



ALGORITMOS DE SUPER-RESOLUÇÃO DE IMAGENS COLORIDAS  
UTILIZANDO FILTRO DE WIENER

Diego Santos Wanderley

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientadores: Mariane Rembold Petraglia  
José Gabriel Rodríguez  
Carneiro Gomes

Rio de Janeiro  
Dezembro de 2013

ALGORITMOS DE SUPER-RESOLUÇÃO DE IMAGENS COLORIDAS  
UTILIZANDO FILTRO DE WIENER

Diego Santos Wanderley

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

---

Prof. Mariane Rembold Petraglia, Ph.D.

---

Prof. José Gabriel Rodríguez Carneiro Gomes, Ph.D.

---

Prof. Felipe Maia Galvão França, Ph.D.

---

Prof. Lisandro Lovisolo, D.Sc.

RIO DE JANEIRO, RJ – BRASIL  
DEZEMBRO DE 2013

Wanderley, Diego Santos

Algoritmos de Super-Resolução de Imagens Coloridas Utilizando Filtro de Wiener/Diego Santos Wanderley. – Rio de Janeiro: UFRJ/COPPE, 2013.

XIII, 85 p.: il.; 29, 7cm.

Orientadores: Mariane Rembold Petraglia

José Gabriel Rodríguez Carneiro Gomes

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2013.

Referências Bibliográficas: p. 72 – 77.

1. Super-Resolução. 2. Filtro de Wiener. 3. Processamento de Imagens. I. Petraglia, Mariane Rembold *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*À minha família.*

# Agradecimentos

Aos meus pais e minha irmã, pelo apoio incondicional que permitiu a realização desta conquista.

Aos meus professores e orientadores Mariane e José Gabriel, pelos ensinamentos e conselhos, além da paciência e da disponibilidade em me orientar durante os últimos anos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro através de bolsa de estudos.

Aos colegas e amigos que me auxiliaram durante a difícil jornada do mestrado. Em especial os amigos Daniel Cayres e Alberto Wagner, que contribuíram de maneira ativa com boas discussões e auxílio acadêmico.

À Universidade Federal do Rio de Janeiro (UFRJ), que foi a minha segunda casa nos últimos 10 anos. Por todas as lições acadêmicas e de vida que me proporcionou, além dos grandes amigos que aqui conheci.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## ALGORITMOS DE SUPER-RESOLUÇÃO DE IMAGENS COLORIDAS UTILIZANDO FILTRO DE WIENER

Diego Santos Wanderley

Dezembro/2013

Orientadores: Mariane Rembold Petraglia

José Gabriel Rodríguez Carneiro Gomes

Programa: Engenharia Elétrica

A super-resolução de imagens digitais visa superar limitações teóricas e práticas que restringem a resolução de qualquer dispositivo de captura de imagem. Este trabalho apresenta um estudo sobre algoritmos de super-resolução para imagens coloridas baseados em filtros de Wiener adaptativos, que apresentam baixa complexidade computacional quando comparados com outros métodos que buscam ampliação de resolução da imagem por super-resolução.

Primeiramente, a super-resolução por filtragem de Wiener monocromática é aplicada a cada canal de cor para realizar a super-resolução de imagens coloridas. Duas variações da proposta inicial são realizadas com base em propriedades da visão. A primeira variação tem por objetivo reduzir a complexidade computacional, enquanto que a segunda utiliza a Transformada Wavelet para diminuir efeitos de ruído presentes nas imagens de baixa resolução. Os métodos estudados nesta dissertação são comparados a uma técnica de super-resolução por filtragem adaptativa com penalidades para os erros de otimização associados a bordas e cor.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

## COLOR IMAGE SUPER-RESOLUTION BASED ON WIENER FILTERS

Diego Santos Wanderley

December/2013

Advisors: Mariane Rembold Petraglia

José Gabriel Rodríguez Carneiro Gomes

Department: Electrical Engineering

Digital image super-resolution aims at overcoming theoretical and practical limitations that restrict image device resolution. This work presents a study on color-image super-resolution algorithms based on adaptive Wiener filtering, which have low computational complexity in comparison to other image resolution enhancement methods.

First, the monochrome super-resolution using Wiener filtering is applied to each color channel to perform the color images super-resolution. Two variations of the initial proposal are made based on the eye properties. The first variation aims at reducing the computational complexity, while the second employs the Wavelet Transform to mitigate the effects of random noise in the low resolution images. The methods studied in this thesis are compared to a super-resolution technique by adaptive filtering with penalties for optimization errors associated with borders and color.

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Delimitação . . . . .	2
1.2 Objetivos . . . . .	3
1.3 Organização . . . . .	3
<b>2 Super-Resolução</b>	<b>5</b>
2.1 Técnicas de Super-Resolução . . . . .	7
<b>3 Super-Resolução com Filtro de Wiener</b>	<b>10</b>
3.1 Algoritmo de SR AWF . . . . .	11
3.2 Modelo Estatístico . . . . .	13
3.3 SR AWF com Wavelet . . . . .	21
3.3.1 Ajuste do Modelo de Autocorrelação por Sub-Banda . . . . .	23
<b>4 Super-Resolução de Imagens Coloridas</b>	<b>27</b>
4.1 SR AWF nos Canais de Cores . . . . .	28
4.1.1 RGB . . . . .	28
4.1.2 Compressão de Crominância . . . . .	30
4.1.3 Emprego da Transformada Wavelet na Luminância . . . . .	32
4.2 SR <i>Multiframe</i> Adaptativo . . . . .	34
4.2.1 Termo de Fidelidade . . . . .	35
4.2.2 Luminância . . . . .	36
4.2.3 Crominância . . . . .	37
4.2.4 Dependência entre Canais de Cores . . . . .	37
4.2.5 Função-Custo Total . . . . .	39
<b>5 Resultados</b>	<b>41</b>
5.1 Métricas de Avaliação . . . . .	41

5.1.1	PSNR . . . . .	42
5.1.2	S-CIELAB . . . . .	42
5.1.3	SSIM . . . . .	43
5.2	Parâmetros Utilizados . . . . .	43
5.3	Efeito do Ruído . . . . .	44
5.4	Complexidade Computacional . . . . .	61
5.5	Grade HR Esparsa . . . . .	63
<b>6</b>	<b>Conclusões</b>	<b>70</b>
	<b>Referências Bibliográficas</b>	<b>72</b>
<b>A</b>	<b>Sensores de Imagem</b>	<b>78</b>
A.1	CCD . . . . .	79
A.2	CMOS . . . . .	80
<b>B</b>	<b><i>Color Filter Array</i></b>	<b>82</b>
B.1	<i>Demosaicing</i> . . . . .	85

# Lista de Figuras

2.1	Condição básica para super-resolução. . . . .	6
2.2	Modelo de aquisição de imagem. . . . .	7
3.1	Pixels de imagens de baixa resolução em uma grade de alta resolução. . . . .	11
3.2	Diagrama de blocos do algoritmo de SR AWF. . . . .	11
3.3	Janela de observação sobre grade HR. . . . .	13
3.4	Autocorrelação de imagens como processo aleatório. . . . .	14
3.5	Autocorrelação de blocos de imagens como processo aleatório. . . . .	15
3.6	Modelo de função de autocorrelação . . . . .	15
3.7	Estimativas de funções bidimensionais de autocorrelação de imagens. . . . .	16
3.8	Super-resolução com filtro de Wiener. . . . .	19
3.9	SR AWF com diferentes modelos de autocorrelação. . . . .	20
3.10	Diagrama do algoritmo SR AWF com Wavelet. . . . .	22
3.11	Comparação entre SR AWF e SR AWF com Wavelet. . . . .	22
3.12	Coefficientes da Wavelet sem decimação. . . . .	24
3.13	Autocorrelação relativa às diferentes sub-bandas da Wavelet. . . . .	25
3.14	Comparação de SR AWF com Wavelet. . . . .	26
4.1	Diagrama do algoritmo de SR AWF para os canais RGB. . . . .	28
4.2	Super-resolução colorida com filtro de Wiener. . . . .	29
4.3	Diagrama do algoritmo de SR AWF com compressão de crominância. . . . .	31
4.4	Comparação da SR AWF com compressão de crominância $C_b$ e $C_r$ . . . . .	32
4.5	Diagrama do algoritmo de SR AWF com Wavelet na luminância. . . . .	33
4.6	Comparação da SR AWF com Wavelet na luminância. . . . .	34
4.7	Imagem HR gerada por algoritmo de SR <i>multiframe</i> adaptativo. . . . .	40
5.1	Imagens HR desejadas. . . . .	45
5.2	Imagem <i>Charlize</i> LR com diferentes níveis de ruído. . . . .	45
5.3	Imagem <i>Lighthouse</i> LR com diferentes níveis de ruído. . . . .	46
5.4	Super-resolução da imagem <i>Charlize</i> sem ruído aditivo. . . . .	48
5.5	Super-resolução da imagem <i>Charlize</i> com SNR de 40 dB. . . . .	49
5.6	Super-resolução da imagem <i>Charlize</i> com SNR de 30 dB. . . . .	50

5.7	Super-resolução da imagem <i>Charlize</i> com SNR de 25 dB. . . . .	51
5.8	Super-resolução da imagem <i>Lighthouse</i> sem ruído aditivo. . . . .	52
5.9	Super-resolução da imagem <i>Lighthouse</i> com SNR de 40 dB. . . . .	53
5.10	Super-resolução da imagem <i>Lighthouse</i> com SNR de 30 dB. . . . .	54
5.11	Super-resolução da imagem <i>Lighthouse</i> com SNR de 25 dB. . . . .	55
5.12	Comparação de complexidade computacional dos algoritmos. . . . .	62
5.13	Evolução da PSNR da super-resolução da imagem <i>Charlize</i> com variação no número de imagens LR. . . . .	64
5.14	Evolução da S-CIELAB da super-resolução da imagem <i>Charlize</i> com variação no número de imagens LR. . . . .	65
5.15	Evolução da SSIM da super-resolução da imagem <i>Charlize</i> com variação no número de imagens LR. . . . .	65
5.16	Super-resolução da imagem <i>Charlize</i> a partir de cinco imagens LR com SNR de 40 dB. . . . .	66
5.17	Evolução da PSNR da super-resolução da imagem <i>Lighthouse</i> com variação no número de imagens LR. . . . .	67
5.18	Evolução da S-CIELAB da super-resolução da imagem <i>Lighthouse</i> com variação no número de imagens LR. . . . .	67
5.19	Evolução da SSIM da super-resolução da imagem <i>Lighthouse</i> com variação no número de imagens LR. . . . .	68
5.20	Super-resolução da imagem <i>Lighthouse</i> a partir de cinco imagens LR e SNR de 40 dB. . . . .	69
A.1	Funcionamento do sensor CCD. . . . .	79
A.2	Funcionamento do sensor CMOS. . . . .	80
B.1	Funcionamento de uma câmera colorida com <i>beam split</i> . . . . .	83
B.2	Microlente, CFA e pixel . . . . .	83
B.3	Padrão de Bayer . . . . .	84
B.4	Exemplo de padrões de CFA. . . . .	84

# Lista de Tabelas

2.1	Dimensões das matrizes do modelo de aquisição de imagem. . . . .	7
3.1	Algoritmo para o cálculo da matriz $\tilde{\mathbf{R}}$ . . . . .	18
3.2	Algoritmo para o cálculo da matriz $\tilde{\mathbf{P}}$ . . . . .	18
5.1	Análise da super-resolução da imagem <i>Charlize</i> sem ruído aditivo. . .	47
5.2	Análise da super-resolução da imagem <i>Charlize</i> com SNR de 40 dB. .	47
5.3	Análise da super-resolução da imagem <i>Charlize</i> com SNR de 30 dB. .	47
5.4	Análise da super-resolução da imagem <i>Charlize</i> com SNR de 25 dB. .	47
5.5	Análise da super-resolução da imagem <i>Lighthouse</i> sem ruído aditivo. .	56
5.6	Análise da super-resolução da imagem <i>Lighthouse</i> com SNR de 40 dB.	56
5.7	Análise da super-resolução da imagem <i>Lighthouse</i> com SNR de 30 dB.	56
5.8	Análise da super-resolução da imagem <i>Lighthouse</i> com SNR de 25 dB.	56
5.9	Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com sem ruído. . . . .	57
5.10	Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB. . . . .	57
5.11	Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 30 dB. . . . .	58
5.12	Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB. . . . .	58
5.13	Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com sem ruído. . . . .	58
5.14	Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB. . . . .	58
5.15	Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 30 dB. . . . .	59
5.16	Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB. . . . .	59
5.17	Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas sem ruído. . . . .	60

5.18	Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB. . . . .	60
5.19	Análise do S-CIELAB da super-resolução do conjunto de 30 imagens com SNR = 30 dB. . . . .	60
5.20	Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB. . . . .	60
A.1	Comparação entre CCD e CMOS. . . . .	81

# Capítulo 1

## Introdução

O avanço tecnológico na área de sensores de imagem possibilitou a difusão de câmeras fotográficas e de vídeo na sociedade atual. Estas câmeras estão incluídas em diversos setores, sendo utilizadas em sistemas complexos ou cotidianos, como satélites, sistemas de segurança, celulares ou computadores portáteis. Apesar da popularização, imagens de alta qualidade podem resultar em sistemas de aquisição caros.

A qualidade de uma imagem (fotografia ou vídeo) está diretamente ligada às condições em que foi realizada a sua aquisição, e principalmente ao modelo adotado para realizar tal processo. Sua resolução é limitada pelo tamanho do sensor responsável pela aquisição, e pelo número de células sensíveis à luz presentes nestes. Uma imagem pode assumir diferentes níveis de qualidade se a resolução de captura for significativamente diferente. Em baixa resolução, a imagem pode ser severamente corrompida, pois a subamostragem do processo acarreta o efeito de *aliasing*.

Diversos fatores afetam a qualidade das imagens capturadas por câmeras digitais comerciais. Uma cena real sofre deformações causadas pela lente e pelo movimento relativo entre a câmera e a cena. As imperfeições de uma lente resultam em problemas de foco e borramento da imagem. Um modelo do processo de aquisição de imagens ainda leva em consideração a subamostragem das imagens capturadas pelo sensor e o ruído dos sensores de leitura dos pixels. A distância entre a cena e o sensor e o tamanho do sensor também são determinantes na qualidade da imagem.

Em determinadas situações queremos aumentar a resolução sem precisar expandir a área do sensor, nem elevar o preço do equipamento ou diminuir a distância até a cena. Imagens de satélites para controle florestal ou meteorológico estão sujeitas às condições atmosféricas, além da grande distância entre o objeto de interesse e o sistema de captura. Outra necessidade é realizar uma impressão de alta qualidade de uma imagem de um vídeo ou de uma câmera de baixa resolução. Este último caso pode ser útil para identificar suspeitos em câmeras de vigilância, que utilizam alta compressão de imagens para armazenamento. Em todos esses casos, pode haver limitações para as dimensões dos sensores [1].

## 1.1 Delimitação

Nas últimas duas décadas, técnicas de processamento de sinais foram utilizadas para obter aproximações de imagens em alta resolução (HR - *High Resolution*), baseando-se em sequências de imagens de baixa resolução (LR - *Low Resolution*) da mesma cena, com ligeiras diferenças. Tais técnicas de melhoramento de imagem são conhecidas na literatura especializada como Super-Resolução (SR), onde o termo super representa a superação das limitações características do sistema de aquisição de imagens. Por este motivo expressões como “reconstrução de imagem por SR” e “reconstrução com SR” (SRR - *Super-Resolution Reconstruction*) são comumente utilizadas.

Devido ao fato de utilizar a combinação de múltiplas imagens LR de uma mesma cena, a SRR utiliza técnicas de estimação de movimento para medir a distância entre objetos da cena e em seguida estabelecer uma relação entre cada um dos *frames* LR e a imagem HR que se deseja encontrar. A estimação de movimento é um procedimento comum em visão computacional e compressão de vídeo, e consiste em encontrar os vetores de deslocamento dos pixels entre duas imagens, através de comparações entre os *frames*.

A super-resolução é, em geral, descrita para imagens em escala de cinza; poucos trabalhos abordam o problema da super-resolução em imagens coloridas. Uma abordagem utilizada em trabalhos como [2, 3] é aplicar algoritmos de super-resolução monocromática em cada um dos canais de cor, de maneira independente. Em [4] a imagem é decomposta em componentes de luminância e crominância, e a SRR é aplicada apenas no primeiro canal. Aparentemente dois problemas podem ser observados. Primeiramente esses métodos não exploram a correlação cruzada entre os canais de cor, por isso são considerados subótimos. O segundo problema deve-se ao fato desses métodos requererem imagens LR coloridas *demosaiiced* (definidas abaixo), não permitindo a utilização de imagens *raw*<sup>1</sup> provenientes da captura dos sensores de imagem mais populares.

Para reduzir custos de fabricação, câmeras fotográficas e filmadoras utilizam uma matriz com pixels sensíveis a diferentes cores, conhecida em inglês como *Color Filter Array* (CFA). Cada um dos pixels do sensor é sensível a uma cor primária, vermelho (R - *Red*), verde (G - *Green*) ou azul (B - *Blue*), formando um mosaico de filtros de cor. Para que a imagem possa ser apresentada e visualizada naturalmente, necessitamos que cada um dos pixels contenha a informação das três cores primárias. Estes novos valores são resultados de técnicas de interpolação e processamento de sinais, utilizando as informações capturadas pelo sensor de imagem. O termo utilizado

---

<sup>1</sup>Em inglês *raw* significa cru. É a imagem como capturada pelo sensor, contendo informação de apenas uma cor primária por pixel.

na literatura para esta combinação é *demosaiçing*, palavra inglesa, e uma tradução aceitável seria “desfazer o mosaico”. Alguns autores utilizam a grafia *demosaiçking*. Explicações mais detalhada sobre sensores de imagem e a reconstrução do CFA por *demosaiçing* são apresentadas no Apêndice A e no Apêndice B, respectivamente.

Em determinadas situações é necessário obter uma resolução de imagem maior do que a área do sensor de imagem. A super-resolução surge como alternativa para este problema pois, conhecendo a deficiência do sistema, podemos fotografar a cena diversas vezes ou utilizar a sequência de quadros de um vídeo para gerar as imagens LR necessárias para o algoritmo.

O processo de aquisição é responsável pelas diferenças necessárias entre as imagens de baixa resolução através de pequenas vibrações da câmera, dos movimentos da cena ou de ruídos produzidos entre a aquisição e a gravação da imagem. Desta forma, proporciona-se uma situação favorável para resolver o problema inicial por SRR. Assim, a partir das imagens de baixa resolução aplica-se algoritmos de processamento de sinais para combiná-las e extrair parâmetros e características, com a finalidade de calcular uma aproximação para a imagem de alta resolução.

## 1.2 Objetivos

O objetivo deste trabalho é estabelecer um método robusto de super-resolução baseado em filtragem de Wiener para imagens coloridas *demosaiçed*. São avaliadas a aplicação direta da super-resolução monocromática com filtros de Wiener nos canais de cores R, G e B, além de duas variações: a primeira visa reduzir o custo computacional utilizando compressão das componentes de crominância; e a segunda tem por objetivo atenuar o efeito do ruído aditivo presente nas imagens LR, aplicando o filtro de Wiener nas componentes da Transformada Wavelet das imagens.

Também é feita uma comparação entre os métodos via filtro de Wiener e um método de super-resolução por algoritmo adaptativo baseado em gradiente descendente. Este método utiliza os diferentes quadros LR como parâmetros de comparação. Desta forma, a imagem HR estimada sofre um processo semelhante ao de aquisição, que visa recriar as imagens LR originais para medir o erro da reconstrução. O método de SR adaptativo também possui parâmetros de otimização para melhorar a qualidade das bordas e das relações entre os canais de cores.

## 1.3 Organização

Inicialmente, este trabalho apresenta a definição, a estrutura e a modelagem matemática do problema de super-resolução de imagens monocromáticas no Capítulo 2. Também é apresentada uma visão geral das principais técnicas de super-resolução

e os diferentes tipos de abordagem, no domínio da frequência e espacial, realizadas por processamento de sinais ou por aprendizagem de máquina.

No Capítulo 3 são apresentadas técnicas de super-resolução através do filtro de Wiener para imagens em escala de cinza. Tais técnicas utilizam um algoritmo rápido e adaptativo para filtragem espacial ao longo da imagem. Um estudo do modelo estatístico utilizado pelo algoritmo é apresentado, de forma a estabelecer os parâmetros para a implementação de um algoritmo rápido. Por fim, a filtragem de Wiener é aplicada aos coeficientes de cada sub-banda da Transformada Wavelet, com a finalidade de reduzir o impacto do ruído presente nas imagens LR.

As técnicas de super-resolução colorida, tema central do trabalho, são apresentadas no Capítulo 4. Na primeira seção, o algoritmo por filtragem de Wiener monocromática é estendido para utilização nos canais de cor R, G e B. Após uma explicação sobre transformação do espaço de cor, uma adaptação do mesmo algoritmo é proposta para o espaço YCbCr, de forma a reduzir o custo computacional na estimação das componentes de crominância. Ainda na Seção 4.1, a Transformada Wavelet é aplicada à componente de luminância da imagem, com o objetivo de reduzir a propagação de erro devido a altos níveis de ruído nas imagens LR.

Também no Capítulo 4, é apresentada a técnica de SR *multiframe* adaptativo. O método corresponde a um algoritmo iterativo baseado na minimização de uma função custo que possui termos de penalidade que relacionam a fidelidade da posição do pixel HR com as variações do sinal na luminância, na crominância e na dependência entre canais de cores.

O Capítulo 5 apresenta resultados para os métodos de super-resolução descritos no Capítulo 4. Inicialmente são apresentadas as métricas de avaliação utilizadas para mensurar a qualidade das imagens de alta resolução estimadas. Os resultados da super-resolução dos diferentes métodos são avaliados quanto a robustez ao ruído e quanto ao tempo de processamento. Também é apresentada uma análise sobre o comportamento dos métodos para diferentes conjuntos de imagens, com diferença no número de imagens LR utilizadas para realizar a super-resolução.

As conclusões, considerações e propostas para projetos futuros são apresentadas no Capítulo 6.

Uma breve descrição sobre sensores de imagem encontra-se no Apêndice A, além de uma comparação entre os dois principais modelos de sensores de imagens digitais: o CCD e o CMOS.

No Apêndice B é abordado o conceito de *Color Filter Array*, seguido pelo seu tratamento, conhecido como *demosaicing*.

# Capítulo 2

## Super-Resolução

Super-resolução consiste em combinar uma sequência de imagens em baixa resolução, possivelmente borradas e com ruído, para produzir uma ou várias imagens de alta resolução. É importante que as imagens LR possuam diferenças de subpixel de deslocamento entre elas. O efeito de *aliasing* nas imagens subamostradas possibilita a recuperação da imagem HR através da fusão, como constatado nos primeiros trabalhos sobre super-resolução [5]. Utilizar técnicas de processamento de sinais é uma alternativa de baixo custo para resolver aproximações impostas por limitações físicas dos sensores de imagem.

No problema de super-resolução, queremos aumentar a resolução da imagem sem perder qualidade, tentando minimizar o erro de reconstrução da cena original. Para aumentar a resolução espacial de um sensor de imagem, precisamos diminuir o tamanho do pixel, de forma que caibam mais pixels na mesma área de sensor. À medida que diminuímos o pixel, diminuímos a sensibilidade à luz, produzindo ruído *shot*, conforme este se aproxima do limite físico da tecnologia [1]. Outra possibilidade seria aumentar o tamanho total do sensor, o que acarretaria o aumento da capacitância [6] e a implementação de um sistema óptico maior e mais preciso, elevando os custos do projeto.

O conjunto de imagens de uma mesma cena deve conter variações de subpixel entre as imagens. Tais diferenças podem ser decorrentes do movimento da cena ou da câmera. Se este movimento é conhecido, ou pode ser estimado com precisão de subpixel, é possível combinar o conjunto de imagens LR para realizar uma reconstrução com super-resolução. Na Fig. 2.1 é exemplificada a condição básica para realizar a super-resolução de uma cena. As imagens LR (à esquerda na Fig. 2.1) são capturadas através de uma câmera em instantes diferentes ou simultaneamente com várias câmeras; em seguida é estimado o movimento dos pixels entre uma imagem de referência e as demais, permitindo a combinação das imagens LR para a SRR.

O processo de aquisição e gravação de uma imagem digital induz a uma subamostragem da imagem original, limitada pela resolução do sensor, o que costuma

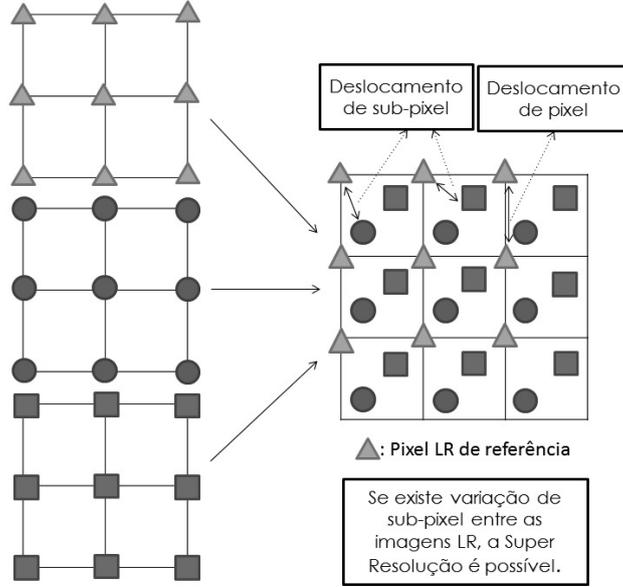


Figura 2.1: Condição básica para super-resolução [1].

acarretar o efeito de *aliasing*, que torna indistinguíveis sinais diferentes. As lentes podem introduzir distorções ópticas, como problemas de foco. Borramentos por movimentos acontecem devido à velocidade limitada do obturador. Após a captura ocorre ruído no sensor de imagem ou durante a leitura dos pixels. Considerando todos os possíveis problemas, podemos modelar uma imagem capturada como [7]:

$$\mathbf{y}(k) = \mathbf{D}(k)\mathbf{H}(k)\mathbf{F}(k)\mathbf{x} + \mathbf{v}(k), \quad 1 \leq k \leq N, \quad (2.1)$$

onde  $\mathbf{y}(k)$  é o vetor correspondente à  $k$ -ésima imagem LR,  $\mathbf{D}(k)$  é uma matriz responsável pela subamostragem (*downsampling*) da imagem,  $\mathbf{H}(k)$  é uma matriz que introduz o borrimento (*blur*), juntamente com a função de espalhamento pontual (PSF – *Point Spread Function*) da câmera.  $\mathbf{F}(k)$  é a matriz responsável por movimentos translacionais e de rotação (*warping*), correspondente ao deslocamento entre as imagens. O vetor  $\mathbf{x}$  contém a informação dos pixels da imagem ideal e  $\mathbf{v}(k)$  é o ruído introduzido pelo sistema de leitura do sensor. O valor  $k$  corresponde ao instante em que a imagem original foi amostrada pela câmera, e  $N$  é o número total de amostras. A Fig. 2.2 apresenta um diagrama do modelo de aquisição de  $\mathbf{y}(k)$ .

A Tabela 2.1 apresenta as dimensões das matrizes e vetores citados acima, considerando a resolução do sensor como  $M_1 \times M_2$  pixels, e  $r$  o fator de ampliação entre a imagem de baixa e a de alta resolução. As matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  correspondentes às imagens e a matriz de ruído  $\mathbf{V}$  são representadas em notação lexicográfica, ou seja, em forma de vetores denotados, respectivamente,  $\mathbf{x}$ ,  $\mathbf{y}$  e  $\mathbf{v}$ . A notação lexicográfica é utilizada para modelar matematicamente as operações do modelo de captura, em uma implementação é comum utilizar uma sequência de operadores nas imagens.

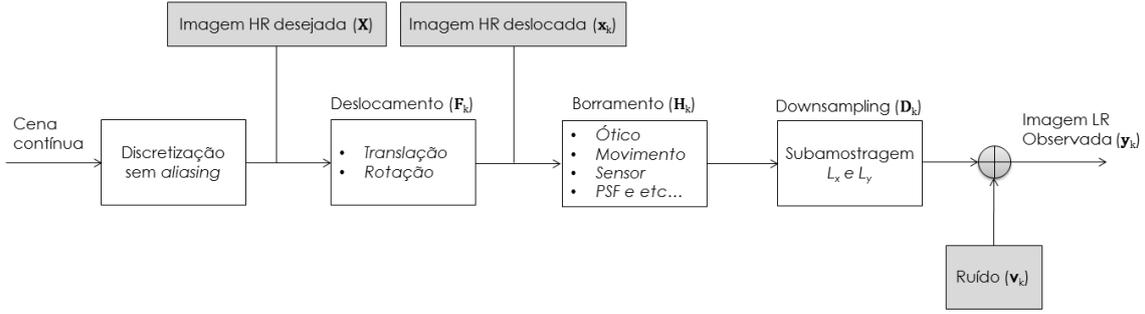


Figura 2.2: Modelo de aquisição de imagem.

Tabela 2.1: Dimensões das matrizes do modelo de aquisição de imagem.

Matriz	Dimensão	Descrição
$\mathbf{Y}$	$[M_1 \times M_2]$	Imagem LR
$\mathbf{y}$	$[M_1 M_2 \times 1]$	Representação lexicográfica de $\mathbf{Y}$
$\mathbf{D}$	$[M_1 M_2 \times r^2 M_1 M_2]$	Subamostragem
$\mathbf{H}$	$[r^2 M_1 M_2 \times r^2 M_1 M_2]$	Borramento
$\mathbf{F}$	$[r^2 M_1 M_2 \times r^2 M_1 M_2]$	Movimento entre quadros
$\mathbf{X}$	$[r M_1 \times r M_2]$	Imagem HR
$\mathbf{x}$	$[r^2 M_1 M_2 \times 1]$	Representação lexicográfica de $\mathbf{X}$
$\mathbf{V}$	$[M_1 \times M_2]$	Ruído
$\mathbf{v}$	$[M_1 M_2 \times 1]$	Representação lexicográfica de $\mathbf{V}$

## 2.1 Técnicas de Super-Resolução

Desde o pioneiro trabalho [5] publicado em 1984, várias técnicas de super-resolução foram propostas, com abordagens tanto no domínio da frequência como no espacial, sob diferentes perspectivas, de processamento de sinais a aprendizagem de máquina. Trabalhos recentes ainda seguem a teoria pioneira, explorando o deslocamento e o *aliasing* no domínio da Transformada de Fourier. Entretanto, a abordagem no domínio da frequência é restrita ao modelo primário de aquisição da imagem apresentado na Fig. 2.2, enquanto os problemas reais são mais complexos. As pesquisas mais recentes tentam resolver o problema no domínio espacial, devido à facilidade de modelar as variadas formas de degradação da imagem [7].

Super-resolução com múltiplos quadros provou ser um método de processamento de imagens bem-sucedido em reduzir o *aliasing* e melhorar a resolução de imagens em sistemas de detecção limitados [1]. Havendo movimento relativo entre os quadros da cena e a câmera, a diversidade de amostras pode ser explorada para reduzir os efeitos da subamostragem. Este processamento pode ser visto como uma troca da resolução temporal pela resolução espacial, o que permite reduzir ou eliminar os efeitos de *aliasing*. Se o *aliasing* puder ser reduzido a um nível mínimo, técnicas de restauração

linear podem ser aplicadas com sucesso para resolver os problemas de borramento da PSF. Se o *aliasing* presente for pequeno ou inexistente no sistema, uma técnica de restauração de quadro único pode ser uma escolha mais adequada. Técnicas de restauração com uma única imagem não são capazes de recuperar componentes de alta frequência perdidas ou degradadas no processo de subamostragem [1].

Considerando a relação entre a imagem HR e os *frames* LR na Eq. (2.1) um sistema linear esparso, muitos estimadores podem ser aplicados à SRR. Máxima Verossimilhança (ML), Máximo a *Posteriori* (MAP) [8] e Projeção em Conjunto Convexo (POCS) [9] podem ser empregados em abordagens de super-resolução no domínio espacial. Técnicas de interpolação espacial não uniforme são empregadas em sequência com métodos de *deblurring* [10–13], com o objetivo de retirar o borramento resultante. As técnicas de restauração por interpolação são em geral intuitivas e computacionalmente eficientes [14], baseada em modelos de observação simples.

Em contraponto aos métodos de restauração por interpolação, abordagens estatísticas tratam a imagem HR e os movimentos entre os *frames* de baixa resolução como variáveis estocásticas, visando criar uma reconstrução ótima. Considerando  $M(\mathbf{f}, \mathbf{h})$  a matriz de degradação definida pelo vetor de movimentos  $\mathbf{f}$  e de borramento  $\mathbf{h}$ , a SRR pode ser modelada em uma estrutura Bayesiana [7]

$$\begin{aligned}
\mathbf{X} &= \arg \max_{\mathbf{X}} Pr(\mathbf{X} | \underline{\mathbf{Y}}) \\
&= \arg \max_{\mathbf{X}} \int_{\mathbf{f}, \mathbf{h}} Pr(\mathbf{X}, M(\mathbf{f}, \mathbf{h}) | \underline{\mathbf{Y}}) d\mathbf{f} \\
&= \arg \max_{\mathbf{X}} \int_{\mathbf{f}, \mathbf{h}} \frac{Pr(\underline{\mathbf{Y}} | \mathbf{X}, M(\mathbf{f}, \mathbf{h})) Pr(\mathbf{X}, M(\mathbf{f}, \mathbf{h}))}{Pr(\underline{\mathbf{Y}})} d\mathbf{f} \\
&= \arg \max_{\mathbf{X}} \int_{\mathbf{f}, \mathbf{h}} Pr(\underline{\mathbf{Y}} | \mathbf{X}, M(\mathbf{f}, \mathbf{h})) Pr(\mathbf{X}) Pr(M(\mathbf{f}, \mathbf{h})) d\mathbf{f},
\end{aligned} \tag{2.2}$$

sendo  $\mathbf{X}$  e  $M(\mathbf{f}, \mathbf{h})$  estatisticamente independentes [8].  $Pr(\underline{\mathbf{Y}} | \mathbf{X}, M(\mathbf{f}, \mathbf{h}))$  é a probabilidade a *posteriori*,  $Pr(\mathbf{X})$  é a probabilidade a priori da imagem desejada em alta resolução e  $Pr(M(\mathbf{f}, \mathbf{h}))$  é a probabilidade a *priori* da estimação de movimento. O vetor  $\mathbf{v}$  na Eq. (2.1) é assumido como um ruído branco Gaussiano com média zero, de modo que:

$$Pr(\underline{\mathbf{Y}} | \mathbf{X}, M(\mathbf{f}, \mathbf{h})) \propto \exp \left\{ -\frac{1}{2\sigma^2} \|\underline{\mathbf{Y}} - M(\mathbf{f}, \mathbf{h})\mathbf{X}\|^2 \right\}. \tag{2.3}$$

A probabilidade  $Pr(\underline{\mathbf{Y}})$  é usualmente definida por uma distribuição de Gibbs na

forma exponencial, dada por:

$$Pr(\mathbf{X}) = \frac{1}{Z} e^{-\alpha A(\mathbf{X})}, \quad (2.4)$$

onde  $A(\mathbf{X})$  é uma função potência não negativa,  $\alpha$  é um ponderador e  $Z$  é apenas um fator de normalização [7].

A forma Bayesiana na Eq. (2.2) é difícil de avaliar devido à integral sobre as estimativas de movimento; porém, se  $M(\mathbf{f}, \mathbf{h})$  for estimado antecipadamente, a seguinte simplificação pode ser feita:

$$\begin{aligned} \mathbf{X} &= \arg \max_{\mathbf{X}} Pr(\underline{\mathbf{Y}} | \mathbf{X}) \\ &= \arg \min_{\mathbf{X}} \{ \|\underline{\mathbf{Y}} - M\mathbf{X}\|^2 + \lambda A(\mathbf{X}) \}, \end{aligned} \quad (2.5)$$

onde  $M$  é a estimativa conhecida de  $M(\mathbf{f}, \mathbf{h})$  e  $\lambda$  absorve a variância do ruído. A Eq. (2.5) é o modelo para super-resolução do método de MAP. As abordagens estatísticas costumam variar na forma como tratam a degradação  $M(\mathbf{f}, \mathbf{h})$ , a probabilidade a *priori*  $Pr(\mathbf{X})$  e a inferência estatística em torno da Eq. (2.2) [7].

Métodos de filtragem adaptativa no tempo podem ser adaptados para o caso espacial da super-resolução [15], com uma mudança de notação no modelo de aquisição. A estrutura do filtro pode ser alterada para melhorar o desempenho, de acordo com um critério desejado, visando a minimização do erro médio quadrático (LMS) ou através do método recursivo dos mínimos quadrados (RLS). O algoritmo *Steepest Descent* (SD) é aplicado para estimar a imagem HR em cada uma das iterações, com uma complexidade computacional baixa, sem necessidade de inversão de matrizes [1]. Tal abordagem se mostrou capaz de tratar qualquer resolução de saída, borramento ou erros de estimação de movimento [14, 16].

Muitos dos métodos de super-resolução utilizam etapas independentes de interpolação e restauração para estimar a imagem HR [17]. Uma desvantagem disso é que uma etapa de restauração independente pode agravar defeitos resultantes da interpolação não uniforme. Uma abordagem utilizando filtros de Wiener adaptativos (AWF) combina a interpolação não uniforme das imagens LR com a restauração da imagem HR em um único passo [17]. Isso proporciona robustez na distribuição espacial dos pixels de baixa resolução, com baixa carga computacional. Os próximos capítulos tratarão com detalhes desta técnica, apresentando variantes e ampliações do modelo original e seu emprego em imagens coloridas.

## Capítulo 3

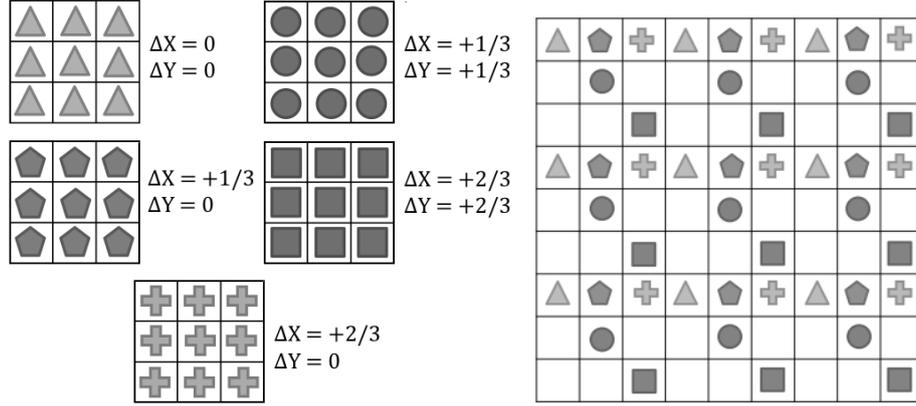
# Super-Resolução com Filtro de Wiener

Métodos de super-resolução com filtro de Wiener são algoritmos de interpolação não uniforme que utilizam somas ponderadas dos pixels de um conjunto de imagens de baixa resolução deslocadas entre si para estimar a imagem HR. Os pesos das somas são determinados com base na teoria do filtro de resposta ao impulso finita (FIR) de Wiener, de forma a adaptar-se à distribuição espacial das amostras LR em uma grade HR populada não uniformemente. Uma das características destes métodos é que a restauração da imagem HR é realizada junto com a interpolação não uniforme dos pixels das imagens LR. Um estudo sobre os métodos de super-resolução utilizando filtro de Wiener adaptativo é apresentado em [18].

No caso particular do algoritmo rápido de super-resolução utilizando filtro de Wiener adaptativo (SR AWF) [17], as imagens de baixa resolução passam por um *upsampling* por preenchimento com zeros e seus pixels são alinhados à mesma grade HR, levando em consideração os deslocamentos relativos a uma imagem de referência. A Fig. 3.1 ilustra os pixels das imagens LR dispostos na grade HR.

O algoritmo SR AWF, em seguida, produz as estimativas dos pixels da imagem HR utilizando somas ponderadas de pixels LR vizinhos. Os pesos atribuídos a cada pixel LR vizinho são estipulados com base na distância espacial relativa entre o vizinho correspondente e o pixel a ser estimado, de forma a minimizar o erro quadrático médio entre o pixel HR desejado e o pixel LR observado na grade HR.

Um modelo estatístico paramétrico para as correlações é pressuposto para definir os pesos do filtro. Os pesos podem ser redefinidos em iterações espaciais e temporais, de acordo com a distribuição dos pixels LR, tornando o filtro robusto mesmo para uma distribuição desfavorável ao modelo. O algoritmo SR AWF é computacionalmente mais simples na filtragem do que a aproximação por quantização vetorial proposta no método SR PWS (*Partition-based Weighted Sum*) [19], além de não necessitar de treino com imagens para o cálculo dos pesos dos filtros.



(a) Grades LR com deslocamentos relativos (b) Grade HR com pixels de LR

Figura 3.1: Pixels de imagens LR em uma grade HR.

### 3.1 Algoritmo de SR AWF

Uma visão geral do algoritmo de SR AWF [17] é ilustrada no diagrama de blocos da Fig. 3.2. As variáveis  $r_{ff}$  e  $r_{df}$  são respectivamente a autocorrelação da imagem observada livre de ruído e a correlação cruzada entre a imagem desejada e a imagem observada livre de ruído, que serão definidas na próxima seção. A variância  $\sigma_n^2$  é relativa ao ruído aditivo das imagens LR. As demais variáveis do diagrama serão definidas a seguir.

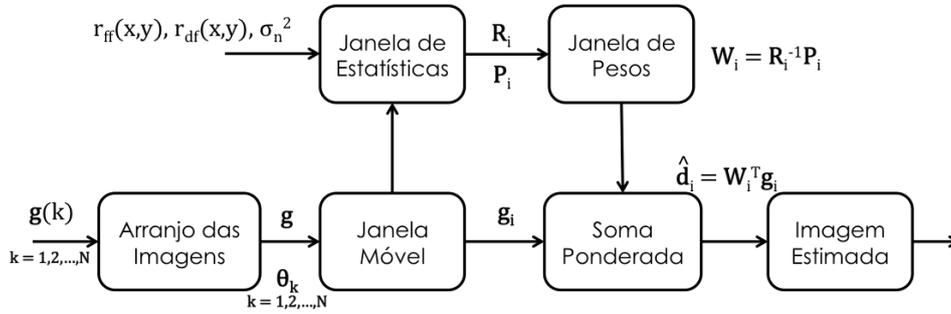


Figura 3.2: Diagrama de blocos do algoritmo de SR AWF.

O primeiro passo é dispor os pixels das imagens de baixa resolução em uma grade comum de alta resolução, levando em consideração o deslocamento relativo entre as imagens LR. Em processamento de vídeo, a imagem média da sequência pode ser usada como referência para melhor estabilidade do movimento sem agregar complexidade computacional. Os parâmetros de deslocamento e movimento entre quadros são armazenados em  $\theta_k$ , onde  $k$  varia de acordo com o número de imagens. Desta forma é possível estabelecer a coordenada de cada pixel LR do vetor de observação  $\mathbf{g}$  em uma grade comum.

A janela de observação, com dimensões  $W_x$  na horizontal e  $W_y$  na vertical, per-

corre a grade HR. Os pixels que se encontram na janela de observação compõem o vetor de observação  $\mathbf{g}_i = [g_{i,1}, g_{i,2}, \dots, g_{i,K}]^T$ , onde  $i$  é o índice referente à posição da janela de observação na grade HR, e  $K$  é o número de pixels LR na  $i$ -ésima janela de observação. Os valores de  $W_x$  e  $W_y$  são inteiros múltiplos dos fatores de interpolação  $L_x$  e  $L_y$ , respectivamente. Neste caso, o número de pixels LR em cada vetor de observação é uma constante definida como:

$$K = \frac{NW_xW_y}{L_xL_y}. \quad (3.1)$$

Para cada índice, é calculada a estimativa para os pixels HR de uma subjanela da janela de observação referente. As dimensões  $D_x$  e  $D_y$  da subjanela podem variar de um pixel às dimensões da janela principal  $W_x$  e  $W_y$ , respectivamente. A Fig. 3.3 ilustra uma situação com  $D_x = D_y = L_x = L_y = 3$  e  $W_x = W_y = 9$ . Os pixels de alta resolução estimados na subjanela são obtidos através da soma ponderada dos pixels de baixa resolução na janela de observação. Desta forma temos o seguinte vetor estimado:

$$\hat{\mathbf{d}}_i = \mathbf{W}_i^T \mathbf{g}_i, \quad (3.2)$$

onde  $\hat{\mathbf{d}}_i = [\hat{d}_{i,1}, \hat{d}_{i,2}, \dots, \hat{d}_{i,D_xD_y}]^T$ , e  $\mathbf{W}_i$  é uma matriz de pesos que possui dimensão  $K \times D_xD_y$ . Cada coluna da matriz  $\mathbf{W}_i$  contém os pesos correspondentes para estimar cada um dos pixels de alta resolução da janela de estimação  $\mathbf{d}_i$ . A mobilidade da janela de observação sobre a grade HR em passos de  $D_x$  e  $D_y$  nas direções horizontal e vertical, respectivamente, permite ao filtro se adaptar às mudanças estatísticas de intensidade para a estimativa de cada um dos pixels HR. As estimativas são combinadas para estimar uma imagem ótima.

Os pesos que minimizam o erro médio quadrático entre a imagem estimada e as imagens de baixa resolução são dados por [17]:

$$\mathbf{W}_i = \mathbf{R}_i^{-1} \mathbf{P}_i, \quad (3.3)$$

onde  $\mathbf{P}_i = E[\mathbf{g}_i \mathbf{d}_i^T]$  é a matriz de correlação cruzada entre o vetor desejado  $\mathbf{d}_i$  e o vetor de observação  $\mathbf{g}_i$  e  $\mathbf{R}_i = E[\mathbf{g}_i \mathbf{g}_i^T]$  é a matriz de autocorrelação do vetor de observação. Como as colunas de  $\mathbf{W}_i$  correspondem aos pesos para estimação de um pixel HR, a soma das linhas de cada coluna deve ser sempre um.

Em [19], uma matriz similar é determinada empiricamente através de um treinamento *offline* baseado em imagens pré-estabelecidas. Um modelo paramétrico contínuo para as funções de autocorrelação e correlação cruzada da imagem desejada, amostrando-as conforme as coordenadas dos pixels LR na grade HR é utilizado em [17]. Isso elimina a necessidade de determinar as matrizes  $\mathbf{R}_i$  e  $\mathbf{P}_i$  com base em imagens pré-estabelecidas.

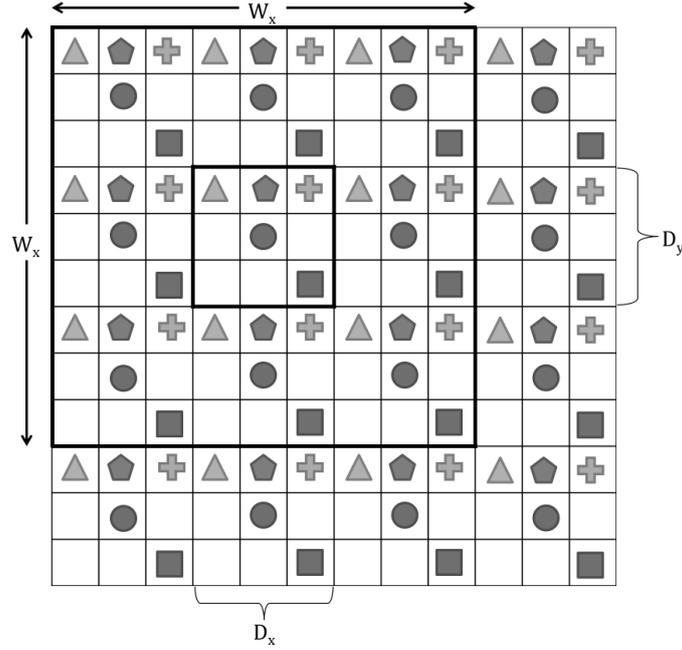


Figura 3.3: Janela de observação sobre grade HR.

## 3.2 Modelo Estatístico

O sinal observado na  $i$ -ésima janela de observação  $\mathbf{g}_i$  pode ser descrito como:

$$\mathbf{g}_i = \mathbf{f}_i + \mathbf{n}_i, \quad (3.4)$$

onde  $\mathbf{f}_i$  é a janela correspondente ao vetor  $\mathbf{g}_i$  em uma imagem livre de ruído e  $\mathbf{n}_i$  é o ruído aleatório associado. Considera-se o ruído branco com média zero e variância  $\sigma_n^2$  e descorrelacionado com  $\mathbf{f}_i$ . Podemos então reescrever as matrizes de autocorrelação  $\mathbf{R}_i$  e de correlação cruzada  $\mathbf{P}_i$  como:

$$\mathbf{R}_i = E[\mathbf{g}_i \mathbf{g}_i^T] = E[\mathbf{f}_i \mathbf{f}_i^T] + \sigma_n^2 \quad (3.5)$$

$$\mathbf{P}_i = E[\mathbf{g}_i \mathbf{d}_i^T] = E[\mathbf{f}_i \mathbf{d}_i^T]. \quad (3.6)$$

Esta abordagem reduz o problema a como determinar  $E[\mathbf{f}_i \mathbf{d}_i^T]$  e  $E[\mathbf{f}_i \mathbf{f}_i^T]$ .

A relação entre a função que descreve a imagem desejada  $d(x, y)$  e a função da imagem livre de ruído  $f(x, y)$  é definida através de

$$f(x, y) = d(x, y) * h(x, y), \quad (3.7)$$

onde  $h(x, y)$  é a PSF do sistema de captura.

Considerando a função de autocorrelação  $r_{dd}(x, y)$  da imagem  $d(x, y)$  como sendo de estacionariedade no sentido amplo (WSS), a correlação cruzada  $r_{df}(x, y)$  entre

$d(x, y)$  e  $f(x, y)$  pode ser reescrita em termos de  $r_{dd}(x, y)$  [20] como:

$$r_{df}(x, y) = r_{dd}(x, y) * h(x, y). \quad (3.8)$$

E a autocorrelação de  $f(x, y)$  é dada por:

$$r_{ff}(x, y) = r_{dd}(x, y) * h(x, y) * h(-x, -h). \quad (3.9)$$

O algoritmo SR AWF permite que um modelo de autocorrelação possa ser utilizado para a imagem desejada. Este modelo pode ser determinado a partir de uma análise empírica de determinadas imagens, do treinamento de um conjunto de imagens, ou através de um modelo estatístico paramétrico.

Em processamento de imagens, podemos tratar uma imagem como um processo aleatório através de dois pontos de vista. No primeiro as linhas correspondem à variável aleatória e as colunas às amostras. No segundo as colunas representam a variável aleatória e as linhas as amostras. Assim, podemos calcular a autocorrelação relativa às linhas e às colunas de uma imagem ou várias imagens. A Fig. 3.4(a) e a Fig. 3.4(b) ilustram respectivamente as autocorrelações das linhas e das colunas de um conjunto de imagens naturais<sup>1</sup> monocromáticas. Ao contrário da correlação do processo WSS, as correlações das imagens não apresentam padrão Toeplitz, como são conhecidas matrizes com diagonais constantes.

Porém, se dividirmos a imagem em blocos de pequenas dimensões, encontraremos resultado semelhante aos processos WSS. A Fig. 3.5(a) e a Fig. 3.5(b) correspondem às autocorrelações das linhas e das colunas de blocos de dimensões  $9 \times 9$  pixels de uma imagem, respectivamente.

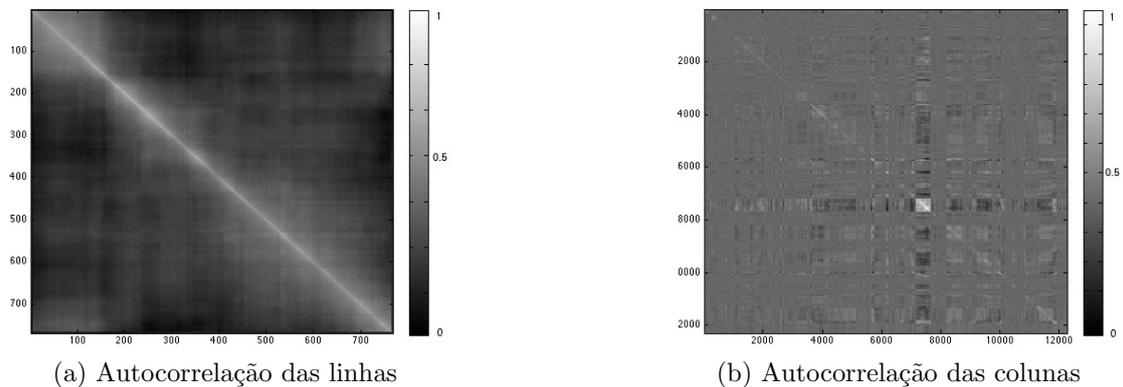
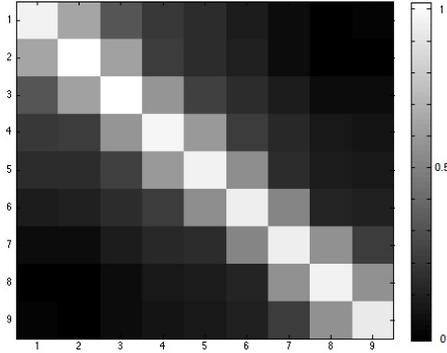


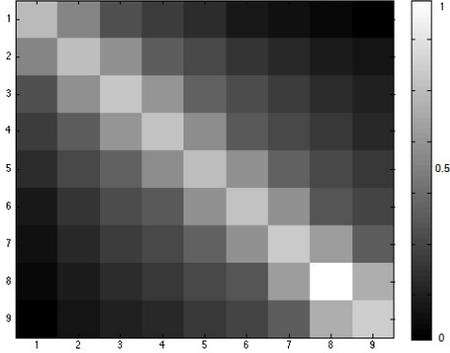
Figura 3.4: Autocorrelação de imagens como processo aleatório.

---

<sup>1</sup>Uma imagem natural apresenta, no domínio da frequência, concentração de energia em frequências baixas. Além disso, normalmente se trata de uma imagem que não sofreu filtragem ou correções artificiais.



(a) Autocorrelação linhas em blocos  $9 \times 9$



(b) Autocorrelação colunas em blocos  $9 \times 9$

Figura 3.5: Autocorrelação de blocos de imagens como processo aleatório.

A análise sobre a estacionariedade no sentido amplo em pequenas regiões de uma imagem nos permite estabelecer um modelo de autocorrelação WSS para as janelas desejadas de uma imagem na grade HR. O modelo paramétrico WSS de autocorrelação com simetria circular [21] utilizado no algoritmo rápido de SR AWF [17] é da forma:

$$r_{dd}(x, y) = \sigma_d^2 \rho \sqrt{(x^2 + y^2)}, \quad (3.10)$$

onde  $\sigma_d^2$  é a variância da imagem desejada e  $\rho$  é um parâmetro de ajuste, que controla o decaimento da autocorrelação com o aumento da distância para o ponto central. A Fig. 3.6 apresenta uma função de autocorrelação com  $\sigma_d^2 = 1$  e  $\rho = 0,75$ , valores sugeridos em [17].

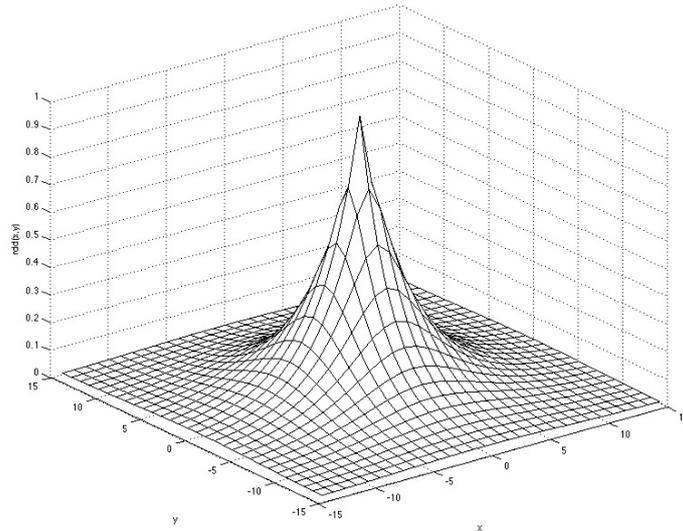


Figura 3.6: Modelo de função de autocorrelação  $r_{dd}(x, y)$  com  $\sigma_d^2 = 1$  e  $\rho = 0,75$ .

Calculando a média das diagonais das matrizes de autocorrelação de linhas e de colunas, podemos aproximar funções que descrevem o comportamento de auto-

correlações das imagens. A Fig. 3.7(b) apresenta a função bidimensional calculada através das médias das diagonais das matrizes da Fig. 3.5(a) e da Fig. 3.5(b). A regressão das médias das diagonais permite calcular valores para  $\rho$  e  $\sigma_d^2$  da Eq. (3.10) e obter uma aproximação para a função de autocorrelação. A Fig. 3.7(b) apresenta a estimativa da função  $r_{dd}(x, y)$  das matrizes da Fig. 3.5 com  $\rho = 0,82$  e  $\sigma_d^2 = 1$ .

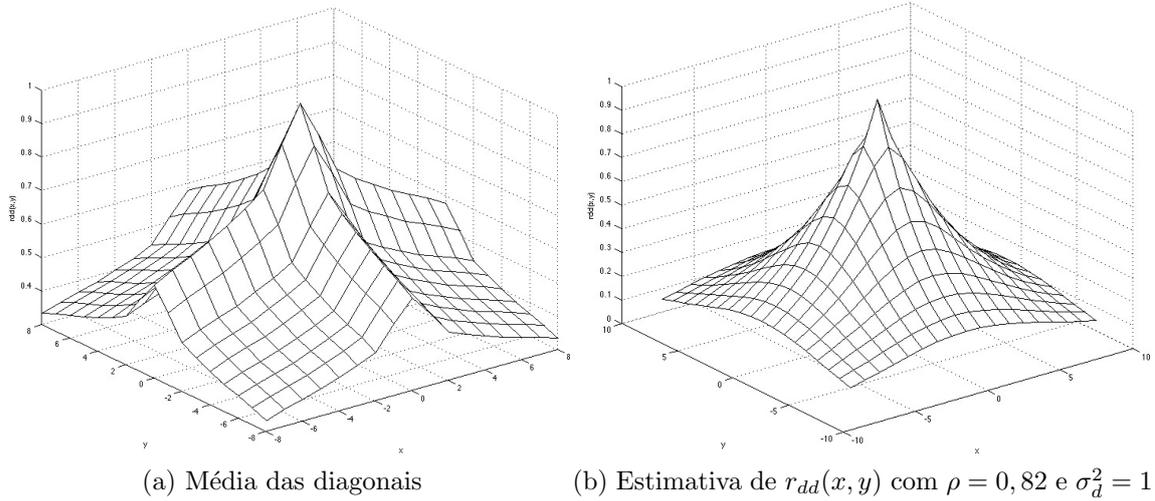


Figura 3.7: Estimativas de funções bidimensionais de autocorrelação de imagens.

Assim, podemos calcular  $r_{dd}$  para cada janela de observação  $\mathbf{g}_i$ , de forma a estabelecer as matrizes  $\mathbf{R}_i$  e  $\mathbf{P}_i$ . Em seguida, através da Eq. (3.3), podemos calcular os pesos  $\mathbf{W}_i$  relativos a cada janela de observação, utilizados para encontrar as estimativas em HR na Eq. (3.2). Ao estabelecer um modelo de autocorrelação que varia com a posição da janela de observação, evitamos o problema de muitas imagens não serem processos aleatórios WSS. Os pesos estarão sujeitos não apenas às variações espaciais dos pixels LR como às estatísticas locais dessas amostras. Para reduzir a complexidade da estimativa, podemos fixar o valor de  $\rho$  na Eq. (3.10) e reescrever o modelo de autocorrelação como:

$$r_{dd_i}(x, y) = \sigma_{d_i}^2 \rho \sqrt{(x^2 + y^2)}, \quad (3.11)$$

onde  $\sigma_{d_i}^2$  é a variância da  $i$ -ésima região da imagem desejada, responsável por  $\mathbf{g}_i$ .

Através da Eq. (3.9), podemos extrair uma relação entre a variância da imagem desejada,  $\sigma_{d_i}^2$ , e a variância da imagem degradada pela PSF,  $\sigma_{f_i}^2$ , dada por [17]:

$$\sigma_{d_i}^2 = \frac{1}{C(\rho)} \sigma_{f_i}^2, \quad (3.12)$$

onde

$$C(\rho) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho \sqrt{x^2 + y^2} \tilde{h}(x, y) dx dy \quad (3.13)$$

e

$$\tilde{h}(x, y) = h(x, y) * h(-x, -y). \quad (3.14)$$

Sendo o ruído aditivo da Eq. (3.4) independente de  $\mathbf{g}_i$ ,  $\sigma_{f_i}^2$  é estimado como:

$$\hat{\sigma}_{f_i}^2 = \hat{\sigma}_{g_i}^2 - \sigma_n^2 \quad (3.15)$$

onde  $\hat{\sigma}_{g_i}^2$  é a variância estimada dos elementos do vetor de observação  $\mathbf{g}_i$ .

Utilizando o modelo da Eq. (3.10) para obtermos  $E[\mathbf{f}_i \mathbf{f}_i^T]$  e  $E[\mathbf{f}_i \mathbf{d}_i^T]$ , as matrizes  $\mathbf{R}_i$  e  $\mathbf{P}_i$  podem ser definidas para qualquer  $\sigma_{d_i}^2$  como:

$$\mathbf{R}_i = E[\mathbf{g}_i \mathbf{g}_i^T] = \sigma_{d_i}^2 E[\mathbf{f}_i \mathbf{f}_i^T] + \sigma_n^2 \mathbf{I} \quad (3.16)$$

e

$$\mathbf{P}_i = E[\mathbf{g}_i \mathbf{d}_i^T] = \sigma_{d_i}^2 E[\mathbf{f}_i \mathbf{d}_i^T], \quad (3.17)$$

respectivamente. Logo, os pesos dos filtros de Wiener são calculados por:

$$\mathbf{W}_i = \left[ E[\mathbf{f}_i \mathbf{f}_i^T] + \frac{\sigma_n^2}{\sigma_{d_i}^2} \mathbf{I} \right]^{-1} E[\mathbf{f}_i \mathbf{d}_i^T]. \quad (3.18)$$

O coeficiente que multiplica a matriz identidade é uma razão ruído-sinal (NSR - *Noise-to-Signal Ratio*). Com a variação de  $\sigma_{d_i}^2$  ao longo dos vetores de observação, permitimos aos pesos dos filtros de Wiener uma adaptação relativa à variação da NSR ao longo da grade HR.

Considerando os modelos das Eqs. (3.8) e (3.9) comuns a todas as janelas de observação, podemos aproximar as matrizes  $\mathbf{R}_i$  e  $\mathbf{P}_i$  usando valores da autocorrelação  $r_{ff}$  e da correlação cruzada  $r_{df}$  pré-definidos, sendo necessário assim calcular apenas uma matriz  $\tilde{\mathbf{R}} = E[\mathbf{f}_i \mathbf{f}_i^T]$  e uma matriz  $\tilde{\mathbf{P}} = E[\mathbf{f}_i \mathbf{d}_i^T]$  para toda a grade HR na Eq. (3.10). Substituindo estas estimativa na Eq. (3.18), temos:

$$\mathbf{W}_i = \left[ \tilde{\mathbf{R}} + \frac{\sigma_n^2}{\sigma_{d_i}^2} \mathbf{I} \right]^{-1} \tilde{\mathbf{P}}. \quad (3.19)$$

A matriz  $\tilde{\mathbf{R}}$  contém as amostras de  $r_{ff}$  com índices iguais às distâncias entre os pixels da janela de observação, possuindo, portanto, dimensões  $K \times K$ , onde  $K$  é definido na Eq. (3.1). A Tabela 3.1 apresenta um algoritmo para o cálculo de  $\tilde{\mathbf{R}}$ .

Por outro lado, a matriz  $\tilde{\mathbf{P}}$  possui dimensões  $K \times D_x D_y$ , pois contém as amostras de  $r_{df}$  com índices iguais às distâncias entre cada pixel da subjanela da grade HR e cada pixel da janela de observação. A Tabela 3.2 apresenta um algoritmo para a implementação base de  $\tilde{\mathbf{P}}$ .

O resultado da SR AWF utilizando os pesos definidos na Eq. (3.19), com os

Tabela 3.1: Algoritmo para o cálculo da matriz  $\tilde{\mathbf{R}}$ .

```

for m = 1:Np
    for n = 1:Np
        % Np - Numero de pixels na janela gi da grade HR
        x = posX(m) - posX(n);
        % posX - coordenadas horizontal dos pixels em gi
        y = posY(m) - posY(n);
        % posY - coordenadas vertical dos pixels em gi
        R(m,n) = rdd(y,x);
        % rdd - funcao autocorrelacao da imagem desejada
    end
end

```

Tabela 3.2: Algoritmo para o cálculo da matriz  $\tilde{\mathbf{P}}$

```

for m = 1:Np
    % Np - Numero de pixels na janela gi da grade HR
    for n = 1:Npj
        % Npj - Numero de pixels na subjanela da grade HR
        x = posX(m) - posXj(n);
        % posX - coordenadas horizontal dos pixels em gi
        % posXj - coordenadas horizontal dos pixels na subjanela
        y = posY(m) - posYj(n);
        % posY - coordenadas vertical dos pixels em gi
        % posYj - coordenadas vertical dos pixels na subjanela
        P(m,n) = rff(y,x);
        % rff - funcao autocorrelacao de f(x, y)
    end
end

```

parâmetros  $\sigma_n = 0,03$  e  $\rho = 0,75$  da Eq. (3.11), é apresentado na Fig. 3.8(d). Uma comparação entre a imagem real desejada, uma imagem LR e uma imagem HR interpolada é observada na Fig. 3.8. A variação do parâmetro  $\rho$  no modelo de autocorrelação  $r_{dd}$  utilizado produz resultados mais suaves para  $\rho$  próximo de um, e imagens com mais detalhes de alta frequência para valores menores deste parâmetro. Valores altos podem borrar a imagem além do necessário, e valores baixos, apesar de ressaltarem as bordas, deixam a super-resolução suscetível ao ruído aditivo. A Fig. 3.9 apresenta uma comparação para diferentes valores de  $\rho$ .

Constatamos então que, com a redução do parâmetro  $\rho$ , a SR AWF produz imagens HR com bordas mais afiadas. Visando realçar as bordas das imagens, utilizaremos ao longo do trabalho um valor de  $\rho$  ligeiramente inferior ao encontrado na análise da autocorrelação das imagens naturais monocromáticas. Para este modelo de SR AWF apresentado nas Figs. 3.8 e 3.9, utilizaremos o valor de  $\rho = 0,75$ , como na Fig. 3.8(d), em vez do valor de  $\rho = 0,82$ , estimado através da autocorrelação das imagens naturais e utilizado na Fig. 3.9(c).



(a) Imagem desejada



(b) Imagem LR



(c) Imagem HR interpolada



(d) Imagem HR com SR AWF ( $\rho = 0,75$ )

Figura 3.8: Super-resolução com filtro de Wiener - ruído:  $\sigma_n = 0,03$ .



(a) Imagem HR com SR AWF ( $\rho = 0, 40$ )



(b) Imagem HR com SR AWF ( $\rho = 0, 50$ )



(c) Imagem HR com SR AWF ( $\rho = 0, 82$ )



(d) Imagem HR com SR AWF ( $\rho = 0, 90$ )

Figura 3.9: SR AWF com diferentes modelos de autocorrelação - ruído:  $\sigma_n = 0, 03$ .

### 3.3 SR AWF com Wavelet

A implementação do método de SR AWF nas sub-bandas da Transformada Wavelet sem subamostragem [22] tem o intuito de reduzir o ruído e aumentar a nitidez da imagem de alta resolução. A Transformada Wavelet [23–26] permite dividir um sinal em diferentes sub-bandas de frequência, a partir de um par de filtros passa-baixas  $H_0(z)$  e passa-altas  $H_1(z)$ . No caso bidimensional, aplicamos os filtros nas direções vertical e horizontal, obtendo quatro sinais resultantes, um de baixas frequências em ambas as direções (LL), e outros três com detalhes verticais (LH), horizontais (HL) e diagonais (HH).

Em [27], é proposto um método híbrido que utiliza Wavelet e interpolação não uniforme para restauração de imagens. Na adaptação do método SR AWF com Wavelet [22], aplica-se um nível de decomposição às imagens LR, e os filtros de Wiener são aplicados separadamente às sub-bandas LL, LH, HL e HH. Portanto, a SR AWF aplicada a um conjunto das mesmas sub-bandas de imagens LR gera uma sub-banda correspondente da imagem HR. A Transformada Wavelet Inversa aplicada às sub-bandas de alta resolução resultará na reconstrução da imagem estimada.

Para conservar o efeito de *aliasing* original das imagens LR, a Transformada Wavelet aplicada às imagens LR da super-resolução não possui subamostragem. O *aliasing* presente no conjunto de imagens LR possibilita a recuperação da imagem HR. Quando as imagens resultantes de aplicação dos filtros da Transformada Wavelet são subamostradas, um novo efeito de *aliasing* é introduzido, corrompendo o *aliasing* original resultante da subamostragem do sistema de captura das imagens LR.

Desta forma, a matriz de pesos dos filtros de Wiener, definida na Eq. (3.19), pode ser reescrita para as sub-bandas da Wavelet com um nível de decomposição sem subamostragem como:

$$\mathbf{W}_{i,l} = \left[ \tilde{\mathbf{R}} + \frac{\sigma_n^2}{\sigma_{d_{i,l}}^2} \mathbf{I} \right]^{-1} \tilde{\mathbf{P}}. \quad (3.20)$$

onde  $l$  é a sub-banda correspondente, variando de 1 a 4, respectivamente LL, LH, HL e HH. O diagrama da Fig. 3.10 ilustra o processo da SR AWF com Wavelet.

A super-resolução por filtragem de Wiener aplicada às sub-bandas da Wavelet possibilita combinar uma técnica de *denoising* à Transformada Wavelet nas imagens de baixa resolução. O objetivo é reduzir o ruído aditivo presente nas imagens LR, através da análise dos histogramas dos detalhes horizontais, verticais e diagonais da Wavelet, aplicando limiares para eliminar os coeficientes com valores fora da faixa de interesse, considerados ruído. Valores altos para os limiares resultam em perda de informação capaz de causar distorções aparentes na imagem. Por outro lado, pequenos limiares podem não ser capazes de extrair o ruído.

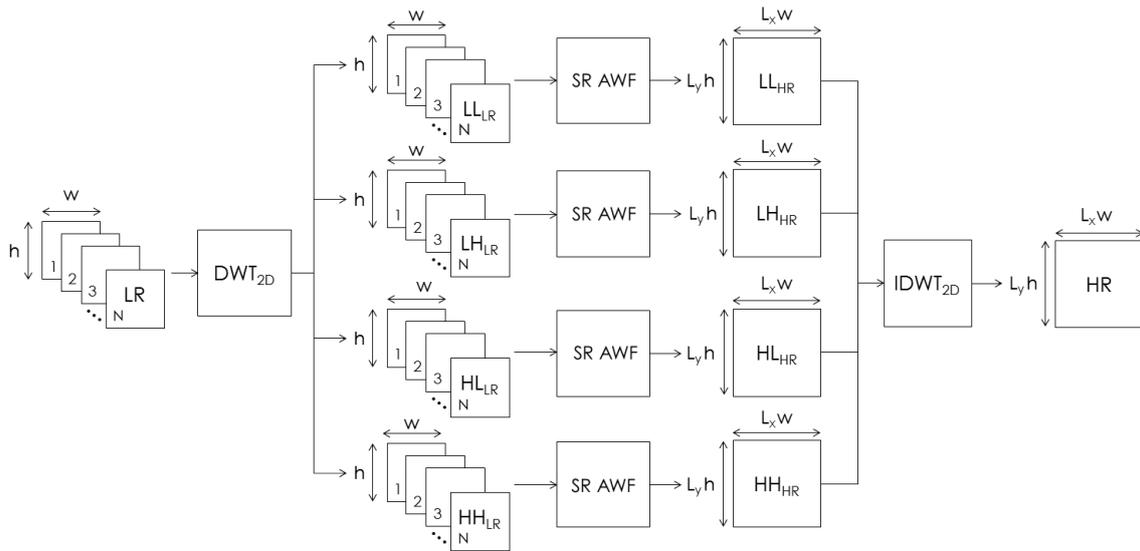


Figura 3.10: Diagrama do algoritmo SR AWF com Wavelet.

Uma comparação entre o método SR AWF convencional e com Wavelets é apresentada na Fig. 3.11. Apesar da redução significativa do ruído utilizando Wavelet, em relação ao método convencional, as imagens resultantes desta adaptação aparecem mais borradas, ou seja, apresentam perda de detalhes de alta frequência. Visando reduzir este efeito, uma modificação é proposta.



(a) Imagem HR com SR AWF



(b) Imagem HR com SR AWF Wavelet

Figura 3.11: Comparação entre SR AWF e SR AWF com Wavelet ( $\rho = 0,75$ ).

### 3.3.1 Ajuste do Modelo de Autocorrelação por Sub-Banda

Um dos principais aspectos da Transformada Wavelet é a possibilidade de interagir diferentemente com as diferentes faixas de frequência, utilizando o conhecimento prévio para aplicar filtros específicos em cada sub-banda. Conforme observado na Fig. 3.9, o ajuste do parâmetro  $\rho$  na Eq. (3.10) influencia no realce ou descarte de altas frequências na imagem resultante em alta resolução. Assim, podemos modelar funções de autocorrelação para cada uma das sub-bandas da Wavelet, utilizando a maior correlação existente na direção da sub-banda.

Observando as imagens da Fig. 3.12, podemos constatar o alongamento das bordas na direção da respectiva alta frequência. Na Fig. 3.12(b), os elementos verticais estão mais destacados do que na Fig. 3.12(c), que tem ênfase nas componentes de alta frequência na horizontal. As componentes diagonais são apresentadas na Fig. 3.12(d) e a Fig. 3.12(a) contém as componentes de baixa frequência.

Podemos reescrever a Eq. (3.10) para cada uma das sub-bandas, de forma a alongar a função de autocorrelação  $r_{dd}$  na direção dos componentes da sub-banda. Para a sub-banda de baixa frequência, o modelo de autocorrelação utilizado é idêntico ao da Eq. (3.10). Os modelos das componentes de alta frequência podem ser reescritos como:

$$r_{dd_{HL}}(x, y) = \sigma_{d_{HL}}^2 \rho \sqrt{(2x)^2 + y^2}, \quad (3.21)$$

$$r_{dd_{LH}}(x, y) = \sigma_{d_{LH}}^2 \rho \sqrt{x^2 + (2y)^2}, \quad (3.22)$$

e

$$r_{dd_{HH}}(x, y) = \sigma_{d_{HH}}^2 \rho \sqrt{(2x)^2 + (2y)^2}, \quad (3.23)$$

onde  $r_{dd_{HL}}$ ,  $r_{dd_{LH}}$  e  $r_{dd_{HH}}$  são respectivamente os modelos de autocorrelação adotados para as sub-bandas de frequências horizontal, vertical e diagonal. A diferença entre os modelos pode ser visualizada na Fig. 3.13. O parâmetro  $\rho$  pode ser ajustado diferentemente para as baixas e as altas frequências. Os valores de  $\rho$  para o cálculo dos gráficos na Fig. 3.13 são 0,6 para a sub-banda de baixa frequência e 0,45 para as sub-bandas de altas frequências.

O objetivo dessas modificações é reduzir o borramento causado pela SR AWF aplicada aos coeficientes da Wavelet. O fato de acelerar o decaimento da função  $r_{dd}$  torna o modelo mais apropriado às altas frequências, que possuem uma autocorrelação mais fraca. No modelo proposto, não apenas o valor de  $\rho$  é reduzido, como a função de autocorrelação  $r_{dd}$  é estreitada na direção ortogonal às altas frequências.

A Fig. 3.14 apresenta uma comparação entre os dois modelos de SR AWF aplicadas nas sub-bandas da Wavelet. Na Fig. 3.14(a) temos um único modelo de autocorrelação aplicado a todas as sub-bandas, com  $\rho = 0,75$ . O resultado da aplicação dos modelos de  $r_{dd}$  da Fig. 3.13 pode ser visualizado na Fig. 3.14(b).



(a) Baixa frequência



(b) Alta frequência horizontal



(c) Alta frequência vertical



(d) Alta frequência diagonal

Figura 3.12: Coeficientes da Wavelet sem decimação.

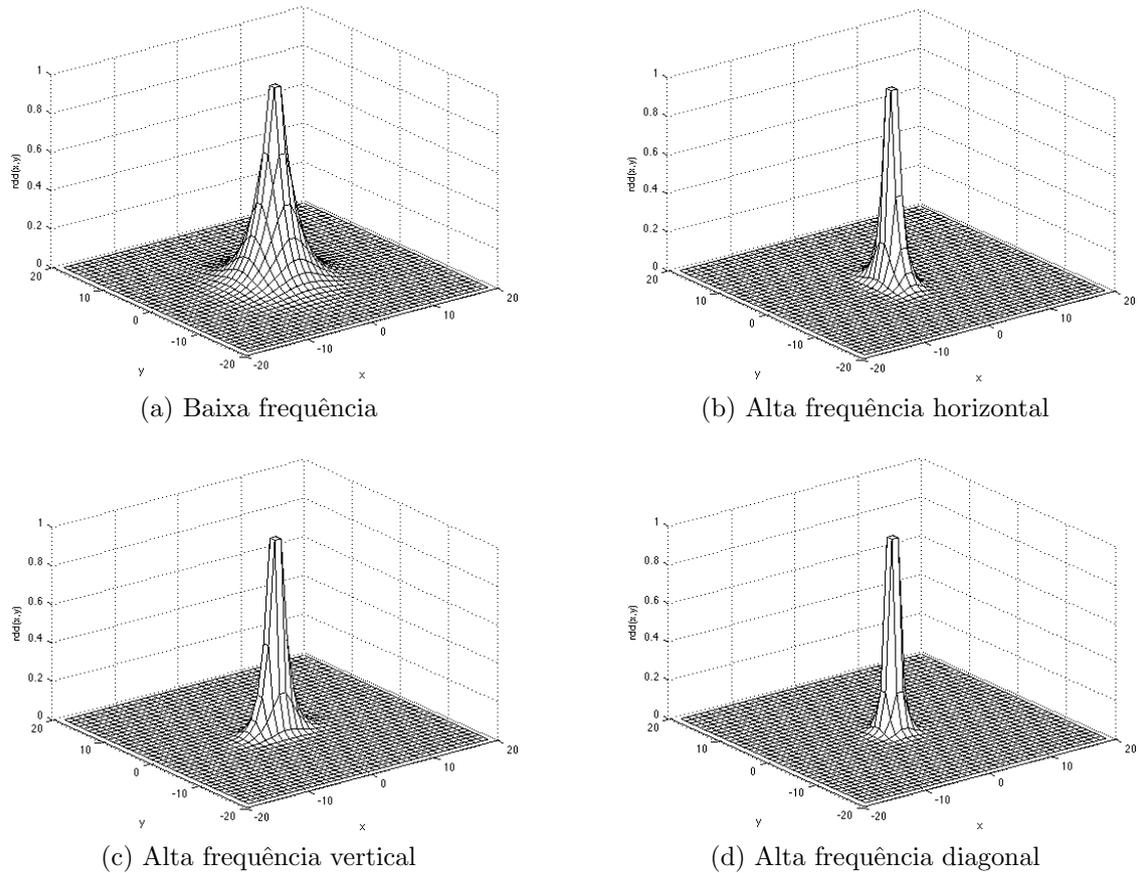


Figura 3.13: Autocorrelação relativa às diferentes sub-bandas da Wavelet.

É perceptível, em uma comparação visual entre as duas imagens da Fig. 3.14, o maior realce de bordas e altas frequências, inclusive ruído, na imagem com diferentes modelos de autocorrelação, que apresentam decaimento mais acentuado. Por outro lado, a imagem estimada com base em um modelo  $r_{dd}$  único apresenta maiores perdas de detalhes em alta frequência, apesar da maior redução de ruído muitos detalhes ficam comprometidos.

Apesar de imagens com predominância de baixas frequências apresentarem forte autocorrelação, observamos que ao empregar altos valores de  $\rho$  às sub-bandas LL, a imagem HR resultante da SR AWF com Wavelets apresenta-se borrada. Para melhorar o resultado, foi necessário estreitar o modelo de autocorrelação  $r_{dd}$  empregado nas baixas frequências, reduzindo assim o valor de  $\rho$ . O valor de  $\rho = 0,6$  para o modelo  $r_{dd_{LL}}$  mostrou-se satisfatório para todos os casos analisados.



(a) SR AWF com Wavelet com  $r_{dd}$  única



(b) SR AWF com Wavelet com diferentes  $r_{dd}$

Figura 3.14: Comparação de SR AWF com Wavelet.

# Capítulo 4

## Super-Resolução de Imagens Coloridas

Os princípios do tratamento de imagens coloridas baseiam-se na forma como o sistema visual humano reage à luz, capaz de estimular uma resposta entre os comprimentos de onda 350 nm e 780 nm. O olho possui dois tipos de células fotorreceptoras na retina, conhecidas como cones e bastonetes. A segunda reage apenas a níveis de luminância. Por outro lado, os cones se encontram em bem menor número na retina e reagem a três faixas de frequências diferentes, consideradas as três cores primárias: vermelho, verde, e azul. A combinação das cores primárias permite a geração de praticamente todas as cores visíveis [21]. Por este motivo é comum a representação de imagens coloridas no padrão RGB, mas modelos que separam luminância e crominância, como o YIQ, também são utilizados.

A super-resolução de imagens coloridas pode ser rapidamente implementada com a expansão de técnicas de SRR monocromáticas aplicadas separadamente aos canais de cor das imagens coloridas. Métodos mais robustos costumam considerar características e propriedades das imagens. Um importante problema a ser considerado é o impacto da interpolação do *Color Filter Array* e a reconstrução da imagem por *demosaicing*. A resolução do problema *demosaicing* super-resolução em duas etapas é considerada subótima, pois ambas as etapas correspondem a problemas de resolução de sensores de câmera, e por isso [28] propõe uma estrutura conjunta.

Este capítulo levará em consideração apenas imagens pós-*demosaicing*, uma vez que as técnicas de super-resolução desenvolvidas para estas imagens se mostraram satisfatórias ao longo dos últimos anos. Uma variação do método SR AWF [17] será aplicada em imagens coloridas, primeiramente com a super-resolução aplicada em cada um dos canais de cor, em seguida com uma adaptação, valorizando a luminância em relação à crominância. Por último, uma modificação na técnica de super-resolução e *demosaicing* com múltiplos quadros [29], que utiliza as relações entre os canais de cores, será apresentada.

## 4.1 SR AWF nos Canais de Cores

O objetivo desta seção é mostrar a adaptação do algoritmo SR AWF para imagens coloridas. De maneira simples, o algoritmo será aplicado independentemente nos diferentes canais de cores, levando em consideração apenas a correlação intracanal, não incluindo relações entre os canais de cor. Inicialmente, a SR AWF será aplicada diretamente aos canais R, G e B. Visando reduzir a carga computacional, o sinal será dividido em luminância e crominância, e o sinal de crominância será decimado por um fator igual a dois em cada dimensão. Uma terceira variação utilizando Wavelets na luminância visa reduzir o erro causado pelo ruído.

### 4.1.1 RGB

O algoritmo proposto consiste em aplicar a técnica SR AWF ilustrada na Fig. 3.2 nos três canais de cores primários. Desta forma teremos três janelas de observação, representadas pelos vetores  $\mathbf{g}_{R_i} = [g_{R_{i,1}}, g_{R_{i,2}}, \dots, g_{R_{i,K_i}}]^T$ ,  $\mathbf{g}_{G_i} = [g_{G_{i,1}}, g_{G_{i,2}}, \dots, g_{G_{i,K_i}}]^T$  e  $\mathbf{g}_{B_i} = [g_{B_{i,1}}, g_{B_{i,2}}, \dots, g_{B_{i,K_i}}]^T$ , onde  $R$ ,  $G$  e  $B$  são as correspondentes grades HR para as respectivas cores, e  $i$  é o índice referente à posição da janela na grade HR. Consequentemente, a estimativa pode ser reescrita como:

$$\hat{\mathbf{d}}_{C,i} = \mathbf{W}_{C,i}^T \mathbf{g}_{C,i}, \quad C = R, G, B; \quad (4.1)$$

onde  $\mathbf{W}_{C,i}$  é uma matriz de pesos com dimensão  $K_i \times D_x D_y$ , relativa a cada um dos canais de cor. A Fig. 4.1 apresenta o diagrama da adaptação do algoritmo SR AWF para RGB. Uma comparação entre as imagens HR original, LR, HR interpolada (interpolação bilinear) e HR com SR AWF em RGB é apresentada na Fig. 4.2.

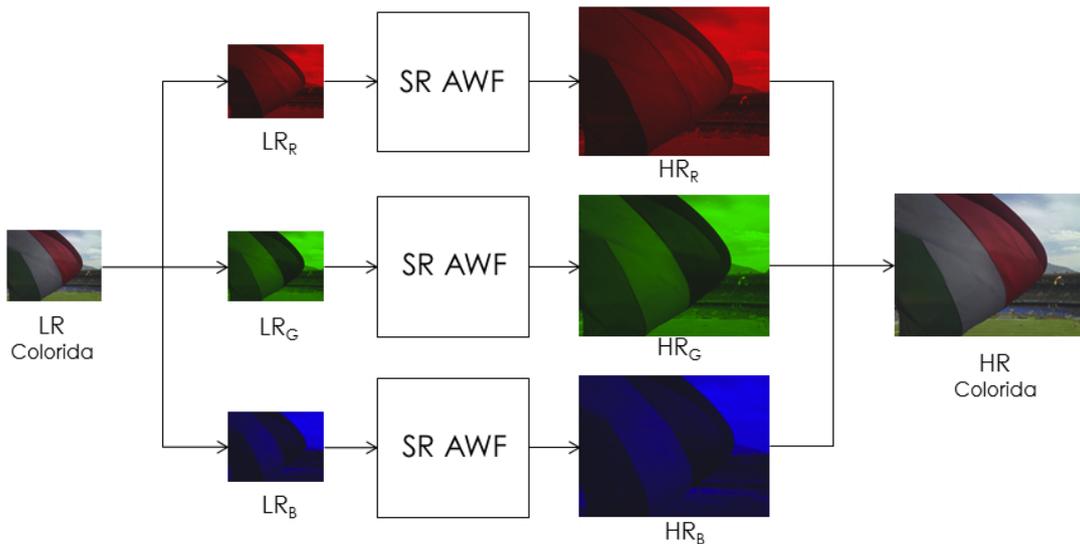


Figura 4.1: Diagrama do algoritmo de SR AWF para os canais RGB.



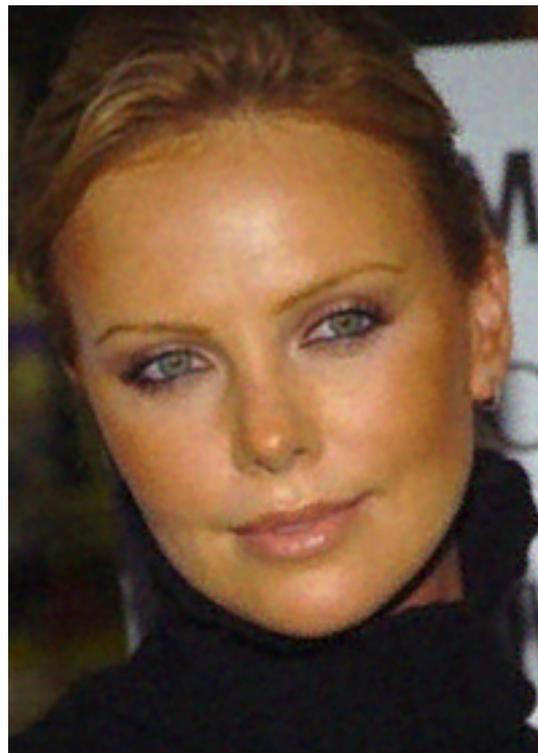
(a) Imagem HR original



(b) Imagem LR



(c) Imagem HR interpolada



(d) Imagem HR com SR AWF em RGB

Figura 4.2: Super-resolução colorida com filtro de Wiener - ruído:  $\sigma_n = 0,03$ .

### 4.1.2 Compressão de Crominância

O modelo YIQ foi inicialmente desenvolvido para a transmissão comercial de TV colorida, permitindo a manutenção e a compatibilidade com o padrão monocromático NTSC (*National Television Standards Committee*) em funcionamento à época. A componente  $Y$ , conhecida como luminância, contém todas as informações necessárias para compor uma imagem monocromática. Informações relativas às cores, também chamadas de crominância, são concentradas nas componentes  $I$  e  $Q$ . A conversão de RGB para YIQ é definida como [30]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (4.2)$$

Analogamente, os sistemas PAL (*Phase Alternating Line*) e SECAM (*Séquentiel Couleur à Mémoire*), empregam o modelo YUV. No contexto de imagens digitais é comum utilizar o modelo YCbCr. A conversão de RGB para YCbCr é definida como:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,1687 & -0,331 & 0,5 \\ 0,5 & -0,419 & 0,081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (4.3)$$

Uma das principais vantagens dos modelos que separam luminância e crominância em processamento de imagens é que a luminância é desacoplada das informações de cores. Assim, processa-se a luminância sem impactar significativamente o conteúdo das cores de uma imagem colorida.

Essas representações permitiram aos sistemas de TV explorar a maior sensibilidade da visão humana às mudanças na luminância, em relação a alterações de matiz ou saturação. Isto possibilita utilizar uma banda de transmissão mais larga para o sinal  $Y$ , idêntica em ambas as representações, e bandas mais estreitas para os sinais de crominância ( $I$  e  $Q$  ou  $Cb$  e  $Cr$ ). Em uma compressão de imagem cada pixel da componente  $Y$  pode ser quantizado em 8 bits, enquanto as componentes  $Cb$  e  $Cr$  podem utilizar apenas 2 bits por pixel original (4:2:0). Ao todo 12 bits serão utilizados, em vez dos 24 bits originais.

Da mesma forma em que foi aplicada aos canais RGB, a técnica SR AWF pode ser aplicada aos canais YCbCr, porém utilizando as vantagens desta representação. A proposta é utilizar apenas 1/4 dos pixels de crominância para estimar uma imagem HR sem perdas de qualidade visual. Após a conversão das imagens LR para o espaço YCbCr, os sinais  $Cb$  e  $Cr$  são subamostrados por um fator de dois em cada dimensão (linhas e colunas). As componentes resultantes  $Cb$  e  $Cr$  em alta resolução terão uma área quatro vezes menor do que a resultante  $Y$ . Um simples método de interpolação

pode ser utilizado para adequar as dimensões. O diagrama do método proposto é ilustrado na Fig. 4.3.

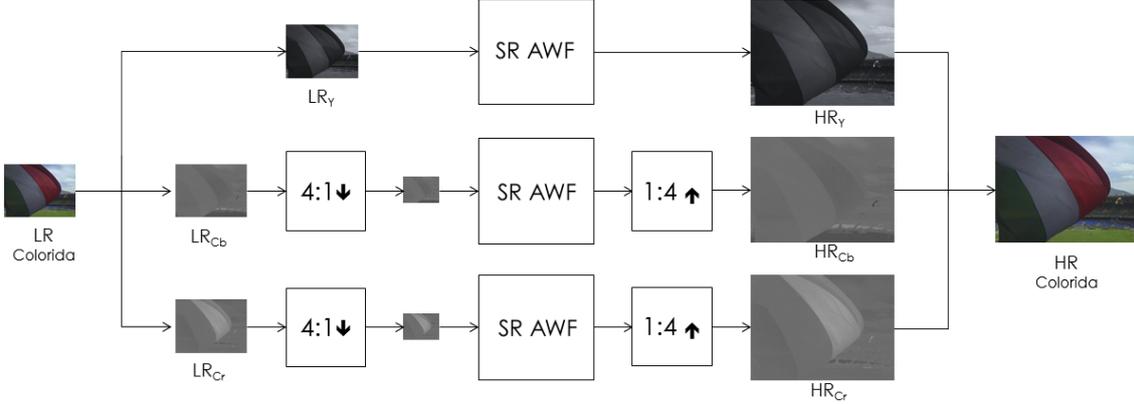


Figura 4.3: Diagrama do algoritmo de SR AWF com compressão de crominância.

A vantagem de processar a crominância em uma resolução menor do que a luminância está na redução do número de operações aritméticas necessárias. Com a redução espacial, apenas 1/4 dos filtros  $\mathbf{W}_i$  utilizados na luminância será necessária em cada um dos canais de crominância, pois o número de filtros necessários está diretamente ligado ao número de pixels processados. Isto significa uma redução de até 50% na quantidade de filtros de Wiener adaptativos, considerando todos os canais da imagem. A definição dos pesos de  $\mathbf{W}_i$  na Eq. (3.19) é a operação mais custosa do algoritmo SR AWF, pois para cada índice  $i$  há a inversão de uma matriz com dimensões  $K \times K$ , sendo  $K$  definido na Eq. (3.1). O aumento do número de pixels na janela de observação eleva quadraticamente o número de operações necessárias para inverter a matriz  $\mathbf{R}_i$ .

Os vetores de observação podem ser descritos como  $\mathbf{g}_{Y_i} = [g_{Y_{i,1}}, g_{Y_{i,2}}, \dots, g_{Y_{i,K_i}}]^T$ , com  $i$  variando de 1 até o número máximo de janelas de observação, e  $\mathbf{g}_{Cb_j} = [g_{Cb_{j,1}}, g_{Cb_{j,2}}, \dots, g_{Cb_{j,K_j}}]^T$  e  $\mathbf{g}_{Cr_j} = [g_{Cr_{j,1}}, g_{Cr_{j,2}}, \dots, g_{Cr_{j,K_j}}]^T$ , sendo o valor máximo de  $j$  igual a 1/4 do valor máximo de  $i$ . Sendo os índices  $Y$ ,  $Cb$  e  $Cr$  correspondentes às respectivas grades HR, a estimativa pode ser reescrita como:

$$\hat{\mathbf{d}}_{Y,i} = \mathbf{W}_{Y,i}^T \mathbf{g}_{Y,i}, \quad (4.4)$$

$$\hat{\mathbf{d}}_{C,j} = \mathbf{W}_{C,j}^T \mathbf{g}_{C,j}, \quad C = Cb, Cr. \quad (4.5)$$

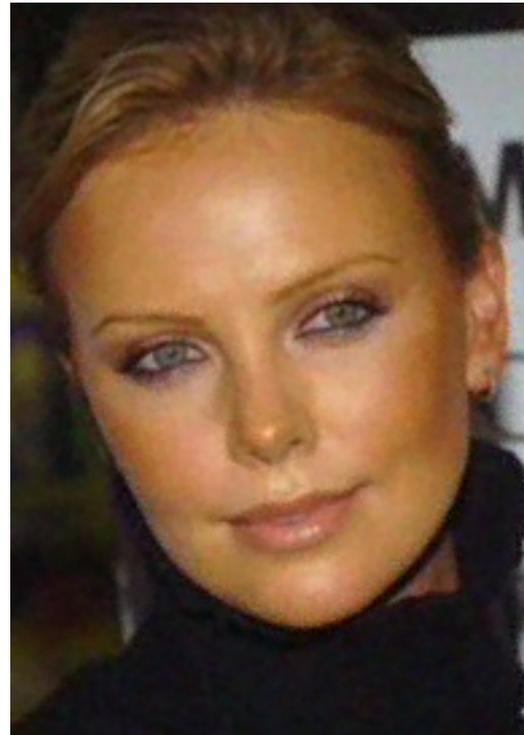
As operações para conversão do espaço RGB para YCbCr são facilmente implementadas, como mostrado na Eq. (4.3), com um custo computacional muito inferior ao ganho produzido com a redução de processamento na crominância. Para retornar ao espaço RGB, basta realizar o produto da inversa da matriz de transformação com os sinais  $Y$ ,  $Cb$  e  $Cr$ . Uma análise sobre a complexidade computacional dos

algoritmos de SR AWF nos diferentes canais de cores é apresentada na Seção 5.4.

Uma comparação entre o algoritmo SR AWF aplicado uniformemente aos canais  $R$ ,  $G$  e  $B$  e aos canais  $Y$ ,  $Cb$  e  $Cr$  com redução espacial de crominância é apresentada na Fig. 4.4. Em uma análise visual, não é possível estabelecer se há perda de qualidade na imagem com compressão de crominância. A comparação dos resultados com base em métricas de avaliação será apresentada no Capítulo 5.



(a) SR AWF em RGB



(b) SR AWF em YCbCr

Figura 4.4: Comparação da SR AWF com compressão de crominância  $Cb$  e  $Cr$ .

### 4.1.3 Emprego da Transformada Wavelet na Luminância

Nesta seção, investigamos o emprego da Transformada Wavelet nas técnicas de super-resolução, com o objetivo de melhorar a qualidade da imagem HR resultante, quando as imagens LR apresentam ruído.

A principal desvantagem da aplicação da Transformada Wavelet à SR AWF é a introdução de mais canais para serem processados. Basicamente o número de operações é quadruplicado, pois é necessário obter a imagem HR de cada uma das quatro imagens obtidas com a Wavelet sem decimação e com um estágio de decomposição. Como descrito na Seção 3.3, a Transformada Wavelet sem decimação é utilizada para evitar a degradação do *aliasing* original das imagens LR através de uma nova subamostragem. Assim, cada sub-banda resultante da Transformada

Wavelet sem decimação possui um número de coeficientes igual à quantidade de pixels da imagem que sofreu a transformação Wavelet.

Com base nas propriedades da visão e das componentes YCbCr da imagem discutidas neste capítulo, podemos aplicar a Transformada Wavelet apenas na luminância, enquanto à crominância é aplicado o filtro adaptativo de Wiener imediatamente.

Desta forma, o ruído e as altas frequências da imagem são tratados na luminância, onde são mais significativos para a resposta visual, aumentando apenas o custo computacional para o processamento da luminância. Se a Transformada Wavelet for aplicada aos três canais RGB, ou YCbCr, um total de 12 operações de SR AWF seriam aplicadas às 12 grades HR existentes. Com a aplicação da Wavelet apenas na luminância, são necessárias apenas seis operações de super-resolução.

O diagrama do algoritmo SR AWF com Wavelet na luminância é apresentado na Fig. 4.5. Ao contrário da compressão de luminância, esta proposta não visa uma redução de operações aritméticas. O principal objetivo é reduzir o impacto do ruído aditivo nas imagens LR, apesar de aumento no custo computacional de aproximadamente o triplo de operações em relação à SR AWF aplicada aos canais RGB. O aumento da complexidade computacional ocorre pois a aplicação da Transformada Wavelet sem decimação na luminância introduz quatro novos conjuntos de imagens com o mesmo número de coeficientes das luminâncias das imagens LR originais.

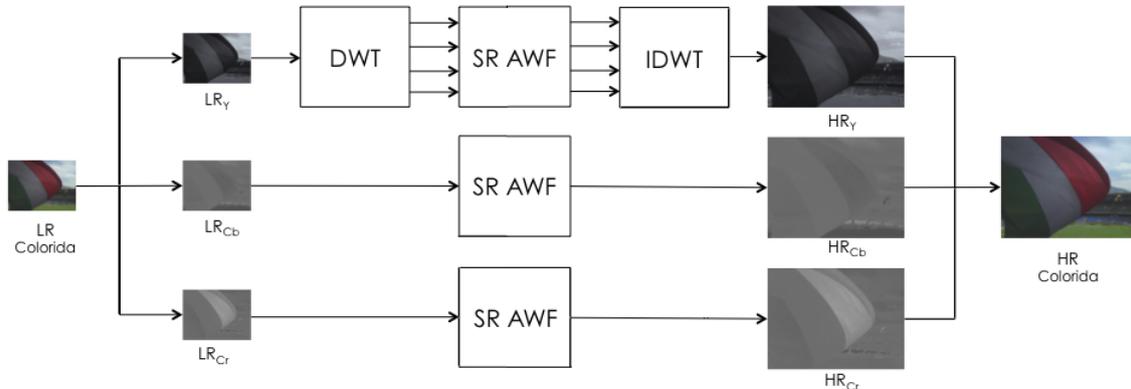
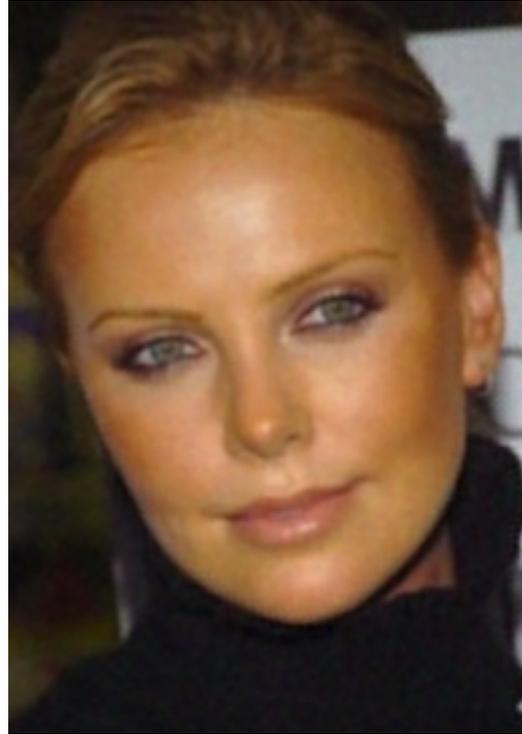


Figura 4.5: Diagrama do algoritmo de SR AWF com Wavelet na luminância.

Uma comparação entre a SR AWF em RGB e a SR AWF com Wavelet na luminância é apresentada na Fig. 4.6. A imagem HR resultante da SR AWF aplicada em RGB na Fig. 4.6(a) apresenta mais componentes de alta frequência, principalmente ruído, do que a imagem HR da Fig. 4.6(b) com SR AWF com Wavelet na luminância. Visualmente, a imagem da Fig. 4.6(b) apresenta menor intensidade de ruído e regiões mais suaves do que a imagem da Fig. 4.6(a). Por outro lado, a SR AWF com Wavelet na luminância apresenta perda de informação visível em regiões de altas frequências, como as íris dos olhos.



(a) SR AWF em RGB



(b) SR AWF com Wavelet na luminância

Figura 4.6: Comparação da SR AWF com Wavelet na luminância.

## 4.2 SR *Multiframe* Adaptativo

Os métodos propostos na seção anterior utilizam características dos sinais de cores para reduzir os custos computacionais, mas não os relacionam para melhorar a eficácia do resultado final. Em [28], um conjunto de métodos robustos é proposto para a melhoria e fusão de imagens. Em [29], é proposto um algoritmo conjunto (super-resolução e *demosaicing* simultâneos) que utiliza pesos que se adaptam com base nos canais de cores e suas relações. Este algoritmo conjunto é baseado em um processo iterativo para obtenção da super-resolução [16].

A função-custo proposta em [29] visa resolver, além da super-resolução, problemas de descontinuidades de cor no processo de *demosaicing*, porém adaptações podem ser realizadas para tratar apenas o caso de super-resolução de imagens coloridas. Os termos da função-custo consistem em:

1. Impor semelhança entre os pixels LR e as estimativas HR;
2. Valorizar bordas afiadas na luminância da imagem HR;
3. Reforçar a suavidade das componentes de croma na imagem HR;
4. Incentivar a homogeneidade da localização e orientação das bordas nos diferentes canais de cor (dependência entre canais de cores).

Estes termos serão discutidos com maiores detalhes a seguir.

### 4.2.1 Termo de Fidelidade

O termo de fidelidade mede a semelhança entre os pixels da imagem resultante HR e as imagens amostradas em LR. Uma das funções-custos mais comuns para medir a proximidade é o método dos mínimos quadrados (LS - *Least Squares*) que minimiza a norma  $L_2$

$$\gamma(\mathbf{y}, \mathbf{D}\mathbf{x}) = \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2, \quad (4.6)$$

onde  $\mathbf{y}$ ,  $\mathbf{x}$  e  $\mathbf{D}$  são definidos na Eq. (2.1). Para o caso de ruído aditivo ( $\mathbf{v}$ ) branco, Gaussiano e com média zero, esta abordagem provê a estimativa de  $\mathbf{x}$  pela máxima verossimilhança [31, 32]. Alguns estudos estatísticos sugerem utilizar um ruído Laplaciano para as técnicas de fusão de imagens [33]. A comparação entre os desempenhos das normas  $L_2$  e  $L_1$  é discutida em [14], onde a norma  $L_1$  mostrou-se mais robusta a erros de estimação de movimento entre as imagens de baixa resolução em relação à norma  $L_2$ .

Considerando o modelo apresentado na Eq. (2.1), com movimento relativo, borramento e subamostragem, o termo de fidelidade é definido como:

$$J_0(\mathbf{x}) = \sum_{i=R,G,B} \sum_{k=1}^N \|\mathbf{D}(k)\mathbf{H}(k)\mathbf{F}(k)\mathbf{x}_i - \mathbf{y}_i(k)\|_2^2. \quad (4.7)$$

Sendo  $J_0(\mathbf{x})$  um somatório baseado nas diferenças entre as  $k$  imagens LR, representadas por  $\mathbf{y}_i(k)$ , e a imagem desejada  $\mathbf{x}$ . Durante o processo de estimação da imagem desejada, o valor de  $\mathbf{x}_i$  será substituído pela imagem estimada  $\hat{\mathbf{x}}_i$ . No caso ideal, quando  $\hat{\mathbf{x}} = \mathbf{x}$ , a diferença  $\mathbf{D}(k)\mathbf{H}(k)\mathbf{F}(k)\mathbf{x} - \mathbf{y}(k)$  será igual a zero, satisfazendo a Eq. 2.1. O termo *multiframe* é empregado no nome do método por este algoritmo utilizar várias imagens para compor os ponderadores que serão utilizados para estimar a imagem HR desejada.

A derivada parcial de  $J_0(\mathbf{x})$  relativa a cada um dos canais de cor, que será utilizada para calcular as estimações da imagem desejada, é

$$\frac{\partial J_0(\mathbf{x})}{\partial \mathbf{x}_i} = \mathbf{F}_{(k)}^T \mathbf{H}_{(k)}^T \mathbf{D}_{(k)}^T [\mathbf{D}_{(k)} \mathbf{H}_{(k)} \mathbf{F}_{(k)} \mathbf{x}_i - \mathbf{y}_i(k)], \quad (4.8)$$

onde  $i = R, G$  ou  $B$ .

## 4.2.2 Luminância

Como citado anteriormente, o olho humano é mais sensível aos detalhes da luminância do que às componentes de crominância de uma imagem. Portanto, é importante durante a reconstrução da imagem HR criar bordas mais definidas na luminância [34].

Um método bem sucedido utilizado para *denoising* e *deblurring* é o *Total Variation* (TV) [35]. Este método penaliza a quantidade total de variações na imagem, utilizando uma medida de norma  $L_1$  da magnitude do gradiente que pode ser definida como:

$$\Gamma_{TV}(\mathbf{x}) = \|\nabla \mathbf{x}\|_1, \quad (4.9)$$

onde  $\nabla$  é o operador gradiente.

Relacionando o TV a um modelo de filtro bilateral [36], foi proposta uma regularização robusta em [14] chamada de Bilateral-TV (B-TV). A aplicação do B-TV à luminância resulta em um termo de regularização definido como:

$$J_1(\mathbf{x}) = \sum_{l=-P}^P \sum_{m=-P}^P \alpha^{|m|+|l|} \|\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y\|_1, \quad (4.10)$$

onde  $S_x^l$  e  $S_y^m$  são operadores correspondentes ao deslocamento da imagem  $\mathbf{x}$  por  $l$  pixels na direção horizontal e  $m$  pixels na vertical. O parâmetro  $P$  define o tamanho do filtro bilateral. Um peso  $0 < \alpha < 1$  é aplicado para efeito de decaimento espacial durante o somatório. Por último,  $\mathbf{x}_Y$  é a componente de luminância da imagem, definida através da Eq. (4.3) como  $\mathbf{x}_Y = 0,299\mathbf{x}_R + 0,587\mathbf{x}_G + 0,114\mathbf{x}_B$ .

Considerando a derivada do termo  $\|\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y\|_1$  na Eq. (4.10) em relação a um dos canais de cor da imagem como:

$$\frac{\partial \|\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y\|_1}{\partial \mathbf{x}_i} = \frac{\partial \|\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y\|_1}{\partial \mathbf{x}_Y} \frac{\partial \mathbf{x}_Y}{\partial \mathbf{x}_i}, \quad (4.11)$$

onde

$$\frac{\partial \|\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y\|_1}{\partial \mathbf{x}_Y} = [I - S_y^{-m} S_x^{-l}] \text{sign}(\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y), \quad (4.12)$$

temos que a derivada parcial de  $J_1(\mathbf{x})$  em relação a qualquer canal de cor RGB é

$$\frac{\partial J_1(\mathbf{x})}{\partial \mathbf{x}_i} = \sum_{l=-P}^P \sum_{m=-P}^P \alpha^{|m|+|l|} [I - S_y^{-m} S_x^{-l}] \text{sign}(\mathbf{x}_Y - S_x^l S_y^m \mathbf{x}_Y) \frac{\partial \mathbf{x}_Y}{\partial \mathbf{x}_i}, \quad (4.13)$$

onde  $S_x^{-l}$  e  $S_y^{-m}$  correspondem às transpostas das matrizes  $S_x^l$  e  $S_y^m$ , respectivamente.

### 4.2.3 Crominância

A crominância também necessita de uma regularização espacial, uma vez que o olho humano é sensível às mudanças de baixas frequências destes canais [37]. A regularização de Tikhonov é amplamente empregada na forma [31]

$$\Gamma_T(\mathbf{x}) = \|\Lambda\mathbf{x}\|_2^2, \quad (4.14)$$

onde  $\Lambda$  é geralmente um operador passa-altas Laplaciano, ou até mesmo uma matriz identidade. Esta regularização visa limitar a energia total da imagem ou penalizar energias nas altas frequências, resultando em uma imagem borrada. Desta forma, os pixels ruidosos e as bordas serão removidos no processo.

A regularização de Tikhonov pode ser empregada na crominância como:

$$J_2(\mathbf{x}) = \|\Lambda\mathbf{x}_{Cb}\|_2^2 + \|\Lambda\mathbf{x}_{Cr}\|_2^2, \quad (4.15)$$

onde  $\mathbf{x}_{Cb}$  e  $\mathbf{x}_{Cr}$  correspondem respectivamente a  $Cb$  e  $Cr$ , como definidos na Eq. (4.3).

A derivada parcial de  $J_2(\mathbf{x})$  em relação a um canal de cor pode ser obtida através da expansão

$$\begin{aligned} \frac{\partial J_2(\mathbf{x})}{\partial \mathbf{x}_i} &= \frac{\partial \|\Lambda\mathbf{x}_{Cb}\|_2^2}{\partial \mathbf{x}_i} + \frac{\partial \|\Lambda\mathbf{x}_{Cr}\|_2^2}{\partial \mathbf{x}_i} \\ &= \frac{\partial \|\Lambda\mathbf{x}_{Cb}\|_2^2}{\partial \mathbf{x}_{Cb}} \frac{\partial \mathbf{x}_{Cb}}{\partial \mathbf{x}_i} + \frac{\partial \|\Lambda\mathbf{x}_{Cr}\|_2^2}{\partial \mathbf{x}_{Cr}} \frac{\partial \mathbf{x}_{Cr}}{\partial \mathbf{x}_i}. \end{aligned} \quad (4.16)$$

Utilizando a identidade

$$\frac{\partial \|\Lambda\mathbf{x}_C\|_2^2}{\partial \mathbf{x}_C} = \frac{\mathbf{x}_C^T \Lambda^T \Lambda \mathbf{x}_C}{\partial \mathbf{x}_C} = 2\Lambda^T \Lambda \mathbf{x}_C, \quad (4.17)$$

na Eq. (4.16), temos que a derivada parcial de  $J_2(\mathbf{x})$  em relação a  $\mathbf{x}_i$  é

$$\begin{aligned} \frac{\partial J_2(\mathbf{x})}{\partial \mathbf{x}_i} &= 2\Lambda^T \Lambda \mathbf{x}_{Cb} \frac{\partial \mathbf{x}_{Cb}}{\partial \mathbf{x}_i} + 2\Lambda^T \Lambda \mathbf{x}_{Cr} \frac{\partial \mathbf{x}_{Cr}}{\partial \mathbf{x}_i} \\ &= 2\Lambda^T \Lambda \left( \frac{\partial \mathbf{x}_{Cb}}{\partial \mathbf{x}_i} \mathbf{x}_{Cb} + \frac{\partial \mathbf{x}_{Cr}}{\partial \mathbf{x}_i} \mathbf{x}_{Cr} \right), \end{aligned} \quad (4.18)$$

onde  $\frac{\partial \mathbf{x}_{Cb}}{\partial \mathbf{x}_i}$  e  $\frac{\partial \mathbf{x}_{Cr}}{\partial \mathbf{x}_i}$  são constantes definidas na matriz de conversão da Eq. (4.3).

### 4.2.4 Dependência entre Canais de Cores

O termo relativo à dependência entre os canais de cores tem por finalidade compensar o descasamento de bordas entre os canais de cores. A minimização da norma do

produto vetorial entre quaisquer dois pixels coloridos adjacentes força os diferentes canais a terem semelhantes localização e orientação de bordas [38]. Uma função-custo a ser minimizada é sugerida em [29] levando em consideração as dependências ponto-a-ponto entre as cores. O termo de dependência entre cores proposto é uma função-custo diferenciável descrita como:

$$J_3(\mathbf{x}) = \sum_{l=-1}^1 \sum_{m=-1}^1 \left[ \|\mathbf{x}_R \odot S_x^l S_y^m \mathbf{x}_G - \mathbf{x}_G \odot S_x^l S_y^m \mathbf{x}_R\|_2^2 + \|\mathbf{x}_G \odot S_x^l S_y^m \mathbf{x}_B - \mathbf{x}_B \odot S_x^l S_y^m \mathbf{x}_G\|_2^2 + \|\mathbf{x}_B \odot S_x^l S_y^m \mathbf{x}_R - \mathbf{x}_R \odot S_x^l S_y^m \mathbf{x}_B\|_2^2 \right] \quad (4.19)$$

onde  $\odot$  é um operador de multiplicação ponto-a-ponto.

Consideraremos os termos relativos a cada um dos canais de cor  $R$ ,  $G$  e  $B$  na Eq. (4.19) como, respectivamente  $\underline{L}_{RG}$ ,  $\underline{L}_{GB}$  e  $\underline{L}_{BR}$ . Desta forma, a Eq. (4.19) pode ser simplificada como:

$$J_3(\mathbf{x}) = \sum_{l=-1}^1 \sum_{m=-1}^1 [\underline{L}_{RG} + \underline{L}_{GB} + \underline{L}_{BR}], \quad (4.20)$$

O multiplicador ponto-a-ponto  $\odot$  permite a aplicação da propriedade comutativa. Para o termo  $\underline{L}_{RG}$  temos:

$$\|\mathbf{x}_R \odot S_x^l S_y^m \mathbf{x}_G - \mathbf{x}_G \odot S_x^l S_y^m \mathbf{x}_R\|_2^2 = \|S_x^l S_y^m \mathbf{x}_G \odot \mathbf{x}_R - \mathbf{x}_G \odot S_x^l S_y^m \mathbf{x}_R\|_2^2. \quad (4.21)$$

Desta forma,  $\underline{L}_{RG} = \underline{L}_{GR}$ ,  $\underline{L}_{GB} = \underline{L}_{BG}$  e  $\underline{L}_{BR} = \underline{L}_{RB}$ .

Podemos substituir o operador  $\odot$  por um produto escalar diferenciado. O vetor  $\mathbf{x}_G$  pode ser interpretado como uma matriz diagonal  $\mathbf{X}_G$  e o produto  $S_x^l S_y^m \mathbf{x}_G$  como uma matriz  $\mathbf{X}_G^{l,m}$ , que é a forma diagonal do vetor  $\mathbf{x}_G$  deslocado por  $l, m$  pixels nas direções horizontal e vertical. Desta forma, o termo  $\underline{L}_{RG}$  pode ser reescrito como:

$$\underline{L}_{RG} = \|\mathbf{X}_G^{l,m} \mathbf{x}_R - \mathbf{X}_G S_x^l S_y^m \mathbf{x}_R\|_2^2, \quad (4.22)$$

onde  $\mathbf{X}_G^{l,m}$  e  $\mathbf{X}_G$  são matrizes simétricas. Consequentemente, os demais termos podem ser escritos como:

$$\underline{L}_{GB} = \|\mathbf{X}_B^{l,m} \mathbf{x}_G - \mathbf{X}_B S_x^l S_y^m \mathbf{x}_G\|_2^2 \quad (4.23)$$

e

$$\underline{L}_{BG} = \|\mathbf{X}_R^{l,m} \mathbf{x}_B - \mathbf{X}_R S_x^l S_y^m \mathbf{x}_B\|_2^2. \quad (4.24)$$

A derivada de  $J_3(\mathbf{x})$  relativa a cada um dos canais de cor  $R$ ,  $G$  e  $B$  será repre-

sentada pelo somatório das derivadas parciais dos termos que contêm o canal de cor considerado. Para o caso da cor vermelha, temos:

$$\frac{\partial J_3(\mathbf{x})}{\partial \mathbf{x}_R} = \sum_{l=-1}^1 \sum_{m=-1}^1 \left[ \frac{\partial \underline{L}_{RG}}{\partial \mathbf{x}_R} + \frac{\partial \underline{L}_{BR}}{\partial \mathbf{x}_R} \right]. \quad (4.25)$$

Manipulando o termo  $\underline{L}_{BR}$  podemos reescrevê-lo em função de  $\mathbf{x}_R$

$$\begin{aligned} \underline{L}_{BR} &= \|\mathbf{x}_B \odot S_x^l S_y^m \mathbf{x}_R - \mathbf{x}_R \odot S_x^l S_y^m \mathbf{x}_B\|_2^2 \\ &= \|\mathbf{x}_R \odot S_x^l S_y^m \mathbf{x}_B + \mathbf{x}_B \odot S_x^l S_y^m \mathbf{x}_R\|_2^2 \\ &= \|\mathbf{X}_B^{l,m} \mathbf{x}_R - \mathbf{X}_B S_x^l S_y^m \mathbf{x}_R\|_2^2, \end{aligned} \quad (4.26)$$

o que nos permite derivar este termo em relação a  $\mathbf{x}_R$  de maneira semelhante ao termo  $\underline{L}_{RG}$ . Logo, a Eq. (4.25) pode ser expandida em

$$\begin{aligned} \frac{\partial J_3(\mathbf{x})}{\partial \mathbf{x}_R} &= \sum_{l=-1}^1 \sum_{m=-1}^1 \left[ 2(\mathbf{X}_G^{l,m} - \mathbf{X}_G S_x^l S_y^m)(\mathbf{X}_G^{l,m} \mathbf{x}_R - \mathbf{X}_G S_x^l S_y^m \mathbf{x}_R) \right. \\ &\quad \left. + 2(\mathbf{X}_B^{l,m} - \mathbf{X}_B S_x^l S_y^m)(\mathbf{X}_B^{l,m} \mathbf{x}_R - \mathbf{X}_B S_x^l S_y^m \mathbf{x}_R) \right]. \end{aligned} \quad (4.27)$$

Procedimento semelhante é adotado para as demais cores.

## 4.2.5 Função-Custo Total

A função-custo total é uma soma ponderada das funções-custo apresentadas anteriormente. O minimizante da função-custo total é descrito por:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} [J_0(\mathbf{x}) + \lambda' J_1(\mathbf{x}) + \lambda'' J_2(\mathbf{x}) + \lambda''' J_3(\mathbf{x})], \quad (4.28)$$

onde  $\lambda'$ ,  $\lambda''$  e  $\lambda'''$  são as constantes escalares ponderadoras das funções-custo  $J_1(\mathbf{x})$ ,  $J_2(\mathbf{x})$  e  $J_3(\mathbf{x})$ , respectivamente.

Uma otimização baseada no algoritmo *steepest descent* pode ser aplicada para minimizar a função-custo. A derivada da função ocorre em referência a cada uma das três cores primárias, assumindo as outras duas como constantes. Desta forma, três derivadas parciais serão calculadas para cada termo da função-custo total. Os gradientes relativos aos canais de cores são definidos como:

$$\nabla_{\hat{\mathbf{x}}_i} = \frac{\partial J_0(\mathbf{x})}{\partial \mathbf{x}_i} + \lambda' \frac{\partial J_1(\mathbf{x})}{\partial \mathbf{x}_i} + \lambda'' \frac{\partial J_2(\mathbf{x})}{\partial \mathbf{x}_i} + \lambda''' \frac{\partial J_3(\mathbf{x})}{\partial \mathbf{x}_i}. \quad (4.29)$$

A partir de um chute inicial para  $\hat{\mathbf{x}}^0$ , a imagem HR estimada será calculada iterativamente, com as atualizações realizadas para cada um dos canais de cor da

imagem como:

$$\hat{\mathbf{x}}_i^{n+1} = \hat{\mathbf{x}}_i^n - \beta \nabla \hat{\mathbf{x}}_i^n, \quad i = R, G, B; \quad (4.30)$$

onde  $\beta$  é um escalar que define o passo das iterações. A imagem HR estimada de índice  $n + 1$  será utilizada na iteração seguinte na local da variável  $\hat{\mathbf{x}}^n$ , de forma a calcular a iteração  $n + 2$ .

Para realizar uma implementação mais eficiente, rápida e com menor consumo de memória, as matrizes  $\mathbf{H}$ ,  $\mathbf{\Lambda}$ ,  $\mathbf{D}$ ,  $S_x^l$  e  $S_y^m$ , e suas transpostas devem ser aplicadas na forma de operadores de borramento, filtro passa-altas, subamostragem e deslocamento horizontal e vertical, respectivamente. A aplicação direta dos efeitos destas matrizes como uma sequência de operadores nas imagens poupa o trabalho de construí-las explicitamente e reduz o número de operações necessárias para o cálculo das operações. As operações matriciais para realizar os processos relativos às matrizes  $\mathbf{H}$ ,  $\mathbf{\Lambda}$ ,  $\mathbf{D}$ ,  $S_x^l$  e  $S_y^m$  necessitam de mais operações aritméticas do que a aplicação de filtros espaciais nas imagens.

O resultado da super-resolução do cenário apresentado na Fig. 4.2 utilizando o algoritmo SR *multiframe* adaptativo é apresentado na Fig. 4.7. Os parâmetros utilizados para a estimação da imagem HR, que minimizam a diferença entre a imagem desejada e a imagem estimada, são:  $\alpha = 0,9$ ,  $\lambda' = 0,01$ ,  $\lambda'' = 0,015$ ,  $\lambda''' = 0,2$  e  $\beta = 0,1$ . Para o chute inicial foi utilizada a interpolação bilinear do *frame* LR de referência para as estimações de movimento entre as imagens LR. Foram necessárias 20 iterações para que, com os parâmetros escolhidos, a diferença da Eq. (4.7) convergisse para seu valor mínimo.



Figura 4.7: Imagem HR gerada por algoritmo de SR *multiframe* adaptativo.

# Capítulo 5

## Resultados

Neste capítulo, comparamos o desempenho dos algoritmos de super-resolução de imagens coloridas apresentados no Capítulo 4. Em todos os casos, utilizamos simulações controladas com o objetivo de realizar diferentes análises. Nestes experimentos, criamos sequências de quadros de baixa resolução utilizando uma imagem de alta resolução, como proposto no modelo de aquisição da Eq. (2.1) e da Fig. 2.2. Foram utilizadas 30 imagens HR originais para gerar os cenários analisados.

Inicialmente, a imagem HR é deslocada com base no movimento relativo estipulado. Então, com o objetivo de simular o efeito de PSF das câmeras, a imagem é convoluída com um filtro passa-baixas de formato Gaussiano e desvio-padrão unitário, utilizando uma janela de tamanho  $4 \times 4$  para realizar a filtragem espacial. A imagem resultante é então subamostrada em cada uma das direções pelo fator de escala que utilizamos na super-resolução. A mesma abordagem é aplicada para os diferentes vetores de movimento, nas direções vertical e horizontal, de forma a produzir o conjunto de imagens LR que será utilizado no processo de super-resolução.

Os métodos são avaliados quanto à capacidade de reduzir o efeito do ruído aditivo e à qualidade da super-resolução com diferentes números de imagens LR, ou seja, uma grade HR esparsa. Três métricas são utilizadas para classificar os resultados: a PSNR, a S-CIELAB e a SSIM. As métricas de avaliação e os resultados dos métodos são apresentados e discutidos nas próximas seções.

### 5.1 Métricas de Avaliação

Métodos objetivos para avaliar a qualidade de imagens, tradicionalmente, tentam quantificar os erros visíveis entre uma imagem degradada e uma imagem de referência. Tais métodos podem utilizar propriedades conhecidas do sistema visual humano para gerar índices quantitativos. Em todos os casos, a imagem estimada em alta resolução será comparada com a imagem HR original utilizada para gerar o conjunto de imagens de baixa resolução utilizadas para realizar a super-resolução.

### 5.1.1 PSNR

Um dos principais métodos de avaliação da qualidade de vídeos e imagens é a PSNR (*Peak Signal-to-Noise Ratio*), correspondente à razão entre a potência máxima possível de um sinal e a potência do ruído. A unidade utilizada para representá-la é o decibel (dB). No contexto de compressão de imagens, a PSNR é a relação entre a entrada e a saída de um processo com perdas, e avalia a introdução de ruído na imagem pela distorção dos valores dos pixels em relação à imagem original.

Existem diferentes abordagens para calcular a PSNR de uma imagem colorida. Como o olho humano é mais sensível à informação de luminância a abordagem mais comum é calcular somente a PSNR da luminância para representar a PSNR da imagem colorida. A PSNR é definida através do erro médio quadrático (MSE), que para uma imagem colorida com dimensões  $m \times n$  é calculado por [39]:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[ I_Y(i, j) - \hat{I}_Y(i, j) \right]^2, \quad (5.1)$$

onde  $I_Y$  é a luminância da imagem original e  $\hat{I}_Y$  a luminância da imagem HR estimada. A PSNR é definida como:

$$\text{PSNR} = 20 \log_{10} \left( \frac{\text{MAX}_I}{\text{MSE}} \right), \quad (5.2)$$

sendo  $\text{MAX}_I$  o valor máximo possível para um pixel na imagem. Como o MSE é uma métrica da diferença entre duas imagens, valores pequenos indicam forte proximidade entre a imagem original e a degradada. Por outro lado, a PSNR é calculada com base no inverso do MSE, por isso valores maiores indicam maior proximidade entre as imagens. A PSNR entre duas imagens idênticas é infinita.

### 5.1.2 S-CIELAB

A métrica S-CIELAB é uma medida de fidelidade na reprodução de cores de imagens digitais, que visa determinar a acurácia da reprodução de uma cor em relação à original, quando visualizada por um observador humano [40]. É uma extensão espacial do padrão CIELAB  $\Delta E^*$ , definido pela *International Commission on Illumination* (CIE), para medir a distância  $\Delta E_{ab}^*$  entre duas cores. Esta métrica é adequada para regiões uniformes de cor, mas não para as variações ao longo das imagens, devido à mudança de sensibilidade de cor em função do espaço. A extensão S-CIELAB inclui uma etapa de processamento espacial, antes do cálculo de CIELAB  $\Delta E^*$ , de modo que os resultados correspondem a uma melhor percepção de diferença de cor ao olho humano.

Para realizar esta medida nos experimentos, utilizamos o código de MATLAB, com parâmetros-padrão, disponível em [41].

### 5.1.3 SSIM

O Índice de Semelhança Estrutural (SSIM - *Structural SIMilarity*) tem por objetivo medir a semelhança entre duas imagens, sob a hipótese de que a percepção visual humana é altamente adaptada para a extração de informações sobre a estrutura de uma cena [42]. Este método relaciona a qualidade das imagens através da degradação de informação estrutural, considerando uma das imagens com qualidade perfeita.

Como a SSIM é definida para uma imagem monocromática, assim como na medida PSNR, a SSIM de imagens coloridas é calculada apenas para a luminância. Uma imagem sem degradação em relação à referência possui valor SSIM unitário; o aumento da distorção entre a imagem avaliada e a referência reduz o índice SSIM. Para realizar esta medida, utilizamos o código de MATLAB fornecido em [43]. O algoritmo calcula medidas de similaridade através de comparações da luminância, do contraste e da estrutura, que combinadas fornecem o grau de similaridade

## 5.2 Parâmetros Utilizados

Os quatro métodos apresentados no capítulo anterior apresentam parâmetros, que permitem configurações de acordo com o cenário. As buscas para encontrar os valores dos parâmetros que otimizam os resultados das três variações da SR AWF e da SR *multiframe* adaptativo são exaustivas e os resultados encontrados variam para cada imagem utilizada e para cada nível de ruído apresentado nas imagens LR.

Abaixo são apresentados os valores dos parâmetros escolhidos para a realização de todos os experimentos deste capítulo, além de uma breve justificativa para suas escolhas. Tais valores foram selecionados por apresentarem resultados satisfatórios para os diferentes casos abordados. Em todos os casos o fator de escala da super-resolução foi  $r = 3$ .

### AWF em RGB

$$L_x = L_y = 3, W_x = W_y = 9, D_x = D_y = 3 \text{ e } \rho = 0,75.$$

Os parâmetros  $L_x$  e  $L_y$  correspondem ao fator de escala da super-resolução. Um estudo sobre o impacto dos parâmetros  $W_x$ ,  $W_y$ ,  $D_x$  e  $D_y$  na SR AWF é apresentado em [17]. O aumento da janela de observação, com dimensões  $W_x \times W_y$ , aumenta o número de operações necessárias para realizar a SR AWF. Os parâmetros escolhidos permitem uma baixa complexidade computacional sem perda de qualidade no resultado da super-resolução.

A opção pelo valor de  $\rho$  é discutida na Seção 3.2 deste trabalho.

### **AWF com compressão de crominância**

$$L_x = L_y = 3, W_x = W_y = 9, D_x = D_y = 3,$$

$$\rho_Y = 0,75, \rho_{Cb} = 0,75 \text{ e } \rho_{Cr} = 0,75.$$

Os valores de  $L_x, L_y, W_x, W_y, D_x, D_y$  e dos três  $\rho$ s são semelhantes aos da SR AWF em RGB.

### **AWF com Wavelet na luminância**

$$L_x = L_y = 3, W_x = W_y = 9, D_x = D_y = 3,$$

$$\rho_{Y_{LL}} = 0,6, \rho_{Y_{LH}} = 0,4, \rho_{Y_{HL}} = 0,4, \rho_{Y_{HH}} = 0,4,$$

$$\rho_{Cb} = 0,75 \text{ e } \rho_{Cr} = 0,75.$$

Os valores de  $L_x, L_y, W_x, W_y, D_x, D_y, \rho_{Cb}$  e  $\rho_{Cr}$  são semelhantes aos da SR AWF em RGB. Os diferentes valores de  $\rho$  utilizados na luminância são discutidos na Seção 3.3.1 deste trabalho.

### **Multiframe adaptativo**

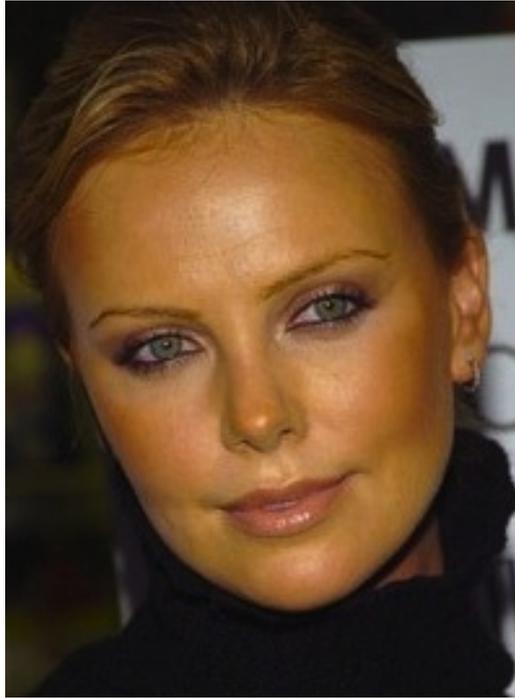
$$\beta = 0,1, \alpha = 0,9, \lambda' = 0,01, \lambda'' = 0,015, \lambda''' = 0,2, n = 20 \text{ iterações.}$$

O parâmetro  $\beta$  corresponde ao passado das iterações e foi escolhido de forma a convergir a função-custo para um valor mínimo com o menor número de iterações. O valor de  $\alpha$  é sugerido em [28]. Os valores de  $\lambda', \lambda''$  e  $\lambda'''$  visam nivelar os valores das funções-custos  $J_1, J_2$ , e  $J_3$ , respectivamente, à ao nível do termo de fidelidade  $J_0$ .

## **5.3 Efeito do Ruído**

Quatro cenários foram criados para a realização dos experimentos. No primeiro caso não houve adição de ruído às imagens LR. Em seguida, adicionamos ruído Gaussiano para obter imagens de baixa resolução com razões sinal-ruído (SNR - *Signal-to-Noise Ratio*) de aproximadamente 40 dB, 30 dB e 25 dB, respectivamente. Ao todo, 30 conjuntos de imagens LR receberam a adição dos diferentes níveis de ruído. Duas imagens foram escolhidas para ilustrar os resultados encontrados. A imagem *Charlize* vinha sendo usada em projetos anteriores. A imagem *Lighthouse* foi incluída porque ter conteúdo significativo de cores em alta frequência.

Na Fig. 5.2 são apresentados exemplos da imagem *Charlize* em baixa resolução, com diferentes intensidades de ruído. Outra cena analisada é a da imagem *Lighthouse*, cujas amostras em baixa resolução com os diferentes níveis de ruído utilizados na análise são visualizadas na Fig. 5.3. Os resultados e as análises da super-resolução desses dois conjuntos de imagens são apresentados a seguir. Após a discussão sobre as duas imagens selecionadas são apresentados os resultados mínimo, máximo, médio e o desvio-padrão para cada uma das métricas de medida aplicada aos resultados da super-resolução das 30 imagens do conjunto analisado.



(a) *Charlize*



(b) *Lighthouse*

Figura 5.1: Imagens HR desejadas.



(a) Sem ruído



(b) SNR = 40 dB



(c) SNR = 30 dB



(d) SNR = 25 dB

Figura 5.2: Imagem *Charlize* LR com diferentes níveis de ruído.

As imagens HR recuperadas foram agrupadas com base na intensidade do ruído aditivo, para possibilitar uma melhor comparação visual entre os diferentes resultados apresentados pelos métodos descritos no capítulo anterior. A Fig. 5.4 apresenta os resultados da super-resolução sem ruído da imagem *Charlize*, a Fig. 5.5 com SNR de 40 dB, e as Figs. 5.6 e 5.7 com SNR de 30 dB e 25 dB, respectivamente. Os resultados da super-resolução da imagem *Lighthouse* são apresentados nas Figs. 5.8, 5.9, 5.10 e 5.11, seguindo a ordem de SNR das figuras da imagem *Charlize*.

Semelhantemente à organização das imagens estimadas, para cada conjunto de reconstrução HR, tabelas com os resultados das métricas PSNR, S-CIELAB e SSIM são apresentadas. Para a imagem *Charlize*, a Tabela 5.1 apresenta os resultados para as imagens da Fig. 5.4; os resultados para as Figs. 5.5, 5.6 e 5.7 são apresentados respectivamente nas Tabelas 5.2, 5.3 e 5.4. As Tabelas 5.5, 5.6, 5.7 e 5.8 apresentam

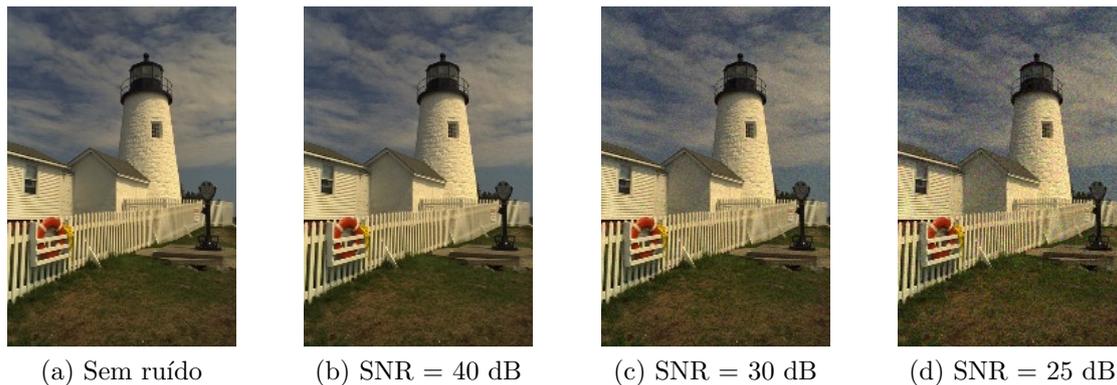


Figura 5.3: Imagem *Lighthouse* LR com diferentes níveis de ruído.

os resultados relativos à super-resolução da imagem *Lighthouse*, correspondendo às Figs. 5.8, 5.9, 5.10 e 5.11, respectivamente.

Levando em consideração os valores das três métricas e a análise visual das imagens HR de *Charlize*, para os diferentes níveis de ruído, observamos que o método SR AWF em RGB, de maneira geral, apresenta os melhores resultados. O modelo de SR AWF com Wavelet na luminância apresentou resultados bem próximos, apesar de uma PSNR e SSIM ligeiramente inferiores, obteve uma imagem HR mais suave e menos ruidosa no cenário com maior intensidade de ruído. Isto se deve ao fato do modelo com Wavelets funcionar como um filtro passa-baixas, reduzindo significativamente as componentes de alta frequência, em especial o ruído. Os resultados apresentados nas Figs. 5.6(c) e 5.7(c) são imagens mais limpas em relação à imagem obtida com a filtragem de Wiener aplicada diretamente aos canais de cor nas Figs. 5.6(a) e 5.7(a).

Observando apenas as métricas de qualidade, notamos que a SR AWF com compressão de crominância apresentou resultados piores do que o caso sem compressão. Porém, os resultados visuais não indicam diferenças significativas entre os dois métodos. Ao comparar a Fig. 5.4(a) com SR AWF em RGB e a Fig. 5.4(b) com SR AWF com compressão de crominância, observamos pequenas apenas diferenças em regiões de altas frequências de cor, como a iris e a maquiagem sob os olhos. Essas regiões aparecem mais borradas no caso com compressão de crominância.

Este borramento pode ser observado na Fig. 5.7(b), onde o ruído aditivo das imagens LR é mais intenso. Neste caso o método SR AWF com compressão de crominância aplicado à imagem *Charlize* apresenta regiões mais limpas de ruído, como a testa e o rosto, em relação à Fig. 5.7(a) com SR AWF em RGB.

O motivo da diferença entre os resultados das métricas e da análise visual pode ser explicado pelo fato da crominância ter pouco impacto no sistema visual. O observador das imagens pode não perceber as diferenças, mas a análise dos valores pixel a pixel indica uma diferença matematicamente relevante.

A técnica de SR *multiframe* adaptativo apresentou os piores resultados para todos os cenários analisados, levando em consideração as métricas de qualidade e o resultado visual. Como observado na Fig. 5.7(d), o método não se mostrou eficiente para filtrar o ruído aditivo das imagens LR. Na Fig. 5.4(d) sem ruído e na Fig. 5.5(d) com ruído de baixa intensidade as imagens apresentaram bordas serrilhadas.

Tabela 5.1: Análise da super-resolução da imagem *Charlize* sem ruído aditivo.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>70,5</b>	<b>0</b>	<b>0,999</b>
AWF com compressão de crominância	46,9	3175	0,919
AWF com Wavelet na luminância	<b>70,1</b>	<b>0</b>	<b>0,999</b>
<i>Multiframe</i> adaptativo	44,1	4006	0,877

Tabela 5.2: Análise da super-resolução da imagem *Charlize* com SNR de 40 dB.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>63,8</b>	<b>0</b>	<b>0,995</b>
AWF com compressão de crominância	46,9	3207	0,918
AWF com Wavelet na luminância	589,0	<b>3</b>	<b>0,988</b>
<i>Multiframe</i> adaptativo	44,0	4042	0,875

Tabela 5.3: Análise da super-resolução da imagem *Charlize* com SNR de 30 dB.

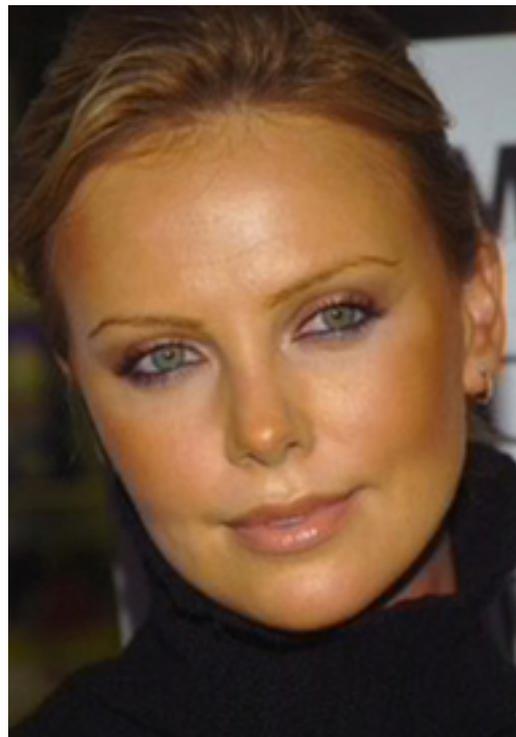
Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>56,5</b>	<b>23</b>	<b>0,979</b>
AWF com compressão de crominância	46,7	3419	0,911
AWF com Wavelet na luminância	52,5	61	<b>0,959</b>
<i>Multiframe</i> adaptativo	43,7	4451	0,863

Tabela 5.4: Análise da super-resolução da imagem *Charlize* com SNR de 25 dB.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>53,4</b>	<b>172</b>	<b>0,961</b>
AWF com compressão de crominância	46,4	3582	0,899
AWF com Wavelet na luminância	50,3	209	<b>0,935</b>
<i>Multiframe</i> adaptativo	43,0	5817	0,840



(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância



(d) SR *multiframe* adaptativo

Figura 5.4: Super-resolução da imagem *Charlize* sem ruído aditivo.



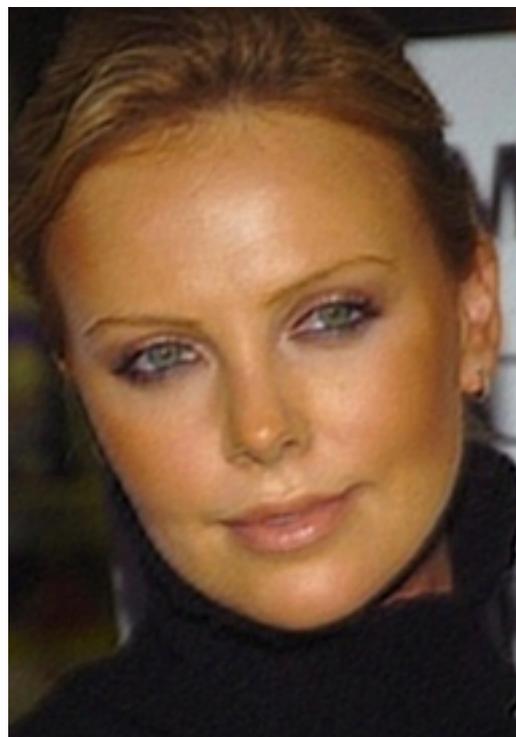
(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância

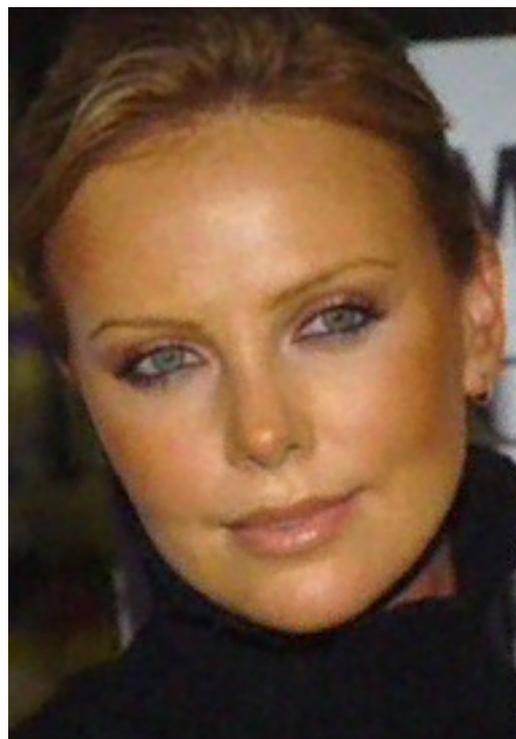


(d) SR *multiframe* adaptativo

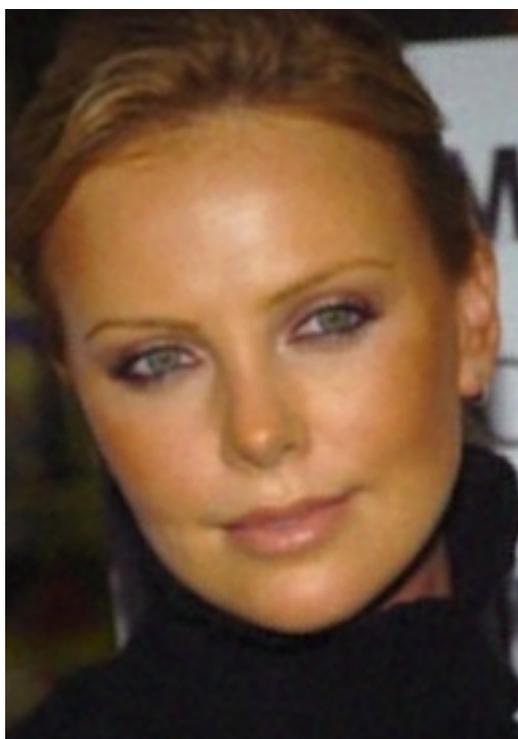
Figura 5.5: Super-resolução da imagem *Charlize* com SNR de 40 dB.



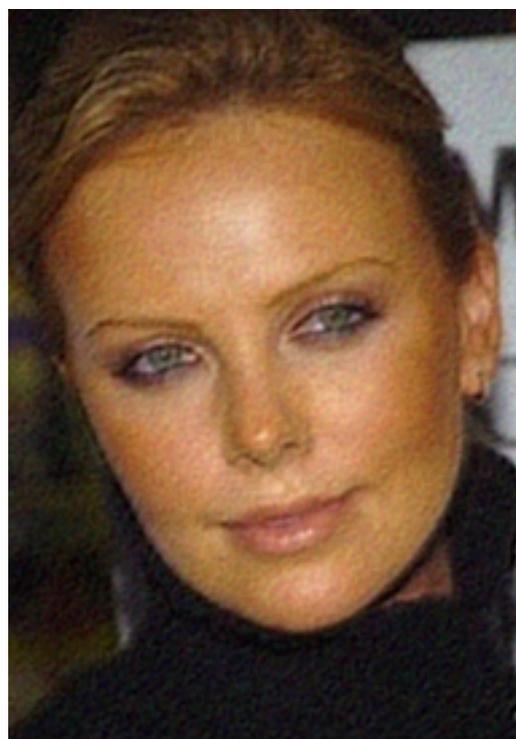
(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade

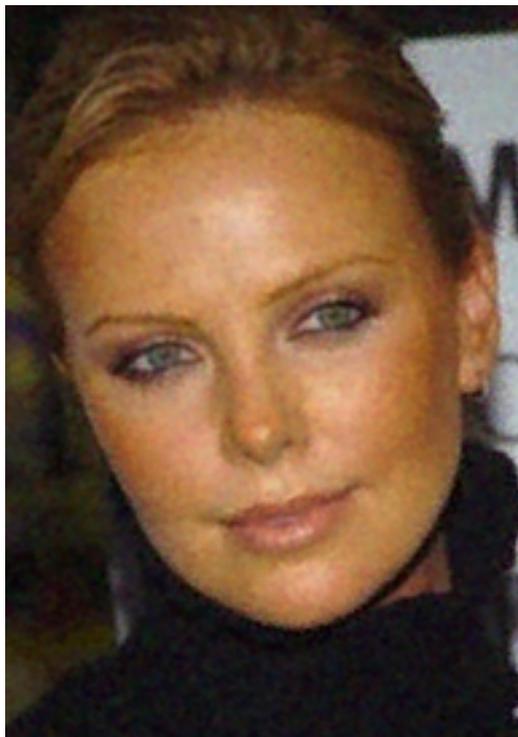


(c) SR AWF com Wavelet na luminância

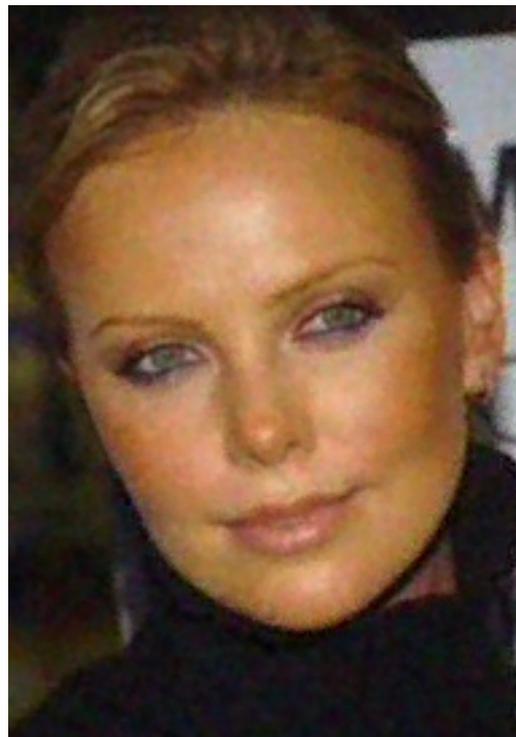


(d) SR *multiframe* adaptativo

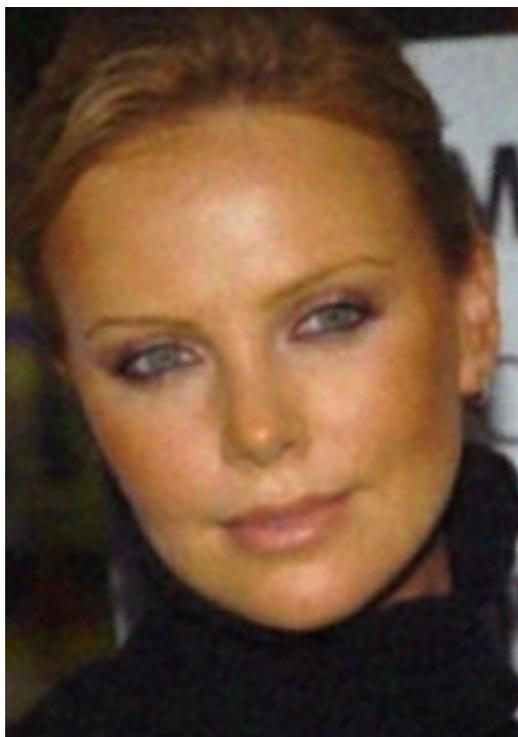
Figura 5.6: Super-resolução da imagem *Charlize* com SNR de 30 dB.



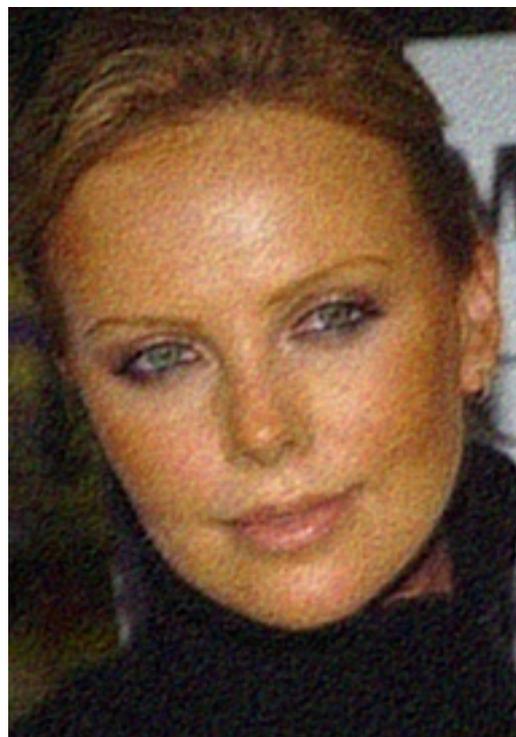
(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância



(d) SR *multiframe* adaptativo

Figura 5.7: Super-resolução da imagem *Charlize* com SNR de 25 dB.



(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância



(d) SR *multiframe* adaptativo

Figura 5.8: Super-resolução da imagem *Lighthouse* sem ruído aditivo.



(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade

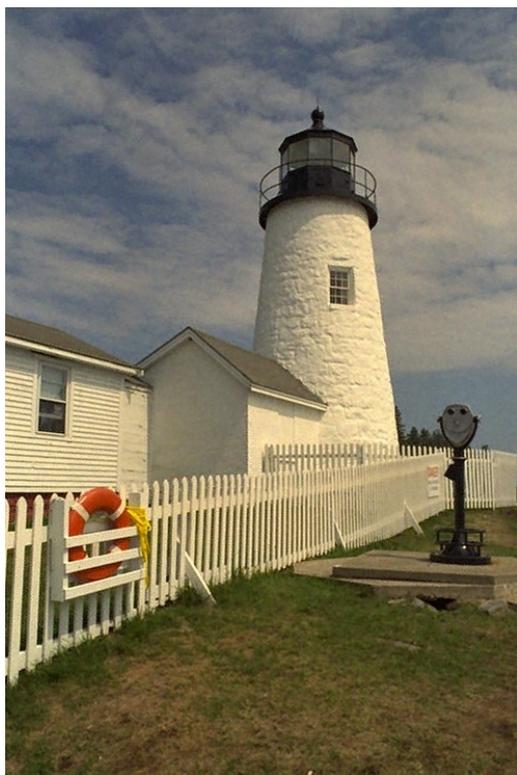


(c) SR AWF com Wavelet na luminância

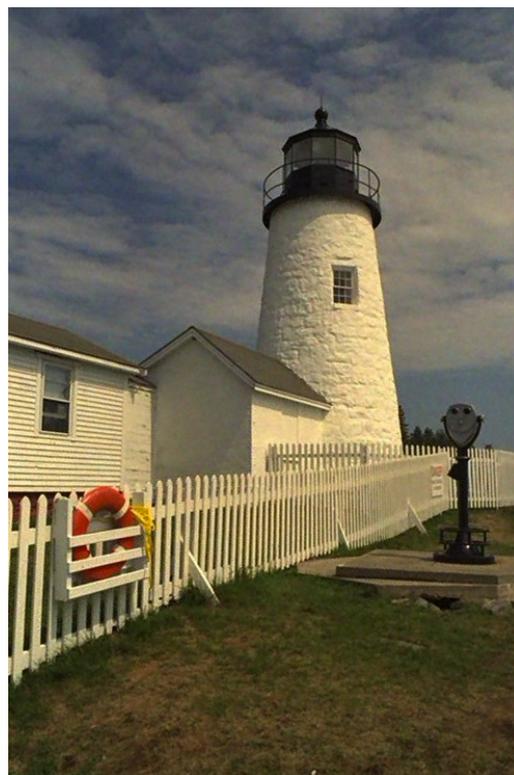


(d) SR *multiframe* adaptativo

Figura 5.9: Super-resolução da imagem *Lighthouse* com SNR de 40 dB.



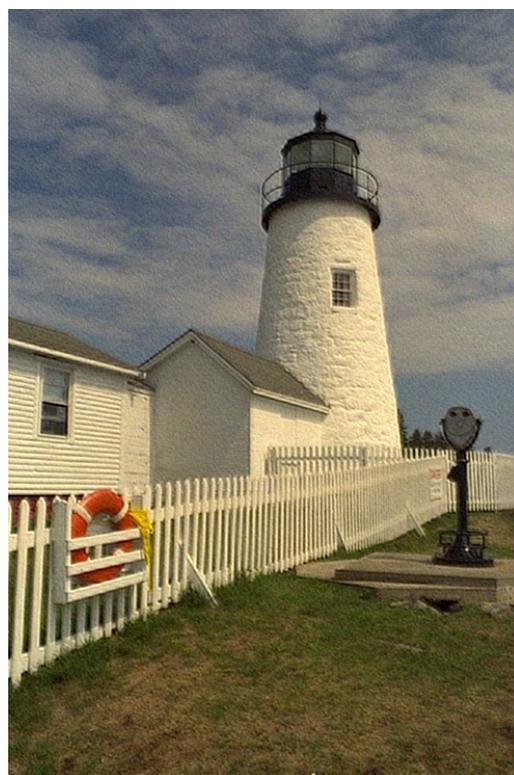
(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância



(d) SR *multiframe* adaptativo

Figura 5.10: Super-resolução da imagem *Lighthouse* com SNR de 30 dB.



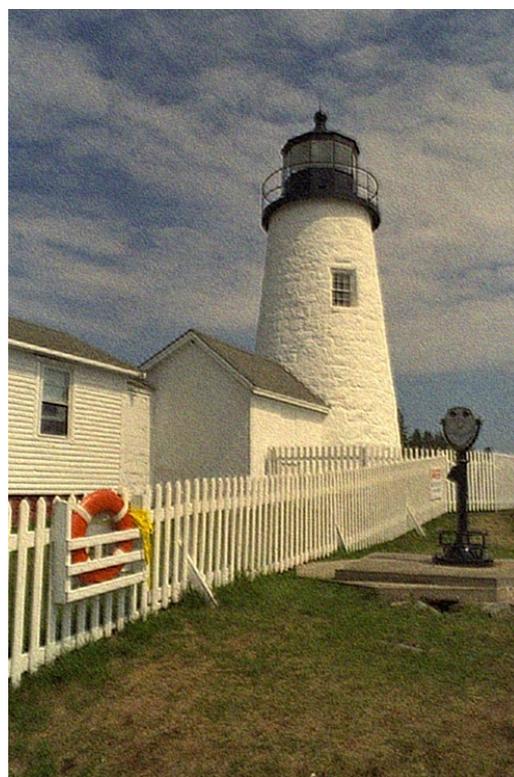
(a) SR AWF em RGB



(b) SR AWF com compressão de cromaticidade



(c) SR AWF com Wavelet na luminância



(d) SR *multiframe* adaptativo

Figura 5.11: Super-resolução da imagem *Lighthouse* com SNR de 25 dB.

Tabela 5.5: Análise da super-resolução da imagem *Lighthouse* sem ruído aditivo.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>79,4</b>	<b>1</b>	<b>0,999</b>
AWF com compressão de crominância	52,6	3337	0,929
AWF com Wavelet na luminância	<b>78,9</b>	<b>1</b>	<b>0,999</b>
<i>Multiframe</i> adaptativo	49,8	2637	0,894

Tabela 5.6: Análise da super-resolução da imagem *Lighthouse* com SNR de 40 dB.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>71,3</b>	<b>1</b>	<b>0,990</b>
AWF com compressão de crominância	52,5	3229	0,926
AWF com Wavelet na luminância	66,6	<b>1</b>	<b>0,981</b>
<i>Multiframe</i> adaptativo	49,8	2663	0,887

Tabela 5.7: Análise da super-resolução da imagem *Lighthouse* com SNR de 30 dB.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF em RGB	<b>63,6</b>	<b>49</b>	<b>0,961</b>
AWF com compressão de crominância	52,1	3391	0,909
AWF com Wavelet na luminância	<b>60,2</b>	52	<b>0,948</b>
<i>Multiframe</i> adaptativo	49,1	2959	0,839

Tabela 5.8: Análise da super-resolução da imagem *Lighthouse* com SNR de 25 dB.

Método de super-resolução	PSNR (dB)	S-CIELAB	SSIM
AWF RGB	<b>59,8</b>	<b>194</b>	<b>0,925</b>
AWF com compressão de crominância	51,7	3700	0,885
AWF com Wavelet na luminância	<b>57,6</b>	144	<b>0,920</b>
<i>Multiframe</i> adaptativo	47,8	3627	0,764

Os resultados para a imagem *Lighthouse* condizem com os da imagem *Charlize*. No caso com maior intensidade de ruído, os resultados das métricas de qualidade das imagens da Fig. 5.11, apresentados na Tabela 5.8, mostram que o método de SR AWF com Wavelet na luminância possui SSIM semelhante à SR AWF em RGB.

A imagem *Lighthouse* é um bom exemplo de super-resolução. Em seus quadros de baixa resolução podemos observar *aliasing* em diversos pontos da imagem,

especialmente na região da cerca, onde há alternância de intensidade luminosa, resultando em altas frequências. Os métodos de SR realizaram a reconstrução dessas áreas, possibilitando correta identificação das cores nas regiões de alta frequência e o correto posicionamento das bordas existentes, independente do nível de ruído das imagens LR.

As análises de PSNR da super-resolução dos 30 conjuntos de imagens processadas podem ser visualizadas nas Tabelas 5.9, 5.10, 5.11 e 5.12. Cada uma dessas tabelas apresenta os resultados para os diferentes níveis de ruídos considerados nesta seção. As tabelas apresentam os valores mínimo, máximo e médio de PSNR da super-resolução do conjunto de imagens analisadas, além do desvio-padrão e do resultado da PNSR das imagens *Charlize* e *Lighthouse*, que foram discutidas acima.

Tabela 5.9: Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com sem ruído.

Método	Mínimo	Máximo	Média	Std <sup>(*)</sup>	<i>Charlize</i>	<i>Lighthouse</i>
#1	68,7	95,0	<b>85,3</b>	7,4	<b>70,5</b>	<b>79,4</b>
#2	43,6	59,1	59,1	6,9	46,9	52,6
#3	68,4	93,5	<b>84,5</b>	7,2	<b>70,1</b>	<b>78,9</b>
#4	40,2	65,4	56,3	7,2	44,1	49,8

#1 - SR AWF em RGB

#2 - SR AWF com compressão de crominância

#3 - SR AWF com Wavelet na luminância

#4 - SR *Multiframe* adaptativo

(\*) Desvio-padrão

Tabela 5.10: Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB.

Método	Mínimo	Máximo	Média	Std	<i>Charlize</i>	<i>Lighthouse</i>
#1	63,7	82,4	<b>74,7</b>	5,0	<b>56,5</b>	<b>63,2</b>
#2	43,6	67,6	59,1	6,3	46,9	52,1
#3	58,2	79,9	<b>70,9</b>	6,1	<b>59,0</b>	<b>60,2</b>
#4	40,1	64,8	56,0	6,2	44,0	49,8

Os resultados das Tabelas 5.9, 5.10, 5.11 e 5.12 mostram que a média da PSNR das imagens HR resultantes da super-resolução possui comportamento parecido com as estimativas das duas imagens analisadas anteriormente. O método de SR AWF em RGB apresentou os melhores resultados médio de PSNR, porém o método de SR AWF com Wavelet na luminância apresentou valores de PSNR próximos, em especial no caso de maior ruído na Tabela 5.12.

Tabela 5.11: Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 30 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	55,8	74,2	<b>66,8</b>	5,1	<b>56,5</b>	<b>63,2</b>
#2	43,3	65,5	57,7	6,3	46,7	52,1
#3	50,8	72,4	<b>64,1</b>	6,1	<b>52,5</b>	<b>60,2</b>
#4	39,9	61,4	54,1	6,2	43,7	49,1

Tabela 5.12: Análise da PSNR (dB) da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	52,2	69,3	<b>61,6</b>	4,9	<b>53,4</b>	<b>59,8</b>
#2	43,1	63,8	56,7	5,7	46,4	51,7
#3	48,4	68,7	<b>61,2</b>	5,9	<b>50,3</b>	<b>57,6</b>
#4	39,5	57,3	51,4	5,1	43,0	47,8

As Tabelas 5.13, 5.14, 5.15 e 5.16 apresentam as análises da métrica SSIM dos resultados da super-resolução das 30 imagens para os quatro cenários com variação da intensidade de ruído nas imagens LR.

Tabela 5.13: Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com sem ruído.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0,978	0,999	<b>0,998</b>	0,004	<b>0,999</b>	<b>0,99</b>
#2	0,846	0,978	0,935	0,029	0,919	0,929
#3	0,978	0,999	<b>0,998</b>	0,004	<b>0,999</b>	<b>0,999</b>
#4	0,757	0,960	0,904	0,048	0,877	0,894

Tabela 5.14: Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0,942	0,996	<b>0,988</b>	0,011	<b>0,995</b>	<b>0,990</b>
#2	0,832	0,974	0,931	0,030	0,198	0,926
#3	0,918	0,994	<b>0,980</b>	0,014	<b>0,988</b>	<b>0,981</b>
#4	0,744	0,960	0,896	0,047	0,875	0,887

O comportamento da SSIM é coerente com a PSNR, tendo o método SR AWF em RGB apresentado os melhores resultados médios nas quatro tabelas. Os resultados das médias também estão de acordo com os valores de SSIM calculados para as imagens *Charlize* e *Lighthouse*. As médias de SSIM do método de SR AWF com Wavelet na luminância apresentou resultados próximos aos resultados da SR AWF em RGB, sendo igual no caso sem ruído e praticamente igual para o caso mais ruidoso, com SNR de 25 dB.

Tabela 5.15: Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 30 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0,882	0,980	<b>0,960</b>	0,019	<b>0,979</b>	<b>0,961</b>
#2	0,800	0,957	0,913	0,032	0,911	0,909
#3	0,858	0,977	<b>0,948</b>	0,022	<b>0,959</b>	<b>0,948</b>
#4	0,664	0,906	0,849	0,047	0,863	0,839

Tabela 5.16: Análise da SSIM da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0,831	0,961	<b>0,926</b>	0,026	<b>0,961</b>	<b>0,925</b>
#2	0,769	0,934	0,890	0,033	0,899	0,885
#3	0,821	0,955	<b>0,920</b>	0,025	<b>0,935</b>	<b>0,920</b>
#4	0,561	0,851	0,778	0,058	0,840	0,764

As análises da métrica S-CIELAB são apresentadas nas Tabelas 5.17, 5.18, 5.19 e 5.20. Assim como nas métricas anteriores, o resultado médio possui comportamento semelhante as imagens *Charlize* e *Lighthouse*, tendo a SR AWF em RGB apresentado os melhores resultados, sendo seguida de perto pela SR AWF com Wavelet na luminância. Porém alguns *outliers* são observados, o que produz valores muito alto de desvio-padrão.

Como o S-CIELAB visa medir a fidelidade na reprodução das cores das imagens, a grande diferença entre os valores da SR AWF em RGB e da SR AWF com compressão de crominância pode ser justificado pelo fato da compressão de crominância descartar informações de cor que não são relevantes para a visão humana mas impactam fortemente na medida de fidelidade de cor entre a imagem estimada e a imagem desejada.

Tabela 5.17: Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas sem ruído.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0	37	<b>1,6</b>	6,7 <sup>(**)</sup>	<b>0</b>	<b>1</b>
#2	59	10376	2547	2562	3175	3357
#3	0	45	<b>18,7</b>	8,2	<b>0</b>	<b>1</b>
#4	98	14440	2866	3454	4006	2637

(\*\*) Desvio-padrão alto se deve à presença de *outlier*.

Tabela 5.18: Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas com SNR = 40 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0	48	<b>4,5</b>	10,8	<b>0</b>	<b>1</b>
#2	56	10422	2545,4	2594,8	307	3229
#3	0	154	<b>13,7</b>	34,1	<b>3</b>	<b>1</b>
#4	101	14528	2903,9	3476,5	4006	2663

Tabela 5.19: Análise do S-CIELAB da super-resolução do conjunto de 30 imagens com SNR = 30 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	0	48	<b>58,7</b>	97,8	<b>23</b>	<b>49</b>
#2	83	10422	2672,5	2721,2	307	3391
#3	5	154	135,2	191,4	<b>61</b>	<b>52</b>
#4	125	14528	3187,5	3717,9	4451	2959

Tabela 5.20: Análise do S-CIELAB da super-resolução do conjunto de 30 imagens processadas com SNR = 25 dB.

Método	Mínimo	Máximo	Média	<i>Std</i>	<i>Charlize</i>	<i>Lighthouse</i>
#1	4	1109	<b>216</b>	268,4	<b>172</b>	<b>194</b>
#2	12138	12138	2980,6	2946,6	3852	3700
#3	9	1586	<b>369,5</b>	444,7	<b>209</b>	<b>144</b>
#4	237	16983	4020,2	4264,1	5817	2627

## 5.4 Complexidade Computacional

A complexidade computacional de cada um dos algoritmos de super-resolução depende da dimensão e do número de imagens LR, do fator de escala da super-resolução e dos parâmetros utilizados no método. No método SR *multiframe* adaptativo o número total de iterações é o principal fator para o aumento da complexidade computacional, sendo que a quantidade de iterações necessárias está diretamente ligada ao passo de iteração ( $\beta$ ) da Eq. (4.30). Um valor pequeno de  $\beta$  necessitará de mais iterações para que o método convirja para o valor mínimo esperado da função-custo. Porém, um valor alto de passo pode acarretar divergência ou uma oscilação em torno do ponto desejado.

Nos modelos de super-resolução com filtro de Wiener adaptativo, o número de operações está diretamente ligado ao tamanho da janela de observação, ao fator de ampliação, ao número e às dimensões das imagens LR utilizadas. O principal gargalo da análise computacional no método é o cálculo da Eq. (3.3). A análise do número de operações de ponto flutuante (flops) necessárias para executar o algoritmo SR AWF é apresentada em [17]. O número de operações para o caso RGB pode ser descrito como:

$$\text{flops} \cong 3 \left( \frac{K^3}{3} + 2D_x D_y K^2 + NK \right), \quad (5.3)$$

onde  $K$  é definido na Eq. (3.1),  $D_x$  e  $D_y$  são inteiros múltiplos dos fatores de escala da super-resolução nas duas dimensões da imagem, respectivamente  $L_x$  e  $L_y$ , e  $N$  é o número total de pixels HR a serem estimados em cada canal de cor.

Para o método SR AWF com Wavelet na luminância, há a adição das sub-bandas da Wavelet, resultando em três quadros a mais do que no caso RGB. Por isso, o número de flops é equivalente ao dobro do número de operações da SR AWF em RGB. No modelo com compressão de crominância, o único parâmetro alterado em relação à Eq. (5.3) é o número de pixels, pois os canais de crominância possuem 1/4 do número de pixels da luminância. O número de operações para o método de SR AWF com compressão de crominância pode ser descrito como:

$$\text{flops} \cong 3 \left( \frac{K^3}{3} + 2D_x D_y K^2 \right) + \frac{3}{2} NK. \quad (5.4)$$

Apesar do uso do tempo de processamento como estimativa de complexidade de um algoritmo estar sujeito a exceções, imprecisões ou problemas, os algoritmos envolvidos nesta dissertação são bem-comportados numericamente e não têm uma relação complicada entre tempo de processamento e número de operações envolvidas. Assumindo então que a relação entre tempo de processamento e complexidade é praticamente linear para os algoritmos em estudo, realizamos a estimativa da

complexidade com base no tempo de processamento.

A análise dos tempos de execução dos algoritmos com as diferentes técnicas de super-resolução indicou que em média, para os conjuntos de imagens utilizados, a SR AWF com compressão de crominância é concluída em 53% do tempo de execução da aplicação da SR AWF em RGB. O resultado é condizente com a diferença entre o número de operações dos dois métodos. Como o número de pixels é a parcela mais significativa no cálculo dos flops, em um caso extremo de  $\lim_{N \rightarrow \infty}$ , temos o dobro de operações em RGB do que no modelo com compressão de crominância.

Como esperado, o método com Wavelet na luminância é executado com o dobro de tempo do método em RGB. A média dos tempos de execução corresponde a 206% do tempo da SR AWF em RGB. Este fato comprova que o número de pixels a serem estimados é o principal gargalo das operações de SR AWF, pois a conversão para o espaço de cores YCbCr e o cálculo da Transformada Wavelet, além de suas operações inversas, impactam de maneira insignificante o tempo total de operação.

O tempo de execução da técnica de SR *multiframe* adaptativo foi em média 154% do tempo de processamento da SR AWF em RGB. Desta forma, o método além de não apresentar melhores resultados, possui um custo de execução maior do que o do método mais simples de super-resolução com filtro de Wiener adaptativo. A Fig. 5.12 apresenta um gráfico com a comparação da complexidade computacional dos algoritmos de SR de imagens coloridas discutidos neste trabalho.

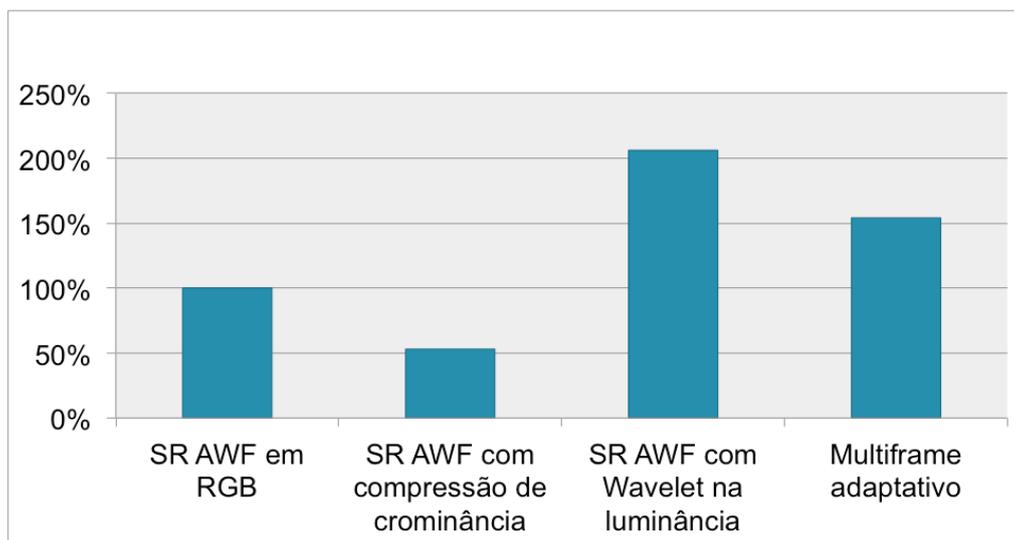


Figura 5.12: Comparação de complexidade computacional dos algoritmos.

## 5.5 Grade HR Esparsa

Um dos principais problemas das técnicas de super-resolução é a falta de imagens LR para compor a grade HR. Desta forma, é comum que a grade seja esparsa, pois em uma captura natural, a variação entre as imagens não é planejada. Nas análises quanto à variação do ruído, nove imagens LR foram utilizadas, preenchendo completamente a grade HR.

Para analisar este problema de grade HR esparsa, variamos de um até nove o número de imagens LR disponíveis para a realização da super-resolução, de forma a preencher toda a grade HR. Todas as imagens LR utilizadas possuem SNR de aproximadamente 40 dB. Assim como na análise anterior, a super-resolução foi implementada com um fator de escala igual a três em ambas as direções das imagens LR, para todos os quatro algoritmos utilizados nas análises anteriores.

A Fig. 5.13 apresenta a evolução da PSNR resultante da super-resolução da imagem *Charlize* para diferentes conjuntos de imagens LR utilizados no processo, variando entre uma e nove imagens LR. Os gráficos com os resultados das métricas S-CIELAB e SSIM são apresentados respectivamente nas Figs. 5.14 e 5.15.

De maneira geral, os resultados apresentam melhoras significativas com o aumento do número de imagens LR na super-resolução. No método SR AWF com compressão de crominância, observamos que a partir de cinco imagens LR distintas, os resultados das três métricas aproximam-se de suas convergências, não apresentando variações significativas, se comparados com cenários mais esparsos. Situação semelhante é percebida nos valores S-CIELAB dos métodos AWF em RGB e com Transformada Wavelet aplicada à luminância. Nestes casos, com apenas seis imagens LR distintas, a super-resolução atingiu a melhor situação permitida pela métrica.

O método de SR *multiframe* adaptativo apresentou para todas as métricas resultados inferiores aos dos modelos com filtro de Wiener. Tais resultados estão de acordo com os encontrados na Seção 5.3, onde o método apresentou os piores resultados de super-resolução. A Fig. 5.16 apresenta a comparação entre os quatro métodos de SR discutidos para o cenário inicial de cinco imagens de baixa resolução distintas. A análise visual é compatível com os resultados das métricas, com o melhor resultado sendo obtido com o da SR AWF em RGB .

Uma análise semelhante foi feita para a imagem *Lighthouse*. A Fig. 5.17 apresenta o gráfico da evolução da PSNR da super-resolução realizada com diferentes quantidades de imagens LR. As Figs. 5.18 e 5.19 ilustram a evolução da S-CIELAB e da SSIM, respectivamente. A comparação dos métodos, para a super-resolução com cinco imagens LR distintas, é apresentada na Fig. 5.20.

Os desempenhos dos métodos de super-resolução para a imagem *Lighthouse* são semelhantes aos apresentados para *Charlize*. Os níveis apresentados pelas métricas

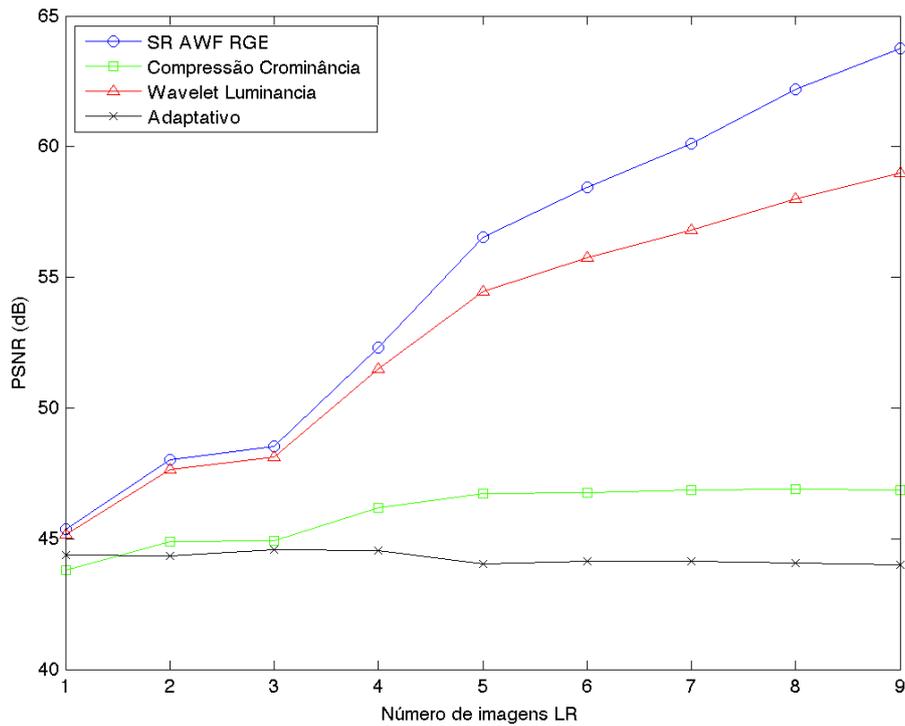


Figura 5.13: Evolução da PSNR da super-resolução da imagem *Charlize* com variação no número de imagens LR.

são próximos e as curvas apresentam padrões análogos para cada relação métrica-método, em especial para os métodos de super-resolução com filtro de Wiener.

Como observado para a imagem *Charlize*, os métodos baseados na SR AWF convergem para seus melhores resultados de S-CIELAB e SSIM após a quinta imagem, apresentando inclinação próxima a zero deste ponto em diante dos gráficos. Por outro lado, os resultados do método de SR *multiframe* adaptativo com cinco imagens LR ainda estão longe do melhor caso com nove imagens LR. Este resultado deve-se principalmente ao fato da correção do *aliasing* não ser bem sucedida, como podemos ver na Fig. 5.20(d). A imagem apresenta diversas falhas na construção de bordas em regiões de altas frequências.

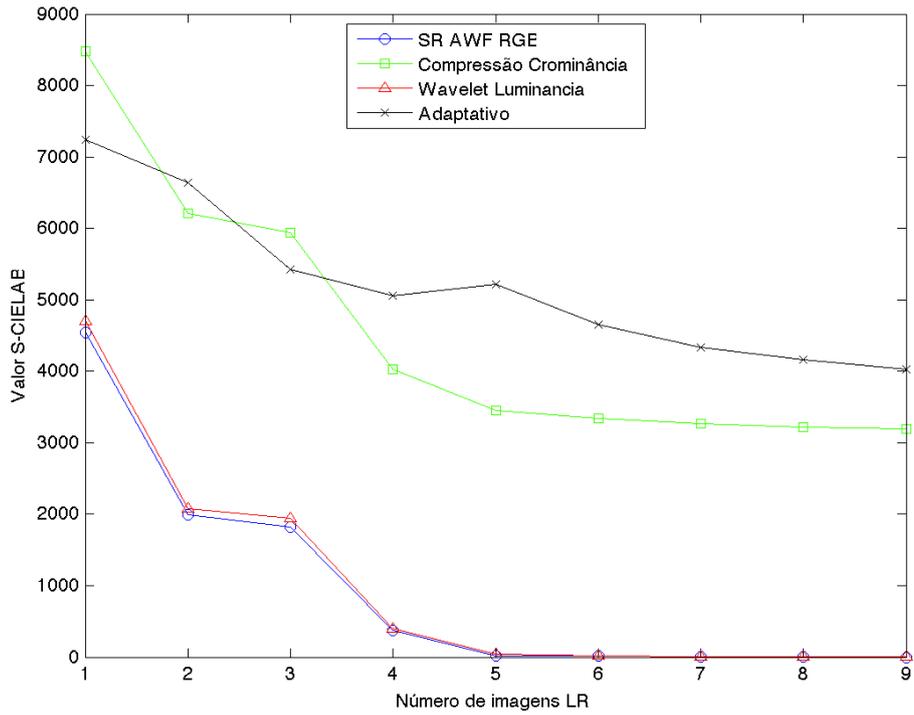


Figura 5.14: Evolução da S-CIELAB da super-resolução da imagem *Charlize* com variação no número de imagens LR.

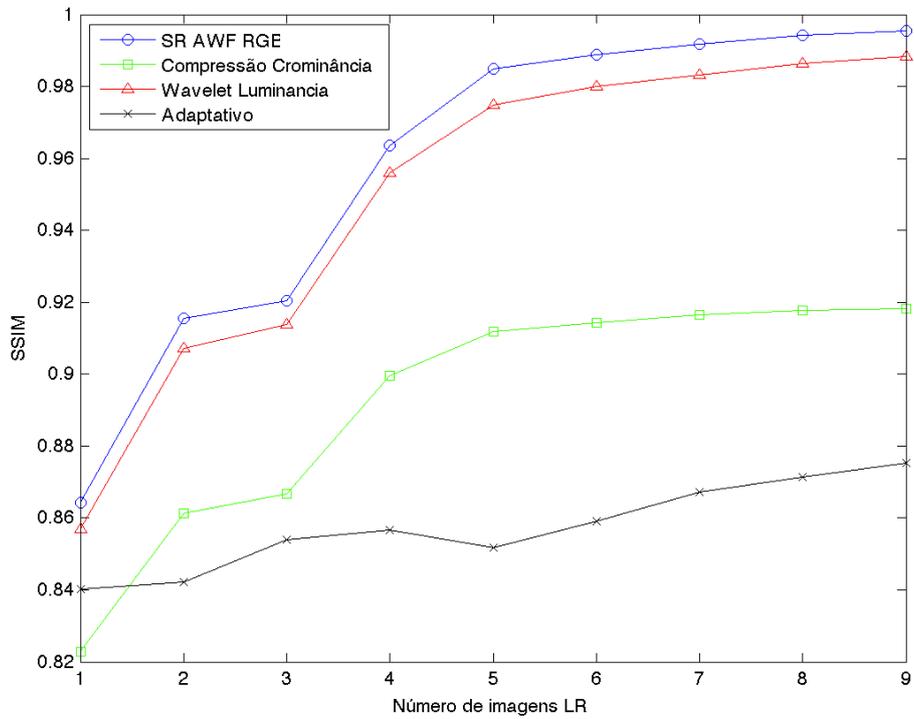
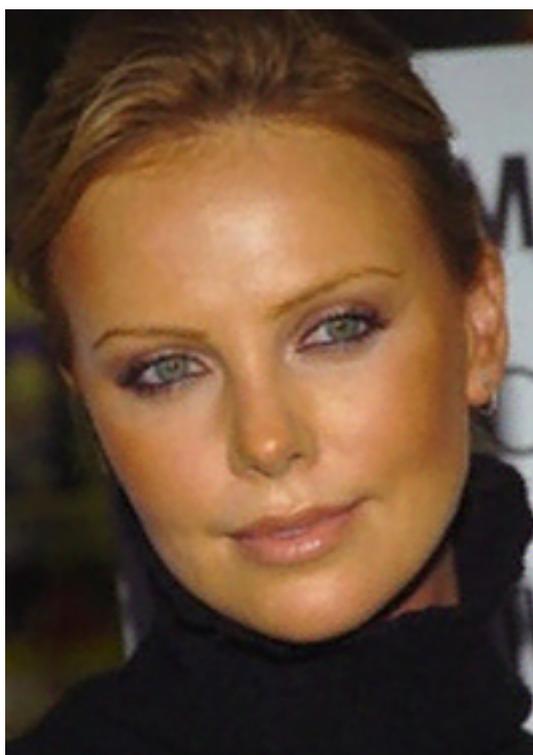
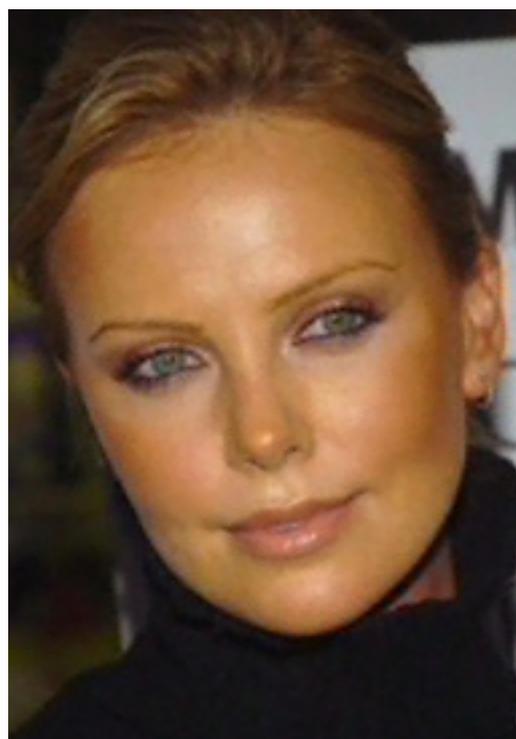


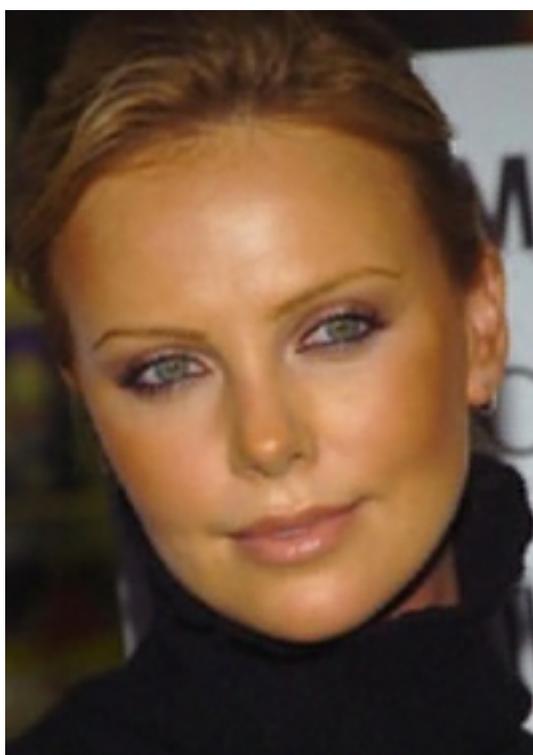
Figura 5.15: Evolução da SSIM da super-resolução da imagem *Charlize* com variação no número de imagens LR.



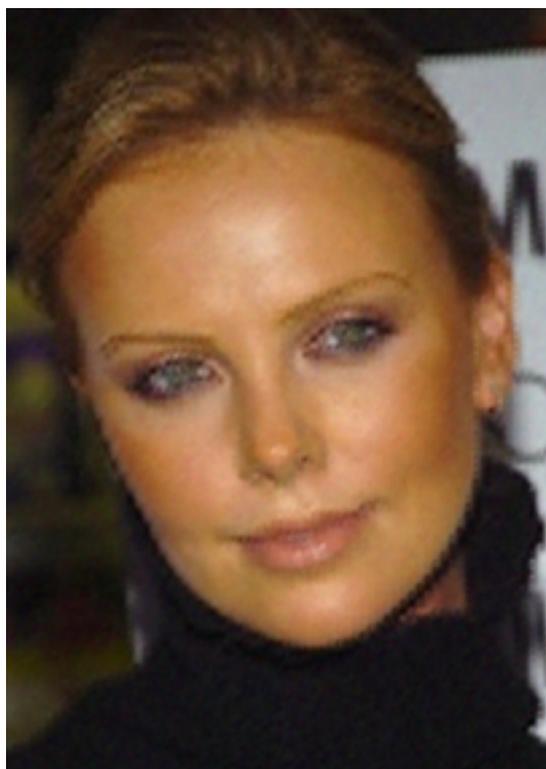
(a) SR AWF RGB



(b) SR AWF com compressão de crominância



(c) SR AWF YCbCr com Wavelet



(d) SR *multiframe* adaptativo

Figura 5.16: Super-resolução da imagem *Charlize* a partir de cinco imagens LR com SNR de 40 dB.

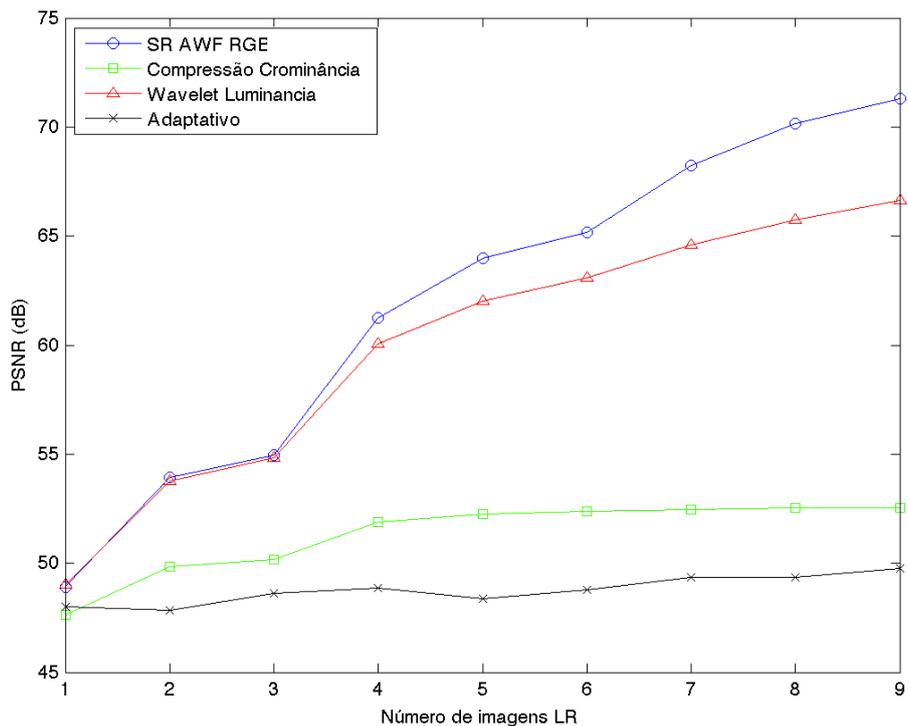


Figura 5.17: Evolução da PSNR da super-resolução da imagem *Lighthouse* com variação no número de imagens LR.

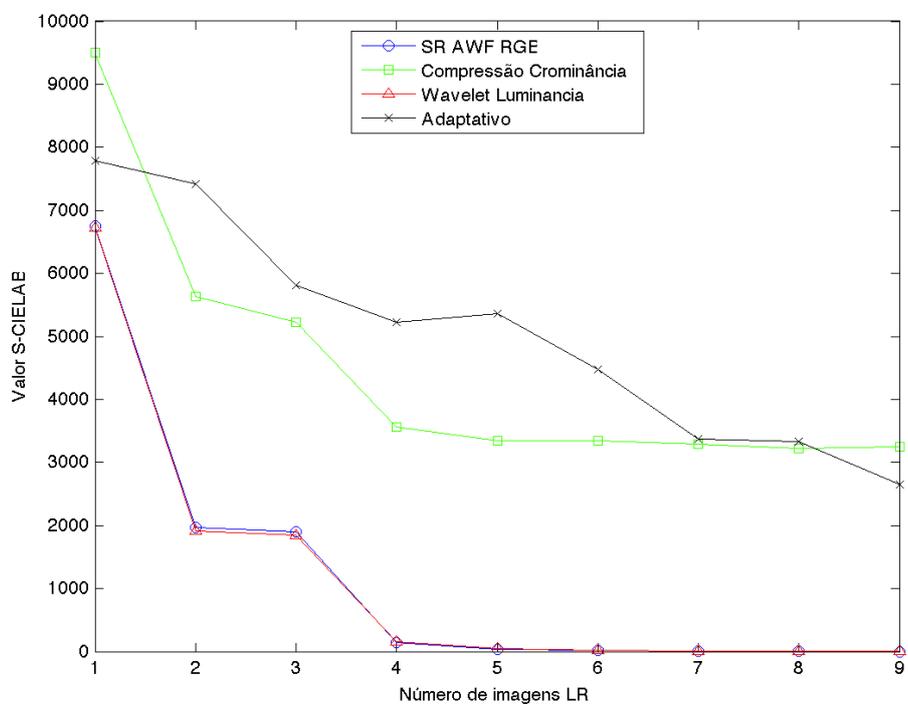


Figura 5.18: Evolução da S-CIELAB da super-resolução da imagem *Lighthouse* com variação no número de imagens LR.

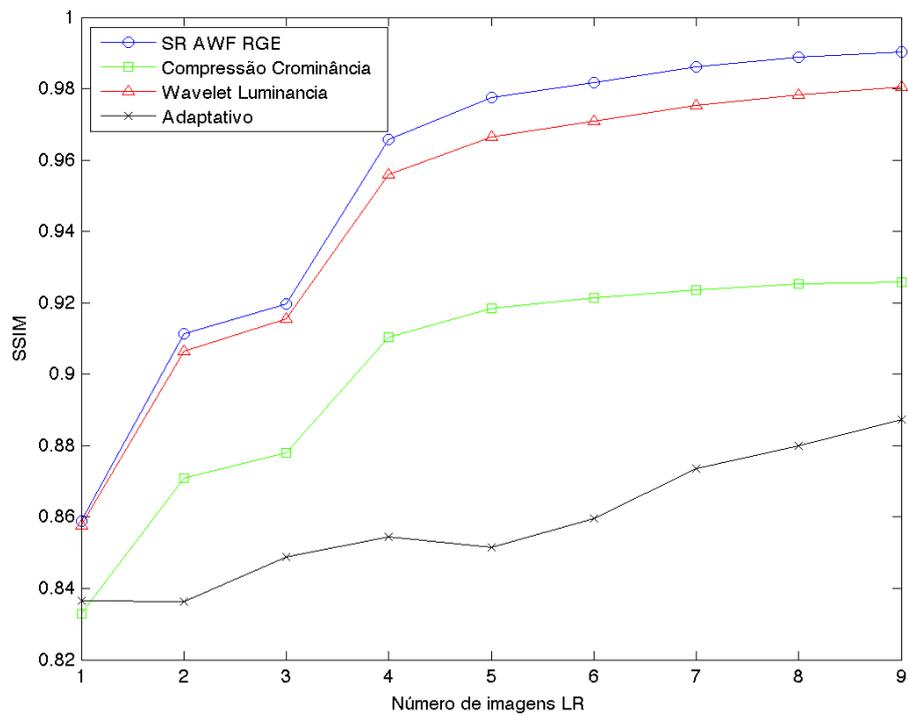


Figura 5.19: Evolução da SSIM da super-resolução da imagem *Lighthouse* com variação no número de imagens LR.



(a) SR AWF RGB



(b) SR AWF com compressão de crominância



(c) SR AWF YCbCr com Wavelet



(d) SR *multiframe* adaptativo

Figura 5.20: Super-resolução da imagem *Lighthouse* a partir de cinco imagens LR e SNR de 40 dB.

# Capítulo 6

## Conclusões

Os resultados deste trabalho confirmam a eficácia da super-resolução de imagens coloridas utilizando filtro de Wiener. Dentre os algoritmos estudados, o SR AWF em RGB é o mais indicado para a super-resolução de imagens *demosaiiced*. Porém, se complexidade computacional for um problema, ou se as imagens LR apresentarem baixa SNR, a aplicação das outras variantes é recomendada.

A variação do método SR AWF com compressão de crominância mostrou-se uma eficiente alternativa para a redução do tempo de processamento, devido ao baixo número de operações aritméticas necessárias em relação aos demais métodos. Apesar dos valores indicados pelas métricas de avaliação sugerirem que o método é inferior às outras duas variações de SR AWF, os resultados visuais não apresentam diferenças significativas que comprometam a qualidade da imagem de alta resolução.

Para imagens LR corrompidas por ruídos de alta intensidade, é indicado o uso da SR AWF com Wavelet na luminância, já que este algoritmo é o mais eficaz para a redução de ruído durante a super-resolução, além de apresentar valores de PSNR e SSIM próximos à SR AWF em RGB. Porém das três variações da SR AWF é o algoritmo menos eficiente. Tem um custo computacional significativamente maior que os demais.

O método de SR *multiframe* adaptativo apresentou resultados visuais desagradáveis e métricas ruins, quando comparado aos concorrentes com filtro de Wiener. Apesar do algoritmo ser baseado na minimização do erro médio quadrático, a evolução da PSNR com as iterações converge para valores inferiores aos dos concorrentes nos casos analisados. Era esperado que os resultados do algoritmo de SR *multiframe* adaptativo fossem inferiores aos do SR AWF, pois os modelos com filtro de Wiener são baseados em um modelo de reconstrução ótima.

Como trabalho futuro, pode-se integrar o *demosaiicing* à super-resolução com filtro de Wiener adaptativo, como proposto recentemente em [44] para a resolução do problema de super-resolução de CFA (uma visão geral sobre o conceito de CFA e a sua reconstrução por *demosaiicing* é apresentada no Apêndice B). Outra abordagem

interessante seria a introdução da correlação cruzada entre os canais de cores no cálculo do filtro de Wiener, com a finalidade de restaurar cores corretamente em casos de super-resolução de CFA. Uma nova opção de trabalho seria aplicar a SR AWF às sub-bandas das Wavelets de métodos que utilizam a Transformada Wavelet no CFA para realizar *demosaiicing*, como em [45].

Uma alternativa no domínio da frequência seria adaptar o modelo AWF para este domínio e aplicá-lo em conjunto com técnicas de reconstrução de imagens coloridas amostradas pelo Padrão de Bayer (ver Apêndice B). Os métodos de *demosaiicing* no domínio da frequência apresentam resultados satisfatórios com rápida execução, como é o caso dos algoritmos propostos em [46].

# Referências Bibliográficas

- [1] PARK, S. C., PARK, M. K., KANG, M. G. “Super-Resolution Image Reconstruction: A Technical Overview”, *IEEE Signal Processing Magazine*, v. 20, n. 3, pp. 21–36, maio 2003.
- [2] SHAH, N. R., ZAKHOR, A. “Resolution Enhancement of Color Video Sequences”, *IEEE Transactions on Image Processing*, v. 8, n. 6, pp. 879–885, junho 1999.
- [3] TOM, B. C., KATSAGGELOS, A. K. “Resolution Enhancement of Monochrome and Color Video Using Motion Compensation”, *IEEE Transactions on Image Processing*, v. 10, n. 2, pp. 278–287, 2001.
- [4] IRANI, M., PELEG, S. “Improving Resolution by Image Registration”, *CVGIP: Graphical Models and Image Processing*, v. 53, n. 3, pp. 231–239, 1991.
- [5] HUANG, T. S., TSAI, R. Y. “Multiframe Image Restoration and Registration”. In: *Advances in Computer Vision and Image Processing*, v. 1, JAI Press, cap. 7, pp. 317–339, Greenwich, Estados Unidos, 1984.
- [6] KIM, S. P., BOSE, N. K., VALENZUELA, H. M. “Recursive Reconstruction of High Resolution Image From Noisy Undersampled Multiframe”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 38, pp. 1013–1027, junho 1990.
- [7] YANG, J., HUANG, T. In: Milanfar, P. (Ed.), *Super-Resolution Imaging*, 1 ed., cap. Image Super-Resolution: Historical Overview and Future Challenges, pp. 1–33, Boca Raton, FL, Estados Unidos, CRC Press, 2011.
- [8] HARDIE, R. C., BARNARD, K. J., ARMSTRONG, E. E. “Join MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images”, *IEEE Transactions on Image Processing*, v. 6, n. 12, pp. 1621–1633, dezembro 1997.

- [9] STARK, H., OSKOUI, P. “High-Resolution Image Recovery from Image-Plane Arrays, Using Convex Projections”, *Journal of Optical Society of America*, v. 6, n. 11, pp. 1715–1726, novembro 1989.
- [10] KEREN, D., PELEG, S., BRADA, R. “Image Sequence Enhancement Using Sub-Pixel Displacements”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, v. 38, pp. 742–746, junho 1988.
- [11] ALAM, M. S., BOGNAR, J. G., HARDIE, R. C., et al. “Infrared Image Registration and High-Resolution Reconstruction Using Multiple Translationally Shifted Aliased Video Frames”, *IEEE Transactions on Instrumentation and Measurement*, v. 49, n. 5, pp. 915–923, outubro 2000.
- [12] ELAD, M., HEL-OR, Y. “A Fast Super-Resolution Reconstruction Algorithm for Pure Translational Motion and Common Space-Invariant Blur”, *IEEE Transactions on Image Processing*, v. 10, n. 8, pp. 1187–1193, agosto 2001.
- [13] LERTRATTANAPANICH, S., BOSE, N. K. “High Resolution Image Formation from Low Resolution Frames Using Delaunay Triangulation”, *IEEE Transactions on Image Processing*, v. 11, n. 12, pp. 1427–1441, dezembro 2002.
- [14] FARSIU, S., ROBINSON, D., ELAD, M., et al. “Fast and Robust Multiframe Super-Resolution”, *IEEE Transactions on Image Processing*, v. 13, n. 10, pp. 1327–1344, outubro 2004.
- [15] ELAD, M., FEUER, A. “Super-Resolution Reconstruction of Image Sequences”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 21, n. 9, pp. 817–834, setembro 1999.
- [16] ZOMET, A., RAV-ACHA, A., PELEG, S. “Robust Super-Resolution”. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, v. 1, pp. 645–650, dezembro 2001.
- [17] HARDIE, R. “A Fast Image Super-Resolution Algorithm Using an Adaptive Wiener Filter”, *IEEE Transactions on Image Processing*, v. 16, n. 12, pp. 2953–2965, dezembro 2007.
- [18] HARDIE, R. C. In: Milanfar, P. (Ed.), *Super-Resolution Imaging*, 1 ed., cap. Super-Resolution Using Adaptive Wiener Filters, pp. 35–61, Boca Raton, FL, Estados Unidos, CRC Press, 2011.
- [19] NARAYANAN, B., HARDIE, R. C., BARNER, K. E., et al. “A Computationally Efficient Super-Resolution Algorithm for Video Processing Using

- Partition Filters”, *IEEE Transactions on Circuits and Systems for Video Technology*, v. 17, n. 5, pp. 621–635, maio 2007.
- [20] THERRIAN, C. W. *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ, Estados Unidos, Prentice Hall, 1992.
- [21] JAIN, A. K. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, Estados Unidos, Prentice-Hall, 1989.
- [22] TAKEMURA, E. S. *Algoritmos para Super-Resolução de Imagens Baseados nas Filtragens de Wiener e Adaptativa Usando a Transformada Wavelet*. Dissertação de Mestrado, UFRJ/COPPE, Rio de Janeiro, RJ, Brasil, 2010.
- [23] MALLAT, S. G. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, n. 7, pp. 174–693, julho 1989.
- [24] MALLAT, S. *A Wavelet Tour of Signal Processing*. 3 ed. Waltham, Massachusetts Waltham, Massachusetts Waltham, MA, Estados Unidos, Academic Press, 1998.
- [25] VAIDYANATHAN, P. P. *Multirate Systems and Filter Banks*. 1 ed. Englewood Cliffs, NJ, Estados Unidos, Prentice Hall, 1992.
- [26] BURRUS, C. S., GOPINATH, R. A., GUO, H. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Upper Saddle River, NJ, Estados Unidos, Prentice Hall, 1997.
- [27] NGUYEN, N., MILANFAR, P. “A Wavelet-Based Interpolation-Restoration Method for Superresolution”, *Circuits, Systems and Signal Process*, v. 19, pp. 321–338, agosto 2000.
- [28] FARSIU, S. *A Fast and Robust Framework for Image Fusion and Enhancement*. Tese de Doutorado, University of California, Santa Cruz, Santa Cruz, CA, Estados Unidos, dezembro 2005.
- [29] FARSIU, S., ELAD, M., MILANFAR, P. “Multiframe Demosaicing and Super-Resolution of Color Images”, *IEEE Transactions on Image Processing*, v. 15, n. 1, pp. 141–159, janeiro 2006.
- [30] GONZALEZ, R. C., WOODS, R. E. *Processamento de Imagens Digitais*. 1 ed. São Paulo, SP, Brasil, Edgard Blucher, 2006.

- [31] ELAD, M., FEUER, A. “Restoration of Single Super-Resolution Image from Severe Blurred, Noise and Down-Sampled Measured Images”, *IEEE Transactions on Image Processing*, v. 6, n. 12, pp. 1646–1658, dezembro 1997.
- [32] ZOMET, A., PELEG, S. “Efficient Super-Resolution and Applications to Mosaics”. In: *Proceedings 15th International Conference on Pattern Recognition*, v. 1, pp. 579 – 583, setembro 2000.
- [33] FARSIU, S., ROBINSON, D., ELAD, M., et al. “Robust Shift and Add Approach to Super-Resolution”. In: *Proceedings of SPIE Conference in Applications of Digital Signal and Image Processing*, pp. 121–130, agosto 2003.
- [34] HEL-OR, Y., KEREN, D. *Demosaicing of Color Images Using Steerable Wavelets*. Relatório técnico, HP Labs Israel, Tech. Rep. HPL-2002-206R1 20020830, 2002.
- [35] RUDIN, L. I., OSHER, S., FATEMI, E. “Nonlinear Total Variation Based Noise Removal Algorithms”, *Physica D*, v. 60, n. 1-4, pp. 259–268, novembro 1992.
- [36] TOMASI, C., MANDUCHI, R. “Bilateral Filtering for Gray and Color Images”. In: *Proceedings of the 1998 IEEE International Conference on Computer Vision*, pp. 836–846, Bombay, India, janeiro 1998.
- [37] PEI, S. C., TAM, I. K. “Effective Color Interpolation in CCD Color Filter Array Using Signal Correlation”, *IEEE Transactions on Image Processing*, v. 13, n. 6, pp. 503–513, setembro 2003.
- [38] KEREN, D., OSADCHY, M. “Restoring Subsampled Color Images”, *Machine Vision and Applications*, v. 11, n. 4, pp. 197–202, 1999.
- [39] THE MATHWORKS, INC. “PSNR - Compute peak signal-to-noise ratio (PSNR) between images”. 2013. Disponível em: <<http://www.mathworks.com/help/vision/ref/psnr.html>>. Acesso em: 23 de dezembro de 2013.
- [40] ZHANG, X., SILVERSTEIN, D. A., FARRELL, J. E., et al. “Color Image Quality Metric S-CIELAB and Its Application on Halftone Texture Visibility”. In: *Proceedings IEEE COMPCON 97. Digest of Papers*, pp. 44–48, 1997.
- [41] ZHANG, X. “S-CIELAB: A Spatial Extension to the CIE L\*a\*b\* DeltaE Color Difference Metric”. maio 1998. Disponível em: <<http://white>.

stanford.edu/~brian/scielab/>. Acesso em: 23 de dezembro de 2013.

- [42] WANG, Z., BOVIK, A. C., SHEIKH, H. R., et al. “Image Quality Assessment: From Error Visibility to Structural Similarity”, *IEEE Transactions on Image Processing*, v. 13, n. 4, abril 2004.
- [43] WANG, Z. “The SSIM Index for Image Quality Assessment”. janeiro 2012. Disponível em: <<https://ece.uwaterloo.ca/~z70wang/research/ssim/>>. Acesso em: 23 de dezembro de 2013.
- [44] KARCH, B. K., HARDIE, R. C. “Adaptive Wiener Filter Super-Resolution of Color Filter Array Images”, *OSA Optics Express*, v. 21, n. 16, pp. 18820–18841, agosto 2013.
- [45] CHEN, L., YAP, K.-H., HE, Y. “Subband Synthesis for Color Filter Array Demosaicking”, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, v. 38, n. 2, pp. 485–493, março 2008.
- [46] DUBOIS, E. “Frequency-Domain Methods for Demosaicking of Bayer-Sampled Color Images”, *IEEE Signal Processing Letters*, v. 12, n. 12, pp. 847–850, dezembro 2005.
- [47] NAKAMURA, J. *Image Sensors and Signal Processing for Digital Still Cameras*. Boca Raton, FL, Estados Unidos, Taylor & Francis Group, 2006.
- [48] BAYER, B. E. *Color Imaging array*. US patent 3,971,065, Eastman Kodak Company, 20 de julho 1976.
- [49] KREIS, T. *Handbook of Holographic Interferometry: Optical and Digital Methods*. 1 ed. Bremen, Alemanha, Wiley-VCH, 31 de janeiro 2005.
- [50] BLU, T., THÉVENAZ, P., UNSER, M. “Linear Interpolation Revitalized”, *IEEE Transactions on Image Processing*, v. 13, n. 5, pp. 710–719, maio 2004.
- [51] ADAMS, J., PARULKI, K., SPAULDING, K. “Color Processing in Digital Cameras”, *IEEE Micro*, v. 18, n. 6, pp. 20–30, novembro-dezembro 1998.
- [52] KEYS, R. G. “Cubic Convolution Interpolation for Digital Image”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 29, n. 6, pp. 1153–1160, dezembro 1991.

- [53] CHANG, E., CHEUNG, S., PAN, D. “Color Filter Array Recovery Using a Threshold-Based Variable Number of Gradients”, *Proceedings SPIE*, v. 3650, pp. 36–43, janeiro 1999.
- [54] GUNTURK, B. K., GLOTZBACH, J., ALTUNBASAK, Y., et al. “Demosaicing: Color Filter Array Interpolation”, *IEEE Signal Processing Magazine*, pp. 44–54, janeiro 2005.
- [55] SAKAMOTO, T., NAKANISHI, C., HASE, T. “Software Pixel Interpolation for Digital Still Cameras Suitable for a 32-bit MCU”, *IEEE Transactions Consumer Electronics*, v. 44, n. 4, pp. 1342–1352, novembro 1998.
- [56] GUNTURK, B. K., ALTUNBASAK, Y., MERSEREAU, R. M. “Color Plane Interpolation Using Alternating Projections”, *IEEE Transactions on Image Processing*, v. 11, n. 9, pp. 997–1013, setembro 2002.

# Apêndice A

## Sensores de Imagem

Uma câmera fotográfica captura variações de intensidade de luz em um plano, que corresponde a uma imagem, e converte essas intensidades em sinais correspondentes que podem ser armazenados. As primeiras câmeras fotográficas se baseavam no fato do halogeneto de prata ser sensível à luz, escurecendo com a exposição. O filme fotográfico era formado com pequenos pedaços de halogeneto de prata colados a uma fita plástica por uma gelatina, formando um mosaico. Conforme a luz incide no filme, cada pedaço de halogeneto de prata “queima” de acordo com a intensidade luminosa. Neste sistema de fotografia a informação da imagem é capturada e convertida por reações químicas, e o filme serve de sensor e armazenador.

As popularmente conhecidas como câmeras digitais, utilizam como sensor de captura um dispositivo semicondutor (CCD ou CMOS), também conhecido como imageador, que converte a informação da imagem em sinais eletrônicos em vez de produzir reações químicas. Neste caso, o sensor funciona como dispositivo de conversão e não tem utilidade para armazenar, como no caso do filme fotográfico. O sistema de aquisição eletrônica requer um dispositivo de armazenamento independente, que pode ser um cartão de memória baseado em semicondutores, discos rígidos ou óticos.

O sensor eletrônico é formado por uma matriz finita de pixels, e a quantidade de pixels corresponde à resolução do sensor. A principal vantagem de imagens capturadas por sensores eletrônicos é a facilidade de processamento, que pode ser feito através de circuitos internos na câmera, ou até mesmo no circuito eletrônico do próprio sensor. Os sinais elétricos de cada pixel são digitalizados para armazenamento. Este trabalho se baseia em processamento de imagens digitais, por este motivo está focado em sensores eletrônicos. Em [47], são abordados os principais assuntos sobre sensores de imagens de câmeras digitais.

## A.1 CCD

*Charge-Coupled Device* (CCD) ou Dispositivo de Carga Acoplada, em uma tradução literal, é um sensor de imagem formado por uma matriz de capacitores acoplados, onde cada capacitor corresponde a um pixel. Este nome deriva de como a carga é lida após a captura da imagem. O processo de fabricação da matriz de capacitores possibilita que a carga de um capacitor possa ser transferida para um capacitor vizinho sem comprometer a qualidade do sinal. Um circuito externo ao sensor é responsável pelo controle da transferência de carga elétrica entre os capacitores.

Os capacitores são expostos à luz (imagem) e se carregam de acordo com a intensidade luminosa e o tempo de exposição. Após este intervalo, a carga armazenada por cada capacitor representa o valor da intensidade luminosa de cada um dos pixels da imagem. A primeira linha da matriz é então lida por um registrador, que possui na saída um amplificador e um conversor analógico-digital. Após a conclusão da leitura da primeira linha, o circuito controlador ativa a transferência de cargas, e a segunda linha é transferida para a primeira e então lida. Os capacitores estão acoplados de tal forma que as cargas possam ser transferidas linha a linha, da última até a primeira e então para o registrador. Após a leitura de todas as linhas, os dados da imagem são então armazenados. Este processo é representado na Fig. A.1.

Nas últimas décadas o CCD foi a tecnologia predominante para aplicações em câmeras digitais. A maturidade do processo de produção permitiu que a transferência de cargas mantenha alta qualidade, com baixo ruído de imagem e alta densidade de pixels. Quando queremos aumentar a resolução de uma câmera, é importante elevar o número de pixels para que a área do sensor aumente.

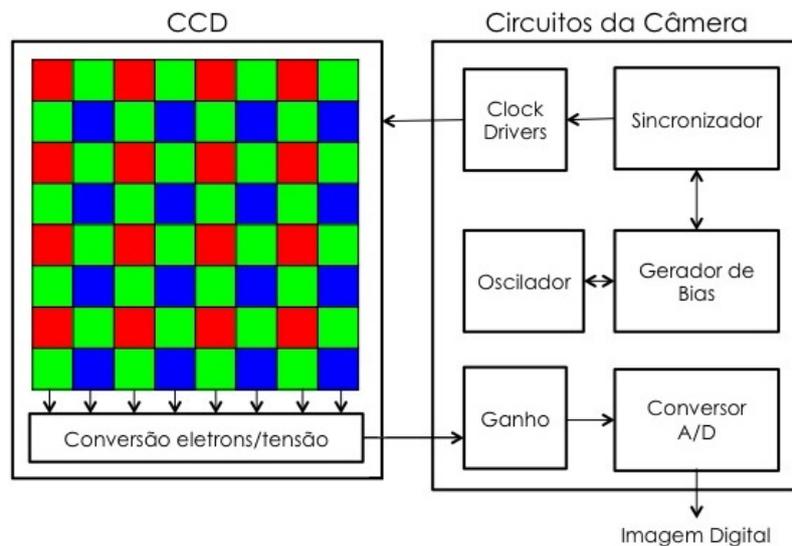


Figura A.1: Funcionamento do sensor CCD.

## A.2 CMOS

Os sensores CMOS (*Complementary Metal-Oxide Semiconductor*) utilizam a mesma tecnologia empregada comumente em fabricação de circuitos integrados (CI), sendo a mais empregada na fabricação de CIs. Os custos de produção deste sensor são bem inferiores aos de CCDs. Cada pixel do sensor é composto por uma série de transistores, onde um fotodiodo é responsável em transformar luz em eletricidade. Como há transistores instalados em cada pixel, a tecnologia CMOS permite que circuitos de amplificação, leitura, processamento e quantização possam ser realizados dentro do mesmo chip do sensor.

Pixels CMOS podem ser acionados individualmente através de um sistema de coordenadas  $X - Y$ , representando linha e coluna. Em geral, este tipo de leitura apresenta menor consumo de energia e maior velocidade de leitura, em relação a leitura global do sensor, pois somente os pixels selecionados ficam ativados. Este esquema de leitura permite uma flexibilidade na escolha de vários modos de leitura. O endereçamento por linha e coluna pode ser observado na Fig. A.2 sobre o funcionamento do sensor CMOS.

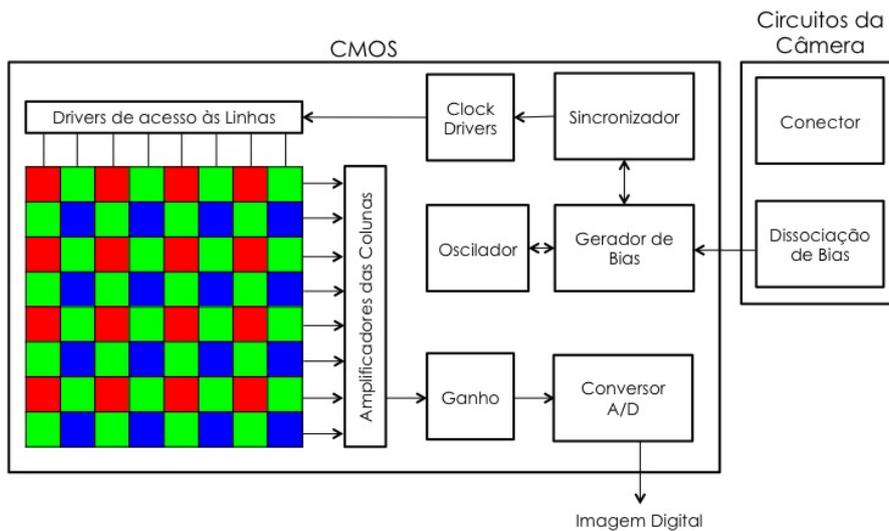


Figura A.2: Funcionamento do sensor CMOS.

APS (*Active Pixel Sensor* – Sensor por Pixel Ativo) é o termo dado a um pixel quando o sinal pode ser amplificado por um circuito interno ao pixel. Quando amplificado no pixel, o sinal fica mais robusto diante do ruído injetado pelo sistema de leitura na saída, enquanto em sistemas passivos (sem amplificação no pixel) o ruído é amplificado junto com o sinal, após a saída. Por possuir pixels passivos, o CCD é afetado por injeção de ruído na transferência de cargas, o que pode acarretar um problema conhecido como *smear* [47]. A relação entre o tamanho do fotodiodo e a superfície do pixel é conhecida como *fill factor*. Quanto menor esta razão, mais

circuitos podemos incluir em cada um dos pixels. Por possuírem muitos transistores, os pixels CMOS convencionais são menos sensíveis à luz do que os de CCD, pois muitos fótons não colidem com o fotodiodo e sim com transistores do pixel. Microlentes são utilizadas sobre a matriz de pixels para compensar este problema e convergir os fótons para o fotodiodo.

Outro benefício é que a tensão de alimentação diminui conforme a escala de tamanho da tecnologia, reduzindo o consumo de energia, consideravelmente menor do que para os CCDs. Os fotodiodos CMOS são imunes ao efeito de *blooming*, problema comum nos capacitores de CCDs, que ocorre quando uma fonte de luz sobrecarrega a sensibilidade do pixel, fazendo com que este transborde carga para os capacitores vizinhos. A Tabela A.1 apresenta uma comparação entre as tecnologias CCD e CMOS.

Tabela A.1: Comparação entre CCD e CMOS.

<b>Parâmetros \ Sensor</b>	<b>CCD</b>	<b>CMOS</b>
<i>Blooming</i>	Suscetível	Inexistente
Custo	Alto	Baixo
Consumo de Energia	Alto	Baixo
Circuitos	Externos	Internos e Externos
Leitura	Linha a linha	Coordenada X-Y
Fabricação	Altamente especializada	Amplamente utilizado
<i>Fill Factor</i>	Alto	Baixo
Ruído de Imagem	Muito Baixo	Baixo
Ruído de Transferência	Suscetível	Insignificante
Sensibilidade à Luz	Alta	Baixa

# Apêndice B

## *Color Filter Array*

Apesar de sensíveis à luminosidade, principalmente aos comprimentos de onda visíveis pelo olho humano, os sensores de imagem CCD e CMOS não são sensíveis à cor. Se visualizássemos a imagem produzida por um destes sensores, a veríamos em tons de cinza, como a luminância. Por este motivo, um método de separação de cores precisa ser empregado no sensor para capturar imagens coloridas.

Câmeras de vídeo de alta qualidade costumam usar três sensores de imagem, correspondentes às cores primárias RGB. A luz capturada pelas lentes da câmera é direcionada para um conjunto de prismas e espelhos dicróicos, conhecido como *beam split*. Os espelhos funcionam como filtros óticos, que refletem uma frequência de luz e permitem a passagem das demais. A função dos prismas é direcionar de maneira precisa as diversas reflexões da luz, direcionando cada uma das cores primárias para seus sensores correspondentes. Entre os prismas e os sensores existe um filtro óptico, que permite a passagem apenas do comprimento de luz desejado. Este processo, ilustrado na Fig. B.1, necessita de grande espaço na câmera para a inclusão de prismas, maior distância entre as lentes e três chips de sensores de imagem, o que torna a sua fabricação muito cara.

Como citado anteriormente, um sensor de imagem eletrônico é composto por uma matriz de pixels, coberta por microlentes que convergem a luz que incide no pixel para o foto-sensor. Em aplicações de câmeras fotográficas digitais de baixo custo, uma matriz de filtros de cor é colocada dentro do chip do sensor, sobre o fotodiodo. As cores são separadas por pixels, permitindo a implementação em câmeras de pequeno porte. Este modelo, conhecido como *Color Filter Array* (CFA – matriz de filtro de cores), é ilustrado na Fig. B.2. Câmeras digitais costumam utilizar cores primárias no CFA, por possuírem alta razão ao sinal-ruído e maior reprodução de cores, devido às suas propriedades de comprimento de onda serem mais seletivas [47].

O padrão de CFA mais utilizado é o padrão de Bayer. Sua configuração possui quadrados compostos por quatro filtros de cor, dois verdes em diagonal, um azul e

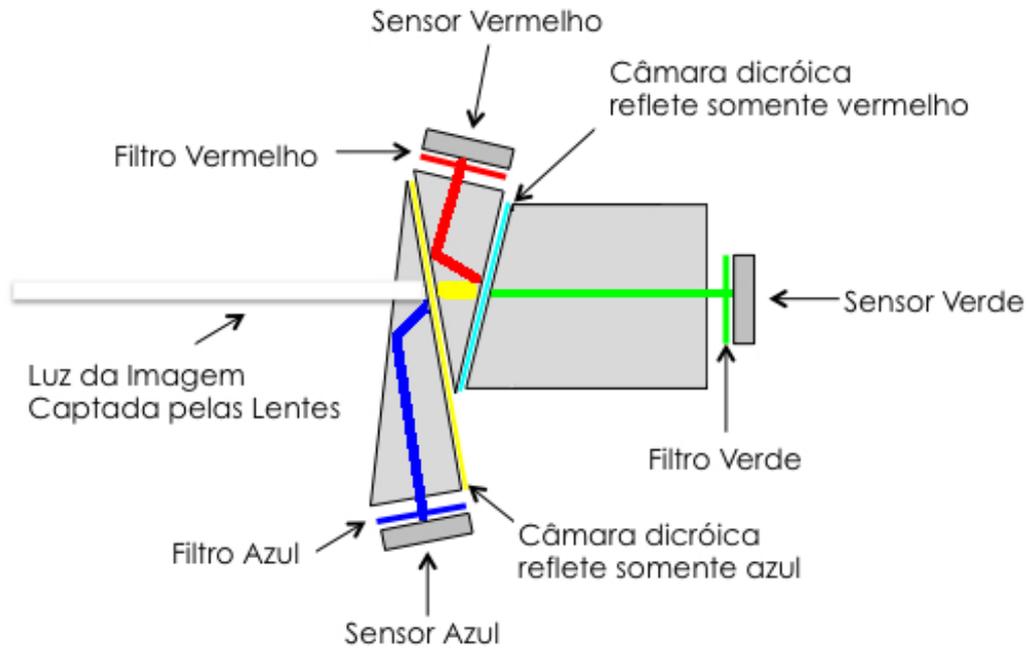


Figura B.1: Funcionamento de uma câmera colorida com *beam split*.

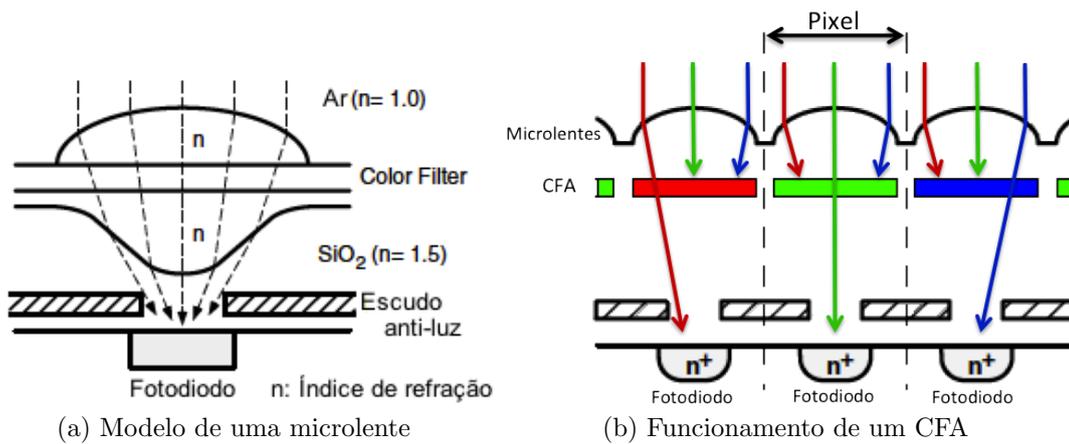


Figura B.2: Microlente, CFA e pixel [47].

um vermelho, que se espalham sobre o sensor, como ilustrado na Fig. B.3, sendo cada filtro de cor sobreposto à um pixel do sensor. O sistema visual humano é mais sensível aos detalhes de cor verde, que se assemelha à luminância [48]. Em contrapartida, as cores vermelha e azul estão associadas à percepção de cor.

Alguns padrões de CFA utilizam as cores complementares ciano (Cy), magenta (Mg) e amarelo (Ye), e também a luz branca (W). Sendo a luz branca a combinação de todas as cores, representamos as cores complementares com as seguintes equações:

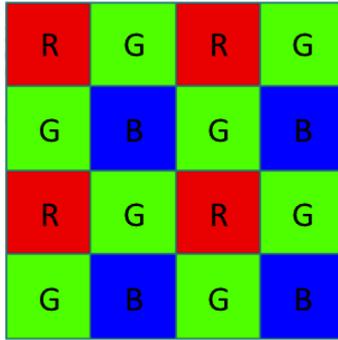


Figura B.3: Padrão de Bayer

$$Cy = G + B = W - R \quad (\text{B.1a})$$

$$Mg = R + B = W - G \quad (\text{B.1b})$$

$$Ye = R + G = W - B. \quad (\text{B.1c})$$

Modelos alternativos de CFA são apresentados na Fig. B.4. A letra ‘E’ representa “emerald”, uma cor parecida com o ciano, estipulada pela Sony.

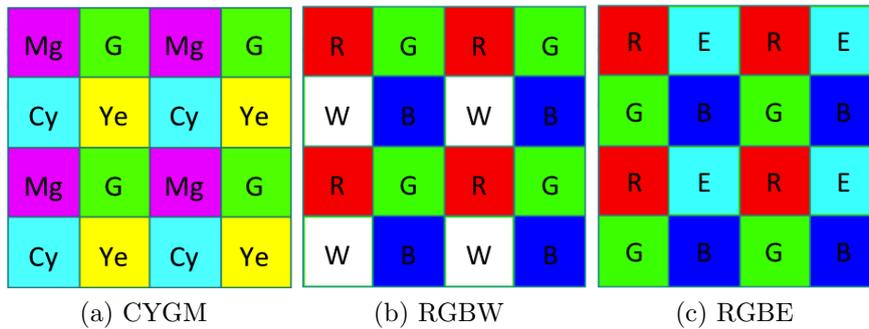


Figura B.4: Exemplo de padrões de CFA.

Estes padrões de CFA apresentados nas Figs. B.3 e B.4 são conhecidos com filtros de cor lateral. Outro tipo de separação de cor é conhecido como filtro de cor vertical. Este filtro é semelhante ao utilizado em filmes coloridos, porém implementado com sensores eletrônicos. Cada posição na matriz do sensor possui três pixels ativos, dispostos um sobre o outro, que utilizam o silício do foto-sensor como filtro de cor, baseando-se nos comprimentos de onda. Primeiramente é filtrado o canal azul, em seguida o verde e por último o vermelho, que possui maior comprimento de onda. Este sensor é conhecido como X3 Foveon [49].

## B.1 *Demosaicing*

Devido ao fato de cada sensor do CFA possuir o valor de apenas uma cor primária, é necessário fazer uma estimativa das demais cores em cada pixel. O conceito de *demosaicing* consiste em interpolar esses pixels para reconstruir a cena original com informação das três cores básicas em cada pixel. As principais técnicas de *demosaicing* são aplicadas no domínio espacial. Os primeiros métodos utilizavam a informação dos pixels vizinhos para estimar os valores ausentes através de médias, interpolação bilinear [50, 51], e bicúbica [52], ou cópia da informação vizinha. São métodos de fácil implementação e baixo custo computacional, e em geral utilizam apenas a informação do mesmo canal de cor para calcular os valores ausentes.

A utilização de algoritmos baseados em médias espaciais é semelhante à aplicação de um filtro passa-baixas. Como em imagens as altas frequências concentram-se em regiões de transição bruscas, a aplicação deste tipo de algoritmo tende a borrar as bordas da nova imagem. Visando evitar este tipo de problema, foram desenvolvidos algoritmos adaptativos que possibilitam aplicar diferentes estimativas para regiões de altas e baixas frequências. Em [53] é introduzido um método que utiliza a informação do gradiente direcional, permitindo interpolações em quatro direções e oito sentidos. Diversos métodos de interpolação de CFA são abordados em [54].

É bem conhecida a existência de correlação entre os canais (intercanal) de cor, vermelho, verde e azul. A correlação entre os canais de cor pode ser utilizada para a realização do *demosaicing* [55]. Nessas técnicas a luminosidade costuma ser aproximada pelo canal verde, que possui boa relação com os demais canais. Quando a luminosidade é conhecida, a falta de crominância vermelho ou azul pode ser estimada a partir de uma interpolação vermelho-verde ou azul-verde, respectivamente [45].

Outro ramo promissor das pesquisas sobre *demosaicing* utiliza técnicas com Transformada Wavelet, decompondo a imagem em uma série de sub-bandas com diferentes componentes de frequências. Em [56], é apresentado um algoritmo de alternância de projeções que utiliza as características de correlação intercanal de alta frequência e atualiza iterativamente os coeficientes Wavelets de acordo com um limiar. Wavelets também são utilizadas para dividir a imagem em sub-bandas e aproveitar as características de forte correlação espacial intracanal nas baixas frequências e a correlação intercanal dominante nas altas frequências [45].

O CFA pode ser interpretado como o somatório dos sinais  $R$ ,  $G$  e  $B$  subamostrados periodicamente. Este somatório, quando reagrupado em torno das frequências de subamostragem, permite a identificação de luminância e crominância da imagem no domínio da Transformada de Fourier bidimensional. Métodos de *demosaicing* no domínio da frequência utilizam um conjunto de filtros para extrair os sinais correspondentes de luminância e crominância na densidade espectral de potência. [46].