

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**MÉTODO DE SUPER-RESOLUÇÃO DE ÚNICA FIGURA
BASEADO EM APRENDIZADO PARA IMAGENS DE PLACAS
BRASILEIRAS DE BAIXÍSSIMA QUALIDADE**

ALEXANDRE NATÃ VICENTE

ORIENTADOR: HÉLIO PEDRINI

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
ÁREA DE CONCENTRAÇÃO INFORMÁTICA FORENSE E
SEGURANÇA DA INFORMAÇÃO**

PUBLICAÇÃO: PPGENE.DM – 640/2016

BRASÍLIA / DF: AGOSTO/2016

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

MÉTODO DE SUPER-RESOLUÇÃO DE ÚNICA FIGURA BASEADO
EM APRENDIZADO PARA IMAGENS DE PLACAS BRASILEIRAS
DE BAIXÍSSIMA QUALIDADE

ALEXANDRE NATÃ VICENTE

DISSERTAÇÃO DE Mestrado SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA
ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:



HÉLIO PEDRINI, Dr., UNICAMP
(ORIENTADOR)



FLÁVIO DE BARROS VIDAL, Dr., CIC/UnB
(EXAMINADOR INTERNO)



FABRIZIO ALPHONSUS ALVES DE MELO NUNES SOARES Dr., INF/UFG
(EXAMINADOR EXTERNO)

Brasília, 22 de agosto de 2016.

FICHA CATALOGRÁFICA

VICENTE, ALEXANDRE NATÃ

Método de Super-Resolução de Única Figura Baseado em Aprendizado para Imagens de Placas Brasileiras de Baixíssima Qualidade
[Distrito Federal] 2016.

xxiii, 52 p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2016).

Dissertação de Mestrado – Universidade de Brasília, Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. Super-resolução 2. Aprendizado de Máquinas 3. Reconstrução de Imagens 4. Placas de Veículos 5. Criminalística 6. Ciências Forenses

I. ENE/FT/UnB. II. Título (Série)

REFERÊNCIA BIBLIOGRÁFICA

Vicente, Alexandre Natã (2016). Método de Super-Resolução de Única Figura Baseado em Aprendizado para Imagens de Placas Brasileiras de Baixíssima Qualidade
Dissertação de Mestrado, Publicação PPGENE.DM – 640/2016, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 56 p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Alexandre Natã Vicente

TÍTULO DA DISSERTAÇÃO: Método de Super-Resolução de Única Figura Baseado em Aprendizado para Imagens de Placas Brasileiras de Baixíssima Qualidade

GRAU/ANO: Mestre/2016.

É concedida à Universidade de Brasília permissão para reproduzir cópias desta Dissertação de Mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Do mesmo modo, a Universidade de Brasília tem permissão para divulgar este documento em biblioteca virtual, em formato que permita o acesso via redes de comunicação e a reprodução de cópias, desde que protegida a integridade do conteúdo dessas cópias e proibido o acesso a partes isoladas desse conteúdo. O autor reserva outros direitos de publicação e nenhuma parte deste documento pode ser reproduzida sem a autorização por escrito do autor.



Alexandre Natã Vicente

SPO, Conjunto A, Lote 23, Complexo da PCDF, Ed. Sede
CEP 70.610-907 – Brasília – DF - Brasil

*Dedico este trabalho
à minha esposa, Lizandra Nascimento Vicente,
à minha mãe, Márcia da Fonseca Vicente,
ao meu pai, Dorival Aparecido Vicente,
e à minha irmã, Gisele Vicente.*

AGRADECIMENTOS

Agradeço à minha esposa Lizandra Nascimento Vicente pelo incentivo e por ficar ao meu lado em todos os momentos, sejam eles felizes ou tristes, e ter compartilhado minhas preocupações, pressões e obrigações.

Agradeço aos meus familiares, Márcia da Fonseca Vicente, Dorival Aparecido Vicente e Gisele Vicente, por me apoiarem mesmo à distância, e por suportarem a saudade.

Agradeço ao meu orientador, Prof. Dr. Hélio Pedrini por todo o apoio, disponibilidade e paciência durante a realização deste trabalho.

Todos os acima citados e aqueles não citados, mas que fizeram parte da minha vida nesses últimos anos, meus sinceros agradecimentos: muito obrigado!

RESUMO

MÉTODO DE SUPER-RESOLUÇÃO DE ÚNICA FIGURA BASEADO EM APRENDIZADO PARA IMAGENS DE PLACAS BRASILEIRAS DE Baixíssima Qualidade

Autor: Alexandre Natã Vicente

Orientador: Hélio Pedrini

Programa de Pós-graduação em Engenharia Elétrica

Brasília, Agosto de 2016

A melhoria da resolução especial de imagens de placas de veículos em cenários reais possui papel importante no campo da investigação criminal e da ciência forense. Neste trabalho, um método de super-resolução de única imagem baseado em aprendizado é proposto, o qual utiliza o conhecimento *a priori* de que a entrada é formada por um conjunto de imagens de algarismos de placas capturadas em baixíssima qualidade e baixíssima resolução. O método proposto emprega uma árvore de decisão para classificar cada imagem de entrada e o resultado da classificação é utilizado para ponderar o peso da imagem do bloco na etapa de reconstrução. Adicionalmente, a equalização de histograma é aplicada para melhorar a eficácia do classificador. Para demonstrar que o método proposto produz resultados satisfatórios, experimentos com imagens sintéticas e imagens do mundo real são realizados e analisados.

ABSTRACT

A LEARNING-BASED SINGLE-IMAGE SUPER-RESOLUTION METHOD FOR VERY LOW QUALITY LICENSE PLATE IMAGES

Author: Alexandre Natã Vicente

Supervisor: Hélio Pedrini

Programa de Pós-graduação em Engenharia Elétrica

Brasília, August 2016

Spatial resolution enhancement of license plate images in real scenarios plays an important role in criminal investigation and forensic science. This work presents a learning-based single-image super-resolution method based on a priori knowledge that the input is composed of license plate digit images captured at poor quality and very low resolution. The proposed method employs a decision tree to classify each input image and the classification results are used to weight the image patches in the reconstruction step. Additionally, the histogram equalization technique is performed to improve the effectiveness of the classifier. Experiments conducted on synthetic and real-world images demonstrate that the proposed method is capable of producing satisfactory results.

SUMÁRIO

1.	INTRODUÇÃO.....	1
1.1.	MOTIVAÇÃO.....	1
1.2.	OBJETIVOS E CONTRIBUIÇÕES	3
1.3.	ORGANIZAÇÃO DO TRABALHO.....	4
2.	CONCEITOS	5
2.1.	CONCEITOS FUNDAMENTAIS DE IMAGENS.....	5
2.2.	FORMAÇÃO DA IMAGEM DE BAIXA RESOLUÇÃO	6
2.3.	SUPER-RESOLUÇÃO DE IMAGENS E REVISÃO DA LITERATURA	6
2.3.1.	CLASSIFICAÇÃO POR NÚMERO DE IMAGENS DE ENTRADA	6
2.3.2.	CLASSIFICAÇÃO DE ACORDO COM O DOMÍNIO DE ATUAÇÃO	7
2.3.2.1.	MÉTODOS NO DOMÍNIO DA FREQUÊNCIA.....	7
2.3.2.2.	MÉTODOS NO DOMÍNIO ESPACIAL	8
2.3.3.	MÉTODOS BASEADOS EM RECONSTRUÇÃO.....	9
2.3.3.1.	ITERATIVE BACK PROJECTION	10
2.3.3.2.	GRADIENT PROFILE PRIOR	11
2.3.4.	MÉTODOS BASEADOS EM APRENDIZADO.....	13
2.4.	APRENDIZADO DE MÁQUINAS	14
2.4.1.	ÁRVORE DE DECISÃO.....	15
2.5.	ÁLGEBRA MATRICIAL	17
2.6.	MÉTRICAS UTILIZADAS NO TRABALHO	17
2.6.1.	CORRELAÇÃO.....	18
2.6.2.	SSIM.....	18
2.6.3.	RMSE.....	19
3.	MÉTODO PROPOSTO	20
3.1.	VISÃO GERAL DO MÉTODO	20

3.2.	PLACAS BRASILEIRAS E A SEGMENTAÇÃO DOS ALGARISMOS	24
3.3.	CRIAÇÃO DO DICIONÁRIO DE BLOCOS.....	26
3.4.	CRIAÇÃO DO CONJUNTO Ω	27
3.5.	CLASSIFICAÇÃO DA IMAGEM GERAL	29
3.6.	CÁLCULO DOS PESOS DOS BLOCOS	30
3.7.	MINIMIZAÇÃO DA FUNÇÃO CUSTO	31
4.	RESULTADOS EXPERIMENTAIS.....	33
4.1.	EFEITO DA MANIPULAÇÃO DE HISTOGRAMA NO DESEMPENHO DA ÁRVORE DE DECISÃO	33
4.2.	IMAGENS SIMULADAS	35
4.3.	IMAGENS DE ALGARISMOS DE PLACAS REAIS DE BAIXA RESOLUÇÃO	45
5.	CONCLUSÕES E TRABALHOS FUTUROS.....	48
	REFERÊNCIAS BIBLIOGRÁFICAS	50
	A – CÁLCULOS PARA A MINIMIZAÇÃO DA FUNÇÃO CUSTO.....	54

LISTA DE TABELAS

Tabela 4.1 – Taxas de erro de classificação média utilizando-se a correlação como medida de similaridade.....	34
Tabela 4.2 – Taxas de erro de classificação média utilizando-se a SSIM como medida de similaridade.....	35
Tabela 4.3 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 5 dB.	43
Tabela 4.4 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 10 dB.	44
Tabela 4.5 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 15 dB.	44

LISTA DE FIGURAS

Figura 1.1: Exemplo de resultado de um algoritmo de super-resolução (<i>Gradient Profile Prior</i> – GPP). À esquerda, as imagens em baixa resolução e, à direita, as imagens super-resolvidas com um fator de ampliação de 8 (Sun, Xu, & Shum, 2008).	2
Figura 2.1: Super-resolução utilizando o algoritmo GPP (Sun, Xu, & Shum, 2008).	12
Figura 2.2: Esquema de criação do dicionário de blocos.	13
Figura 2.3: Esquema geral de funcionamento de algoritmo baseado em exemplo. Imagem adaptada de Bevilacqua (2014).	14
Figura 2.4: Exemplo de árvore de decisão. Imagem adaptada de Mitchell (1997).	16
Figura 3.1: Fluxograma geral do método proposto	21
Figura 3.2: Árvore de decisão obtida na seção 4.2. Observação: “10” refere-se ao algarismo 0.	22
Figura 3.3: Árvore de decisão obtida na seção 4.2 com as probabilidades das folhas. Observação: “10” refere-se ao algarismo 0.	22
Figura 3.4: Caracteres com a fonte “Mandatory”.	25
Figura 3.5: Imagem em baixa resolução.	25
Figura 3.6: Exemplo de segmentação dos algarismos de uma placa em baixa resolução. A placa foi extraída da imagem ilustrada na Figura 3.5 e os retângulos coloridos mostram a segmentação de cada um dos quatro algarismos. Cada retângulo gerou uma imagem individual de baixa resolução.	26
Figura 3.7: Esquema de criação do dicionário de blocos utilizado no método proposto.	26
Figura 3.8: Árvore de decisão obtida na seção 4.3. Observação: “10” refere-se ao algarismo 0.	30
Figura 3.9: Ilustração do funcionamento do método do Gradiente Descendente.	32
Figura 4.1: As imagens padrões com fundo branco e com fundo cinza são mostradas, respectivamente, na linha superior e inferior.	34
Figura 4.2 – Amostra de imagens em baixa resolução do banco de imagens simuladas (as imagens apresentadas foram ampliadas por um fator de 8 pelo método do vizinho mais próximo para facilitar a visualização).	36
Figura 4.3 – Amostra do resultado das imagens do conjunto de teste super-resolvidas pelo método proposto, IBP, KRR, GPP e interpolação bicúbica, todas por um fator de ampliação de 4. Para facilitar a visualização, os resultados obtidos pelos métodos citados foram ampliados por um fator de 2 utilizando o método do vizinho mais próximo e a imagem original em baixa resolução foi ampliada por um fator de 8 utilizando o método do vizinho mais próximo.	37
Figura 4.4: Processo de cálculo das métricas RMSE e SSIM.	38
Figura 4.5: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 5 dB.	39
Figura 4.6: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 5 dB.	39

Figura 4.7: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 5 dB.	40
Figura 4.8: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 10 dB.	40
Figura 4.9: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 10 dB.	41
Figura 4.10: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 10 dB.	41
Figura 4.11: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 15 dB.	42
Figura 4.12: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 15 dB.	42
Figura 4.13: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 15 dB.	43
Figura 4.14 – Amostra de imagens do banco de imagens de algarismos de placas de automóveis reais de baixa resolução (as imagens apresentadas foram ampliadas por um fator de 8 pelo método do vizinho mais próximo para facilitar a visualização).45	
Figura 4.15 – Amostra do resultado das imagens do conjunto teste super-resolvidas pelo método proposto, IBP, KRR, GPP e a interpolação bicúbica, todas por um fator de ampliação de 4. Para facilitar a visualização, os resultados obtidos pelos métodos citados foram ampliados por um fator de 2 utilizando o método do vizinho mais próximo e a imagem original em baixa resolução foi ampliada por um fator de 8 utilizando o método do vizinho mais próximo.	46

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES

CFTV	Circuitos Fechados de Televisão
CLS	Mínimos Quadrados com Restrição (<i>Constrained Least Squares</i>)
CONTRAN	Conselho Nacional de Trânsito
GPP	Perfil de Gradiente a Priori (<i>Gradient Profile Prior</i>)
HR	Alta Resolução (<i>High Resolution</i>)
IBP	Retroprojeção Iterativa (<i>Iterative Back Projection</i>)
KRR	<i>Kernel Ridge Regression</i>
LOOCV	Validação Cruzada <i>Leave-One-Out</i> (<i>Leave-One-Out Cross-Validation</i>)
LR	Baixa Resolução (<i>Low Resolution</i>)
MAP	Máxima A Posteriori (<i>Maximum A Posteriori</i>)
MERCOSUL	Mercado Comum do Sul
ML	Máxima Verossimilhança (<i>Maximum Likelihood</i>)
POCS	Projeção em Conjuntos Convexos (<i>Projection Onto Convex Sets</i>)
px	pixel
RMSE	Raiz do Erro Quadrático Médio (<i>Root Mean Squared Error</i>)
SR	Super-Resolução (<i>Super-Resolution</i>)
SSIM	Similaridade Estrutural (<i>Structural Similarity</i>)

1. INTRODUÇÃO

Este capítulo apresenta as principais motivações para a investigação realizada neste trabalho (seção 1.1), os objetivos planejados e as contribuições atingidas (seção 1.2), bem como a estrutura geral e a organização do texto (seção 1.3).

1.1. MOTIVAÇÃO

A utilização de imagens digitais é crescente, tanto na vida pessoal como profissional das pessoas e empresas. Em 2015, o número de celulares no mundo passou dos 7 bilhões (Júnior, 2015) e, como os celulares mais modernos são equipados com câmeras fotográficas, é fácil ver a disseminação das imagens digitais. Por consequência, assuntos relacionados à melhoria da qualidade das imagens possuem grande interesse acadêmico e apelo econômico.

As imagens sofrem degradação de sua qualidade devido à diversos fatores como distorções de lente, distorções de foco, borramento, movimento do objeto, movimento da câmera, iluminação inadequada, incidência inadequada e pouca resolução. Apesar disso, obter uma imagem de boa qualidade é de interesse de diversas áreas como medicina, biologia, vigilância e monitoramento, detecção de alvo e reconhecimento, sensoriamento remoto e microscopia.

Em muitas das aplicações em que se deseja uma imagem de melhor resolução é impossível a substituição do sensor de aquisição da imagem (como é o caso de satélites, telescópios em órbita, sondas em locais de difícil acesso e imagens capturadas no passado) ou o custo de substituição das câmeras e sensores é economicamente inviável. Neste cenário, técnicas que melhoram a resolução de uma imagem através de processamento, após a aquisição da imagem, ganham força. Estas técnicas que buscam melhorar a qualidade das imagens aumentando sua resolução são denominadas de técnicas de super-resolução.

Super-resolução é uma família de métodos utilizados para aumentar a resolução espacial de imagens. A finalidade de um método de super-resolução é aumentar os detalhes perceptíveis de uma imagem e, se possível, que a imagem em super-resolvida (alta resolução), revele detalhes não perceptíveis na imagem em baixa resolução. A Figura 1.1 ilustra a aplicação de um algoritmo de super-resolução em uma imagem de baixa resolução.



Figura 1.1: Exemplo de resultado de um algoritmo de super-resolução (*Gradient Profile Prior – GPP*). À esquerda, as imagens em baixa resolução e, à direita, as imagens super-resolvidas com um fator de ampliação de 8 (Sun, Xu, & Shum, 2008).

Conforme demonstrado por Nasrollahi & Moeslund (2014), diversos artigos e livros foram publicados nas últimas duas décadas sobre super-resolução e diversos caminhos, por vezes, bastante distintos, foram seguidos no intuito de melhorar a resolução de uma imagem e acrescentar detalhes que antes não eram perceptíveis.

No âmbito da perícia e da investigação criminal, há grande necessidade de melhoria da qualidade de imagens após estas terem sido capturadas. Pela própria natureza do crime, este tipicamente acontece em locais e horas imprevisíveis e a captura de imagens que registram o crime ocorre sob circunstâncias pouco favoráveis, ou seja, em situações em que a iluminação e a incidência são inadequadas, as imagens são capturadas por câmeras de baixa qualidade (Circuitos Fechados de TV – CFTV, câmeras móveis como as de celulares) e há grande distância entre os objetos de interesse e a câmera. Como não é possível se alterar os equipamentos e circunstâncias em que o delito foi gravado, resta a perícia criminal trabalhar com técnicas que melhorem a qualidade da imagem via processamento, buscando revelar informações que eram de difícil percepção nas imagens de baixa qualidade. Dentre os

trabalhos de super-resolução de especial interesse forense por tratar de imagens de faces, placas de veículos ou oriundas de CFTV, pode-se citar: (Zou & Yuen, 2012), (Zeng & Huang, 2012), (Yoshida, Takahashi, Deguchi, Ide, & Murase, 2012), (Zhou, 2012), (Zhang L. , Zhang, Shen, & Li, 2010) e (Yang, Su, Chen, & Moon, 2010).

Um item de particular interesse para a perícia criminal são as placas de veículos. Muitas vezes o veículo utilizado em determinado crime é gravado, porém, devido à pouca qualidade das imagens, não é possível identificar a sua placa. Dessa forma, algoritmos que busquem melhorar a qualidade das imagens de placas têm aplicação direta na atividade de perícia criminal e de investigação criminal. Algoritmos voltados à melhoria de placas também são de interesse do controle de tráfego (por exemplo, para verificação de fuga de pedágios ou aplicações de multas) e controle de acesso entre outras aplicações. Segundo Nasrollahi & Moeslund (2014), até o ano de 2012 havia pelo menos vinte e dois artigos publicados tratando especificamente de algoritmos de super-resolução aplicados à leitura de placas de veículos.

Este trabalho foi desenvolvido em torno deste interesse da perícia criminal em algoritmos para melhoramento de placas de veículos. Mais especificamente, este trabalho focou na melhoria dos algoritmos presentes em placas de veículos, uma vez que, de posse dos algoritmos, o universo de placas possíveis já fica bastante restrito, já que as letras presentes nas placas seguem uma lógica de distribuição por região geográfica e temporal, de tal forma que, conhecendo-se os números da placa, a região geográfica do crime e quando este aconteceu, a placa do veículo fica praticamente individualizada.

Este trabalho buscou desenvolver um algoritmo de super-resolução específico para melhorar a qualidade de algoritmos de placas de veículos brasileiros com baixíssima resolução (cada algoritmo com menos de 20 x 20 pixels). O algoritmo desenvolvido se trata de uma abordagem para uma única imagem baseado em aprendizado e pode ser classificado como um algoritmo de super-resolução baseado em exemplos.

1.2. OBJETIVOS E CONTRIBUIÇÕES

Este trabalho apresenta um método de super-resolução de imagem única baseado em aprendizado e em dicionário de blocos para imagens de placas de automóveis brasileiros de baixíssima qualidade, o qual explora o conhecimento *a priori* de que as imagens contêm apenas um dígito de uma fonte de caracteres conhecida (“Mandatory”).

O método proposto foi submetido a experimentos em dois cenários: um com imagens de baixa resolução obtidas a partir de imagens de alta resolução conhecidas (seção 4.2), permitindo uma análise quantitativa e qualitativa do desempenho do método em comparação com outros métodos de super-resolução, e um cenário mais realista em que as imagens de baixa resolução foram nativamente capturadas em baixa resolução (seção 4.3), permitindo apenas uma análise qualitativa, porém mais realista, do desempenho do método em comparação com outros.

Como objetivos e contribuições secundários, o trabalho evidenciou a importância e o impacto que a manipulação do histograma acarreta no desempenho de algoritmos de super-resolução que utilizam dicionário de blocos. Também foi criado e se deixou à disposição do público os bancos de imagens utilizados nos experimentos, permitindo uma comparação de métodos neste domínio específico de imagens de algoritmos de placas brasileiras. O trabalho contribui para a comunidade científica forense ao demonstrar a potencialidade de algoritmo de super-resolução para a melhoria da qualidade visual de imagens de placas, mesmo sob condições severas de deterioração, e ao demonstrar que algoritmos de super-resolução podem explorar conhecimentos *a priori* de domínios específicos. Na mesma esteira, a abordagem didática do funcionamento do método proposto facilita a sua compreensão e o desenvolvimento de novas técnicas na área.

1.3. ORGANIZAÇÃO DO TRABALHO

O trabalho está organizado da seguinte forma: o Capítulo 2 apresenta os conceitos fundamentais de imagens e sua formação, apresenta o problema de super-resolução e faz uma revisão de sua literatura, discorre brevemente sobre aprendizado de máquinas, especialmente sobre árvore de decisão e álgebra matricial e termina apresentando a definição das métricas utilizadas neste trabalho; o Capítulo 3 descreve o método proposto neste trabalho, passando pelas normas sobre placas de automóveis brasileiros e por uma visão geral do método, seguida de uma visão mais detalhada de cada passo do método proposto; o Capítulo 4 apresenta e discute os resultados obtidos pelo método em um cenário com imagens de algoritmos de baixa resolução simuladas e em um cenário com imagens de algoritmos de placas de automóveis reais de baixa resolução e compara os resultados do método obtido com outros métodos de super-resolução; por fim, o Capítulo 5 apresenta as conclusões do trabalho.

2. CONCEITOS

Neste capítulo, os conceitos fundamentais utilizados no trabalho são apresentados. Primeiramente, os fundamentos de imagens (seção 2.1) e o modelo de formação de imagem adotado no trabalho (seção 2.2) são descritos. A seguir, o problema de super-resolução e uma revisão da literatura da área (seção 2.3) são apresentados. Em seguida, conceitos sobre aprendizado de máquina (seção 2.4) e álgebra matricial (seção 2.5) são brevemente descritos. Finalmente, as medidas utilizadas no trabalho são definidas (seção 2.6).

2.1. CONCEITOS FUNDAMENTAIS DE IMAGENS

Nesta seção o autor deste trabalho define alguns conceitos fundamentais de imagens utilizados neste trabalho.

Incidência é o ângulo entre o equipamento de captura da imagem e o objeto de interesse. Diz-se que uma incidência é boa ou adequada quando ela permite que se observe as características relevantes do objeto de interesse.

Iluminação é a forma como a cena ou objeto de interesse está sendo iluminada durante a captura da imagem. A iluminação tem papel relevante na captura da imagem, sendo que uma boa iluminação normalmente é aquela que proporciona bom contraste, boa definição das cores e dos detalhes do objeto de interesse.

Frequência de amostragem espacial é a quantidade de amostras (pixels) por unidade de comprimento. Dessa forma, uma boa frequência de amostragem espacial corresponde a uma alta frequência de amostragem, ou seja, uma alta quantidade de pixels por unidade de comprimento.

Resolução de uma imagem digital é a quantidade de pixels que uma imagem possui. Em geral, é apresentada na forma $w \times h$, em que w é o número de pixels na horizontal e h o número de pixels na vertical (e.g. 1920 x 1080 e 4000 x 2000) ou na forma de número total de pixels da imagem (e.g. 1 MP e 4 MP).

Distância é corresponde à distância entre o equipamento de captura da imagem e o objeto de interesse. A distância adequada para a captura da imagem de um objeto depende da incidência, da iluminação e da resolução/frequência de amostragem espacial.

2.2. FORMAÇÃO DA IMAGEM DE BAIXA RESOLUÇÃO

Uma imagem obtida por uma câmera digital tradicional é capturada a partir da luz proveniente da cena, a qual passa pelo conjunto de lentes e espelhos da câmera e atinge um sensor matricial, que transforma a intensidade da luz recebida em sinal elétrico, o qual é lido e forma a imagem digital obtida. Esses processos podem ser modelados pela seguinte equação (Tian, Yap, & He, 2011):

$$LR = (Blur * HR) \downarrow + n \quad \text{Eq. 2.1}$$

em que **LR** é a imagem digital obtida em baixa-resolução, **Blur** é o filtro que representa o espalhamento/borramento e as distorções de lente e de foco, **HR** é a imagem em alta resolução, \downarrow é o operador de subamostragem e **n** é um ruído aditivo.

2.3. SUPER-RESOLUÇÃO DE IMAGENS E REVISÃO DA LITERATURA

Super-resolução é uma família de métodos utilizados para aumentar a resolução espacial de imagens. A finalidade de um método de super-resolução é aumentar os detalhes perceptíveis de uma imagem e não apenas aumentar a quantidade de pixels de uma imagem sem ganho nos detalhes ou nitidez e, se possível, que a imagem super-resolvida (alta resolução) revele detalhes não perceptíveis na imagem em baixa resolução.

Existem diversos métodos de super-resolução com abordagens bastante distintas na literatura. Nas subseções de 2.3.1 a 2.3.4 são apresentadas as classificações mais comuns para os algoritmos de super-resolução.

2.3.1. CLASSIFICAÇÃO POR NÚMERO DE IMAGENS DE ENTRADA

Uma grande divisão na família de métodos de super-resolução diz respeito à quantidade de imagens utilizadas como entrada no algoritmo, podendo ser uma única imagem (métodos de imagem única) ou múltiplas imagens (método de múltiplas imagens) (Nasrollahi & Moeslund, 2014). Os princípios dos métodos de imagem única e de múltiplas imagens diferem profundamente.

Em linhas gerais, métodos de múltiplas imagens tentam combinar detalhes presentes nas diversas imagens e utilizar a informação redundante bem como as pequenas diferenças de informação entre as múltiplas imagens a fim de gerar uma ou várias imagens em alta resolução.

A premissa básica em métodos de múltiplas imagens é que se dispõe de uma sequência de imagens em baixa resolução de uma mesma cena e que diferem pouco entre si. Então, modela-se as imagens em baixa resolução como versões degradadas da cena em alta resolução conforme exposto na Eq. 2.1. A diferença entre as imagens em baixa resolução deve ser de tal forma que o deslocamento possua parte fracionária de pixel, pois, se o deslocamento corresponder a múltiplos inteiros de pixel, as imagens não possuem informações distintas.

Já os métodos de imagem única, por ter como informação apenas uma única entrada, buscam outras fontes de informação para adicionar na imagem de alta resolução. O problema de super-resolução de imagem única é intrinsecamente mal posto, já que diversas imagens em alta resolução podem gerar a imagem em baixa resolução observada.

Métodos de interpolação tradicionais (por exemplo, interpolação bilinear e interpolação bicúbica) estimam os pixels da imagem em alta resolução como uma combinação dos pixels de baixa resolução, agindo como uma suavização. Tal suavização pode gerar artefatos de borramento e de eco na imagem em alta resolução quando a imagem natural apresenta descontinuidades como bordas e cantos (Sun, Xu, & Shum, 2008). Por isso, o objetivo dos métodos de super-resolução de uma única imagem é obter resultados melhores que esses por meio do conhecimento *a priori* de estatísticas.

Os métodos de super-resolução de uma única imagem podem ser divididos em duas famílias principais: métodos baseados em reconstrução e métodos baseados em aprendizado, os quais estão descritos, respectivamente, nas seções 2.3.3 e 2.3.4 (Nasrollahi & Moeslund, 2014).

2.3.2. CLASSIFICAÇÃO DE ACORDO COM O DOMÍNIO DE ATUAÇÃO

Uma outra grande divisão na família de métodos de super-resolução diz respeito ao domínio em que os métodos atuam, ou seja, se atuam no domínio da frequência ou se atuam no domínio espacial.

2.3.2.1. MÉTODOS NO DOMÍNIO DA FREQUÊNCIA

Os métodos desta família primeiramente transformam a(s) imagem(ns) em baixa resolução para o domínio da frequência e estimam a imagem em alta resolução neste domínio. Posteriormente, transformam de volta a imagem estimada em alta resolução para o domínio espacial. A transformação para o domínio da frequência é feita, via de regra, ou pela

transformada de Fourier - (Gerchberg, 1974) e (Santis & Gori, 1975) - ou por transformadas Wavelet (Demirel & Anbarjafari, 2011).

A título de exemplo, os passos principais do método de múltiplas imagens baseado na transformada de Fourier proposto por Tsai & Huang (1984) são mostrados a seguir.

Seja g_k um conjunto de imagens transladadas de uma mesma região da Terra, f a imagem contínua da Terra, tal que $g_k(m, n) = f(x, y)$, em que $x = m + \Delta_{m_k}$ e $y = n + \Delta_{n_k}$, F_{g_k} a transformada contínua de Fourier da k -ésima imagem em baixa resolução e F_f a transformada contínua de Fourier da imagem em alta resolução. Então, pela propriedade de translação da transformada de Fourier:

$$F_{g_k}(m, n) = e^{i2\pi(m\Delta_{m_k} + n\Delta_{n_k})} F_f(m, n) \quad \text{Eq. 2.2}$$

Como as imagens em baixa resolução são versões amostradas da imagem contínua, então $g_k(m, n) = f(mT_m + \Delta_{m_k}, nT_n + \Delta_{n_k})$, em que T_m e T_n são os períodos de amostragens. Sendo G_k a transformada discreta de Fourier da imagem em g_k , a relação entre as transformada de Fourier contínua e discreta é dada por:

$$G_k(m, n) = \frac{1}{T_m} \frac{1}{T_n} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} F_{g_k} \left(\frac{m}{MT_m} + i \frac{1}{T_m}, \frac{n}{NT_n} + j \frac{1}{T_n} \right) \quad \text{Eq. 2.3}$$

em que M e N são os valores máximos das dimensões m e n , respectivamente, das imagens em baixa resolução. Postulando que as imagens em alta resolução são limitadas em banda e substituindo a Eq. 2.3 na Eq. 2.2, obtém-se a seguinte equação na forma matricial, em que Φ relaciona as transformadas discreta de Fourier das imagens em baixa resolução com a transformada contínua de Fourier da cena:

$$\mathbf{G} = \Phi \mathbf{F}_f \quad \text{Eq. 2.4}$$

Portanto, para estimar a imagem em alta resolução é necessário resolver a Eq. 2.4 e depois realizar a transformada inversa de Fourier para se ter a imagem em alta resolução no domínio espacial.

2.3.2.2. MÉTODOS NO DOMÍNIO ESPACIAL

Os métodos desta família atuam diretamente no domínio espacial. Esta família de métodos é mais populosa do que a família dos métodos no domínio da frequência. Exemplos de métodos no domínio espacial são: *Iterative Back Projections*, *Projection Onto Convex*

Sets, Maximum Likelihood, Maximum a Posteriori, métodos baseado em aprendizado e métodos baseados em reconstrução.

2.3.3. MÉTODOS BASEADOS EM RECONSTRUÇÃO

Métodos baseados em reconstrução são a maioria dos métodos para múltiplas imagens de entradas e também para uma única imagem de entrada e consistem em definir uma restrição para que a imagem em alta resolução, de tal forma que a qualidade da reconstrução seja aumentada (Nasrollahi & Moeslund, 2014). Há uma variedade de métodos nesta família e, em geral, a restrição utilizada é uma forma explícita para a função distribuição ou para a função energia. Alguns métodos dessa família são: *Iterative Back Projection* (IBP) (Irani & Peleg, 1990), *Projection Onto Convex Sets* (POCS) (Stark & Oskoui, 1989) (Malczewski & Stasinski, 2008), *Maximum Likelihood* (ML) (Cheeseman, Kanefsky, Kraft, & Stutz, 1994) (Jung, Jiao, Liu, & Gong, 2011), *Maximum A Posteriori* (MAP) (Zhang, Zhang, & Shen, 2012), *Gradient Profile Prior* (GPP) (Sun, Xu, & Shum, 2008) e *Fields of Experts* (Guo, Yang, Zhang, & Yu, 2009).

Alguns métodos dessa família exploram bordas, buscando reconstruí-las de forma mais nítida, como descrito por Sun, Xu, & Shum (2008). Outros métodos tentam resolver o fato do problema ser mal posto utilizando métodos de regularização.

Métodos de regularização utilizam um parâmetro de regularização o qual contém um conhecimento *a priori* para a imagem em alta resolução que permita reduzir o universo de soluções possíveis do problema mal posto. Um exemplo típico é o problema de mínimos quadrados com restrição (CLS – *Constrained Least Squares*). O CLS pode ser formulado de acordo com a equação a seguir:

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \sum_{k=1}^p \|\mathbf{LR}_k - \mathbf{Blur}_k \cdot \mathbf{HR}\|^2 + \alpha \|\mathbf{C} \cdot \mathbf{HR}\|^2 \right\} \quad \text{Eq. 2.5}$$

em que α é o multiplicador de Lagrange e \mathbf{C} é um filtro, usualmente um filtro passa-altas. Dessa forma, o elemento $\alpha \|\mathbf{C} \cdot \mathbf{HR}\|^2$ atua como uma restrição e permite uma solução única tendendo à solução, no caso, para uma imagem suave devido ao filtro passa-altas.

Outros exemplos comuns de regularização são os métodos baseados em estatísticas como o *Maximum A Posteriori* (MAP) e o *Maximum Likelihood* (ML).

2.3.3.1. ITERATIVE BACK PROJECTION

O método *Iterative Back Projection* (IBP) é um dos primeiros métodos de super-resolução no domínio espacial, podendo ser utilizado tanto para uma única imagem quanto para múltiplas imagens de entrada. Nesta subsecção, seu funcionamento para múltiplas imagens de entrada é descrito de forma sucinta (Irani & Peleg, 1990).

Primeiramente, define-se o modelo de formação da imagem, como o descrito na Eq. 2.1. Então, o IBP irá minimizar o seguinte custo:

$$\|A(HR) - LR\|_2^2 \quad \text{Eq. 2.6}$$

em que o operador A é tal qual $A(x) = (Blur * x) \downarrow$. Então, inicializa-se a estimativa da imagem em alta resolução e a cada iteração tal estimativa é refinada. Esta estimativa pode ser obtida ao se realizar o processo de registro das imagens em baixa resolução sobre uma grade de alta resolução e, depois, aplicar a média sobre esta grade.

Então, calcula-se o erro entre as imagens simuladas via Eq. 2.1 ($LR_simulada_k^{(t)}, k = 1, \dots, K$) e as imagens em baixa resolução de entrada, segundo a Eq. 2.7, em que t é o número da iteração:

$$Erro = \sqrt{\frac{1}{K} \sum_{k=1}^K \|LR_k - LR_simulada_k^{(t)}\|_2^2} \quad \text{Eq. 2.7}$$

Para o refinamento da estimativa da imagem em alta resolução usualmente se utiliza o método de Richardson dado pela seguinte expressão:

$$HR^{(t+1)}(x, y) = HR^{(t)}(x, y) + \frac{1}{K} \sum_{k=1}^K ((LR_k - LR_simulada_k^{(t)}) \uparrow) * Blur \quad \text{Eq. 2.8}$$

em que \uparrow é o operador de aumento de amostragem e $Blur$ é o operador de desborramento, o qual possui relação com o operador de borramento dado pela Eq. 2.9.

$$\forall \omega, \quad 0 < |1 - BLUR(\omega)BLUR(\omega)| < 1 \quad \text{Eq. 2.9}$$

em que $BLUR(\omega)$ e $BLUR(\omega)$ representam, respectivamente, as transformadas de Fourier do $Blur$ e $Blur$.

Repete-se as iterações até um critério de parada ser atingido, seja ele um número máximo de iterações ou algum critério que envolva as alterações ocorridas entre uma iteração e outra.

2.3.3.2. GRADIENT PROFILE PRIOR

O método *Gradient Profile Prior* (GPP) (Sun, Xu, & Shum, 2008) explora as bordas da imagem e tenta reconstruí-las de forma mais nítida. Em resumo, este método busca aprender como é o perfil do gradiente em torno de uma borda a partir de um grande conjunto de imagens naturais e, uma vez aprendida a forma deste perfil, esta informação é utilizada para forçar as bordas da imagem em alta resolução a ter um perfil semelhante ao aprendido.

A nitidez de um perfil de gradiente que passa pelo pixel x_0 de borda é definida da seguinte forma:

$$\sigma(p(x_0)) = \sqrt{\sum_{x \in p(x_0)} \frac{m(x)}{\sum_{s \in p(x_0)} m(s)} d^2(x, x_0)} \quad \text{Eq. 2.10}$$

em que m é a intensidade do gradiente, $p(x_0)$ é o perfil de gradiente que passa pelo pixel x_0 , $d(x, x_0)$ é a distância entre o pixel x e o pixel x_0 , σ variância, também denominada nitidez do perfil (quanto menor σ , mais nítida e bem definida é a borda).

Uma imagem interpolada tende a suavizar a imagem em alta resolução, por consequência, gera bordas borradas e pouco nítidas. Observou-se que os perfis dos gradientes em torno das bordas de imagens naturais são estáveis e invariantes à resolução da imagem, por isso, a distribuição dos perfis dos gradientes em torno das bordas é modelada como uma distribuição gaussiana generalizada e seus parâmetros (λ e σ) são obtidos após se analisar um grande conjunto de imagens naturais. Depois, analisa-se como se comportam os perfis dos gradientes em torno das bordas da imagem quando se realiza a operação de aumento de amostragem. Então, de posse dos parâmetros da distribuição gaussiana generalizada (λ e σ), pode-se calcular o campo gradiente transformado, ou seja, o campo gradiente que a imagem super-resolvida deve ter para apresentar bordas mais nítidas. Este cálculo é feito por meio das seguintes equações:

$$\nabla I_h^T(x) = \frac{\lambda_h \cdot \alpha(\lambda_h) \cdot \sigma_l \cdot \Gamma\left(\frac{1}{\lambda_l}\right)}{\lambda_l \cdot \alpha(\lambda_l) \cdot \sigma_h \cdot \Gamma\left(\frac{1}{\lambda_h}\right)} \exp\left\{-\left(\frac{\alpha(\lambda_h) \cdot |d(x, x_0)|}{\sigma_h}\right)^{\lambda_h} + \left(\frac{\alpha(\lambda_l) \cdot |d(x, x_0)|}{\sigma_l}\right)^{\lambda_l}\right\} \nabla I_l^u(x) \quad \text{Eq. 2.11}$$

$$\alpha(\lambda) = \sqrt{\Gamma\left(\frac{3}{\lambda}\right) / \Gamma\left(\frac{1}{\lambda}\right)} \quad \text{Eq. 2.12}$$

em que λ e σ são os parâmetros da distribuição gaussiana generalizada aprendido para o conjunto de imagens naturais em alta resolução e em baixa resolução, $d(x, x_0)$ é a distância entre o pixel x e o pixel x_0 , ∇I_l^H é o campo gradiente da imagem super-resolvida, ∇I_h^T é o campo gradiente transformado.

Uma vez calculado o campo gradiente transformado, determina-se a imagem super-resolvida buscando-se a imagem que minimiza a função energia definida na Eq. 2.13.

$$E(I_h|I_l, \nabla I_h^T) = E_i(I_h|I_l) + \beta E_g(\nabla I_h|\nabla I_h^T) \quad \text{Eq. 2.13}$$

$$E_i(I_h|I_l) = \|(Blur * I_h) \downarrow - I_l\|^2 \quad \text{Eq. 2.14}$$

$$E_g(\nabla I_h|\nabla I_h^T) = \|\nabla I_h - \nabla I_h^T\|^2 \quad \text{Eq. 2.15}$$

em que I_h é a imagem em alta resolução (super-resolvida), I_l é a imagem em baixa resolução, $Blur$ é o filtro que representa o espalhamento/borramento e as distorções de lente e de foco, \downarrow é o operador de subamostragem, ∇I_h é o campo gradiente da imagem em alta resolução, ∇I_h^T é o campo gradiente transformado.

Aplicando-se o método GPP, obtém-se bordas bem mais nítidas do que as originais, como ilustrado na Figura 2.1.

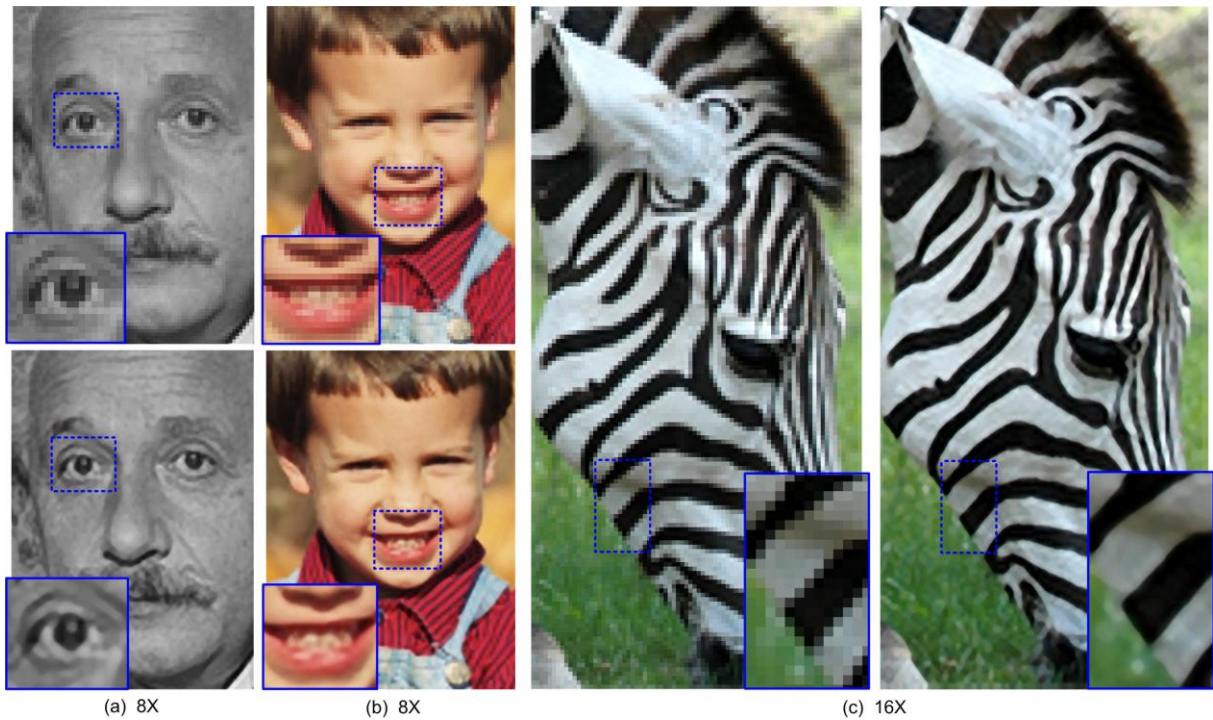


Figura 2.1: Super-resolução utilizando o algoritmo GPP (Sun, Xu, & Shum, 2008).

2.3.4. MÉTODOS BASEADOS EM APRENDIZADO

Os métodos de super-resolução baseados em aprendizado, os quais são também conhecidos como métodos de alucinação, utilizam um passo de aprendizado de máquinas a fim de estimar os detalhes da imagem super-resolvida. Esses métodos consistem em procedimentos baseados nos pixels, em que o valor dos pixels da imagem super-resolvida é estimado via aprendizado estatístico, ou em procedimentos baseados em blocos (também conhecidos como baseados em exemplos), em que a imagem super-resolvida é estimada utilizando-se um dicionário de blocos correspondentes em baixa e alta resolução. Nestes métodos baseados em blocos, após gerado o dicionário, analisa-se a imagem em baixa resolução e busca-se no dicionário os blocos em alta resolução mais prováveis de terem originados os blocos em baixa resolução observados, portanto, é clara a relação entre a qualidade do dicionário gerado e o resultado obtido por esses métodos.

Métodos baseados em blocos ou métodos baseados em exemplos para uma única imagem utilizam um dicionário de exemplos o qual mapeia a relação entre imagens em alta resolução e o seu correspondente em baixa resolução. Em geral, os blocos utilizados são blocos de $n \times n$ pixels com n ímpar (tais como 3×3 , 5×5 ou 7×7). Então, constrói-se um dicionário que relaciona os blocos em alta e baixa resolução, formando pares contendo o bloco em baixa resolução e seu correspondente em alta resolução (Bevilacqua, 2014). A Figura 2.2 mostra o esquema de criação do dicionário de blocos.

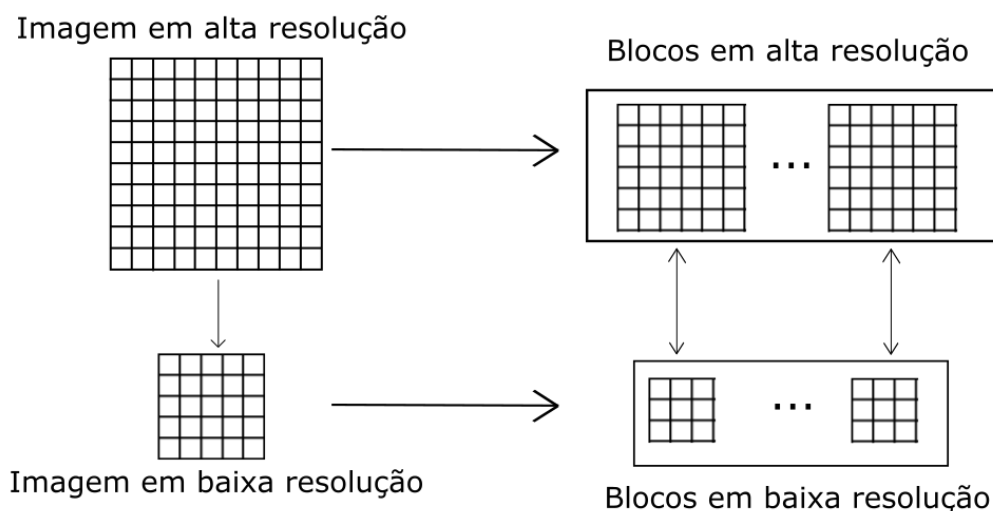


Figura 2.2: Esquema de criação do dicionário de blocos.

Dessa forma, primeiro se realiza a etapa de treinamento, na qual o dicionário de blocos é construído, e, depois, se utiliza o dicionário criado para super-resolver a imagem. A

construção do dicionário pode utilizar um conjunto externo de imagens, ou seja, utilizar imagens que não a que se deseja super-resolver, ou utilizar apenas a própria imagem e suas versões em diversas escalas. Esta última abordagem explora a similaridade existente entre elementos de uma mesma cena, já que é comum que em cenas naturais diversos elementos se repitam. Além disso, na construção do dicionário podem ser utilizados esquemas de um único passo ou esquemas baseados em pirâmide, os quais utilizam diversas escalas das imagens.

Após construído o dicionário, na etapa de se utilizar o dicionário construído para super-resolver a imagem, o que se faz é particionar a imagem em baixa resolução que será super-resolvida em blocos e, para cada um desses blocos, busca-se no dicionário o bloco em alta resolução correspondente. De posse de todos os blocos de alta resolução, a imagem super-resolvida é formada ao se juntar os blocos (Bevilacqua, 2014). A Figura 2.3 ilustra esse processo.

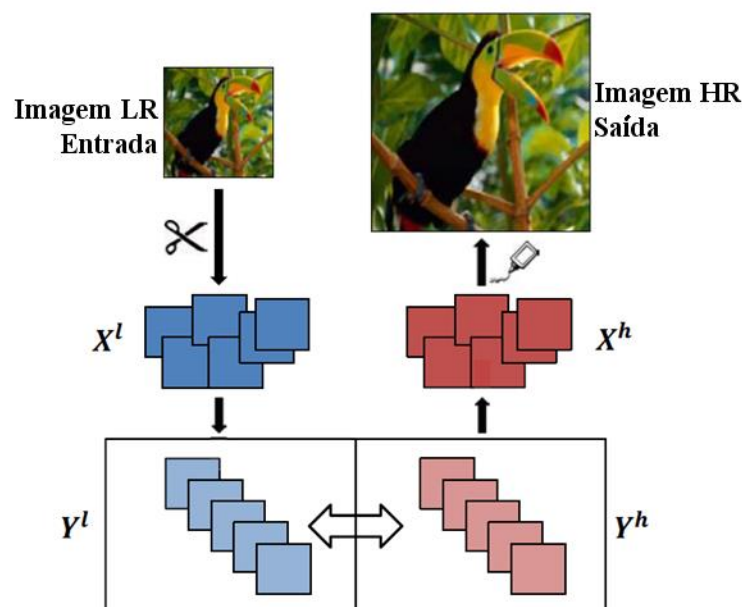


Figura 2.3: Esquema geral de funcionamento de algoritmo baseado em exemplo. Imagem adaptada de Bevilacqua (2014).

2.4. APRENDIZADO DE MÁQUINAS

Aprendizado de máquinas envolve o estudo de reconhecimento de padrões e teoria da aprendizagem computacional. Aprendizado de máquinas foi definido por Arthur Samuel como: “programar computadores para aprender a partir de experiências pode eventualmente eliminar a necessidade de um esforço de programação detalhada” (Samuel, 1959). Já Mitchell definiu aprendizado de máquinas de maneira mais formal: “diz-se que um programa de

computador aprendeu da experiência E com relação a uma classe de tarefas T e medida de desempenho P se o desempenho da tarefa T , medida por P , melhorou com a experiência E ” (Mitchell, 1997). O aprendizado de máquinas trata do estudo e da construção de algoritmos que aprendem a partir de um conjunto de dados e fazem previsão ou classificação para novos dados. Dessa forma, nota-se a característica desses algoritmos de serem dependentes dos dados de treinamento e não apenas de instruções e modelos previamente programados.

Aprendizado de máquinas é vastamente utilizado para se fazer classificações ou previsões de dados, principalmente quando o modelo de previsão ou classificação não é facilmente visualizado explicitamente como em aplicações de reconhecimento de caracteres (OCR), algoritmos de busca e visão computacional.

Os algoritmos de aprendizado de máquinas são normalmente classificados nas seguintes categorias, sob o ponto de vista de utilizar ou não a resposta esperada para o conjunto de dados de treinamento (Stuart & Norvig, 2009):

- **Aprendizado supervisionado:** é apresentado um conjunto de treinamento de entradas com as correspondentes saídas esperadas; dessa forma, o algoritmo visa aprender a regra que mapeia o conjunto de entradas recebidas para as saídas também conhecidas.
- **Aprendizado não-supervisionado:** é apresentado um conjunto de treinamento de entrada sem as correspondentes saídas esperadas; dessa forma, cabe ao algoritmo aprender a estrutura dos dados.
- **Aprendizado por reforço:** o programa interage dinamicamente com o ambiente que realiza certo objetivo, sem um “professor” dizendo explicitamente se o programa está próximo ou não de seu objetivo, como, por exemplo, um programa aprendendo como jogar contra um oponente.

2.4.1. ÁRVORE DE DECISÃO

Árvore de decisão é um dos métodos de aprendizado de máquinas mais vastamente utilizados e práticos para inferência indutiva, tendo sido aplicados com sucesso em diagnósticos médicos, análise de crédito e risco financeiro, entre várias outras áreas (Mitchell, 1997). Uma das grandes vantagens das árvores de decisão é que elas podem ser representadas por um conjunto de regras do tipo “se - então”, o que permite uma compreensão mais fácil da árvore de decisão obtida.

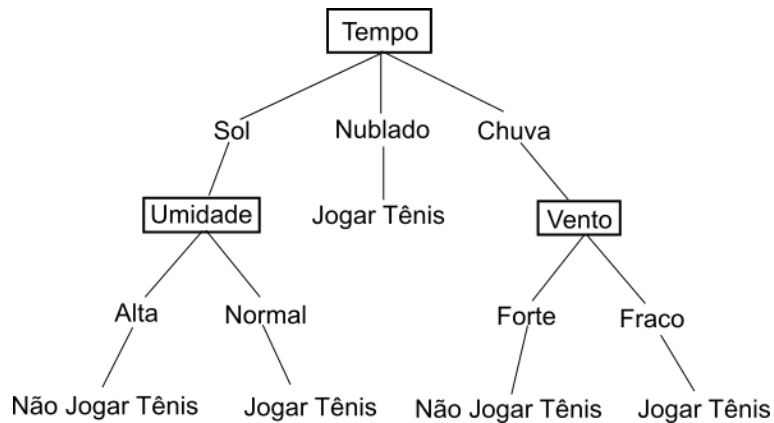


Figura 2.4: Exemplo de árvore de decisão. Imagem adaptada de Mitchell (1997).

Uma árvore de decisão classifica as entradas ordenando-as da raiz até uma de suas folhas, as quais fornecem a classificação para a entrada. Cada nó da árvore especifica um atributo da entrada que será testado e cada ramo corresponde a cada possibilidade de resultado do teste ou do valor atributo testado. A Figura 2.4 apresenta uma árvore de decisão, exemplificando a raiz, os nós e as folhas de uma árvore de decisão.

De acordo com Mitchell (1997), árvores de decisão geralmente apresentam bons resultados em problemas com as seguintes características:

- Instâncias são representadas por pares de valores de atributos. Instâncias são descritas por um conjunto fixo de atributos (por exemplo, vento) e seus valores (por exemplo, forte). A situação mais fácil para uma árvore de decisão é quando os valores dos atributos possuem apenas poucos valores disjuntos (por exemplo, forte ou fraco), entretanto, há algoritmos de árvore de decisão que lidam bem com valores pertencentes ao corpo dos reais.
- A função alvo possui um conjunto discreto de valores. Na árvore da Figura 2.4, a classificação é booleana (jogar ou não jogar tênis). Árvores de decisão lidam bem com um conjunto discreto de valores maior que dois; já para lidar com funções alvo de valores pertencentes ao corpo dos reais é necessária uma alteração mais substancial nos algoritmos geradores das árvores, tal que aplicações desse tipo são menos comuns.
- Pode ser necessário que as descrições sejam disjuntas.

- Os dados de treinamento podem conter erros. Este método de aprendizado de máquinas é robusto a erros, tanto erros na classificação do conjunto de treinamento quanto erros nos atributos das instâncias.
- Os dados de treinamento podem conter valores de atributos faltantes. Árvores de decisão podem ser aplicadas mesmo quando não se conhecem alguns atributos das instâncias de treinamento.

2.5. ÁLGEBRA MATRICIAL

Manipulações de equações envolvendo operadores de subamostragem, de filtragem, entre outros, são rotineiras no processamento digital de imagens. Uma forma de facilitar a manipulação envolvendo tais operadores é utilizar álgebra matricial. A maioria dos operadores utilizados em processamento digital de imagens pode ser representada por uma matriz e, assim, toda a álgebra matricial pode ser aplicada. Ao se modelar um problema de processamento digital de imagens de forma matricial, as imagens costumam ser modeladas como um vetor $I = [i_1, i_2, \dots, i_n]^T$, em que i_1, i_2, \dots, i_n são os valores dos pixels da imagem digital.

Uma aplicação do modelamento por álgebra matricial é a possibilidade do cálculo da derivada envolvendo uma imagem utilizando os teoremas a seguir.

Teorema 2.1 – Seja $u = a^T I = I^T a$, em que $I = [i_1, i_2, \dots, i_n]^T$ e $a = [a_1, a_2, \dots, a_n]^T$ é um vetor de constantes, então:

$$\frac{\partial u}{\partial I} = \frac{\partial(a^T I)}{\partial I} = \frac{\partial(I^T a)}{\partial I} = a$$

Teorema 2.2 – Seja $u = I^T A I$, em que $I = [i_1, i_2, \dots, i_n]^T$ e A uma matriz $n \times n$ de constantes, então:

$$\frac{\partial u}{\partial I} = \frac{\partial(I^T A I)}{\partial I} = (A + A^T)I$$

2.6. MÉTRICAS UTILIZADAS NO TRABALHO

Nesta seção, as métricas utilizadas no trabalho são brevemente descritas. A subseção 2.6.1 define a correlação (também denominada coeficiente de correlação), a subseção 2.6.2 define o SSIM, uma métrica baseada em semelhança estrutural que busca se aproximar do sistema visual humano e subseção 2.6.3 define a clássica métrica RMSE.

2.6.1. CORRELAÇÃO

A correlação (também denominada de coeficiente de correlação), mais especificamente o coeficiente de correlação de Pearson, entre duas imagens é definida pela Eq. 2.16 (Gonzalez & Woods, 2007):

$$\text{correlação}(\mathbf{A}, \mathbf{B}) = \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\mathbf{A}(i, j) - \bar{A})(\mathbf{B}(i, j) - \bar{B})}{\sqrt{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\mathbf{A}(i, j) - \bar{A})^2} \sqrt{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\mathbf{B}(i, j) - \bar{B})^2}} \quad \text{Eq. 2.16}$$

$$\bar{A} = \frac{1}{n_i n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \mathbf{A}(i, j) \quad \text{Eq. 2.17}$$

em que \mathbf{A} e \mathbf{B} são duas imagens com n_i linhas e n_j colunas.

A correlação é uma medida muito utilizada para se verificar se há relação linear entre duas variáveis aleatórias, embora uma alta correlação não implique em relação de causalidade (Aldrich, 1995). No contexto de imagens, espera-se que imagens parecidas apresentem alta correlação (valor próximo de -1 ou 1), enquanto imagens pouco semelhantes apresentem baixa correlação (valor próximo de 0).

2.6.2. SSIM

O SSIM (*Structural Similarity Index*) (Wang, Bovik, Sheikh, & Simoncelli, 2004) é uma métrica de similaridade mais consistente com o sistema visual e cognitivo humano, que se baseia na informação estrutural. O SSIM é dado pela Eq. 2.18:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad \text{Eq. 2.18}$$

em que μ_x e μ_y são as médias das imagens \mathbf{x} e \mathbf{y} , respectivamente, σ_x^2 e σ_y^2 são as variâncias das imagens \mathbf{x} e \mathbf{y} , respectivamente, σ_{xy} é a covariância de \mathbf{x} e \mathbf{y} , c_1 e c_2 são constantes utilizadas para estabilizar a equação quando o denominador $(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)$ é próximo de zero.

O SSIM foi desenvolvido para prever a qualidade de televisão digital e imagens cinematográficas, bem como outros tipos de imagens digitais e de vídeos. No contexto deste trabalho, espera-se que imagens parecidas apresentem alto SSIM (valor próximo de 1), enquanto imagens pouco semelhantes apresentem baixo SSIM (valor próximo de 0).

2.6.3. RMSE

A Raiz do Erro Quadrático Médio (*Root Mean Squared Error* - RMSE) é uma medida vastamente utilizada para cálculo de diferença entre dois sinais. Seu cálculo é feito calculando-se a raiz quadrada da média das diferenças entre os sinais ponto a ponto elevadas ao quadrado, como o próprio nome da medida sugere. A Eq. 2.19 define o RMSE (Gonzalez & Woods, 2007):

$$\text{RMSE}(\mathbf{A}, \mathbf{B}) = \sqrt{\frac{1}{n_i n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\mathbf{A}(i, j) - \mathbf{B}(i, j))^2} \quad \text{Eq. 2.19}$$

em que \mathbf{A} e \mathbf{B} são duas imagens com n_i linhas e n_j colunas.

No contexto deste trabalho, espera-se que imagens parecidas apresentem baixo RMSE (valor próximo de 0), enquanto imagens pouco semelhantes apresentem alto RMSE (valor próximo de 1).

3. MÉTODO PROPOSTO

Neste capítulo, o algoritmo de super-resolução proposto para super-resolver imagens de algarismos individuais oriundos de placas de veículos brasileiros é apresentado. Este capítulo está organizado da seguinte forma: a seção 3.1 mostra a visão geral do método proposto utilizando o fluxograma do método como guia para tal explanação, a seção 3.2 apresenta a normatização brasileira de placas de automóveis e como foram segmentadas as imagens de placas utilizadas no trabalho, a seção 3.3 explica de forma detalhada como foi realizada a criação do dicionário de blocos, a seção 3.4 descreve a criação do conjunto Ω (que armazena os possíveis blocos de alta resolução correspondentes para cada bloco da imagem de baixa resolução), a seção 3.5 apresenta como foi feita a classificação da imagem geral e explora o efeito da manipulação do histograma no desempenho do classificador, a seção 3.6 apresenta o cálculo dos pesos dos blocos, apresentando uma equação que leva em consideração tanto a informação de semelhança local do bloco quanto a informação global da imagem advinda da classificação global da imagem, enquanto a seção 3.7 explica o método Gradiente Descendente utilizado para minimizar a função custo.

3.1. VISÃO GERAL DO MÉTODO

A Figura 3.1 mostra o esquema geral do método proposto. Em linhas gerais, primeiro é necessário segmentar os algarismos da placa em imagens individuais. Tal segmentação pode ser feita de maneira manual ou utilizando algum programa para automatizar esta tarefa. Como o método foi criado visando à aplicação forense, que tipicamente envolve um pequeno volume de imagens e em que normalmente o tempo de processamento não é limitante, assumiu-se que a segmentação dos algarismos será feita de maneira manual antes de se aplicar o método de super-resolução aqui apresentado.

Uma vez segmentados os algarismos da placa em imagens distintas, cada imagem de algarismo de baixa resolução (LR) é fornecida como entrada do algoritmo. Realiza-se um preenchimento na imagem de entrada utilizando-se a cor de fundo como valores dos pixels de preenchimento. Em seguida, faz-se a primeira estimativa da imagem em alta resolução (HR) utilizando-se interpolação bicúbica e se redimensionam as Imagens Padrões (imagens em alta resolução dos algarismos 0, 1, 2, ..., 8, 9) para as mesmas dimensões da imagem HR e, nesta imagem redimensionada, é aplicado o mesmo processo de preenchimento aplicado à imagem LR de entrada.

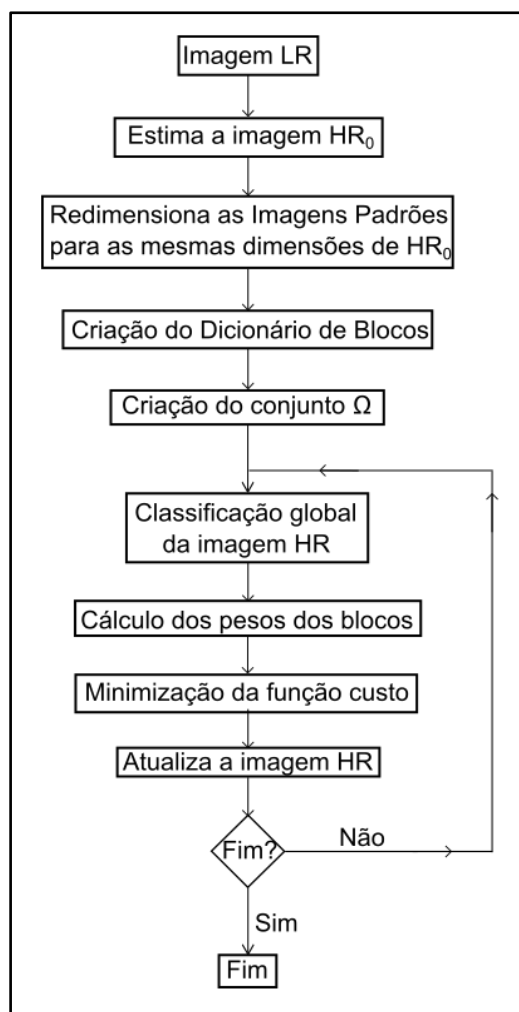


Figura 3.1: Fluxograma geral do método proposto

O passo seguinte consiste na criação do dicionário de blocos. Os blocos de baixa resolução do dicionário de blocos são obtidos aplicando-se o filtro Blur e o operador de subamostragem, ambos da Eq. 3.1, às Imagens Padrões. O dicionário de blocos é obtido fazendo-se a associação entre os blocos de alta resolução e os blocos de baixa resolução. A seção 3.3 detalha a criação do dicionário de blocos.

De posse do dicionário de blocos, cria-se o conjunto Ω , o qual, para cada bloco da imagem de baixa resolução, armazena os possíveis blocos de alta resolução correspondentes. A determinação de quais blocos de alta resolução podem corresponder ao bloco da imagem de baixa resolução observada é feita por meio de uma medida de similaridade entre o bloco de baixa resolução da imagem a ser super-resolvida e todos os blocos de baixa resolução pertencentes ao dicionário previamente construído. A seção 3.4 detalha a criação do conjunto Ω .

Então, aplica-se um classificador na imagem HR estimada até então. O classificador escolhido para ser utilizado no algoritmo foi uma árvore de decisão. Dessa forma, a imagem HR estimada é fornecida para a árvore de decisão e o classificador fornece como retorno tanto a classificação obtida quanto as probabilidades de a classificação estar correta obtida no treinamento. A título de exemplo, as Figuras 3.2 e 3.3 ilustram a árvore de decisão obtida na seção 4.2, a partir de um treinamento com 300 imagens com SNR de 5 dB, 10 dB e 15 dB.

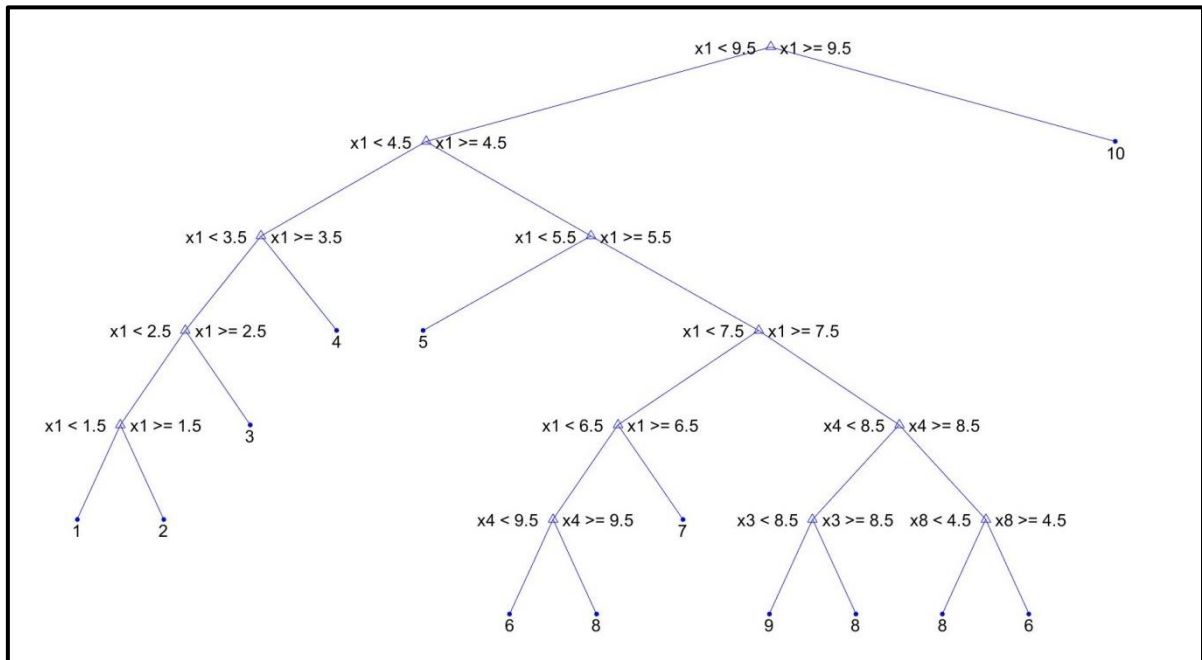


Figura 3.2: Árvore de decisão obtida na seção 4.2. Observação: “10” refere-se ao algoritmo 0.

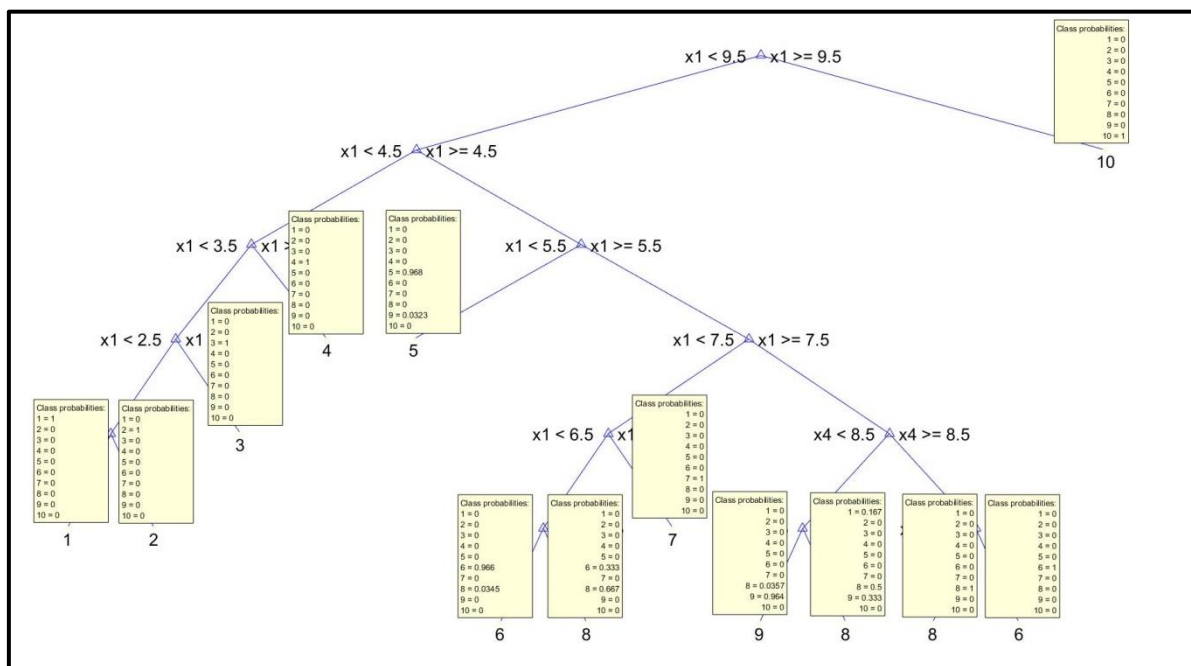


Figura 3.3: Árvore de decisão obtida na seção 4.2 com as probabilidades das folhas. Observação: “10” refere-se ao algoritmo 0.

A Figura 3.3 mostra as probabilidades das folhas da árvore. Essas probabilidades são utilizadas no algoritmo de super-resolução, na etapa de teste, como aproximação para a probabilidade da classificação estar correta. Por exemplo, se a imagem que está sendo super-resolvida segue a árvore de decisão e termina na folha relativa ao “5”, considera-se que a probabilidade do algarismo ser “5” é 0,968 e a probabilidade do algarismo ser “9” é 0,0323.

De posse dos resultados obtidos no classificador, calculam-se os pesos para os blocos que devem ser utilizados na Eq. 3.1. Então, minimiza-se a Eq. 3.1, a qual funciona como uma função de custo. Para esta tarefa de minimização foi utilizado o algoritmo Gradiente Descendente. Encontrado o valor da imagem HR estimada que minimiza a equação Eq. 3.1, atualiza-se a imagem HR estimada e verifica-se se o critério de parada do algoritmo foi atingido (número máximo de iterações); em caso positivo, finaliza-se o algoritmo e a imagem super-resolvida é a imagem HR estimada; em caso negativo, retorna-se ao passo da classificação da imagem HR estimada.

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \begin{aligned} & \|\mathbf{LR} - (\mathbf{Blur} * \mathbf{HR}) \downarrow\|_2^2 \\ & + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} \mathbf{w}(i,j,k) \|\mathbf{R}(i,j)\mathbf{HR} - \mathbf{x}(i,j,k)\|_2^2 \end{aligned} \right\} \quad \text{Eq. 3.1}$$

em que $\widehat{\mathbf{HR}}$ é a imagem \mathbf{HR} estimada, \mathbf{LR} imagem em baixa-resolução (entrada do algoritmo), \mathbf{Blur} é o filtro que representa o espalhamento/borramento e as distorções de lente e de foco do sistema que capturou a imagem \mathbf{LR} , \downarrow é o operador de subamostragem, λ é o parâmetro de regularização que controla a contribuição relativa entre o fator de fidelidade à entrada ($\|\mathbf{LR} - (\mathbf{Blur} * \mathbf{HR}) \downarrow\|_2^2$) e o fator de aprendizado ($\sum_{i,j} \sum_{k \in \Omega(i,j)} \mathbf{w}(i,j,k) \|\mathbf{R}(i,j)\mathbf{HR} - \mathbf{x}(i,j,k)\|_2^2$), (i,j) é a posição do pixel em baixa resolução correspondente ao canto superior esquerdo do bloco de baixa resolução, Ω é um conjunto que, para cada bloco da imagem de baixa resolução, armazena os possíveis blocos de alta resolução correspondentes (ver seção 3.4 para maiores detalhes), $\mathbf{w}(i,j,k)$ são os pesos calculados para cada bloco de alta resolução (índice k) pertencente ao $\Omega(i,j)$, $\mathbf{R}(i,j)$ é o operador que seleciona o bloco de alta resolução (tamanho $m \times m$) da imagem \mathbf{HR} na posição correspondente a (i,j) de baixa resolução (é necessário converter a posição (i,j) no seu correspondente na imagem em alta resolução), $\mathbf{x}(i,j,k)$ é o k -ésimo bloco de alta resolução pertencente ao $\Omega(i,j)$.

A função custo (Eq. 3.1) foi escolhida por permitir um equilíbrio, controlado por λ , entre a fidelidade e o aprendizado e por permitir que os pesos dos blocos utilizados na

aprendizagem sejam calculados e suavizados levando em consideração tanto a similaridade entre os blocos quanto à similaridade/classificação global da imagem. Além disso, a utilização da norma L_2 torna a equação diferenciável, o que é desejável para diversos métodos de minimização.

3.2. PLACAS BRASILEIRAS E A SEGMENTAÇÃO DOS ALGARISMOS

Placas de veículos automotores são padronizadas, porém, tal padronização varia de país para país e, às vezes, de estado para estado dentro de um mesmo país. No Brasil, a padronização das placas é nacional e regulamentada pelo Conselho Nacional de Trânsito (CONTRAN). As principais resoluções do CONTRAN relativas às placas são:

- Resolução 231: esta resolução é a base das especificações das placas brasileiras, trazendo definições de forma e tamanho tanto das placas quanto dos caracteres presentes nas placas. As informações mais relevantes para o presente trabalho são: a fonte dos caracteres é a “Mandatory”, a altura dos caracteres para placas de automóveis é de 63 mm, a espessura do traço é de 10 mm e a largura dos caracteres segue tabela apresentada na resolução.
- Resolução 241: esta resolução altera algumas especificações da resolução 231, porém, mantém a fonte dos caracteres como sendo a “Mandatory” e a altura dos caracteres para placas de automóveis em 63 mm.
- Resolução 309: esta resolução altera as dimensões das placas, mas não altera as dimensões dos caracteres.
- Resolução 372: esta resolução altera as dimensões das placas e dos caracteres de placas para motocicletas, motonetas, ciclomotores e triciclos motorizados, porém, não alterou as especificações de placas para automóveis, que é o foco deste trabalho.
- Resolução 510: esta resolução altera completamente a resolução 231 e suas alterações e reformula completamente a forma e especificações das placas brasileiras para compatibilizá-las com um padrão comum do MERCOSUL, alterando, inclusive, a fonte dos caracteres para a fonte “FE Engschrift” e a altura dos caracteres para 65 mm e 53 mm para automóveis e motocicletas, respectivamente. Porém, como esta resolução só entrará em vigor em 2017, neste

trabalho esta resolução não foi considerada, já que não se encontram disponíveis placas que seguem esta especificação para testes.

O fato de as placas serem padronizadas e utilizarem uma fonte conhecida (a Figura 3.4 exemplifica a fonte “Mandatory”) proporciona um ambiente favorável para a aplicação de um algoritmo de super-resolução que explore as características de um domínio específico de aplicação.

**0123456789
ABCDEFGHIJKLM
NOPQRSTUVWXYZ**

Figura 3.4: Caracteres com a fonte “Mandatory”.

Utilizando-se a fonte “Mandatory” são geradas dez Imagens Padrões em alta resolução, cada uma contendo um único algarismo em fonte “Mandatory” (0, 1, 2, ..., 8, 9).

O método proposto neste trabalho aceita, como entrada, uma imagem contendo apenas um único algarismo de placa de veículo. Dessa forma, é necessário segmentar os algarismos de uma placa antes de submetê-los ao algoritmo de super-resolução proposto. Essa segmentação pode ser feita manualmente ou com o auxílio de alguma rotina computacional. Neste trabalho, a segmentação foi realizada manualmente e foram seguidas as seguintes regras: não incluir o que se pôde identificar como pixel oriundo dos algarismos vizinhos e utilizar uma “margem” de 1 ou 2 pixels em relação ao que se identificou como sendo de fato o algarismo. A Figura 3.6 ilustra a segmentação realizada.



Figura 3.5: Imagem em baixa resolução.

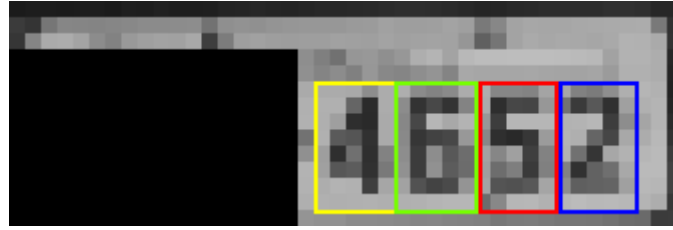


Figura 3.6: Exemplo de segmentação dos algarismos de uma placa em baixa resolução. A placa foi extraída da imagem ilustrada na Figura 3.5 e os retângulos coloridos mostram a segmentação de cada um dos quatro algarismos. Cada retângulo gerou uma imagem individual de baixa resolução.

3.3. CRIAÇÃO DO DICIONÁRIO DE BLOCOS

A criação de blocos é feita de acordo com o esquema mostrado na Figura 3.7.

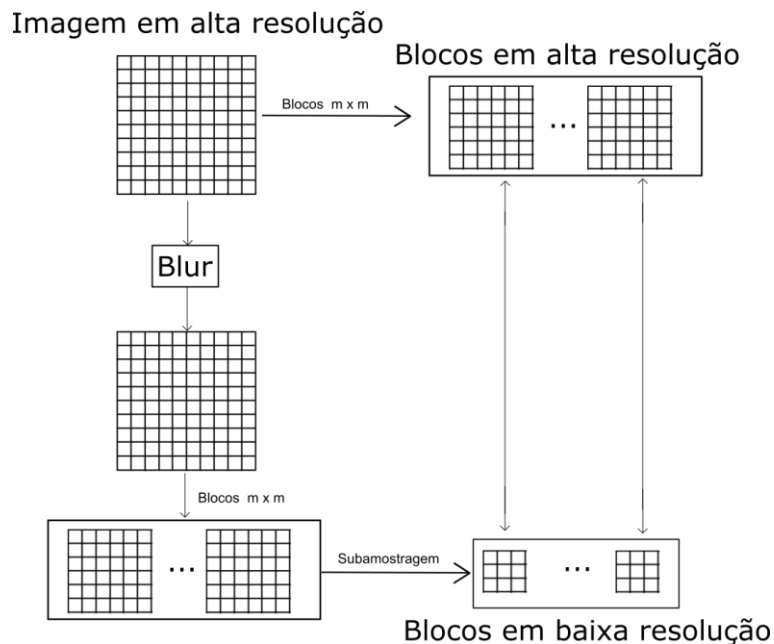


Figura 3.7: Esquema de criação do dicionário de blocos utilizado no método proposto.

Dada uma imagem de alta resolução, criam-se os blocos de alta resolução de dimensões $m \times m$ simplesmente copiando os pixels pertencentes ao bloco de dimensões $m \times m$ (por exemplo, o bloco 1 é composto pelos pixels $(1:m-1, 1:m-1)$, o bloco 2 é composto pelos pixels $(2:m, 1:m-1)$, e assim por diante).

Para criar os blocos de baixa resolução, primeiro se aplica o filtro Blur (filtro que representa o espalhamento/borramento e as distorções de lente e de foco do sistema que capturou a imagem LR) da Eq. 3.1. Então, criam-se blocos de dimensões $m \times m$ de forma análoga à supracitada a partir da imagem resultante da aplicação do filtro Blur. A estes blocos se aplica o operador de subamostragem (\downarrow) da Eq. 3.1 e se obtém os blocos de baixa

resolução. Por fim, o dicionário é construído fazendo-se a associação entre os blocos de baixa resolução e seu bloco de alta resolução correspondente.

Para a criação do dicionário de blocos, cada uma das Imagens Padrões é utilizada como imagem de alta resolução de entrada na rotina descrita nesta seção. Os blocos gerados por todas as Imagens Padrões compõem o dicionário de blocos final.

3.4. CRIAÇÃO DO CONJUNTO Ω

O conjunto Ω é um conjunto que, para cada bloco da imagem de baixa resolução, armazena os possíveis blocos de alta resolução correspondentes. A determinação de quais blocos de alta resolução podem corresponder ao bloco da imagem de baixa resolução observada é feita por meio de uma medida de similaridade (correlação) entre o bloco de baixa resolução da imagem a ser super-resolvida e todos os blocos de baixa resolução pertencentes ao dicionário previamente construído. Caso a medida de similaridade seja maior que um limiar (gatilho) previamente definido (T), então, o bloco de alta resolução correspondente ao bloco de baixa resolução pertencente ao dicionário é incluído no conjunto Ω como possível bloco de alta resolução correspondente. Para cada posição (i, j) da imagem em baixa resolução, mantém-se no conjunto apenas os n Posição blocos com as maiores medidas de similaridade (correlação).

Assim, o conjunto Ω pode ser representado pela Eq. 3.2:

$$\Omega(i, j) = \left\{ \begin{array}{l} \{x_{m \times m}^1, x_{m \times m}^2, \dots, x_{m \times m}^{k_0}\}_{\text{padrão } 0'} \\ \{x_{m \times m}^1, x_{m \times m}^2, \dots, x_{m \times m}^{k_1}\}_{\text{padrão } 1'} \\ \dots \\ \{x_{m \times m}^1, x_{m \times m}^2, \dots, x_{m \times m}^{k_9}\}_{\text{padrão } 9} \end{array} \right\} \quad \text{Eq. 3.2}$$

em que (i, j) é a posição do pixel em baixa resolução correspondente ao canto superior esquerdo do bloco de baixa resolução, $x_{m \times m}^1, x_{m \times m}^2, \dots, x_{m \times m}^{k_j}$ são os blocos de alta resolução pertencentes ao dicionário cujo seu correspondente de baixa resolução satisfaz o critério de similaridade adotado. Por questões de implementação, é útil separar os blocos de alta resolução de acordo com o padrão que o originou (0, 1, 2, ..., 9).

No método proposto, a medida de similaridade adotada para este passo foi a correlação definida na Eq. 2.16, pois seu cálculo é bastante rápido, característica bastante desejável já que tal medida será calculada muitas vezes nesta etapa.

Além disso, para evitar um conjunto Ω excessivamente grande, limitou-se o número de elementos pertencentes ao conjunto Ω a N_Ω . Assim, caso o número de elementos do conjunto Ω seja maior que N_Ω , descartam-se os elementos com menores medidas de similaridade até que só restem N_Ω elementos.

A seguir, os principais passos da criação do conjunto Ω são apresentados na forma de um pseudocódigo.

Criação do conjunto Ω	
entradas:	LR: imagem em LR n : dimensão do bloco () em baixa resolução T : gatilho de similaridade $nPosição$: quantidade $n \times n$ máxima de blocos para cada posição (i, j) da imagem em LR N_Ω : quantidade máxima de blocos do conjunto Ω
saída:	conjunto Ω
<pre> for i = 1 até total de linhas da imagem LR - n + 1 do for j = 1 até total de colunas LR - n + 1 do blocoLRAtual ← bloco LR cujo pixel do canto superior esquerdo é (i, j) for padrão = 1 até número de padrões do k ← 0 for blocoBibliotecaLR ∈ Biblioteca(padrão) S ← medida_de_Similaridade(padrão, blocoBibliotecaLR) if S > T $\Omega(i, j, k) \leftarrow [blocoBibliotecaLR, S]$ k ← k+1 end end end end end for i = 1 até total de linhas da imagem LR - n + 1 do for j = 1 até total de colunas LR - n + 1 do if quantidade de elementos em $\Omega(i, j, :)$ > $nPosição$ Ordena $\Omega(i, j, :)$ de acordo com S (medidas de similaridades) $\Omega(i, j, :)$ ← apenas os $nPosição$ primeiros elementos end end end if número de elementos de Ω > N_Ω Ordena Ω de acordo com S (medidas de similaridades) Ω ← apenas os N_Ω primeiros elementos end </pre>	

3.5. CLASSIFICAÇÃO DA IMAGEM GERAL

Para a tarefa de classificação da imagem, utilizou-se um classificador do tipo árvore de decisão. Este classificador foi escolhido pois seu resultado é de fácil interpretação e explicação, característica esta desejável e útil para a área criminalística, uma vez que os resultados das ciências forenses são destinados às pessoas leigas, usualmente da área do Direito. Conseqüentemente, não é raro que se precise explicar a técnica utilizada para esses profissionais de outras áreas. Além disso, a árvore de decisão fornece facilmente, para cada folha, as probabilidades de cada classe obtidas no conjunto de treinamento, informações essas que são utilizadas como pesos ($w(i, j, k)$ da Eq. 3.1) dos blocos do conjunto Ω .

Como descritores das imagens para o classificador, utilizou-se a ordem decrescente de uma métrica de similaridade entre a imagem a ser super-resolvida e cada uma das imagens padrões. A seguir é apresentado um pseudocódigo da obtenção dos descritores.

Obtenção dos descritores para o classificador	
entradas:	IPresize{i}: Imagem padrão referente ao padrão “i” redimensionada para dimensões iguais à HR_image HR_image: imagem para a qual os descritores estão sendo calculados
saída:	descritores
for i = 1 até número de padrões do similaridade(i) ← medida_de_Similaridade(HR_image, IPresize{i})	
end [similaridadeOrdenado, sortIndex] ← sort(similaridade, 'descend') % sortIndex(i) corresponde à ordem que similaridade(i) ocupado no vetor ordenado descritores ← sortIndex;	

Para ilustrar a obtenção dos descritores, suponha uma imagem (HR_image) cujo vetor similaridade para uma dada métrica de similaridade tenha resultado no seguinte vetor: 0,12, 0,85, 0,34, 0,22, 0,15, 0,36, 0,72, 0,11, 0,28, 0,30, para $i=1, 2, \dots, 10$ respectivamente. Então, o vetor ordenado de maneira descendente (similaridadeOrdenado) seria: 0,85, 0,72, 0,36, 0,34, 0,30, 0,28, 0,22, 0,15, 0,12, 0,11 e o sortIndex seria: 2, 7, 6, 3, 10, 9, 4, 5, 1, 8.

Comparou-se o desempenho do classificador utilizando como métrica de similaridade a correlação e o valor de SSIM (*Structural Similarity*), conforme detalhado na seção 4.1. Além disso, na seção 4.1 também foi estudado o impacto da manipulação de histograma antes do cálculo dos descritores.

Dos resultados obtidos na seção 4.1, pode-se concluir que o SSIM demonstrou ser uma medida mais promissora do que a correlação como medida de similaridade para fins de obter

os descritores para a classificação da árvore de decisão. Desta forma, nesta etapa do método proposto foi utilizado o SSIM como métrica de similaridade.

A árvore de decisão obtida na seção 4.3, em que o treinamento da árvore foi feito utilizando imagens de Algarismos de Placas de Automóveis Reais de baixa resolução com 180 imagens sendo 18 de cada algarismo, é ilustrada na Figura 3.8.

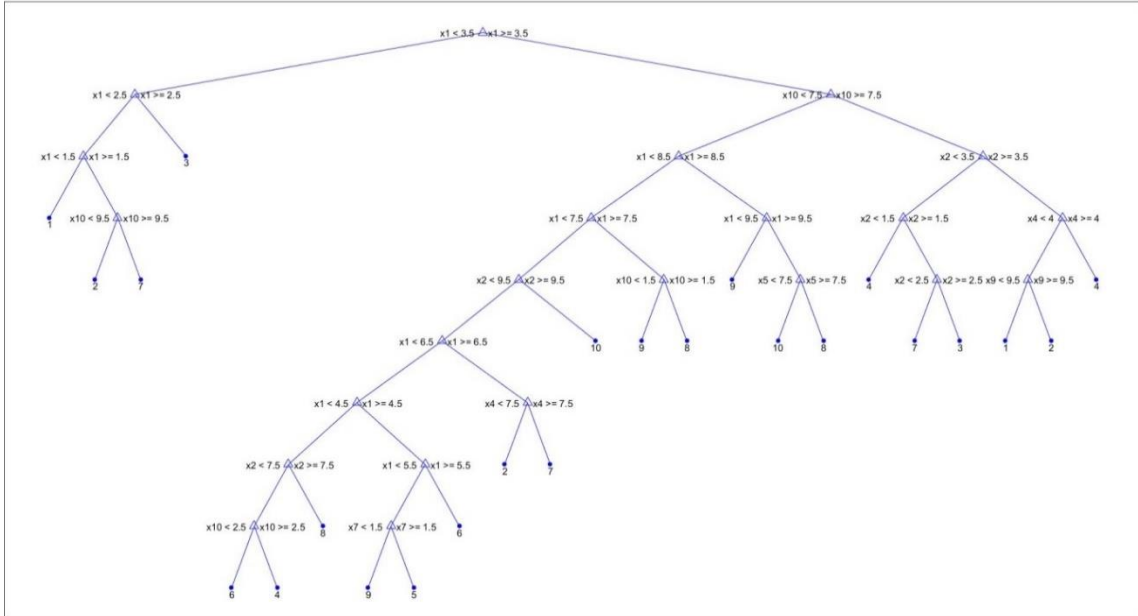


Figura 3.8: Árvore de decisão obtida na seção 4.3. Observação: “10” refere-se ao algarismo 0.

3.6. CÁLCULO DOS PESOS DOS BLOCOS

O cálculo dos pesos $w(i, j, k)$ é feito de acordo com a Eq. 3.3:

$$w(i, j, k) = P(r) \frac{S(k)}{\sum_{n \in \Omega_r(i, j)} S(n)} \quad \text{Eq. 3.3}$$

em que $w(i, j, k)$ é o peso, $P(c), c = 0, 1, \dots, 9$, a probabilidade de cada classe obtida do classificador no conjunto de treinamento correspondente à folha da classificação da imagem em alta resolução estimada, r o padrão que originou o elemento $k \in \Omega(i, j)$, $S(k)$ a medida de similaridade calculada para o bloco k na construção do conjunto Ω e $\Omega_r(i, j)$ o subconjunto de $\Omega(i, j)$ formado apenas pelos blocos originados pelo padrão r , ou seja, $\Omega_r(i, j) = \{\{x_{m \times m}^1, x_{m \times m}^2, \dots, x_{m \times m}^{kr}\}_{padr\tilde{a}o\ r}\}$. Como a medida de similaridade adotada na construção do conjunto Ω foi a correlação, então $S(k)$ corresponde à correlação.

Esta equação de pesos foi escolhida por utilizar a informação da semelhança entre o bloco da imagem de baixa resolução e o bloco oriundo do dicionário de blocos (termo $\frac{S(k)}{\sum_{n \in \Omega_r(i,j)} S(n)}$) ponderada pela informação de classificação global da imagem (termo $P(r)$).

3.7. MINIMIZAÇÃO DA FUNÇÃO CUSTO

Para determinar o mínimo da função custo, utilizou-se o método do Gradiente Descendente. Este método foi escolhido porque ele é de fácil implementação, já que a função custo é diferenciável. Como o objetivo do método proposto é demonstrar a funcionalidade e as possibilidades de um algoritmo de super-resolução baseado em aprendizado e dicionário de blocos para um domínio específico, e não obter um algoritmo otimizado para ser utilizado em alguma aplicação final, não foram investigados outros algoritmos para calcular o mínimo da função custo.

O método do Gradiente Descendente é um método de otimização de primeira ordem baseado no fato do gradiente de uma função apontar para a direção de maior crescimento da função. Consequentemente, a direção de maior decaimento é a direção oposta do gradiente. Dessa forma, o algoritmo busca o mínimo da função de custo seguindo a direção oposta do gradiente em pequenos passos.

Mais formalmente, seja $F(\mathbf{x}): \mathbb{R}^n \rightarrow \mathbb{R}$, diferenciável. Para qualquer ponto \mathbf{x} , a direção de maior decaimento da função é dada por $-\nabla F(\mathbf{x})$. Então, seja $\gamma \in \mathbb{R}$ pequeno o suficiente, se andarmos $-\gamma \nabla F(\mathbf{x}_0)$ a partir de \mathbf{x}_0 chegaremos a um ponto em que F é menor que $F(\mathbf{x}_0)$. Ou seja, fazendo $\mathbf{x}_1 = \mathbf{x}_0 - \gamma \nabla F(\mathbf{x}_0)$, temos que $F(\mathbf{x}_0) \geq F(\mathbf{x}_1)$. Então, expressando esse processo iterativamente, tem-se:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma \nabla F(\mathbf{x}_n), n \geq 0 \quad \text{Eq. 3.4}$$

$$F(\mathbf{x}_0) \geq F(\mathbf{x}_1) \geq F(\mathbf{x}_2) \geq \dots \quad \text{Eq. 3.5}$$

Para um número suficiente de iterações, a solução converge para um mínimo local. Sob certas condições da função F e do passo γ (que pode variar a cada iteração), pode-se garantir a convergência ao mínimo local. A Figura 3.9 ilustra graficamente o funcionamento do método do Gradiente Descendente.

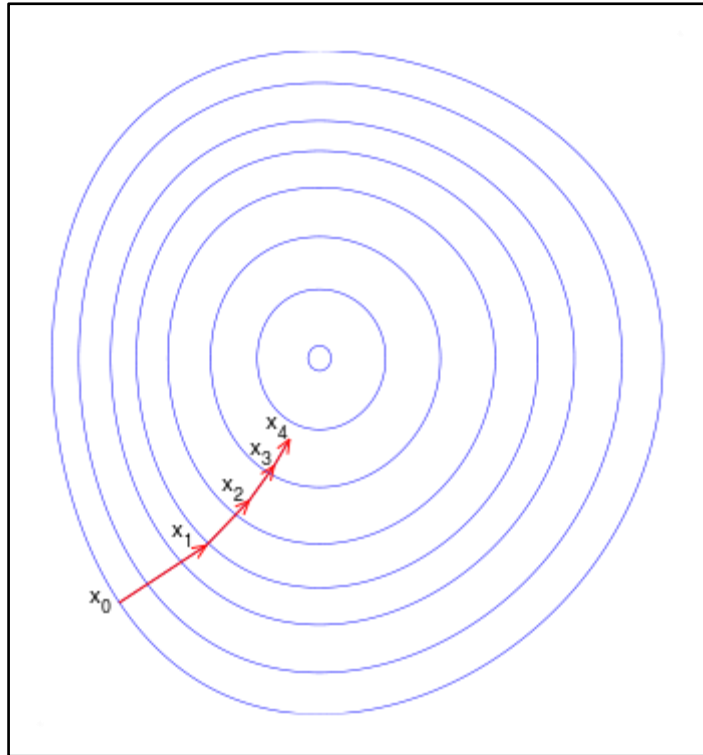


Figura 3.9: Ilustração do funcionamento do método do Gradiente Descendente.

O gradiente da função custo definida na Eq. 3.1 se resume à derivada da função custo em relação à HR. O cálculo desta derivada é detalhadamente apresentado no Anexo A – Cálculos para a minimização da função custo.

4. RESULTADOS EXPERIMENTAIS

Este capítulo apresenta e discute os resultados experimentais obtidos com o método proposto. Na seção 4.1, é investigado o efeito da manipulação de histograma e da métrica de similaridade adotada no desempenho do classificador do tipo árvore de decisão apresentado na seção 2.4.1. Na seção 4.2, são mostrados os resultados do algoritmo quando aplicado às imagens sintéticas de baixa resolução obtidas a partir de imagens de alta resolução conhecidas. Neste cenário, como se conhece as imagens em alta resolução que originaram as versões em baixa resolução é possível demonstrar quantitativamente, por meio das métricas RMSE e SSIM, a melhora obtida pelo método proposto. Já na seção 4.3, são mostrados resultados do algoritmo quando aplicados às imagens reais, capturadas originalmente em baixa resolução. Dessa forma, como não se possui as imagens de referência em alta resolução, neste cenário apenas pode-se avaliar os resultados em termos qualitativos.

4.1. EFEITO DA MANIPULAÇÃO DE HISTOGRAMA NO DESEMPENHO DA ÁRVORE DE DECISÃO

A fim de investigar o efeito da manipulação de histograma e da métrica de similaridade adotada no desempenho do classificador do tipo árvore de decisão apresentado na seção 2.4.1 e, conseqüentemente, no desempenho do método de super-resolução proposto, analisou-se o comportamento da árvore de classificação para um fator de escala igual a quatro utilizando validação cruzada, mais especificamente utilizando-se LOOCV (*Leave-One-Out Cross-Validation*) e banco de imagens de algarismos de placas de automóveis reais de baixa resolução composto por 200 imagens, sendo 20 imagens de cada algarismo (0, 1, ..., 9) descrito na seção 4.3. Os seguintes cenários foram submetidos ao LOOCV e foram obtidas as respectivas taxas de erro médias de classificação:

1. Sem nenhuma manipulação de histograma.
2. Manipulação de histograma do tipo equalização de histograma (Pedrini & Schwartz, 2007), aplicado às imagens de treinamento e teste da árvore de classificação.
3. Manipulação de histograma do tipo equalização de histograma aplicado tanto às imagens de treinamento e teste da árvore de classificação quanto às imagens padrões.

4. Manipulação de histograma do tipo aumento de contraste (utilizando a função $imadjust([low_in\ high_in], [low_out\ high_out], gamma)$ com os seguintes parâmetros $imadjust(Imagem, [min(min(Imagem)); max(max(Imagem))], [0; 1], 0)$ aplicado tanto às imagens de treinamento e teste da árvore de classificação quanto às imagens padrões.
5. Manipulação de histograma do tipo normalização da imagem (divide-se os valores dos pixels pela norma L_1 da imagem) aplicado tanto às imagens de treinamento e teste da árvore de classificação quanto às imagens padrões.

Cada um desses cenários foi testado utilizando-se imagens padrões com fundo branco (IP Fundo Branco) e imagens padrões com fundo cinza (IP Fundo Cinza), ilustradas na Figura 4.1, e utilizando-se como descritores a ordem decrescente da medida de similaridade correlação e da medida de similaridade SSIM. As taxas de erro média obtidas são mostradas nas Tabelas 4.1 e 4.2.



Figura 4.1: As imagens padrões com fundo branco e com fundo cinza são mostradas, respectivamente, na linha superior e inferior.

Tabela 4.1 – Taxas de erro de classificação média utilizando-se a correlação como medida de similaridade.

Cenário	IP Fundo Branco	IP Fundo Cinza
1	46,0 %	45,0 %
2	35,5 %	38,5 %
3	38,5 %	30,5 %
4	46,0 %	44,5 %
5	46,0 %	45,0 %

Tabela 4.2 – Taxas de erro de classificação média utilizando-se a SSIM como medida de similaridade.

Cenário	IP Fundo Branco	IP Fundo Cinza
1	48,5 %	43,5 %
2	19,5 %	21,0 %
3	30,5 %	33,5 %
4	37,0 %	35,0 %
5	45,5 %	52,5 %

Das taxas de erro de classificação média obtidas, conclui-se que o melhor desempenho do classificador ocorre quando se manipula o histograma por meio de equalização de histograma aplicado às imagens de treinamento e teste da árvore de classificação (ou seja, sem aplicar equalização de histograma às imagens padrões), quando se utilizam imagens padrões com fundo branco e quando se utiliza a medida de similaridade SSIM.

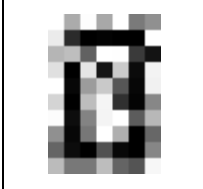

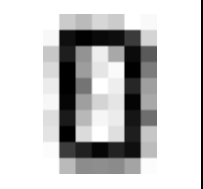



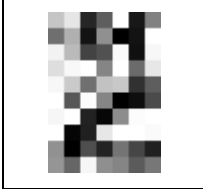
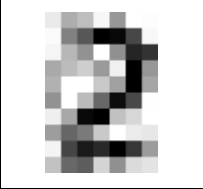
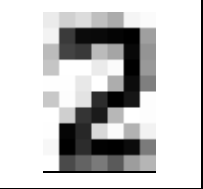
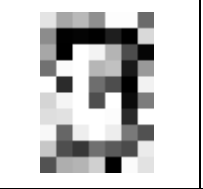
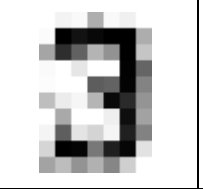
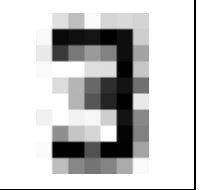
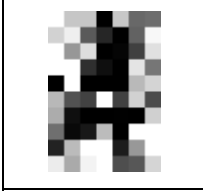
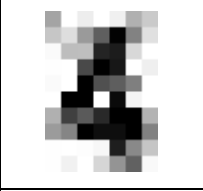
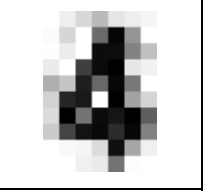
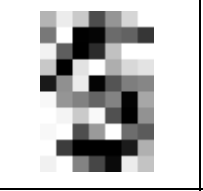
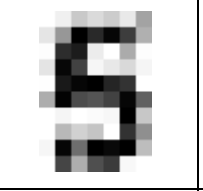
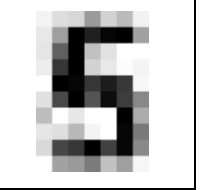
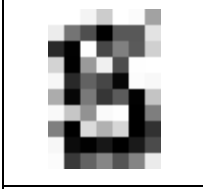
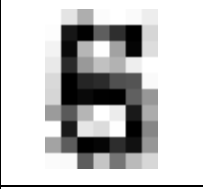
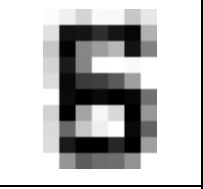
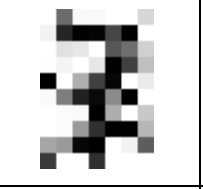
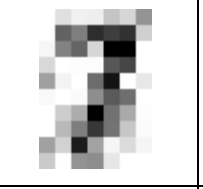
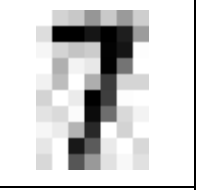
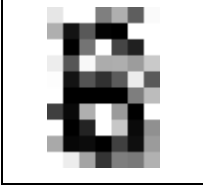
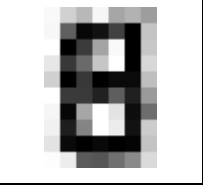
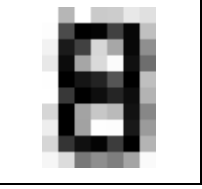
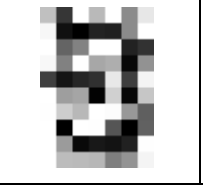
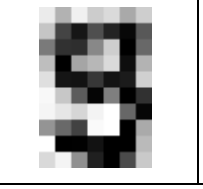
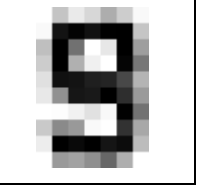
4.2. IMAGENS SIMULADAS

Apesar do objetivo principal do método proposto é ser aplicado a imagens de placas de automóveis brasileiros reais de baixa resolução em ambiente não controlado (incidência e iluminação não controladas), cenário no qual não se tem a imagem de alta resolução para comparação do resultado obtido e uma referência de alta resolução, para fins de comparação com outros métodos torna-se desejável que o método seja aplicado em um cenário controlado, em que se tenham imagens de baixa resolução e os seus correspondentes em alta resolução para referência. Nesta seção, foi realizada a simulação de tal cenário controlado.

A partir de imagens dos algarismos de 0 a 9 em alta resolução (dimensões: 40 pixels de altura e 28 pixels de largura) foram geradas imagens em baixa resolução (dimensões: 10 pixels de altura e 7 pixels de largura) com diferentes valores de relação sinal-ruído. Para se obter as diferentes relações de sinal-ruído, adicionou-se um ruído gaussiano branco às imagens de baixa resolução. Os valores da relação sinal-ruído utilizados foram 5 dB, 10 dB e 15 dB. Para cada um desses valores de relação sinal ruído, foram geradas 20 imagens de baixa resolução para cada algarismo de 0 a 9, resultando num total de 600 imagens de baixa resolução. As 600 imagens de baixa resolução foram divididas em dois grupos, treinamento e teste, cada um com 300 imagens com distribuição igual com relação ao algarismo e as relações sinal ruído. Esse banco de imagens está disponível em

<https://github.com/alevicente/SR-LicensePlate/tree/master/Datasets/> sob o nome 4.A – Synthetic Images. A Figura 4.2 ilustra algumas imagens desse banco de imagens criado.

Figura 4.2 – Amostra de imagens em baixa resolução do banco de imagens simuladas (as imagens apresentadas foram ampliadas por um fator de 8 pelo método do vizinho mais próximo para facilitar a visualização).



















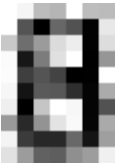








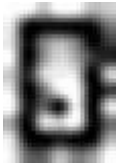

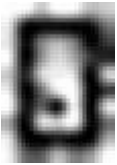





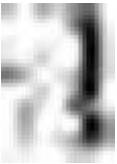






SNR = 5 dB	SNR = 10 dB	SNR = 15 dB	SNR = 5 dB	SNR = 10 dB	SNR = 15 dB
					
					
					
					
					

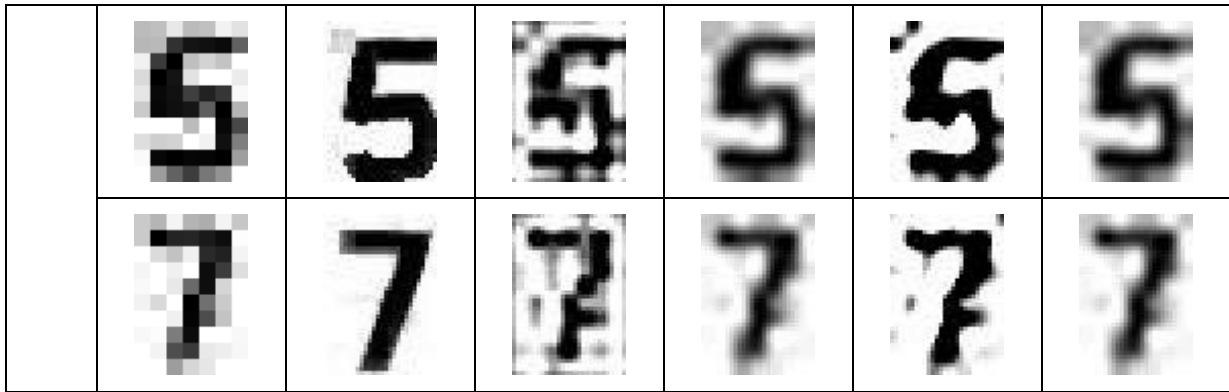
O classificador foi treinado no conjunto de treinamento, e depois o método proposto foi aplicado ao conjunto de teste. Para fins de comparação, também foram aplicados ao conjunto de teste os seguintes métodos: *Iterative Back Projection* (IBP) (Irani & Peleg, 1991), o *Kernel Ridge Regression* (KRR) (Kim & Kwon, 2008), *Gradient Profile Prior* (GPP) (Sun, Xu, & Shum, 2008) e a interpolação bicúbica.

Nesta seção, o método proposto utilizou as seguintes configurações: imagens padrões de fundo branco, fator de escala: 4, tamanho dos blocos em alta definição: 16×16 , filtro de Blur: filtro média 3×3 , $T = 0,8$, $n_{Posição} = 3$, número máximo de blocos $n = 1000$ para o

conjunto Ω , $\lambda = 0,2$, número máximo de iterações $N = 20$. Uma amostra dos resultados obtidos é ilustrada na Figura 4.3.

Figura 4.3 – Amostra do resultado das imagens do conjunto de teste super-resolvidas pelo método proposto, IBP, KRR, GPP e interpolação bicúbica, todas por um fator de ampliação de 4. Para facilitar a visualização, os resultados obtidos pelos métodos citados foram ampliados por um fator de 2 utilizando o método do vizinho mais próximo e a imagem original em baixa resolução foi ampliada por um fator de 8 utilizando o método do vizinho mais próximo.

SNR [dB]	Vizinho mais próximo	Método proposto	IBP	KRR	GPP	Interpolação Bicúbica
5						
						
						
10						
						
						
15						



Para a análise quantitativa do resultado obtido foram utilizadas a métrica SSIM, definida na Eq. 2.18, e a Raiz do Erro Quadrático Médio (RMSE), definida na Eq. 2.19.

Como as métricas são sensíveis a translações, ainda que pequenas translações, a fim de utilizar uma medida que representa mais fielmente a visão humana, adotou-se o seguinte procedimento: para cada imagem de referência em alta resolução (que gerou as imagens em baixa resolução com as diferentes relações sinais ruído) foi recortado um retângulo de tal modo a manter apenas os pixels do algarismo (consequentemente, excluindo qualquer borda que contenha apenas pixels do fundo). Depois, calculou-se as métricas RMSE e SSIM movendo-se esse recorte em todas as posições possíveis dentro da imagem super-resolvida. O menor valor obtido de RMSE e o maior valor obtido de SSIM foram utilizados como valores finais das medidas. A Figura 4.4 ilustra esse processo.

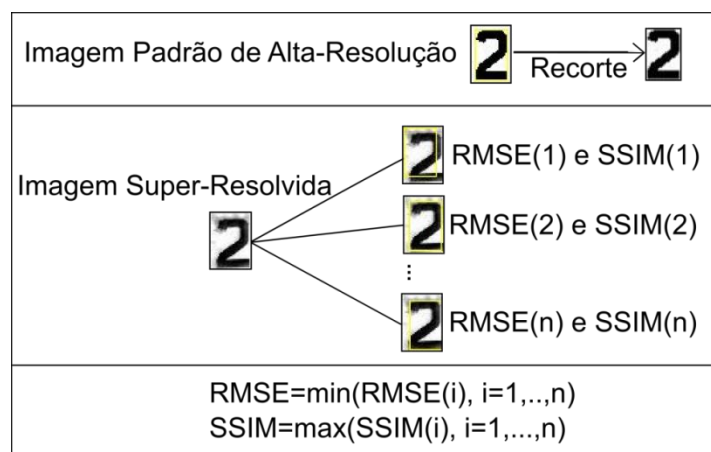


Figura 4.4: Processo de cálculo das métricas RMSE e SSIM.

Foram calculados o RMSE e SSIM para cada uma das imagens do conjunto de teste, de acordo com esse processo descrito. Os resultados obtidos são apresentados nas Figuras 4.5 a 4.13 e nas Tabelas 4.3 a 4.5.

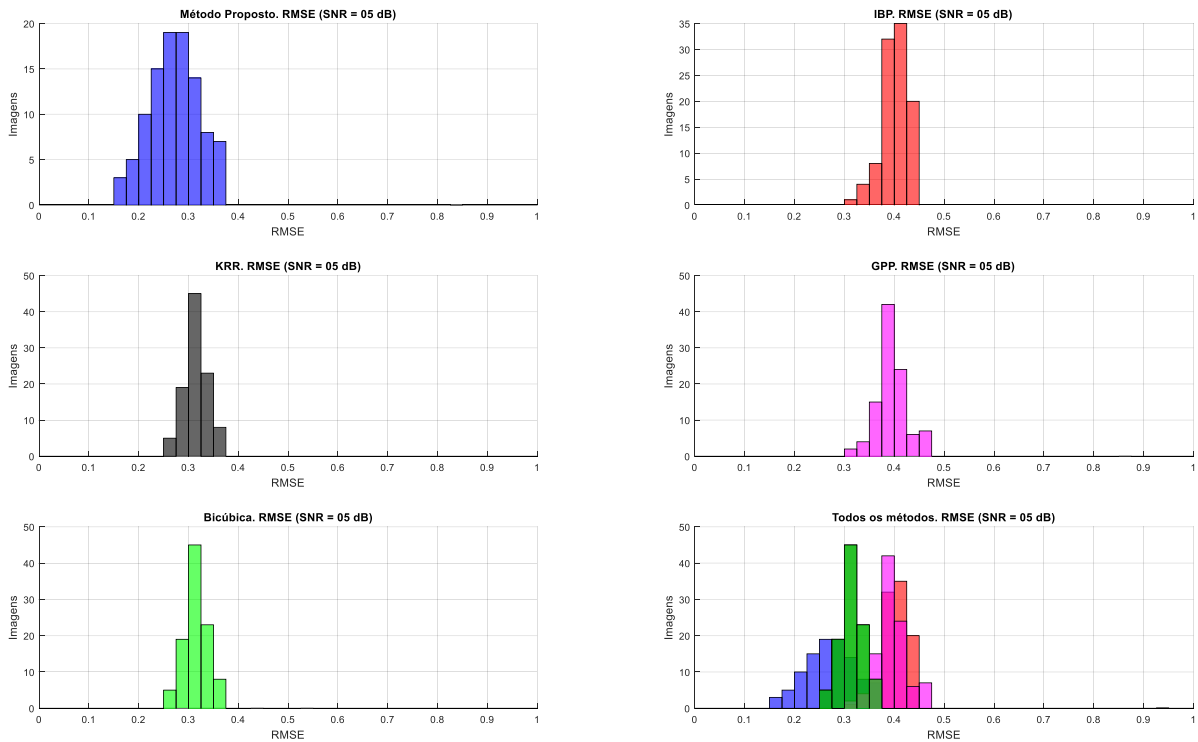


Figura 4.5: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 5 dB.

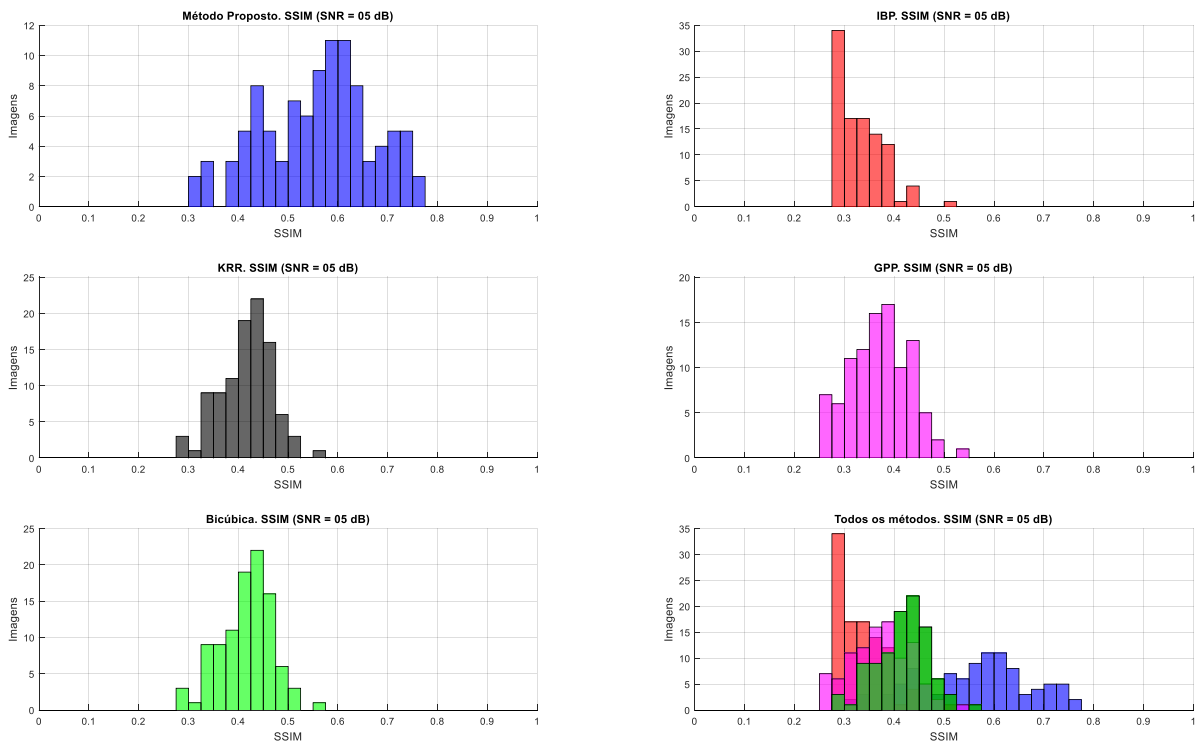


Figura 4.6: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 5 dB.

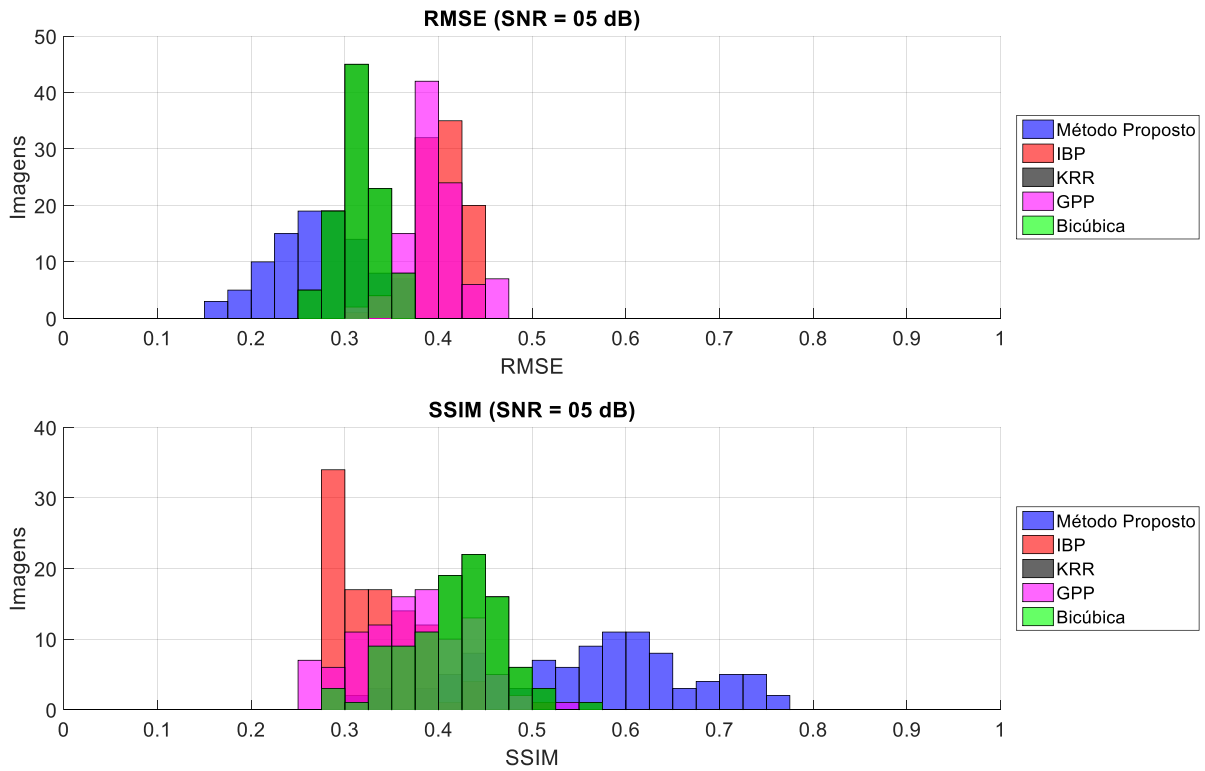


Figura 4.7: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 5 dB.

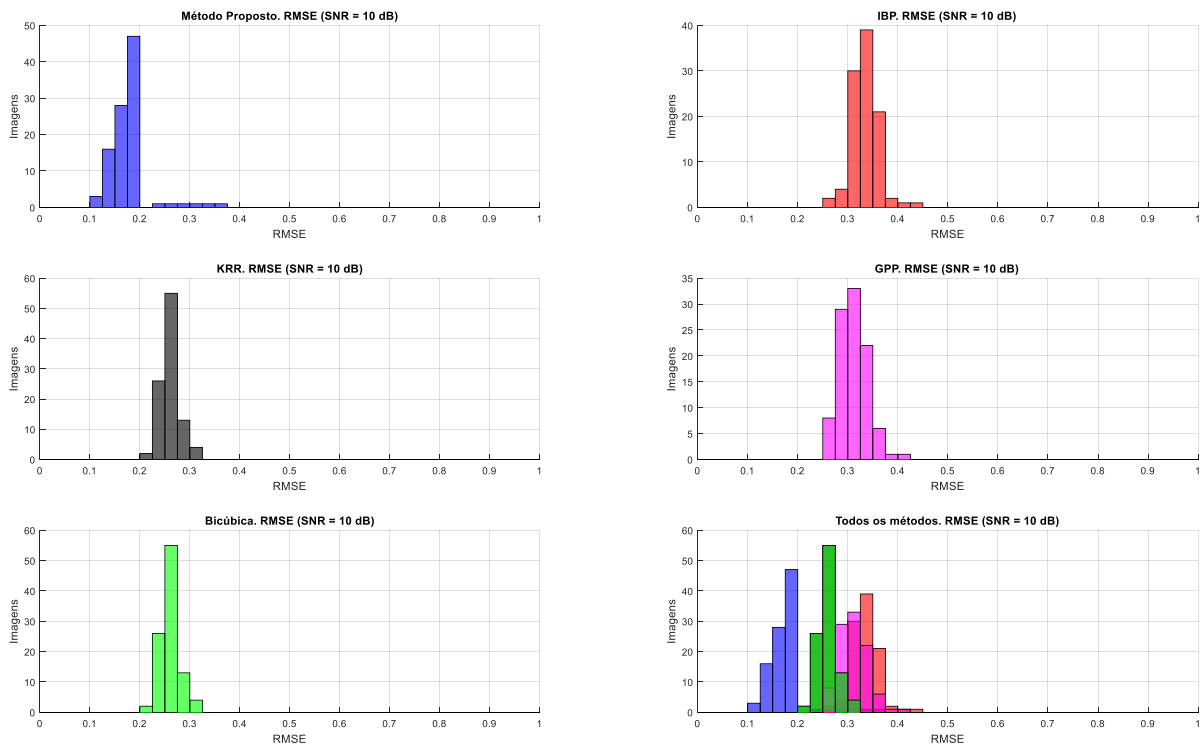


Figura 4.8: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 10 dB.

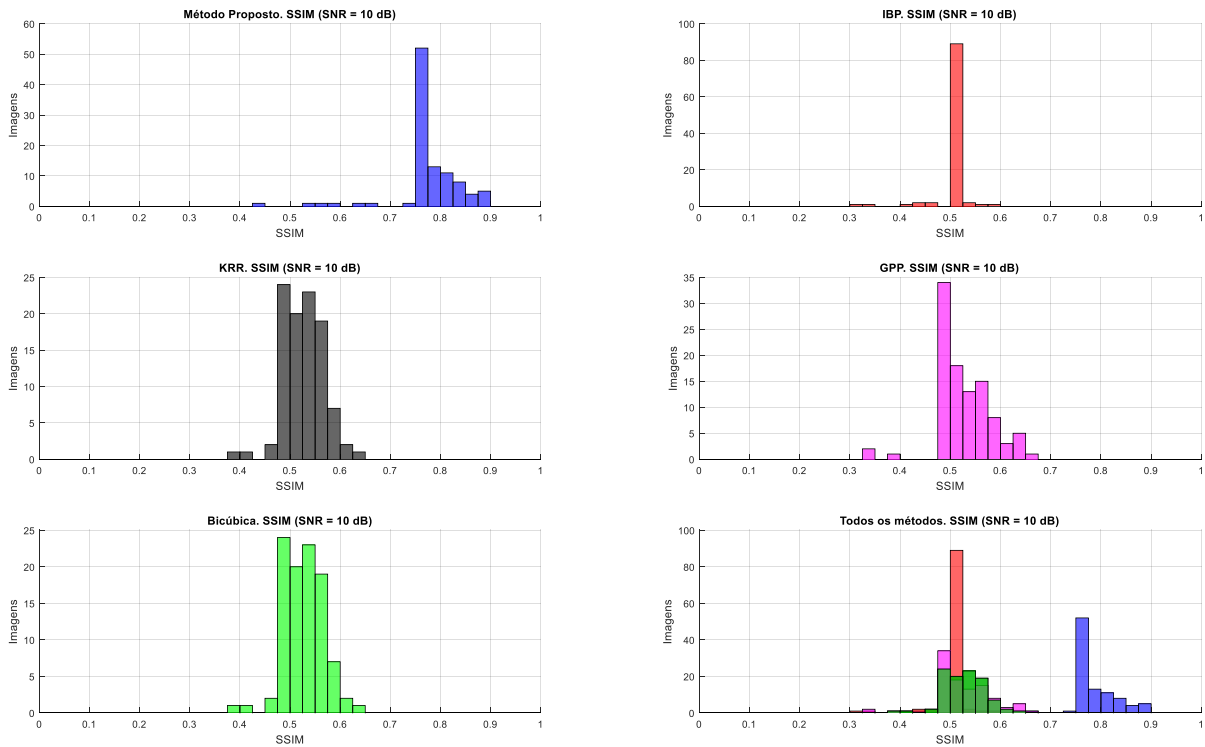


Figura 4.9: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 10 dB.

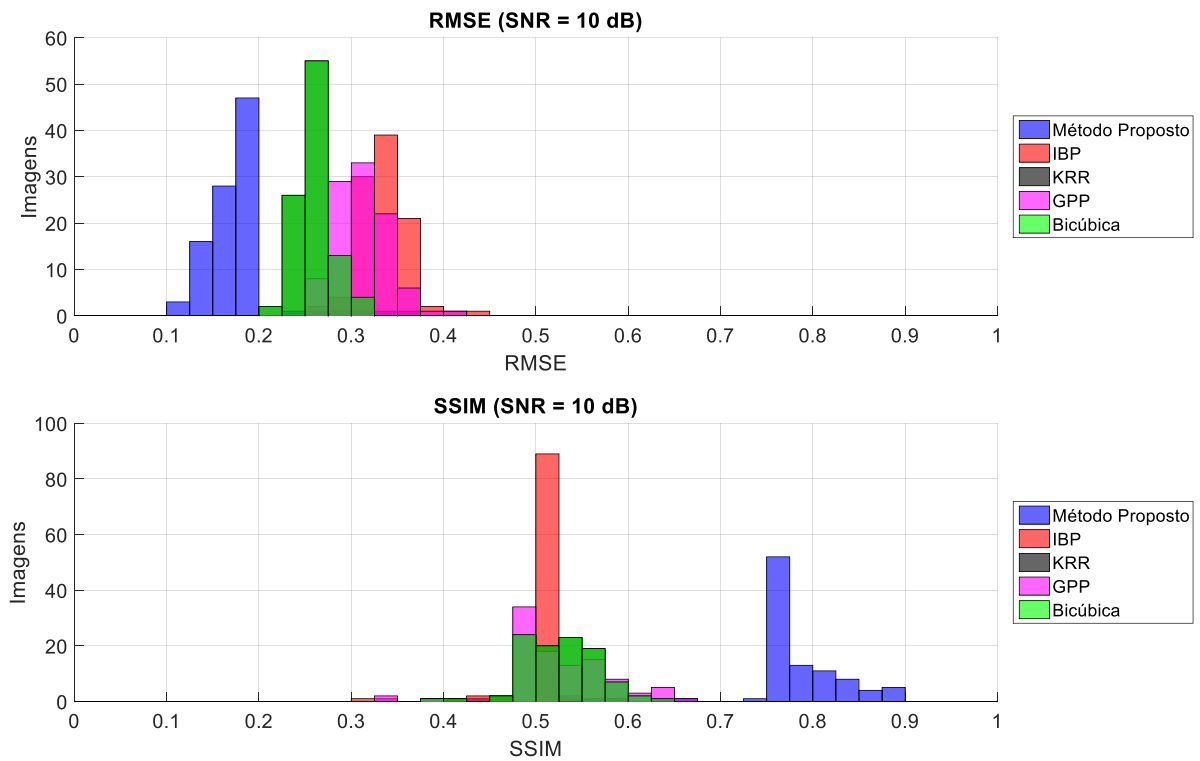


Figura 4.10: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 10 dB.

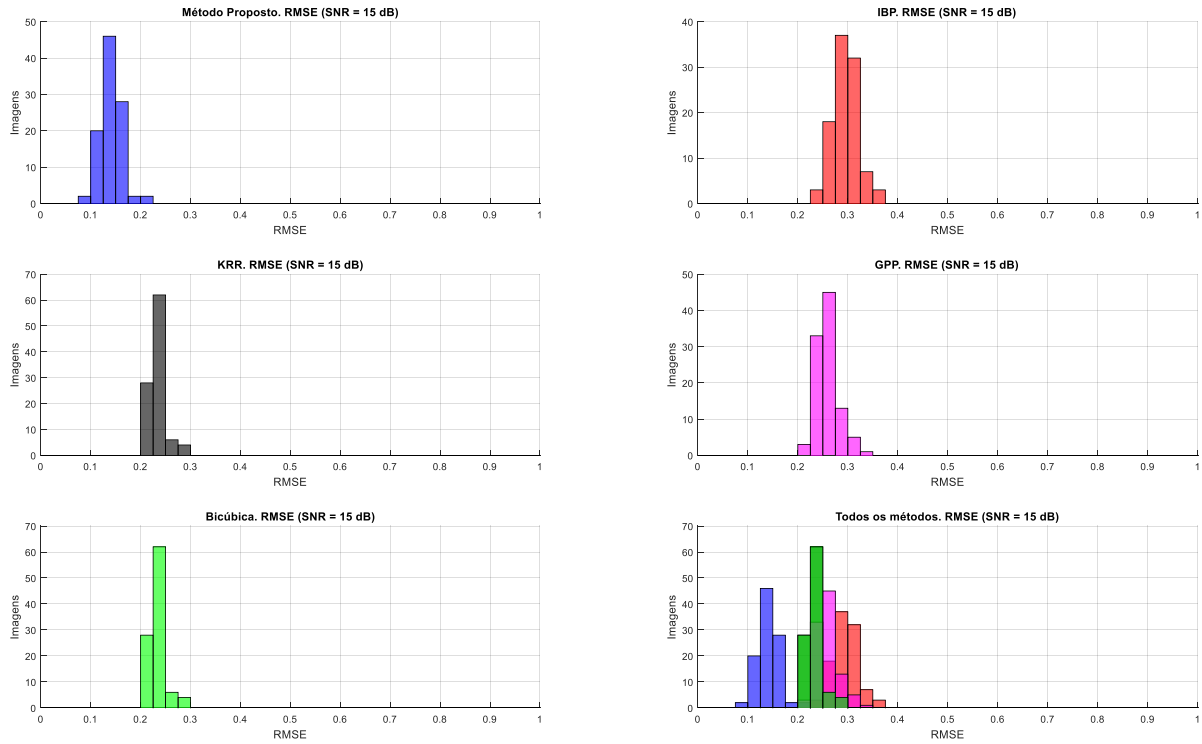


Figura 4.11: Histogramas dos resultados das medidas de RMSE para as 100 imagens com SNR igual a 15 dB.

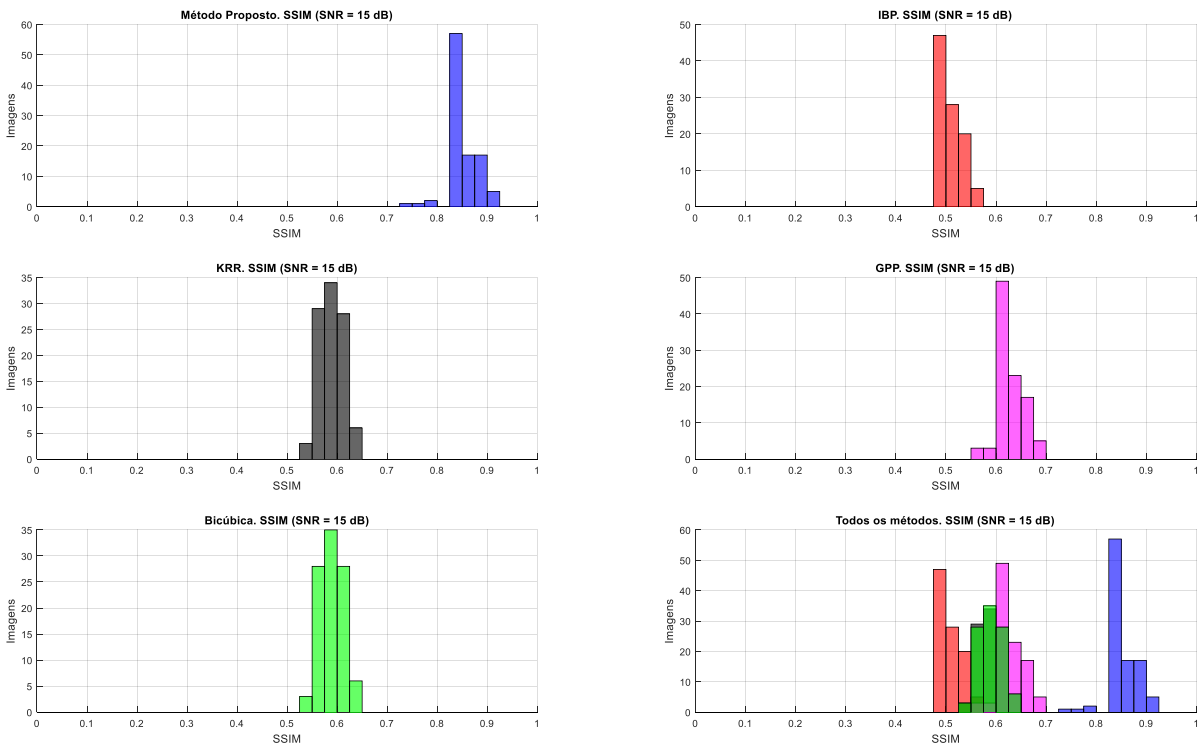


Figura 4.12: Histogramas dos resultados das medidas de SSIM para as 100 imagens com SNR igual a 15 dB.

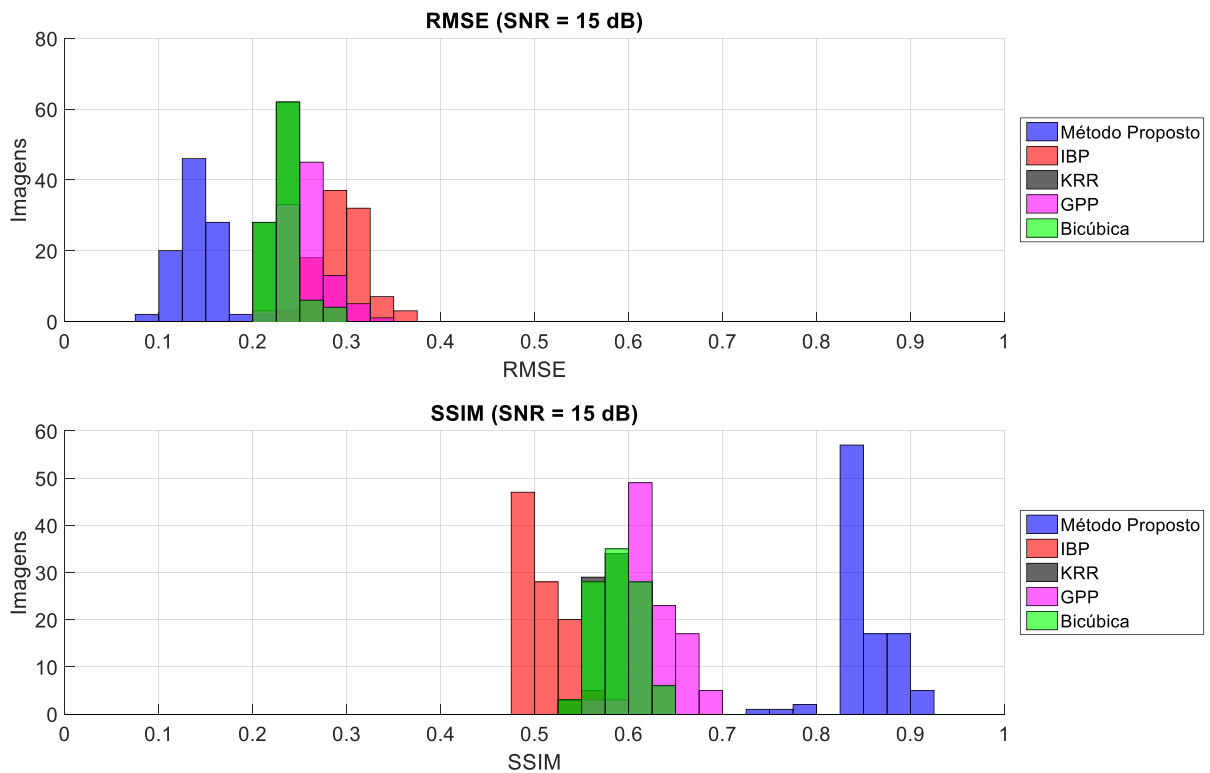


Figura 4.13: Histogramas dos resultados das medidas RMSE e SSIM para as 100 imagens com SNR igual a 15 dB.

Tabela 4.3 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 5 dB.

Método	SNR = 5 dB			
	RMSE		SSIM	
	Média	Desvio Padrão	Média	Desvio Padrão
Método Proposto	27,24 %	4,89 %	55,71 %	11,03%
IBP	40,18 %	2,69 %	33,39 %	4,16 %
KRR	31,46 %	2,36 %	41,60 %	5,26 %
GPP	39,53 %	3,04 %	37,08 %	5,94 %
Interpolação Bicúbica	31,46 %	2,35 %	41,61 %	5,26 %

Tabela 4.4 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 10 dB.

Método	SNR = 10 dB			
	RMSE		SSIM	
	Média	Desvio Padrão	Média	Desvio Padrão
Método Proposto	17,32 %	3,72 %	77,28 %	6,72 %
IBP	33,53 %	2,61 %	51,02 %	3,28 %
KRR	26,03 %	1,87 %	52,93 %	3,82 %
GPP	31,12 %	2,79 %	52,96 %	5,15 %
Interpolação Bicúbica	26,03 %	1,87 %	52,93 %	3,82 %

Tabela 4.5 – Valores médios e desvio padrão das medidas de RMSE e SSIM para todo o conjunto de teste com SNR de 15 dB.

Método	SNR = 15 dB			
	RMSE		SSIM	
	Média	Desvio Padrão	Média	Desvio Padrão
Método Proposto	14,15 %	2,22 %	85,70 %	2,57 %
IBP	29,53 %	2,39 %	51,17 %	2,06 %
KRR	23,42 %	1,67 %	58,77 %	2,48 %
GPP	25,85 %	2,17 %	63,07 %	2,64 %
Interpolação Bicúbica	23,42 %	1,67 %	58,77 %	2,48 %

Analisando-se os resultados quantitativos obtidos, conclui-se que o método proposto superou os resultados dos demais métodos, apresentando uma média de aumento de aproximadamente 15% no SSIM em relação à interpolação bicúbica para imagens com SNR de 5 e 10 dB e de aproximadamente 25% no SSIM em relação à interpolação bicúbica para imagens com SNR de 15 dB.

Também se verifica que, em relação à métrica SSIM, a melhora foi menor para o algoritmo 1 em comparação com os outros algoritmos em todos os três cenários de SNR. Isto provavelmente se deve ao fato do algoritmo 1 poder ser confundido com parte de outros algoritmos, como 0, 8, 6 ou 4, o que faz com que o algoritmo escolha blocos originados em outros algoritmos que não o 1.

Nota-se que o algoritmo não mostrou ganho quantitativo (RMSE ou SSIM) para imagens do algarismo 0 com SNR igual a 5 dB. Isto provavelmente se deu porque o algarismo 0 é muito semelhante a outros algarismos, como 8 e 6, e com uma relação sinal ruído tão baixa, o algoritmo não foi capaz de identificar corretamente o algarismo.

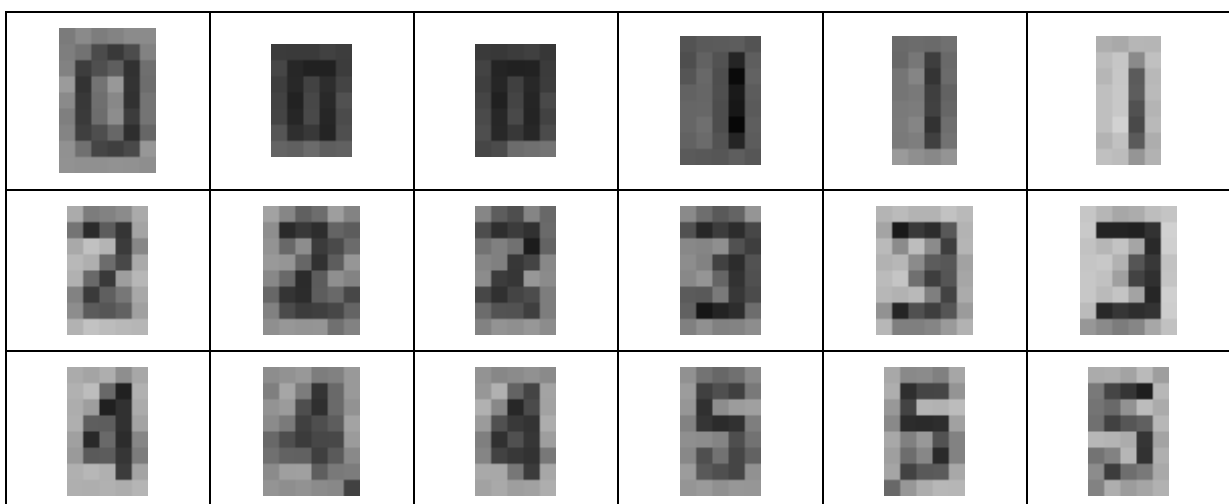
Numa análise qualitativa, o método proposto também se mostrou superior aos demais métodos confrontados. Qualitativamente nota-se que o método proporcionou uma melhor definição das bordas, um melhor contraste e melhor nitidez da imagem.

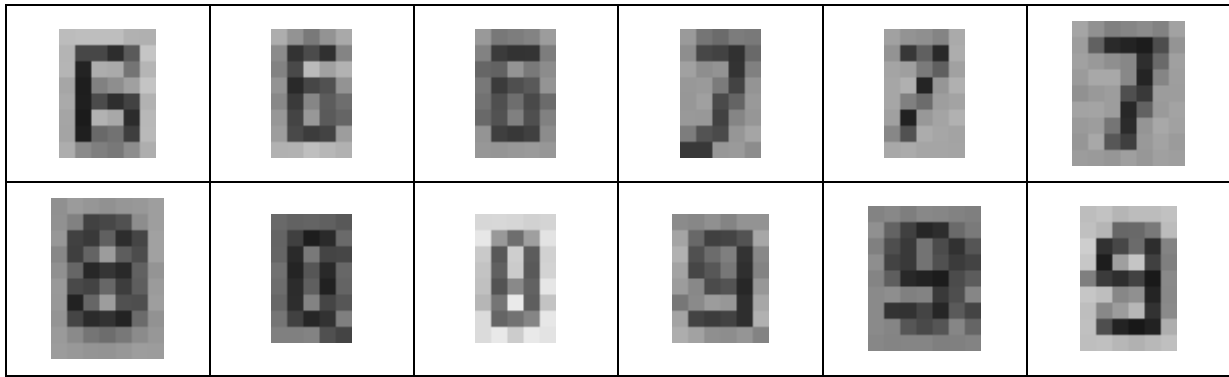
4.3. IMAGENS DE ALGARISMOS DE PLACAS REAIS DE BAIXA RESOLUÇÃO

O objetivo principal do método proposto é ser aplicado a imagens de placas de automóveis brasileiros reais de baixa resolução em ambiente não controlado (incidência e iluminação não controladas). Nesta seção, os resultados obtidos em ensaios exatamente neste cenário são mostrados e discutidos.

Para esse conjunto de imagens, criou-se um banco de imagens de algarismos de placas de automóveis reais de baixa resolução composto por 200 imagens, sendo 20 imagens de cada algarismo (0, 1, ..., 9). Esse banco de imagens está disponível em <https://github.com/alevicente/SR-LicensePlate/tree/master/Datasets/> sob o nome 4.B – Real Images. A Figura 4.14 ilustra algumas imagens desse banco de imagens criado.

Figura 4.14 – Amostra de imagens do banco de imagens de algarismos de placas de automóveis reais de baixa resolução (as imagens apresentadas foram ampliadas por um fator de 8 pelo método do vizinho mais próximo para facilitar a visualização).



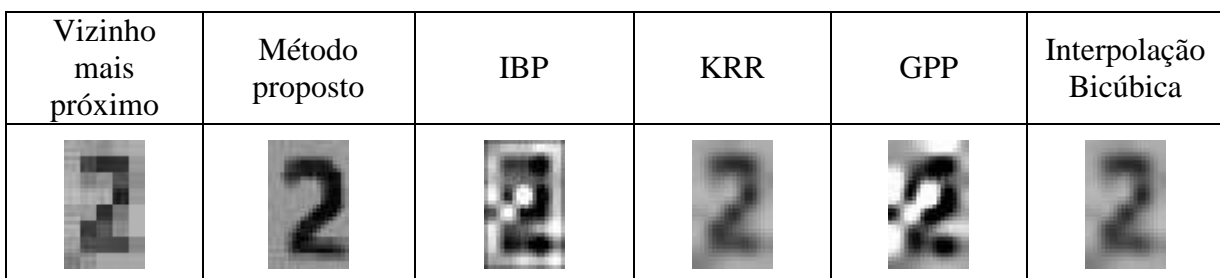




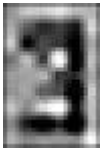





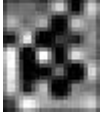

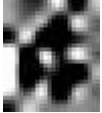







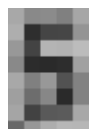







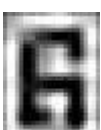


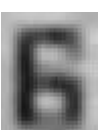


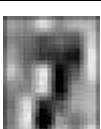





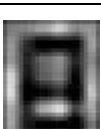


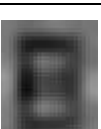

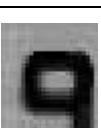




As 200 imagens do banco de imagens de algarismos de placas de automóveis reais de baixa resolução foram divididas em dois grupos, treinamento e teste. No conjunto de treinamento foram mantidas 180 imagens sendo 18 de cada algarismo. Já o conjunto de teste possui as 20 imagens restantes, sendo 2 de cada algarismo.

O classificador foi treinado no conjunto de treinamento, e depois o método proposto foi aplicado ao conjunto de teste. Para fins de comparação, também foram aplicados ao conjunto de teste os seguintes métodos: *Iterative Back Projection* (IBP) (Irani & Peleg, 1991), o *Kernel Ridge Regression* (KRR) (Kim & Kwon, 2008), *Gradient Profile Prior* (GPP) (Sun, Xu, & Shum, 2008) e a interpolação bicúbica.

O método proposto utilizou as seguintes configurações: imagens padrões de fundo cinza, fator de escala: 4, tamanho dos blocos em alta definição: 16 x 16, filtro de Blur: filtro média 3 x 3, $T = 0,8$, $n_{Posição} = 3$, número máximo de blocos $n = 1000$ para o conjunto Ω , $\lambda = 0,2$, número máximo de iterações $N = 50$. Uma amostra dos resultados obtidos é ilustrada na Figura 4.15.

Figura 4.15 – Amostra do resultado das imagens do conjunto teste super-resolvidas pelo método proposto, IBP, KRR, GPP e a interpolação bicúbica, todas por um fator de ampliação de 4. Para facilitar a visualização, os resultados obtidos pelos métodos citados foram ampliados por um fator de 2 utilizando o método do vizinho mais próximo e a imagem original em baixa resolução foi ampliada por um fator de 8 utilizando o método do vizinho mais próximo.



Vizinho mais próximo	Método proposto	IBP	KRR	GPP	Interpolação Bicúbica
					
					
					
					
					
					
					
					

Analisando-se os resultados obtidos de forma qualitativa, nota-se que o método proposto apresenta resultados visualmente superiores aos demais métodos confrontados. Percebe-se uma melhora na qualidade visual da imagem, uma maior facilidade de reconhecimento dos algarismos, uma melhor definição das bordas, uma melhor nitidez e melhor contraste.

5. CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho investigou o problema de super-resolução em imagens de placas de automóveis que seguem o padrão definido em normas brasileiras. Um método de imagem única baseado em aprendizado para imagens de placas de baixíssima qualidade foi proposto e analisado, o qual explora o conhecimento *a priori* de que as imagens contêm apenas um dígito de uma fonte de caracteres conhecida (“Mandatory”).

Esta dissertação resultou na produção e publicação do artigo científico intitulado “A Learning-Based Single-Image Super-Resolution Method for Very Low Quality License Plate Images”, apresentado na IEEE International Conference on Systems, Man, and Cybernetics (SMC), em Budapeste, Hungria, de 09 a 12 de outubro de 2016.

O algoritmo proposto demonstrou desempenho satisfatório mesmo em cenários muito degradados, com baixíssima resolução, baixa qualidade, presença de ruído e condições de iluminação não controladas, em que os demais algoritmos de super-resolução testados (IBP, KRR e GPP) falharam. Na seção 4.2, em que foi aplicado o algoritmo em um cenário controlado, cujas imagens de baixa resolução foram obtidas a partir de uma imagem em alta resolução conhecida e depois degradada com a adição de um ruído branco gaussiano, os resultados quantitativos se mostraram muito superiores aos obtidos pelos demais métodos, tendo apresentado uma média de aumento de aproximadamente 15% no SSIM em relação à interpolação bicúbica para imagens com SNR de 5 dB e 10 dB e de aproximadamente 25% no SSIM em relação à interpolação bicúbica para imagens com SNR de 15 dB. Já na seção 4.3, em que foi aplicado o algoritmo em imagens de algoritmos de placas de automóveis reais de baixa resolução, o resultado qualitativo se mostrou muito superior ao obtido pelos demais métodos, incluindo melhora significativa na definição das bordas, do contraste e da nitidez.

O método proposto emprega árvore de decisão como classificador, o qual permite uma fácil interpretação pelos seres humanos. Tal característica permite que o método possa ser explicado para pessoas leigas, como as pessoas destinatárias de Laudos Periciais Criminais (advogados, magistrados, promotores, júri, entre outros).

O trabalho evidenciou a importância e o impacto que a manipulação do histograma acarreta no desempenho de algoritmos de super-resolução que utilizam dicionário de blocos. Esta influência da manipulação de histograma é pouco explorada na literatura de super-

resolução e pode ser um caminho para melhorar a eficácia dos métodos baseados em dicionário de blocos.

Não há um banco de imagens de placas brasileiras disponíveis ao público para comparação de desempenho de algoritmos. Então, uma contribuição adicional deste trabalho foi a criação e disponibilização ao público da base de imagens utilizadas no trabalho, permitindo uma comparação de métodos neste domínio específico de imagens de algarismos de placas brasileiras.

O presente trabalho também contribuiu para a comunidade científica, mais particularmente para a comunidade científica forense, ao demonstrar a potencialidade de algoritmos de super-resolução para melhoria da qualidade visual de imagens de placas, mesmo quando a qualidade da imagem é muito baixa, e ao demonstrar que algoritmos de super-resolução podem ser desenvolvidos para domínios específicos ao se explorar as características conhecidas *a priori* desses domínios.

Por fim, o presente trabalho buscou demonstrar de maneira didática o funcionamento de um algoritmo de super-resolução de única imagem baseado em aprendizado e dicionário de blocos de uma maneira didática, facilitando a compreensão e o desenvolvimento de novas técnicas na área.

Como proposta para trabalhos futuros, a partir da metodologia aqui apresentada, modificações e refinamentos podem ser propostos como: alterações no classificador, nos descritores do classificador e no algoritmo de cálculo do mínimo da função custo, além da expansão do conjunto de imagens padrões para qualquer conjunto limitado de imagens – como incluir os caracteres de letras para se aplicar a todos os caracteres alfanuméricos de uma placa.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aldrich, J. (1995). Correlations Genuine and Spurious in Pearson and Yule. *Statistical Science*, pp. 364-376.
- Bevilacqua, M. (2014). Tese de Doutorado, Algorithms for super-resolution of images and videos based on learning methods. *Université de Rennes*.
- Cheeseman, P., Kanefsky, B., Kraft, R., & Stutz, J. (1994). Super-resolved surface reconstruction from multiple images. *Technical Report FIA9412, NASA*.
- Demirel, H., & Anbarjafari, G. (2011). Image resolution enhancement by using discrete and stationary wavelet decomposition. *IEEE Trans. Image Process.* 20, pp. 1458-1460.
- Gerchberg, R. (1974). Super-resolution through error energy rection. *J. Mod. Opt.* 21, (pp. 709-720).
- Gonzalez, R. C., & Woods, R. (2007). *Digital Image Processing* (3^a ed.). Pearson Prentice Hall.
- Guo, K., Yang, X., Zhang, R., & Yu, S. (2009). Learning super resolution with global and local constraints. *Proceedings of IEEE International Conference on Multimedia and Expo*, (pp. 590-593).
- Irani, M., & Peleg, S. (1990). Super-resolution from image sequences. *Proceedings of IEEE International Conference on Pattern Recognition*, (pp. 115-120).
- Irani, M., & Peleg, S. (1991, Maio). Improving Resolution by Image Registration. *GRAPHICAL MODELS AND IMAGE PROCESSING (CVGIP)*, pp. 231-239.
- Jung, C., Jiao, L., Liu, B., & Gong, M. (2011). Position-patch based face hallucination using convex optimization. *IEEE Signal Process Lett.* 18(6), pp. 367-369.
- Júnior, E. (2015). Retrieved 05 23, 2016, from Empresa Brasil de Comunicação: <http://www.ebc.com.br/tecnologia/2015/05/uit-diz-que-numero-de-celulares-no-mundo-passou-dos-7-bilhoes-em-2015>
- Kim, K., & Kwon, Y. (2008). Example-based Learning for Single-Image Super-resolution. *Proceedings of the 30th DAGM symposium on Pattern Recognition*, (pp. 456-465).
- Malczewski, K., & Stasinski, R. (2008). Toeplitz-based iterative image fusion scheme for MRI. *Proceedings of IEEE International Conference on Image Processing*, (pp. 341-344).
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- Nasrollahi, K., & Moeslund, T. (2014). Super-resolution: a comprehensive survey. *Machine Vision and Applications*, 1423-1468.

- Pedrini, H., & Schwartz, W. (2007). *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. São Paulo: Thomson Learning.
- Samuel, A. (1959, Julho). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, pp. 210-229.
- Santis, P. D., & Gori, F. (1975). On an iterative method for super-resolution. *J. Mod. Opt.* 22, (pp. 691-695).
- Stark, H., & Oskoui, P. (1989). High-resolution image recovery from imageplane arrays, using convex projections. *J. Opt. Soc. Am.*, pp. 1715-1726.
- Stuart, R., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach (2^a Ed)*. Prentice Hall.
- Sun, J., Xu, Z., & Shum, H.-Y. (2008). Image Super-Resolution using Gradient Profile Prior. *Computer Vision and Pattern Recognition*, (pp. 1-8).
- Tian, Y., Yap, K.-H., & He, Y. (2011). High Resolution Vehicle License Plate Reconstruction Using Soft Recognition Learning. *International Conference on Information, Communications, and Signal Processing (ICICS)*, (pp. 1-5).
- Tsai, R., & Huang, T. (1984). Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing*, pp. 317-339.
- Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E. (2004, Abril). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612.
- Yang, X., Su, G., Chen, J., & Moon, Y. (2010). Restoration of lowresolution car plate images using PCA based image super-resolution. *Proceedings of IEEE International Conference on Image Processing*, (pp. 2789–2792). Hong Kong.
- Yoshida, T., Takahashi, T., Deguchi, D., Ide, I., & Murase, H. (2012). Robust face super-resolution using free-form deformations for low-quality surveillance video. *Proceedings of IEEE International Conference on Multimedia and Expo*.
- Zeng, X., & Huang, H. (2012). Super-resolution method for multiview face recognition from a single image per person using nonlinear mappings on coherent features. *IEEE Signal Process. Lett.* 19(4), 195-198.
- Zhang, H., Zhang, L., & Shen, H. (2012). A super-resolution reconstruction algorithm for hyper spectral images. *Signal Process.* 92(9), pp. 2082-2096.
- Zhang, L., Zhang, H., Shen, H., & Li, P. (2010). A super-resolution reconstruction algorithm for surveillance images. *Signal Process.* 90(3), pp. 848-859.
- Zhou, F. (2012). A coarse-to-fine subpixel registration method to recover local perspective deformation in the application of image superresolution. *IEEE Trans. Image Process.* 21(1), pp. 53-66.

Zou, W., & Yuen, P. (2012). Very low resolution face recognition problem. *IEEE Trans. Image Process.* 21(1), 327-340.

ANEXOS

A – CÁLCULOS PARA A MINIMIZAÇÃO DA FUNÇÃO CUSTO

Neste anexo, o cálculo da derivada da função custo utilizada no método proposto é detalhado. A Eq. 3.1 (copiada na Eq. A.1) define a função custo a ser minimizada.

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \|\mathbf{LR} - (\mathbf{Blur} * \mathbf{HR}) \downarrow\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} \mathbf{w}(i,j,k) \|\mathbf{R}(i,j)\mathbf{HR} - \mathbf{x}(i,j,k)\|_2^2 \right\} \quad \text{Eq. A.1}$$

Para simplificar a notação, serão adotadas as seguintes notações:

$$\mathbf{M} \cdot \mathbf{HR} = (\mathbf{Blur} * \mathbf{HR}) \downarrow$$

$$w = \mathbf{w}(i,j,k)$$

$$\mathbf{R} = \mathbf{R}(i,j)$$

$$\mathbf{x} = \mathbf{x}(i,j,k)$$

Utilizando-se essas simplificações, a equação pode ser reescrita como:

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \|\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \|\mathbf{R} \cdot \mathbf{HR} - \mathbf{x}\|_2^2 \right\} \quad \text{Eq. A.2}$$

Todos os operadores e variáveis desta equação serão considerados matrizes com as seguintes dimensões: $\mathbf{LR}_{nLR \times 1}$; $\mathbf{M}_{nLR \times nHR}$; $\mathbf{HR}_{nHR \times 1}$, $\lambda_{1 \times 1}$; $W_{1 \times 1}$; $\mathbf{R}_{nBloco \times nHR}$; $\mathbf{x}_{nBloco \times 1}$, em que nLR e nHR são, respectivamente, o número de pixels das imagens em baixa resolução e em alta resolução e $nBloco$ é o número de pixels do bloco em alta resolução.

Desenvolvendo o componente $\|\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}\|_2^2$ da equação, tem-se:

$$\begin{aligned} \|\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}\|_2^2 &= [\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}]^T [\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}] = [\mathbf{LR}^T - \mathbf{HR}^T \cdot \mathbf{M}^T] [\mathbf{LR} - \mathbf{M} \cdot \mathbf{HR}] \\ &= \mathbf{LR}^T \cdot \mathbf{LR} - \mathbf{LR}^T \cdot \mathbf{M} \cdot \mathbf{HR} - \mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{LR} + \mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{M} \cdot \mathbf{HR} \end{aligned}$$

Como $\mathbf{LR}_{1 \times nLR}^T \cdot \mathbf{M}_{nLR \times nHR} \cdot \mathbf{HR}_{nHR \times 1}$ tem dimensão 1×1 , então:

$$\mathbf{LR}_{1 \times nLR}^T \cdot \mathbf{M}_{nLR \times nHR} \cdot \mathbf{HR}_{nHR \times 1} = (\mathbf{LR}_{1 \times nLR}^T \cdot \mathbf{M}_{nLR \times nHR} \cdot \mathbf{HR}_{nHR \times 1})^T = \mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{LR}$$

Portanto,

$$\|\mathbf{LR} - \mathbf{M.HR}\|_2^2 = \mathbf{LR}^T \cdot \mathbf{LR} - 2\mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{LR} + \mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{M} \cdot \mathbf{HR} \quad \text{Eq. A.3}$$

Analogamente, o componente $\|\mathbf{R.HR} - \mathbf{x}\|_2^2$ da equação pode ser escrito como:

$$\|\mathbf{R.HR} - \mathbf{x}\|_2^2 = \mathbf{HR}^T \cdot \mathbf{R}^T \cdot \mathbf{R} \cdot \mathbf{HR} - 2\mathbf{HR}^T \cdot \mathbf{R}^T \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{x} \quad \text{Eq. A.4}$$

Substituindo as Eq. A.3 e Eq. A.4 em Eq. A.2, tem-se:

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \mathbf{LR}^T \cdot \mathbf{LR} - 2\mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{LR} + \mathbf{HR}^T \cdot \mathbf{M}^T \cdot \mathbf{M} \cdot \mathbf{HR} \right. \\ \left. + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w(\mathbf{HR}^T \cdot \mathbf{R}^T \cdot \mathbf{R} \cdot \mathbf{HR} - 2\mathbf{HR}^T \cdot \mathbf{R}^T \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{x}) \right\} \quad \text{Eq. A.5}$$

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \mathbf{HR}^T \left(\mathbf{M}^T \cdot \mathbf{M} + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \mathbf{R}^T \cdot \mathbf{R} \right) \mathbf{HR} \right. \\ \left. - 2 \cdot \mathbf{HR}^T \left(\mathbf{M}^T \cdot \mathbf{LR} + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \cdot \mathbf{R}^T \cdot \mathbf{x} \right) + \mathbf{LR}^T \cdot \mathbf{LR} \right. \\ \left. + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \cdot \mathbf{x}^T \cdot \mathbf{x} \right\} \quad \text{Eq. A.6}$$

Para simplificar a notação, define-se:

$$\mathbf{A}_{n_{HR} \times n_{HR}} := \mathbf{M}^T \cdot \mathbf{M} + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \mathbf{R}^T \cdot \mathbf{R} \quad \text{Eq. A.7}$$

$$\mathbf{b}_{n_{HR} \times 1} := \mathbf{M}^T \cdot \mathbf{LR} + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \cdot \mathbf{R}^T \cdot \mathbf{x} \quad \text{Eq. A.8}$$

Utilizando-se essas novas definições, a Eq. A.6 pode ser reescrita da seguinte forma:

$$\widehat{\mathbf{HR}} = \arg \min_{\mathbf{HR}} \left\{ \mathbf{HR}^T \cdot \mathbf{A} \cdot \mathbf{HR} - \mathbf{HR}^T \cdot 2 \cdot \mathbf{b} + \mathbf{LR}^T \cdot \mathbf{LR} \right. \\ \left. + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \cdot \mathbf{x}^T \cdot \mathbf{x} \right\} \quad \text{Eq. A.9}$$

Para o método de minimização utilizado, Gradiente Descendente, é necessário se calcular a derivada da função custo a ser minimizada com relação à \mathbf{HR} . Definindo $\phi(\mathbf{HR})$ como a função custo a ser minimizada, tem-se:

$$\phi(\mathbf{HR}) := \mathbf{HR}^T \cdot \mathbf{A} \cdot \mathbf{HR} - \mathbf{HR}^T \cdot 2 \cdot \mathbf{b} + \mathbf{LR}^T \cdot \mathbf{LR} + \lambda \sum_{i,j} \sum_{k \in \Omega(i,j)} w \cdot \mathbf{x}^T \cdot \mathbf{x} \quad \text{Eq. A.10}$$

Utilizando os teoremas de derivadas de matriz definidos na seção 2.5, tem-se:

$$\frac{d\phi(\mathbf{HR})}{d\mathbf{HR}} = (\mathbf{A} + \mathbf{A}^T) \cdot \mathbf{HR} - 2 \cdot \mathbf{b} \quad \text{Eq. A.11}$$

Como a matriz \mathbf{A} é simétrica (fácil de se observar pela própria definição de \mathbf{A}), então:

$$\frac{d\phi(\mathbf{HR})}{d\mathbf{HR}} = 2(\mathbf{A} \cdot \mathbf{HR} - \mathbf{b}) \quad \text{Eq. A.12}$$

Portanto, a Eq. A.12 define o gradiente (derivada) que deve ser utilizado no método Gradiente Descendente na minimização da função custo.