

MARCELO OLIVEIRA CAMPONEZ

**RECONSTRUÇÃO DE IMAGENS POR SUPERRESOLUÇÃO
UTILIZANDO INFERÊNCIA BAYESIANA APROXIMADA**

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

Orientador: Prof. Dr. Mário Sarcinelli Filho.

Orientador: Prof. Dr. Evandro Ottoni Teatini Salles.

VITÓRIA
2012

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

C198r Camponez, Marcelo Oliveira, 1973-
Reconstrução de imagens por superresolução utilizando
inferência Bayesiana aproximada / Marcelo Oliveira Camponez. –
2012.
92 f. : il.

Orientador: Mário Sarcinelli Filho.

Coorientador: Evandro Ottoni Teatini Salles.

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

1. Processamento de imagens 2. Teoria bayesiana de
decisão estatística. I. Sarcinelli Filho, Mário. II. Salles, Evandro
Ottoni Teatini. III. Universidade Federal do Espírito Santo. Centro
Tecnológico. IV. Título.

CDU: 621.3

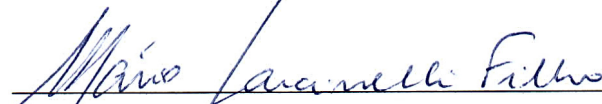
MARCELO OLIVEIRA CAMPONEZ

**RECONSTRUÇÃO DE IMAGENS POR SUPERRESOLUÇÃO
UTILIZANDO INFERÊNCIA BAYESIANA APROXIMADA**

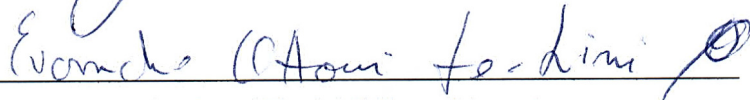
Tese submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Doutor em Engenharia Elétrica.

Aprovada em 26 de outubro de 2012.

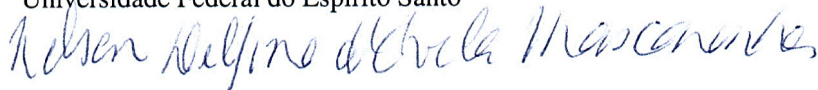
COMISSÃO EXAMINADORA



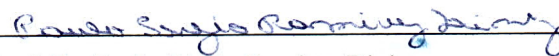
Prof. Dr. Mário Sarcinelli Filho - Orientador
Universidade Federal do Espírito Santo



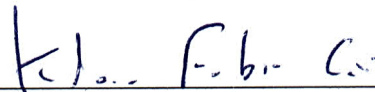
Prof. Dr. Evandro Ottoni Teatini Salles - Orientador
Universidade Federal do Espírito Santo



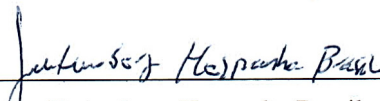
Prof. Dr. Nelson Delfino D'Ávila Mascarenhas
Universidade Federal de São Carlos



Prof. Dr. Paulo Sérgio Ramirez Diniz
Universidade Federal do Rio de Janeiro



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



Prof. Dr. Gutemberg Hespanha Brasil
Universidade Federal do Espírito Santo

Dedico esse trabalho aos meus filhos Matheus e Luísa e minha esposa Leticia.

Agradecimentos

Agradeço a Deus em primeiro lugar, pela oportunidade de concluir mais esse sonho.

Agradeço aos meus orientadores Mário Sarcinelli Filho e Evandro Ottoni Teatini Salles pelas saudáveis e, muitas vezes divertidas, discussões, pelas dicas e valiosas sugestões durante o desenvolvimento desse trabalho.

Não poderia deixar de agradecer também, à Universidade Vila Velha pelo suporte financeiro, pela flexibilidade e o apoio para participação em congressos internacionais.

Agradeço aos colegas de trabalho do Laboratório de Computação e Sistemas Neurais - CISNE do PPGEE da Universidade Federal do Espírito Santo, pelas várias discussões interessantes e a oportunidade de aprendermos e pesquisarmos juntos.

Agradeço ao Dr. Rue Harvard do Departamento de Ciências Matemáticas da *Norwegian University of Science and Technology*, por todo suporte e disponibilidade para tirar dúvidas sobre o método *Integrated Nested Laplace Approximation*, de sua autoria.

Sumário

1	Introdução ao Problema de Superresolução	16
1.1	Algoritmos de Superresolução	19
1.2	Objetivos desta Tese	22
1.3	Metodologia Utilizada	23
1.4	Organização da Tese	23
2	Solução em Forma Fechada para Superresolução	25
2.1	Notação	25
2.2	Modelo de Aquisição de Imagens de Baixa Resolução	26
2.2.1	Modelo de Aquisição <i>Warp-Blur</i>	27
2.2.2	Sistema de Aquisição	30
2.2.3	Modelagem de Movimento	31
2.3	Modelagem Hierárquica Bayesiana	32
2.3.1	Função de Probabilidade Condicional de Verossimilhança	33
2.3.2	Função de Probabilidade Condicional a Priori da Imagem HR	34
2.3.3	Função de Probabilidade Marginal do Hiperparâmetro	35
2.3.4	Modelo Hierárquico Bayesiano	35
2.4	Derivando uma Forma Fechada para o algoritmo de Superresolução	35
2.5	Implementação do Algoritmo de Superresolução	37
2.6	Experimentos	41

3	Superresolução Usando Integrated Nested Laplace Approximation	46
3.1	Integrated Nested Laplace Approximation	46
3.2	Aplicando o INLA ao Problema da Superresolução	49
3.2.1	Uma Equação em Forma Fechada para $\bar{\mathbf{X}}$	49
3.2.2	Selecionando um Conjunto de Valores para λ	50
3.2.3	Derivando uma Forma Fechada para $\tilde{p}_G(\mathbf{X} \lambda, \mathbf{Y})$	51
3.2.4	Uma Forma Fechada para $\tilde{p}(\lambda \mathbf{Y})$	51
3.2.5	Derivando uma Forma Fechada para $\tilde{p}(\mathbf{X} \mathbf{Y})$	54
3.2.6	Gerando a Imagem de Alta Resolução através do Valor Esperado	55
3.2.7	Gerando a Imagem de Alta Resolução através de MAP	55
3.3	Experimentos	56
3.3.1	Imagens Simuladas	56
3.3.2	Imagens Capturadas	61
3.4	Conclusões do Capítulo	63
4	Superresolução Rápida Baseada em DFT-2D	64
4.1	Alterações no Modelo de Aquisição de Imagens Warp-Blur	65
4.2	Superresolução Baseada em DFT-2D	66
4.2.1	Matrizes Toeplitz Circulante	66
4.2.2	Matrizes Toeplitz Circulante em Bloco	67
4.2.3	Aproximação BTC para a Matriz \mathbf{A}	68
4.2.4	Algoritmo DFT Closed Form SR	70
4.2.5	Algoritmo DFT INLA SR	71
4.3	Experimentos	71
4.3.1	Imagem Simulada	72

4.3.2	Imagem Real	73
4.3.3	Usando Mais Imagens LR	75
4.4	Conclusões do Capítulo	77
5	Conclusões e Proposta de Continuação	80
5.1	Contribuições	80
5.1.1	Closed Form SR	81
5.1.2	INLA SR e INLA MAP SR	82
5.1.3	DFT Closed Form DR e DFT INLA SR	83
5.2	Propostas Para Trabalhos Futuros	84
5.3	Conclusão	85
A	Publicações Feitas até o Momento	92

Lista de Tabelas

2.1	Avaliação das imagens HR reconstruídas considerando PSNR (dB)	43
3.1	Avaliação das imagens HR reconstruídas considerando PSNR (dB)	61
4.1	Avaliação das imagens de alta resolução reconstruídas considerando-se PSNR (dB) e o tempo (s) de convergência do algoritmo	76
4.2	Avaliação de imagens de alta resolução reconstruídas considerando-se o tempo (s) de convergência do algoritmo	77
4.3	Avaliação de imagens de alta resolução reconstruídas considerando-se PSNR (dB), o tempo (s) de convergência do algoritmo e o número de imagens de baixa resolução de entrada.	78

Lista de Figuras

1.1	(a) Aquisição de imagens por múltiplas câmeras, (b) aquisição por meio de uma filmadora.	17
1.2	Exemplo ilustrativo de Superresolução (a) imagem de alta resolução de 4 pixels, (b)-(e) imagens de baixa resolução, com movimento subpixel, cada uma com um único pixel.	18
1.3	Exemplo ilustrativo de Superresolução (a)-(d) imagens de baixa resolução, com movimento subpixel, (e) imagem de alta resolução com grid irregular, (f) imagem de alta resolução com grid regular.	19
2.1	Nomenclatura usada em Superresolução	26
2.2	Modelo de Aquisição de Imagem	27
2.3	Modelo Warp-Blur para aquisição de Imagem	28
2.4	Efeito de Degradação na Aquisição de Imagens	29
2.5	Modelo de Aquisição	30
2.6	Exemplo de Transformação de Movimento	32
2.7	O conjunto de imagens HR reconstruídas usando a imagem de teste Barco. (a) A imagem original (<i>ground truth</i>), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) Interpolação, (d) em (Vandewalle et al., 2006) (e) em (Pham et al., 2006) (f) Closed Form SR proposto aqui.	44
2.8	O conjunto de imagens HR reconstruídas usando a imagem de teste Texto. (a) A imagem original (<i>ground truth</i>), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) Interpolação, (d) em (Vandewalle et al., 2006) (e) em (Pham et al., 2006) (f) Closed Form SR proposto aqui.	45

3.1	O gráfico da PDF $p(\mathbf{X} \mathbf{Y})$ (equação (3.52)) versus λ , para o primeiro experimento com quatro quadros LR da imagem Barco	58
3.2	O conjunto de imagens HR reconstruídas usando a imagem de teste Barco. (a) A imagem original (<i>ground truth</i>), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) em (Pham et al., 2006), (d) em (Farsiu et al., 2004b), (e) INLA MAP proposto aqui (f) INLA SR proposto aqui.	59
3.3	O conjunto de imagens HR reconstruídas usando a imagem de teste Texto. (a) A imagem original (<i>ground truth</i>), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) em (Pham et al., 2006), (d) em (Farsiu et al., 2004b), (e) INLA MAP proposto aqui (f) INLA SR proposto aqui.	60
3.4	O conjunto de imagens HR reconstruídas a partir de uma sequência de quatro imagens reais LR. (a) Uma das quatro imagens LR. Imagens HR reconstruídas usando a abordagem: (b) em (Farsiu et al., 2004b), (c) INLA SR proposto aqui (d) INLA MAP proposto aqui.	62
4.1	O conjunto de imagens reconstruídas usando-se a imagem teste Barco. (a) A imagem original (<i>ground truth</i>), (b) uma das imagens LR simuladas. Reconstrução de imagem HR usando-se o método: (c) em (Vandewalle et al., 2006), (d) em (Pham et al., 2006), (e) in (Farsiu et al., 2004b), (f) Closed Form SR (Camponez et al., 2011b), (g) INLA SR (Camponez et al., 2012c), (h) DFT Closed Form (Camponez et al., 2012b; Camponez et al., 2012a) e (i) DFT INLA SR (Camponez et al., 2012b; Camponez et al., 2012a).	74
4.2	O conjunto de imagens reconstruídas usando-se uma sequência de 4 imagens reais. (a) Uma das quatro imagens LR. Imagens HR reconstruídas usando-se o método: (b) em (Farsiu et al., 2004b), (c) Closed Form SR (Camponez et al., 2011b), (d) INLA SR (Camponez et al., 2011a; Camponez et al., 2012c), (e) DFT Closed Form SR (Camponez et al., 2012b; Camponez et al., 2012a) e (f) DFT INLA SR (Camponez et al., 2012b; Camponez et al., 2012a).	76
4.3	O conjunto de imagens reconstruídas usando-se a imagem de teste Barco. (a) A imagem original (<i>ground truth</i>), (b) uma das quatro imagens LR. Imagens HR reconstruídas usando-se o método: (c) Closed Form SR (13 entradas LR), (d) DFT Closed SR (13 entradas LR), (e) Closed Form SR (4 entradas LR) e (f) DFT Closed Form SR (4 entradas LR).	78

Glossário

BTC - Matriz Toeplitz Circulante em Blocos.

CCD - Charge Coupled Device.

DFT - Discrete Fourier Transform.

GCV - Generalized Cross-Validation.

GMRF - Gaussian Markov Random Field.

HR - High Resolution (Alta Resolução).

HR - Low Resolution (Baixa Resolução).

INLA - Integrated Nested Laplace Approximation.

MAP - Máximo a Posteriori.

MCMC - Métodos de Monte Carlo via Cadeias de Markov.

PDF - Função de Densidade de Probabilidade.

PGM - Modelo Gráfico Probabilístico.

POCS - Projection Onto Convex Sets.

PSF - Point Spread Function.

PSNR - Relação Sinal-Ruído de Pico.

SR - Superresolução.

Toeplitz - Matriz de diagonais constantes.

Toeplitz Circulante - Matriz Toeplitz onde cada linha é um rotação a direita da linha superior.

Resumo

Reconstrução de imagens por Superresolução (SR) *multiframe* são técnicas capazes de aumentar a resolução de uma imagem, sem alterar a resolução da câmera que está sendo usada para captá-las, através do uso de algoritmos de processamento de imagens. Um conjunto de imagens de baixa resolução observável é primeiramente registrado em relação a uma das imagens e, posteriormente, fundido em uma imagem de alta resolução, que posteriormente é filtrada para melhorar sua nitidez. Nesse contexto, essa Tese de Doutorado propõe cinco algoritmos de Superresolução, baseados em modelagem estatística Bayesiana.

No primeiro algoritmo desenvolvido, Closed Form SR, uma equação em forma fechada é derivada para o cálculo da imagem de alta resolução. No entanto, o algoritmo depende de um parâmetro λ , que corresponde ao parâmetro de regularização nos métodos espaciais. Tal parâmetro deve ser ajustado de forma heurística.

No segundo algoritmo proposto, INLA SR, e no terceiro algoritmo, INLA MAP SR, um método de inferência Bayesiana aproximado, não paramétrico, recentemente desenvolvido, conhecido na literatura estatística como *Integrated Nested Laplace Approximation* (INLA), foi aplicado ao problema de Superresolução. Nesses algoritmos, o parâmetro λ é ajustado automaticamente.

Finalmente, os dois últimos algoritmos propostos, DFT INLA SR e Closed Form DFT, são versões mais rápidas dos algoritmos INLA SR e Closed Form SR, baseados na transformada 2D de Fourier. Nessa nova implementação o custo computacional, originalmente $O(n^3)$, é reduzido para $O(n^2 \log(n))$, com uma pequena perda da qualidade da imagem HR. Adicionalmente, a dimensão das matrizes manipuladas pelos algoritmos é reduzida de $n^2 \times n^2$ para $n \times n$ (o tamanho da imagem HR).

Experimentos demonstram que os algoritmos propostos nessa Tese são competitivos quando comparados a outros algoritmos do estado da arte.

Abstract

Superresolution (SR) multiframe image enhancement are techniques to increase the resolution of an image without increasing the resolution of the camera used to capture them. Image processing algorithms are used, instead. A set of observable lower resolution images is first registered (with respect to one of those images) and subsequently fused into a single high resolution image, which is then filtered to improve their sharpness. In such context, this PhD Thesis proposes five Superresolution algorithms based on Bayesian statistical modeling.

In the first algorithm developed, referred to as Closed Form SR, an equation is derived in closed form to calculate the high resolution image. However, the algorithm depends on a parameter λ , which corresponds to the regularization parameter in the spatial SR methods. In the case of this algorithm, however, such parameter should be heuristically set.

In the second and third algorithms, referred to as SR INLA and INLA MAP SR, respectively, a nonparametric method to approximate Bayesian inference recently developed, known in the statistical literature as Integrated Nested Laplace Approximation (INLA), was applied to the Superresolution problem. The advantage of such algorithms is that the parameter λ is automatically adjusted.

Finally, the last two proposed algorithms, referred to as DFT INLA and SR Closed Form DFT, respectively, are faster versions of the algorithms INLA SR and Closed Form SR, based on 2D Fourier transform. In this new implementation the computational cost, originally $O(n^3)$ is reduced to $O(n^2 \log(n))$, with a slight loss of quality of the HR image. Additionally, the size of the arrays manipulated by the algorithm is reduced from $n^2 \times n^2$ to $n \times n$ (the HR image size).

Experiments demonstrate that the proposed algorithms are competitive, when compared to other state-of-the-art algorithms.

Capítulo 1

Introdução ao Problema de Superresolução

Essa Tese tem como foco o problema de Superresolução *multiframe* de imagens. Esse termo descreve o processo, muitas vezes referenciado na literatura como melhoria de resolução, de se obter imagens de alta resolução a partir de um conjunto de imagens de baixa resolução observadas. Nesse contexto, o termo Superresolução é usado para descrever a capacidade que os métodos *multiframe* têm para aumentar a resolução espacial, gerando uma imagem com melhor qualidade do que é possível utilizando-se as técnicas tradicionais de restauração com uma única imagem. Assim, um algoritmo de interpolação, por exemplo, não é considerado um método de Superresolução.

O interesse acadêmico e comercial nas técnicas de Superresolução tem aumentado muito nos últimos anos. Isso se deve, principalmente, aos elevados custos envolvidos na fabricação dos sensores eletrônicos de imagem (*charge coupled device* - CCD) e dos componentes ópticos necessários para capturar imagens de alta resolução, que os torna proibitivamente caros para algumas aplicações.

As imagens de baixa resolução, utilizadas pelos algoritmos de Superresolução, são subamostradas e podem ter sido adquiridas por múltiplos sensores em uma única cena (Figura 1.1a) ou por um único sensor que captura uma cena por um período de tempo (Figura 1.1b) (Katsaggelos et al., 2007). Para cenas estáticas as observações são correlacionadas por um movimento global subpixel, causado, por exemplo, pelas posições relativas entre câmeras. Por outro lado, em cenas dinâmicas eles são correlacionados por um movimento local subpixel, devido ao movimento do objeto e, possivelmente, uma mudança de posicionamento da câmera. Nos dois casos, o algoritmo de Superresolução utiliza o conjunto de imagens de baixa resolução (Figura 1.1a) ou os quadros do vídeo (Figura 1.1b) para gerar uma imagem

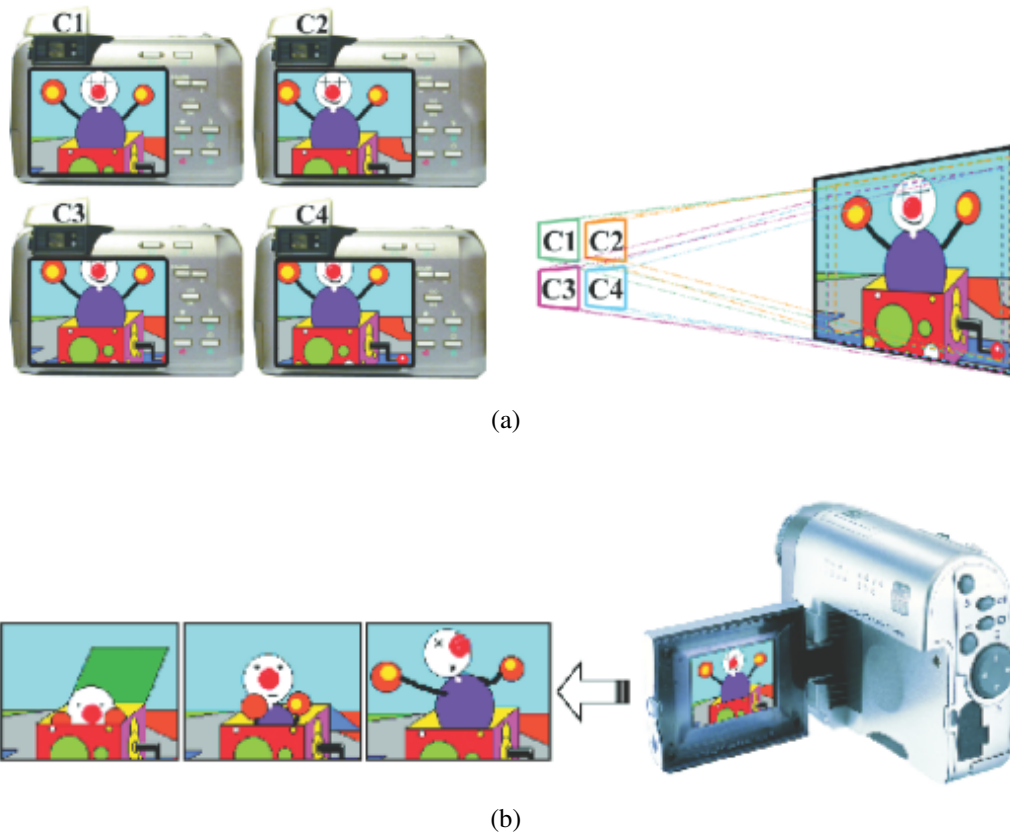


Figura 1.1: (a) Aquisição de imagens por múltiplas câmeras, (b) aquisição por meio de uma filmadora.

com maior resolução espacial.

De uma maneira geral, um algoritmo de Superresolução envolve três fases distintas. A primeira, conhecida como registro de imagens ou estimação de movimento, estima o movimento subpixel existente entre as imagens de baixa resolução. Assim, é possível trazer todas as imagens de baixa resolução para um mesmo referencial espacial. A segunda, diz respeito a fusão da imagem de alta resolução. Finalmente, a terceira está relacionada à melhoria final da imagem através da aplicação de filtros. Alguns algoritmos implementam somente as duas primeiras etapas como é o caso em (Vandewalle et al., 2006) e (Pham et al., 2006). Há casos também onde as etapas de fusão e filtragem são realizadas conjuntamente, como em (Farsiu et al., 2004b).

De uma maneira muito simples, o funcionamento de um algoritmo de Superresolução é ilustrado na Figura 1.2. O problema consiste em inferir a imagem de alta resolução com quatro pixels (Figura 1.2a) a partir das imagens de baixa resolução observadas, de um pixel cada (Figura 1.2b até 1.2e). Os quadros tracejados que aparecem nessa ilustração representam como as imagens de baixa resolução foram registradas, relativamente ao movimento subpixel entre elas, tomando como referência a Figura 1.2b. Após o registro é feita a fusão

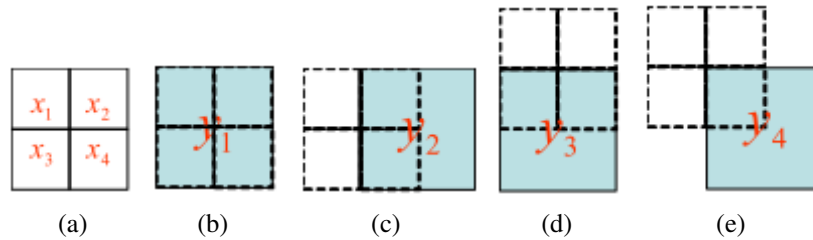


Figura 1.2: Exemplo ilustrativo de Superresolução (a) imagem de alta resolução de 4 pixels, (b)-(e) imagens de baixa resolução, com movimento subpixel, cada uma com um único pixel.

da imagem de alta resolução e, posteriormente, é aplicado um filtro para restauração da imagem.

A Figura 1.3 traz outro exemplo que ilustra a Superresolução. Intuitivamente, cada imagem de baixa resolução observada, onde as figuras geométricas nas Figuras 1.3a a 1.3d correspondem as posições relativas dos pixels, representa uma versão subamostrada (serrilhada) da versão original da cena. Esse serrilhamento não pode ser removido se somente uma imagem é processada. Por causa dos deslocamentos subpixel, cada imagem observada contém informações complementares. Com o conhecimento exato dos deslocamentos, as imagens podem ser combinadas para remover o serrilhamento e gerar uma imagem com maior resolução (Katsaggelos et al., 2007). Assumindo que a imagem da Figura 1.3e está disponível, isso é, assumindo que os movimentos subpixel são conhecidos, o problema de Superresolução se torna um problema de reamostragem, que converte uma imagem com amostragem não uniforme (Figura 1.3e) em uma uniformemente amostrada (Figura 1.3f).

As Aplicações práticas para os métodos de restauração de imagem por Superresolução estão crescendo rapidamente, conforme o tema ganha exposição. As técnicas de Superresolução podem ser aplicadas a uma grande variedade de campos, como:

- melhoria de imagens a partir de sequências de vídeo;
- TV de alta definição;
- melhoraria de imagens coloridas adquiridas por um CCD;
- monitoramento de vídeo;
- sensoriamento remoto;
- imagens médicas;
- pré-processamento de imagens para fins de reconhecimento de padrões.

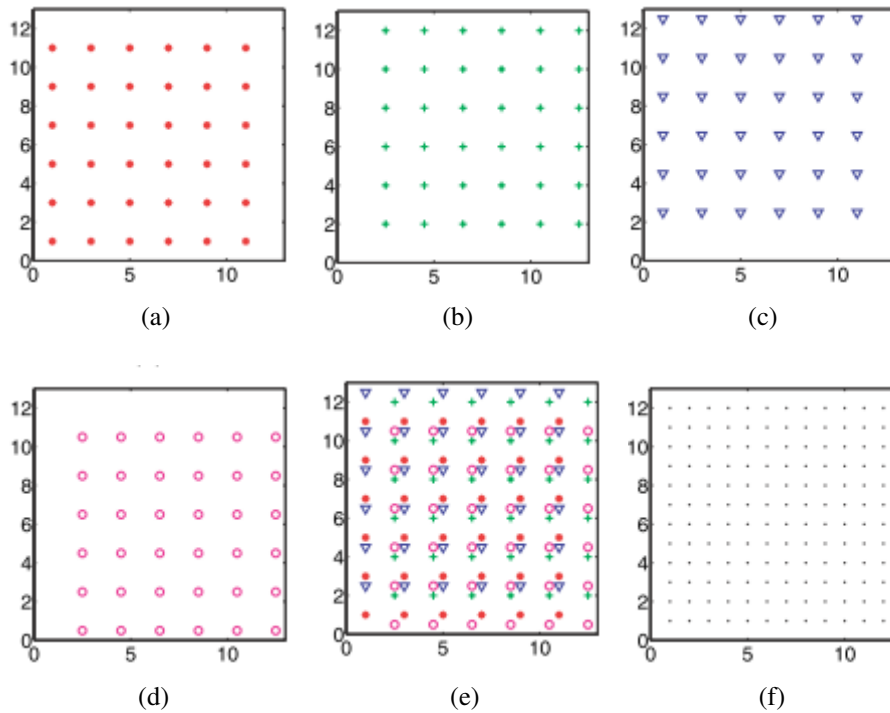


Figura 1.3: Exemplo ilustrativo de Superresolução (a)-(d) imagens de baixa resolução, com movimento subpixel, (e) imagem de alta resolução com grid irregular, (f) imagem de alta resolução com grid regular.

1.1 Algoritmos de Superresolução

Superresolução tem sido uma área muito ativa de pesquisa desde que no artigo (Tsai e Huang, 1984) foi publicado um algoritmo no domínio da frequência. Métodos no domínio da frequência são baseados em três princípios fundamentais:

- a propriedade de *shifting* da Transformada de Fourier (FT);
- a relação de *aliasing* entre a Transformada Contínua de Fourier (CFT) e a Transformada Discreta de Fourier (DFT);
- a cena original é limitada em banda. Essa propriedade permite a formulação de um sistema de equações relacionando os coeficientes *aliased* da DFT das imagens observadas com as amostras da CFT da cena desconhecida.

Essas equações são resolvidas produzindo os coeficientes no domínio da frequência da cena original, que podem ser recuperados pela inversa da DFT. Desde então, várias extensões ao

método básico de Tsai-Huang foram propostas (Tekalp et al., 1992), (Kim et al., 1990), (Kim e Su, 1993) e (Bose et al., 1993).

Algumas limitações dos métodos no domínio da frequência, como a limitada habilidade de incluir conhecimento *a priori* para regularização (Borman e Stevenson, 1999), causaram a gradual substituição destes métodos por algoritmos no domínio espacial. Seguindo essa linha, diversos algoritmos foram propostos, como *Interpolation of Non-Uniformly Spaced Samples* (Komatsu et al., 1993), *Iterated Backprojection* (Irani e Peleg, 1993) e *Projection Onto Convex Sets* (POCS) (Patti et al., 1997), (Tom e Katsaggelos, 1996), (Eren et al., 1997), entre outros.

A reconstrução de imagens por Superresolução é um exemplo de problema inverso mal posto, uma vez que múltiplas soluções existem para um dado conjunto de imagens observadas. Por causa disso, métodos de reconstrução por Superresolução utilizando regularização Tikhonov-Arsenin foram desenvolvidos (Hong et al., 1997). As funções de regularização que caracterizam essas soluções são tipicamente casos especiais de funções de densidade de probabilidade (PDF) *a priori* modeladas por campos Markovianos aleatórios no âmbito dos métodos Bayesianos.

Métodos estatísticos Bayesianos, que lidam com a reconstrução por Superresolução como um problema de estimação estatística, têm ganhado destaque nos últimos anos, por prover uma poderosa ferramenta teórica para a inclusão de condições de contorno *a priori* a fim de satisfazer a solução do problema inverso mal posto.

Diversos trabalhos recentes propõem métodos baseados em inferência Bayesiana. Em (Tipping e Bishop, 2003), por exemplo, *Evidence Approximation* é usada para resolver o problema de Superresolução. Em (Tian e Ma, 2010) um algoritmo baseado em Métodos de Monte Carlo via Cadeias de Markov (MCMC) é usado. De fato, os métodos conhecidos como *Evidence* e *Empirical Analysis*, as aproximações Laplacianas e *Variational* e a simulação de distribuições *a posteriori* têm grande potencial para reconstrução de imagens por Superresolução (Katsaggelos et al., 2007).

Concomitantemente, Havard Rue, Sara Martino and Nicolas Chopin (Rue et al., 2009) desenvolveram um novo método aproximado para inferência Bayesiana em modelos Gaussianos latentes, que ficou conhecido como *Integrated Nested Laplace Approximation* - INLA, que é um método não paramétrico, baseado em integração numérica.

Nesse contexto, esta Tese de Doutorado propõe um novo método de reconstrução de imagens por Superresolução, baseado em inferência Bayesiana aproximada. Primeiramente, a partir da modelagem do problema através de um modelo Bayesiano hierárquico, uma equação em forma fechada foi derivada para a fusão das imagens de baixa resolução (LR) em

uma imagem de alta resolução (HR), publicado em (Camponez et al., 2011b). A partir dessa equação, um novo algoritmo foi desenvolvido. Várias simulações mostraram que o método proposto resulta em imagens mais nítidas e menos ruidosas quando comparado a outros métodos recentemente desenvolvidos. Essa foi a primeira contribuição desta Tese. Entretanto, a qualidade da imagem de alta resolução gerada depende da escolha do valor de um hiperparâmetro λ , que corresponde, nos métodos no domínio espacial, ao coeficiente de regularização.

Assim, o passo seguinte, que é a segunda contribuição dessa Tese, é o desenvolvimento de um método totalmente automático capaz de estimar λ . Para isso, um método de inferência Bayesiana aproximado, não paramétrico, recentemente desenvolvido (Rue et al., 2009), conhecido na literatura estatística com *Integrated Nested Laplace Approximation* (INLA), foi aplicado ao problema de Superresolução. Até onde este autor conhece, é a primeira vez que a técnica INLA é aplicada em processamento de imagens. No desenvolvimento teórico do novo algoritmo é demonstrado matematicamente que todas as equações necessárias para a aplicação do INLA ao problema de Superresolução, entre elas a equação para fusão da imagem de alta resolução supracitada, podem ser escritas em forma fechada (Camponez et al., 2011a). As equações deduzidas permitiram o cálculo de dois vetores, \mathbf{T} que contém realizações da imagem de alta resolução e \mathbf{P} que contém as probabilidades associadas a cada uma delas, a partir dos quais dois algoritmos foram desenvolvidos. No primeiro, chamado de INLA SR, a imagem de alta resolução corresponde à esperança estatística de \mathbf{T} , que pode ser calculada através da média de seus elementos, ponderados por \mathbf{P} . No segundo, chamado de INLA MAP (Máximo a Posteriori), através de uma busca em \mathbf{P} , a realização com maior probabilidade é escolhida como imagem de alta resolução. O INLA MAP corresponde ao algoritmo de Superresolução inicialmente proposto onde a escolha de λ é automática. Por outro lado, o INLA SR é uma abordagem nova para o problema de Superresolução. Vários experimentos mostram que tanto o INLA MAP quanto o INLA SR, aqui propostos, quando comparados a outros algoritmos do estado da arte em Superresolução, alcançam excelentes resultados (Camponez et al., 2012c).

De certa forma, os métodos de simulação MCMC e o algoritmo de inferência INLA são bastante semelhantes: ambos são métodos de inferência estatística Bayesiana. Entretanto, enquanto o MCMC pode ser aplicado a qualquer que seja o modelo da PDF *a priori*, *a posteriori* ou a função de Verossimilhança, o INLA somente pode ser aplicado a modelos Gaussianos latentes, onde a matriz de covariância é governada por poucos parâmetros e o campo latente é um campo aleatório Gaussiano Markoviano - GMRF, com matriz de precisão \mathbf{Q} (inversa da matriz de variância/covariância) esparsa. Então, no caso geral de inferência estatística Bayesiana o método MCMC deve ser usado. Somente em casos especiais onde o modelo estatístico é Gaussiano latente, como no problema de reconstrução de imagens por Superresolução, o INLA pode ser usado, como proposto nesta Tese.

A aplicação dos métodos MCMC (Tian e Ma, 2010) e INLA ao problema da reconstrução por Superresolução merece alguns comentários adicionais. No caso do MCMC, a imagem de alta resolução é gerada pela média aritmética das amostras de alta resolução realizadas, depois de descartadas as amostras do período de *burn*¹. Então, um número muito grande de amostras é necessário para a geração de uma boa imagem de alta resolução. No caso do INLA, por outro lado, somente as amostras que correspondem às maiores massas de probabilidade são geradas, e a imagem de alta resolução pode ser computada de duas formas: como a média ponderada das imagens realizadas, usando suas respectivas probabilidades como peso; ou pode-se usar como imagem HR a amostra que corresponde a maior probabilidade. Então, somente poucas imagens precisam ser consideradas.

Entretanto, as matrizes manipuladas pelos algoritmos INLA SR e INLA MAP são muito grandes, embora esparsas. Então, mesmo usando-se estruturas de dados apropriadas para armazenar matrizes esparsas, o uso de memória e o esforço computacional são intensos. Nesse sentido, a terceira contribuição dessa Tese, publicada em (Camponez et al., 2012b; Componez et al., 2012a), é na direção de melhorar o desempenho desses algoritmos. Assim, envolvendo a imagem HR em um toróide \mathcal{T}_n e relaxando uma equação em (Camponez et al., 2011a; Componez et al., 2012c), dois algoritmos para reconstrução de imagens por Superresolução muito rápidos, chamados nessa Tese de DFT Closed Form SR e DFT INLA SR, baseado na transformada de Fourier 2D (2D DFT), são desenvolvidos. Nessa nova implementação o custo computacional, originalmente $O(n^3)$, é reduzido para $O(n^2 \log(n))$, com uma pequena perda da qualidade da imagem HR. Adicionalmente, a dimensão das matrizes manipuladas pelo algoritmo é reduzida de $n^2 \times n^2$ para $n \times n$ (o tamanho da imagem HR).

1.2 Objetivos desta Tese

Essa Tese tem como objetivos:

- a) desenvolver algoritmos de Superresolução, cujo objetivo é a obtenção de uma única imagem de alta resolução a partir de um conjunto de imagens capturadas, baseado em inferência estatística aproximada, usando um método Bayesiano não paramétrico, recentemente desenvolvido, conhecido como *Integrated Nested Laplace Approximation*;
- b) mostrar, através de experimentos, que o método proposto, quando comparado ao estado da arte em Superresolução, obtém resultados competitivos.

¹O período de *burn* corresponde às primeiras amostras geradas no MCMC, que devem ser descartadas.

1.3 Metodologia Utilizada

Nessa subseção será descrita a metodologia utilizada de forma a alcançar os objetivos propostos nessa Tese:

- o primeiro passo é a descrição do problema, seguido de uma revisão bibliográfica dos principais métodos de Superresolução, passando a definição do modelo de aquisição de imagens e, também, de inferência estatística Bayesiana;
- em seguida, o desenvolvimento teórico que envolve a derivação matemática da equação em forma fechada para o cálculo da imagem de alta resolução, a derivação de todas as equações, em forma fechada, necessárias para aplicação do método INLA à Superresolução, assim como o desenvolvimento matemático para possibilitar explorar uma versão rápida dos algoritmos propostos, baseada na Transformada de Fourier 2D;
- uma vez pronto o desenvolvimento teórico, segue-se a implementação dos algoritmos propostos, dentro dos melhores padrões e técnicas disponíveis de programação e de Álgebra Numérica Computacional. A seguir, através de vários experimentos, utilizando-se imagens de baixa resolução, umas simuladas e outras capturadas por câmeras, os algoritmos desenvolvidos serão comparados com outros do estado da arte em Superresolução;
- finalmente, encerra-se essa etapa da pesquisa com a análise dos resultados encontrados, com proposições de possíveis melhorias e continuações e a publicação de artigos científicos em congressos e em revistas da área de processamento de imagens.

1.4 Organização da Tese

No Capítulo 2 são apresentados o modelo de aquisição de imagens de baixa resolução, a modelagem Bayesiana hierárquica, a dedução da equação em forma fechada para geração da imagem de alta resolução, simulações e comparações do algoritmo proposto com outros do estado da arte em Superresolução. No Capítulo 3 é feita uma revisão bibliográfica do método de inferência estatística Bayesiana INLA. Em seguida, o INLA é aplicado ao problema de Superresolução. Para isso, todas as equações necessárias são derivadas em forma fechada, gerando um novo algoritmo. Finalmente, alguns experimentos são apresentados e o novo algoritmo é comparado a outros do estado da arte em Superresolução. No Capítulo 4, relaxando-se uma equação no modelo matemático e envolvendo a imagem HR em um toróide \mathcal{T}_n , uma versão muito rápida do algoritmo, baseada na transformada de Fourier

2D, é desenvolvida. Alguns experimentos mostram que, ao preço de uma pequena perda na qualidade da imagem HR gerada, o custo computacional pode ser reduzido de $O(n^3)$ para $O(n^2 \log(n))$. Além disso, a dimensão das matrizes manipuladas pelo algoritmo é reduzida de $n^2 \times n^2$ para $n \times n$, que corresponde ao tamanho da imagem HR. Finalmente, no Capítulo 5 são apontadas as conclusões dessa Tese e apresentadas propostas para continuação desta pesquisa.

Capítulo 2

Solução em Forma Fechada para Superresolução

O objetivo desse Capítulo é desenvolver um algoritmo de reconstrução de imagens por Superresolução baseado em uma solução analítica em forma fechada. Para isso, primeiramente a notação utilizada durante esse trabalho é definida. Então, são descritos o modelo de aquisição de imagens e, em seguida, o modelo Hierárquico Bayesiano, que serão utilizados. A partir desse ponto, uma equação em forma fechada para a reconstrução de uma imagem de alta resolução (HR), tendo como entrada um conjunto de imagens de baixa resolução (LR) observáveis, é derivada. No final do Capítulo alguns experimentos são apresentados, e o algoritmo proposto tem seu desempenho comparado a outros do estado da arte em Superresolução.

2.1 Notação

A Figura 2.1 mostra uma aplicação típica para Superresolução. Nesse exemplo, uma sequência de imagens LR, capturada por uma câmera de vídeo, é usada para reconstruir imagens HR. Assim, X_i representa a i -ésima imagem HR. O conjunto de imagens LR pode ser agrupado em um vetor \mathbf{Y} . Cada elemento y_i em \mathbf{Y} é uma imagem LR no instante i . Todos os parâmetros requeridos para gerar as imagens LR, a partir da imagem HR, conforme modelo de aquisição de imagens adotado, são assumidos conhecidos, exceto os parâmetros do modelo translacional/rotacional (deformações), inferidos a partir das imagens de baixa resolução por algum algoritmo de registro de imagens.

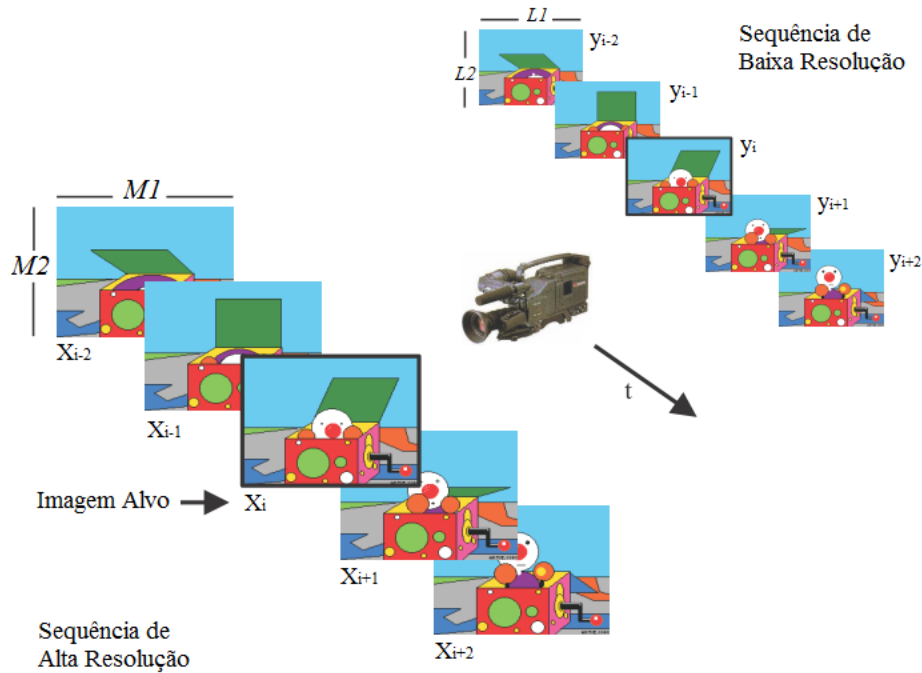


Figura 2.1: Nomenclatura usada em Superresolução

2.2 Modelo de Aquisição de Imagens de Baixa Resolução

Para efeito de avaliação dos algoritmos desenvolvidos nesta Tese é necessário conhecer, a priori, a imagem HR, para compará-la à imagem HR que será construída com base nos algoritmos propostos. Daí, as imagens de LR, as observações disponíveis, serão geradas a partir da imagem HR disponível. Em seguida, partindo das imagens LR, a imagem HR será reconstruída, e o resultado dessa reconstrução é comparado com a imagem HR original.

A relação entre imagens HR e LR é descrita por um modelo de aquisição de imagens, conforme ilustrado na Figura 2.2. De acordo com esse modelo, a imagem original, que é uma cena contínua, pode ser representada por uma imagem HR discretizada sem degradação (*aliasing*), resultante de uma amostragem do espaço contínuo numa taxa superior à de Nyquist. Essa imagem é então deformada e distorcida. As deformações representam translações, rotações ou outros mapeamentos mais sofisticados que são requeridos para gerar as imagens LR observadas. A modelagem dessas deformações é feita através de algoritmos de registro de imagem que estimam o movimento relativo entre as imagens HR através dos quadros LR. As distorções são causadas pelo sistema ótico nas lentes, pelos fotossensores, e pela subamostragem no sensor (CCD). Através de uma operação de decimação, isto é, de redução da resolução ou subamostragem, como é conhecido na literatura de processamento de sinais, é gerada a imagem LR, que nesse caso contém degradação. Finalmente, é adicionado ruído à saída do sistema, gerando as imagens LR observadas. O ruído modela vários elementos

na cadeia para formação de imagens, tais como: erro de registro de imagens; ruído térmico ou da própria eletrônica utilizada, ou mesmo erros no armazenamento ou transmissão das imagens.

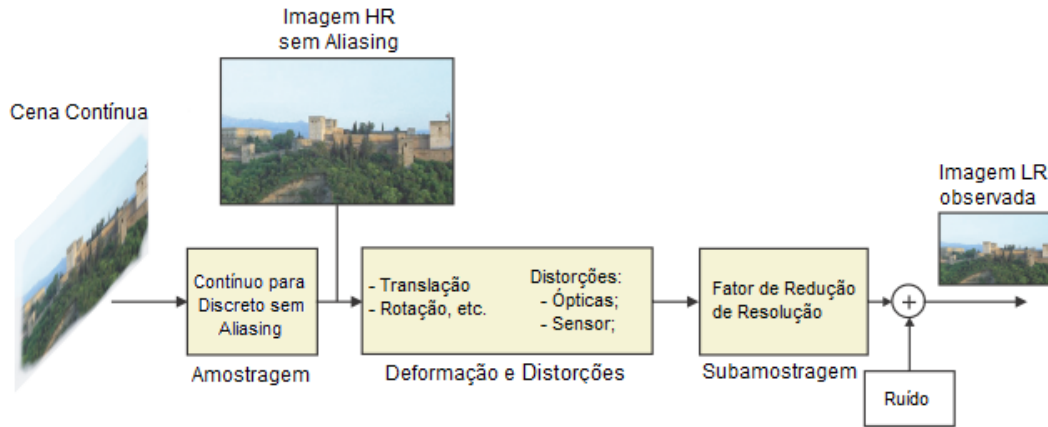


Figura 2.2: Modelo de Aquisição de Imagem

Uma questão interessante em relação ao modelo de aquisição é a ordem na qual as operações de deformação (*warp*) e distorção (*blur*) são aplicadas à imagem de alta resolução, o que acaba gerando, na verdade, dois modelos distintos, que ficaram conhecidos na literatura como: *warp-blur* e *blur-warp*, conforme (Wang e Qi, 2004). Os autores afirmam que, quando o movimento deve ser estimado a partir das imagens LR, o uso do modelo *warp-blur* pode causar erros sistemáticos e, portanto, nesse caso, é mais apropriado o uso do modelo *blur-warp*. Foi mostrado que quando as deformações na imagem são espaço-temporalmente invariantes e o movimento tem apenas um componente global de translação, os dois modelos coincidem. Nesse caso, as matrizes de degradação e deformação correspondem à matrizes de convolução e, então, podem ser comutadas. Nessa Tese as deformações nas imagens são assumidas espaço-temporalmente invariantes e o modelo *warp-blur* é adotado.

2.2.1 Modelo de Aquisição *Warp-Blur*

Como o nome já indica, nesse modelo de aquisição de imagens o efeito de deformação (*warp*), causado pelo movimento relativo entre as imagens, é aplicado anteriormente ao de distorção/degradação (*blur*), causado pelas distorções ópticas e pela subamostragem.

No modelo *warp-blur*, apresentado na Figura 2.3, \mathbf{y}_k é um vetor, de tamanho $L_1 L_2 \times 1$, cujos elementos correspondem aos pixels da imagem LR, $y_k[l_1, l_2]$, no instante k , ordenados lexicograficamente¹. L_1 e L_2 representam o número de pixels na horizontal e vertical,

¹Ordenados na forma de um vetor coluna.

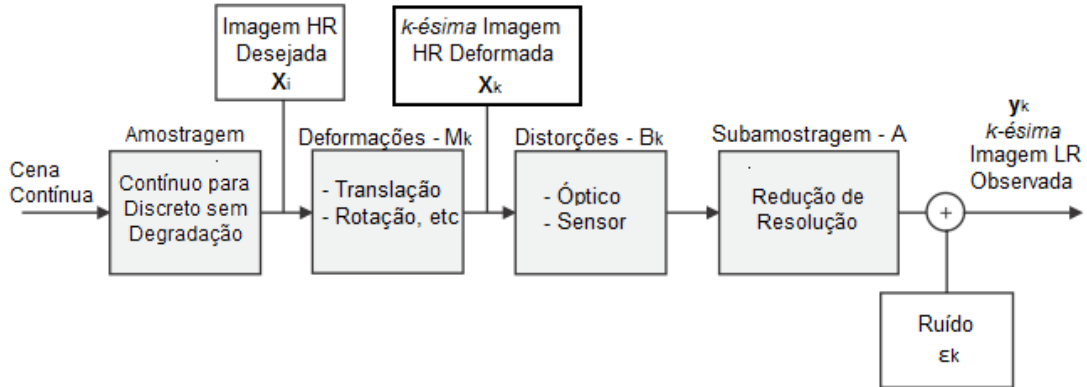


Figura 2.3: Modelo Warp-Blur para aquisição de Imagem

respectivamente. \mathbf{X}_k é um vetor, de tamanho $M_1 M_2 \times 1$, cujos elementos correspondem aos pixels da imagem HR, $X_k[m_1, m_2]$, no instante k , ordenados lexicograficamente. Da mesma forma, M_1 e M_2 representam o número de pixels na horizontal e vertical. Finalmente, ϵ_k é um vetor $L_1 L_2 \times 1$ que representa o efeito conjunto dos erros de aquisição e de estimação de movimento.

A matriz \mathbf{D}_k , de tamanho $L_1 L_2 \times M_1 M_2$, que representa o efeito combinado das distorções \mathbf{B}_k e da subamostragem \mathbf{A} , modela o sistema de aquisição para a imagem k . A coluna m , em \mathbf{D}_k , representa a Função de Espalhamento do pixel m da imagem HR durante a aquisição. A linha n , que representa a Função de Aquisição de pixel, contém a ponderação de cada pixel da imagem HR na formação do pixel n na imagem LR. Assim,

$$\mathbf{D}_k = \mathbf{A} \mathbf{B}_k. \quad (2.1)$$

No caso em que as imagens LR e HR são temporalmente coincidentes, o índice k , que representa a posição temporal das imagens, é igual a i , conforme a modelo da Figura 2.1. Além disso, nesse caso \mathbf{M}_k é a matriz identidade. Assim, a relação matemática entre as imagens LR e HR pode ser modelada como

$$\mathbf{y}_i = \mathbf{D}_i \mathbf{X}_i + \eta_i. \quad (2.2)$$

onde, η_i é um vetor $L_1 L_2 \times 1$ que representa o ruído de aquisição.

A Figura 2.4 mostra um exemplo do efeito da degradação ocorrido em uma imagem HR pela aquisição da imagem LR.

Quando as imagens LR e a HR não estão no mesmo instante temporal, por causa do movimento relativo entre as imagens da sequência, um modelo de movimento, que descreve as

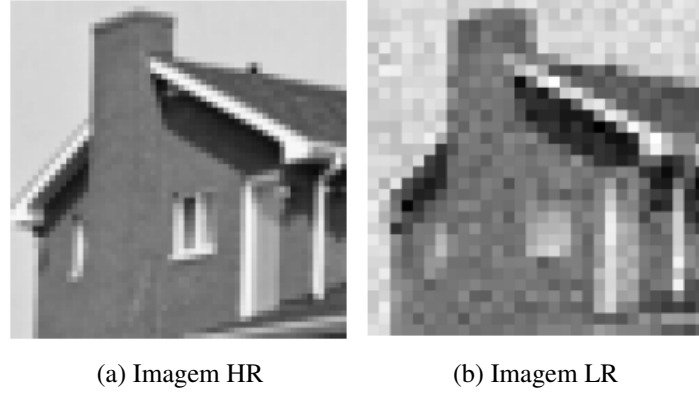


Figura 2.4: Efeito de Degradação na Aquisição de Imagens

transformações entre imagens, deve ser usado. Assume-se que a imagem no instante temporal k pode ser composta pela imagem do instante temporal i , com movimento compensado, mais uma nova informação, que não pode ser obtida da imagem no instante k . A equação que descreve o modelo é

$$\mathbf{X}_k = \mathbf{M}_{k,i} \mathbf{X}_i + \boldsymbol{\psi}_{k,i} \quad (2.3)$$

onde, a matriz $\mathbf{M}_{k,i}$, de tamanho $M_1 M_2 \times M_1 M_2$, representa a transformação de movimento do instante temporal i para o k , e $\boldsymbol{\psi}_{k,i}$ é um vetor de tamanho $M_1 M_2 \times 1$, que modela o erro na inferência dos parâmetros de compensação do movimento entre as imagens.

No *warp-blur* os modelos de aquisição de imagens e de movimento são combinados. Dessa forma, é possível relacionar imagens LR e HR em instantes temporais diferentes. Assim, substituindo-se (2.3) em (2.2), obtém-se

$$\mathbf{y}_k = \mathbf{D}_k (\mathbf{M}_{k,i} \mathbf{X}_i + \boldsymbol{\psi}_{k,i}) + \boldsymbol{\eta}_k, \quad (2.4)$$

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{M}_{k,i} \mathbf{X}_i + \mathbf{D}_k \boldsymbol{\psi}_{k,i} + \boldsymbol{\eta}_k, \quad (2.5)$$

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{M}_{k,i} \mathbf{X}_i + \boldsymbol{\xi}_{k,i} + \boldsymbol{\eta}_k, \quad (2.6)$$

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{M}_{k,i} \mathbf{X}_i + \boldsymbol{\varepsilon}_{k,i}, \quad (2.7)$$

$$\mathbf{y}_k = \mathbf{H}_{k,i} \mathbf{X}_i + \boldsymbol{\varepsilon}_{k,i}. \quad (2.8)$$

A matriz $\mathbf{H}_{k,i}$ em (2.8), de tamanho $L_1 L_2 \times M_1 M_2$, representa a modelagem conjunta de movimento e aquisição. O erro $\boldsymbol{\varepsilon}_{k,i}$, que é o efeito conjunto dos erros de aquisição $\boldsymbol{\eta}_k$ e de movimento $\boldsymbol{\xi}_{k,i}$, é assumido Gaussiano com média zero e variância $\boldsymbol{\sigma}_k^2$

$$\boldsymbol{\varepsilon}_{k,i} \sim NID[0, \boldsymbol{\sigma}_k^2]. \quad (2.9)$$

A notação em (2.8) pode ser simplificada, sem perda de generalidade, de forma a tornar o texto mais limpo e facilitar o entendimento dos desenvolvimentos matemáticos que serão apresentados no decorrer dessa Tese. Assim, considerando o índice i , que representa o instante temporal da imagem HR, igual a zero, tem-se a nova notação

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{X} + \boldsymbol{\varepsilon}_k, \quad (2.10)$$

que descreve o modelo *warp-blur* de formação de imagens LR observáveis a partir de imagens HR, adotado nesse trabalho.

Como pode ser visto, a reconstrução de imagens por Superresolução é um problema inverso e mal-posto. É inverso pois o modelo de aquisição, descrito pela equação (2.10), tem como entrada a imagem HR e como saída as imagens LR, e o que se deseja em um algoritmo de Superresolução é exatamente o inverso, isto é, a construção de uma imagem HR a partir de uma sequência de imagens observáveis LR, é mal-posto pois não tem solução única. Assim, observada uma sequência de imagens LR, existem várias imagens \mathbf{X} , diferentes, que representam uma possível solução para o problema.

2.2.2 Sistema de Aquisição

A Figura 2.5 mostra um exemplo de um sistema de aquisição, que modela simultaneamente as operações de distorção e decimação (Zibetti, 2007).

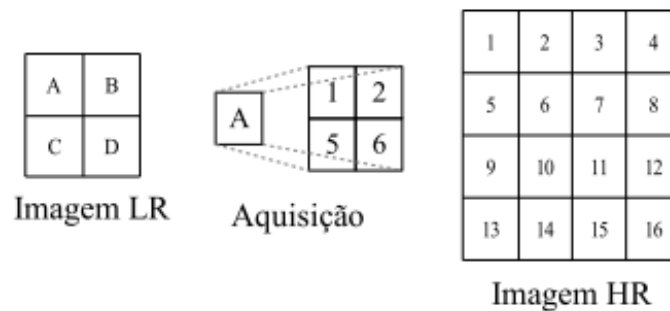


Figura 2.5: Modelo de Aquisição

As imagens LR e HR, na Figura 2.5, são temporalmente coincidentes. Cada uma delas é representada lexicograficamente pelos vetores \mathbf{y}_1 e \mathbf{X} , de dimensões (4×1) e (16×1) , em pixels. Elas se relacionam, conforme equação (2.2), através de

$$\mathbf{y}_1 = \mathbf{D}_1 \mathbf{X} + \boldsymbol{\eta}_1, \quad (2.11)$$

onde,

$$\begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \end{bmatrix} + \eta_1, \quad (2.12)$$

e

1. $[A, B, C, D]^T$ representa os pixels da imagem LR;
2. $[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]^T$ representa os pixels da imagem HR;
3. \mathbf{D}_1 representa o sistema de aquisição de imagens, onde cada pixel da imagem de baixa resolução corresponde ao valor médio de um quadrado de área 2×2 pixels na imagem de alta resolução. Isso corresponde a aplicar um filtro passa-baixas na imagem de alta resolução, seguido de uma operação de decimação com um fator dois em ambas as direções;
4. η_1 representa um erro aleatório do sistema de aquisição.

2.2.3 Modelagem de Movimento

A Figura 2.6 apresenta um exemplo de transformação de movimento, considerando um movimento translacional global de um pixel para a direita, onde as bordas da imagem são assumidas periódicas.

Dessa forma, as imagens 1 e 2, de tamanho 9 pixels cada, representadas lexicograficamente pelos vetores \mathbf{X}_1 e \mathbf{X}_2 , de dimensões 9×1 , se relacionam através da transformação

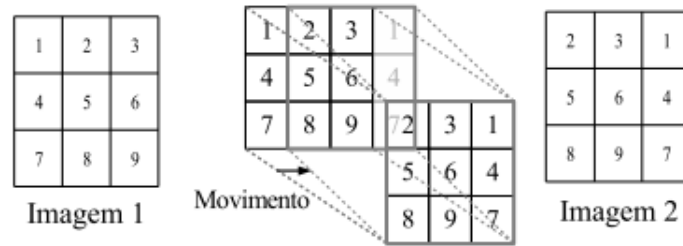


Figura 2.6: Exemplo de Transformação de Movimento

$$\mathbf{X}_1 = \mathbf{M}_{1,2} \mathbf{X}_2, \quad (2.13)$$

onde $\mathbf{M}_{1,2}$ é uma matriz, de tamanho 9×9 , que mapeia a transformação da imagem 2 para 1, ou seja, a mudança de posição dos pixels de uma imagem para outra, e pode ser escrito como

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \\ 5 \\ 6 \\ 4 \\ 8 \\ 9 \\ 7 \end{bmatrix}. \quad (2.14)$$

Embora a dimensão de $\mathbf{M}_{1,2}$, comparada à imagem HR, seja muito grande, uma significativa proporção de seus elementos são iguais a zero. Em (2.14), por exemplo, de 81 elementos somente 9 (11,11%) são diferentes de zero. Essa característica de esparsidade pode ser aproveitada no algoritmo de Superresolução, a fim de reduzir o consumo de memória necessária para armazenar tal matriz.

2.3 Modelagem Hierárquica Bayesiana

Nesta seção, uma abordagem de inferência Bayesiana (Hilbe et al., 2007; Izenman, 2008) é explorada para estabelecer a articulação entre a densidade de probabilidade condicional (PDF) *a posteriori* da imagem \mathbf{X} HR e do hiperparâmetro λ , que faz parte do modelo da PDF *a priori* da imagem HR, dado um conjunto de imagens LR observáveis \mathbf{Y} . O escalar

λ equivale, nos métodos de Superresolução Regularizados e, também, nos algoritmos de restauração de imagens, ao coeficiente de regularização. Assim, a PDF conjunta pode ser escrita como

$$p(\mathbf{X}, \mathbf{Y}, \lambda) = p(\mathbf{Y}|\mathbf{X}, \lambda)p(\mathbf{X}, \lambda), \quad (2.15)$$

ou

$$p(\mathbf{X}, \mathbf{Y}, \lambda) = p(\mathbf{Y}|\mathbf{X}, \lambda)p(\mathbf{X}|\lambda)p(\lambda), \quad (2.16)$$

ou, de outra forma,

$$p(\mathbf{X}, \mathbf{Y}, \lambda) = p(\mathbf{X}, \lambda|\mathbf{Y})p(\mathbf{Y}). \quad (2.17)$$

Assim, igualando as equações (2.16) e (2.17), tem-se

$$p(\mathbf{X}, \lambda|\mathbf{Y})p(\mathbf{Y}) = p(\mathbf{Y}|\mathbf{X}, \lambda)p(\mathbf{X}|\lambda)p(\lambda), \quad (2.18)$$

que pode ser reescrita como

$$p(\mathbf{X}, \lambda|\mathbf{Y}) \propto p(\mathbf{Y}|\mathbf{X}, \lambda)p(\mathbf{X}|\lambda)p(\lambda). \quad (2.19)$$

Nesse caso, o sinal de proporcionalidade aparece em (2.19) pois essa equação difere de (2.18) apenas por uma constante $p(\mathbf{Y})$.

A equação (2.19) representa o modelo Hierárquico Bayesiano adotado, onde $p(\mathbf{Y}|\mathbf{X}, \lambda)$ é a PDF condicional das imagens LR dada a imagem HR e o hiperparâmetro do modelo *a priori* da imagem HR; $p(\mathbf{X}|\lambda)$ é a PDF *a priori* da imagem HR dado o hiperparâmetro e, finalmente, $p(\lambda)$ é a PDF marginal do hiperparâmetro.

2.3.1 Função de Probabilidade Condicional de Verossimilhança

O primeiro termo da equação (2.19), $p(\mathbf{Y}|\mathbf{X}, \lambda)$, é conhecido na literatura estatística como Função de Probabilidade Condicional de Verossimilhança ou *Likelihood*. Assumindo que as imagens LR são obtidas independentemente a partir da imagem original HR, e, também, que não dependem do hiperparâmetro λ , essa PDF pode ser escrita como

$$p(\mathbf{Y}|\mathbf{X}, \lambda) = \prod_{k=1}^m p(\mathbf{y}_k|\mathbf{X}), \quad (2.20)$$

onde m é o número de imagens LR.

Além disso, de acordo com (2.10), a PDF condicional $p(\mathbf{y}_k|\mathbf{X})$ pode ser expressa como

$$p(\mathbf{y}_k|\mathbf{X}) = p(\mathbf{H}_k \mathbf{X} + \boldsymbol{\varepsilon}_k|\mathbf{X}). \quad (2.21)$$

Uma vez que \mathbf{H}_k é uma matriz determinística, o termo $\mathbf{H}_k\mathbf{X}$, dado \mathbf{X} , também é determinístico. Por conseguinte, (2.21) pode ser ainda mais simplificada, removendo o termo $\mathbf{H}_k\mathbf{X}$, para se chegar a

$$p(\mathbf{y}_k|\mathbf{X}) = p(\boldsymbol{\varepsilon}_k|\mathbf{X}) = p(\boldsymbol{\varepsilon}_k) \propto \exp\left(-\frac{1}{2\sigma_k^2}\|\mathbf{y}_k - \mathbf{H}_k\mathbf{X}\|^2\right). \quad (2.22)$$

Substituindo (2.22) em (2.20) obtém-se

$$p(\mathbf{Y}|\mathbf{X}, \lambda) \propto \prod_{k=1}^m \exp\left(-\frac{1}{2\sigma_k^2}\|\mathbf{y}_k - \mathbf{H}_k\mathbf{X}\|^2\right) = \exp\left(-\sum_{k=1}^m \frac{1}{2\sigma_k^2}\|\mathbf{y}_k - \mathbf{H}_k\mathbf{X}\|^2\right), \quad (2.23)$$

que representa a Função de Probabilidade Condicional de Verossimilhança ou Verossimilhança, $p(\mathbf{Y}|\mathbf{X}, \lambda)$, adotada nessa Tese.

2.3.2 Função de Probabilidade Condicional a Priori da Imagem HR

O segundo termo da equação (2.19), $p(\mathbf{X}|\lambda)$, é conhecido na literatura estatística como PDF condicional *a priori* da imagem HR, e, em geral, é modelado como um vetor aleatório localmente suave. Um campo aleatório Markoviano Gaussiano (Gaussian Markov *random field* - GMRF) (Li, 1995), é considerado uma aproximação razoável para esse modelo, uma vez que o GMRF tem mostrado ser eficaz para o problema de reconstrução de imagem por SR (Hardie et al., 1997). Tal modelo tem como formulação matemática (Rue e Held, 2005)

$$p(\mathbf{X}|\lambda) \propto \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} |\lambda\mathbf{Q}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\lambda\mathbf{X}^T\mathbf{Q}\mathbf{X}\right), \quad (2.24)$$

onde \mathbf{Q} é uma matriz de precisão $N \times N$ ($N = M_1M_2$), que é a inversa da matriz de covariância, cujas entradas, considerando uma vizinhança 4, como em (Rue e Held, 2005), são dadas por

$$Q_{ij} = \begin{cases} 4 & \text{se } i = j; \\ -1 & \text{se } i, j \text{ estão em uma vizinhança 4;} \\ 0 & \text{outros casos.} \end{cases} \quad (2.25)$$

A matriz \mathbf{Q} gerada por (2.25) é esparsa, o que possibilita a aplicação de algoritmos computacionais com custo reduzido (Rue e Held, 2005) e a aplicação de um método estatístico conhecido como *Integrated Nested Laplace Approximation* (INLA), que será explorado no próximo Capítulo.

2.3.3 Função de Probabilidade Marginal do Hiperparâmetro

O último termo em (2.19), a PDF do hiperparâmetro, $p(\lambda)$, é definida como uma distribuição uniforme, que, conforme (Galatsanos et al., 2000) e (Figueiredo e Nowak, 2001), é uma suposição razoável para o problema de reconstrução de imagem. Isso implica que

$$p(\lambda) = \frac{1}{\lambda_{max} - \lambda_{min}}, \text{ para } \lambda \in [\lambda_{min}, \lambda_{max}], \quad (2.26)$$

onde $[\lambda_{min}, \lambda_{max}]$ representa a faixa dinâmica. No Capítulo 3, a Seção 3.2.2 apresenta um detalhamento de como essa faixa é determinada.

2.3.4 Modelo Hierárquico Bayesiano

Finalmente, introduzindo (2.23), (2.24) e (2.26) em (2.19), obtém-se

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto \frac{1}{\lambda_{max} - \lambda_{min}} |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(-\frac{1}{2} \lambda \mathbf{X}^T \mathbf{Q} \mathbf{X} - \sum_{k=1}^m \frac{1}{2\sigma_k^2} \|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 \right), \quad (2.27)$$

que é o Modelo Hierárquico Bayesiano da PDF $p(\mathbf{X}, \lambda | \mathbf{Y})$ adotado nessa Tese.

2.4 Derivando uma Forma Fechada para o algoritmo de Superresolução

Nesta seção a partir da equação (2.27), uma solução analítica é derivada para a reconstrução da imagem HR, como segue. O ponto inicial é

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(-\frac{1}{2} \lambda \mathbf{X}^T \mathbf{Q} \mathbf{X} - \sum_{k=1}^m \frac{1}{2\sigma_k^2} \|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 \right), \quad (2.28)$$

onde

$$\|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = (\mathbf{y}_k - \mathbf{H}_k \mathbf{X})^T (\mathbf{y}_k - \mathbf{H}_k \mathbf{X}), \quad (2.29)$$

ou

$$\|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = \mathbf{y}_k^T \mathbf{y}_k - 2\mathbf{y}_k^T \mathbf{H}_k \mathbf{X} + \mathbf{X}^T \mathbf{H}_k^T \mathbf{H}_k \mathbf{X}. \quad (2.30)$$

Considerando as variáveis

$$\mathbf{m}_k = \mathbf{H}_k^T \mathbf{H}_k, \quad (2.31)$$

$$\mathbf{b}_k = -2\mathbf{H}_k^T \mathbf{y}_k, \quad (2.32)$$

e

$$\mathbf{c}_k = \mathbf{y}_k^T \mathbf{y}_k, \quad (2.33)$$

obtém-se

$$\|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = \mathbf{X}^T \mathbf{m}_k \mathbf{X} + \mathbf{b}_k^T \mathbf{X} + \mathbf{c}_k. \quad (2.34)$$

De tal resultado,

$$\sum_{k=1}^m \|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = \sum_{k=1}^m \mathbf{X}^T \mathbf{m}_k \mathbf{X} + \sum_{k=1}^m \mathbf{b}_k^T \mathbf{X} + \sum_{k=1}^m \mathbf{c}_k, \quad (2.35)$$

cujos termos são desenvolvidos como

$$\sum_{k=1}^m \mathbf{X}^T \mathbf{m}_k \mathbf{X} = \mathbf{X}^T (\mathbf{m}_1 + \dots + \mathbf{m}_m) \mathbf{X} = \mathbf{X}^T \mathbf{M} \mathbf{X}, \quad (2.36)$$

$$\mathbf{M} = \sum_{k=1}^m \mathbf{m}_k, \quad (2.37)$$

$$\sum_{k=1}^m \mathbf{b}_k^T \mathbf{X} = (\mathbf{b}_1 + \dots + \mathbf{b}_m)^T \mathbf{X} = \mathcal{B}^T \mathbf{X}, \quad (2.38)$$

$$\mathcal{B} = \sum_{k=1}^m \mathbf{b}_k, \quad (2.39)$$

e

$$\mathbf{C} = \sum_{k=1}^m \mathbf{c}_k. \quad (2.40)$$

Agora, introduzindo (2.37), (2.39) e (2.40) em (2.35), e supondo que todas as imagens de baixa resolução têm a mesma variância do ruído, obtém-se

$$\sum_{k=1}^m \frac{1}{2\sigma_k^2} \|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = \frac{1}{2\sigma^2} (\mathbf{X}^T \mathbf{M} \mathbf{X} + \mathcal{B}^T \mathbf{X} + \mathbf{C}), \quad (2.41)$$

onde aqui se assumiu $\sigma_k = \sigma$ para $k = 1, \dots, n$. Então, introduzindo (2.41) em (2.28),

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(- \left(\frac{1}{2} \mathbf{X}^T (\lambda \mathbf{Q} + \frac{\mathbf{M}}{\sigma^2}) \mathbf{X} + \frac{\mathcal{B}^T \mathbf{X}}{2\sigma^2} + \frac{\mathbf{C}}{2\sigma^2} \right) \right). \quad (2.42)$$

Substituindo as variáveis

$$\mathbf{B}^T = \frac{\mathcal{B}^T}{2\sigma^2}, \quad (2.43)$$

$$\mathbf{K} = \frac{\mathbf{C}}{2\sigma^2}, \quad (2.44)$$

$$\mathbf{A} = \left(\lambda \mathbf{Q} + \frac{\mathbf{M}}{\sigma^2} \right), \quad (2.45)$$

obtém-se

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(- \left(\frac{1}{2} \mathbf{X}^T \mathbf{A} \mathbf{X} + \mathbf{B}^T \mathbf{X} + \mathbf{K} \right) \right), \quad (2.46)$$

que é uma equação bem conhecida na literatura. De (Bishop, 2006), a identidade

$$\frac{1}{2}\mathbf{X}^T\mathbf{A}\mathbf{X} + \mathbf{B}^T\mathbf{X} + \mathbf{K} = \frac{1}{2}(\mathbf{X} + \mathbf{A}^{-1}\mathbf{B})^T(\mathbf{X} + \mathbf{A}^{-1}\mathbf{B}) + \mathbf{K} - \frac{1}{2}\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \quad (2.47)$$

é válida, e então

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{X} + \mathbf{A}^{-1}\mathbf{B})^T(\mathbf{X} + \mathbf{A}^{-1}\mathbf{B}) - \mathbf{K} + \frac{1}{2}\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}\right). \quad (2.48)$$

Da equação (2.48) pode-se notar que se λ é dado, $p(\mathbf{X}|\mathbf{Y}, \lambda)$ é uma PDF Gaussiana. Então,

$$\bar{\mathbf{X}} = E(\mathbf{X}) = -\mathbf{A}^{-1}\mathbf{B}, \quad (2.49)$$

que, associada a

$$\bar{\mathbf{X}} = \left(\lambda \mathbf{Q} + \frac{\mathbf{H}_1^T \mathbf{H}_1 + \dots + \mathbf{H}_m^T \mathbf{H}_m}{\sigma^2}\right)^{-1} \left(\frac{\mathbf{H}_1^T \mathbf{y}_1 + \dots + \mathbf{H}_m^T \mathbf{y}_m}{\sigma^2}\right), \quad (2.50)$$

corresponde a forma fechada, ou solução analítica, para a restauração da imagem HR a partir do conjunto de imagens LR observável. No contexto de inferência estatística INLA, que será explorado no próximo Capítulo, $\bar{\mathbf{X}}$ corresponde a uma solução em forma fechada para o cálculo da moda da PDF condicional de $p(\mathbf{X}|\lambda, \mathbf{Y})$.

2.5 Implementação do Algoritmo de Superresolução

Nessa seção são discutidos alguns aspectos computacionais relativos à implementação do algoritmo de Superresolução proposto em (2.50). Basicamente dois assuntos serão abordados: o método de inversão da matriz \mathbf{A} e o aproveitamento das características de esparsidade das matrizes envolvidas no problema.

O resultado apresentado em (2.50), que é uma das contribuições dessa Tese, publicada em (Camponez et al., 2011b), corresponde a uma solução de Máximo *a Posteriori* (MAP) conforme descrito, por exemplo, em (Hardie et al., 1997). A diferença é que em vez de usar (2.27) como uma função custo a ser minimizada de forma iterativa, em geral por um método numérico não linear como o Gradiente Conjugado, a solução aqui proposta desenvolve uma fórmula fechada para o cálculo da imagem HR.

Em (2.49), \mathbf{B} é uma imagem HR resultante da fusão de todas as imagens LR no *grid* de alta resolução, e \mathbf{A} é um filtro inverso regularizado, comumente encontrado na literatura de Superresolução. \mathbf{B} e parte de \mathbf{A} são computados uma única vez e reusados para diferentes

valores de λ . Como pode ser observado, o maior custo computacional, na implementação de tal algoritmo, corresponde à inversão da matriz \mathbf{A} .

Conforme (2.45), o primeiro termo da matriz \mathbf{A} , $\lambda\mathbf{Q}$, é a matriz de precisão da PDF *a priori* da imagem HR, multiplicada por um escalar. As distribuições de probabilidade podem ser representadas por diagramas chamados modelos gráficos probabilísticos (PGM), onde os nós representam as variáveis aleatórias, e os *links* expressam a relação probabilística entre essas variáveis (Bishop, 2006). Há uma classe de modelo gráfico, chamada de campo aleatório Markoviano Gaussiano (GMRF), que desperta especial interesse nesse trabalho, já que foi usada para modelar a PDF da imagem HR, conforme (2.24).

Uma GMRF é um vetor aleatório seguindo uma distribuição Gaussiana multivariável, com os pressupostos de independência condicional, daí o termo Markov. Conforme (Rue e Held, 2005), matrizes esparsas aparecem naturalmente para GMRFs, com $Q_{ij} \neq 0$ somente se i and j são vizinhos, conforme (2.25). Pela construção, a maioria das matrizes de precisão para GMRFs são esparsas, onde somente poucos termos de \mathbf{Q} são diferentes de zero.

A segunda parcela de \mathbf{A} , $\frac{\mathbf{M}}{\sigma^2}$, é uma somatória das normas quadráticas das matrizes \mathbf{H}_k que modelam os efeitos de distorção, degradação e subamostragem na aquisição da imagem LR a partir da HR. Conforme visto na seção 2.2, as matrizes \mathbf{H}_k são esparsas, o que faz com que \mathbf{M} também seja esparsa. Assim, a matriz \mathbf{A} também é esparsa, já que é formada pela soma de duas matrizes esparsas. Dessa forma, conforme (Rue e Held, 2005), o método com melhor desempenho computacional para a inversão dessa matriz é a fatoração de Cholesky.

Em álgebra linear, a fatoração de Cholesky é uma decomposição de uma matriz Hermítica definida positiva \mathbf{W} no produto de uma matriz triangular inferior e o seu conjugado transposto para o caso complexo ou, no presente caso,

$$\mathbf{W} = \mathbf{L}\mathbf{L}^T, \quad (2.51)$$

onde \mathbf{L} é uma matriz triangular inferior. Quando aplicável, essa fatoração é aproximadamente duas vezes mais eficiente que a decomposição LU para resolver sistemas de equações lineares (Press et al., 2007). No caso de GMRFs, onde as matrizes envolvidas são esparsas, principalmente em sistemas da ordem de 10.000 a 100.000 nós, como é o caso de processamento de imagens e, em especial, nos algoritmos de Superresolução, a aplicação da fatoração de Cholesky resulta em uma aceleração ainda maior, pelo fato de apenas os termos diferentes de zero em \mathbf{L} serem calculados e armazenados, com baixo uso de memória.

Uma condição suficiente e necessária para que seja possível aplicar uma fatoração de Cholesky em \mathbf{A} , uma matriz de números reais, é que ela seja positiva definida simétrica (SPD).

Definição 1 Segundo (Rue e Held, 2005), uma matriz \mathbf{W} é positiva definida se

$$\mathbf{x}^T \mathbf{W} \mathbf{x} > 0, \quad \forall \mathbf{x} \neq 0. \quad (2.52)$$

Se \mathbf{W} é simétrica, então é chamada de matriz positiva definida simétrica (SPD), que é representada pela notação $\mathbf{W} > 0$.

Algumas das propriedades de uma matriz SPD são:

1. $\text{rank}(\mathbf{W}) = n$.
2. $|\mathbf{W}| > 0$.
3. $W_{ii} > 0$.
4. $W_{ii}W_{jj} > 0$, para $i \neq j$.
5. $W_{ii} + W_{jj} - 2|W_{ij}| > 0$, para $i \neq j$.
6. $\max W_{ii} > \max_{i \neq j} |W_{ij}|$.
7. \mathbf{W}^{-1} é SPD.
8. Todas as sub-matrizes de \mathbf{W} são SPD.
9. Se \mathbf{W} e \mathbf{Z} são SPD, então $\mathbf{W} + \mathbf{Z}$ também é SPD.

As condições a seguir são todas suficientes e necessárias para que uma matriz simétrica \mathbf{W} seja SPD:

1. Todos os autovalores ϕ_1, \dots, ϕ_n de \mathbf{W} são estritamente positivos.
2. Existe uma matriz \mathbf{P} tal que $\mathbf{W} = \mathbf{P}\mathbf{P}^T$. Se \mathbf{P} é triangular inferior, então é chamada de triângulo de Cholesky de \mathbf{W} .
3. Todas as submatrizes têm determinantes estritamente positivos.

Uma condição suficiente mas não necessária para que uma matriz simétrica seja SPD é que ela seja diagonal dominante, ou seja,

$$W_{ii} - \sum_{j:j \neq i} |W_{ij}| > 0, \forall i. \quad (2.53)$$

Como pode ser observado em (2.25), \mathbf{Q} é uma matriz diagonal dominante, e, por isso, conforme (2.53), SPD.

$\mathbf{H}_k \mathbf{H}_k^T$ é uma forma quadrática, cujo determinante

$$|\mathbf{H}_k \mathbf{H}_k^T| = |\mathbf{H}_k|^2 \quad (2.54)$$

é sempre positivo. Assim, matrizes na forma quadrática são SPD. A soma de matrizes SPD também é SPD, conforme a propriedade 9 da Definição 1. Logo, \mathbf{M} é SPD.

Dessa forma, fica demonstrado que a matriz \mathbf{A} composta pela soma $(\mathbf{Q} + \mathbf{M})$, isto é, a soma de duas SPD, também é SPD, e que, portanto, uma decomposição de Cholesky pode ser aplicada.

Assim, o algoritmo de Superresolução proposto, chamado a partir desse ponto de Closed Form SR, foi desenvolvido da seguinte maneira:

1. O programa foi escrito em Matlab[®], versão 7.8.0.347.
2. As matrizes \mathbf{B} , \mathbf{Q} e \mathbf{M} são computadas uma única vez e reusadas para diferentes valores de λ .
3. O cálculo de \mathbf{A}^{-1} é feito usando-se fatoraçoão de Cholesky.
4. Todas as matrizes foram representadas na forma esparsa, utilizando-se a função *sparse* do Matlab[®].

A forma mais tradicional de avaliação da qualidade de um algoritmo de processamento de imagens é através do cálculo da relação sinal-ruído (SNR) e/ou da relação sinal-ruído de pico (PSNR) entre a imagem original e a imagem processada. O PSNR é a métrica mais amplamente utilizada de medida de qualidade objetiva de imagens e, por isso, foi adotada nessa Tese para comparar o algoritmo proposto com outros do estado da arte em Superresolução (Huynh-Thu e Ghanbari, 2008). Para imagens $m \times n$, em 256 níveis cinza, o PSNR é definido como

$$PSNR = 20 \log \left(\frac{255}{\sqrt{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [X(i, j) - X_p(i, j)]^2}} \right), \quad (2.55)$$

onde $X(i, j)$ e $X_p(i, j)$ representam, respectivamente, os pixels da imagem original e os da imagem processada.

Assim, quando comparadas várias imagens, quanto maior o valor em PSNR melhor a qualidade da imagem.

2.6 Experimentos

O objetivo dessa seção é comparar o desempenho do algoritmo de Superresolução baseado na equação em forma fechada em (2.50) com outros algoritmos para aumento de resolução. No primeiro experimento, uma sequência de quadros LR foi criada, a partir de uma imagem HR de um Barco, de tamanho 256×256 (Figura 2.7a). Para criar tais imagens LR, essa imagem do Barco foi deslocada por uma quantidade aleatória de pixels a partir de uma distribuição contínua uniforme ao longo do intervalo $(-2, 2)$, em ambas direções. A imagem foi considerada periódica nas duas direções. Então, para simular o efeito *Point Spread Function* (PSF) da câmera, uma função de aquisição, que é o valor médio de um quadrado de (2×2) pixels da imagem HR, foi aplicada para formar um pixel da imagem LR. Essa função de aquisição corresponde a aplicar um filtro passa-baixas e em seguida uma subamostragem com fator de decimação 2 nas direções horizontal e vertical. Finalmente, um ruído Gaussiano com média zero e desvio padrão 8, como adotado em (Tian e Ma, 2010), é adicionado a cada imagem processada, para se obter uma imagem de baixa resolução observável. Tal abordagem foi utilizada para gerar 16 imagens LR do mesmo Barco, a partir da cena original. Uma dessas imagens LR, de tamanho 128×128 , é apresentada na Figura 2.7b. Dessas imagens LR, quatro experimentos são executados com o objetivo de restaurar a imagem HR, cada um considerando quatro imagens LR.

O resultado usando interpolação bicúbica é mostrado na Figura 2.7c. O resultado implementando o método de Superresolução não iterativo descrito em (Vandewalle et al., 2006) é mostrado na Figura 2.7d. O resultado implementando o método proposto em (Pham et al., 2006) é mostrado na Figura 2.7e. Para gerar os últimos dois resultados foi usado o *software* "Superresolution 2.0", disponível para *download* em <http://lcav.epfl.ch/software/superresolution>.

Finalmente, o resultado mostrado na Figura 2.7f corresponde à abordagem de Superresolução em forma fechada proposta nessa Tese, com hiperparâmetro tendo sido ajustado, após várias simulações, para $\lambda = 0,001$.

Analisando as imagens da Figura 2.7 pode ser visto que o algoritmo de Superresolução aqui proposto obteve bons resultados em recuperar os detalhes que foram perdidos nos quadros LR. Isso fica claro pela melhor visibilidade dos fios que atravessam o mastro do barco, e também pela melhora na visibilidade da palavra "LACOR", no casco do barco. Na Tabela 2.1 os vários algoritmos são comparados utilizando-se PSNR. Por essa métrica, o algoritmo de Superresolução em forma fechada, aqui proposto, apresenta os melhores resultados.

No segundo experimento, imagens LR foram criadas a partir de uma imagem de Texto HR (Figura 2.8a), de tamanho 200×200 , usando o mesmo procedimento descrito na primeira

experiência. A mesma abordagem foi utilizada para gerar 16 imagens de Texto LR, a partir da cena original. Um desses quadros LR, de tamanho 100×100 , é apresentado na Figura 2.8b. Dessas imagens LR, quatro experimentos são executados, com o objetivo de restaurar a imagem HR Texto, cada um considerando quatro imagens LR.

O resultado usando interpolação bicúbica é mostrado na Figura 2.8c. O resultado implementando o método de Superresolução não iterativo descrito em (Vandewalle et al., 2006) é mostrado na Figura 2.8d. O resultado implementando o método proposto em (Pham et al., 2006) é mostrado na Figura 2.8e. Para gerar esses dois últimos resultados foi utilizado o mesmo *software* indicado no primeiro experimento.

O resultado mostrado na Figura 2.8f corresponde à abordagem de Superresolução em forma fechada proposta nessa Tese, com hiperparâmetro tendo sido ajustado, após várias simulações, para $\lambda = 0,0004$.

Analisando a Figura 2.8 pode ser visto que a imagem gerada pelo algoritmo de Superresolução em forma fechada (Figura 2.8f), aqui proposto, apresenta melhores resultados que as outras, pois, além de ser a mais nítida, onde os componentes de alta frequência foram melhor recuperados, não tem artefatos e nem erros de bordas. A imagem da Figura 2.8c é muito ruidosa. A imagem da Figura 2.8d, por exemplo, é muito mais suavizada, e apresenta efeitos de borda. A imagem na Figura 2.8e é muito suavizada e também apresenta efeitos de borda. Na Tabela 2.1 os vários algoritmos são comparados utilizando-se como medida objetiva o PSNR. Por essa métrica, mais uma vez, o algoritmo proposto apresenta resultados que superam os outros algoritmos.

Em (Tian e Ma, 2010) um algoritmo para recuperação de imagens por Superresolution usando MCMC foi proposto. Nos experimentos com imagens simuladas, os autores utilizaram as mesmas imagens HR apresentados nas Figuras 2.7a e 2.8a. O mesmo procedimento para gerar os quadros LR também foi utilizado. Assim como nessa seção, imagens HR foram recuperadas a partir de um conjunto de quatro imagens LR. Em seguida, a métrica de desempenho PSNR foi utilizada para comparar a abordagem MCMC com outros algoritmos. O método MCMC SR apresentou PSNR de 28,02 e 20,17, respectivamente, para as imagens do Barco e do Texto. Nos dois casos, a Tabela 2.1 sugere que o desempenho do algoritmo aqui proposto é superior. Nós não tivemos acesso às imagens LR utilizados em tal artigo, mas em nossas simulações, mesmo com diferentes conjuntos de quadros LR os valores de PSNR na Tabela 2.1 não se alteraram significativamente. Portanto, o resultado dessa comparação é confiável.

Tabela 2.1: Avaliação das imagens HR reconstruídas considerando PSNR (dB)

Imagem		Bi-	Vandewalle	Pham	MCMC	Closed Form
Teste	Exp.	cubic	et al.	et al.		SR
		spline				(novo)
Boat	1	20,29	24,88	27,20		30,18
	2	22,77	24,73	26,98		30,13
	3	25,05	24,73	27,15	28,02	30,21
	4	20,77	24,97	26,79		30,22
Média		22,22	24,83	27,03	28,02	30,19
Text	1	13,15	16,61	17,93		22,08
	2	14,52	16,41	17,51		22,05
	3	16,43	16,32	17,57	20,17	21,89
	4	12,81	16,34	17,51		22,16
Média		14,23	16,42	17,63	20,17	22,05

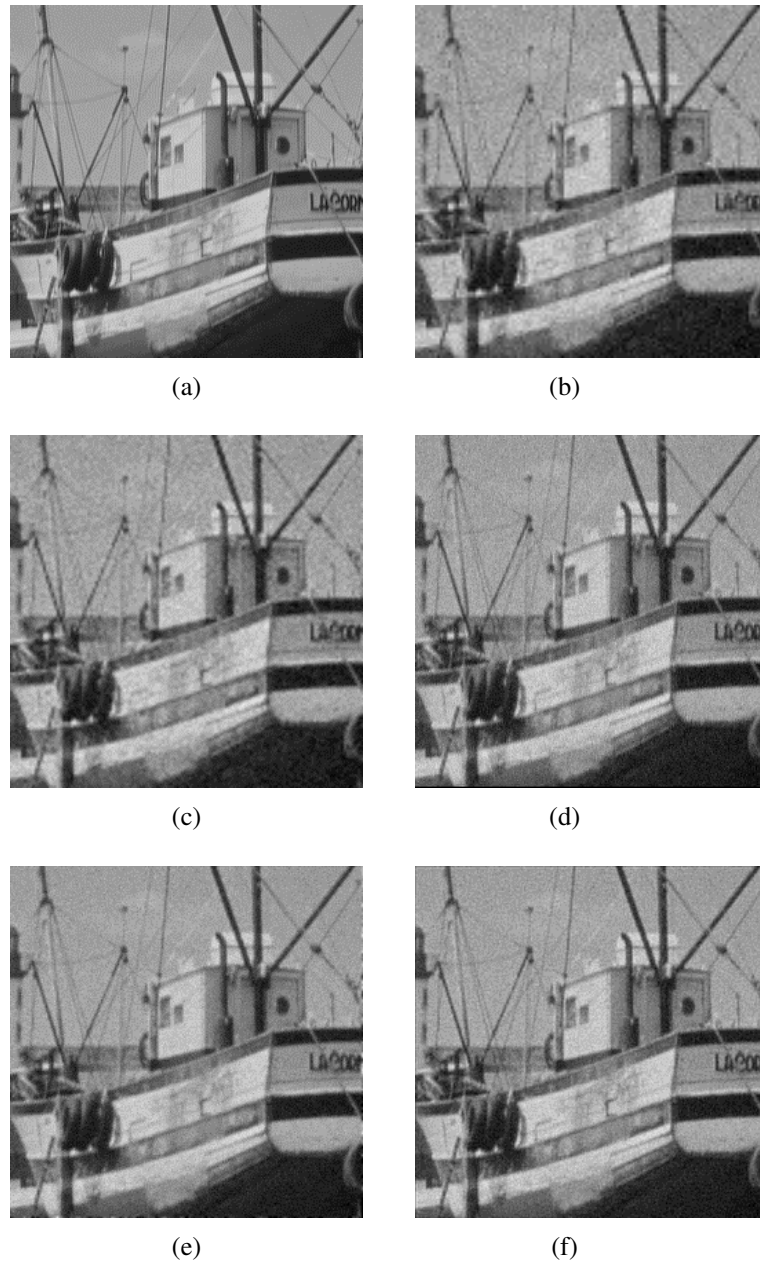


Figura 2.7: O conjunto de imagens HR reconstruídas usando a imagem de teste Barco. (a) A imagem original (*ground truth*), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) Interpolação, (d) em (Vandewalle et al., 2006) (e) em (Pham et al., 2006) (f) Closed Form SR proposto aqui.

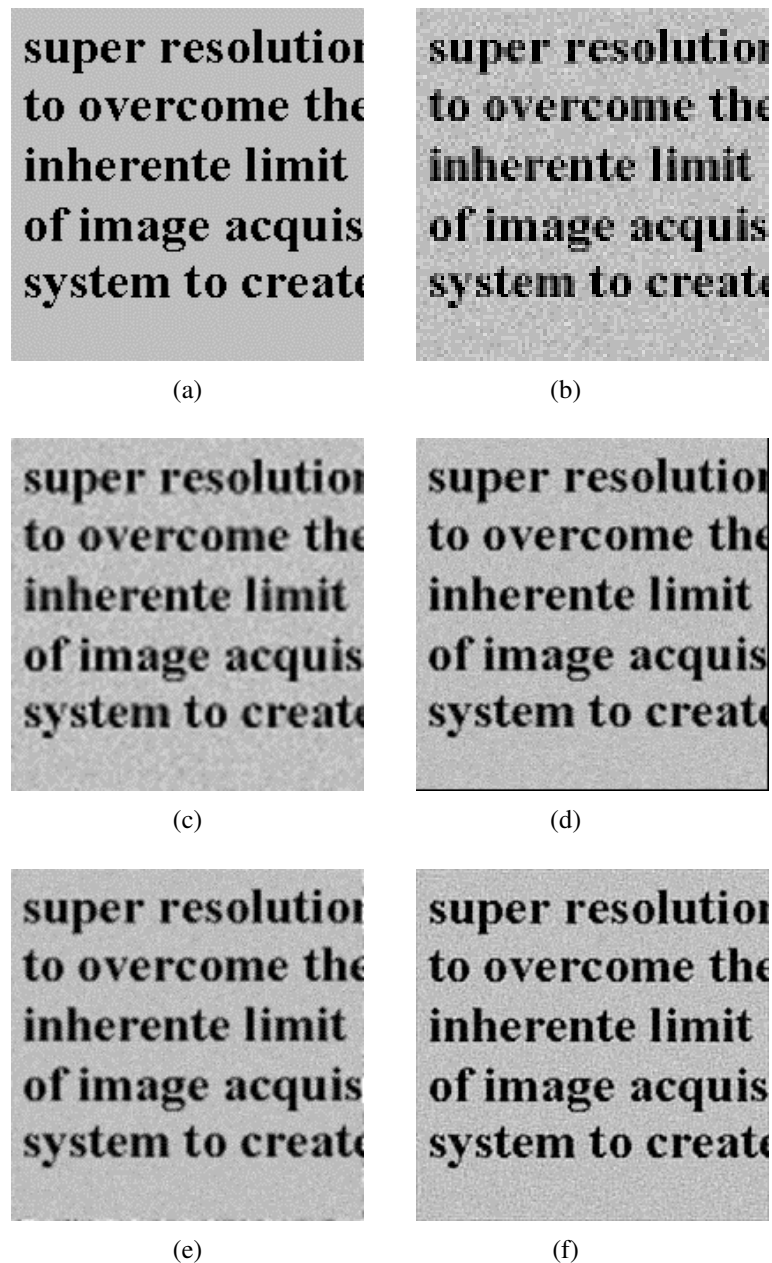


Figura 2.8: O conjunto de imagens HR reconstruídas usando a imagem de teste Texto. (a) A imagem original (*ground truth*), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) Interpolação, (d) em (Vandewalle et al., 2006) (e) em (Pham et al., 2006) (f) Closed Form SR proposto aqui.

Capítulo 3

Superresolução Usando Integrated Nested Laplace Approximation

No capítulo anterior um novo algoritmo para recuperação de imagens HR por Superresolução, chamado de Closed Form SR, foi proposto. Entretanto, o desempenho desse novo método depende da escolha de um hiperparâmetro. Dessa forma, a geração de imagens HR de boa qualidade depende da habilidade do usuário em escolher um valor adequado para λ . Nesse capítulo, é desenvolvida uma forma automática para a escolha de λ . Para isso, um método de inferência Bayesiana conhecido como *Integrated Nested Laplace Approximation*-INLA, recentemente desenvolvido, e, até onde o autor sabe, inédito na área de processamento de imagens, é aplicado ao problema de Superresolução e explorado, com sucesso, na implementação de um algoritmo totalmente automático para recuperação de imagens por Superresolução.

3.1 Integrated Nested Laplace Approximation

Havard Rue, Sara Martino and Nicolas Chopin (Rue et al., 2009) desenvolveram um método aproximado de inferência Bayesiana em modelos Gaussianos latentes, conhecido como *Integrated Nested Laplace Approximation* (INLA).

Modelos Gaussianos latentes são amplamente utilizados na análise Bayesiana. Tais modelos assumem um campo Gaussiano latente $\mathbf{X} = (x_1, \dots, x_p)^T$, que é pontualmente observado através de m dados condicionais independentes \mathbf{Y} . A matriz de covariância do campo Gaussiano latente e sua função de verosimilhança são governadas por alguns poucos parâmetros $\lambda = (\lambda_1, \dots, \lambda_\alpha)^T$, usualmente $\alpha \leq 6$ (Rue et al., 2009). A PDF marginal posterior

de \mathbf{X} e λ condicionada a \mathbf{Y} , então, é

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto p(\lambda) p(\mathbf{X} | \lambda) \prod_{l \in \tau} p(y_l | x_l, \lambda), \quad (3.1)$$

onde τ representa o domínio da variável l .

Muitos, embora nem todos, os modelos Gaussianos latentes na literatura satisfazem duas propriedades básicas, que também são assumidas no método INLA. A primeira propriedade é que o campo latente \mathbf{X} , que é muitas vezes de grande dimensão ($p = 10^2$ até 10^5) admite propriedades de independência condicional. Assim, o campo latente é um campo aleatório Markoviano Gaussiano (GMRF) com uma matriz de precisão \mathbf{Q} esparsa. A segunda propriedade é que o número α de hiperparâmetros λ é pequeno.

Uma abordagem comum para inferência em modelos Gaussianos latentes é o uso de aproximações probabilísticas, através do uso de Métodos de Monte Carlo via Cadeias de Markov (MCMC) (MacKay, 1999). Outra abordagem é o uso de aproximações determinísticas, através do uso de inferência Variacional Bayesiana (Bishop, 2006). É bem conhecido, no entanto, que os métodos MCMC tendem a apresentar desempenho ruim quando aplicados a estes modelos (Rue et al., 2009).

O INLA é uma alternativa ao uso de métodos MCMC e de inferência Variacional. Em contraste com MCMC, o método INLA não gera amostras a partir da PDF posterior. Ele aproxima a PDF posterior de uma forma não-paramétrica. Usando INLA é possível calcular diretamente aproximações muito precisas para as marginais posteriores. A principal vantagem dessas aproximações é uma redução significativa do esforço computacional: enquanto os algoritmos MCMC precisam de horas, até dias, para convergirem para um resultado, a aproximação INLA fornece estimativas mais precisas em questão de minutos, ou mesmo segundos.

A abordagem INLA usa o fato de que as PDFs marginais posteriores de interesse podem ser escritas como

$$p(x_l | \mathbf{Y}) = \int p(x_l | \lambda, \mathbf{Y}) p(\lambda | \mathbf{Y}) d\lambda, \quad (3.2)$$

$$p(\lambda_j | \mathbf{Y}) = \int p(\lambda | \mathbf{Y}) d\lambda_{-j}, \quad (3.3)$$

e usa essas formas para construir aproximações aninhadas

$$\tilde{p}(x_l | \mathbf{Y}) = \int \tilde{p}(x_l | \lambda, \mathbf{Y}) \tilde{p}(\lambda | \mathbf{Y}) d\lambda, \quad (3.4)$$

$$\tilde{p}(\lambda_j | \mathbf{Y}) = \int \tilde{p}(\lambda | \mathbf{Y}) d\lambda_{-j}. \quad (3.5)$$

Aqui, $\tilde{p}(\cdot | \cdot)$ é uma função de densidade (condicional) de seus argumentos. A aproximação $\tilde{p}(\lambda_j | \mathbf{Y})$ é computada integrando-se $\tilde{p}(\lambda | \mathbf{Y})$ em λ_{-j} , onde λ_{-j} são todos os λ dentro do

intervalo $[\lambda_{min}, \lambda_{max}]$, exceto λ_j . Essa integração é possível pois a dimensão de λ é pequena. Aproximações para $p(x_l|\mathbf{Y})$ são computadas pela aproximação de $p(\lambda|\mathbf{Y})$ e $p(x_l|\lambda, \mathbf{Y})$, e usando integração numérica, isto é, soma finita para integrar em λ ,

$$\tilde{p}(x_l|\mathbf{Y}) = \sum_z \tilde{p}(x_l|\lambda_z, \mathbf{Y}) \tilde{p}(\lambda_z|\mathbf{Y}) \Delta\lambda_z. \quad (3.6)$$

Um aspecto crítico dessa abordagem é explorar e manipular $\tilde{p}(\lambda|\mathbf{Y})$ e $\tilde{p}(x_l|\mathbf{Y})$ de uma maneira não paramétrica.

A aproximação usada em (Rue et al., 2009) para a marginal posterior de λ , $p(\lambda|\mathbf{Y})$, é

$$\tilde{p}(\lambda|\mathbf{Y}) \propto \left. \frac{p(\mathbf{X}, \lambda, \mathbf{Y})}{\tilde{p}_G(\mathbf{X}|\lambda, \mathbf{Y})} \right|_{\mathbf{X}=\bar{\mathbf{X}}}, \quad (3.7)$$

onde $\tilde{p}_G(\mathbf{X}|\lambda, \mathbf{Y})$ é uma aproximação Gaussiana da PDF da marginal condicional \mathbf{X} , e $\mathbf{X} = \bar{\mathbf{X}}$ é a moda da PDF da marginal condicional \mathbf{X} (que corresponde a uma estimativa de MAP de \mathbf{X} (Hardie et al., 1997)), para um dado λ . Quando $\tilde{p}(\mathbf{X}|\lambda, \mathbf{Y})$ é Gaussiana, como, por exemplo, no modelo usado para Superresolução no Capítulo 2, a moda, a mediana e o valor médio são iguais. Assim, nesse caso, $\bar{\mathbf{X}}$ corresponde ao valor médio de \mathbf{X} .

Para a aproximação de $\tilde{p}(x_l|\lambda, \mathbf{Y})$ há três possibilidades:

1. aproximação Gaussiana - é computacionalmente menos custosa e, ainda, apresenta bons resultados. Além disso, essa aproximação já foi calculada no denominador de (3.7);
2. aproximação de Laplace - apresenta resultados melhores que a primeira, porém com custo computacional alto;
3. aproximação de Laplace simplificada - apresenta bons resultados e não é tão custosa como a de Laplace.

Nesta Tese a aproximação usada foi a Gaussiana. Na verdade, como será demonstrado a seguir, uma equação em forma fechada foi derivada para o cálculo de $\tilde{p}(x_l|\lambda, \mathbf{Y})$.

O algoritmo INLA, que implementa a primeira opção, pode ser resumido como:

1. selecione um conjunto de $\lambda = (\lambda_1, \dots, \lambda_\alpha)$;
2. para $z = 1$ até α faça:
3. calcule $\tilde{p}(\lambda_z|\mathbf{Y})$ como em (3.7);

4. calcule $\tilde{p}(x_l|\lambda_z, \mathbf{Y})$ como uma função de x_l ;
5. fim do laço de repetição;
6. calcule $\tilde{p}(x_l|\mathbf{Y}) = \sum_z \tilde{p}(x_l|\lambda_z, \mathbf{Y}) \tilde{p}(\lambda_z|\mathbf{Y}) \Delta\lambda_z$;
7. re-escala $\tilde{p}(x_l|\mathbf{Y})$.

O algoritmo começa pela seleção de um pequeno conjunto de valores de λ , baseado na distribuição $\tilde{p}(\lambda|\mathbf{Y})$. Primeiro, a moda de (3.7) é encontrada, e, então, alguns pontos igualmente espaçados são espalhados sobre o domínio. Para cada ponto λ em tal *grid* dois vetores são calculados, a saber $\tilde{p}(\lambda_z|\mathbf{Y})$ e $\tilde{p}(x_l|\lambda_z, \mathbf{Y})$. Então, uma integração numérica $\sum_k \tilde{p}(x_l|\lambda_z, \mathbf{Y}) \tilde{p}(\lambda_z|\mathbf{Y}) \Delta\lambda_z$ é calculada, de uma maneira não paramétrica.

Por fim, $\tilde{p}(x_l|\mathbf{Y})$ é reescalado, dividindo-se cada elemento de tal vetor pela soma de todos os elementos, de forma que a somatória das probabilidades se torne igual a um.

3.2 Aplicando o INLA ao Problema da Superresolução

Nessa seção a técnica de inferência estatística Bayesiana INLA é adaptada para ser aplicada ao problema de recuperação de imagens por Superresolução. Todas as equações necessárias para a fusão e filtragem da imagem HR, a partir das imagens LR observáveis, serão derivadas. A fim de tornar mais fácil o entendimento dessa adaptação, as próximas subseções seguirão o mais próximo possível a ordem adotada no desenvolvimento do algoritmo INLA da última seção.

O principal objetivo do método de inferência INLA é computar todas as, ou algumas das, p posteriores marginais para x_l mais, possivelmente, a marginal posterior para λ ou algum λ_j .

Diferentemente, no problema Superresolução o principal objetivo é encontrar a imagem HR. Isto pode ser realizado através do cálculo do valor esperado através do pós-processamento da marginal $p(\mathbf{X}|\mathbf{Y})$.

3.2.1 Uma Equação em Forma Fechada para $\bar{\mathbf{X}}$

Em (3.7) $\tilde{p}(\lambda|\mathbf{Y})$ é avaliada na moda da PDF da marginal condicional de \mathbf{X} , isto é, $\mathbf{X} = \bar{\mathbf{X}}$, para um dado λ .

Em (Camponez et al., 2011b) mostra-se que (2.28) pode ser escrita, conforme demonstrado no Capítulo 2, como

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(-\frac{1}{2} (\mathbf{X} + \mathbf{A}^{-1} \mathbf{B})^T \mathbf{A} (\mathbf{X} + \mathbf{A}^{-1} \mathbf{B}) + \mathbf{V} \right), \quad (3.8)$$

onde

$$\mathbf{V} = -\mathbf{K} + \frac{1}{2} \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}, \quad (3.9)$$

$$\mathbf{A} = \left(\lambda \mathbf{Q} + \sum_{k=1}^m \frac{\mathbf{H}_k^T \mathbf{H}_k}{\sigma^2} \right), \quad (3.10)$$

$$\mathbf{B} = -\sum_{k=1}^m \frac{H_k^T \mathbf{y}_k}{\sigma^2}, \quad (3.11)$$

e

$$\mathbf{K} = \sum_{k=1}^m \frac{\mathbf{y}_k^T \mathbf{y}_k}{2\sigma^2}. \quad (3.12)$$

De (3.8) pode-se notar que se λ é dado, $p(\mathbf{X} | \mathbf{Y}, \lambda)$ é uma função Gaussiana. Então,

$$\bar{\mathbf{X}} = E(\mathbf{X}) = -\mathbf{A}^{-1} \mathbf{B}, \quad (3.13)$$

que, associada a

$$\bar{\mathbf{X}} = \left(\lambda \mathbf{Q} + \frac{\mathbf{H}_1^T \mathbf{H}_1 + \dots + \mathbf{H}_m^T \mathbf{H}_m}{\sigma^2} \right)^{-1} \left(\frac{\mathbf{H}_1^T \mathbf{y}_1 + \dots + \mathbf{H}_m^T \mathbf{y}_m}{\sigma^2} \right), \quad (3.14)$$

corresponde a uma forma fechada para o cálculo da imagem HR. No contexto do INLA, $\bar{\mathbf{X}}$ corresponde a uma forma fechada para o cálculo da moda da PDF da marginal condicional de \mathbf{X} .

3.2.2 Selecionando um Conjunto de Valores para λ

O próximo passo no algoritmo INLA é escolher um conjunto de valores para λ , para que se possa calcular a aproximação em (3.7) e, *a posteriori*, a integral em (3.6).

No Capítulo 2, a PDF do hiperparâmetro foi definida como uma distribuição uniforme. Então, definir um *grid* é equivalente a escolher os valores máximo e mínimo para a distribuição uniforme e, além disso, escolher o passo para formação do *grid*. Na definição de tais limites, o efeito de λ na formação da imagem HR através de Superresolução foi estudado.

A equação (3.14), que pode ser usada para gerar realizações de imagens HR, tem dois termos. O primeiro atua como um filtro passa-baixas: valores elevados de λ suavizam a

imagem. Em oposição, pequenos valores de λ fazem com que a imagem fique ruidosa. Por sua vez, o segundo termo atua como uma regra para interpolar as imagens de baixa resolução. Várias simulações foram realizadas, mostrando que os valores de λ maiores do que 0,01 suavizam demais as imagens, enquanto valores abaixo de 0,0005 tornam a imagem demasiadamente ruidosa.

Assim, na implementação do algoritmo a faixa de valores de λ

$$0,0005 < \lambda < 0,01 \quad (3.15)$$

foi adotada, usando 0,0005 como o passo para a formação do *grid*.

3.2.3 Derivando uma Forma Fechada para $\tilde{p}_G(\mathbf{X}|\lambda, \mathbf{Y})$

De (3.8) pode-se notar que se λ é dado, $p(\mathbf{X}|\lambda, \mathbf{Y})$ é Gaussiano. Além do mais, o valor de \mathbf{X} aplicado à aproximação (3.7), onde $\mathbf{X} = \bar{\mathbf{X}}$, é a moda da PDF marginal condicional de \mathbf{X} , para um dado λ . Então, usando (3.13) e (3.9), pode-se reescrever (3.8) como

$$\tilde{p}_G(\mathbf{X}|\lambda, \mathbf{Y}) = p(\bar{\mathbf{X}}|\lambda, \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp\left(-\mathbf{K} - \frac{1}{2} \mathbf{B}^T \bar{\mathbf{X}}\right). \quad (3.16)$$

3.2.4 Uma Forma Fechada para $\tilde{p}(\lambda|\mathbf{Y})$

O método INLA usa a aproximação $\tilde{p}(\lambda|\mathbf{Y})$ em (3.7) para a marginal posterior de λ . Quando o INLA é aplicado ao problema de Superresolução, ao invés de se usar tal abordagem uma equação em forma fechada pode ser derivada como

$$p(\lambda|\mathbf{Y}) \propto |\mathbf{A}|^{-\frac{1}{2}} |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp\left(-\mathbf{K} - \frac{1}{2} \mathbf{B}^T \bar{\mathbf{X}}\right). \quad (3.17)$$

A função de densidade de probabilidade $p(\lambda|\mathbf{Y})$ depende do nível de ruído em \mathbf{Y} e de uma estimativa da imagem \mathbf{X} . Embora essa PDF possa fazer um *link* com técnicas automáticas de estimação do parâmetro de regularização, como aquela em (Galatsanos e Katsaggelos, 1992), no INLA o λ não precisa ser estimado, porque ele é removido do problema por uma integração numérica, como mostrado na subseção seguinte.

O ponto inicial para derivar (3.17) é o termo

$$\|\mathbf{y}_k - \mathbf{H}_k \mathbf{X}\|^2 = \mathbf{y}_k^T \mathbf{y}_k - 2\mathbf{y}_k^T \mathbf{H}_k \mathbf{X} + \mathbf{X}^T \mathbf{H}_k^T \mathbf{H}_k \mathbf{X} \quad (3.18)$$

da equação (2.27). Considerando as variáveis

$$\mathbf{m}_k = \mathbf{H}_k^T \mathbf{H}_k, \quad (3.19)$$

$$\mathbf{u}_k = \mathbf{H}_k^T \mathbf{y}_k, \quad (3.20)$$

e

$$\mathbf{c}_k = \mathbf{y}_k^T \mathbf{y}_k, \quad (3.21)$$

obtém-se

$$\|\mathbf{y}_k - H_k \mathbf{X}\|^2 = \mathbf{X}^T \mathbf{m}_k \mathbf{X} - 2\mathbf{u}_k^T \mathbf{X} + \mathbf{c}_k. \quad (3.22)$$

De tais equações,

$$\sum_{k=1}^m \|\mathbf{y}_k - H_k \mathbf{X}\|^2 = \sum_{k=1}^m \mathbf{X}^T \mathbf{m}_k \mathbf{X} - \sum_{k=1}^m 2\mathbf{u}_k^T \mathbf{X} + \sum_{k=1}^m \mathbf{c}_k, \quad (3.23)$$

cujos termos são desenvolvidos como

$$\sum_{k=1}^m \mathbf{X}^T \mathbf{m}_k \mathbf{X} = \mathbf{X}^T (\mathbf{m}_1 + \dots + \mathbf{m}_m) \mathbf{X} = \mathbf{X}^T \mathbf{M} \mathbf{X}, \quad (3.24)$$

$$\mathbf{M} = \sum_{k=1}^m \mathbf{m}_k, \quad (3.25)$$

$$- \sum_{k=1}^m 2\mathbf{u}_k^T \mathbf{X} = -2(\mathbf{u}_1 + \dots + \mathbf{u}_m)^T \mathbf{X} = -2\mathbf{U}^T \mathbf{X}, \quad (3.26)$$

$$\mathbf{U} = \sum_{k=1}^m \mathbf{u}_k, \quad (3.27)$$

e

$$\mathbf{C} = \sum_{k=1}^m \mathbf{c}_k. \quad (3.28)$$

Agora, introduzindo (3.25), (3.27) e (3.28) em (3.23), e supondo que todas as imagens LR têm a mesma variância, obtém-se

$$\sum_{k=1}^m \frac{1}{2\sigma^2} \|\mathbf{y}_k - H_k \mathbf{X}\|^2 = \frac{1}{2\sigma^2} (\mathbf{X}^T \mathbf{M} \mathbf{X} - 2\mathbf{U}^T \mathbf{X} + \mathbf{C}), \quad (3.29)$$

e, introduzindo (3.29) em (2.27),

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(- \left(\frac{1}{2} \mathbf{X}^T (\lambda \mathbf{Q} + \frac{\mathbf{M}}{\sigma^2}) \mathbf{X} - 2 \frac{\mathbf{U}^T}{2\sigma^2} \mathbf{X} + \frac{\mathbf{C}}{2\sigma^2} \right) \right). \quad (3.30)$$

Substituindo as variáveis de forma que

$$\mathbf{E}^T = \frac{\mathbf{U}^T}{\sigma^2}, \quad (3.31)$$

$$\mathbf{W} = \frac{\mathbf{C}}{\sigma^2}, \quad (3.32)$$

e

$$\mathbf{A} = \left(\lambda \mathbf{Q} + \frac{\mathbf{M}}{\sigma^2} \right), \quad (3.33)$$

obtem-se

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto |\lambda \mathbf{Q}|^{\frac{1}{2}} \exp \left(-\frac{1}{2} (\mathbf{X}^T \mathbf{A} \mathbf{X} - 2 \mathbf{E}^T \mathbf{X} + \mathbf{W}) \right). \quad (3.34)$$

Manipulando o argumento de (3.34),

$$-\frac{1}{2} (\mathbf{X}^T \mathbf{A} \mathbf{X} - 2 \mathbf{E}^T \mathbf{A}^{-1} \mathbf{A} \mathbf{X} + \mathbf{W}), \quad (3.35)$$

e introduzindo (3.13), (3.20), (3.27) e (3.31), obtém-se

$$-\frac{1}{2} (\mathbf{X}^T \mathbf{A} \mathbf{X} - 2 \bar{\mathbf{X}}^T \mathbf{A} \mathbf{X} + \mathbf{W}). \quad (3.36)$$

Então, depois de mais algumas manipulações algébricas,

$$-\frac{1}{2} (\mathbf{X}^T \mathbf{A} \mathbf{X} - 2 \bar{\mathbf{X}}^T \mathbf{A} \mathbf{X} + \mathbf{W} + \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}} - \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}}), \quad (3.37)$$

$$-\frac{1}{2} (\mathbf{W} - \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}}) - \frac{1}{2} (\mathbf{X}^T \mathbf{A} \mathbf{X} - 2 \bar{\mathbf{X}}^T \mathbf{A} \mathbf{X} + \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}}), \quad (3.38)$$

$$-\frac{1}{2} (\mathbf{W} - 2 \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}} + \bar{\mathbf{X}}^T \mathbf{A} \bar{\mathbf{X}}) - \frac{1}{2} ((\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A} \mathbf{X} - \bar{\mathbf{X}}^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}})), \quad (3.39)$$

e substituindo (3.13), (3.20), (3.27), (3.31) e (3.33) em (3.39), obtém-se

$$-\frac{1}{2} \left(\mathbf{W} - 2 \bar{\mathbf{X}}^T \mathbf{A} (\mathbf{A}^{-1} \frac{\mathbf{U}}{\sigma^2}) + \bar{\mathbf{X}}^T (\lambda \mathbf{Q} + \frac{\mathbf{M}}{\sigma^2}) \bar{\mathbf{X}} \right) - \frac{1}{2} \left(\mathbf{X}^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}) - \bar{\mathbf{X}}^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}) \right), \quad (3.40)$$

e

$$-\frac{1}{2} \left(\mathbf{W} - 2 \bar{\mathbf{X}}^T \frac{\mathbf{U}}{\sigma^2} + \bar{\mathbf{X}}^T \lambda \mathbf{Q} \bar{\mathbf{X}} + \bar{\mathbf{X}}^T \frac{\mathbf{M}}{\sigma^2} \bar{\mathbf{X}} \right) - \frac{1}{2} \left((\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}) \right). \quad (3.41)$$

Introduzindo (3.32), obtém-se

$$-\frac{1}{2} \bar{\mathbf{X}}^T \lambda \mathbf{Q} \bar{\mathbf{X}} - \frac{1}{2 \sigma^2} (\bar{\mathbf{X}}^T \mathbf{M} \bar{\mathbf{X}} - 2 \mathbf{U}^T \bar{\mathbf{X}} + \mathbf{C}) - \frac{1}{2} (\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}), \quad (3.42)$$

e agora, introduzindo (3.29),

$$-\frac{1}{2} \bar{\mathbf{X}}^T \lambda \mathbf{Q} \bar{\mathbf{X}} - \sum_{k=1}^m \frac{1}{2 \sigma^2} \|y_k - H_k \bar{\mathbf{X}}\|^2 - \frac{1}{2} (\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}). \quad (3.43)$$

Reintroduzindo (3.43) em (3.34), e usando (2.27), obtém-se

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto p(\bar{\mathbf{X}}, \lambda | \mathbf{Y}) \exp \left(-\frac{1}{2} (\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A} (\mathbf{X} - \bar{\mathbf{X}}) \right), \quad (3.44)$$

onde $p(\bar{\mathbf{X}}, \lambda | \mathbf{Y})$ é dado por (3.16).

Finalmente, a marginal $p(\lambda|\mathbf{Y})$ pode ser encontrada pela integração de (3.44) em \mathbf{X} , resultando em

$$p(\lambda|\mathbf{Y}) = \int p(\mathbf{X}, \lambda|\mathbf{Y}) d\mathbf{X}, \quad (3.45)$$

de modo que

$$p(\lambda|\mathbf{Y}) \propto p(\bar{\mathbf{X}}, \lambda|\mathbf{Y}) \int \exp\left(-\frac{1}{2}(\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{A}(\mathbf{X} - \bar{\mathbf{X}})\right) d\mathbf{X}. \quad (3.46)$$

A integral sobre \mathbf{X} pode ser avaliada simplesmente observando-se um resultado padrão para o coeficiente de normalização de uma Gaussiana multivariada (Johnson e Wichern, 2002), (Bishop, 2006), resultando em

$$p(\lambda|\mathbf{Y}) \propto (2\pi)^{\frac{N}{2}} |\mathbf{A}|^{-\frac{1}{2}} p(\bar{\mathbf{X}}, \lambda|\mathbf{Y}). \quad (3.47)$$

Finalmente, substituindo (3.16) em (3.47), a forma fechada

$$p(\lambda|\mathbf{Y}) \propto |\mathbf{A}|^{-\frac{1}{2}} |\lambda\mathbf{Q}|^{\frac{1}{2}} \exp\left(-\mathbf{K} - \frac{1}{2}\mathbf{B}^T \bar{\mathbf{X}}\right) \quad (3.48)$$

é derivada para $p(\lambda|\mathbf{Y})$.

3.2.5 Derivando uma Forma Fechada para $\tilde{p}(\mathbf{X}|\mathbf{Y})$

Nessa seção, da equação (3.4) uma forma fechada para $p(\mathbf{X}|\mathbf{Y})$, que corresponde à aproximação de $\tilde{p}(\mathbf{X}|\mathbf{Y})$ no método INLA, é derivada. Note-se que

$$p(\mathbf{X}|\mathbf{Y}) = \int p(\mathbf{X}|\lambda, \mathbf{Y}) p(\lambda|\mathbf{Y}) d\lambda. \quad (3.49)$$

Então, essa equação pode ser aproximada por

$$p(\mathbf{X}|\mathbf{Y}) = \sum_{\lambda \in \text{grid}} p(\mathbf{X}|\lambda, \mathbf{Y}) p(\lambda|\mathbf{Y}) \Delta\lambda. \quad (3.50)$$

No método INLA, $p(\mathbf{X}|\mathbf{Y})$ é avaliada somente nos valores da moda de \mathbf{X} , isto é, $\mathbf{X} = \bar{\mathbf{X}}$. Assim, substituindo (3.48) e (3.16) o resultado torna-se

$$p(\mathbf{X}|\mathbf{Y}) \propto \sum |\mathbf{A}|^{-\frac{1}{2}} |\lambda\mathbf{Q}| \exp(-2\mathbf{K} - \mathbf{B}^T \bar{\mathbf{X}}) \Delta\lambda. \quad (3.51)$$

A função de densidade de probabilidade $\tilde{p}(\mathbf{X}|\mathbf{Y})$ somente é avaliada nos pontos do domínio que fazem parte do conjunto de pontos λ selecionados, que correspondem aos pontos de maior probabilidade. Dessa forma, como λ é um escalar, a soma em (3.51), avaliada em cada ponto em questão, é o produto de $p(\mathbf{X}|\lambda, \mathbf{Y})$, $p(\lambda, \mathbf{Y})$ e $\Delta\lambda$ em cada ponto. Além disso, como o *grid* escolhido é uniforme e $p(\mathbf{X}|\mathbf{Y})$ não depende de λ , essa expressão pode ser simplificada como

$$p(\mathbf{X}|\mathbf{Y}) \propto \exp(-2\mathbf{K} - \mathbf{B}^T \bar{\mathbf{X}}), \quad (3.52)$$

que corresponde a uma equação em forma fechada para se obter $p(\mathbf{X}|\mathbf{Y})$.

3.2.6 Gerando a Imagem de Alta Resolução através do Valor Esperado

Através de (3.52), a PDF $p(\mathbf{X}|\mathbf{Y})$ pode ser determinada, e, subsequentemente, pós-processada a fim de se encontrar o valor esperado de \mathbf{X} , que corresponde a imagem HR.

Então, o algoritmo para reconstrução de imagens por Superresolução usando o método não paramétrico *Integrated Nested Laplace Approximation*, aqui denominado de INLA SR, que é uma das contribuições dessa Tese, cujos resultados foram publicados em (Camponez et al., 2011a; Camponez et al., 2012c), pode ser resumido como:

1. selecione um conjunto de $\lambda = (0,0005 \ 0,0010 \ \dots \ 0,01)$;
2. para $z = 1$ até 20 faça;
3. calcule o vetor $\mathbf{T}(z) = \bar{\mathbf{X}}(\lambda(z))$ usando (3.14);
4. calcule o vetor $\mathbf{P}(z) = p(\mathbf{X}(z)|\mathbf{Y})$ usando (3.52);
5. fim do laço de repetição;
6. re-escale $\mathbf{P}(z)$, pela divisão de cada elemento do vetor pela soma de todos os elementos;
7. calcule a imagem HR como $\bar{\mathbf{X}} = \sum_{z=1}^{20} \mathbf{T}(z)\mathbf{P}(z)$.

Deve-se notar que no algoritmo acima \mathbf{T} é um vetor contendo 20 realizações da imagem HR, e \mathbf{P} é um vetor contendo a probabilidade associada a cada uma dessas realizações.

3.2.7 Gerando a Imagem de Alta Resolução através de MAP

Alternativamente, a imagem de alta resolução pode ser encontrada de uma outra maneira. Através de (3.52), a PDF $p(\mathbf{X}|\mathbf{Y})$ pode ser determinada. Usando-se (3.14) as realizações das imagens de alta resolução, que correspondem aos elementos do *grid* de λ , podem ser calculadas. Posteriormente, dentro do conjunto de imagens de alta resolução realizadas, a que corresponder a maior probabilidade é escolhida como solução do problema. Esse algoritmo, chamado de INLA MAP (*Máximo a Posteriori*), que é outra contribuição dessa Tese publicada em (Camponez et al., 2012c; Camponez et al., 2012a), pode ser sumariado como:

1. selecione um conjunto de $\lambda = (0,0005 \ 0,0010 \ \dots \ 0,01)$;

2. para $z = 1$ até 20 faça;
3. calcule o vetor $\mathbf{T}(z) = \bar{\mathbf{X}}(\lambda(z))$ usando (3.14);
4. calcule o vetor $\mathbf{P}(z) = p(\mathbf{X}(z)|\mathbf{Y})$ usando (3.52);
5. fim do laço de repetição;
6. reescale $\mathbf{P}(z)$, pela divisão de cada elemento do vetor pela soma de todos os elementos;
7. encontre o valor máximo de $\mathbf{P}(z)$. A imagem em $\mathbf{T}(z)$ correspondente a esse ponto é a imagem HR.

Note que o método INLA MAP corresponde a uma versão automática do algoritmo de Superresolução em forma fechada, Closed Form SR, proposto nessa Tese, apresentado no Capítulo 2, publicado em (Camponez et al., 2011b).

3.3 Experimentos

A fim de demonstrar o desempenho do algoritmo de SR proposto neste documento, alguns dos resultados experimentais são aqui apresentados. O primeiro conjunto de experimentos envolve dados simulados (imagens LR) derivados de duas imagens HR. Esses dados permitem uma análise quantitativa de erro, usando o índice de desempenho PSNR, frequentemente utilizado na literatura. O segundo conjunto de resultados experimentais considera uma sequência de imagens reais, capturadas com uma câmara digital comercial.

3.3.1 Imagens Simuladas

O objetivo dessa seção é comparar o desempenho do algoritmo INLA SR, proposto nessa Tese, com outros algoritmos para aumento de resolução. No primeiro experimento, uma sequência de quadros LR foi criada, a partir de uma imagem HR de um Barco, de tamanho 256×256 HR (Figura 3.2a). Para criar tais imagens LR, essa imagem do Barco foi deslocada, por uma quantidade aleatória de pixels a partir de uma distribuição contínua uniforme ao longo do intervalo $(-2, 2)$, em ambas direções. A imagem foi considerada periódica nas duas direções. Então, para simular o efeito *Point Spread Function* (PSF) da câmera, uma função de aquisição, que é o valor médio de um quadrado de 2×2 pixels da imagem HR, foi aplicada para formar um pixel da imagem LR. Essa função de aquisição corresponde a

aplicar um filtro passa-baixas e, em seguida, uma subamostragem com fator de decimação 2 nas direções horizontal e vertical. Finalmente, um ruído Gaussiano com média zero e desvio padrão 8, como adotado em (Tian e Ma, 2010), é adicionado a cada imagem processada, para se obter uma imagem de baixa resolução observável. Tal abordagem foi utilizada para gerar 16 imagens LR do mesmo Barco, a partir da cena original. Uma dessas imagens LR, de tamanho 128×128 , é apresentada na Figura 3.2b. Dessas imagens LR, quatro experimentos são executados com o objetivo de restaurar a imagem HR, cada um deles considerando quatro imagens LR.

O resultado implementando o método não iterativo de Superresolução descrito em (Pham et al., 2006) é mostrado na Figura 3.2c. Para gerar tal resultado foi usado o programa "Superresolution 2.0" (Vandewalle et al., 2006), disponível em <http://lcamtuf.ch/software/superresolution> para *download*. O resultado implementando o algoritmo de Superresolução Robusto com norma L1, descrito em (Farsiu et al., 2004b), é mostrado na Figura 3.2d. Para gerar tal resultado foi usado o programa "MDSP Resolution Enhancement Software", disponível para *download* em <http://users.soe.ucsc.edu/~milanfar/software>, com parâmetros configurados como: *regularization factor* = 0,01; *number of interactions* = 50; *regularization spatial decaying coefficient* = 0,8; *regularization kernel size* = 2; *step size* = 20; *start point = shift an add*.

O resultado mostrado na Figura 3.2e é uma variação do INLA SR aqui proposto, que a partir desse ponto é referenciado como INLA MAP, conforme definido na subseção 3.2.7. Nesse algoritmo \mathbf{T} foi definido como um vetor contendo 20 realizações da imagem HR, e \mathbf{P} como um vetor contendo a probabilidade associada a cada uma dessas realizações. O algoritmo INLA MAP corresponde à imagem HR com maior probabilidade no vetor \mathbf{T} . Na Figura 3.1 $p(\mathbf{X}|\mathbf{Y})$ versus λ foi plotado para o primeiro experimento. O ponto de máximo desse gráfico ($\lambda = 0,0005$) corresponde à solução INLA MAP. Assim, a solução INLA MAP, que é outra contribuição dessa Tese (Camponez et al., 2012c), corresponde a uma versão automática do algoritmo de Superresolução em forma fechada proposto nessa Tese, apresentado no Capítulo 2, e publicado em (Camponez et al., 2011b).

Finalmente, o resultado mostrado na Figura 3.2f corresponde ao método INLA SR, com os parâmetros estabelecidos conforme a subseção 3.2.6.

Analisando as imagens da Figura 3.2 pode ser visto que tanto o INLA MAP quanto o INLA SR obtiveram resultados muito bons em recuperar os detalhes que foram perdidos nos quadros LR. Isso fica claro pela melhor visibilidade dos fios que atravessam o mastro do barco, e, também, pela melhora na visibilidade da palavra "LACOR", no casco do barco. No entanto, a imagem gerada pela abordagem INLA SR parece ter menos ruído. Na Tabela 3.1 os vários algoritmos são comparados, usando a PSNR como métrica. Por essa métrica,

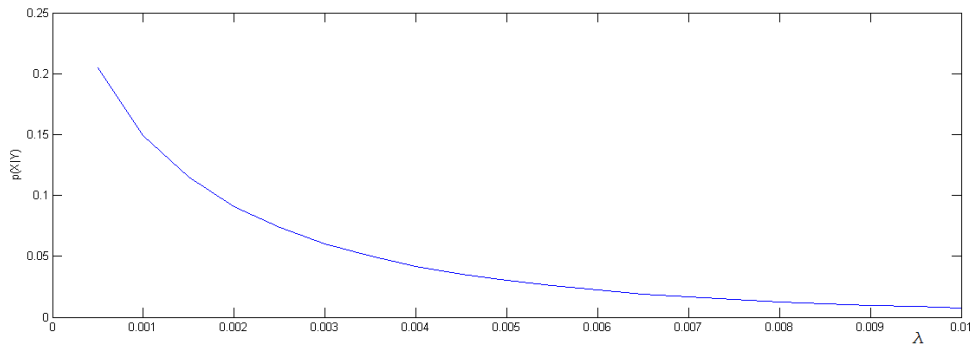


Figura 3.1: O gráfico da PDF $p(\mathbf{X}|\mathbf{Y})$ (equação (3.52)) versus λ , para o primeiro experimento com quatro quadros LR da imagem Barco

o método INLA SR, aqui proposto, apresenta os melhores resultados, seguido de perto pelo método INLA MAP, também aqui proposto.

Na segunda experiência, uma imagem de LR foi criada usando uma imagem HR Texto (Figura 3.3a), de tamanho 200×200 , usando o mesmo procedimento descrito na primeira experiência. A mesma abordagem foi utilizada para gerar 16 imagens LR Texto, a partir da cena original. Uma das imagens LR, de tamanho 100×100 LR, é apresentada na Figura 3.3b. Dessas imagens LR, quatro experimentos são executados, com o objetivo de restaurar a imagem HR Texto, cada um considerando quatro imagens LR.

O resultado implementando o algoritmo de Superresolução não iterativo descrito em (Pham et al., 2006) é mostrado na Figura 3.3c. Para gerar tal resultado foi usado o mesmo programa do primeiro experimento. O resultado implementando o método de Superresolução Robusto com norma L1 descrito em (Farsiu et al., 2004b) é mostrado na Figura 3.3d. Foi usado o mesmo programa descrito no primeiro experimento, com os parâmetros configurados como: *regularization factor* = 0,005; *number of interactions* = 50; *regularization spatial decaying coefficient* = 0,8; *regularization kernel size* = 2; *step size* = 20; *start point* = *shift an add*.

O resultado implementando INLA MAP é mostrado na Figura 3.3e, e o resultado do método INLA SR, com parâmetros configurados conforme a subseção 3.2.6, é mostrado na Figura 3.3f.

Analisando os resultados da Figura 3.3 pode-se notar que as imagens geradas pelos algoritmos INLA MAP (Figura 3.3e) e INLA SR (Figura 3.3f) são melhores que os outros. A imagem na Figura 3.3c, por exemplo, é muito suavizada. Por sua vez, a imagem na Fig. 3.3d apresenta alguns artefatos em sua margem esquerda, parecendo que está rotacionada para a direita. Na Tabela 3.1 os vários algoritmos são comparados usando-se PSNR. Por

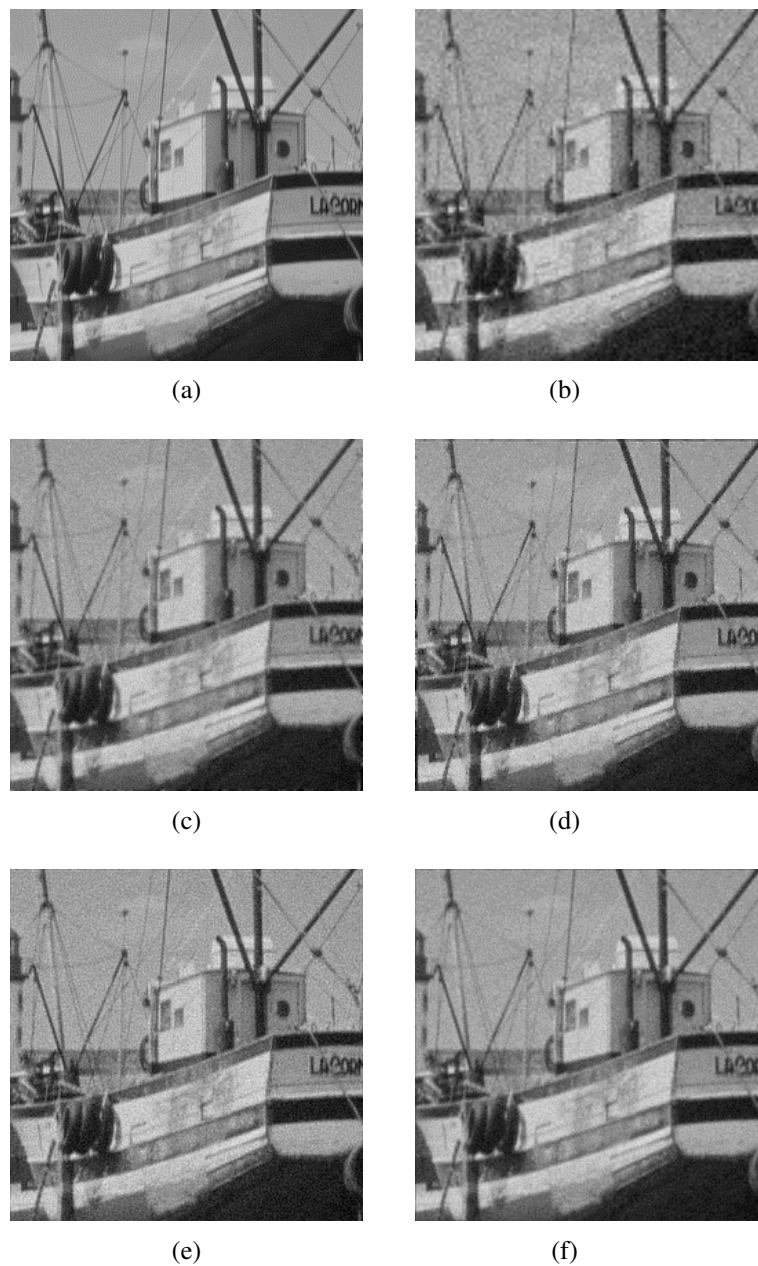


Figura 3.2: O conjunto de imagens HR reconstruídas usando a imagem de teste Barco. (a) A imagem original (*ground truth*), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) em (Pham et al., 2006), (d) em (Farsiu et al., 2004b), (e) INLA MAP proposto aqui (f) INLA SR proposto aqui.

essa métrica, novamente, os métodos INLA SR e INLA MAP apresentaram os melhores resultados.

Em (Tian e Ma, 2010) um algoritmo para recuperação de imagens por Superresolução usando MCMC foi proposto. Nos experimentos com imagens simuladas, os autores utili-

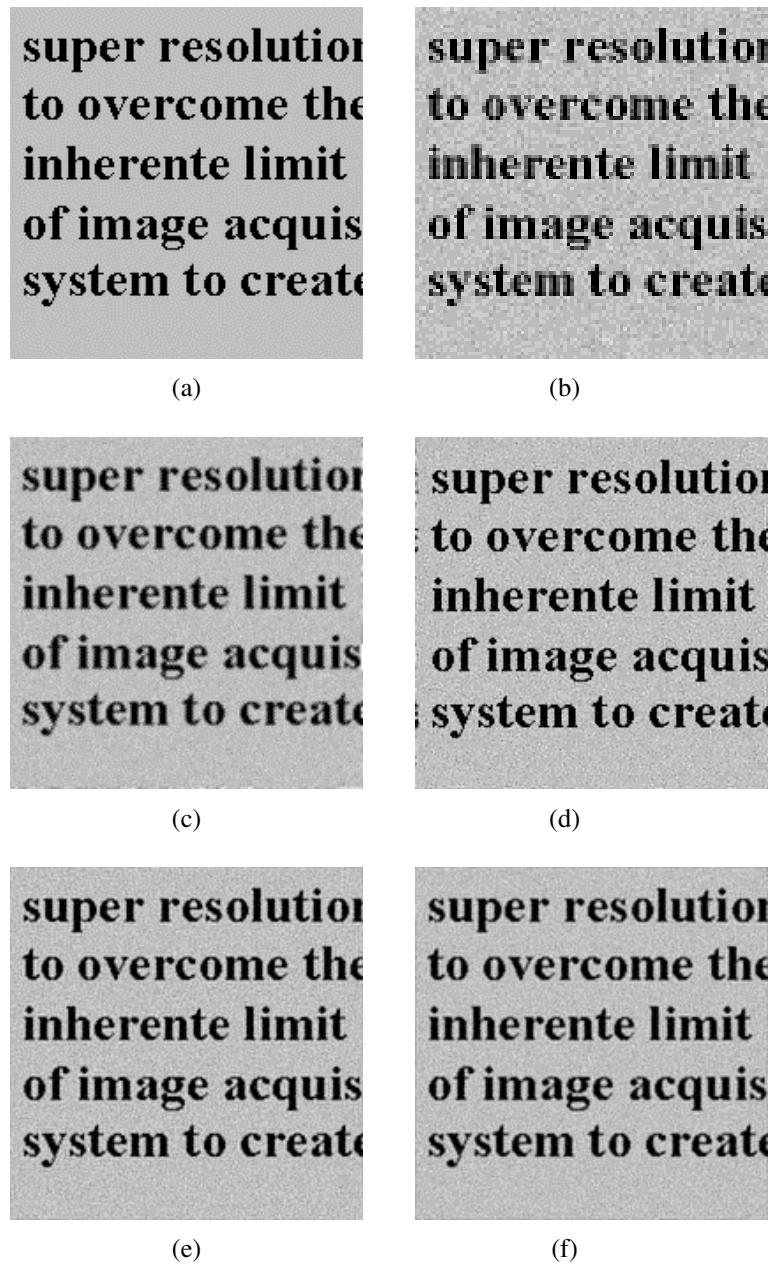


Figura 3.3: O conjunto de imagens HR reconstruídas usando a imagem de teste Texto. (a) A imagem original (*ground truth*), (b) uma das quatro imagens simuladas. Imagens HR reconstruídas usando a abordagem: (c) em (Pham et al., 2006), (d) em (Farsiu et al., 2004b), (e) INLA MAP proposto aqui (f) INLA SR proposto aqui.

zaram as mesmas imagens HR apresentadas na Figura 3.2a e na Figura 3.3a. O mesmo procedimento para gerar os quadros LR também foi utilizado. Assim como nessa seção, imagens HR foram recuperadas a partir de um conjunto de quatro imagens LR. Em seguida, a métrica de desempenho PSNR foi utilizada para comparar a abordagem MCMC com outros algoritmos. O método MCMC SR apresentou PSNR de 28,02 e 20,17, respectivamente,

para as imagens do Barco e do Texto. Nos dois casos, a Tabela 3.1 sugere que o desempenho dos algoritmos INLA SR e INLA MAP, propostos nessa Tese, é superior. Nós não tivemos acesso às imagens LR utilizadas em tal artigo, mas em nossas simulações, mesmo com diferentes conjuntos de quadros LR, os valores de PSNR na Tabela 3.1 não se alteraram significativamente.

Tabela 3.1: Avaliação das imagens HR reconstruídas considerando PSNR (dB)

Imagem		Pham	Farsiu	INLA	INLA
Test	Run	et al.	et al.	MAP	SR
		(Pham et al., 2006)	(Farsiu et al., 2004b)	(novo)	(novo)
Barco	1	27,20	19,44	28,88	30,45
	2	26,98	19,77	29,08	30,26
	3	27,15	18,23	29,17	30,27
	4	26,79	18,26	29,20	30,31
Média		27,03	18,93	29,08	30,32
Texto	1	17,93	10,76	22,67	21,50
	2	17,51	11,89	21,88	21,44
	3	17,57	10,54	22,05	21,63
	4	17,51	10,53	22,04	21,64
Média		17,63	10,93	21,99	21,55

3.3.2 Imagens Capturadas

A terceira avaliação de desempenho dos algoritmos de Superresolução é realizada usando-se imagens reais LR. Um conjunto de imagens chamadas de Disk, composto por 26 imagens não comprimidas em 256 tons de cinza, cada uma com tamanho 57×49 pixels, capturadas por uma câmera Olympus C-4000, disponível em <http://www.soe.ucsc.edu/milanfar/software/sr-datasets.html>, foi usado. Os primeiros 20 quadros seguem aproximadamente

o modelo de movimento global translacional. Destes 20 primeiros quadros, 4 foram escolhidos para esse experimento. Em seguida, essas 4 imagens foram recortadas de maneira a ficarem de tamanho 48×48 (uma delas é apresentada na Figura 3.4a). O experimento consiste em recuperar uma imagem HR, de tamanho 96×96 , a partir dessas quatro imagens LR.

Nessa simulação, a função PSF da câmera é assumida como uma função de aquisição, onde cada pixel da imagem LR é gerado pelo valor médio de um quadrado 2×2 de pixels da imagem HR, como na seção anterior. Além disso, as imagens são consideradas periódicas em ambas as direções.

O resultado implementando o método proposto em (Farsiu et al., 2004b) é mostrado na Figura 3.4b. Para gerar tal resultado foi usado o programa "MDSP Resolution Enhancement Software", disponível para *download* em <http://users.soe.ucsc.edu/~milanfar/software>, com parâmetros configurados como: *regularization factor* = 0,005; *number of interactions* = 50; *regularization spatial decaying coefficient* = 0,8; *regularization kernel size* = 2; *step size* = 20; *start point* = *shift an add*.

O resultado do método INLA MAP é mostrado na Figura 3.4c e o resultado para o INLA SR, com parâmetros configurados como: $\sigma = 24$; o conjunto de valores de $\lambda = [0,0001 \ 0,0002 \ 0,0003 \ 0,0004 \ 0,0005]$, é mostrado na Figura 3.4d.

Analisando a Figura 3.4 pode-se observar que os três algoritmos obtiveram sucesso ao recuperar detalhes que não apareciam nas imagens LR. As imagens da Figura 3.4b e Figura 3.4c são muito similares, enquanto a imagem da Figura 3.4d é um pouco mais suavizada. Esses resultados foram publicados em (Campanez et al., 2011a; Campanez et al., 2012c).

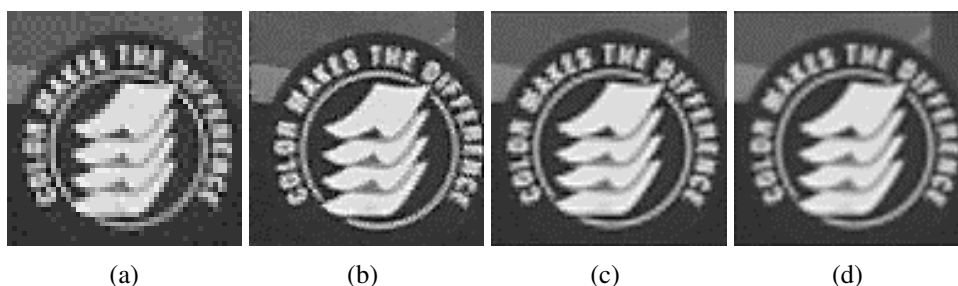


Figura 3.4: O conjunto de imagens HR reconstruídas a partir de uma sequência de quatro imagens reais LR. (a) Uma das quatro imagens LR. Imagens HR reconstruídas usando a abordagem: (b) em (Farsiu et al., 2004b), (c) INLA SR proposto aqui (d) INLA MAP proposto aqui.

3.4 Conclusões do Capítulo

Neste Capítulo um método de inferência não-paramétrico, novo e com grande potencial, não usado até esse momento, até onde o autor conhece, na área de processamento de imagens, foi aplicado ao problema de Superresolução, e explorado com sucesso em dois algoritmos: INLA SR e INLA MAP. O resultado é uma abordagem que alcança consistentemente desempenhos superiores na reconstrução de imagens HR, quando comparada a outros métodos do estado da arte.

Tal método de inferência, referido como INLA, é um método novo de inferência Bayesiana estatística, com base em aproximações das funções densidade de probabilidade (PDF) e integração numérica, de uma forma não-paramétrica. Ao aplicar INLA ao problema de Superresolução, essa Tese demonstra que todas as equações necessárias para implementar a técnica proposta podem ser escritas em forma fechada.

Experimentos com imagens simuladas mostram que os dois algoritmos propostos nesse Capítulo, o INLA SR e o INLA MAP, atingem valores mais elevados de PSNR, e, portanto, alcançam melhores resultados, na comparação objetiva com outros algoritmos do estado da arte. As imagens geradas, como pode ser verificado nas Figuras 3.2 e 3.3, corroboram essa superioridade. Os resultados da experiência com imagens reais, como pode ser visto na Figura 3.4, confirmam, mais uma vez, o bom desempenho dos dois métodos aqui propostos.

Comparando os dois algoritmos propostos, o INLA MAP é o que, em geral, gera imagens com maior nitidez de bordas. Mas as experiências mostraram que o INLA SR pode ser muito útil em situações em que as imagens LR são muito ruidosas, particularmente onde o ruído é Gaussiano, porque pode obter imagens HR nítidas, como boa transição de bordas e com menos ruído.

Entretanto, as matrizes manipulados pelos algoritmos de Superresolução INLA SR e INLA MAP são, geralmente, grandes, embora esparsas. Assim, mesmo utilizando estruturas de dados adequadas para armazenar e manipular matrizes esparsas, o uso de memória é grande e o esforço computacional intenso. Assim, soluções que reduzam o custo computacional e a demanda por alocação de memória, otimizando o desempenho, são desejáveis.

Capítulo 4

Superresolução Rápida Baseada em DFT-2D

Nos dois Capítulos anteriores foram propostos três algoritmos para reconstrução de imagens com o uso de Superresolução. No primeiro algoritmo (Closed Form SR), proposto no Capítulo 2, a imagem de alta resolução é reconstruída usando-se uma equação em forma fechada. Nas outras duas implementações (INLA SR and INLA MAP) um método de inferência Bayesiana recentemente desenvolvido, baseado em integração numérica, conhecido na literatura estatística como *Integrated Nested Laplace Approximation* (INLA), foi aplicado ao problema de recuperação de imagens por Superresolução.

Tais algoritmos, quando comparados a outras técnicas do estado da arte, na recuperação de imagens por Superresolução, alcançam resultados superiores, conforme demonstram os experimentos mostrados nos Capítulos 2 e 3. Entretanto, as matrizes manipuladas são muito grandes, embora esparsas. Assim, mesmo utilizando-se estruturas de dados adequadas para armazenar e manipular matrizes esparsas, a demanda por alocação de memória RAM e o custo computacional são altos.

Neste Capítulo, esses problemas são tratados. Relaxando-se uma equação no modelo matemático original e envolvendo a imagem HR em um toróide \mathcal{T}_n mostra-se que o custo computacional dos algoritmos de Superresolução Closed Form SR, INLA SR e INLA MAP pode ser significativamente reduzido, bem como as dimensões das matrizes manipuladas (de $n^2 \times n^2$ para $n \times n$, o tamanho das imagens HR), ao custo de uma pequena redução da qualidade da imagem HR. Essa nova proposta é explorada em dois novos algoritmos baseados na Transformada Discreta de Fourier 2D (2D DFT).

4.1 Alterações no Modelo de Aquisição de Imagens Warp-Blur

O modelo de observação descreve a relação matemática entre as imagens LR e a imagem HR durante o processo de aquisição de imagens. Conforme descrito na Seção 2.2.1, o modelo de aquisição adotado nesta Tese é o Warp-Blur. Há alguns resultados analíticos muito interessantes a respeito de matrizes em bloco, mais especificamente nas matrizes do tipo Toeplitz circulantes em bloco (BTC), que são explorados neste Capítulo. Então, sem perda de generalidade, visto que a mesma metodologia pode ser aplicada qualquer que seja a dimensão da imagem HR e qualquer que seja o fator de decimação utilizado, dois pressupostos, que modificam o modelo de aquisição de imagens originalmente usado, são adotados:

1. considera-se que a imagem (HR) é quadrada;
2. e que um fator de decimação de dois, em ambas as direções, horizontal e vertical, é aplicado para geração das imagens LR.

Esses pressupostos são adotados porque facilitam quantificar o ganho, em termos de desempenho computacional e uso de memória RAM, dos novos algoritmos desenvolvidos neste Capítulo.

Esse modelo de aquisição Warp-Blur simplificado pode ser matematicamente expresso como:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{X} + \boldsymbol{\varepsilon}_k, \quad (4.1)$$

onde \mathbf{y}_k representa a k -ésima imagem LR, de dimensões $\frac{n}{2} \times \frac{n}{2}$ e \mathbf{X} representa a k -ésima imagem HR, de dimensão $n \times n$, ambas ordenadas lexicograficamente como um vetor, com tamanhos $\frac{n^2}{4} \times 1$ e $n^2 \times 1$, respectivamente. \mathbf{H}_k é uma matriz de dimensões $\frac{n^2}{4} \times n^2$ que representa o produto de outras duas matrizes, \mathbf{D} uma matriz $\frac{n^2}{4} \times n^2$ que modela as operações de convolução e sub-amostragem, e \mathbf{M}_k uma matriz $n^2 \times n^2$ que modela o efeitos de deformação (isto é, translações e rotações), conforme previamente definidas no Capítulo 2, Seção 2.2.1. Finalmente, $\boldsymbol{\varepsilon}_k$ é um vetor $\frac{n^2}{4} \times 1$, representando o ruído aditivo Gaussiano branco, com média zero e variância σ_k^2 .

O modelo Bayesiano Hierárquico usado continua o mesmo do Capítulo 2, e por isso não necessita de ser repetido aqui.

4.2 Superresolução Baseada em DFT-2D

A equação (2.50) tem papel chave na implementação de um algoritmo rápido. Note-se que a inversão no primeiro termo corresponde ao maior custo computacional envolvido nos diversos métodos propostos nesta Tese. Nesta Seção, um novo algoritmo para fusão da imagem HR, muito mais rápido, é desenvolvido.

As distribuições de probabilidade podem ser representadas por diagramas conhecidos como modelos gráficos probabilístico (PGM) (Johnson e Wichern, 2002). Em um PGM, cada nó representa uma variável aleatória, e os *links* expressam as relações de probabilidade entre tais variáveis. Uma classe especial de modelos gráficos, chamada de campo aleatório Markoviano Gaussiano (GMRF) desperta um interesse especial nesse trabalho. Um GMRF é um vetor aleatório seguindo uma distribuição multivariável Gaussiana onde são assumidas certas condições de independência entre as variáveis aleatórias, daí o aparecimento do termo Markoviano como uma referência às cadeias de Markov (Rue e Held, 2005).

Neste Capítulo uma classe especial de GMRF, conhecida com GMRF estacionária, é usada. GMRF estacionárias podem ser obtidas assumindo-se dois pressupostos. O primeiro, as entradas da matriz de precisão \mathbf{Q} são parâmetros constantes, como em (2.25). Segundo, o grafo é frequentemente um toróide \mathcal{T}_n .

Dessa forma, a imagem HR é envolvida em uma grafo toroidal \mathcal{T}_n . Isso faz com que a matriz \mathbf{Q} assuma a forma BTC, para a qual existem resultados analíticos muito interessantes a respeito da estrutura dos autovalores. A vantagem prática é que operações típicas em GMRF estacionários podem ser feitas através do uso da Transformada de Fourier Discreta (DFT). A complexidade computacional típica para essas operações é de $O(n \log(n))$, para n termos a serem transformados, e não depende do número de vizinhos no grafo (Rue e Held, 2005).

4.2.1 Matrizes Toeplitz Circulante

Qualquer matriz \mathbf{C} , de tamanho $n \times n$, é chamada de Toeplitz circulante (Rue e Held, 2005; Gray, 2006) se assume a forma

$$\mathbf{C} = \begin{pmatrix} c_0 & c_1 & c_2 & \dots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \dots & c_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ c_1 & c_2 & c_3 & \dots & c_0 \end{pmatrix}. \quad (4.2)$$

Então, tal matriz é totalmente especificada por seu vetor base

$$\mathbf{c} = (c_0, c_1, \dots, c_{n-1})^T. \quad (4.3)$$

Assim, uma matriz circulante pode ser expressa como

$$\mathbf{C} = \mathbf{F}\Lambda\mathbf{F}^H, \quad (4.4)$$

para alguma matriz diagonal Λ contendo os autovalores de \mathbf{C} . As matrizes \mathbf{F} e \mathbf{F}^H são, respectivamente, a transformada discreta de Fourier e sua inversa. Note-se que \mathbf{F} é uma matriz unitária, isto é, $\mathbf{F}^{-1} = \mathbf{F}^H$, onde \mathbf{F}^H é a matriz transposta conjugada de \mathbf{F} , e

$$\Lambda = \sqrt{n} \text{diag}(\mathbf{F}\mathbf{c}), \quad (4.5)$$

onde $\text{diag}(\cdot)$ é uma matriz diagonal.

4.2.2 Matrizes Toeplitz Circulante em Bloco

Qualquer matriz \mathbf{C} , de tamanho $n^2 \times n^2$, é BTC (Rue e Held, 2005; Gray, 2006), com blocos $n \times n$ se tem a forma

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_1 & \mathbf{C}_2 & \dots & \mathbf{C}_{n-1} \\ \mathbf{C}_{n-1} & \mathbf{C}_0 & \mathbf{C}_1 & \dots & \mathbf{C}_{n-2} \\ \mathbf{C}_{n-2} & \mathbf{C}_{n-1} & \mathbf{C}_0 & \dots & \mathbf{C}_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 & \dots & \mathbf{C}_0 \end{pmatrix}, \quad (4.6)$$

onde \mathbf{C}_i é uma matriz $n \times n$ Toeplitz circulante com base \mathbf{c}_i . A base de \mathbf{C} é uma matriz $n \times n$

$$\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{n-1}). \quad (4.7)$$

Similarmente às matrizes circulantes, matrizes BTC podem ser expressas como

$$\mathbf{C} = (\text{DFT2}) \Gamma (\text{IDFT2}), \quad (4.8)$$

para alguma matriz diagonal Γ contendo os autovalores de \mathbf{C} . DFT2 e IDFT2 são, respectivamente, a transformada de Fourier discreta em duas dimensões (2D DFT) e sua inversa.

Embora (4.8) mostre como computar os autovalores, essa expressão não será usada diretamente. Ao invés disso, será usada a relação entre a transformada de Fourier discreta 2D e os autovetores de uma matriz BTC (Rue e Held, 2005). Seja Γ uma matriz bloco diagonal contendo todos os $n \times n$ autovalores da matriz \mathbf{C} na diagonal. Armazene esses autovalores em uma matriz Π , de tamanho $n \times n$, onde a linha i de Π corresponde à diagonal Γ_i em Γ .

Considerando que \mathbf{F} é uma matriz que corresponde à aplicação da transformada de Fourier, Π pode ser calculada da seguinte maneira: compute a DFT de cada linha da matriz base \mathbf{c} , compute a DFT de cada coluna e escale ambas por n . O resultado é que Π é a transformada de Fourier discreta 2D da base \mathbf{c} .

Usando-se essa notação, a matriz $n \times n$

$$\Pi = n \text{DFT2}(\mathbf{c}), \quad (4.9)$$

contém todos os autovalores de \mathbf{C} , uma matriz BTC de base \mathbf{c} .

Seja \mathbf{B} uma matriz $n \times n$, \mathbf{X} uma matriz $n^2 \times n^2$ e $\bar{\mathbf{A}}$ uma matriz BTC, $n^2 \times n^2$, com base $\bar{\mathbf{a}}$, respectivamente. Além disso, define-se $\text{vec}(\cdot)$ como uma representação vetorial obtida sobrepondo-se as colunas de seus argumentos uma sobre a outra. De acordo com (Rue e Held, 2005), a solução de

$$\text{vec}(\mathbf{X}) = (\bar{\mathbf{A}})^{-1} \text{vec}(\mathbf{B}) \quad (4.10)$$

é

$$\mathbf{X} = \frac{1}{n} \text{DFT2}(\text{IDFT2}(\mathbf{B}) \oslash \text{DFT2}(\bar{\mathbf{a}})), \quad (4.11)$$

onde ‘ \oslash ’ denota uma divisão elemento por elemento.

Comparando-se (4.10) com a equação (2.49), pode-se notar que se \mathbf{A} for uma matriz BTC basta uma aplicação direta de (4.11) para a resolução, de uma maneira muito eficiente, do problema da reconstrução de imagens por Superresolução. Tal solução reduz tanto o custo computacional quanto o consumo de memória. A matriz \mathbf{A} , como pode ser visto em (2.50), é formada por um conjunto de matrizes, algumas das quais são BTC e outras que são bloco Toeplitz circulantes em blocos (uma hierarquia superior de matrizes Toeplitz circulantes). Entretanto, \mathbf{A} não é BTC e, portanto, a aplicação direta de (4.11) não é possível. Na próxima seção, uma aproximação que torna essa matriz BTC é proposta.

4.2.3 Aproximação BTC para a Matriz \mathbf{A}

Nessa seção uma aproximação é feita na matriz \mathbf{A} in (2.49) de maneira a torná-la BTC, possibilitando o desenvolvimento de um algoritmo de reconstrução de imagens mais eficiente. A equação (2.50) pode ser reescrita como

$$\bar{\mathbf{X}} = \left(\lambda \mathbf{Q} + \frac{(\mathbf{D}\mathbf{M}_1)^T (\mathbf{D}\mathbf{M}_1) + \dots + (\mathbf{D}\mathbf{M}_m)^T (\mathbf{D}\mathbf{M}_m)}{\sigma^2} \right)^{-1} \mathbf{B} \quad (4.12)$$

para m imagens LR, ou,

$$\bar{\mathbf{X}} = \left(\lambda \mathbf{Q} + \frac{(\mathbf{M}_1^T \mathbf{D}^T \mathbf{D} \mathbf{M}_1) + \dots + (\mathbf{M}_m^T \mathbf{D}^T \mathbf{D} \mathbf{M}_m)}{\sigma^2} \right)^{-1} \mathbf{B}. \quad (4.13)$$

O próximo passo é envolver a imagem HR em um grafo toroidal \mathcal{T}_n . Isso faz com que a matriz \mathbf{Q} , de tamanho $n^2 \times n^2$, se torne BTC, com blocos $n \times n$, tendo como base a matriz \mathbf{q} , de tamanho $n \times n$. Então, conforme (4.8) essa matriz pode ser expressa como

$$\mathbf{Q} = (DFT2) \Gamma_q (IDFT2), \quad (4.14)$$

para alguma matriz diagonal Γ_q contendo os autovalores de \mathbf{Q} .

As matrizes \mathbf{M}_k , de tamanho $n^2 \times n^2$, conforme definidas na Subseção 2.2.3, também são BTC, com blocos $n \times n$, e tendo como base a matriz \mathbf{v}_k , de tamanho $n \times n$. Então,

$$\mathbf{M}_k = (DFT2) \Gamma_{v_k} (IDFT2), \quad (4.15)$$

para alguma matriz diagonal Γ_{v_k} contendo os autovalores de \mathbf{M}_k .

No caso geral, a matriz $\mathbf{D}^T \mathbf{D}$, de tamanho $n^2 \times n^2$, é bloco Toeplitz circulante em blocos, o que não possibilita a aplicação de algoritmos como em (4.11). Somente em casos especiais é que essa matriz pode tomar a forma BTC, com blocos $n \times n$. Estamos interessados em explorar exatamente esse caso onde a matriz é BTC, com blocos $n \times n$, tendo como base a matriz \mathbf{d} , de tamanho $n \times n$. Neste caso,

$$\mathbf{D}^T \mathbf{D} = (DFT2) \Gamma_d (IDFT2), \quad (4.16)$$

para alguma matriz diagonal Γ_d contendo os autovalores de $\mathbf{D}^T \mathbf{D}$.

Então, substituindo-se (4.14), (4.15) e (4.16) em (4.13)

$$\begin{aligned} \bar{\mathbf{X}} = & \left((DFT2) \lambda \Gamma_q (IDFT2) + (DFT2) \frac{\Gamma_{v1}^T \Gamma_d \Gamma_{v1}}{\sigma^2} (IDFT2) \right. \\ & \left. + \dots + (DFT2) \frac{\Gamma_{vm}^T \Gamma_d \Gamma_{vm}}{\sigma^2} (IDFT2) \right)^{-1} \mathbf{B}, \end{aligned} \quad (4.17)$$

mas $\Gamma_{vk}^T \Gamma_d \Gamma_{vk}$, para $k = 1 \dots m$, são matrizes diagonais, onde a multiplicação é comutativa. Além disso, Γ_{vk}^T e Γ_{vk} são matrizes ortogonais, de maneira que o produto das duas é igual à matriz identidade. Então,

$$\bar{\mathbf{X}} = \left((DFT2) \lambda \Gamma_q (IDFT2) + (DFT2) \frac{\Gamma_d}{\sigma^2} (IDFT2) + \dots + (DFT2) \frac{\Gamma_d}{\sigma^2} (IDFT2) \right)^{-1} \mathbf{B}, \quad (4.18)$$

de onde se obtém,

$$\bar{\mathbf{X}} = \left((DFT2) \left(\lambda \Gamma_q + \frac{\Gamma_d}{\sigma^2} + \dots + \frac{\Gamma_d}{\sigma^2} \right) (IDFT2) \right)^{-1} \mathbf{B}, \quad (4.19)$$

e

$$\bar{\mathbf{X}} = \left((DFT2) \left(\lambda \Gamma_q + \frac{m \Gamma_d}{\sigma^2} \right)^{-1} (IDFT2) \right) \mathbf{B}, \quad (4.20)$$

onde m é o número de imagens LR.

Embora a equação (4.20) só possa ser usada em um caso particular, ela foi importante para guiar a escolha da aproximação usada na matriz \mathbf{A} , de forma a torná-la BTC. Note-se que a primeira parte da equação (4.12) $\lambda \mathbf{Q}$ é BTC, uma vez que a imagem HR foi envolvida em um toróide \mathcal{T}_n . A segunda parte dessa equação também é BTC. Entretanto o tamanho dos blocos em cada uma das partes é diferente, o que não possibilita a implementação de algoritmos baseados em DFT-2D. A aproximação proposta nesta Tese consiste em substituir o somatório de matrizes da segunda parte da equação (4.12) pela matriz identidade \mathbf{I} . Tal matriz além de ser Toeplitz Circulante, também é BTC qualquer que seja o tamanho do bloco, o que permitirá a implementação de algoritmos mais rápidos baseados em DFT-2D. Assim a aproximação proposta, baseada no resultado em (4.20), pode ser escrita como

$$\bar{\mathbf{X}} \approx - \left(\lambda \mathbf{Q} + \frac{m \mathbf{I}}{4\sigma^2} \right)^{-1} \mathbf{B}, \quad (4.21)$$

onde $\frac{m}{4}$ é uma constante empiricamente definida como o número de imagens LR dividido por quatro, e \mathbf{I} é uma matriz identidade de tamanho $n^2 \times n^2$.

4.2.4 Algoritmo DFT Closed Form SR

Um novo algoritmo, chamado aqui de DFT Closed Form SR, pode ser desenvolvido aplicando-se (4.9) e (4.11) a (4.21). Assim, obtém-se

$$\bar{\bar{\mathbf{X}}} = \frac{1}{n} DFT2 \left(IDFT2(\bar{\mathbf{B}}) \oslash \left(DFT2(\lambda \mathbf{q}) + \frac{m \Pi_I}{4\sigma^2} \right) \right), \quad (4.22)$$

que corresponde a uma equação em forma fechada, baseada em DFT-2D, para calcular $\bar{\bar{\mathbf{X}}}$, uma matriz $n \times n$, que corresponde à imagem HR, onde \mathbf{q} é uma matriz $n \times n$, base de \mathbf{Q} , que tem a forma

$$\mathbf{q} = \begin{pmatrix} 4 & -1 & 0 & \dots & 0 & -1 \\ -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \quad (4.23)$$

Π_I é uma matriz, de tamanho $n \times n$, contendo os autovalores de \mathbf{I} , isto é,

$$\Pi_{I(i,j)} = 1; \forall i e j, i e j \in [n, n] \quad (4.24)$$

e $\bar{\mathbf{B}}$ é uma matriz, de tamanho $n \times n$, cujas colunas são tomadas de \mathbf{B} em (2.49).

4.2.5 Algoritmo DFT INLA SR

Outro novo algoritmo, baseado no método de inferência Bayesiana INLA, denominado, daqui por diante, DFT INLA SR, pode ser sumarizado como:

1. defina o grid de λ como $\lambda = (0, 0005 \ 0, 0010 \ \dots \ 0, 01)$;
2. para $i = 1$ para 20 faça;
3. calcule o vetor $\bar{\mathbf{T}}(i)$ usando (4.22);
4. calcule o vetor $\bar{\mathbf{P}}(i) = p(\bar{\mathbf{T}}(i)|\mathbf{Y})$ usando (3.52);
5. fim do loop de repetição;
6. reescale $\bar{\mathbf{P}}(i)$, dividindo-se cada elemento pela soma de todos os elementos;

Então, no algoritmo DFT INLA SR a imagem HR $\bar{\bar{\mathbf{X}}}$ é calculada como o valor esperado de \mathbf{X} , isto é,

$$\bar{\bar{\mathbf{X}}} = \sum_{i=1}^{20} \bar{\mathbf{T}}(i) \bar{\mathbf{P}}(i). \quad (4.25)$$

4.3 Experimentos

A fim de demonstrar o desempenho dos algoritmos baseados em DFT-2D, propostos neste Capítulo, alguns experimentos são apresentados nesta Seção. O primeiro conjunto de experimentos envolve dados simulados (imagens LR) derivados a partir de uma imagem HR. Esse experimento permite análise quantitativa de erros, usando-se como métrica o PSNR, frequentemente usada na literatura, e o tempo para convergência dos algoritmos. O segundo conjunto de experimentos considera uma sequência de imagens reais capturadas com uma câmera digital comercial. O terceiro experimento considera a mesma sequência de imagens LR do primeiro experimento, mas agora usando-se mais imagens LR como entrada dos algoritmos para geração da imagem HR. Esses dados permitem comparar o desempenho dos algoritmos através de PSNR e do tempo de convergência, usando-se quatro e, depois, treze imagens LR como entrada.

4.3.1 Imagem Simulada

O objetivo dessa subseção é comparar o desempenho dos algoritmos DFT Closed Form SR e do DFT INLA SR, ambos propostos nesta Tese, com outros algoritmos para melhoria de imagens. No primeiro experimento, uma sequência de quadros LR é criada a partir da imagem HR Barco de tamanho 256×256 (Fig. 4.1a). Para criar tais imagens LR, a imagem HR Barco foi deslocada por uma quantidade aleatória de pixels a partir de uma distribuição contínua uniforme ao longo do intervalo $(-2, 2)$, em ambas direções.

A imagem foi considerada periódica nas duas direções. Então, para simular o efeito *Point Spread Function* (PSF) da câmera fotográfica, uma função de aquisição, que equivale ao valor médio de uma área quadrada de 2×2 , em pixels, da imagem HR, foi aplicada para formar um elemento da imagem LR. A função de aquisição usada corresponde a se aplicar um filtro passa-baixas seguido de uma operação de subamostragem com fator de decimação de dois em ambas as direções, vertical e horizontal.

Finalmente, um ruído Gaussiano com média zero e desvio padrão 8, como adotado em (Tian e Ma, 2010), é adicionado a cada imagem processada, para se obter uma imagem de baixa resolução observável. Uma dessas imagens LR, de tamanho 128×128 , é apresentada na Figura 4.1b. De tais imagens LR, são executados quatro experimentos com o objetivo de restaurar a imagem HR, cada um considerando quatro imagens LR.

O resultado da implementação dos métodos de Superresolução descritos em (Vandewalle et al., 2006) e (Pham et al., 2006) são mostrados nas Figura 4.1c e 4.1d, respectivamente. Para gerar tais resultados foi usado o *software* "Superresolution 2.0", disponível para *download* em <http://lcav.epfl.ch/software/superresolution>. O resultado implementando o método iterativo *Robust norm L1*, descrito em (Farsiu et al., 2004b), é mostrado na Figura 4.1e. Para gerar tal resultado foi usado o *software* "MDSP Resolution Enhancement Software", disponível para *download* em <http://users.soe.ucsc.edu/~milanfar/software>, com parâmetros configurados como: *regularization factor* = 0,05; *number of interactions* = 50; *regularization spatial decaying coefficient* = 0,8; *regularization kernel size* = 2; *step size* = 20; *start point* = *shift an add*.

O resultado mostrado na Figura 4.1f corresponde ao algoritmo Closed Form SR descrito em (Camponez et al., 2011b) e o resultado mostrado na Figura 4.1g corresponde ao método INLA SR descrito em (Camponez et al., 2012c).

Finalmente, o resultado implementando o método DFT Closed Form SR é mostrado na Figura 4.1h, e o resultado implementando o algoritmo DFT INLA SR é mostrado na Figura 4.1i.

Analisando as imagens da Figura 4.1 pode-se notar que tanto o algoritmo DFT Closed Form quanto o DFT INLA SR obtiveram bons resultados em recuperar o detalhes que haviam sido perdidos nos quadros LR. Isso pode ser notado pela melhoria da visibilidade dos cabos que passam pelo mastro do barco e, também, pela melhoria da visibilidade da palavra "LACOR" no casco do barco. Isso mostra que os dois algoritmos são competitivos quando comparados a outros do estado da arte, em uma avaliação subjetiva considerando-se a qualidade das imagens que eles geram. Na Tabela 4.1 os vários algoritmos são comparados usando-se a métrica PSNR (em dB) e o tempo para convergência do algoritmo (em segundos) (somente o tempo para fusão e filtragem das imagens é considerado, sendo o tempo gasto para inferir os parâmetros de registro das imagens desconsiderado). Considerando-se tais aspectos, os algoritmos DFT Closed Form SR e DFT INLA SR, propostos nesta Tese, apresentam resultados muito bons de PSNR, além de serem muito rápidos.

Em (Tian e Ma, 2010) um algoritmo de Superresolução usando MCMC foi proposto. Nos experimentos, usando imagens simuladas, os autores utilizaram a mesma imagem HR apresentada na Figura 4.1a, e o mesmo processo de geração de imagens LR. Então, a métrica PSNR foi usada para comparar o método usando MCMC com outros algoritmos. O método MCMC apresentou PSNR de 28,02 para a imagem do barco, conforme reproduzido na Tabela 4.1. Nesse caso, a Tabela 4.1 sugere que o desempenho de ambos DFT Closed Form SR e DFT INLA MAP foram superiores. Não foi possível o acesso às imagens LR usadas em tal artigo, mas em nossas simulações os valores de PSNR na Tabela 4.1 não se alteraram muito para diferentes conjuntos de imagens LR. Então, acredita-se que o resultado de tal comparação é confiável.

O algoritmo MCMC, de acordo com (Tian e Ma, 2010), foi escrito em linguagem C, e as simulações foram executadas em um computador do tipo PC com uma CPU Pentium 2.4 GHz e 512 MB de memória RAM. Os *softwares* "Superresolution 2.0", "MDSP Resolution Enhancement Software" e os algoritmos em (Camponez et al., 2011b), (Camponez et al., 2012c), os algoritmos DFT Closed Form SR e o DFT INLA SR foram programados em MATLAB[®], em um PC com CPU Pentium 1.86 GHz e 1 GB de memória RAM.

4.3.2 Imagem Real

A segunda avaliação de desempenho dos algoritmos de Superresolução é realizada usando-se imagens reais LR. Um conjunto de imagens, chamadas de Disk, composto por 26 imagens não comprimidas em 256 tons de cinza, cada uma com tamanho 57×49 pixels, capturadas por uma câmera Olympus C-4000, disponível em <http://www.soe.ucsc.edu/milanfar/software/sr-datasets.html>, foi usado. Os primeiros 20 quadros seguem aproximadamente

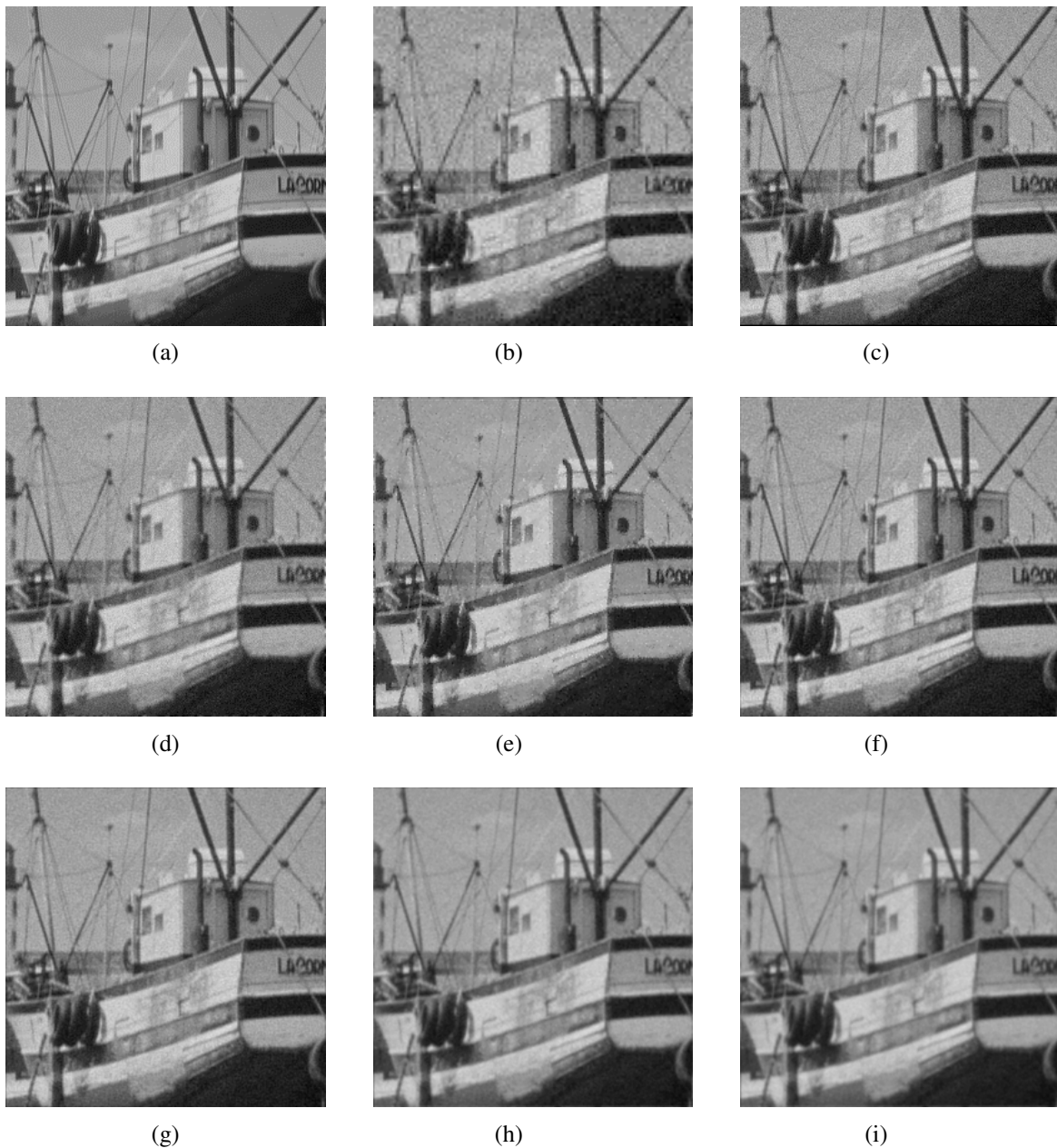


Figura 4.1: O conjunto de imagens reconstruídas usando-se a imagem teste Barco. (a) A imagem original (*ground truth*), (b) uma das imagens LR simuladas. Reconstrução de imagem HR usando-se o método: (c) em (Vandewalle et al., 2006), (d) em (Pham et al., 2006), (e) in (Farsiu et al., 2004b), (f) Closed Form SR (Camponez et al., 2011b), (g) INLA SR (Camponez et al., 2012c), (h) DFT Closed Form (Camponez et al., 2012b; Camponez et al., 2012a) e (i) DFT INLA SR (Camponez et al., 2012b; Camponez et al., 2012a).

o modelo de movimento global translacional. Destes 20 primeiros quadros, 4 foram escolhidos para esse experimento. Em seguida, essas 4 imagens foram recortadas de maneira a ficarem de tamanho 48×48 (uma delas é apresentada na Figura 4.2a). O experimento

consiste em recuperar uma imagem HR, de tamanho 96×96 , a partir dessas quatro imagens LR.

Nessa simulação, a função PSF da câmera é assumida como uma função de aquisição, onde cada pixel da imagem LR é gerado pelo valor médio de um quadrado 2×2 de pixels da imagem HR, como na seção anterior. Além disso, as imagens são consideradas periódicas em ambas as direções.

O resultado implementando o método proposto em (Farsiu et al., 2004b) é mostrado na Figura 4.2b. Para gerar tal resultado foi usado o programa "MDSP Resolution Enhancement Software", disponível para *download* em <http://users.soe.ucsc.edu/~milanfar/software>, com parâmetros configurados como: *regularization factor* = 0,005; *number of interactions* = 50; *regularization spatial decaying coefficient* = 0,8; *regularization kernel size* = 2; *step size* = 20; *start point* = *shift an add*.

O resultado do método Closed Form SR com $\lambda = 0,001$ é mostrado na Figura 4.2c e o resultado para o INLA SR, com parâmetros configurados como: $\sigma = 24$; o conjunto de valores de $\lambda = [0,0001 \ 0,0002 \ 0,0003 \ 0,0004 \ 0,0005]$, é mostrado na Figura 4.2d.

O resultado da implementação do método DFT Closed Form SR com $\lambda = 0,001$ é mostrado na Figura 4.2e e o resultado do algoritmo DFT INLA SR, com parâmetros $\sigma = 8$; o conjunto de valores de $\lambda = [0,0001 \ 0,0002 \ 0,0003 \ 0,0004 \ 0,0005]$, é mostrado na Figura 4.2f.

Analisando a Figura 4.2 pode-se notar que todos os algoritmos obtiveram sucesso em recuperar detalhes que não estavam presentes nas imagens LR. Em nossa avaliação as imagens das Figuras 4.2b e 4.2c são muito similares, enquanto as imagens apresentadas nas Figuras 4.2d, 4.2e e 4.2f são mais suavizadas. Na Tabela 4.2 os vários métodos, todos implementados em Matlab[®], são comparados usando-se o tempo de convergência do algoritmo (em segundos). Mais uma vez, os algoritmos DFT Closed Form SR e o DFT INLA SR são muito mais rápidos que os outros.

4.3.3 Usando Mais Imagens LR

Conforme mostrado nas subseções anteriores, os algoritmos propostos neste Capítulo reduzem o custo computacional e a demanda por alocação de memória, ao custo de uma pequena degradação na imagem HR. Nessa subseção estamos interessados em testar se é possível melhorar a imagem HR gerada, através do uso de mais imagens LR de entrada, e se ainda assim os algoritmos propostos baseados em DFT 2D permanecem competitivos em termos de desempenho.

Tabela 4.1: Avaliação das imagens de alta resolução reconstruídas considerando-se PSNR (dB) e o tempo (s) de convergência do algoritmo

Núm.	Método		Vanderwalle		Phan		Farsiu		Closed Form		INLA		DFT		DFT	
	Exec.	MCMC	et. all		et. all		et. all		SR		SR		Closed Form		INLA	
	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)
1	28,02	648	24,75	13,24	26,76	59,09	19,44	20,03	30,25	3,02	30,45	60,54	29,93	0,17	29,14	1,61
2			24,79	13,41	27,16	59,54	19,77	19,76	30,19	3,02	30,26	55,66	29,36	0,18	28,92	1,61
3			24,68	13,43	27,14	59,09	18,23	19,39	30,19	3,02	30,27	53,04	29,68	0,17	29,08	1,56
4			24,64	13,36	26,43	60,01	18,26	18,87	30,23	3,03	30,31	52,60	29,71	0,17	29,11	1,61
Média	28,02	648	24,72	13,36	26,87	59,43	18,93	19,51	30,22	3,02	30,32	54,71	29,67	0,17	29,06	1,60

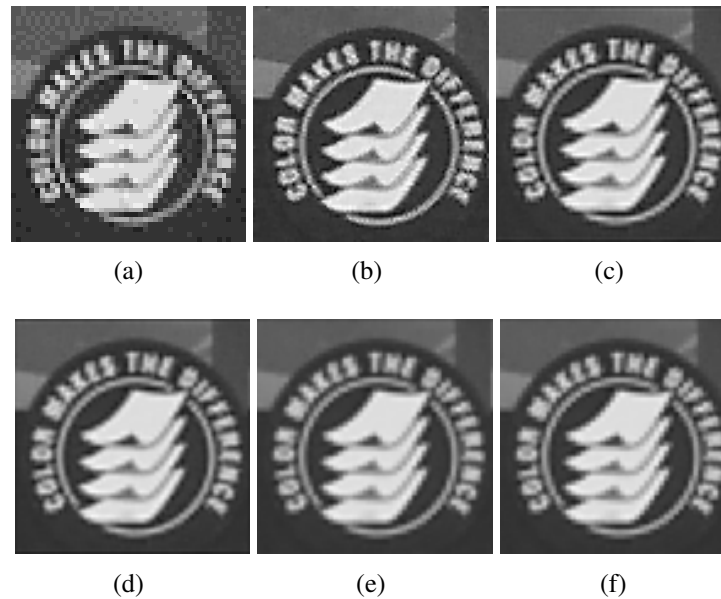


Figura 4.2: O conjunto de imagens reconstruídas usando-se uma sequência de 4 imagens reais. (a) Uma das quatro imagens LR. Imagens HR reconstruídas usando-se o método: (b) em (Farsiu et al., 2004b), (c) Closed Form SR (Camponez et al., 2011b), (d) INLA SR (Camponez et al., 2011a; Camponez et al., 2012c), (e) DFT Closed Form SR (Camponez et al., 2012b; Camponez et al., 2012a) e (f) DFT INLA SR (Camponez et al., 2012b; Camponez et al., 2012a).

Nesse experimento dois métodos são comparados: o Closed Form SR, publicado em (Camponez et al., 2011b), e o DFT Closed Form SR algorithm, publicado em (Camponez

et al., 2012b; Camponez et al., 2012a). Para tanto, as mesmas imagens usadas no primeiro experimento deste Capítulo são usadas mais uma vez. A Figura 4.3a corresponde à imagem HR e a Figura 4.3b corresponde a uma das imagens LR, de tamanho 128×128 .

Dois experimentos com objetivo de recuperar a imagem HR são executados para a imagem do Barco. No primeiro, quatro imagens LR foram utilizadas como entrada dos algoritmos Closed Form SR (Camponez et al., 2011b), e para o método DFT Closed Form SR (Camponez et al., 2012b; Camponez et al., 2012a), ambos implementados com $\lambda = 0,001$. No segundo, treze imagens LR foram utilizadas como entrada dos mesmos dois algoritmos.

Os resultados para o algoritmo Closed Form SR (Camponez et al., 2011b) são mostrados na Figura 4.3e e na Figura 4.3c, considerando-se quatro e treze imagens LR, respectivamente. Por outro lado, o resultado na Figura 4.3f corresponde ao método DFT Closed Form SR, considerando-se quatro imagens LR de entrada, e o resultado na Figura 4.3d corresponde ao DFT Closed Form SR, mais uma vez, agora considerando-se treze imagens LR de entrada. Tais resultados são também disponibilizados na Tabela 4.3, onde a qualidade da imagem HR recuperada é comparada, usando-se a métrica PSNR, assim como o tempo (em segundos) necessário para a convergência dos algoritmos. Pode ser visto, a partir de tais resultados, que é possível compensar a degradação na imagem HR causada pelo uso dos algoritmos de recuperação de imagens HR baseados em DFT 2D, propostos neste Capítulo, aumentando-se o número de imagens LR de entrada, e, ainda assim, tais métodos se mantêm competitivos em termos de desempenho.

Tabela 4.2: Avaliação de imagens de alta resolução reconstruídas considerando-se o tempo (s) de convergência do algoritmo

Farsiu	Closed Form	INLA	DFT Closed	DFT
SR	SR	SR	Form SR	INLA SR
(s)	(s)	(s)	(s)	(s)
4,52	0,50	1,31	0,09	0,10

4.4 Conclusões do Capítulo

Este Capítulo propõe uma modificação no modelo matemático dos algoritmos Closed Form SR (Camponez et al., 2011b) e INLA SR (Camponez et al., 2012c; Camponez et al., 2011a) propostos nesta Tese. A imagem HR é envolvida em um grafo toroidal \mathcal{T}_n e uma

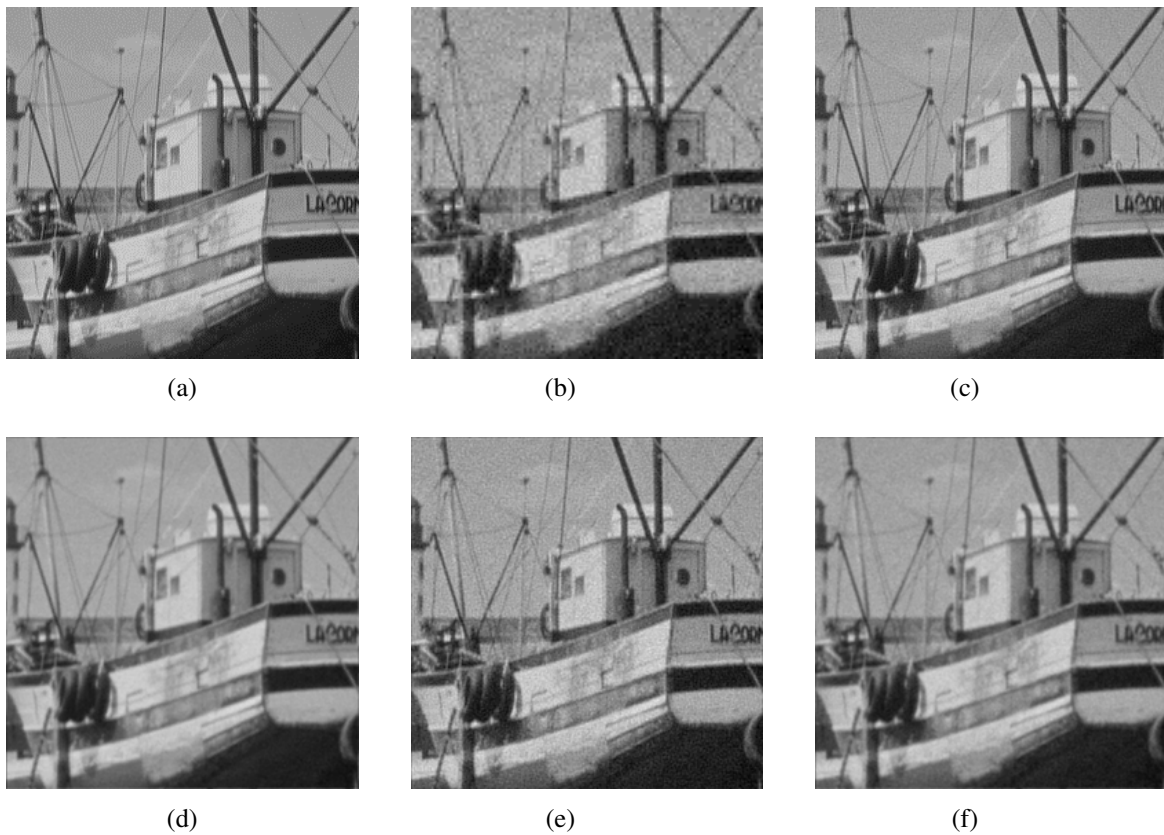


Figura 4.3: O conjunto de imagens reconstruídas usando-se a imagem de teste Barco. (a) A imagem original (*ground truth*), (b) uma das quatro imagens LR. Imagens HR reconstruídas usando-se o método: (c) Closed Form SR (13 entradas LR), (d) DFT Closed Form SR (13 entradas LR), (e) Closed Form SR (4 entradas LR) e (f) DFT Closed Form SR (4 entradas LR).

Tabela 4.3: Avaliação de imagens de alta resolução reconstruídas considerando-se PSNR (dB), o tempo (s) de convergência do algoritmo e o número de imagens de baixa resolução de entrada.

Algoritmo	Número de Entradas LR	Tempo (s)	PSNR (dB)
Closed Form SR	04	3,02	30,22
Closed Form SR	13	3,87	32,26
DFT Closed Form SR	04	0,17	29,67
DFT Closed Form SR	13	0,22	30,54

restrição (em (2.50)) é relaxada, o que possibilita aproximar a matriz \mathbf{A} de forma a torná-la Bloco Toeplitz Circulante - BTC. A vantagem de tal abordagem é que algoritmos para matrizes BTC podem ser computados usando-se a transformada discreta de Fourier 2D (2D DFT), como demonstrado neste Capítulo, o que reduz grandemente o custo computacional. Mesmo com o relaxamento da equação (2.50) as imagens geradas ainda apresentam bons resultados quando comparadas a outros algoritmos do estado da arte. Isso acontece pois a aproximação usada não altera a regra de interpolação das imagens LR no grid de alta resolução, que corresponde a matriz \mathbf{B} da equação (2.49). Somente o filtro inverso regularizado, que corresponde à matriz \mathbf{A} é alterado com o relaxamento proposto.

Os resultados experimentais demonstram que os algoritmos propostos neste Capítulo, DFT Closed Form SR e o DFT INLA SR, são muito mais rápidos que as versões originais Closed Form SR e INLA SR, embora apresentem uma pequena queda na qualidade da imagem HR. Além disso, apresentam altos valores de PSNR, quando comparados a outros algoritmos do estado da arte. O ganho em termos de tempo de execução associado a tais algoritmos é uma consequência da redução do custo computacional de $O(n^3)$, que é o custo computacional aproximado para inversão de uma matriz $n^2 \times n^2$ usando-se decomposição de Cholesky (que é usada nos métodos originais em (Camponez et al., 2011a), (Camponez et al., 2011b) e (Camponez et al., 2012c)), para $O(n^2 \log(n))$, que é o custo associado ao cálculo da 2D-DFT, usando-se a Transformada Rápida de Fourier (FFT). Além do mais, a dimensão das matrizes manipuladas pelo algoritmo é reduzida de $n^2 \times n^2$ para $n \times n$, que é o tamanho da imagem HR.

O novo método aumenta consideravelmente o desempenho dos métodos INLA SR e Closed Form SR, ao custo de uma pequena redução na qualidade da imagem HR. Como demonstrado no terceiro experimento, o algoritmo DFT Closed Form SR, quando usando 13 imagens LR como entrada, é ainda aproximadamente quatorze vezes mais rápido que o método Closed Form SR usando quatro imagens LR de entrada. Então, é possível usar mais imagens de entrada para compensar a perda de qualidade dos algoritmos baseados em DFT 2D, se necessário, ainda obtendo uma magnífica redução do esforço computacional.

Capítulo 5

Conclusões e Proposta de Continuação

Esse Capítulo traz um sumário das contribuições desta Tese no campo da reconstrução de imagens de alta resolução, por Superresolução, a partir de um conjunto de imagens de baixa resolução. Além disso, vários problemas ainda em aberto relacionados a esta Tese são descritos, possibilitando um mapeamento de possíveis trabalhos futuros.

5.1 Contribuições

Nesta Tese um conjunto de cinco algoritmos, aplicáveis ao problema da recuperação de imagens por superresolução, a partir de um conjunto de imagens de baixa resolução, foi desenvolvido.

Primeiramente, uma equação em forma fechada, isto é, uma solução analítica, para o problema de Superresolução foi desenvolvida a partir do modelo de observação adotado, e explorada no algoritmo Closed Form SR. Embora tal algoritmo tenha gerado imagens de alta qualidade, o método ainda depende de um parâmetro λ heurísticamente definido.

Então, aplicou-se um método de inferência Bayesiano aproximado, não paramétrico, conhecido na literatura estatística como *Integrated Nested Laplace Approximation* (INLA), até então não utilizado na área de processamento de imagens, ao problema. Isso possibilitou o desenvolvimento de dois algoritmos: INLA MAP (que corresponde a uma versão totalmente automática do algoritmo Closed Form SR) e o INLA SR (uma proposta nova onde a imagem de alta-resolução é calculada a partir de uma média ponderada). Experimentos demonstraram que esses dois últimos algoritmos geram imagens de alta qualidade, porém com um elevado custo computacional e uma alta demanda por memória.

Finalmente, envolvendo a imagem de alta resolução em um grafo Toroidal \mathcal{T}_n e fazendo uma aproximação Bloco Toeplitz Circulante na matriz de precisão \mathbf{Q} , do modelo a priori da imagem de alta resolução, outros dois algoritmos foram propostos: DFT Closed Form SR e DFT INLA SR. Tais algoritmos conseguem uma redução magnífica do custo computacional e da demanda por alocação de memória, ao preço de uma pequena redução da qualidade da imagem de alta resolução.

Nas próximas Subseções essas contribuições serão detalhadas.

5.1.1 Closed Form SR

No Capítulo 2, primeiro foi definido o modelo de aquisição de imagens de baixa resolução a partir da imagem de alta resolução. Em seguida, uma abordagem de inferência Bayesiana baseada em um modelo hierárquico é desenvolvida.

Nesse modelo, a PDF a priori da imagem foi definida como um campo aleatório Gaussiano Markoviano (GMRF), assumindo a forma proposta em (Rue et al., 2009) que é ligeiramente diferente, por exemplo, dos modelos usados em (Tian e Ma, 2010), (Hardie et al., 1997) e (Schultz e Stevenson, 1996). Foi justamente essa escolha para a PDF a priori que permitiu um tratamento matemático mais adequado, possibilitando o desenvolvimento de uma solução analítica para o problema de recuperação de imagens HR por Superresolução a partir de uma sequência de imagens LR, que é a primeira contribuição desta Tese. Essa equação foi explorada em um algoritmo, chamado nesta Tese de Closed Form SR.

A solução analítica desenvolvida envolve a inversão de uma matriz, chamada de \mathbf{A} na modelagem matemática. Demonstrou-se matematicamente que essa matriz é positiva definida e, portanto, sua inversão, que corresponde ao maior custo computacional do método, pode ser realizada através de uma decomposição de Cholesky. Assim, foi possível melhorar consideravelmente o desempenho do método.

Experimentos demonstraram que o algoritmo proposto é muito competitivo, em relação a outros métodos do estado da arte, quando comparado objetivamente usando-se PSNR como métrica, e também quando analisado subjetivamente através da comparação das imagens geradas.

Os resultados alcançados foram apresentados em Setembro de 2011 no *7th International Symposium on Visual Computing*, em Las Vegas, Nevada, EUA (Camponez et al., 2011b).

Embora o desempenho do algoritmo Closed Form SR dependa da escolha heurística de um parâmetro λ , que até então não era feita de forma automática pelo algoritmo, os experimentos mostraram que valores de lambda entre $\lambda = 0,001$ e $\lambda = 0,005$ funcionam bem para

todos os casos testados e que não há uma diferença significativa em PSNR nas imagens HR geradas quando o parâmetro está dentro dessa faixa.

5.1.2 INLA SR e INLA MAP SR

No Capítulo 3, um segundo algoritmo, chamado de INLA SR, que é a segunda contribuição desta Tese, foi desenvolvido. Aplicou-se ao problema de Superresolução um método de inferência estatística Bayesiana recente (Rue et al., 2009), não paramétrico, conhecido como *Integrated Nested Laplace Approximation* - INLA, baseado em integração numérica, até então nunca utilizado na área de processamento de imagens. Além disso, demonstrou-se que todas as equações necessárias para implementação de tal método, de uma maneira totalmente automática, podem ser escritas em forma fechada.

No método INLA SR a imagem de alta resolução é gerada por uma média ponderada baseada em dois vetores. O primeiro contém um conjunto de realizações de imagens de alta resolução, onde cada realização é gerada pela equação em forma fechada (2.50) desenvolvida anteriormente no método Closed Form SR. O segundo contém a probabilidade associada a cada uma dessas realizações, probabilidades estas que são calculadas através de uma equação em forma fechada, também desenvolvida nesta Tese.

Um terceiro algoritmo, INLA MAP, que é a terceira contribuição desta Tese, foi derivado do INLA SR, e corresponde a uma versão totalmente automática do primeiro algoritmo, Closed Form SR. O INLA MAP corresponde a escolher como imagem de alta resolução a realização que tiver maior probabilidade associada.

Experimentos demonstraram que os métodos propostos superam outros algoritmos do estado da arte em Superresolução. Esses resultados foram apresentados no *IEEE International Symposium on Signal Processing and Information Technology*, em Dezembro de 2011, em Bilbao, Espanha (Camponez et al., 2011a).

Um artigo com o desenvolvimento dos três algoritmos, experimentos e comparações com outros métodos do estado da arte foi publicado na *IEEE Transactions on Image Processing* (Camponez et al., 2012c).

Tais experimentos mostram que os métodos propostos conseguem recuperar imagens HR de excelente qualidade. Entretanto, as matrizes manipuladas por esses algoritmos são grandes, embora com grau de esparsidade muito alto. Isso causa dois problemas:

1. o consumo na alocação de memória é alto, mesmo usando-se estruturas de dados apropriadas a matrizes esparsas;

2. o custo computacional para inversão da matriz \mathbf{A} na equação (2.49) é muito alto, mesmo usando-se um algoritmo eficiente como a decomposição de Cholesky.

5.1.3 DFT Closed Form DR e DFT INLA SR

No Capítulo 4, o alto custo computacional associado aos algoritmos anteriormente propostos e a alta demanda para alocação de memória são tratados. Relaxando-se uma equação no modelo matemático original e envolvendo a imagem HR em um grafo Toroidal \mathcal{T}_n , mostra-se que o custo computacional dos algoritmos de Superresolução Closed Form SR, INLA SR e INLA MAP pode ser significativamente reduzido, bem como as dimensões das matrizes manipuladas (de $n^2 \times n^2$ para $n \times n$, o tamanho das imagens HR), ao custo de uma pequena redução da qualidade da imagem HR. Essa nova proposta é explorada em dois novos algoritmos baseados na Transformada Discreta de Fourier 2D (2D DFT): DFT Closed Form SR e o DFT INLA SR, que são outras duas contribuições desta Tese.

O algoritmo DFT Closed Form corresponde a uma versão extremamente rápida e com pouco consumo de memória do Closed Form SR. O algoritmo DFT INLA SR é uma versão mais rápida do algoritmo INLA SR. Essas contribuições foram apresentadas em Outubro de 2012 no *International Conference on Image Processing - ICIP 2012* (Camponez et al., 2012b), Orlando, Florida, EUA.

Posteriormente, o autor mostra que é possível compensar a perda de qualidade na imagem HR gerada pelos algoritmos baseados em DFT-2D aumentando-se a quantidade de imagens LR na entrada, e, ainda assim, manter um desempenho muito melhor que outros algoritmos do estado da arte, mantendo um baixo consumo de memória. Esse resultado foi apresentado em Setembro de 2012 no XIX Congresso Brasileiro de Automática - CBA (Camponez et al., 2012a), Campina Grande, Paraíba, Brasil.

Uma artigo com o desenvolvimento destes dois últimos algoritmos, novos experimentos e comparações com outros métodos do estado da arte foi submetido para publicação na *IEEE Transactions on Image Processing* (Camponez et al., 2012d).

Assim, com a melhoria do desempenho e a redução do consumo de memória, os métodos propostos nesta Tese conseguem unir qualidades que são muito difíceis de serem alcançadas em algoritmos de processamento de imagens:

- geração de imagens de alta qualidade;
- uso racional de memória;
- baixo custo computacional.

5.2 Propostas Para Trabalhos Futuros

Os trabalhos desenvolvidos nesta Tese abriram novas questões para futuras investigações, que serão brevemente descritas nessa seção:

- Em geral, Superresolução é um problema inverso e mal posto, isto é, não tem solução; ou tem infinitas soluções; ou a solução é instável em relação à entrada (Burger, 2003; Hansen, 1998; Park et al., 2003; Borman e Stevenson, 1998). Para resolver tal problema e obter uma solução única e estável, os algoritmos de Superresolução usam informações adicionais sobre a imagem. Nos algoritmos estatísticos Bayesianos, como os aqui desenvolvidos, a informação adicional é definida na forma de uma distribuição a priori (Winkler, 2006; Mohammad-Djafari, 1998). Nos algoritmos determinísticos no domínio do tempo a informação adicional é tratada como uma restrição (Bazaraa et al., 2006; Katsaggelos, 1989) ou como uma penalidade de regularização (Burger, 2003; Gibson e Bovik, 2000; Barrett et al., 2005). Na literatura de problemas inversos, diversos métodos têm sido desenvolvidos para encontrar de forma automática o coeficiente de regularização, como o Generalized Cross-Validation (GCV) (Burger, 2003) e os métodos baseados na Curva-L (Hansen, 1998; Reginska, 1996; Belge et al., 2002). Um estudo mais aprofundado da equação (3.51), Seção 3.2.5 do Capítulo 3, pode suscitar o desenvolvimento de um novo método de regularização, onde se possa aproveitar o formato em Bloco das matrizes \mathbf{Q} e \mathbf{A} , que, conforme descrito no Capítulo 4, podem se tornar Bloco Toeplitz Circulante. Assim, pode-se buscar uma forma matemática para essa equação baseada em DFT-2D, que funcionaria como um método rápido e automático para determinação de coeficientes de regularização.
- No Capítulo 4, Seção 4.2.3, é feita uma proposta para aproximar a matriz \mathbf{A} de forma a torná-la Bloco Toeplitz Circulante (Gray, 2006; Rue e Held, 2005). Nessa proposta as matrizes que modelam a decimação e o movimento são aproximadas. Um estudo complementar pode ser feito, na perspectiva de se encontrar outras possíveis aproximações para essa matriz, que conduzam a imagens HR melhores. Uma primeira possibilidade é buscar uma aproximação para a matriz \mathbf{Q} , deixando inalteradas as outras matrizes.
- Adaptar os modelos desenvolvidos para o uso com imagens coloridas. Há alguns trabalhos discutindo o tema de Superresolução com imagens coloridas (Farsiu et al., 2004a). A solução mais frequentemente aplicada é separar a imagem em 3 canais e aplicar um algoritmo monocromático a cada canal independentemente (Tom e Katsaggelos, 2001). Outra possibilidade é transferir o problema para um outro espaço de cores onde os canais de crominância são separados do de luminância, e onde a Superresolução somente é aplicada ao canal de Luminância (Irani e Peleg, 1991).

- Os algoritmos desenvolvidos nessa Tese pertencem ao grupo de métodos classificados na literatura como tradicionais, que são aqueles que aplicam as técnicas para a obtenção de uma única imagem HR. Quando é necessário gerar mais que uma imagem HR, como no caso do aumento da resolução de um vídeo, esse procedimento é repetido diversas vezes. Esses métodos não aproveitam os resultados previamente obtidos na Superresolução de uma imagem HR para obter a imagem subsequente (Park et al., 2003; Chaudhuri, 2001; Borman e Stevenson, 1998). Por outro lado, os métodos conhecidos como Sequenciais são capazes de aplicar a Superresolução nas imagens da seqüência, uma a uma, utilizando os resultados previamente obtidos (Elad e Feuer, 1999; Mateos et al., 2000). Finalmente, os métodos chamados de Simultâneos têm a capacidade de estimar toda a seqüência de imagens em um único processo (Zibetti e Mayer, 2005; Borman et al., 1999). À luz dessas três categorias de métodos, um estudo pode ser realizado vislumbrando o aproveitamento das técnicas desenvolvidas nesta Tese na formulação de métodos Sequenciais e/ou Simultâneos.
- A aplicação dos algoritmos de Superresolução em sistemas onde as imagens são comprimidas também é necessária (Park et al., 2003; Chaudhuri, 2001; Chen e Schultz, 1998; Altunbasak e Patti, 2000; Martins e Forchhammer, 2002; Segall et al., 2002; Park et al., 2002; Gunturk et al., 2002), uma vez que é muito comum que imagens ou vídeos sejam comprimidos antes da transmissão e/ou do armazenamento. Nesse caso, os algoritmos de Superresolução precisam adaptar seus modelos para trabalharem com tal paradigma. Por exemplo, é importante analisar e modelar o erro de compressão causado pela quantização. Nesses casos, um modelo de erro Gaussiano muitas vezes não é aceitável, especialmente quando uma taxa de compressão alta é aplicada. Dessa forma, um estudo aprofundado do desempenho dos algoritmos propostos em sistemas com compressão pode ser realizado, buscando adequar as técnicas desenvolvidas nesta Tese para serem aplicadas a imagens comprimidas.

5.3 Conclusão

Nesta Tese foram desenvolvidas novas técnicas e análises que possibilitaram o desenvolvimento de vários algoritmos eficientes, que podem ser aplicados a uma enorme gama de aplicações na fusão de imagens de alta resolução a partir de um conjunto de imagens de baixa resolução. Experimentos envolvendo imagens simuladas e, também, imagens reais demonstram a eficiência dos métodos propostos e o potencial que estes têm para aplicações reais. Espera-se, pois, que este trabalho motive outros estudantes e pesquisadores nessa desafiadora e excitante linha de pesquisa.

Referências Bibliográficas

- Altunbasak, Y. e Patti, A. J. (2000). A maximum a posteriori estimator for high resolution video reconstruction from mpeg video. Em *Image Processing (ICIP), 2000 IEEE International Conference on*.
- Barrett, H. H., Myers, K. J., e Dhurjaty, S. (2005). Foundations of image science. *J. Electronic Imaging*, 14(2):029901.
- Bazaraa, M. S., Sherali, H. D., e Shetty, C. M. (2006). *Nonlinear Programming: Theory And Algorithms*. Wiley-Interscience.
- Belge, M., Kilmer, M. E., Eric, e Miller, E. L. (2002). Efficient determination of multiple regularization parameters in a generalized l-curve framework. *Inverse Problems*, 18:2002.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York, NY, USA.
- Borman, S., , Borman, S., e Stevenson, R. L. (1999). Simultaneous multi-frame map super-resolution video enhancement using spatio-temporal priors. Em *In Proceedings of the IEEE International Conference on Image Processing*, páginas 469–473.
- Borman, S. e Stevenson, R. (1998). Spatial Resolution Enhancement of Low-Resolution Image Sequences - A Comprehensive Review with Directions for Future Research.
- Borman, S. e Stevenson, R. (1999). Super-resolution from image sequences-a review. Em *Proceedings of the 1998 Midwest Symposium on Circuits and Systems*, páginas 374–378.
- Bose, N. K., Kim, H. C., e Valenzuela, H. M. (1993). Recursive total least squares algorithm for image reconstruction from noisy, undersampled frames. *Multidimensional Systems and Signal Processing*, 4(3):253–268.

- Burger, M. (2003). Review and description of "computational methods for inverse problems", by curtis r. vogel. siam, philadelphia, pa, 2002. *Math. Comput.*, 72(243):1574–1575.
- Camponez, M., Salles, E. O. T., e Sarcinelli, M. (2012a). Dft-2d-based fast superresoluition. Em *CBA2012*, Campina Grande, PB.
- Camponez, M. O., Salles, E., e Sarcinelli-Filho, M. (2011a). Applying integrated nested laplace approximation to the superresolution problem. Em *Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on*, páginas 246–252.
- Camponez, M. O., Salles, E., e Sarcinelli-Filho, M. (2011b). A closed form algorithm for superresolution. Em Bebis, G., Boyle, R. D., Parvin, B., Koracin, D., Wang, S., Kim, K., Benes, B., Moreland, K., Borst, C. W., DiVerdi, S., Chiang, Y.-J., e Ming, J., editors, *ISVC (2)*, volume 6939 of *Lecture Notes in Computer Science*, páginas 338–347. Springer.
- Camponez, M. O., Salles, E., e Sarcinelli-Filho, M. (2012b). Dft-based fast superresolution image reconstruction using inla approximation. Em *Image Processing (ICIP), 2012 IEEE International Conference on*.
- Camponez, M. O., Salles, E. O. T., e Sarcinelli-Filho, M. (2012c). Super-resolution image reconstruction using nonparametric bayesian inla approximation. *Image Processing, IEEE Transactions on*, 21(8):3491–3501.
- Camponez, M. O., Sarcinelli-Filho, M., e Salles, E. (2012d). Dft-based fast superresolution image reconstruction using non-parametric bayesian inla approximation. *Image Processing, IEEE Transactions on*, PP(99):1.
- Chaudhuri, S. (2001). *Super-resolution Imaging*. Springer.
- Chen, D. e Schultz, R. R. (1998). Extraction of high-resolution video stills from mpeg image sequences. Em *Image Processing (ICIP), 1998 IEEE International Conference on*, páginas 465–469.
- Elad, M. e Feuer, A. (1999). Super-resolution reconstruction of image sequences. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(9):817–834.
- Eren, P., Sezan, M., e Tekalp, A. (1997). Robust, object-based high-resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing*, 6(10):1446–1451.

- Farsiu, S., Farsiu, S., Farsiu, S., Robinson, D., Robinson, D., Elad, M., Elad, M., Milanfar, P., e Milanfar, P. (2004a). Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, 14:47–57.
- Farsiu, S., Robinson, M. D., Elad, M., e Milanfar, P. (2004b). Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344.
- Figueiredo, M. A. T. e Nowak, R. D. (2001). Wavelet-based image estimation: an empirical bayes approach using jeffrey’s noninformative prior. *IEEE Transactions on Image Processing*, 10:1322–1331.
- Galatsanos, N. P. e Katsaggelos, A. K. (1992). Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *Image Processing, IEEE Transactions on*, 1(3):322–336.
- Galatsanos, N. P., Mesarovic, V. Z., Molina, R., e Katsaggelos, A. K. (2000). Hierarchical bayesian image restoration from partially-known blurs. *IEEE Transactions on Image Processing*, 9:1784–1797.
- Gibson, J. D. e Bovik, A., editors (2000). *Handbook of Image and Video Processing*. Academic Press, Inc., Orlando, FL, USA, 1st edição.
- Gray, R. M. (2006). *Toeplitz and Circulant Matrices: A review*. Stanford University.
- Gunturk, B. K., Altunbasak, Y., Member, S., e Mersereau, R. M. (2002). Multiframe resolution-enhancement methods for compressed video. *IEEE Signal Processing Lett*, 9:170–174.
- Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Hardie, R., Barnard, K., e Armstrong, E. (1997). Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633.
- Hilbe, R. J., michel Marin, J., e Robert, C. P. (2007). *Bayesian Core: A Practical Approach to Bayesian Computational Statistics*. Springer Texts in Statistics. Springer, New York.
- Hong, M., Kang, M., e Katsaggelos, A. (1997). Regularized multichannel restoration approach for globally optimal high-resolution video sequence. Em *Proceedings of SPIE*, volume 3024, páginas 1306–1316.

- Huynh-Thu, Q. e Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44(13):800–801.
- Irani, M. e Peleg, S. (1991). Improving resolution by image registration. *CVGIP: Graph. Models Image Process.*, 53(3):231–239.
- Irani, M. e Peleg, S. (1993). Motion analysis for image enhancement: Resolution, occlusion, and transparency. *Journal of Visual Communication and Image Representation*, 4(4):1993–12.
- Izenman, A. J. (2008). *Modern Multivariate Statistical Techniques : Regression, Classification, and Manifold Learning*. Springer Texts in Statistics. Springer, New York.
- Johnson, R. A. e Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis*. Prentice Hall, Inc., Upper Saddle River, NJ, USA, 5 edição.
- Katsaggelos, A., Molina, R., e Mateos, J. (2007). *Super Resolution of Images and Video*. Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool.
- Katsaggelos, A. K. (1989). Iterative image restoration algorithms. *Optical Engineering*, 28:735–748.
- Kim, S. P., Bose, N. K., e Valenzuela, H. M. (1990). Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38:1013–1027.
- Kim, S. P. e Su, W.-Y. (1993). Recursive high-resolution reconstruction of blurred multi-frame images. *IEEE Transactions on Image Processing*, 2:534–539.
- Komatsu, T., Igarashi, T., Aizawa, K., e Saito, T. (1993). Very high resolution imaging scheme with multiple different-aperture cameras. *Signal Processing: Image Communication*, 5(5-6):511–526.
- Li, S. Z. (1995). *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, New York, NY, USA.
- MacKay, D. J. C. (1999). Learning in graphical models. chapter Introduction to Monte Carlo methods, páginas 175–204. MIT Press, Cambridge, MA, USA.
- Martins, B. e Forchhammer, S. (2002). A unified approach to restoration, deinterlacing and resolution enhancement in decoding mpeg-2 video. *IEEE Trans. Circuits Syst. Video Techn.*, 12(9):803–811.

- Mateos, J., Mateos, J., e Molina, R. (2000). Resolution enhancement of compressed low resolution video. Em *Acoustics, Speech, and Signal Processing, 2000. ICASSP 00. Proceedings. 2000 IEEE International Conference on*.
- Mohammad-Djafari, A. (1998). From deterministic to probabilistic approaches to solve inverse problems. Em *Proc. SPIE 3459, Bayesian Inference for Inverse Problems*, páginas 2–11.
- Park, S. C., Kang, M. G., Segall, C. A., e Katsaggelos, A. K. (2002). Spatially adaptive high-resolution image reconstruction of low-resolution dct-based compressed images. Em *Image Processing (ICIP), 2002 IEEE International Conference on*, volume II, páginas 861–864, Rochester, New York.
- Park, S. C., Park, M. K., e Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36.
- Patti, A., Sezan, M., e Murat Tekalp, A. (1997). Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Transactions on Image Processing*, 6(8):1064–1076.
- Pham, T., Van Vliet, L., e Schutte, K. (2006). Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Applied Signal Processing*, 2006:1–12.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., e Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3 edição.
- Reginska, T. (1996). A regularization parameter in discrete ill-posed problems. *SIAM J. Sci. Comput.*, 17(3):740–749.
- Rue, H. e Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*, volume 104 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, UK.
- Rue, H., Martino, S., e Chopin, N. (2009). Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.
- Schultz, R. e Stevenson, R. (1996). Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6):996–1011.
- Segall, C. A., Molina, R., Katsaggelos, A. K., e Mateos, J. (2002). Reconstruction of high-resolution image frames from a sequence of low-resolution and compressed observations. Em *ICASSP*, páginas III–1701–1704, Orlando, Florida.

- Tekalp, A. M., Ozkan, M. K., e Sezan, M. I. (1992). High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. Em *Acoustics, Speech, and Signal Processing, 1992. ICASSP 92. Proceedings. 1992 IEEE International Conference on*.
- Tian, J. e Ma, K. (2010). Stochastic super-resolution image reconstruction. *Journal of Visual Communication and Image Representation*, 21(3):232–244.
- Tipping, M. e Bishop, C. (2003). Bayesian image super-resolution. *Advances in Neural Information Processing Systems*, páginas 1303–1310.
- Tom, B. e Katsaggelos, A. (1996). An iterative algorithm for improving the resolution of video sequences. Em *Proceedings of SPIE, the International Society for Optical Engineering*, volume 2727, páginas 1430–1438. Society of Photo-Optical Instrumentation Engineers.
- Tom, B. C. e Katsaggelos, A. K. (2001). Resolution enhancement of monochrome and color video using motion compensation. *IEEE Trans. Image Processing*, 10:278–287.
- Tsai, R. Y. e Huang, T. S. (1984). Multiframe image restoration and registration. Em Huang, T. S., editor, *Advances in Computer Vision and Image Processing*, volume 1, páginas 317–339. JAI Press, Greenwich, CT, USA.
- Vandewalle, P., Süsstrunk, S., e Vetterli, M. (2006). A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP Journal on Applied Signal Processing*, 2006:1–14.
- Wang, Z. e Qi, F. (2004). On ambiguities in super-resolution modeling. *SPLetters*, 11(8):678–681.
- Winkler, G. (2006). *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (Stochastic Modelling and Applied Probability)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Zibetti, M. V. W. (2007). Super-resolução simultânea para sequência de imagens. Em *Tese de Doutorado*, Santa Catarina, BR. UFSC.
- Zibetti, M. V. W. e Mayer, J. (2005). Simultaneous super-resolution for video sequences. Em *Image Processing (ICIP), 2005 IEEE International Conference on*, páginas 877–880.

Apêndice A

Publicações Feitas até o Momento

Os resultados apresentados nessa Tese foram publicados em:

1. *Superresolution Image Reconstruction Using Non-parametric Bayesian INLA Approximation*, publicado na *IEEE Transactions on Image Processing* e disponível para acesso na base de dados do IEEE, desde 01/05/2012 (Camponez et al., 2012c);
2. *DFT-2D-Based Fast Superresoluiton*, publicado nos *proceedings* do 19th Congresso Brasileiro de Automação, Campina Grande, Paraíba, Brasil, 2012 (Camponez et al., 2012a);
3. *DFT-Based Fast Superresolution Image Reconstruction Using INLA Approximation*, publicado nos *proceedings* do *19th IEEE International Conference on Image Processing*, Orlando, Florida, EUA, 2012 (Camponez et al., 2012b);
4. *Applying Integrated Nested Laplace Approximation to the Superresolution Problem*, publicado nos *proceedings* do *11th IEEE International Symposium on Signal Processing and Information Technology*, Bilbao, Espanha, 2011 (Camponez et al., 2011a);
5. *A Closed Form Algorithm for Superresolution*, publicado nos *proceedings* do *7th International Symposium on Visual Computing*, Las Vegas, Nevada, EUA, 2011 (Camponez et al., 2011b).