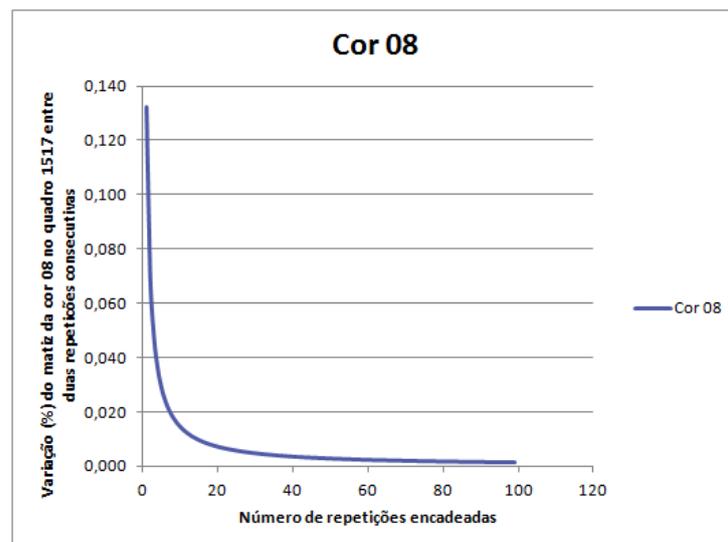
Figura 3.16: Resultado do matiz das cores corrigidas do quadro 1517 em 100 repetições encadeadas para uma região de 1 *pixel*.



Fonte: Elaborada pelo autor

A Figura 3.17 mostra o gráfico da variação percentual do matiz da cor 08 entre repetições consecutivas e evidencia que estas variações estão diminuindo sendo de apenas 0,00142% entre as duas últimas repetições.

Figura 3.17: Variação (%) do matiz da cor 08 entre duas repetições consecutivas do quadro 1517.



Fonte: Elaborada pelo autor

É possível identificar que a filtragem temporal pela média dos valores corrigidos é capaz de anular o efeito da convergência observado no MAIA. Além disso, o fato desse efeito ser anulado mesmo para uma região de 1 *pixel* significa que a filtragem espacial não é necessária para anular o efeito da convergência. A influência da filtragem espacial está apenas na dispersão dos valores corrigidos gerando curvas do matiz com mais ou menos ruídos.

3.4 Aplicação do MFTA em sequência de imagens sem referências fixas de cores

O MFTA apresentou um bom resultado na correção das cores, comparado com o Método Simples (*ground truth*), quando sequência de imagens sem movimento relativo entre a cena e a câmera são utilizadas. Esta situação é bastante favorável uma vez que não é necessário fazer o *tracking* das regiões coloridas. Isto evita os ruídos devido aos desvios de posicionamento e garante a utilização sempre das mesmas referências durante toda a sequência de imagens. No entanto, uma evolução natural é a aplicação do MFTA em sequências de imagens com movimento relativo entre a cena e a câmera. Nesta situação, as regiões coloridas estão em movimento e precisam ser rastreadas.

Para se fazer o *tracking* das regiões coloridas primeiramente elas precisam ser caracterizadas e identificadas. Assim, o primeiro quadro de uma sequência de imagens deve ser segmentado em regiões que apresentem uma uniformidade mínima de cor. A escolha de qual estratégia a ser utilizada na segmentação está baseada no resultado da comparação feita por Simão et al. (2015).

Neste trabalho, duas estratégias de segmentação são comparadas utilizando o MFTA em sequências sem movimento relativo entre a câmera e a cena. A primeira estratégia utiliza um quadrado de tamanho fixo e a segunda estratégia é chamada de *graph-based image segmentation algorithm* (Felzenszwalb e Huttenlocher, 2004). Os resultados mostraram que a segmentação utilizando a primeira estratégia apresenta um melhor desempenho em relação à segunda estratégia. Além disso, ela possui uma menor complexidade de implementação.

A segmentação por quadrados de tamanho fixo é uma rotina de varredura *pixel a pixel* da imagem. Esta varredura procura *pixels* com uma área quadrada uniforme em torno deles. O critério de uniformidade é verdadeiro quando a diferença entre os valores absolutos e a média aritmética dos canais de cor, para os *pixels* de uma região quadrada, é menor que um limiar definido. A Figura 3.18 mostra o resultado dessa segmentação para um quadrado de, por exemplo, 5x5 *pixels*.

Figura 3.18: Primeiro quadro da sequência de imagens e a imagem segmentada com quadrados de 5×5 pixels.



Fonte: Elaborada pelo autor

Uma vez que as regiões coloridas tenham sido definidas, a próxima etapa é fazer o *tracking* dessas regiões. Supondo que os deslocamentos dessas regiões sejam pequenos entre dois quadros consecutivos, este *tracking* pode ser implementado por uma simples busca na área ao redor de cada região segmentada. A busca consiste na procura de uma área que possui a maior correlação com a região colorida segmentada. Quando essa área é localizada, a posição e a cor média são registradas.

Para manter um processo iterativo, a cada iteração, a imagem capturada é segmentada e as novas áreas são incluídas no cálculo da matriz de mapeamento de cores. Simultaneamente, as áreas obtidas na segmentação do quadro anterior e que não possuem correspondentes localizadas dentro da nova imagem são retiradas desse cálculo. Outro fato importante para o processo iterativo é a escolha do tamanho das áreas quadradas. Acontece que existe uma relação de compromisso entre a redução do ruído e a diminuição da quantidade de regiões uniformes localizadas devido ao aumento da área dessas regiões. Uma vez que o tamanho das áreas quadradas tenha sido definido, ele permanece inalterado durante todo o processo de correção. Toda a estratégia de *tracking* que foi definida é inserida no algoritmo do MFTA e a versão modificada para sequências com movimento relativo entre a câmera e a cena é descrita abaixo.

Algoritmo MFTA (com tracking de regiões coloridas)

Entrada: Dimensão dos quadrados utilizados para segmentar a imagem $l = 5$

Entrada: Dimensão dos blobs utilizados para buscas nas imagens $m = 25$

Entrada: Uma sequência de imagens de entrada V_e

Saída: Uma sequência de imagens de saída V_s

1. $I' \leftarrow$ primeira imagem da sequência V_e
2. segmenta a imagem I' em quadrados de $l \times l$ pixels gerando uma lista de quadrados P contendo coordenadas centrais, cor média e um blob para cada quadrado

3. C recebe a lista de cores médias dos quadrados da lista P
4. $i \leftarrow 0$
5. insere I' na sequência V_s
6. **para outros** $I' \in V_e$
7. **faça** localize os blobs da list P na imagem I' gerando uma lista Q contendo coordenadas centrais e cor média dos quadrados localizados
8. C' recebe a lista de cores médias dos quadrados da lista Q
9. exclua das listas P e Q os quadrados sem correspondentes, os que estão fora da região de interesse, os quadrados com sobreposição e aqueles com desvio de posição e as respectivas cores médias em C e C'
10. $T \leftarrow C * C'^+$; calcula a matriz de transformação usando pseudo-inversa
11. $C \leftarrow ((C \times i) + T * C') / (i + 1)$; atualiza as cores de referência utilizando média ponderada
12. atualiza as coordenadas centrais da lista P com as correspondentes da lista Q
13. segmenta a imagem I' em quadrados de $l \times l$ pixels gerando uma lista de quadrados R
14. compara os quadrados da lista P com os quadrados da lista R , exclua os quadrados da lista P sem correspondentes na lista R e inclua quadrados novos da lista R na lista P
15. inclua em C as cores médias dos novos quadrados da lista P
16. $i \leftarrow i + 1$
17. $I \leftarrow T * I'$; aplica a matriz de transformação a toda a imagem
18. insere o quadro corrigido I na sequência V_s
19. avança para a próximo quadro da sequência V_e
20. **retorna** V_s

Assim, com esta estrutura preparada, o MFTA foi aplicado na sequência com movimentos relativos definida na Seção 3.1.3. Foi escolhido um tamanho mediano de 5×5 pixels para segmentação das regiões coloridas de forma a não comprometer a quantidade de regiões coloridas e permitir uma razoável filtragem dos ruídos. Apesar de não haver variação significativa na iluminação ambiente o resultado da correção das imagens apresenta uma saturação de cor no decorrer do tempo. Na Figura 3.19 têm-se o quadro inicial e o quadro 700 corrigidos da sequência com movimentos relativos. É possível perceber que após 700 quadros a cor vermelha da parede ficou ligeiramente mais saturada.

A região em vermelho foi a que apresentou o maior desvio durante a correção enquanto as demais regiões não apresentaram efeitos tão significativos. Assim, uma atenção especial deve ser dada a esta região buscando identificar as possíveis causas. Para isso, esta região será tomada como foco nos ensaios e experimentos sendo analisada mais detalhadamente

Figura 3.19: Quadro inicial e quadro 700 da sequência corrigidos pelo MFTA.



Fonte: Elaborada pelo autor

uma vez que mostrou ser o pior caso em toda a imagem.

O efeito da distorção da saturação é indesejável e sua origem pode estar relacionada a algumas considerações que foram feitas. Por exemplo, a influência dos ângulos de incidência e reflexão da luz, em uma superfície, foi desconsiderada. No entanto, durante o deslocamento da câmera estes ângulos estão variando enquanto é feita a captura das imagens. Assim, se a influência desses ângulos na intensidade da luz refletida for significativa, as regiões capturadas sofrerão com essas variações e conseqüentemente a obtenção da matriz de mapeamento de cores ficará comprometida. Isto pode se tornar crítico dependendo da quantidade de superfícies planas identificadas. Outra hipótese está relacionada com a composição da luz refletida. Foi assumido que as superfícies são lambertianas, ou seja, aquelas que não possuem reflexão especular. Acontece que em ambientes externos há muitas superfícies espelhadas que podem contrariar esta suposição. Por exemplo, folhas lisas apresentam este tipo de reflexão bem como vidros e outros objetos pintados com tintas brilhantes. Todas essas superfícies podem interferir nos resultados. Estes fenômenos não estão contemplados na transformação linear utilizada e que foi escolhida devido sua simplicidade de aplicação.

Além dessas considerações relativas às superfícies presentes no ambiente, existem outras relativas ao movimento e ao processo de *tracking*. Estas podem promover a inclusão de, no mínimo, duas novas fontes de ruídos. A primeira seria referente ao posicionamento das regiões coloridas. Acontece que a estratégia de *tracking*, utilizada para localizar a nova posição das regiões em cada quadro, ao sofrer o desvio de pelo menos um *pixel*, pode gerar ruídos. Isto pode acontecer mesmo que a região esteja parada. A segunda fonte de ruído pode ser devido a inclusão e exclusão de cores no cálculo da matriz de mapeamento de cores. Devido ao movimento, algumas regiões coloridas saem da cena enquanto outras regiões vão sendo acrescentadas. Isto significa que as regiões coloridas utilizadas nos cálculos não são as mesmas durante toda a sequência, pois elas possuem um tempo de permanência nas cenas. Esta situação pode causar uma polarização na correção das cores.

3.5 Conclusão

O MFTA apresenta boa estabilidade para sequências de imagem sem movimento relativo entre a cena e a câmera. Ele resolve o problema causado pela presença de ruídos no MAIA. No entanto, este bom resultado não foi observado quando sequências com movimento relativo foram utilizadas. Isto mostra que toda a dinâmica envolvendo a segmentação e o *tracking* das regiões coloridas introduz ruídos que não são atenuados pela filtragem temporal existente no MFTA. Estes ruídos afetam a qualidade das cores utilizadas na obtenção da matriz de mapeamento de cores, \mathbf{T} , comprometendo o processo iterativo.

Para se alcançar um melhoria do MFTA duas alternativas são possíveis. A primeira diz respeito a um avanço na qualidade das cores utilizadas no processo iterativo com o objetivo de reduzir os ruídos gerados durante o *tracking* das regiões coloridas. A segunda alternativa é uma evolução no processo de filtragem temporal capaz de filtrar os ruídos introduzidos no processo de segmentação e *tracking*. As duas alternativas são investigadas no próximo capítulo.

Capítulo 4

Método do Filtro Temporal para Transformações (MFTT)

4.1 Introdução

Como pôde ser observado nos resultados obtidos com o MAIA, o processo iterativo, utilizado na correção das cores das imagens de uma sequência, apresenta uma grande sensibilidade aos ruídos. Esta sensibilidade é responsável por uma indesejável convergência que gera imagens com somente valores de cinza. No entanto, quando estes ruídos são atenuados, utilizando filtros, a convergência indesejável se torna mais lenta. Este resultado pode ser observado na Seção 3.2, onde uma filtragem espacial é aplicada para obtenção das amostras das cores utilizadas no cálculo da transformação T . Esta filtragem corresponde a um filtro passa-baixa implementado pelo cálculo da cor média de uma região quadrada de $n \times n$ pixels.

Os resultados do MAIA indicaram uma redução significativa no efeito da convergência indesejável como mostrado nas Figuras 3.9, 3.10(a) e 3.10(b) para áreas de 3×3 , 7×7 e 13×13 pixels, respectivamente. No entanto, apesar da filtragem espacial produzir melhores resultados à medida que as áreas aumentam de tamanho, ela não foi suficiente para anular o efeito da convergência indesejável observado no MAIA.

Dado que no MAIA foram utilizadas sequências de imagens sem movimento relativo e que as amostras corrigidas de referência são sempre geradas para as mesmas regiões com posicionamento fixo, uma filtragem temporal pôde ser aplicada. Esta filtragem também é um filtro passa-baixa implementado pela média aritmética das amostras corrigidas nos quadros anteriores da sequência de imagens.

A inclusão do filtro temporal deu origem ao MFTA que foi capaz de estabilizar as cor-

reções das cores anulando o efeito da convergência indesejável mesmo para uma região de 1 *pixel*. Isto ficou evidente na correção de 10 repetições encadeadas da sequência sem movimento relativo (Figura 3.15). Este resultado, Seção 3.3, mostrou que a filtragem espacial não é mais necessária para anular o efeito da convergência indesejável, mas ainda pode ser utilizada para atenuar o desvio padrão das cores corrigidas.

No entanto, este bom resultado alcançado para a sequência sem movimento relativo não se verificou quando as regiões coloridas não estão em posições fixas. Este é o caso da sequência com movimento relativo entre a cena e a câmera. Para que a filtragem temporal fosse aplicada, um esquema de *tracking* das regiões coloridas precisou ser utilizado. Entretanto, devido aos desvios de localização e a inclusão e exclusão dinâmica de regiões coloridas, este esquema se torna uma fonte de ruídos que impactam na qualidade das cores utilizadas na correção.

Para que o processo iterativo apresente um bom resultado em sequências de imagens com movimento relativo entre a cena e a câmera é necessário que uma compensação do efeito dos ruídos seja implementada. No mínimo, dois caminhos são possíveis para se alcançar esta característica. O primeiro consiste em buscar a melhoria da qualidade das amostras de cores utilizadas para correção e o segundo caminho é buscar uma filtragem mais robusta. Estas duas alternativas são analisadas nas próximas seções.

4.2 Melhoria da qualidade das cores de referência

Os valores das amostras de cor, utilizados para obtenção da matriz de mapeamento de cores \mathbf{T} , foram obtidos pela média aritmética de uma região quadrada de $n \times n$ *pixels*. Entretanto, durante o processo de *tracking*, os erros de posicionamento provocam a inserção de pontos que não fazem parte originalmente da área rastreada. Esta inserção de pontos espúrios se traduz em um ruído presente nos valores das amostras de cor das regiões coloridas.

Estes pontos espúrios devem ser considerados como uma contaminação da região e não devem fazer parte do cálculo da amostra de cor. Para evitar a influência destes pontos, duas alternativas são avaliadas para excluí-los do cálculo do valor das amostras de cor. A primeira alternativa substitui a média pela mediana na caracterização da amostra de cor de uma região colorida e a segunda calcula a cor média da região colorida utilizando uma máscara binária de *pixels*. Para aproveitar a estrutura definida no MFTA, o processo de segmentação das imagens permanece o mesmo. Apenas a definição dos valores das amostras de cor é alterada.

As duas alternativas apresentadas buscam obter valores das amostras de cor com melhor qualidade e que sejam mais estáveis em relação aos ruídos.

4.2.1 Avaliação da filtragem com mediana

A primeira alternativa avaliada faz uso da mediana no lugar da média dos valores das cores dos *pixels* para caracterização da amostra de cor. A determinação da cor mediana depende de um critério de ordenação das cores dos *pixels* pertencentes a uma região colorida. Dois critérios são utilizados: o matiz das cores dos *pixels* e os valores absolutos dos canais de cor R, G e B. Neste segundo critério, a mediana é encontrada separadamente para cada canal de cor.

O primeiro critério de ordenação testado foi pelo matiz das cores dos *pixels* da região segmentada. A Figura 4.1 mostra a imagem ampliada de uma região colorida com área de 5×5 *pixels*. Esta região foi definida pelo critério de uniformidade definido na Seção 3.4 e utilizado no processo de segmentação das imagens da sequência.

Figura 4.1: Região colorida com uma área de 5×5 *pixels*.



Fonte: Elaborada pelo autor

Os valores R, G e B de cada *pixel* são convertidos para o sistema de cores HSV e os valores do matiz tabelados em ordem crescente, como no exemplo da Tabela 4.1, para obtenção da mediana. Nesta tabela é possível observar alguns valores de matiz repetidos. Isto se deve aos *pixels* presentes na região 5×5 que possuem o mesmo matiz. Assim, os valores R, G e B do *pixel* da posição H_{14} são tomados para definição do valor da amostra de cor da referida região colorida. Este procedimento é executado para cada região colorida segmentada da imagem sendo que os valores das amostras são definidos a partir de diferentes posições dentro de cada região.

Os valores espúrios de cor dos *pixels*, que eventualmente contaminarem a região colorida, ficam localizados nas extremidades do vetor ordenado. Por outro lado, os valores mais frequentes e que efetivamente caracterizam a cor da região ficam centralizados no vetor. Dessa forma, há uma tendência da mediana ser mais estável que a média aritmética, uma vez que no cálculo desta última estão incluídos todos os valores inclusive os espúrios.

Tabela 4.1: Exemplo do matiz ordenado para definição da cor mediana

Posição	Matiz
H_{16}	0,138889
H_{21}	0,140351
H_{22}	0,141667
H_{17}	0,142857
H_{11}	0,155556
H_{18}	0,159420
H_{20}	0,159420
H_{12}	0,166667
H_{19}	0,166667
H_{23}	0,166667
H_{24}	0,166667
H_{25}	0,166667
H_{14}	0,175926
H_{10}	0,175926
H_{06}	0,178571
H_{03}	0,182540
H_{01}	0,182540
H_{15}	0,182540
H_{02}	0,182540
H_{04}	0,182540
H_{05}	0,182540
H_{13}	0,184211
H_{08}	0,186275
H_{09}	0,187500
H_{07}	0,190476

O teste com o critério de ordenação pelo matiz utilizou uma área quadrada de 5×5 *pixels* como foi feito na Seção 3.4. O resultado da correção do quadro 700 da sequência com movimento relativo utilizando o valor da mediana para caracterização das amostras das cores pode ser visto na Figura 4.2. Nesta figura, é possível observar que a distorção na saturação da cor vermelha ainda está presente. Esta distorção é a mesma observada na Figura 3.19 quando as amostras de cor foram caracterizadas pela cor média da região colorida. O quadro 700 foi escolhido para avaliação por estar a uma quantidade significativa de quadros decorridos desde o início da sequência e por ser suficiente para se observar a manifestação dos efeitos dos ruídos, como visto na Seção 3.4.

O segundo critério de ordenação para obtenção da cor mediana utiliza os valores absolutos dos canais de cor R, G e B. Neste critério, os canais R, G e B dos *pixels* de cada região

Figura 4.2: Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA com mediana ordenada pelo matiz.



Fonte: Elaborada pelo autor

colorida são ordenados separadamente e os valores medianos R_{mdn} , G_{mdn} e B_{mdn} são obtidos de cada vetor ordenado. A cor $[R_{mdn} G_{mdn} B_{mdn}]$ é, então, utilizada para caracterizar o valor das amostras de cor de cada região. Deve ser observado que os canais de cor não são completamente independentes devido à sobreposição das curvas dos sensores. Assim, este critério é uma aproximação e considera sensores sem sobreposição.

O resultado do teste com este segundo critério pode ser visto na Figura 4.3. Ele apresenta uma pequena redução na saturação, mas ainda distante do desejado.

Figura 4.3: Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA com mediana ordenada pelos canais R, G e B.



Fonte: Elaborada pelo autor

Neste ponto, um questionamento relativo ao processo de segmentação se torna evidente: Qual seria o resultado da correção de cores se o processo de segmentação utilizasse a mediana ao invés da média aritmética na composição do critério de uniformidade? Este resultado pode ser visto na Figura 4.4. Nele é possível observar uma redução da saturação da cor vermelha.

Os resultados apresentados nas Figuras 4.2, 4.3 e 4.4 permitem uma avaliação qualitativa dos diferentes critérios empregados no MFTA que buscam uma melhoria na qualidade das

Figura 4.4: Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA em que a segmentação utiliza a mediana ao invés da média aritmética na composição do critério de uniformidade.



Fonte: Elaborada pelo autor

cores de referência. No entanto, esta avaliação está sujeita às interferências dos processos psicofísicos presentes na visão humana quando as cores são comparadas visualmente. Para se evitar estes efeitos e alcançar uma maior precisão, uma avaliação quantitativa é apresentada para comparar estes resultados.

Inicialmente, uma área de 5×10 *pixels*, presente na primeira imagem da sequência, é separada (Figura 4.5) e o valor da cor média é calculado $R_m = 123$, $G_m = 59$ e $B_m = 51$. Convertendo estes valores para o espaço HSV obtêm-se uma saturação $S_m = 0,5854$. Este valor de saturação é uma referência para avaliação das imagens corrigidas.

Figura 4.5: Região de 5×10 *pixels* separada do primeiro quadro.

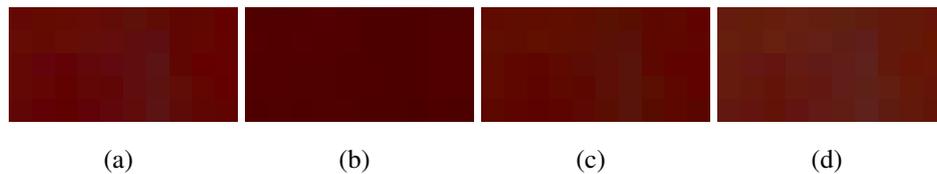


Fonte: Elaborada pelo autor

A Figura 4.6 mostra, lado a lado, a mesma área localizada nas imagens do quadro 700 corrigidas pelos quatro critérios vistos até então: (a) a segmentação e a definição do valor da amostra são feitos pela média aritmética (Seção 3.4); (b) segmentação pela média e defi-

nição do valor pela mediana do matiz; (c) segmentação pela média e definição do valor pela mediana dos canais R, G e B e (d) segmentação e definição do valor pela mediana dos canais R, G e B.

Figura 4.6: Região de 5x10 *pixels* para comparação dos critérios: (a) segmentação e valor da amostra pela média aritmética; (b) segmentação pela média e valor pela mediana do matiz; (c) segmentação pela média e valor pela mediana dos canais R, G e B; (d) segmentação e valor pela mediana dos canais R, G e B.



Fonte: Elaborada pelo autor

A Tabela 4.2 apresenta os respectivos valores RGB e HSV da cor média da região 5x10 *pixels* do quadro 700 original e aqueles corrigidos por cada um dos critérios definidos. Estes valores permitem verificar que o resultado da correção, (b), utilizando o critério da mediana ordenada pelo matiz foi mais saturado que o critério da Seção 3.4, (a). O critério da mediana ordenada pelos canais R, G e B produziu um resultado, (c), próximo do resultado (a). Por fim, o critério, (d), da mediana ordenada pelos canais R, G e B quando utilizado na segmentação das regiões coloridas e na caracterização do valor das amostras apresenta a menor saturação da cor vermelha. Isto novamente reforça a hipótese da sensibilidade do método aos ruídos e que a mediana produz resultados mais estáveis que a média aritmética.

Tabela 4.2: Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 700 para os critérios (a), (b), (c) e (d)

Critério	R_m	G_m	B_m	H_m	S_m	V_m
original	123	59	51	0,0185	0,5854	123,00
(a)	97	8	6	0,0037	0,9381	97,00
(b)	79	0	0	0,0000	1,0000	79,00
(c)	93	10	1	0,0163	0,9892	93,00
(d)	99	26	15	0,0218	0,8485	99,00

Apesar do uso da mediana no processo de segmentação contribuir para uma diminuição da saturação, numericamente perceptível, ela ainda não é capaz de anular totalmente a distorção provocada pelos ruídos. Isto pode ser verificado pelo cálculo do módulo do vetor que

é a diferença entre a cor média de cada região e a cor média da região do primeiro quadro. Estes valores foram calculados no espaço de cor RGB e são apresentados na Tabela 4.3.

Tabela 4.3: Valores do módulo do erro entre a cor média de cada região do quadro 700 e a cor média do primeiro quadro

Critério	Erro
(a)	72,8148
(b)	89,5433
(c)	76,1643
(d)	54,4151

Por outro lado, a busca pela melhoria da qualidade das cores se mostrou como um caminho viável e apto a produzir contribuições para a melhoria das correções.

Os resultados apresentam uma independência do quadro utilizado na avaliação. As Tabelas 4.4 e 4.5 mostram os resultados para o quadro 800 e as Tabela 4.6 e 4.7 para o quadro 900.

Tabela 4.4: Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 800 para os critérios (a), (b), (c) e (d)

Critério	R_m	G_m	B_m	H_m	S_m	V_m
(a)	98	6	6	0,0000	0,9388	98,00
(b)	79	0	0	0,0000	1,0000	79,00
(c)	94	10	2	0,0145	0,9787	94,00
(d)	100	26	15	0,0216	0,8500	100,00

Tabela 4.5: Valores do módulo do erro entre a cor média de cada região do quadro 800 e a cor média do primeiro quadro

Critério	Erro
(a)	73,8850
(b)	89,5433
(c)	75,1199
(d)	53,9815

Tabela 4.6: Valores RGB e HSV médios da mesma região das imagens corrigidas do quadro 900 para os critérios (a), (b), (c) e (d)

Critério	R_m	G_m	B_m	H_m	S_m	V_m
(a)	97	5	5	0,0000	0,9485	97,00
(b)	74	0	0	0,0000	1,0000	74,00
(c)	92	4	1	0,0055	0,9891	92,00
(d)	93	13	5	0,0147	0,9479	96,00

Tabela 4.7: Valores do módulo do erro entre a cor média de cada região do quadro 900 e a cor média do primeiro quadro

Critério	Erro
(a)	75,5513
(b)	92,1032
(c)	80,5357
(d)	71,6380

É possível observar um aumento no erro entre a cor média do primeiro quadro e a cor média dos quadros 800 e 900. Isso é esperado uma vez que os efeitos dos ruídos continuam agindo durante toda a sequência. Entretanto, os resultados relativos entre os critérios testados se mantêm os mesmos. O critério (d) para o quadro 900 apresentou um aumento no erro mais significativo. Isso corresponde a um aumento na saturação da cor vermelha fazendo com que se aproximasse da saturação do critério (a).

4.2.2 Avaliação da máscara binária

A segunda alternativa para melhoria da qualidade das cores também busca uma forma de diminuir a influência dos pontos espúrios na caracterização dos valores das amostras de cor das regiões coloridas. Para isso, uma máscara binária é definida para cada região colorida determinando quais *pixels* serão utilizados. Estas máscaras possuem o mesmo tamanho que as regiões segmentadas e são atualizadas a cada iteração.

Dois critérios para definição da máscara binária são avaliados: os valores de brilho dos *pixels* e os valores absolutos dos canais R, G e B. O primeiro critério gera uma única máscara para uso imediato e o segundo gera, inicialmente, três máscaras separadas, uma para cada canal de cor, e em seguida aplica uma conjunção entre elas, gerando uma única máscara.

Para o primeiro critério, o valor em cada posição da máscara será verdadeiro se a diferença entre o brilho do respectivo *pixel* da região colorida e o brilho médio desta região for menor ou igual à um limiar $k\sigma$, onde, k é um fator de ajuste e σ é o desvio padrão do brilho da referida região.

O procedimento adotado exclui do cálculo do valor das amostra de cor os *pixels* cujo desvio do brilho em relação ao brilho médio seja maior que k vezes o desvio padrão σ . Esta exclusão pode, em alguns casos, gerar uma máscara só de zeros, impedindo o uso desta amostra de cor no cálculo da transformação **T**. Assim, para evitar esta perda de informação, o menor valor de ajuste k que não gera máscaras vazias é escolhido.

O teste da correção de cor utilizando o critério de brilho para geração das máscaras binárias é realizado com $k = 0.98$. O resultado da correção para o quadro 700 pode ser visto na Figura 4.7. A cor vermelha ainda se apresenta ligeiramente menos saturada quando comparada com o resultado da Seção 3.4. Isto pode ser verificado numericamente observando a região 5×10 *pixels* do quadro 700 corrigido que possui uma cor média com valores $R_m = 100$, $G_m = 18$ e $B_m = 17$. Isto equivale a uma saturação $S_m = 0,8300$.

Figura 4.7: Quadros 1 e 700 da sequência com movimento relativo corrigidos pelo MFTA utilizando uma máscara binária definida pelo brilho dos *pixels* de cada região colorida.



Fonte: Elaborada pelo autor

O segundo critério para geração das máscaras binárias aplica o mesmo procedimento de comparação visto no primeiro critério. A diferença em relação ao primeiro critério está apenas nos valores utilizados que, neste caso, são os valores absolutos dos canais R, G e B de cor enquanto, no primeiro, são os valores de brilho. A máscara efetivamente utilizada foi computada por uma conjunção entre as três máscaras obtidas para cada canal. Por causa dessa operação lógica, a chance de ocorrer uma máscara vazia aumenta. Isto pôde ser verificado pelo valor de ajuste k que precisou ser maior para evitar esta possibilidade.

A Figura 4.8 mostra o resultado da correção do quadro 700 utilizando as máscaras geradas pelo segundo critério para um valor de ajuste $k = 1.35$. Neste caso, os valores RGB da cor média da região 5×10 *pixels* são $R_m = 99$, $G_m = 16$ e $B_m = 14$ e a saturação $S_m = 0,8586$. A

distorção da cor vermelha está muito próxima daquela observada quando o primeiro critério é utilizado.

Figura 4.8: Quadros 1 e 700 da segunda sequência corrigidos pelo MFTA utilizando uma máscara binária definida pelos três canais R, G e B.



Fonte: Elaborada pelo autor

Os dois critérios para geração de máscaras binárias apresentam resultados muito semelhantes àquele do critério (d) da avaliação da mediana presente na Tabela 4.2. Isto significa que o uso da máscara binária, semelhantemente à segmentação pela mediana, critério (d), apresentou melhores resultados que os critérios (b) e (c) da Seção 4.2.1. Novamente fica claro o benefício da melhoria da qualidade das cores de referência, pois houve uma redução na saturação. No entanto, esta redução não é suficiente para anular completamente este efeito de distorção de cores.

Os resultados para os quadros 800 e 900 estão apresentados na Tabelas 4.8 e 4.9. Vale lembrar que o matiz próximo dos valores zero e um são referentes à mesma região de cores. Isto justifica a diferenças entre os valores do matiz para a Tabela 4.8. Os valores alcançados para os quadros 800 e 900 estão próximos dos resultados obtidos para o quadro 700. No entanto, eles evidenciam duas tendências distintas para os dois critérios avaliados. O critério de brilho para geração da máscara binária apresentou resultados mais estáveis entre os quadros 800 e 900. Por outro lado, o critério que utiliza os canais RGB mostrou uma maior variação na saturação.

Tabela 4.8: Valores RGB e HSV médios da região 5×10 pixels da imagem corrigidas do quadro 800 para os dois critérios utilizados

Critério	R_m	G_m	B_m	H_m	S_m	V_m
Brilho	101	17	18	0,9980	0,8317	101,00
Canais RGB	101	15	14	0,0019	0,8614	101,00

Tabela 4.9: Valores RGB e HSV médios da região 5×10 *pixels* da imagem corrigidas do quadro 900 para os dois critérios utilizados

Critério	R_m	G_m	B_m	H_m	S_m	V_m
Brilho	101	17	18	0,9980	0,8317	101,00
Canais RGB	99	11	12	0,9981	0,8889	99,00

Para completar a análise da melhoria da qualidade das cores, uma última verificação é necessária. Esta verificação diz respeito ao problema de registro (Olsson et al., 2006) que está relacionado com a localização precisa das regiões coloridas. Para isso, uma avaliação do erro de posicionamento causado pelo processo de *tracking* é feita.

O processo de *tracking* localiza as regiões coloridas, em uma nova imagem, através de uma busca em torno da posição em que elas foram localizadas na imagem anterior. A nova localização de cada região colorida é definida como sendo o ponto onde ocorre a maior correlação entre esta região e a região de busca. Após as regiões serem localizadas, as posições são atualizadas para que uma nova busca possa ser feita na imagem seguinte da sequência.

Para verificar a precisão do procedimento de *tracking*, o método detector de pontos de interesse, *Speeded-Up Robust Features-SURF* (Bay et al., 2008) é utilizado para localizar pontos característicos de referência. Estes pontos são comparados com os resultados de localização feitos pelo processo de *tracking*. Inicialmente, os pontos característicos de duas imagens consecutivas são obtidos pelo SURF, os pares de pontos que possuem correspondência são identificados e suas coordenadas são armazenadas.

As coordenadas dos pontos característicos identificados na primeira imagem são utilizadas para definir um conjunto de regiões quadradas centradas neste pontos. Estas regiões são, então, localizadas pelo procedimento de *tracking* na imagem seguinte e suas coordenadas registradas. Estas coordenadas localizadas pelo *tracking* são comparadas com as coordenadas dos pontos característicos identificados pelo SURF na segunda imagem.

A Figura 4.9 mostra os pontos característicos identificados na primeira imagem (esquerda). Estes pontos estão marcados pelo centro de um quadro amarelo que identifica a região utilizada pelo *tracking* e pelo centro de um círculo vermelho que mostra os pontos identificados pelo SURF. Na segunda imagem (direita), a posição onde estes pontos são localizados é identificada por um quadrado amarelo para o *tracking* e por um círculo vermelho para o SURF.

Uma linha vermelha e outra verde são plotadas ligando os pontos da primeira imagem aos

Figura 4.9: Localização de pontos característicos em duas imagens consecutivas da sequência.



Fonte: Elaborada pelo autor

pontos localizados na segunda imagem. A linha vermelha associa os pontos localizados pelo *tracking* e a linha verde os pontos localizados pelo SURF. Para os pontos localizados nas mesmas posições pelos dois procedimentos, a linha vermelha sempre cobre a linha verde. Por outro lado, quando há um desvio na localização, a linha verde se torna aparente. É possível observar na Figura 4.9 que as linhas verdes não são frequentes. Isto indica que as localização pelos dois procedimentos são, em sua maioria, coincidentes.

Um resultado mais preciso é alcançado por uma avaliação quantitativa do desvio de localização. Esta avaliação é realizada utilizando a sequência com movimento relativo entre a cena e a câmera. Para cada imagem da sequência, os desvios entre os pontos localizados pelo *tracking* e os pontos localizados pelo SURF são registrados e acumulados. Este desvios são medidos em valores inteiros de *pixels*. A Tabela 4.10 mostra a frequência dos desvios ocorridos na horizontal e na vertical em valores absolutos e percentual.

Tabela 4.10: Frequência dos desvios entre os pontos localizados pelo *tracking* e os pontos localizados pelo SURF

Desvio (<i>pixels</i>)	Frequência horizontal		Frequência vertical	
0	66.169	70,851%	69502	74,420%
1	25.583	27,393%	23132	24,769%
2	1289	1,380%	719	0,770%
3	176	0,188%	20	0,021%
4	79	0,085%	8	0,009%
5	50	0,054%	5	0,005%
6	17	0,018%	5	0,005%
7	9	0,010%	0	0%
8	11	0,012%	0	0%
9	3	0,003%	0	0%
10	2	0,002%	1	0,001%
11	0	0%	0	0%
12	1	0,001%	0	0%
13	0	0%	0	0%
14	2	0,002%	0	0%
15	1	0,001%	0	0%

É possível observar que o desvio mais frequente é zero e que mais de 98% deles são de até um *pixel* de deslocamento, ou seja, o procedimento de *tracking* localizou com boa precisão a maioria das regiões em relação à localização feita pelo SURF. Isto significa que

o procedimento de *tracking* é relativamente preciso na localização das regiões em torno dos pontos característicos. Devido as características intrínsecas destes pontos, as regiões ao redor deles são normalmente texturizadas facilitando a sua identificação pelo *tracking*. No entanto, isto não se repete, necessariamente, para regiões que apresentem um determinado grau de uniformidade como aquelas utilizadas no processo de correção de cores.

Os resultados analisados até aqui mostram que as interferências de origem espacial, como o desvio de localização e a inserção de pontos espúrio nas regiões coloridas, não são fáceis de serem suprimidas. Podem até serem atenuadas, por exemplo, pela substituição da média pela mediana no processo de segmentação, mas não são completamente eliminadas. Este comportamento é similar ao efeito da mudança no tamanho da área das regiões coloridas feita no MAIA. Naquela situação, o aumento da área reduzia o efeito da convergência indesejável, mas não o anulava por completo.

4.3 Uma abordagem baseada em filtragem temporal

A persistência dos efeitos dos ruídos, mesmo depois da aplicação de métodos para melhoria da qualidade das cores de referência, exige que uma outra abordagem seja investigada para se alcançar a atenuação destes efeitos. Para tanto, deve-se considerar que a caracterização das amostras de cor pela média aritmética de uma região colorida, com tamanhos variados, não foi suficiente para conter a convergência indesejável observada no MAIA. Além disso, a solução para conter este efeito só foi alcançada quando a filtragem saiu do contexto espacial para o contexto temporal presente no MFTA. A partir disso, é possível supor que uma melhoria na filtragem temporal pode produzir resultados mais estáveis.

A filtragem temporal presente no MFTA resultou na eliminação da convergência indesejável observada no MAIA quando referências fixas de cor foram utilizadas. No entanto, quando as referências de cor localizadas nas imagens apresentam um deslocamento em relação à câmera, os resultados não são tão bons. Isto torna evidente que o deslocamento espacial das referências influencia na qualidade das cores e consequentemente na qualidade das transformações.

Este fato sugere que o ruído presente nas cores esteja sendo transferido para as transformações. Assim, em uma situação onde haja pouca ou nenhuma variação da fonte de iluminação deveria ser obtido uma sequência de transformações com poucas diferenças entre si. No entanto, sob a influência destes ruídos, as transformações geradas estariam mapeando espaços de cores espúrios. Esta falta de uniformidade entre transformações consecutivas deve, então, ser atenuada.

As cores de referência, geradas no MAIA a cada correção de cores, também sofriam influência dos ruídos e comportavam-se como referências espúrias sujeitando o MAIA a uma convergência indesejável. No entanto, estas referências tornaram-se mais estáveis quando foram calculadas pela média temporal das referências anteriores. Extrapolando este conceito para as transformações, tem-se que estas sofrem influência dos ruídos e comportam-se como transformações espúrias. Diante disso, é possível que um esquema de filtragem temporal das transformações seja capaz de gerar transformações menos susceptíveis aos ruídos atenuando a distorção das cores.

Este contexto evidencia uma estrutura que envolve três tipos de filtragem que atuam em três domínios distintos:

- A primeira filtragem é espacial e atua no domínio dos *pixels* de uma região.
- A segunda filtragem é temporal aplicada no domínio dos valores de cor de uma mesma região ao longo do tempo.
- A terceira filtragem também é uma filtragem temporal que atua no domínio das transformações.

Cada um destes domínios envolvidos emprega uma quantidade crescente de informação para obtenção das correções de cor. Isto é coerente à medida em que se busca maior precisão nas correções.

4.4 O MFTT

Este método apresenta uma abordagem para implementação de uma filtragem temporal das transformações \mathbf{T} obtidas em cada iteração. Esta filtragem é caracterizada por um filtro passa-baixa que produz uma transformação que é uma média das transformações anteriores. As transformações utilizadas devem ser aquelas obtidas sob variações relativamente suaves da iluminação. Por exemplo, variações obtidas em ambiente externos devido à iluminação solar e aquelas devido às transições entre ambientes adjacentes.

Assim, uma janela de w iterações é definida considerando que houve pouca ou nenhuma variação de iluminação neste espaço de tempo. A filtragem temporal deve ser aplicada em um conjunto de transformações anteriores $[\mathbf{T}_{i-w+1}, \dots, \mathbf{T}_i]$, dentro de uma janela w de quadros, para obter a transformação da iteração atual \mathbf{T}_i .

Considerando um conjunto de w iterações, as cores de referência utilizadas em cada iteração são armazenadas formando um grande conjunto de referência S_i e da mesma forma é feito para os conjuntos de cores capturadas formando um grande conjunto S'_i . A partir desses dois grandes conjuntos a transformação da iteração i é determinada. A cada iteração a janela de largura w se desloca inserindo os novos conjuntos de cores e removendo os mais antigos. Assim, a transformação \mathbf{T}_i é uma transformação média de uma janela de w transformações definida pela Equação 4.1.

$$\mathbf{T}_i = [\mathbf{S}_i][\mathbf{S}'_i]^+ \quad (4.1)$$

Nesta equação, $[\mathbf{S}_i]$ é a matriz com os valores das cores de referência e $[\mathbf{S}'_i]$ é a matriz com os valores das cores capturadas. As duas matrizes são obtidas em w quadros e dadas pelas Equações 4.2 e 4.3, respectivamente. Os valores k_{i-w+1} até k_i são os tamanhos dos conjuntos em cada iteração.

$$\mathbf{S}_i = \{[\mathbf{c}_1, \dots, \mathbf{c}_{k_{i-w+1}}]_{i-w+1}, \dots, [\mathbf{c}_1, \dots, \mathbf{c}_{k_i}]_i\} \quad (4.2)$$

$$\mathbf{S}'_i = \{[\mathbf{c}'_1, \dots, \mathbf{c}'_{k_{i-w+1}}]_{i-w+1}, \dots, [\mathbf{c}'_1, \dots, \mathbf{c}'_{k_i}]_i\}. \quad (4.3)$$

O algoritmo do MFTT apresenta pequenas alterações em relação ao algoritmo do MFTA como pode ser visto na descrição abaixo. Nesta descrição é possível observar que os conjuntos de cores C e C' são inseridos na primeira posição dos conjuntos S e S' , respectivamente. As demais posições destes conjuntos são deslocadas de uma posição e os grupos de cores da última posição são descartados.

Algoritmo *Método do Filtro Temporal para Transformações (MFTT)*

Entrada: Dimensão da janela de iterações w

Entrada: Dimensão dos quadrados utilizados para segmentar a imagem $l = 5$

Entrada: Dimensão dos blobs utilizados para buscas nas imagens $m = 25$

Entrada: Uma sequência de imagens de entrada V_e

Saída: Uma sequência de imagens de saída V_s

(* versão com *tracking* de regiões coloridas *)

1. $I' \leftarrow$ primeira imagem da sequência V_e
2. segmenta a imagem I' em quadrados de $l \times l$ pixels gerando uma lista de quadrados P contendo coordenadas centrais, cor média e um blob para cada quadrado
3. C recebe a lista de cores médias dos quadrados da lista P

4. **para** $j \leftarrow 1$ **até** w
5. **faça** $S_j \leftarrow C$;
6. $S'_j \leftarrow C$
7. $i \leftarrow 0$
8. insere I' na sequência V_s
9. **para outros** $I' \in V_e$
10. **faça** localize os blobs da list P na imagem I' gerando uma lista Q contendo coordenadas centrais e cor média dos quadrados localizados
11. C' recebe a lista de cores médias dos quadrados da lista Q
12. exclua das listas P e Q os quadrados sem correspondentes, os que estão fora da região de interesse, os quadrados com sobreposição e aqueles com desvio de posição e as respectivas cores médias em C e C'
13. $S_{2:w} \leftarrow S_{1:w-1}$
14. $S_1 \leftarrow C$
15. $S'_{2:w} \leftarrow S'_{1:w-1}$
16. $S'_1 \leftarrow C'$
17. $T \leftarrow S * S'^+$; calcula a matriz de transformação usando pseudo-inversa
18. $C \leftarrow ((C \times i) + T * C') / (i + 1)$; atualiza as cores de referência utilizando média ponderada
19. atualiza as coordenadas centrais da lista P com as correspondentes da lista Q
20. segmenta a imagem I' em quadrados de $l \times l$ pixels gerando uma lista de quadrados R
21. compara os quadrados da lista P com os quadrados da lista R , exclua os quadrados da lista P sem correspondentes na lista R e inclua quadrados novos da lista R na lista P
22. inclua em C as cores médias dos novos quadrados da lista P
23. $i \leftarrow i + 1$
24. $I \leftarrow T * I'$; aplica a matriz de transformação a toda a imagem
25. insere o quadro corrigido I na sequência V_s
26. avança para a próximo quadro da sequência V_e
27. **retorna** V_s

O MFTT apresenta um resultado melhor que o MFTA utilizando a mediana no processo de segmentação. Na Figura 4.10 pode ser observado que a cor vermelha saturada, presente quando o MFTA foi utilizado, não é mais percebida com o MFTT.

A Tabela 4.11 mostra os valores da cor vermelha da mesma região 5×10 pixels definida na Figura 4.5. Estes valores são relativos ao quadro um e ao quadro 700. É possível observar que a diferença máxima dos canais RGB ocorre para o canal B e vale 13,73%. A saturação

Figura 4.10: Primeiro e último quadros da sequência corrigidos pelo MFTT para $w=30$.

Fonte: Elaborada pelo autor

da cor vermelha no quadro 700 está muito próxima da saturação original presente no quadro um. A diferença entre estes valores de saturação é de 7,12%. A diferença entre os dois vetores de cor no espaço de cor RGB é de 11,7473.

Tabela 4.11: Valores RGB e HSV médios da região 5×10 pixels do quadro um capturado e do quadro 700 corrigido

Quadro	R_m	G_m	B_m	H_m	S_m	V_m
1	123	59	51	0,0185	0,5854	123,00
700	118	51	44	0,0158	0,6271	118,00

Os resultados dos quadros 800 e 900, Tabela 4.12, são muito próximos evidenciando uma maior estabilidade na correção das cores. a variação da saturação entre os quadros 700 e 900 corrigidos foi de apenas 0,86%.

Tabela 4.12: Valores RGB e HSV médios da região 5×10 pixels do quadro 800 e 900 corrigidos

Quadro	R_m	G_m	B_m	H_m	S_m	V_m
800	118	51	44	0,0158	0,6271	118,00
900	117	51	43	0,0180	0,6325	117,00

4.5 Conclusão

O MFTT transfere o conceito da filtragem temporal, aplicada nos valores de referência, para os valores das transformações obtidas em cada iteração. Isto permite a utilização de

uma número maior de informações atenuando o efeito dos pontos espúrios introduzidos pelo ruído. Este comportamento pode ser observado pelo resultado do teste feito com a sequência com movimento relativo. Este resultado foi significativamente melhor que àqueles obtidos pela melhoria da qualidade das cores de referência.

Para uma sequência de imagens contendo uma grande variedade de efeitos presentes em cenas comuns, a filtragem temporal das transformações, presente no MFTT, mostrou-se capaz de atenuar a distorção das cores causada pelo processo iterativo de correção. Uma consequência desta filtragem é uma insensibilidade às variações abruptas da iluminação. No entanto, as variações de iluminação em ambiente externos devido à iluminação solar e aquelas devido às transições entre ambientes adjacentes são relativamente suaves, não impactando significativamente nos resultados.

Capítulo 5

Experimentos Realizados

5.1 Ferramentas de programação

O uso da linguagem do MATLAB foi adotado nos *scripts* de programação. O MATLAB possui um ambiente de fácil uso e que integra a análise numérica, o cálculo com matrizes, o processamento de sinais e a construção de gráficos. Possui também um vasto conjunto de funções nativas e outras provenientes de suas extensões (*toolboxes*). Além disso, os problemas e soluções podem ser expressos similar à forma como são descritos matematicamente.

O *Computer Vision System Toolbox* fornece algoritmos e ferramentas para projeto e simulação de visão computacional e sistemas de processamento de imagens. Este *toolbox* inclui algoritmos de extração de características, detecção de movimento, detecção de objetos, *tracking* de objetos, visão estéreo, processamento e análise de vídeos. Ele também inclui ferramentas para entrada e saída de arquivos de vídeo, exibição de vídeos, desenhos gráficos e composições.

As principais funções utilizadas foram aquelas relativas à extração de quadros a partir de sequências de imagens, à correção gama e à exibição de marcas como quadrados, círculos e cruzes nas imagens para identificação de pontos e áreas de interesse. Outra característica do MATLAB utilizada foi a capacidade de execução de funções escritas em C++. Esta interface está baseada nas *mex-functions* que são acessadas diretamente pela linha de comando ou pelos *scripts* do MATLAB. A vantagem da utilização de *mex-functions* está no aumento do desempenho computacional.

Duas funções foram implementadas em C++ utilizando o interfaceamento através das *mex-functions*. A primeira é a função que implementa a segmentação das imagens em pe-

quenos quadrados de tamanho fixo. Esta segmentação está descrita na Seção 3.4. A segunda função faz o *tracking* das regiões coloridas como descrito na mesma seção.

5.2 Experimentos de segmentação da imagem

A segmentação de imagens desempenha um importante papel na área da visão computacional (Deshmukh, 2010). O objetivo da segmentação é particionar uma imagem em seus objetos ou regiões constituintes. A segmentação de imagens não triviais é uma das tarefas mais difíceis em processamento de imagem e sua precisão determina o eventual sucesso ou falha dos processos automáticos (Gonzalez e Woods, 2008).

No processo de correção de cores, apresentado nesta tese, a segmentação das imagens em regiões coloridas representa uma importante etapa. Esta segmentação é responsável pela definição dos valores que serão utilizados para obtenção da matriz de mapeamento de cores. Inicialmente, dois métodos de segmentação foram avaliados. O primeiro método é a segmentação por crescimento de região (Gonzalez e Woods, 2008) e o segundo método é chamado de *graph-based image segmentation algorithm* (Felzenszwalb e Huttenlocher, 2004). No entanto, os resultados das avaliações destes métodos apresentaram uma significativa instabilidade na quantidade e nos formatos das regiões segmentadas entre quadros consecutivos. Assim, um terceiro método de segmentação por quadrados de tamanho fixo foi escolhido para utilização no processo de correção de cores. Este método escolhido segmenta as imagens em regiões quadradas de tamanho fixo como visto na Seção 3.4. Os métodos são avaliados nas seções seguintes.

5.2.1 Segmentação utilizando crescimento de região

Este método de segmentação utiliza informações provenientes do domínio espacial da imagem como, por exemplo, a intensidade ou a cor dos *pixels*. O princípio fundamental do crescimento de região é a agregação de *pixels* que atendam critérios de conectividade e similaridade. Estes critérios são relativos a um *pixel* tomado como "semente", em torno do qual ocorre o processo de crescimento de região. Além desses dois critérios, outros três são igualmente importantes: (i) critério de escolha das sementes; (ii) critério de desambiguação de múltiplas sementes válidas para um mesmo *pixel*; e (iii) critério de parada.

O critério de escolha das sementes iniciais tem grande importância no resultado final da segmentação (Huang et al., 2010). Assim, diversas abordagens para a escolha automática das sementes têm sido propostas (Ikonomakis et al., 2000; Singh e Singh, 2013). Durante a

agregação dos *pixels*, aqueles que satisfizerem o critério de conectividade poderão, eventualmente, atender ao critério de similaridade para mais de uma semente. Então, é necessário definir um critério para resolver esta ambiguidade. Por fim, um critério de parada deve ser definido quando não houverem mais *pixels* a serem agregados (Gonzalez e Woods, 2008).

O experimento realizado com o método de segmentação por crescimento de região utiliza a implementação em C++ chamada *Fast 3D/2D Region Growing (MEX)*¹. Este código recebe como entrada uma imagem cinza e as coordenadas de um *pixels* que servirá como semente. A saída deste procedimento é uma imagem binária contendo a região crescida. Para gerar a lista de sementes, o conceito básico de um *grid*, empregado por Ikonomakis et al. (2000), foi utilizado. Este *grid*, espaçado por uma distância L , define os pontos que são utilizados como sementes.

O processo de segmentação das regiões segue a ordem com que as sementes foram geradas. Assim, as sementes são utilizadas, uma a uma, para gerar as regiões e sempre que uma região gerada sobrepuser as coordenadas das sementes que ainda não tenham sido utilizadas estas são, então, retiradas da lista. As regiões geradas recebem um rótulo, que é registrado em uma matriz de rótulos. O processo termina quando não há mais sementes a serem utilizadas.

Como dito anteriormente, o código utilizado recebe uma imagem monocromática como entrada. No entanto, as imagens a serem segmentadas são coloridas. Assim, a etapa de geração da região é executada para cada um dos canais de cor R, G e B separadamente. Em seguida, algumas operações morfológicas de dilatação seguidas de erosão, caracterizando uma operação de fechamento, são realizadas para remover os buracos. Finalmente uma operação lógica E é realizada entre as três regiões geradas para cada canal de cor. A região resultante é tomada como a região segmentada.

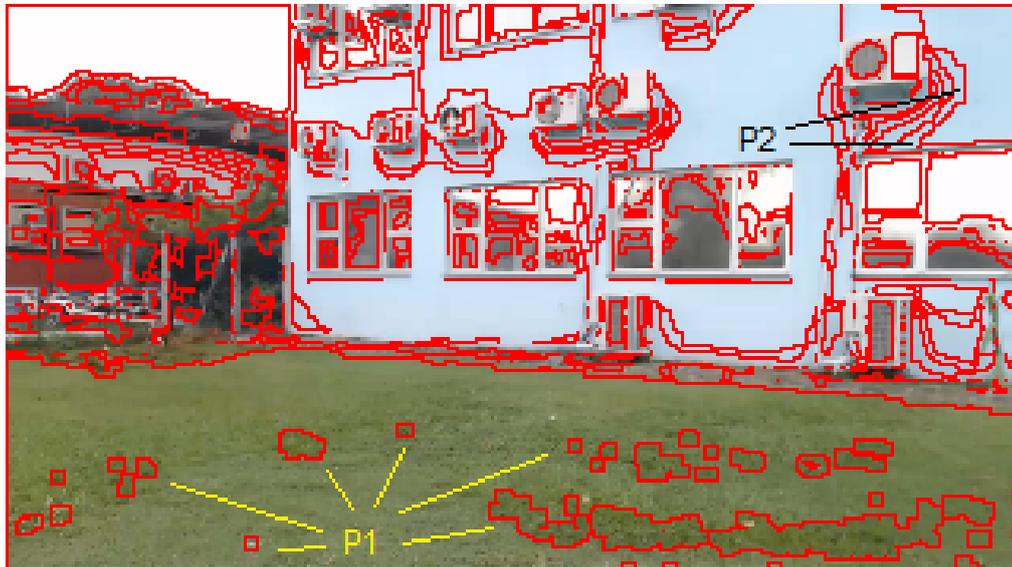
As Figuras 5.1, 5.2 e 5.3 mostram, respectivamente, os resultados da segmentação por crescimento de região do primeiro, segundo e terceiro quadros da sequência com movimento relativo entre câmera e cena.

A segmentação por crescimento de regiões produz grandes áreas uniformes como pode ser observado na Figura 5.1. Esta é uma característica muito importante para redução do ruído espacial que contamina a cor média da área em questão. No entanto, quando os três quadros consecutivos são comparados (Figuras 5.1, 5.2 e 5.3) é possível perceber uma instabilidade tanto na quantidade como na forma das regiões geradas.

Por exemplo, na posição $P1$ da Figura 5.1 o algoritmo de segmentação identificou regiões

¹O código pode ser obtido no endereço eletrônico <http://www.mathworks.com/matlabcentral/fileexchange/41666-fast-3d-2d-region-growing--mex->

Figura 5.1: Resultado da segmentação do primeiro quadro por crescimento de região (utiliza a implementação em C++ chamada *Fast 3D/2D Region Growing (MEX)*).



Fonte: Elaborada pelo autor

Figura 5.2: Resultado da segmentação do segundo quadro por crescimento de região (utiliza a implementação em C++ chamada *Fast 3D/2D Region Growing (MEX)*).



Fonte: Elaborada pelo autor

que desapareceram no quadro seguinte (Figura 5.2). Também é possível observar na posição *P2* que duas regiões distintas se fundiram em uma única região.

Figura 5.3: Resultado da segmentação do terceiro quadro por crescimento de região (utiliza a implementação em C++ chamada *Fast 3D/2D Region Growing (MEX)*).



Fonte: Elaborada pelo autor

A região segmentada na posição *P3* da Figura 5.2 se dividiu em duas regiões no quadro seguinte (Figura 5.3). O surgimento de novas regiões, que antes não existiam, pode ser observado na posição *P4* da Figura 5.3.

Esta instabilidade na forma e na quantidade das regiões geradas impõem a necessidade de um esquema mais sofisticado e complexo de controle e identificação das regiões segmentadas. Regiões que antes não existiam podem surgir por serem novas regiões ou devido à divisão de uma região existente. De forma similar, regiões podem desaparecer por se fundirem a outras regiões ou simplesmente por não terem sido identificadas durante a segmentação.

Considerando que as diferenças entre as imagens consecutivas de uma sequência sejam mínimas, a instabilidade na geração das regiões segmentadas não deve estar relacionada com mudanças de formas causadas pelas mudanças de ângulos durante o movimento. Assim, as causas mais prováveis desta instabilidade recaem sobre o ruído presente na imagem e também sobre a lógica *E* empregada no método de segmentação envolvendo os três canais de cor *R*, *G* e *B*. Para investigar esta hipótese um outro método de segmentação para imagens coloridas é avaliado na próxima seção.

5.2.2 Segmentação utilizando *graph-based image segmentation algorithm*

Um dos objetivos deste método de segmentação é identificar agrupamentos ou regiões perceptualmente importantes refletindo aspectos globais da imagens (Felzenszwalb e Huttenlocher, 2004). Estes aspectos se diferem dos aspectos locais como, por exemplo, intensidade e cor dos *pixels*. Para isso, ele utiliza uma representação baseada em grafos para as estruturas de *clusters* identificadas nas imagens (Zahn, 1971; Urquhart, 1982). Para obter esta representação, os princípios de agrupamento que integram o agrupamento perceptivo da visão humana são utilizados na caracterização dos algoritmos para geração dos *clusters*.

O agrupamento perceptivo (*Perceptual Grouping*) está relacionado com a percepção visual humana da forma dos objetos. Ele está apoiado nos princípios de agrupamento de proximidade (*proximity*), de semelhança (*similarity*), de destino comum (*common fate*), de simetria (*symmetry*), de paralelismo (*parallelism*), de continuidade (*continuity*) e de fechamento (*closure*) que fundamentam a chamada psicologia de Gestalt (Wagemans et al., 2012). Outra característica deste método é que sua implementação² é computacionalmente eficiente possuindo uma complexidade computacional $O(n \log n)$ para n *pixels* da imagem.

As Figuras 5.4, 5.5 e 5.6 mostram, respectivamente, os resultados da segmentação pelo método *graph-based image segmentation algorithm* para o primeiro, segundo e terceiro quadros da sequência com movimento relativo entre câmera e cena. É possível observar nas posições *P5* na Figura 5.4 duas regiões geradas durante a segmentação do primeiro quadro. Estas duas regiões aparecem divididas em outras regiões menores durante a segmentação do segundo quadro, como pode ser visto na Figura 5.5. O resultado da segmentação do terceiro quadro (Figura 5.6) revela que estas mesmas regiões voltaram a se unir em regiões maiores.

Estes resultados evidenciam que os mesmos problemas identificados na segmentação por crescimento de região, visto na seção anterior, estão também presentes na segmentação utilizando o método *graph-based image segmentation algorithm*. Estes problemas estão relacionados às regiões segmentadas que ora se unem em regiões maiores e ora se dividem em regiões menores.

Este comportamento é consequência do ruído presente nas imagens fazendo com que as fronteiras entre as regiões segmentadas se alterem constantemente. Este efeito no processo de segmentação dificulta o controle da lista de regiões exigindo uma estrutura mais complexa para identificação e associação dessas regiões nos quadros consecutivos.

Os experimentos realizados evidenciam a necessidade de um método que seja mais robusto quanto à variação dos limites e tamanhos das áreas segmentadas. Esta estabilidade

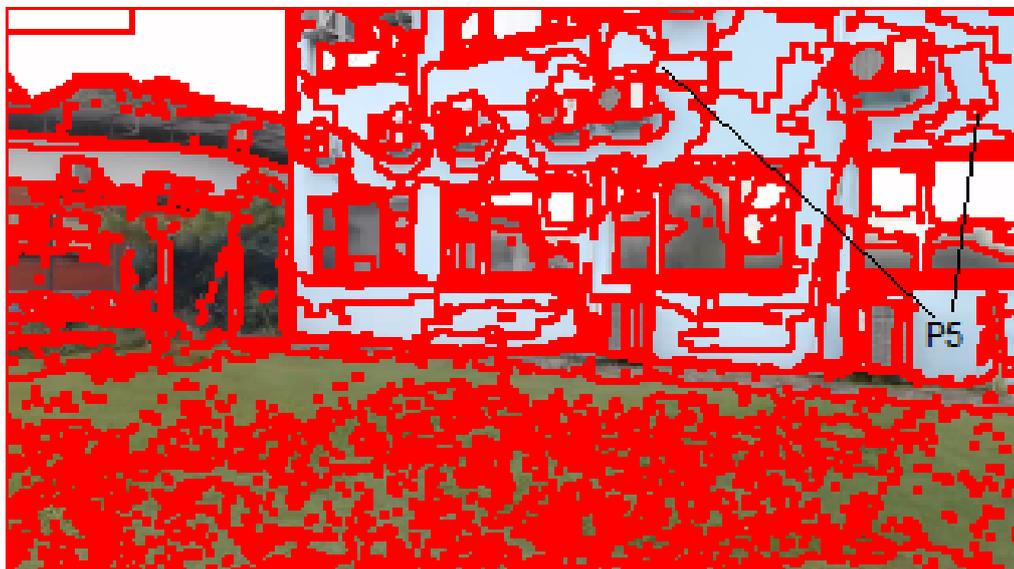
²O código pode ser obtido no endereço eletrônico <http://cs.brown.edu/~pff/segment/>

Figura 5.4: Resultado da segmentação por *graph-based* do primeiro quadro.



Fonte: Elaborada pelo autor

Figura 5.5: Resultado da segmentação por *graph-based* do segundo quadro.



Fonte: Elaborada pelo autor

desejada foi alcançada pela utilização do método de segmentação por quadrados de tamanho fixo. Este método garante uma maior estabilidade na identificação e associação das regiões segmentadas nos quadros consecutivos da sequência de imagens.

Figura 5.6: Resultado da segmentação por *graph-based* do terceiro quadro.

Fonte: Elaborada pelo autor

5.2.3 Segmentação utilizando quadrados de tamanho fixo

Os dois métodos de segmentação vistos anteriormente apresentam uma instabilidade na geração das regiões segmentadas que demanda um esquema mais sofisticado para controle da identificação, da localização, dos limites, dos tamanhos e das quantidades dessas regiões. Para viabilizar o uso de um controle mais simples, o método de segmentação por quadrados de tamanho fixo, citado na Seção 3.4, foi utilizado. Este método de segmentação não atende o caso mais geral de segmentação de imagens coloridas para utilização em problemas de visão computacional, mas permite uma simplificação no controle das regiões segmentadas e no cálculo da cor característica dessas regiões.

Na segmentação por quadrados de tamanho fixo, uma rotina de busca analisa os *pixels* diferentes de zero procurando por aqueles que possuem uma região quadrada em seu entorno que satisfaça o critério de homogeneidade. Quando uma região é encontrada a sua posição é registrada e a busca continua a partir do próximo *pixel* que é capaz de produzir uma nova região sem sobreposições com as regiões já encontradas.

A segmentação da imagem por regiões quadradas possibilita várias facilidades. A primeira delas é relativa à posição e localização. Para registrar a posição da região encontrada pode ser utilizado um critério simples como, por exemplo, o ponto central da região. Por outro lado, a definição da posição de regiões com formas variadas já não é tão simples. Se o centro de massa for utilizado como critério para esta definição ele estará sujeito às variações

de forma da região segmentada e conseqüentemente ficará variando a posição.

Outra diferença da segmentação por quadrados de tamanho fixo em relação aos dois outros métodos citados está associada à quantidade e tamanho das regiões segmentadas que eles podem produzir. O primeiro método é capaz de produzir uma grande quantidade de áreas pequenas com tamanho fixo. Por outro lado, os dois outros métodos são capazes de produzirem grandes áreas, mas tipicamente em pequenas quantidades. Por fim, a forma quadrada das regiões segmentadas facilita o cálculo das cores representativas dessas regiões enquanto as formas irregulares das regiões obtidas pelos outros dois métodos impactam no desempenho deste cálculo.

5.3 Experimentos de comparação do MFTA

Na Seção 3.3, a análise dos resultados alcançados pelo MFTA concentrava-se na verificação da estabilidade das correções de cor quando havia somente variação da iluminação sobre as regiões coloridas. Para que somente a influência da iluminação pudesse ser avaliada, o efeito do movimento foi evitado pela utilização de uma seqüência de imagens sem movimento relativo entre a câmera e a cena.

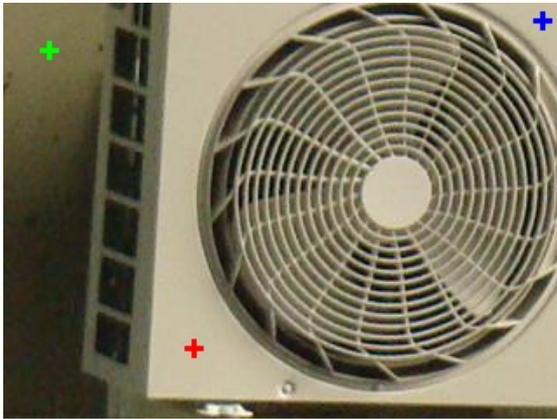
Os resultados do MFTA foram comparados com aqueles obtidos pelo Método Simples (*ground truth*), Seção 3.1.2, e pelo MAIA, Seção 3.2. A comparação entre estes métodos permitiu verificar que o MFTA alcança, para a seqüência de imagens utilizada, uma estabilidade na correção das cores similar àquela alcançada pelo Método Simples sem a necessidade de utilização das mesmas referências de cor, mas utilizando referências atualizadas a cada quadro.

Nesta seção, a comparação do MFTA com outros algoritmos de constância de cor comumente encontrados na literatura permite avaliar a efetividade do método apresentado em relação a estes algoritmos. Além disso, esta comparação permite avaliar o desempenho da transformação linear com nove parâmetros em relação ao simples ajuste de ganho nos três canais R, G e B aplicados nas correções de cor. Para execução dos experimentos são utilizados seis seqüências de imagens sem movimento relativo entre câmera e cena. O primeiro quadro de cada uma das seqüências pode ser visto na Figura 5.7.

A primeira seqüência foi obtida por uma câmera SONY DSC-H9 e a segunda por uma *webcam* PHILIPS SPC530NC. As outras quatro seqüências foram obtidas da internet. Todas elas³ foram capturadas utilizando a técnica *time lapse* sob variação da iluminação do pôr do

³Os arquivos contendo as seqüências de imagens podem ser obtidos no endereço eletrônico <https://nuvem.ufes.br/index.php/s/wocdqSL8Uhak7g0>

Figura 5.7: Seis sequências (1-6) para análise. As marcas em cruz vermelha, verde e azul representam os pontos fixos escolhidos para análise do matiz.



(a) Sequência 1



(b) Sequência 2



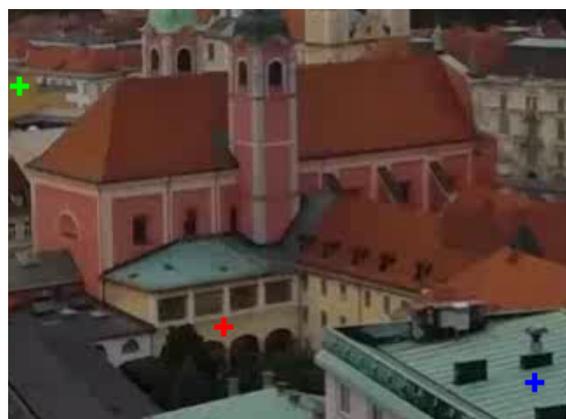
(c) Sequência 3



(d) Sequência 4



(e) Sequência 5



(f) Sequência 6

Fonte: Elaborada pelo autor

sol. Este é um clássico exemplo de variação de iluminação encontrada na natureza. Em cada sequência, uma área de 240×320 pixels das imagens é separada para a análise. A segmentação do primeiro quadro é feita para gerar regiões de 3×3 pixels. A comparação utiliza os algoritmos *Gray-World* (Buchsbaum, 1980), *Max-RGB* (Land, 1977) e *Gray-Edge* (van de Weijer et al., 2007a) e está dividida em duas etapas sendo uma qualitativa e outra quantitativa. A quantidade de quadros das sequências de 1 a 6 é 201, 550, 190, 170, 350 e 525, respectivamente.

5.3.1 Comparação qualitativa

A comparação qualitativa é feita graficamente pelas curvas do matiz de três pontos coloridos presentes nas imagens de cada sequência. Estes pontos permanecem fixos durante toda a sequência e os respectivos valores do matiz são registrados para cada quadro corrigido das sequências. Uma linha base também é plotada utilizando o valor do matiz do primeiro quadro como referência. Os pontos escolhidos estão marcados com cruces coloridas na Figura 5.7. A análise gráfica permite verificar visualmente como os algoritmos ajustam a curva de variação do matiz trazendo-as para próximo das linhas base.

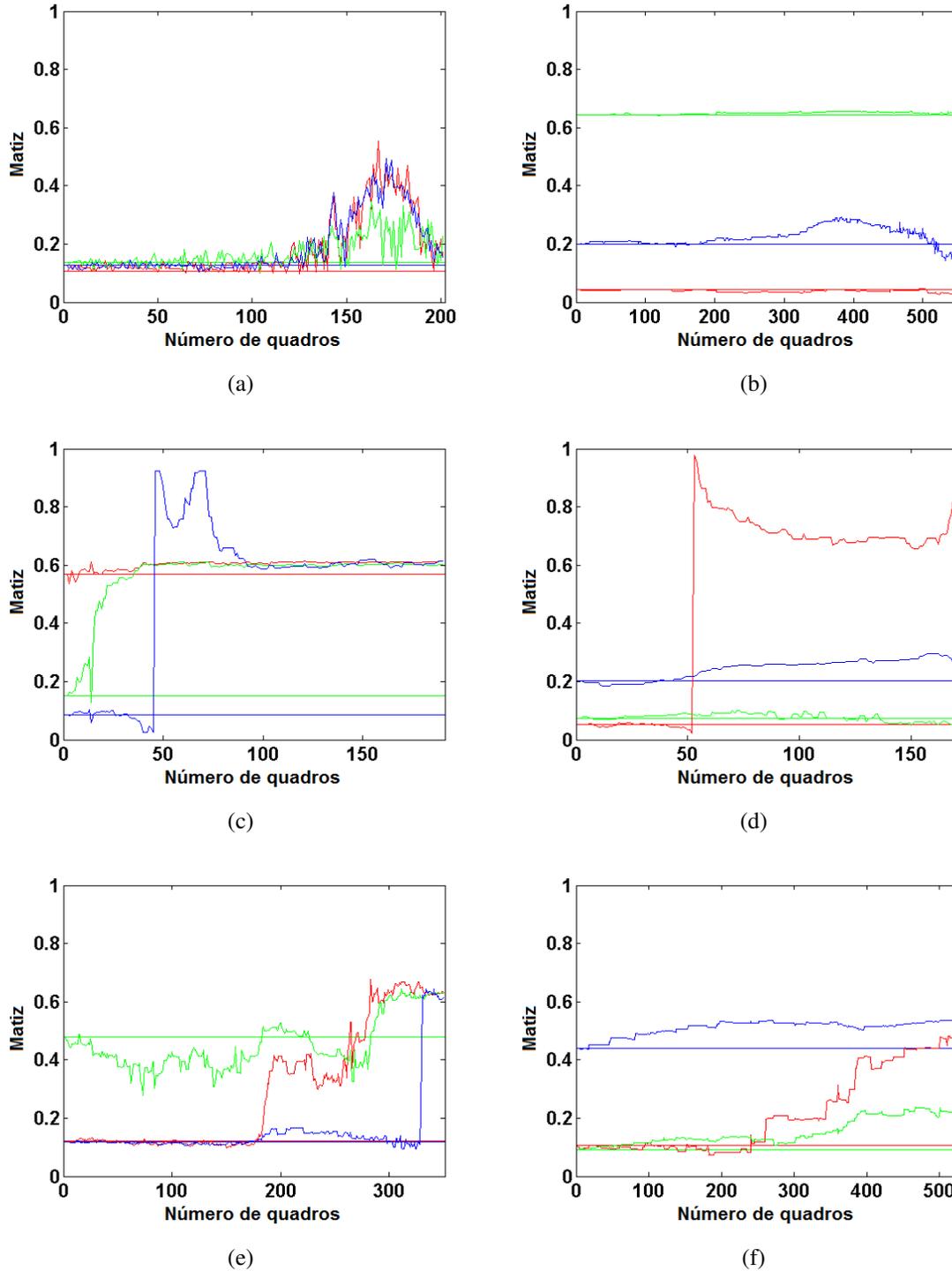
A Figura 5.8 mostra a variação do matiz proveniente das imagens originais das sequências e suas linhas base. A primeira sequência é a mais ruidosa e os valores do matiz dos pontos observados são similares. Por outro lado, a segunda e a terceira sequências possuem valores de matiz mais separados. Na segunda sequência, estes valores permanecem separados durante todos os quadros enquanto na terceira, eles convergem para um mesmo valor de matiz a partir do centésimo quadro. Finalmente, as três últimas sequências possuem dois dos três valores de matiz que iniciam muito semelhantes e em algum ponto eles se separam em valores diferentes. Este conjunto de pontos forneceu uma boa variação de valores do matiz para avaliar os diferentes métodos.

A Figura 5.9 mostra o resultado da correção de cores da primeira sequência alcançado pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*, respectivamente. Para a primeira sequência, o resultado do algoritmo *Max-RGB* é visualmente o mais ruidoso enquanto o MFTA é o mais estável. Os algoritmos *Gray-World* e *Gray-Edge* também apresentaram resultados visualmente ruidosos.

Para a segunda sequência, o MFTA apresenta um resultado semelhante ao resultados dos algoritmos *Gray-World* e *Gray-Edge* (Figura 5.10). Nesta sequência, o algoritmo *Max-RGB* também apresentou o pior resultado.

A terceira sequência apresenta um resultado interessante como pode ser visto na Figura 5.11. Os algoritmos *Gray-World* e *Gray-Edge* apresentam uma capacidade para separar os

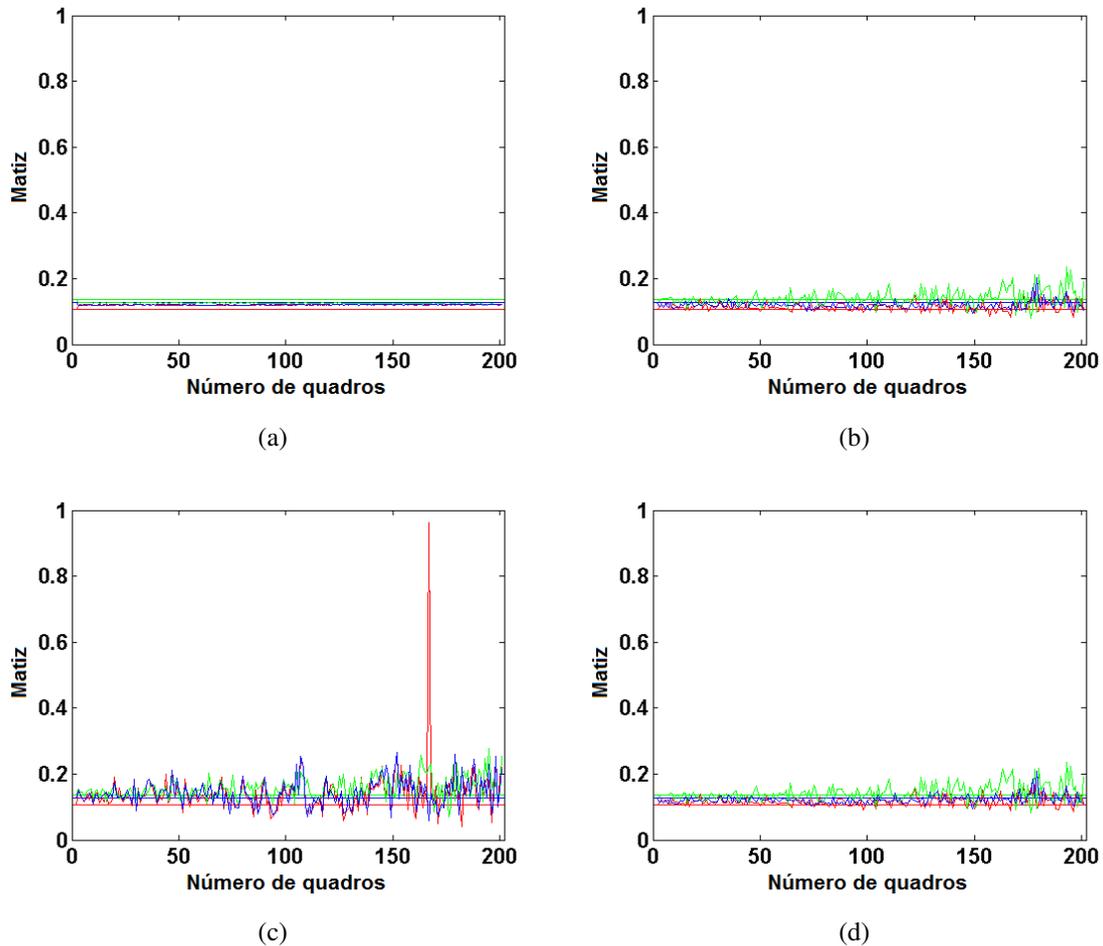
Figura 5.8: Variação do matiz das imagens capturadas. (a)-(f) Sequência de 1 a 6.



Fonte: Elaborada pelo autor

valores do matiz das curvas vermelha e azul permitindo um bom ajuste dessas curvas em relação às linhas base como pode ser visto nas Figuras 5.11(b) e 5.11(c). Estas curvas con-

Figura 5.9: Variação do matiz das imagens corrigidas da Sequência 1 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.



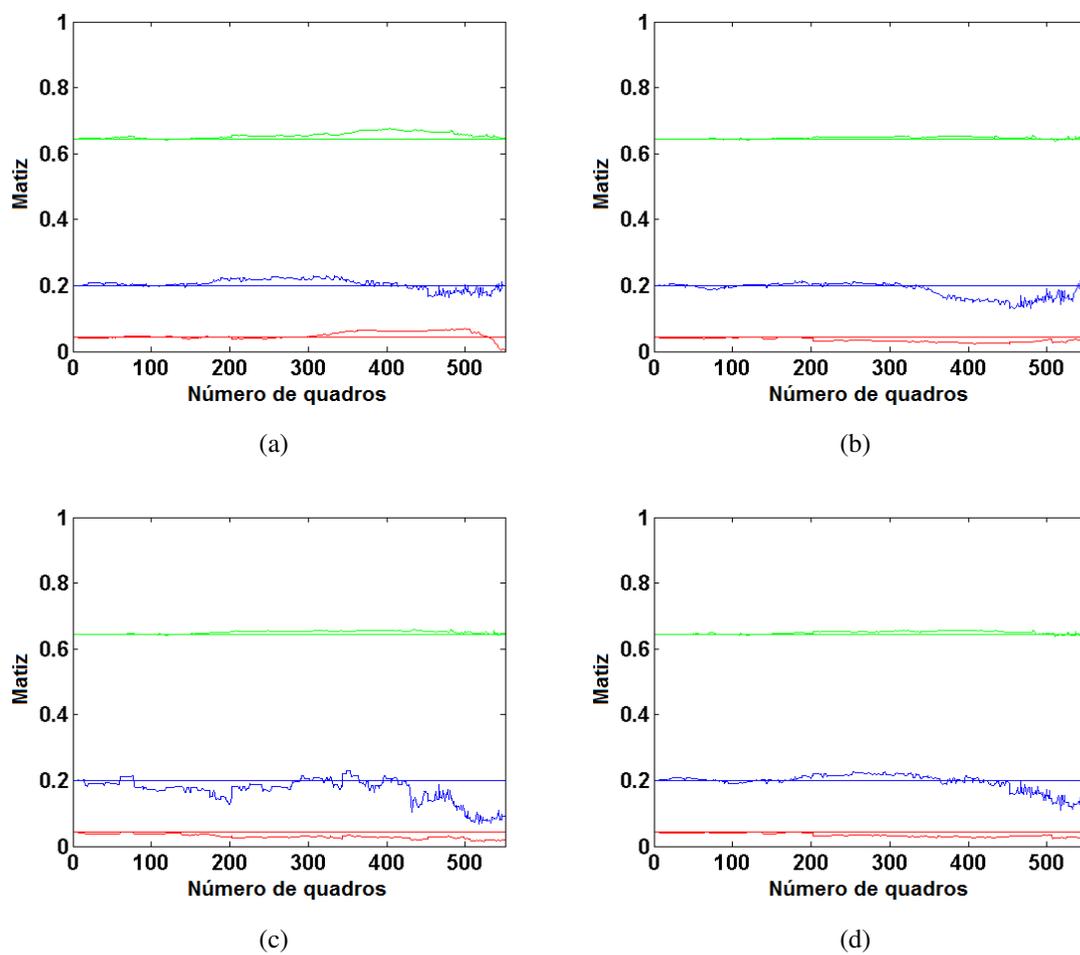
Fonte: Elaborada pelo autor

vergiavam para um único valor de matiz a partir do centésimo quadro nas imagens originais da sequência (Figura 5.8(c)). No entanto, os valores corrigidos do matiz da curva verde ficaram variando entre 0,1 e 0,5 sem se aproximarem da linha base. Isto evidencia a capacidade limitada do ajuste feito por ganhos individuais nos canais R, G e B.

Por outro lado, o MFTA aproximou as três curvas das linhas base satisfatoriamente (Figura 5.11(a)). Apenas a partir do quadro 150 um pequeno desvio pode ser notado. Uma correção de cores utilizando nove parâmetros permite um ajuste mais abrangente. O algoritmo *Max-RGB* seguiu a tendência do *Gray-World* e *Gray-Edge*, mas com resultados visualmente piores.

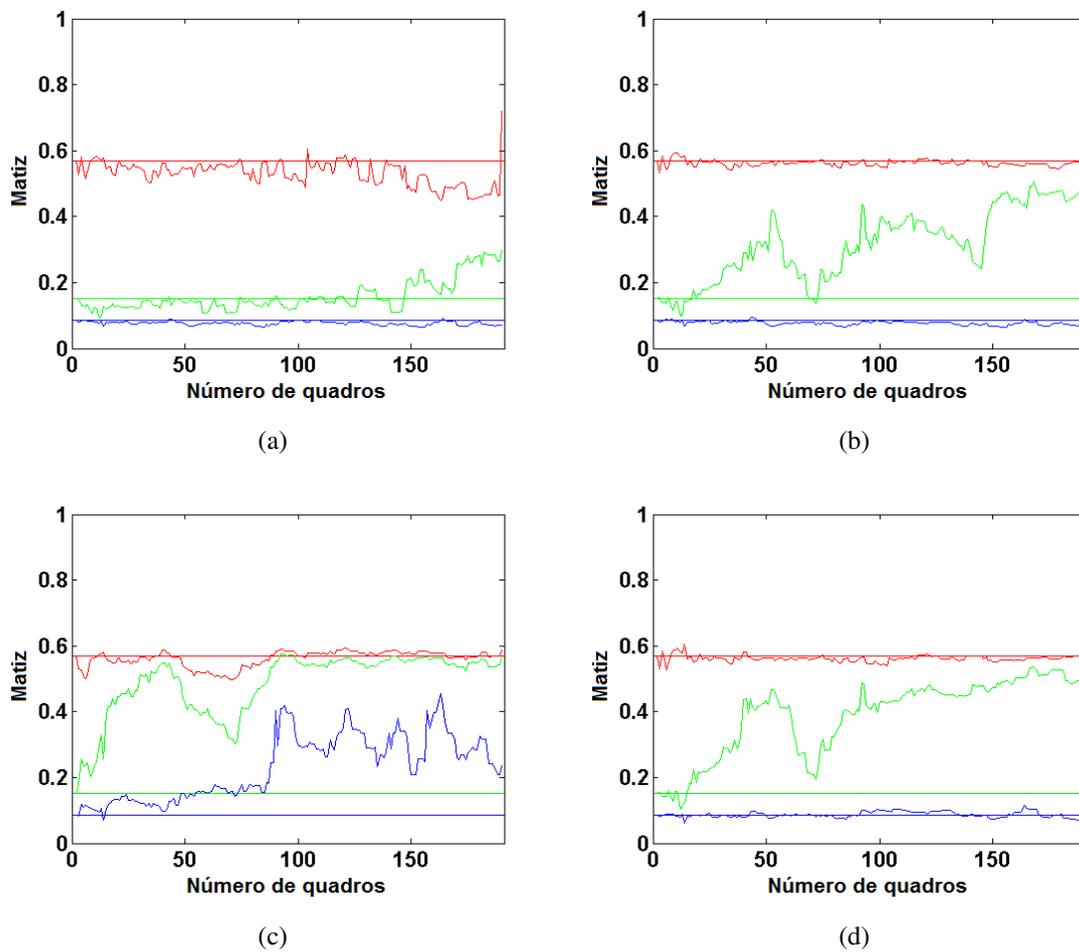
O resultados obtidos para a quarta sequência podem ser vistos na Figura 5.12. O MFTA e o *Gray-World* apresentaram os melhores ajustes das curvas como podem ser vistos nas

Figura 5.10: Variação do matiz das imagens corrigidas da Sequência 2 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.



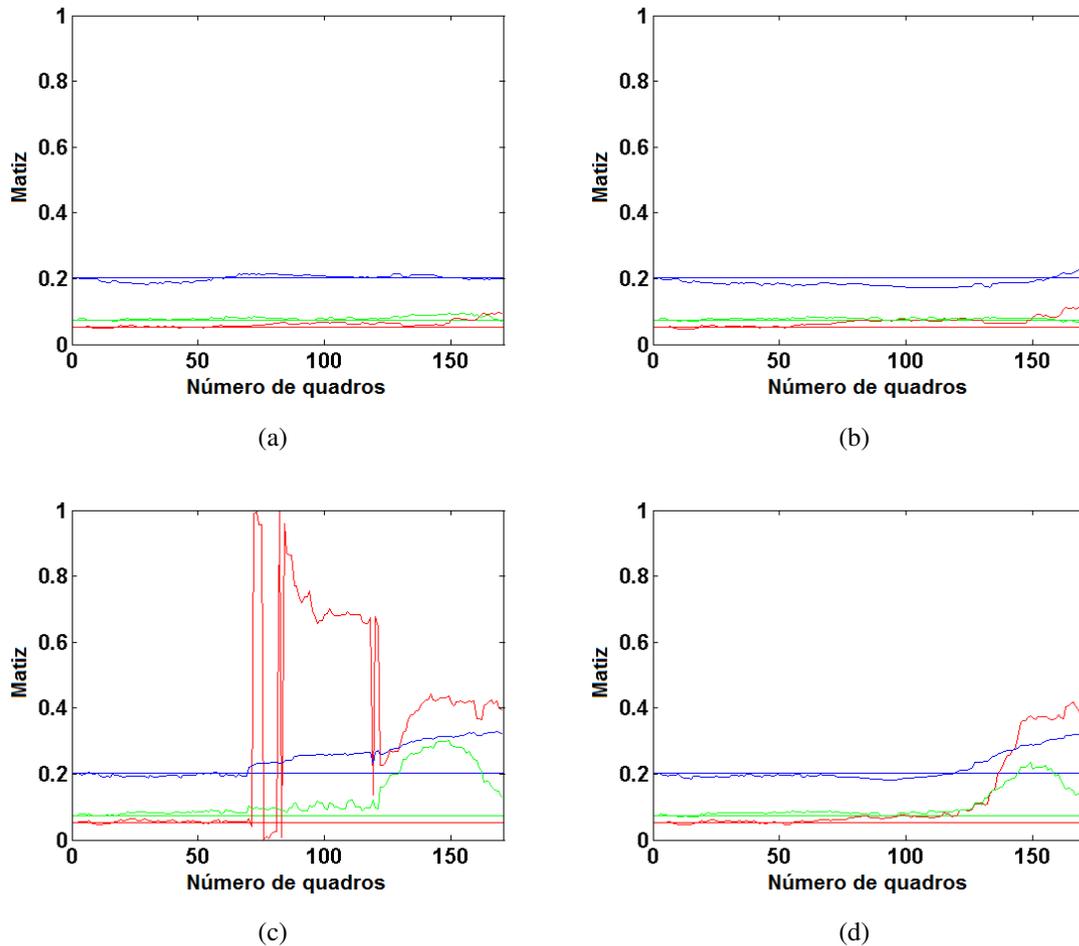
Fonte: Elaborada pelo autor

Figura 5.11: Variação do matiz das imagens corrigidas da Sequência 3 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.



Fonte: Elaborada pelo autor

Figura 5.12: Variação do matiz das imagens corrigidas da Sequência 4 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.

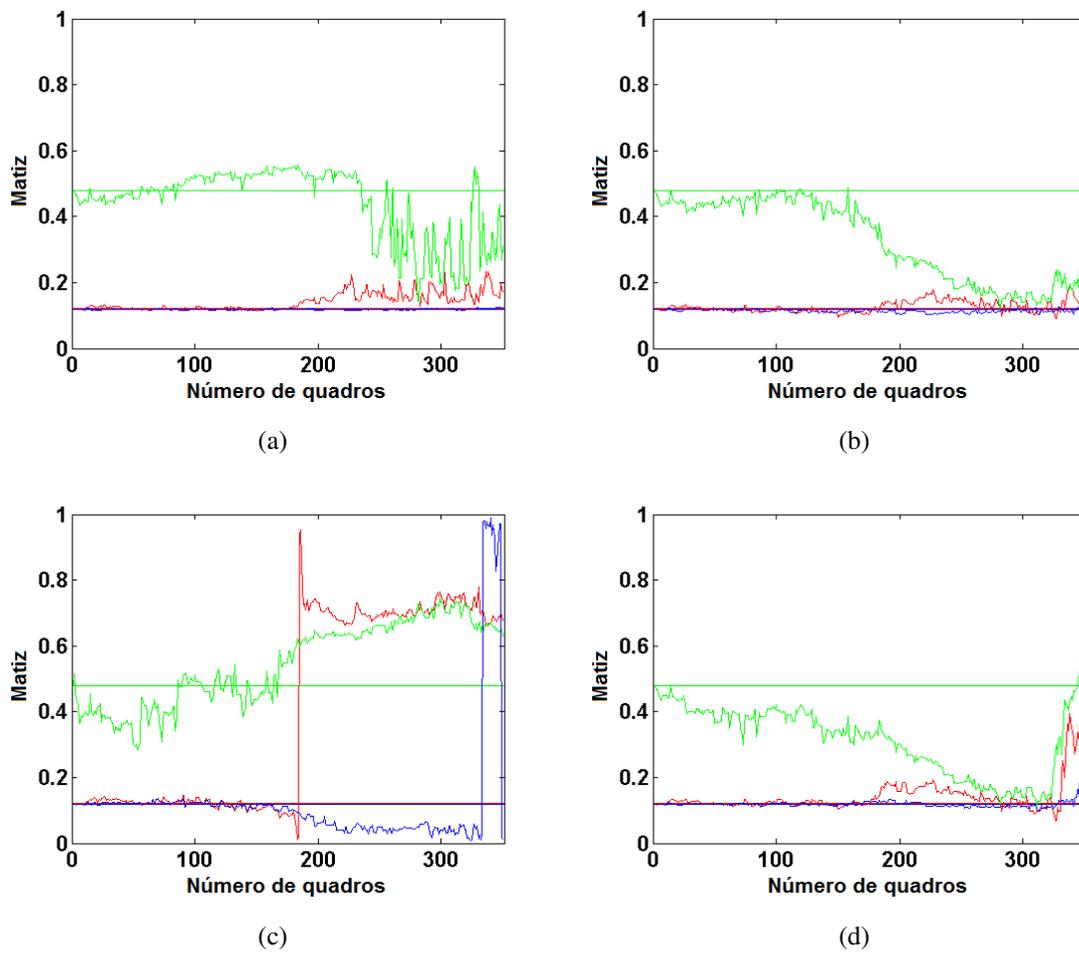


Fonte: Elaborada pelo autor

Figuras 5.12(a) e 5.12(b), respectivamente. Por outro lado, o algoritmo *Max-RGB* apresentou o resultado mais instável (Figura 5.12(c)) enquanto o *Gray-Edge* conseguiu um bom ajuste das curvas até o quadro 125 (Figura 5.12(d)).

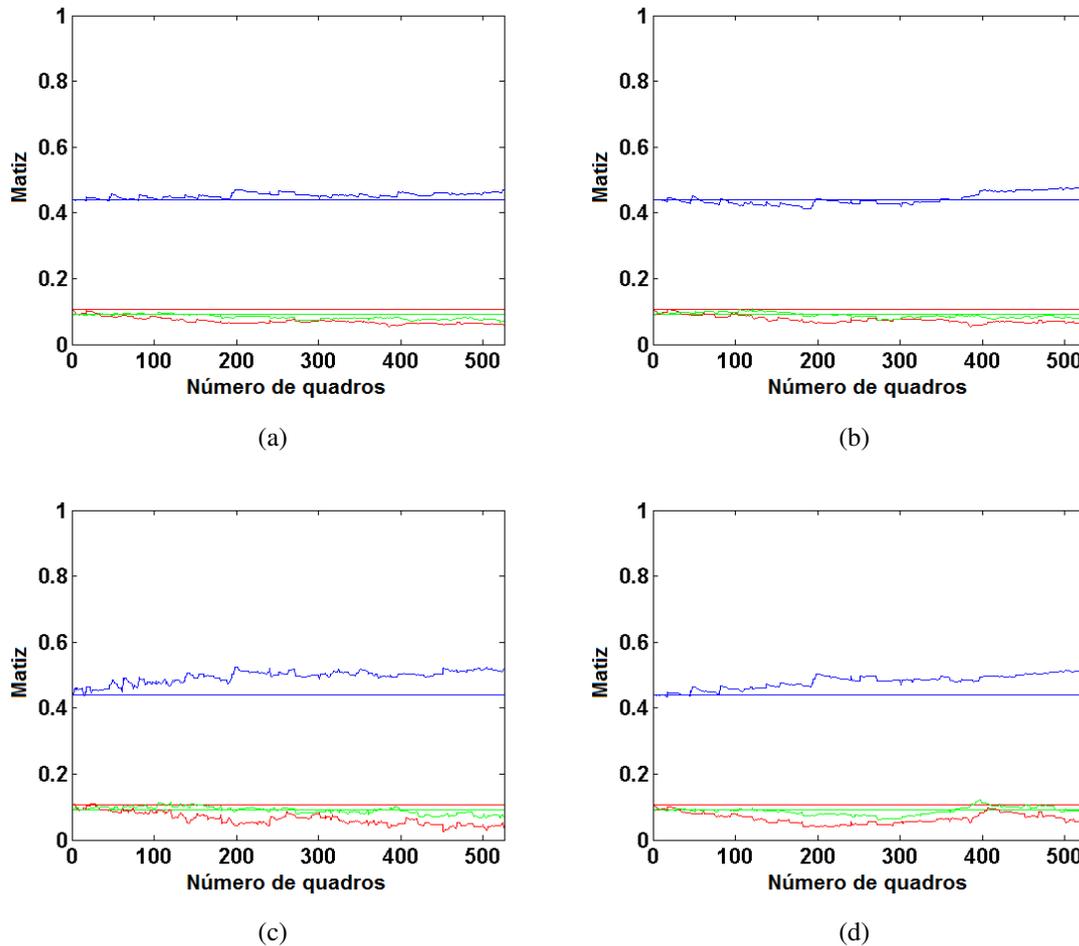
Nos resultados da quinta sequência (Figura 5.13), os algoritmos *Gray-World* e *Gray-Edge* apresentaram uma boa correção para as curvas vermelha e azul, mas não tão boa para a curva verde. O algoritmo *Max-RGB* aumentou o desvio das curvas em relação aos valores originais (Figura 5.13(c)). Diferentemente, o MFTA apresenta os melhores resultados. Ele corrigiu bem as curvas vermelha e azul enquanto manteve a curva verde estável até o quadro 230 (Figura 5.13(a)). A piora no resultado, observada a partir do quadro 230, se deve à diminuição dos níveis de iluminação. Esta redução da luminosidade afetou o resultado de todos os métodos, mas o MFTA foi que se manteve mais estável.

Figura 5.13: Variação do matiz das imagens corrigidas da Sequência 5 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.



Fonte: Elaborada pelo autor

Figura 5.14: Variação do matiz das imagens corrigidas da Sequência 6 pelo (a) MFTA, (b) *Gray-World*, (c) *Max-RGB* e (d) *Gray-Edge*.



Fonte: Elaborada pelo autor

As curvas com os valores originais do matiz para a última sequência não parecem ser tão complexas como pode ser visto na Figura 5.14. No entanto, os algoritmos *Max-RGB* e *Gray-Edge* não apresentaram uma boa correção como pode ser visto nas Figuras 5.14(c) e 5.14(d). O MFTA e o algoritmo *Gray-World* apresentaram melhores resultados (Figuras 5.14(a) e 5.14(b)).

5.3.2 Comparação quantitativa

Uma avaliação mais precisa do desempenho alcançado pelos métodos comparados é obtida por uma análise quantitativa. Para isso, o erro normalizado entre o matiz das cores corrigidas e o matiz do primeiro quadro, dado pela Equação 3.14, é utilizado como medida

de comparação. A Tabela 5.1 mostra o erro normalizado para as seis sequências de imagens avaliadas. A segunda coluna possui os valores do erro normalizado para os valores de matiz originais em relação às linhas base.

Tabela 5.1: Erro normalizado para comparação dos algoritmos

Sequência	Original	MFTA	Gray World	Max RGB	Gray Edge
1	0,18184	0,03106	0,03907	0,10018	0,04181
		(-82,92%)	(-78,51%)	(-44,91%)	(-77,01%)
2	0,03855	0,03167	0,02977	0,05089	0,03017
		(-17,86%)	(-22,78%)	+32,03%	(-21,72%)
3	0,77857	0,08446	0,18686	0,49665	0,25551
		(-89,15%)	(-76,00%)	(-36,21%)	(-67,18%)
4	0,28824	0,02198	0,03652	0,27825	0,13040
		(-92,37%)	(-87,33%)	(-3,47%)	(-54,76%)
5	0,29271	0,10161	0,16912	0,36548	0,19795
		(-65,29%)	(-42,22%)	+24,86%	(-32,38%)
6	0,25150	0,05927	0,05340	0,10922	0,09350
		(-76,43%)	(-78,77%)	(-56,57%)	(-62,82)

Entre os métodos avaliados, o algoritmo *Max-RGB* é o que apresenta o pior resultado em todas as sequências. Para as Sequências 2 e 5 este algoritmo produz um resultado em que o erro normalizado é maior que o erro normalizado da variação do matiz original. Os resultados do erro normalizado da primeira sequência mostram que o MFTA é mais eficiente para reduzir a variação do matiz que os algoritmos *Gray-World* e *Gray-Edge*. Isto por que os erros normalizados desses algoritmos são 25,77% e 34,59%, respectivamente, maiores que o erro normalizado do MFTA.

Para a segunda sequência de imagens, o MFTA e os algoritmos *Gray-World* e *Gray-Edge* mostraram um resultado similar. A máxima diferença entre os valores do erro normalizado para estes algoritmos é de apenas 6,38%. A terceira sequência alcançou uma grande correção quando o MFTA foi utilizado. Neste caso, o erro normalizado reduziu cerca de 89,15%. Por

outro lado, uma menor redução do erro normalizado, 76,00% and 67,18%, foram obtidas para os algoritmos *Gray-World* e *Gray-Edge*, respectivamente.

A maior redução do erro normalizado ocorreu para a Sequência 4 quando o MFTA apresentou um valor de apenas 3,737 em números absolutos. Este valor representa uma redução de 92,37% em relação à variação original do matiz. O segundo melhor resultado foi alcançado pelo algoritmo *Gray-World* com 6,209, mas este valor representa um erro normalizado que é 66,15% maior que aquele alcançado pelo MFTA. O algoritmo *Gray-World* apresenta um erro normalizado de 59,192 na Sequência 5 enquanto o MFTA apresentou um erro normalizado que é 39,92% menor. Este resultado confirma a evidência de que o MFTA apresentou um desempenho melhor que os outros métodos quando o nível de iluminação é reduzido.

Para a última sequência de imagens, o algoritmo *Gray-World* apresentou o melhor resultado reduzindo o erro normalizado em 78,77%. O MFTA também produziu um bom resultado com uma redução de 76,43% que é muito similar à redução alcançada pelo algoritmo *Gray-World*.

5.4 Experimentos com o MFTT

Na Seção 4.4, o desempenho do MFTT foi avaliado para uma sequência de imagens com movimento relativo entre a câmera e a cena. Esta sequência é de pequena duração, cerca de 50s e 1500 quadros, e foi capturada na condição de céu nublado. Estas características permitem considerar a iluminação ambiente como sendo constante e uniforme durante toda esta sequência. Nestas condições, quaisquer distorções que sejam identificadas nas cores corrigidas podem ser atribuídas ao processo de *tracking* necessário em sequências com deslocamento.

Distorções desse tipo foram observadas na Seção 3.4 quando o MFTA com *tracking* das regiões coloridas foi utilizado para a mesma sequência de imagens. Por outro lado, o MFTT apresentou melhor resultado que o MFTA e foi capaz de atenuar a distorção identificada. No entanto, o uso de uma sequência de imagens sem variação da iluminação não permite identificar de forma ampla o comportamento deste método e todos os seus efeitos. Assim, nesta seção, alguns experimentos são apresentados nos quais o MFTT é aplicado em uma sequência de imagens capturadas sob variação de iluminação enquanto ocorre deslocamento da câmera. Estes experimentos são realizados para diferentes tamanhos de janela de quadros, w , definida na Seção 4.4.

As Figuras 5.15(a) e 5.15(b) mostram, respectivamente, o primeiro e o último quadros da

nova sequência de imagens. Esta sequência foi capturada no mesmo ambiente que a sequência citada anteriormente. Ela é composta, aproximadamente, de 31500 quadros e possui um tempo total de 17 minutos e 30 segundos. Este longo período possibilitou o registro de uma grande variação da intensidade da iluminação durante o por do sol. O movimento executado durante a captura das imagens se caracteriza por uma translação em direção às construções que aparecem na cena. A câmera estava segura pelas mãos enquanto se caminhava lentamente. A sequência original foi capturada com resolução 1920x1080 *pixels*. No entanto, imagens reduzidas, possuindo 480x270 *pixels*, foram utilizadas na aplicação do método. Estas imagens são segmentadas em regiões quadradas de 9x9 *pixels* para gerar as amostras de cor.

Figura 5.15: Primeiro e último quadros da sequência com movimento e variação de iluminação.



(a)



(b)

Fonte: Elaborada pelo autor

A sequência de imagens com movimento relativo não possui um ponto fixo de onde se possa observar a variação da cor enquanto ocorre a variação da iluminação. Então, para se conseguir avaliar os resultados, três pontos localizados no primeiro quadro e que aparecem nos quadros seguintes foram escolhidos para terem suas cores medidas. A Figura 5.16 mostra a posição desses pontos no primeiro quadro. A posição destes pontos foi escolhida de

forma a se obter uma área uniforme com uma cor característica. No entanto, a área de cor vermelha e a área de cor verde apresentam uma textura que pode interferir nos resultados. Estes mesmos pontos são selecionados manualmente nas imagens corrigidas subsequentes e suas coordenadas são registradas. O registros destes pontos é feito a cada 30 segundos. Isto equivale a um intervalo de 900 quadros entre dois registros consecutivos.

Figura 5.16: Pontos escolhidos para avaliação das cores.



Fonte: Elaborada pelo autor

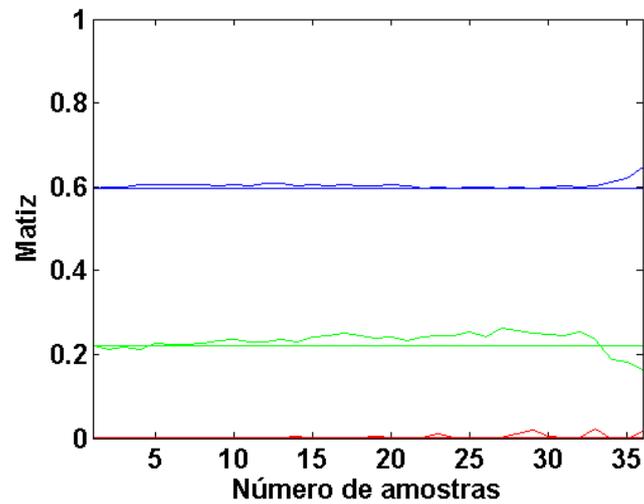
Este esquema gerou 36 amostras de cor para cada ponto marcado na Figura 5.16. O matiz original das amostras de cor, provenientes das imagens capturadas, estão plotados na Figura 5.17.

É possível observar que a variação do matiz não é tão significativa como variação de intensidade sugerida pelas imagens das Figuras 5.15(a) e 5.15(b). Isto pode estar associado à densidade das nuvens que encobriam o céu durante a captura das imagens. As nuvens funcionam como um difusor espalhando de forma mais uniforme a luminosidade, mas também funcionam com um filtro bloqueando determinados comprimentos de onda. De qualquer forma, a variação do matiz é suficiente para a análise do MFTT com diversos tamanhos de janela de quadros w .

Assim, seis experimentos foram realizados com o valor da janela de quadros, w , valendo 15, 30, 150, 450, 900 e 1800 quadros relativos à 1/2, 1, 5, 15, 30 e 60 segundos. Os resultados obtidos mostram que o tamanho da janela de quadros, w , interfere na estabilidade da correção das cores. Para $w = 15$ e $w = 30$, os resultados da variação do matiz dos três pontos marcados na Figura 5.16, provenientes das imagens processadas, são mostrados nas Figuras 5.18(a) e 5.18(b), respectivamente. É possível observar que o método apresenta uma significativa distorção do matiz das cores para pequenos valores de w .

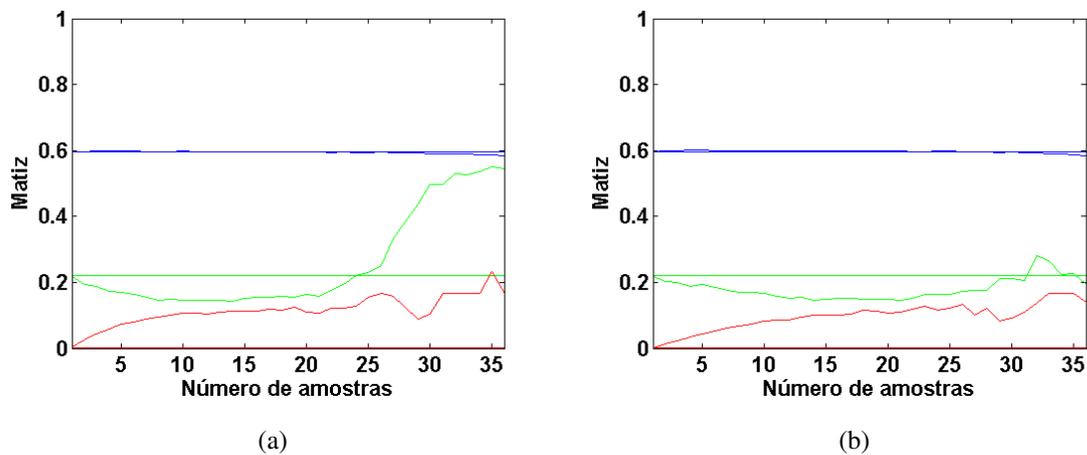
Esta distorção diminui progressivamente com o aumento do tamanho da janela. Para

Figura 5.17: Variação do matiz original proveniente das imagens capturadas para os três pontos escolhidos.



Fonte: Elaborada pelo autor

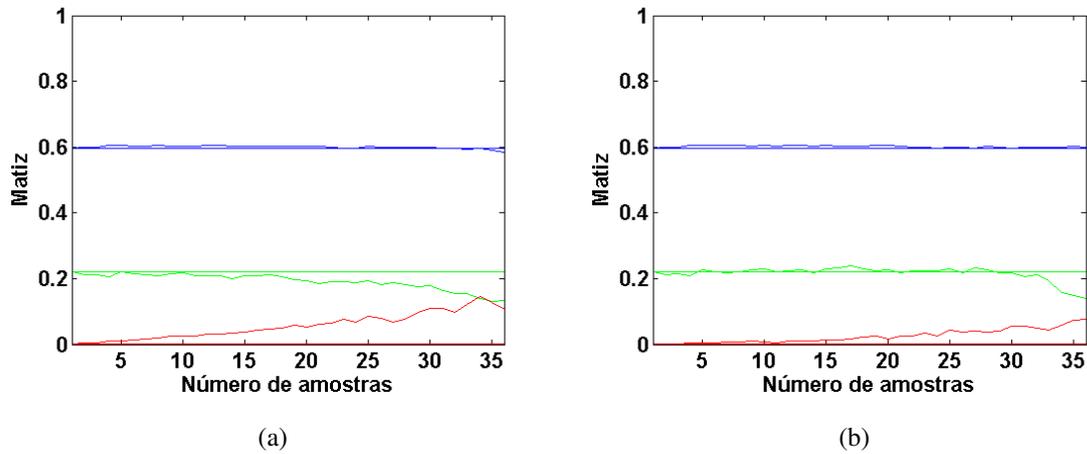
Figura 5.18: Variação do matiz das imagens corrigidas para (a) $w = 15$, (b) $w = 30$ quadros.



Fonte: Elaborada pelo autor

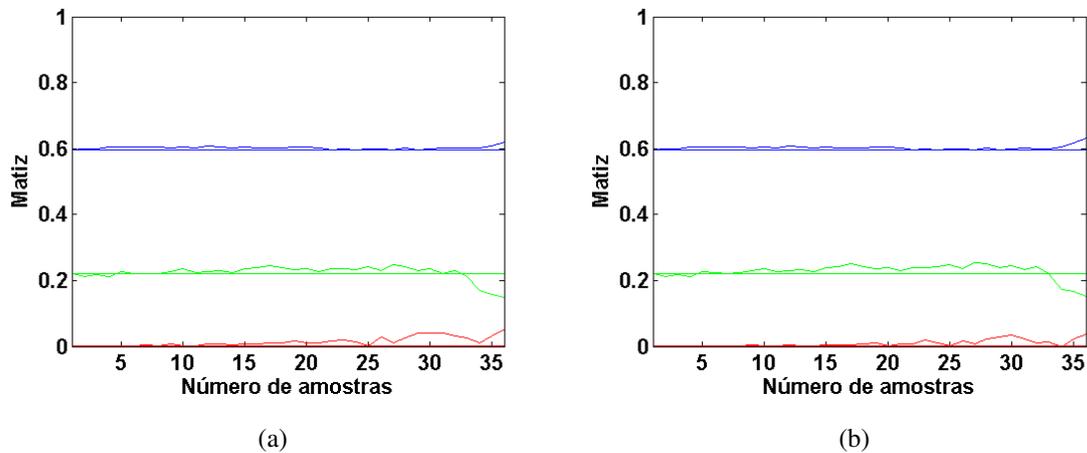
$w = 150$ e $w = 450$ quadros os resultados são melhores como mostram as Figuras 5.19(a) e 5.19(b), respectivamente. Na Figura 5.20(a) é possível observar que um aumento na janela para $w = 900$ ainda promove uma melhoria na estabilidade do matiz das cores. No entanto, para $w = 1800$ o acréscimo na qualidade da correção não é tão significativo, como mostrado a Figura 5.20(b). Isto indica que há um limite para tamanho da janela de quadro em que melhorias na qualidade da correção possam ser alcançadas.

Figura 5.19: Variação do matiz das imagens corrigidas para (a) $w = 150$, (b) $w = 450$ quadros.



Fonte: Elaborada pelo autor

Figura 5.20: Variação do matiz das imagens corrigidas para (a) $w = 900$ e (b) $w = 1800$ quadros.



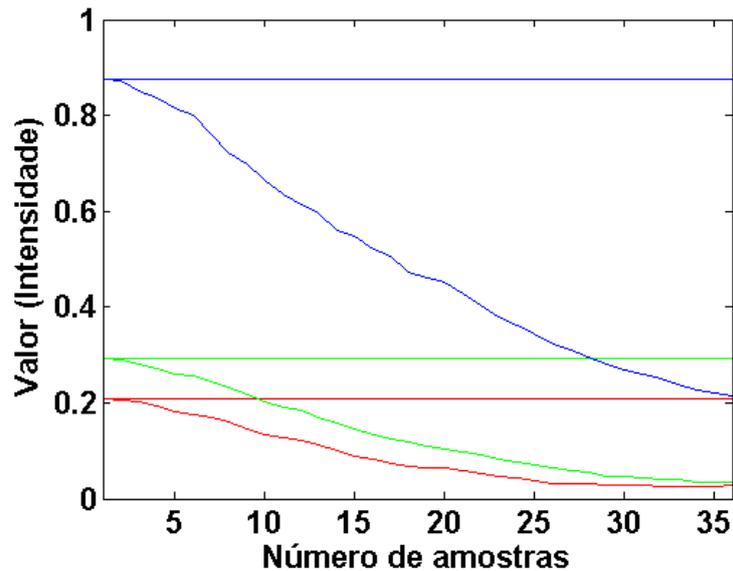
Fonte: Elaborada pelo autor

Estes resultados evidenciam a capacidade do MFTT de neutralizar os efeitos da introdução de ruídos quando há movimento relativo entre a câmera e a cena. Estes ruídos aparecem devido à erros de localização do processo de *tracking* e ao processo dinâmico de inserção e retirada de regiões coloridas. Estes dois processos são necessários para sequências de imagens com movimento relativo.

Na Figura 5.20(b) é possível observar, para as últimas amostras, um certo desvio na correção do matiz. Este desvio na correção do matiz das três cores pode estar relacionado

com a intensidade das cores capturadas uma vez que elas alcançam valores muito baixos para os quadros finais da sequência como pode ser visto na Figura 5.21.

Figura 5.21: Variação da intensidade das cores proveniente das imagens capturadas para os três pontos escolhidos.

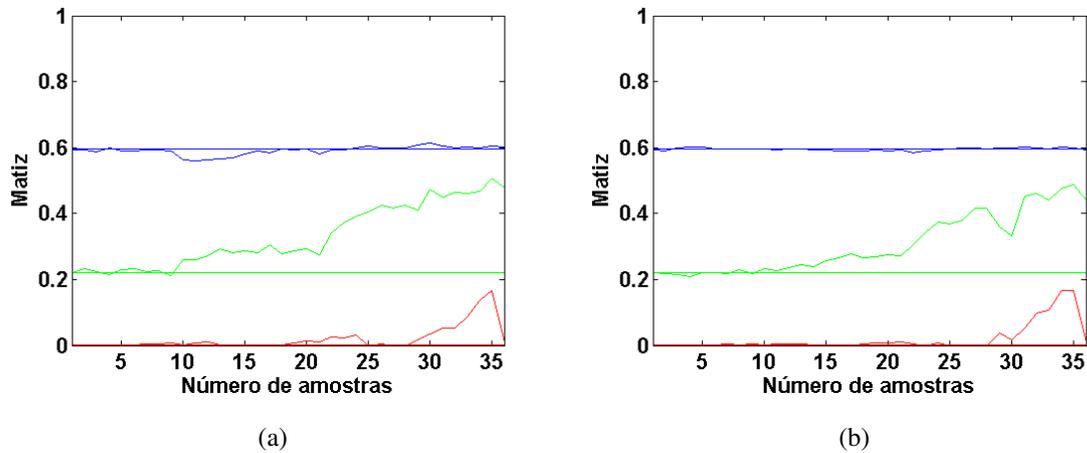


Fonte: Elaborada pelo autor

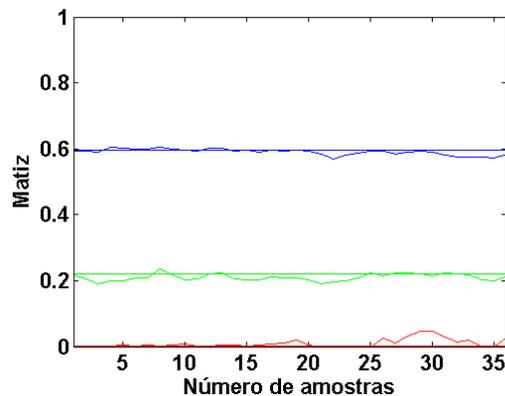
A dificuldade de identificação das cores em baixos níveis de intensidade é uma característica presente inclusive no sistema visual humano. No olho humano, as células sensíveis às cores, chamadas de cones, ficam desativadas para baixos níveis de iluminação. Apenas os bastonetes respondem nestas condições. A influência dos baixos níveis de iluminação também podem ser observada em outros métodos de correção de cores. As Figuras 5.22(a) e 5.22(b) mostram o resultado do matiz das cores corrigidas pelos métodos *Max-RGB* e *Gray-Edge*. É possível observar como a baixa luminosidade interfere nos resultados da correção feita por eles.

O método *Gray-World* mostrou-se mais menos suscetível a esta variação de luminosidade e apresentou um resultado na correção próximo daquele alcançado pelo método proposto aqui. Este resultado pode ser visto na Figura 5.23

Os resultados apresentados mostram a importância da escolha do tamanho da janela de quadros. Eles também evidenciam que a faixa de valores é relativamente ampla permitindo que seja ajustada para atender um compromisso com o desempenho computacional sem que haja distorção nas cores corrigidas.

Figura 5.22: (a) Correção pelo método *Max-RGB* e (b) correção pelo método *Gray-Edge*.

Fonte: Elaborada pelo autor

Figura 5.23: Correção pelo método *Gray-World*.

Fonte: Elaborada pelo autor

5.5 Conclusão

Neste capítulo foram mostrados alguns experimentos que justificam a escolha do processo de segmentação adotado. Estes experimentos mostraram uma variação intrínseca na quantidade e na forma das regiões segmentadas quando a segmentação por crescimento de regiões e o método *graph-based image segmentation algorithm* (Felzenszwalb e Huttenlocher, 2004) são utilizados. Esta variação foi controlada quando a segmentação por quadrados de tamanho fixo foi utilizada.

Os experimentos realizados com o MFTA mostraram bons resultados, para os exemplos

testados, quando comparados àqueles obtidos com alguns dos principais algoritmos conhecidos na literatura quando sequência de imagens sem movimento relativo são utilizadas. Por fim, os experimentos com o MFTT permitiram verificar que o ajuste do tamanho da janela de quadros anulou a distorção introduzida nas cores durante o processo de correção em sequências com movimento relativo entre a câmera e a cena.

Capítulo 6

Conclusões e Trabalhos Futuros

A constância de cor é um problema mal posto e de difícil solução. Por conta disso, diversas suposições são feitas em relação à iluminação, aos sensores e às superfícies para a formulação de métodos computacionais que corrijam as variações das cores sob variação do espectro de frequências da fonte de iluminação. No entanto, quando estas suposições falham, por exemplo, em imagens de cenas naturais, os algoritmos não apresentam estabilidade satisfatória na correção das cores. Uma dessas suposições é relativa à largura de banda do espectro de frequência dos sensores de imagem.

A maioria dos algoritmos consideram as curvas espectrais dos sensores como sendo banda estreita e as aproximam de funções delta. Esta suposição permite uma simplificação no processo de correção de cores caracterizada pela multiplicação dos canais de cor R, G e B de uma imagem capturada por três fatores de ganho distintos para produzir uma imagem corrigida. Estes fatores podem ser obtidos pela razão direta entre os valores R, G, e B da cor representativa de um iluminante padrão pelo valores R, G e B da cor representativa do iluminante da cena que originou a imagem capturada. Assim, a maioria dos algoritmos de constância de cor adota um iluminante de referência e busca estimar a cor do iluminante da cena se valendo de outras suposições relativas ao processo de formação da imagem.

Nesta tese, uma abordagem menos restritiva em relação às curvas espectrais dos sensores é adotada de forma que o processo de correção de cores alcance um aspecto mais adaptativo por meio da utilização de uma matriz linear de transformação com nove parâmetros. Para que isto seja possível, um conjunto de cores identificadas na cena é utilizado em um processo iterativo que gera cores corrigidas tomando como referência as cores corrigidas em iterações anteriores aplicadas em uma sequência de imagens. Este conceito foi, primeiramente, aplicado no MAIA, Seção 3.2, que apresentou uma significativa sensibilidade aos ruídos provocando uma convergência indesejável que produzia imagens com somente valores de cinza. Este efeito somente foi anulado quando o conceito da filtragem temporal foi

introduzido no Método usando Amostras de referência da Imagem Anterior (MAIA) dando origem ao Método do Filtro Temporal para Amostras de referência (MFTA) na Seção 3.3.

Para as sequências utilizadas nos experimentos, os benefícios do uso desta matriz de transformação ficaram evidentes na apresentação do MFTA na Seção 3.3. Os resultados alcançados por este método, vistos na Seção 5.3, são melhores que aqueles apresentados pelos algoritmos comumente apresentados na literatura tais como *Gray-World*, *Max-RGB* e *Gray-Edge* quando sequências de imagens sem movimento relativo entre a câmera e a cena são utilizadas.

No entanto, a introdução de movimento relativo entre a câmera e a cena, Seção 3.4, evidenciou a sensibilidade do MFTA aos ruídos proveniente dos desvios de localização quando um processo de *tracking* é utilizado para localizar as regiões coloridas. Estes ruídos exerceram influência mesmo após as tentativas de melhoria da qualidade das cores de referência descritas nas Seções 4.2.1 e 4.2.2.

Para alcançar uma maior estabilidade no processo de correção de cores em sequências de imagens com movimento relativo entre a câmera e a cena, o conceito da filtragem temporal, inicialmente aplicado sobre amostras de cor contaminadas com ruídos, foi extrapolado para ser aplicado nas transformações obtidas em cada iteração. Esta extrapolação está fundamentada no fato de que amostras de cor contaminadas geram transformações também contaminadas. Assim, para remover os ruídos presentes nas transformações, uma filtragem temporal foi utilizada buscando uma transformação média, como visto na Seção 4.4.

Os resultados obtidos na Seção 5.4, em que o Método do Filtro Temporal para Transformações (MFTT) foi avaliado para uma sequência de imagens obtidas sob variação de iluminação e com deslocamento da câmera, mostram a capacidade de estabilização na correção de cores na presença de ruídos introduzidos pelo processo de *tracking*. Este processo é necessário em sequência com movimento para que as mesmas regiões coloridas seja utilizadas durante o deslocamento.

A grande variedade de situações que envolvem a iluminação e a refletância das superfícies presentes nos ambientes comuns sofre com a. Neste contexto, a luminosidade destes ambientes pode assumir níveis variados de intensidade criando regiões mais claras e outras mais escuras. As regiões mais claras, por exemplo, aquelas que recebem incidência direta da luz solar estão sujeitas a uma composição espectral da radiação eletromagnética diferente das regiões mais escuras que são iluminadas por reflexões múltiplas. Estas zonas mais claras e mais escuras são conhecidas na fotografia como regiões de alta luz e baixa luz. Esta composição de iluminação está presente em ambientes relativamente comuns, mas por questão de simplicidade ela foi desconsiderada quando a condição de iluminação uniforme foi admitida na definição do problema tratado nesta tese.

Assim, uma primeira direção para os trabalhos futuros é a adequação dos métodos apresentados de forma que considerem a existência de regiões de alta luz e baixa luz. Isto pode ser feito pela segmentação de cada imagem capturada para gerar duas imagens distintas que contenham estas regiões permitindo que as correções de cor sejam feitas separadamente. Outras considerações feitas como a existência de superfícies foscas e a pouca influência dos ângulos de incidência e reflexão das superfícies podem ser tratadas como problemas de reflexão especular. O tratamento destes casos poderia ser feito pela remoção das regiões identificadas com este tipo de reflexão.

A segunda direção para os trabalhos futuros está relacionada aos ruídos. Como visto anteriormente, a eficiência dos métodos apresentados nesta tese está associada diretamente à capacidade de atenuação dos ruídos que são introduzidos no processo de correção das cores. Estes ruídos basicamente possuem três origens até agora identificadas considerando o modelo de formação de imagem adotado. A primeira origem é a própria imagem que, dependendo do processo de captura, apresenta melhor ou pior qualidade. A segunda origem de ruídos está no processo de cálculo das correções que introduz variações que se acumulam progressivamente. Por fim, a terceira origem está associada ao processo de *tracking*. Este processo introduz ruídos, em princípio de duas formas, pelo erro de localização e pelo processo dinâmico de inclusão e exclusão de regiões coloridas. Outras fontes de ruídos podem, eventualmente, serem identificadas caso diferentes modelos de formação da imagem sejam adotados. Por exemplo, adotando o modelo de refletâncias especulares a influência dos ângulos de inclinação deve ser considerada.

Dessa forma, alguns trabalhos futuros podem ser feitos direcionados à identificação e classificação da influência das diversas configurações de movimento e de iluminação possíveis nas sequências de imagens. Isto permite caracterizar para quais situações os métodos apresentados são mais eficazes. Este resultado, uma vez encontrado, permitiria um estudo que envolvesse a composição dos métodos apresentados de forma que suas correções pudessem ser sobrepostas.

A terceira direção para os trabalhos futuros aponta para a definição de heurísticas capazes de orientar a escolha de parâmetros intrínsecos dos métodos apresentados. Um destes parâmetros é o tamanho dos quadrados utilizados na segmentação das imagens. Uma heurística que leve em consideração as texturas e os artefatos existentes nas imagens pode determinar o tamanho mais adequado para a segmentação em regiões quadradas mantendo uma relação de compromisso entre a eficiência computacional e a quantidade de regiões identificadas. Esta heurística também pode evoluir para uma abordagem dinâmica da escolha do tamanho dos quadrados. Inicialmente com tamanhos adaptativos entre quadros subsequentes e posteriormente dentro do mesmo quadro.

Outro parâmetro importante é a largura da janela de quadros utilizada no MFTT. Este parâmetro pode ser determinado por uma heurística que considere, por exemplo, a velocidade dos movimentos, ou seja, a variação espacial para que a atenuação dos ruídos introduzidos seja mais eficaz durante o deslocamento da câmera. A frequência de segmentação das imagens também é um parâmetro que pode ser definido por uma heurística baseada na velocidade de deslocamento.

Uma quarta direção está relacionada com a melhoria de processos incorporados aos métodos apresentados. Estas melhorias seriam para substituir os algoritmos não tão eficientes empregados em diversos pontos do código. Um exemplo é o caso de procura linear que poderia ser substituída por algoritmos mais eficientes. Estas melhorias permitiriam, por exemplo, a aplicação do método em tempo real. Os três principais processos sujeitos a esta otimização são:

- Segmentação: melhorias no processo de segmentação das imagens podem ser buscadas para uma melhor identificação de áreas coloridas úteis no processo de correção de cores. Além disso, a eficiência computacional deve ser considerada para melhoria dos resultados.
- *Tracking*: O processo de *tracking* utilizado não apresenta nenhuma forma de otimização. Ele busca as regiões em um espaço definido ao redor de um ponto fixo. Uma estimativa do deslocamento da imagem pode contribuir para que este espaço de busca seja melhor adaptado.
- Filtragem: O processo de filtragem também pode ser melhor estudado para que um filtro com uma sintonia mais adequada e com maior ordem seja implementado utilizando um menor número de amostras.

Experimentos que envolvam uma transformação não-linear capaz de obter um mapeamento entre cores de quadros consecutivos mais precisos podem integrar um quinta direção para trabalhos futuros. Por exemplo, o uso de redes neurais capazes de aprender a transformação mais apropriada em um processo iterativo de correção é uma alternativa ainda a ser explorada.

Referências Bibliográficas

- Agarwal, V., Abidi, B. R., Koschan, A., e Abidi, M. A. (2006a). An overview of color constancy algorithms. *Journal of Pattern Recognition Research*, 1(1):42–54.
- Agarwal, V., Gribok, A., Koschan, A., e Abidi, M. (2006b). Estimating illumination chromaticity via kernel regression. Em *Image Processing, 2006 IEEE International Conference on*, páginas 981–984.
- Agarwal, V., Gribok, A. V., e Abidi, M. A. (2007). Machine learning approach to color constancy. *Neural Netw.*, 20(5):559–563.
- Almonfrey, D., Vassallo, R. F., Konzen, A., Schneebeli, H.-J. A., e Covre, V. B. (2011). Uso da constância de cor na robótica móvel. Em *X Simpósio Brasileiro de Automação Inteligente (SBAI'11)*, páginas 1256–1261.
- Bajcsy, R., Lee, S., e Leonardis, A. (1989). *Computational Aspects of Color Constancy*. GRASP LAB: General Robotics and Active Sensory Perception Laboratory. University of Pennsylvania, School of Engineering and Applied Science, Department of Computer and Information Science.
- Barnard, K. (2000). Improvements to gamut mapping colour constancy algorithms. Em *Computer Vision - ECCV 2000*, volume 1842 of *Lecture Notes in Computer Science*, páginas 390–403. Springer Berlin Heidelberg.
- Barnard, K., Ciurea, F., e Funt, B. (2001). Sensor sharpening for computational color constancy. *J. Opt. Soc. Am. A*, 18(11):2728–2743.
- Barnard, K. e Funt, B. (1998). Experiments in sensor sharpening for color constancy. *Color and Imaging Conference*, 1998(1):43–46.
- Bay, H., Ess, A., Tuytelaars, T., e Gool, L. V. (2008). Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346 – 359. Similarity Matching in Computer Vision and Multimedia.

- Ben-Israel, A. e Greville, T. (2003). *Generalized Inverses: Theory and Applications*. CMS Books in Mathematics. Springer.
- Bianco, S. e Schettini, R. (2011). Computational color constancy. Em *Visual Information Processing (EUVIP), 2011 3rd European Workshop on*, páginas 1–7.
- Brainard, D. H. e Freeman, W. T. (1997). Bayesian color constancy. *J. Opt. Soc. Am. A*, 14(7):1393–1411.
- Brainard, D. H. e Wandell, B. A. (1986). Analysis of the retinex theory of color vision. *J. Opt. Soc. Am. A*, 3(10):1651–1661.
- Bril, M. H. (1980). Computer-simulated object-color recognizer. *MIT Research Laboratory of Electronics Progress Report*, 122. Disponível em: <http://dspace.mit.edu/bitstream/handle/1721.1/56775/RLE_PR_122_XXVIII.pdf?sequence=1>. Acesso em: 25 jan 2014.
- Brill, M. H. (1978). A device performing illuminant-invariant assessment of chromatic relations. *Journal of Theoretical Biology*, 71(3):473 – 478.
- Brouk, I., Nemirovsky, A., e Nemirovsky, Y. (2008). Analysis of noise in cmos image sensor. Em *Microwaves, Communications, Antennas and Electronic Systems, 2008. COMCAS 2008. IEEE International Conference on*, páginas 1–8.
- Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26.
- Cardei, V. C., Funt, B., e Barnard, K. (2002). Estimating the scene illumination chromaticity by using a neural network. *J. Opt. Soc. Am. A*, 19(12):2374–2386.
- Celebi, M. E. e Bogdan, S. (2014). *Advances in low-level color image processing*, volume 11 of *Lecture notes in computational vision and biomechanics*. Springer, Dordrecht.
- CIE (2004). *Colorimetry CIE 15:2004*. CIE technical report. Commission internationale de l’Eclairage, CIE Central Bureau, 3 edição.
- Cohen, J. (1964). Dependency of the spectral reflectance curves of the munsell color chips. *Psychonomic Science*, 1(1-12):369–370.
- Dartnall, H. J. A., Bowmaker, J. K., e Mollon, J. D. (1983). Human visual pigments: Microspectrophotometric results from the eyes of seven persons. *Proceedings of the Royal Society of London B: Biological Sciences*, 220(1218):115–130.
- Demaagd, K., Oliver, A., Oostendorp, N., e Scott, K. (2012). *Practical Computer Vision with SimpleCV: The Simple Way to Make Technology See*. O’Reilly Media, Inc.

- Deshmukh, K. S. (2010). Color image segmentation: a review.
- Dixon, E. R. (1978). Spectral distribution of australian daylight. *J. Opt. Soc. Am.*, 68(4):437–450.
- D’Zmura, M. e Iverson, G. (1993a). Color constancy. i. basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *J. Opt. Soc. Am. A*, 10(10):2148–2165.
- D’Zmura, M. e Iverson, G. (1993b). Color constancy. ii. results for two-stage linear recovery of spectral descriptions for lights and surfaces. *J. Opt. Soc. Am. A*, 10(10):2166–2180.
- D’Zmura, M. e Iverson, G. (1994). Color constancy. iii. general linear recovery of spectral descriptions for lights and surfaces. *J. Opt. Soc. Am. A*, 11(9):2389–2400.
- D’Zmura, M. e Lennie, P. (1986). Mechanisms of color constancy. *J. Opt. Soc. Am. A*, 3(10):1662–1672.
- Ebner, M. (2004). A parallel algorithm for color constancy. *J. Parallel Distrib. Comput.*, 64(1):79–88.
- Ebner, M. (2007). *Color Constancy*. Wiley Publishing.
- Ebner, M. (2009). Color constancy based on local space average color. *Machine Vision and Applications*, 20(5):283–301.
- Ebner, M. e Herrmann, C. (2005). On determining the color of the illuminant using the dichromatic reflection model. Em *Pattern Recognition*, volume 3663 of *Lecture Notes in Computer Science*, páginas 1–8. Springer Berlin Heidelberg.
- Felzenszwalb, P. e Huttenlocher, D. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181.
- Finlayson, G. (1996). Color in perspective. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(10):1034–1038.
- Finlayson, G., Chatterjee, S., e Funt, B. (1996). Color angular indexing. Em *Computer Vision - ECCV ’96*, volume 1065 of *Lecture Notes in Computer Science*, páginas 16–27. Springer Berlin Heidelberg.
- Finlayson, G. e Hordley, S. (2000). Improving gamut mapping color constancy. *Image Processing, IEEE Transactions on*, 9(10):1774–1783.

- Finlayson, G., Hordley, S., e Hubel, P. (2001). Color by correlation: a simple, unifying framework for color constancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1209–1221.
- Finlayson, G., Hordley, S., e Morovic, P. (2005). Colour constancy using the chromagenic constraint. Em *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, páginas 1079–1086 vol. 1.
- Finlayson, G., Hordley, S., e Tastl, I. (2006). Gamut constrained illuminant estimation. *International Journal of Computer Vision*, 67(1):93–109.
- Finlayson, G. D., Drew, M. S., e Funt, B. V. (1994a). Color constancy: generalized diagonal transforms suffice. *J. Opt. Soc. Am. A*, 11(11):3011–3019.
- Finlayson, G. D., Drew, M. S., e Funt, B. V. (1994b). Spectral sharpening: sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563.
- Finlayson, G. D. e Hordley, S. D. (1999). Selection for gamut mapping colour constancy. *Image Vision Comput.*, 17(8):597–604.
- Finlayson, G. D. e Hordley, S. D. (2001). Color constancy at a pixel. *J. Opt. Soc. Am. A*, 18(2):253–264.
- Fischer, G. e Sajjaa, M. (2008). whitebalpr - a new method for automatic white balance. *Conference on Colour in Graphics, Imaging, and Vision*, 2008(1):202–207.
- Forsyth, D. A. (1990). A novel algorithm for color constancy. *Int. J. Comput. Vision*, 5(1):5–36.
- Foster, D. H. (2011). Color constancy. *Vision Research*, 51(7):674 – 700.
- Fredembach, C. e Finlayson, G. (2008). The bright-chromagenic algorithm for illuminant estimation. *Journal of Imaging Science and Technology*, 52(4).
- Funt, B., Drew, M., e Ho, J. (1991). Color constancy from mutual reflection. *International Journal of Computer Vision*, 6(1):5–24.
- Funt, B. e Finlayson, G. (1995). Color constant color indexing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(5):522–529.
- Funt, B. e Jiang, H. (2003). Nondiagonal color correction. Em *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, páginas I–481–4 vol.1.

- Funt, B. e Xiong, W. (2004). Estimating illumination chromaticity via support vector regression. *Color and Imaging Conference*, 2004(1):47–52.
- Funt, B. V., Barnard, K., e Martin, L. (1998). Is machine colour constancy good enough? Em *Proceedings of the 5th European Conference on Computer Vision-Volume I - Volume I*, ECCV '98, páginas 445–459, London, UK, UK. Springer-Verlag.
- Garud, H., Pudipeddi, U., Desappan, K., e Nagori, S. (2014). A fast color constancy scheme for automobile video cameras. Em *Signal Processing and Communications (SPCOM), 2014 International Conference on*, páginas 1–6.
- Gehler, P. V., Rother, C., Blake, A., Minka, T., e Sharp, T. (2008). Bayesian color constancy revisited. Em *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, páginas 1–8.
- Gershon, R., Jepson, A. D., e Tsotsos, J. K. (1987). From [r,g,b] to surface reflectance: computing color constant descriptors in images. Em *Proceedings of the 10th international joint conference on Artificial intelligence - Volume 2*, IJCAI'87, páginas 755–758. Morgan Kaufmann Publishers Inc.
- Geusebroek, J. M., van den Boomgaard, R., Smeulders, A. W. M., e Gevers, T. (2003). Color constancy from physical principles. *Pattern Recognition Letters*, 24(11):1653–1662.
- Gevers, T. e Smeulders, A. W. (1999). Color-based object recognition. *Pattern Recognition*, 32(3):453 – 464.
- Gijsenij, A. e Gevers, T. (2007). Color constancy by local averaging. Em *Image Analysis and Processing Workshops, 2007. ICIAPW 2007. 14th International Conference on*, páginas 171–174.
- Gijsenij, A. e Gevers, T. (2011). Color constancy using natural image statistics and scene semantics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(4):687–698.
- Gijsenij, A., Gevers, T., e van de Weijer, J. (2009). Physics-based edge evaluation for improved color constancy. Em *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, páginas 581–588.
- Gijsenij, A., Gevers, T., e van de Weijer, J. (2011). Computational color constancy: Survey and experiments. *Trans. Img. Proc.*, 20(9):2475–2489.
- Gijsenij, A., Gevers, T., e van de Weijer, J. (2012a). Improving color constancy by photometric edge weighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):918–929.

- Gijsenij, A., Lu, R., e Gevers, T. (2012b). Color constancy for multiple light sources. *Trans. Img. Proc.*, 21(2):697–707.
- Gonzalez, R. C. e Woods, R. E. (2008). *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Gu, J., Jiang, J., Susstrunk, S., e Liu, D. (2013). What is the space of spectral sensitivity functions for digital color cameras? Em *Proceedings of the 2013 IEEE Workshop on Applications of Computer Vision (WACV)*, WACV '13, páginas 168–179, Washington, DC, USA. IEEE Computer Society.
- Healey, G. e Kondepudy, R. (1994). Radiometric ccd camera calibration and noise estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(3):267–276.
- Healey, G. e Slater, D. (1994). Global color constancy: recognition of objects by use of illumination-invariant properties of color distributions. *J. Opt. Soc. Am. A*, 11(11):3003–3010.
- Huang, C., Liu, Q., e Li, X. (2010). Color image segmentation by seeded region growing and region merging. Em *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*, volume 2, páginas 533–536.
- Humod, A. H., Leman, Z., Samin, R., e Ashrafi, N. (2015). Improving productivity and quality in manufacturing by applying computer vision systems (image processing technique). *Journal of Electrical and Electronics Engineering*, 10(5):01–09.
- Hurlbert, A. (1999). Colour vision: Is colour constancy real? *Current Biology*, 9(15):R558–R561.
- Ikonomakis, N., Pataniotis, K. N., e Venetsanopoulos, A. N. (2000). Unsupervised seed determination for a region-based color image segmentation scheme. Em *Image Processing, 2000. Proceedings. 2000 International Conference on*, volume 1, páginas 537–540 vol.1.
- Jameson, K. A., Highnote, S. M., e Wasserman, L. M. (2001). Richer color experience in observers with multiple photopigment opsin genes. *Psychonomic Bulletin & Review*, 8(2):244–261.
- Judd, D. B., Macadam, D. L., Wyszecki, G., Budde, H. W., Condit, H. R., Henderson, S. T., e Simonds, J. L. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54(8):1031–1040.

- Kamijo, S., Ikeuchi, K., e Sakauchi, M. (2002). Illumination invariant segmentation of spatio-temporal images by spatio-temporal markov random field model. Em *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 2, páginas 617–622 vol.2.
- Kang, H. (2006). *Computational Color Technology*. Press Monographs. Society of Photo Optical.
- Kawakami, R., Hongxun, Z., Tan, R. T., e Ikeuchi, K. (2013). Camera spectral sensitivity and white balance estimation from sky images. *International Journal of Computer Vision*.
- Kawakami, R., Ikeuchi, K., e Tan, R. (2005). Consistent surface color for texturing large objects in outdoor scenes. Em *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, páginas 1200–1207 Vol. 2.
- Kelly, K. L. (1963). Lines of constant correlated color temperature based on macadam's (u,v) uniform chromaticity transformation of the cie diagram. *J. Opt. Soc. Am.*, 53(8):999–1002.
- Konzen, A. e Schneebeli, H.-J. A. (2007). O problema da constância de cor em visão computacional. Em *VIII Simpósio Brasileiro de Automação Inteligente*, Florianópolis, Brazil.
- Kraft, J. M. e Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences*, 96(1):307–312.
- Kunimitsu, S., Kawabata, K., Asama, H., e Mishima, T. (2003). Detection of object under outdoor environment with binary edge template and its application to automatic cranes. Em *Industrial Technology, 2003 IEEE International Conference on*, volume 1, páginas 140–145 Vol.1.
- Land, E. H. (1977). The Retinex Theory of Color Vision. *Scientific American*, 237(6):108–128.
- Land, E. H. e Mccann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, páginas 1–11.
- Landa, E. e Fairchild, M. (2005). Charting color from the eye of the beholder. *American Scientist*, 93(5):436–443.
- Lee, H.-C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *J. Opt. Soc. Am. A*, 3(10):1694–1699.
- Li, B., Xiong, W., Hu, W., e Funt, B. (2013). Evaluating combinational illumination estimation methods on real-world images.

- Maloney, L. T. e Wandell, B. A. (1986). Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A*, 3(1):29–33.
- Mollon, J. (2006). Monge: The verriest lecture, lyon, july 2005. *Visual Neuroscience*, 23:297–309.
- NASA (2014). Nasa solar system exploration - sun: Facts & figures. Disponível em: <<http://solarsystem.nasa.gov/planets/profile.cfm?Display=Facts&Object=Sun>>. Acesso em: 08 mai. 2014.
- Nieves, J. L., Plata, C., Valero, E. M., e Romero, J. (2008). Unsupervised illuminant estimation from natural scenes: an rgb digital camera suffices. *Appl. Opt.*, 47(20):3574–3584.
- Novak, C. L. e Shafer, S. A. (1991). Supervised color constancy for machine vision. *Proc. SPIE*, 1453:353–368.
- O'Connor, J. J. e Robertson, E. F. (2014). The mactutor history of mathematics archive. Disponível em: <<http://www-history.mcs.st-andrews.ac.uk/Biographies/Monge.html>>. Acesso em: 7 mai. 2014.
- Olkkonen, M., Witzel, C., Hansen, T., e Gegenfurtner, K. R. (2010). Categorical color constancy for real surfaces. *Journal of Vision*, 10(9):16.
- Olsson, C., Kahl, F., e Oskarsson, M. (2006). The registration problem revisited: Optimal solutions from points, lines and planes. Em *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, páginas 1206–1213.
- Parkkinen, J. P. S., Hallikainen, J., e Jaaskelainen, T. (1989). Characteristic spectra of munsell colors. *J. Opt. Soc. Am. A*, 6(2):318–322.
- Prinet, V., Lischinski, D., e Werman, M. (2013). Illuminant chromaticity from image sequences. Em *Computer Vision (ICCV), 2013 IEEE International Conference on*, páginas 3320–3327.
- Ratnasingam, S. e McGinnity, T. (2012). Chromaticity space for illuminant invariant recognition. *Image Processing, IEEE Transactions on*, 21(8):3612–3623.
- Renno, J.-P., Makris, D., Ellis, T., e Jones, G. A. (2005). Application and evaluation of colour constancy in visual surveillance. Em *Proceedings of the 14th International Conference on Computer Communications and Networks, ICCCN '05*, páginas 301–308. IEEE Computer Society.
- Sapiro, G. (1999). Color and illuminant voting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(11):1210–1215.

- Schaefer, G., Hordley, S., e Finlayson, G. (2005). A combined physical and statistical approach to colour constancy. Em *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, páginas 148–153 vol. 1.
- Semo, S. e Spitzer, H. (2000). Color constancy: a biological model and its application for still and video images. Em *Electrical and Electronic Engineers in Israel, 2000. The 21st IEEE Convention of the*, páginas 198–201.
- Simão, J. (2003). *Inspeção visual automática de mamões*. Dissertação (Mestrado em Engenharia Elétrica), Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Espírito Santo, Vitória.
- Simão, J., Andreas Schneebeli, H., e Frizzera Vassallo, R. (2014). An iterative approach for color constancy. Em *Robotics: SBR-LARS Robotics Symposium and Robocontrol (SBR LARS Robocontrol), 2014 Joint Conference on*, páginas 130–135.
- Simão, J., Schneebeli, H.-J. A., e Vassallo, R. F. (2013). Constância de cor em sequências de imagens. Em *XI Simpósio Brasileiro de Automação Inteligente (SBAI'13)*, volume X, páginas 1256–1261, São João del-Rey, Brazil.
- Simão, J., Schneebeli, H.-J. A., e Vassallo, R. F. (2015). Iterative color constancy with temporal filtering for an image sequence with no relative motion between the camera and the scene. *J. Opt. Soc. Am. A*, 32(11):2033–2043.
- Singh, J. e Singh, P. (2013). Automatic seed placement in region growing image segmentation. *Journal of Engineering Computers & Applied Sciences*, 2(7):55–58.
- Sridharan, M. e Stone, P. (2004). Towards on-board color constancy on mobile robots. Em *CRV*, páginas 130–137. IEEE Computer Society.
- Tan, R. T., Nishino, K., e Ikeuchi, K. (2004). Color constancy through inverse-intensity chromaticity space. *J. Opt. Soc. Am. A*, 21(3):321–334.
- Thuillier, G., Hersé, M., Simon, P. C., Labs, D., Mandel, H., Gillotay, D., e Foujols, T. (1998). The visible solar spectral irradiance from 350 to 850 nm as measured by the solspec spectrometer during the atlas i mission. Em Pap, J., Fröhlich, C., e Ulrich, R., editors, *Solar Electromagnetic Radiation Study for Solar Cycle 22*, páginas 41–61. Springer Netherlands.
- Tian, H. (2000). *Noise analysis in CMOS image sensors*. Tese de Doutorado, Stanford University.
- Tominaga, S. e Wandell, B. A. (1989). Standard surface-reflectance model and illuminant estimation. *J. Opt. Soc. Am. A*, 6(4):576–584.

- Toro, J. e Funt, B. (2007). A multilinear constraint on dichromatic planes for illumination estimation. *Image Processing, IEEE Transactions on*, 16(1):92–97.
- Tsukada, M. e Ohta, Y. (1990). An approach to color constancy using multiple images. Em *Computer Vision, 1990. Proceedings, Third International Conference on*, páginas 385–389.
- Tu, K.-Y. (2009). Analysis of camera's images influenced by varying light illumination for design of color segmentation. *J. Inf. Sci. Eng.*, 25(6):1885–1899.
- Upcroft, B., McManus, C., Churchill, W., Maddern, W., e Newman, P. (2014). Lighting invariant urban street classification. Em *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, páginas 1712–1718.
- Urquhart, R. (1982). Graph theoretical clustering based on limited neighbourhood sets. *Pattern Recognition*, 15(3):173 – 187.
- van de Weijer, J., Gevers, T., e Gijsenij, A. (2007a). Edge-based color constancy. *Trans. Img. Proc.*, 16(9):2207–2214.
- van de Weijer, J., Schmid, C., e Verbeek, J. (2007b). Using high-level visual information for color constancy. Em *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, páginas 1–8.
- Vazquez-Corral, J., Vanrell, M., Baldrich, R., e Tous, F. (2012). Color constancy by category correlation. *Image Processing, IEEE Transactions on*, 21(4):1997–2007.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., e von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138:1172–1217.
- Wang, N., Funt, B., Lang, C., e Xu, D. (2011). Video-based illumination estimation. Em *Proceedings of the Third international conference on Computational color imaging, CCIW'11*, páginas 188–198, Berlin, Heidelberg. Springer-Verlag.
- Withagen, P., Groen, F. C. A., e Schutte, K. (2005). Ccd characterization for a range of color cameras. Em *Instrumentation and Measurement Technology Conference, 2005. IMTC 2005. Proceedings of the IEEE*, volume 3, páginas 2232–2235.
- Wyszecki, G. e Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley-Interscience, 2 edição.
- X-Rite (2009). Colorchecker classic. Available at: <http://xritephoto.com/documents/literature/en/ColorData-1p_EN.pdf>. Accessed May 13, 2015.

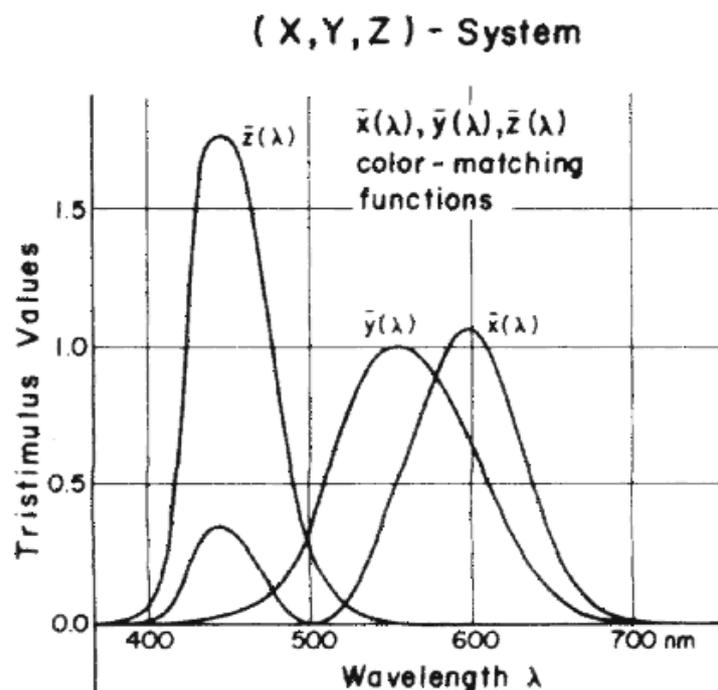
-
- Xiong, W. e Funt, B. (2009). Stereo retinex. *Image and Vision Computing*, 27(1-2):178 – 188.
- Zahn, C. T. (1971). Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions on Computers*, C-20(1):68–86.
- Zaraga, F. e Langfelder, G. (2010). White balance by tunable spectral responsivities. *J. Opt. Soc. Am. A*, 27(1):31–39.

Apêndice A

Espaços de Cores

Em 1931, a *Commission Internationale de l'Éclairage* (CIE) desenvolveu um espaço de cor chamado XYZ. Este espaço de cor é baseado na sensibilidade média de um grupo de indivíduos e representa a resposta visual que caracteriza um observador padrão. A partir de um experimento de comparação de cores feito com indivíduos de visão normal construiu-se as *Color Matching Functions* $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ e $\bar{z}(\lambda)$ que são curvas de correspondência de cor. Estas curvas são mostradas na Figura A.1.

Figura A.1: Curvas de correspondência de cor para o sistema XYZ.



Fonte: (Wyszecki e Stiles, 2000)

Essas curvas representam a resposta espectral de percepção do observador padrão. Os valores X , Y e Z são obtidos através da integral do produto entre as curvas de comparação de cor e a curva espectral de uma cor (Wyszecki e Stiles, 2000). Por exemplo, dado uma cor com uma curva de distribuição espectral $C(\lambda)$, os valores X , Y e Z são obtidos pelas Equações A.1, A.2 e A.3, respectivamente e K é uma constante que normaliza os valores em relação à intensidade do iluminante.

$$X = K \int_{380}^{780} C(\lambda) \bar{x}(\lambda) d\lambda \quad (\text{A.1})$$

$$Y = K \int_{380}^{780} C(\lambda) \bar{y}(\lambda) d\lambda \quad (\text{A.2})$$

$$Z = K \int_{380}^{780} C(\lambda) \bar{z}(\lambda) d\lambda \quad (\text{A.3})$$

Este espaço não possui uma distribuição uniforme em relação à percepção das cores, ou seja, cores que são significativamente distintas para a percepção humana podem possuir valores XYZ muito próximos. Para resolver esse problema foi criado um novo espaço de cor, que possui canais duplos como na teoria de Hering, chamado CIE-LAB.

Os espaços de cor comumente utilizados para a representação em sistemas digitais são: o RGB, formado pelas cores vermelho, verde e azul e usado principalmente para monitores e câmeras digitais, o CMY para impressoras coloridas e o YIQ para transmissão de TV colorida. Usualmente a manipulação de imagens digitais é feita utilizando-se modelos de representação que separam as informações de luminância e crominância. Essa representação é adequada, pois se aproxima da forma como a visão humana percebe as cores. Os modelos de representação mais utilizados são o HSI e o HSV. Nestes modelos a componente H (matiz) representa a cor espectral predominante, a componente S (saturação) representa a quantidade de branco diluída na cor (cores muito saturadas são cores espectrais puras e não possuem influência da cor branca) e as componentes I (intensidade) e V (valor) caracterizam o efeito da luminância. As componentes de cor do espaço RGB podem ser convertidas para os espaços HSI e HSV como descrito abaixo.

- Conversão RGB para HSI:

A conversão de cores no espaço RGB, definido como um cubo unitário, para o espaço HSI é obtida a partir dos valores normalizados do vermelho, verde e azul que satisfazem $r + g + b = 1$ (Gonzalez e Woods, 2008). Os valores HSI no intervalo $[0,1]$ podem

ser obtidos a partir de um conjunto de valores RGB no mesmo intervalo através das Equações A.4, A.5 e A.6.

$$H = \cos^{-1} \left(\frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right) \quad (\text{A.4})$$

$$S = 1 - \frac{3}{R+G+B} \min(R, G, B) \quad (\text{A.5})$$

$$I = \frac{1}{3}(R+G+B) \quad (\text{A.6})$$

- Conversão RGB para HSV:

Os valores de H estão no intervalo [0,360] e as componentes S, V, R, G e B no intervalo [0,1]. Se *Max* for zero, então $S = 0$ e H é não definido.

$$Max = \max(R, G, B) \quad (\text{A.7})$$

$$Min = \min(R, G, B) \quad (\text{A.8})$$

se $Max = R$

$$ht = \frac{G-B}{Max-Min} \quad (\text{A.9})$$

se $Max = G$

$$ht = 2,0 + \frac{G-B}{Max-Min} \quad (\text{A.10})$$

se $Max = B$

$$ht = 4,0 + \frac{R-G}{Max-Min} \quad (\text{A.11})$$

$$H = 60,0 * ht \quad (\text{A.12})$$

se ($H < 0$) corrige H para

$$H = H + 360,0 \quad (\text{A.13})$$

$$S = \frac{Max-Min}{Max} \quad (\text{A.14})$$

$$V = Max \quad (\text{A.15})$$

Computacionalmente a conversão RGB para HSV é mais simples que a conversão RGB para HSI, pois não precisa calcular \cos^{-1} . Assim, o modelo HSV muitas vezes é preferido ao invés do modelo HSI.

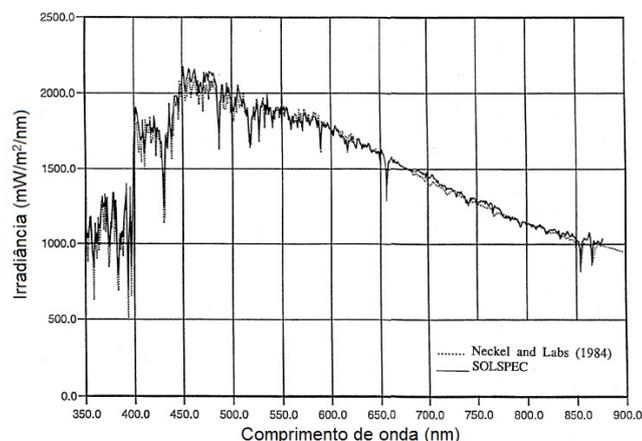
Apêndice B

Modelo dos Iluminantes

As fontes de luz, também chamadas de iluminantes, são caracterizadas por suas curvas de distribuição espectral de potência, $L(\lambda)$, que definem o valor da densidade de potência por comprimento de onda. Entre os principais iluminantes estão o Sol, um iluminante natural, e os diversos tipos de lâmpadas que são iluminantes artificiais.

O espectro da radiação solar está próximo do espectro emitido por um corpo negro com temperatura de 5.800K (NASA, 2014) e sua curva de distribuição espectral de potência na faixa do espectro visível, medida a partir de três espectrômetros (Thuillier et al., 1998), pode se vista na Figura B.1. Na superfície da terra, esta curva espectral varia em função da hora do dia e das condições climática e sofre influências de poeira, nuvens, chuva e neve.

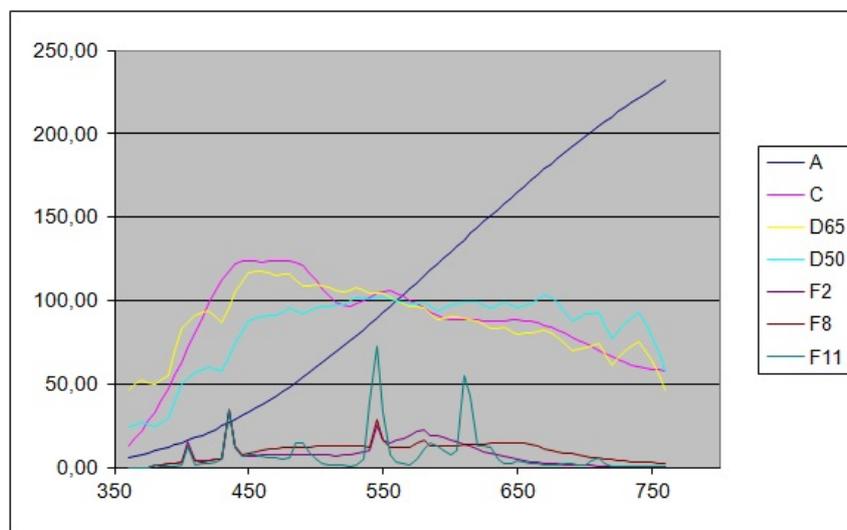
Figura B.1: Curvas espectrais da irradiância solar na faixa visível obtidos por Neckel & Labs e por SOLSPEC/ATLAS 1.



Fonte: (Thuillier et al., 1998)

O CIE, Commission Internationale de L'Eclairage, define um conjunto de iluminantes padronizados (CIE, 2004) que servem de referência para caracterização da luz solar e de lâmpadas artificiais como as incandescentes e fluorescentes. Estes iluminantes são nomeados por letras e números como A, C, D65, F2 e outros. As curvas espectrais de alguns destes iluminantes são mostradas na Figura B.2. Estes iluminantes são difíceis de serem produzidos de forma artificial, mas são fáceis de serem caracterizados matematicamente. Os iluminantes da série D são os mais utilizados, pois caracterizam as diversas fases da luz do dia natural.

Figura B.2: Distribuição espectral relativa de alguns iluminantes padronizados.



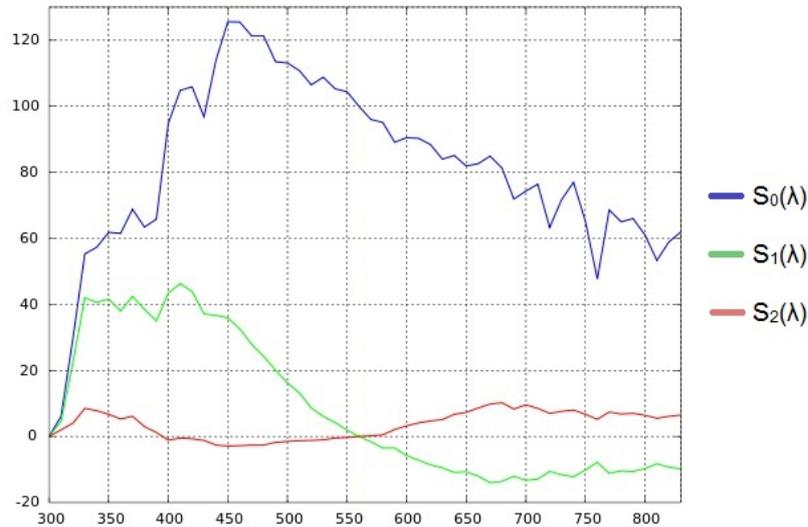
Fonte: (CIE, 2004)

A padronização da série D de iluminantes, feita pelo CIE, deriva do trabalho realizado por Judd et al. (1964). Eles utilizaram um conjunto de 622 amostras da distribuição espectral da luz do dia em uma Análise de Componentes Principais, também conhecida como *principal component analysis* (PCA), mostrando que a curva de distribuição espectral de potência da luz do dia $S(\lambda)$ poderia ser aproximada, de maneira satisfatória, por uma curva média $S_0(\lambda)$ e os dois primeiros componentes característicos $S_1(\lambda)$ e $S_2(\lambda)$, ou seja, uma combinação linear de funções base (Equação B.1). As curvas da distribuição espectral de potência dessas funções base podem ser vista na Figura B.3.

$$S(\lambda) = S_0(\lambda) + M_1 S_1(\lambda) + M_2 S_2(\lambda) \quad (\text{B.1})$$

Judd et al. (1964) mostraram também que as coordenadas de cromaticidade (x, y) da cor do iluminante possuem uma simples relação quadrática entre si (Equação B.2). As coordenadas de cromaticidade x , y e z são uma representação normalizada dos valores X , Y e Z

Figura B.3: Funções base para caracterização dos iluminantes da série D.



Fonte: (CIE, 2004)

do espaço de cor CIE 1931 XYZ. Quando apenas as coordenadas x e y são plotadas em um plano se obtém uma figura chamada diagrama de cromaticidades.

$$y = 2,870x - 3,000x^2 - 0,275 \quad (\text{B.2})$$

Os valores dos coeficientes M_1 e M_2 , definidos a partir dos valores das cromaticidades x e y , são dados pela Equações B.3 e B.4

$$M_1 = \frac{-1,3515 - 1,7703x + 5,9114y}{0,0241 + 0,2562x - 0,7341y} \quad (\text{B.3})$$

$$M_2 = \frac{0,0300 - 31,4424x + 30,0717y}{0,0241 + 0,2562x - 0,7341y} \quad (\text{B.4})$$

Os valores de cromaticidade x e y para uma determinada fase da luz do dia não podem ser determinados de forma direta. Judd et al. (1964) tabelaram estas cromaticidades para alguns valores de temperatura de cor comumente utilizados, tais como 4.800K, 5.500 K, 6.500 K, 7.500 K e 10.000K. Valores de cromaticidade para outras temperaturas de cor devem ser obtidos a partir de figuras definidas por Kelly (1963). O CIE definiu uma aproximação da cromaticidade x em função do valor da temperatura de cor. Para um valor de temperatura de cor T maior ou igual à 4.000K e menor ou igual à 7.000K a componente x de cromaticidade é

dada pela Equação B.5 e para uma temperatura de cor entre 7.000K e 25.000K a componente x é dada pela Equação B.6.

$$x = 0,244063 + 0,09911 \frac{10^3}{T} + 2,9678 \frac{10^6}{T^2} - 4,6070 \frac{10^9}{T^3} \quad (\text{B.5})$$

$$x = 0,237040 + 0,24748 \frac{10^3}{T} + 1,9018 \frac{10^6}{T^2} - 2,0064 \frac{10^9}{T^3} \quad (\text{B.6})$$

Uma vez que o valor da cromaticidade x tenha sido determinado deve ser estimado o valor da cromaticidade y pela expressão quadrática definida na Equação B.2. Com os valores das coordenadas de cromaticidades x e y , os valores de M_1 e M_2 podem ser determinados. Assim, a curva espectral estimada da luz do dia para uma determinada temperatura de cor T pode ser determinada pela Equação B.1.

Muitos trabalhos sobre constância de cor buscam determinar uma representação do iluminante da cena como uma soma ponderada de funções base em um modelo de dimensão finita (Tsukada e Ohta, 1990; Funt et al., 1991; Finlayson et al., 1994a). No entanto, nem todo conjunto de funções espectrais pode servir como funções base (Buchsbaum, 1980). As funções definidas por Judd et al. (1964) modelam com certa precisão as fase da luz do dia, mas falharão para outras fonte de luz como incandescente e fluorescentes. Dixon (1978) também mediu e analisou a distribuição espectral de potência da luz do dia na Austrália e chegou a resultados semelhantes aos de Judd et al. (1964).

Apêndice C

Modelo das Refletâncias

A luz que chega até os sensores de imagem é diretamente proporcional à luz que emerge de uma cena. Esta luz que emerge da cena é resultado da interação entre a luz emitida por um iluminante e a superfície dos objetos presentes na cena. A luz emitida por um iluminante alcança os objetos da cena e é, em parte, absorvida por eles. O restante dessa luz é refletida em diferentes direções. A quantidade de luz refletida é dependente da direção da fonte de luz bem como da direção de quem observa. Superfícies foscas refletem a luz igualmente em todas as direções caracterizando, assim, a reflexão difusa. Neste caso a luz refletida depende somente da direção do iluminante em relação à normal da superfície (Ebner, 2007). Por outro lado, quando a luz incidente é refletida predominantemente em uma única direção a reflexão é chamada de especular. Marcas brilhosas observadas em algumas superfícies são devido à reflexão especular.

A razão entre a luz refletida e a luz incidente em uma superfície caracteriza sua refletância. A refletância de uma superfície é dependente do comprimento de onda λ da radiação eletromagnética sendo refletida, ou seja, a refletância é função do comprimento de onda da luz incidente $r(\lambda)$. Apesar da interação entre luz e superfície ser um fenômeno complexo envolvendo, além da absorção e reflexão, efeitos como refração e difração (Wyszecki e Stiles, 2000), a maioria das superfícies pode ser modelada pela combinação linear da reflexão difusa e da reflexão especular. Assim, considerando $e(\lambda)$ o espectro de luz incidente sobre uma superfície e $r(\lambda)$ a função espectral de refletância dessa superfície, a luz refletida $l(\lambda)$ é dada pela Equação C.1.

$$l(\lambda) = \alpha r(\lambda)e(\lambda) + \beta e(\lambda) \quad (\text{C.1})$$

Nesta equação, α é o fator geométrico associado à reflexão difusa e β é o fator de escala relacionado à reflexão especular. Esta modelagem é conhecida na literatura como modelo de

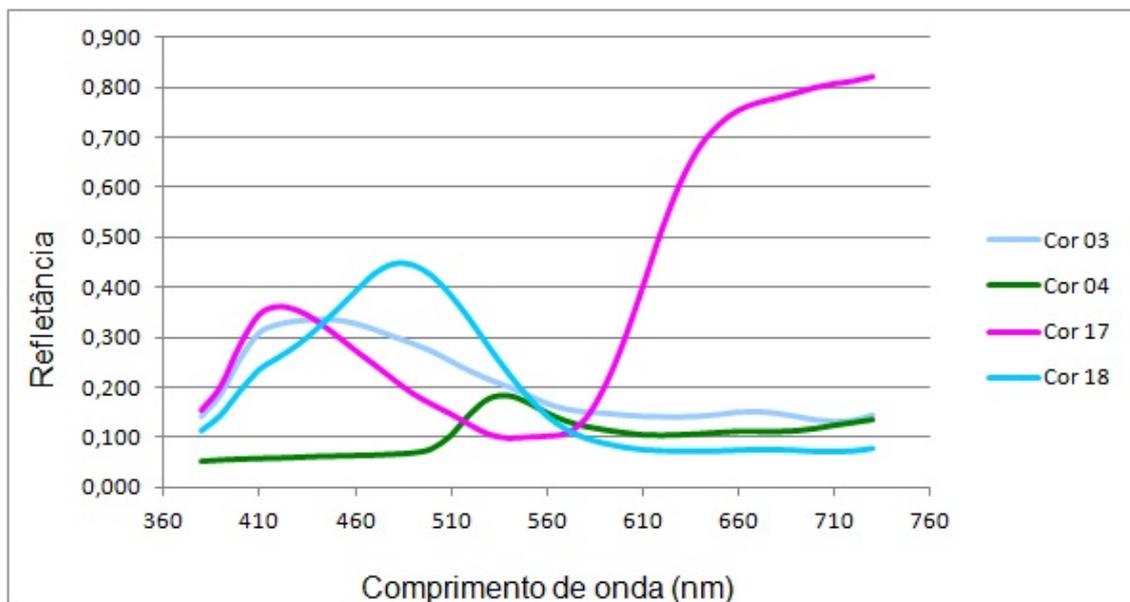
reflexão dicromático (Ebner, 2007). Para superfícies foscas com pouca ou nenhuma reflexão especular o modelo de reflexão difusa ou lambertiano é mais adequado (Equação C.2).

$$I(\lambda) = r(\lambda)e(\lambda) \quad (\text{C.2})$$

O modelo de reflexão difusa, que considera as superfícies como sendo foscas, é o mais utilizado em trabalhos de constância de cor devido sua simplicidade (Gehler et al., 2008; Finlayson et al., 2001; Forsyth, 1990; Finlayson et al., 2006; Ebner, 2009). Entretanto, há também trabalhos que tiram proveito da reflexão especular e consideram o modelo dicromático de reflexão (Ebner e Herrmann, 2005; Schaefer et al., 2005).

A Figura C.1 mostra alguns exemplos de curvas espectrais¹ de refletâncias de superfícies foscas presentes no cartão *Colorchecker*.

Figura C.1: Funções espectrais de refletância de superfícies do *Colorchecker*.



Fonte: Elaborada pelo autor

A curva espectral de uma superfície é uma propriedade intrínseca do objeto observado. Em princípio, estas curvas poderiam assumir a forma de qualquer função contínua. No entanto, a maioria das curvas espectrais naturais existentes podem ser agrupadas formando um conjunto de curvas aproximadamente suaves. Assim, é possível representar este conjunto de refletância por um número finito de funções base. Cohen (1964) examinou um conjunto

¹A base de dados numérica pode ser obtida no endereço eletrônico http://www.babelcolor.com/download/ColorChecker_RGB_and_spectra-Avg20_2006.xls

de 150 fichas de cor selecionadas aleatoriamente de um total de 433 fichas presentes no livro Munsell de cores (Landa e Fairchild, 2005). As curvas de refletância espectral dessas fichas foram medidas utilizando um espectrômetro na faixa de 380 nm a 770 nm em intervalos de 10 nm. Após uma análise de componentes lineares ele concluiu que a soma ponderada de três funções base pode representar com precisão o conjunto inicial de 433 fichas coloridas. No entanto, um estudo similar feito por Parkkinen et al. (1989) produziu uma representação mais precisa das curvas espectrais de refletâncias. Eles utilizaram um conjunto significativamente maior de amostras de refletâncias, um total de 1257 fichas foram utilizadas. Além disso, o intervalo de medição do espectro de frequências foi a metade daquele utilizado por Cohen (1964), 5 nm, na faixa de 400 nm a 700 nm, gerando um conjunto de dados mais precisos. A conclusão foi de que até oito funções base são necessárias para se conseguir uma boa representação das curvas espectrais de refletância das fichas do livro Munsell de cores.

Há também trabalhos que consideram a refletância como a soma ponderada de funções base provenientes de uma expansão de série de Fourier. D’Zmura e Lennie (1986), por exemplo, consideram três funções base sendo uma de valor constante e as outras duas baseadas em cores oponentes, sendo a primeira *Red-Green* e a segunda *Yellow-Blue*.

Apêndice D

Modelo dos Sensores

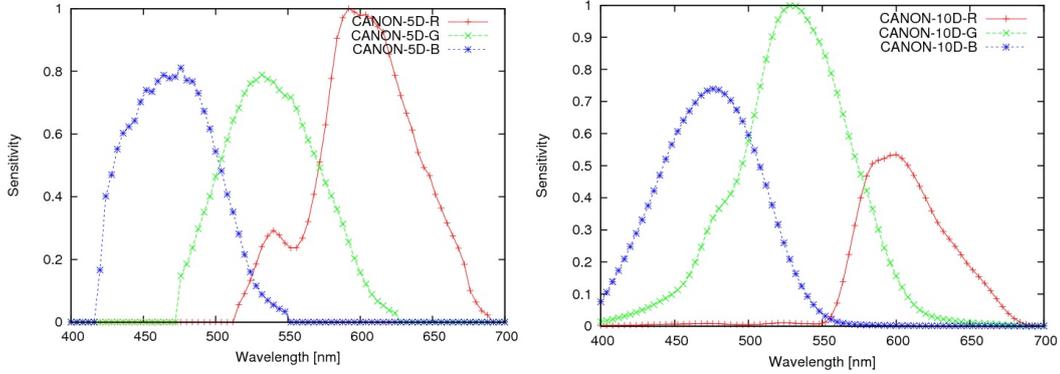
Como citado anteriormente, os sensores das câmeras são construídos baseados no modelo da visão humana. Eles são formados por três canais sensíveis ao espectro visível de frequências. Cada um destes canais possui uma banda distinta de sensibilidade. Eles são chamados *Red*, *Green* e *Blue* e definem o modelo tricromático de cor RGB (Ebner, 2007). Uma função de sensibilidade $S_m(\lambda)$ é definida para cada canal de cor, onde $m \in \{r, g, b\}$. As Figuras D.1, D.2, D.3 e D.4 mostram as curvas de sensibilidade espectral de algumas câmeras¹. Como se pode verificar nestas figuras, as curvas de sensibilidade de todas as câmeras possuem uma região de sobreposição, sendo umas mais significativas que outras. A câmera KODAK DCS 460, por exemplo, possui o canal *blue* sensível a uma faixa de comprimento de onda que ocupa quase todo o espectro visível de frequência. Isto significa que mesmo uma cor saturada com predominância na faixa do verde ou vermelho possuirá uma componente azul significativa. Na verdade, esta cor será capturada como sendo uma cor menos saturada e com um desvio no matiz.

Uma suposição muito comum nos algoritmos de constância de cor considera as curvas espectrais dos sensores como sendo banda estreita. No limite dessa consideração a curva espectral do sensor torna-se uma função Delta de Dirac, ou seja, $S_m(\lambda) = \delta(\lambda - \lambda_m)$, onde $S_m(\lambda)$ é a curva de sensibilidade espectral do sensor para os canais $m \in \{r, g, b\}$ e λ_m é a frequência espectral em que a resposta é diferente de zero (Ebner, 2007). Nesta situação, os valores medidos para os canais R, G e B no ponto x_I da imagem são dados pela Equação D.1 e cuja simplificação é dada conforme apresenta a Equação D.2.

$$C_m(x_I) = \int_{\lambda_1}^{\lambda_2} \delta(\lambda - \lambda_m) L(\lambda) R(\lambda, x_O) d\lambda \quad (\text{D.1})$$

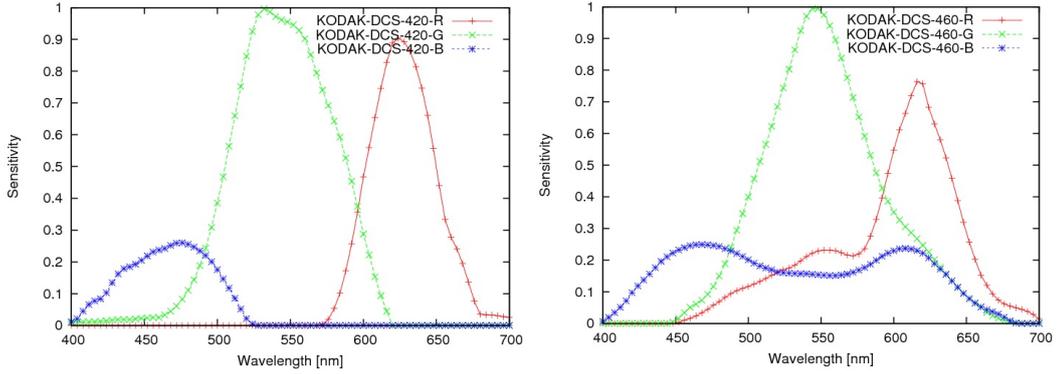
¹A base de dados numérica e as imagens das curvas podem ser obtidas no endereço eletrônico <http://www.cvl.iis.u-tokyo.ac.jp/~rei/research/cs/zhao/database.html>

Figura D.1: Funções de sensibilidade espectral das câmeras CANON 5D e CANON 10D.



Fonte: (Kawakami et al., 2013)

Figura D.2: Funções de sensibilidade espectral das câmeras KODAK DCS 420 e KODAK DCS 460.



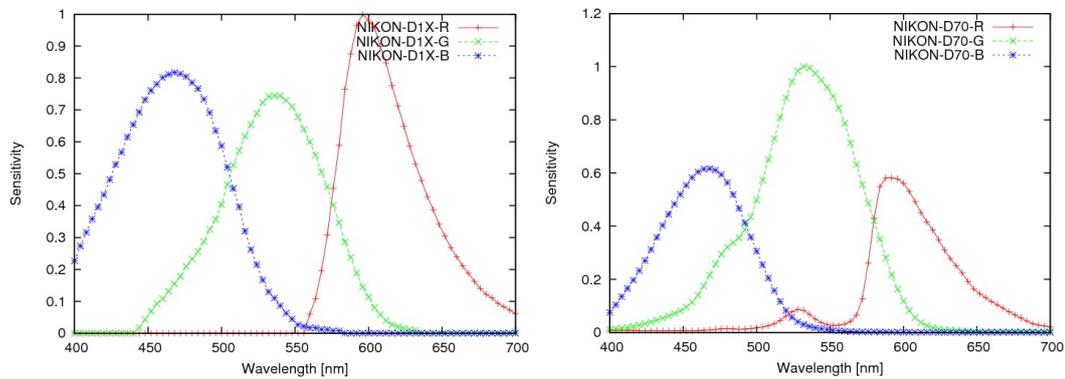
Fonte: (Kawakami et al., 2013)

$$C_m(x_I) = L(\lambda_m)R(\lambda_m, x_O) \quad (D.2)$$

Na Equação D.1, $[\lambda_1, \lambda_2]$ é o intervalo de comprimento de onda de resposta do sensor, $L(\lambda)$ é a radiância emitida pela fonte de luz e $R(\lambda, x_O)$ é a refletância da superfície de um objeto no ponto x_O .

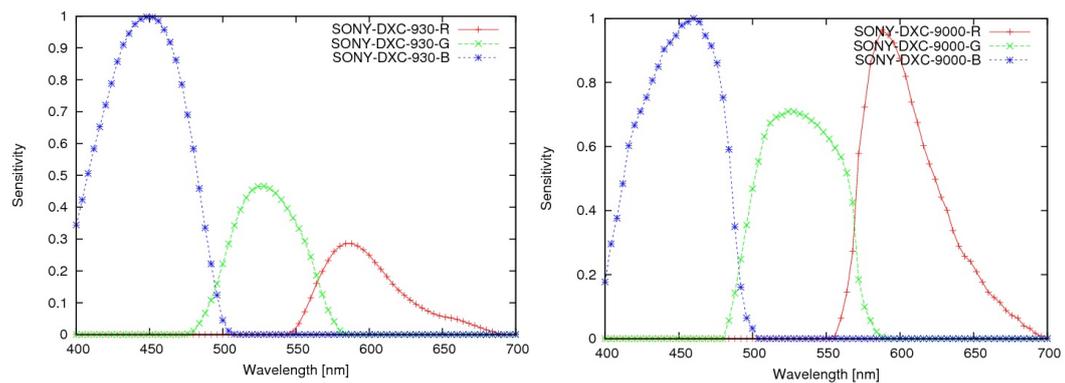
A Equação D.2 define que a cor medida, C_m , na posição x_I da imagem, relativa à posição x_O do objeto, depende apenas da intensidade do iluminante e da refletância da superfície no comprimento de onda λ_m . Esta consideração permite caracterizar a mudança de cor devido à mudança do iluminante como sendo uma simples mudança de escala em cada canal. Acontece que poucas câmeras possuem sensores que se aproximam de funções delta. A Figura

Figura D.3: Funções de sensibilidade espectral das câmeras NIKON D1X e NIKON D70.



Fonte: (Kawakami et al., 2013)

Figura D.4: Funções de sensibilidade espectral das câmeras NIKON D1X e NIKON D70.



Fonte: (Kawakami et al., 2013)

D.4 mostra as curva espectrais de sensibilidade das câmeras SONY DXC 930 e DXC 9000 que se comportam relativamente bem neste contexto. De fato, a câmera SONY DXC 930 é muito utilizada em trabalhos sobre constância de cor (Finlayson e Hordley, 2001; Finlayson et al., 2005; Funt e Jiang, 2003; Ratnasingam e McGinnity, 2012). No entanto, estas câmeras possuem alto custo e não são fáceis de serem encontradas no mercado.