

UNIVERSIDADE FEDERAL DE SÃO CARLOS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**Uso do Algoritmo ICM Adaptativo a
Descontinuidades para o Aumento da
Resolução de Imagens Digitais por Técnicas de
Reconstrução por Super Resolução**

Ana Luísa Dine Martins

São Carlos – SP
Maió/2007

UNIVERSIDADE FEDERAL DE SÃO CARLOS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**Uso do Algoritmo ICM Adaptativo a Descontinuidades para o
Aumento da Resolução de Imagens Digitais por Técnicas de
Reconstrução por Super Resolução**

Ana Luísa Dine Martins

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos como parte dos requisitos necessários à obtenção de título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Nelson Delfino D'Ávila Mascarenhas

Co-orientador: Prof. Dr. Paulo Estevão Cruvinel

São Carlos – SP

Mai/2007

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

M386ua

Martins, Ana Luísa Dine.

Uso do algoritmo ICM adaptativo a descontinuidades para o aumento da resolução de imagens digitais por técnicas de reconstrução por super resolução / Ana Luísa Dine Martins. -- São Carlos : UFSCar, 2007.
76 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2007.

1. Processamento de imagens. 2. Restauração de imagens. 3. Campos aleatórios de Markov. 4. Reconstrução por super resolução. I. Título.

CDD: 006.42 (20^a)

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

“Uso do Algoritmo ICM Adaptativo a Descontinuidades para o Aumento da Resolução de Imagens Digitais por Técnicas de Reconstrução por Super Resolução”

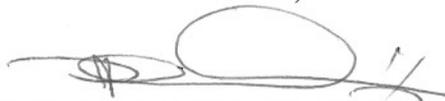
ANA LUISA DINE MARTINS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

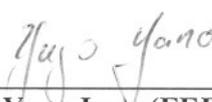
Membros da Banca:



Prof. Dr. Nelson Delfino d'Ávila Mascarenhas
(Orientador – DC/UFSCar)



Prof. Dr. Paulo Estevão Cruvinel
(Co-Orientador - EMBRAPA/Instrumentação Agropecuária)



Prof. Dr. Yuzo Iano (FEEC/UNICAMP)

São Carlos
Maio/2007

A Walter, Virginia, Raphael, Orlanda e Antonio, meu alicerce.

Agradecimentos

Primeiramente e acima de tudo, agradeço à Deus. Aos meus pais, avós e irmão, pessoas sem as quais eu nunca teria conseguido. Ao Otávio, por ser praticamente uma fonte de energia nos longos dias de trabalho. Ao Murillo, por toda a paciência, ajuda e dedicação. À Aretha, pelo companheirismo e amizade nas horas mais difíceis. Ao meu orientador e mestre Nelson Delfino D'Ávila Mascarenhas, por toda dedicação e atenção. Ao meu co-orientador Paulo Estevão Cruvinel, pelo incentivo desde o princípio. À minha madrinha Marília, minha tia Lourdes, minha prima Luciana, meus tios Janio e Clarice e meus primos Nathália e Rodrigo. Aos meus companheiros de laboratório: Alexandre, Moacir, Fernando, João, Gustavo, Marcelo, Marcos, Gabriel, Michelle, João Paulo, Denis, Yuji e Talita. Aos meus amigos: Powerful, Cidinho, Cris e Escova, Deodato, Célia, Alê, Camila Fran e Léo, seu Silmar, Robson, Peixe, Paulinho, Fauze e todos os outros que esqueci no momento mas que são não menos importantes.

Agradeço também a todos os funcionários e professores do DC, e a todas as pessoas que de alguma maneira contribuíram para a realização deste trabalho.

Agradeço à CAPES pelo auxílio financeiro.

Sumário

Lista de Figuras	v
Lista de Tabelas	vii
Resumo	viii
Abstract	ix
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	4
1.3 Objetivo	4
1.4 Organização	5
2 Reconstrução por Super Resolução	6
2.1 Considerações Iniciais	6
2.2 Resolução Espacial da Imagem	7
2.3 Reconstrução por Super Resolução	9
2.4 Modelo de Formação das Imagens	11
2.5 Abordagens e Aplicações Existentes	13
2.5.1 Trabalhos Pioneiros	14
2.5.2 Outras Abordagens	16
2.5.3 Utilização de Informações Contextuais	18
2.6 Considerações Finais	20
3 Campos Aleatórios de Markov e o Modelo de Potts-Strauss	21
3.1 Considerações Iniciais	21
3.2 Campos Aleatórios de Markov	22
3.3 Distribuição de Gibbs	24
3.4 Modelo de Potts-Strauss	26
3.4.1 Estimação por Máxima Pseudoverossimilhança	28
3.5 Adaptação a Descontinuidades	30

3.6	Considerações Finais	34
4	Inferência Estatística e o Algoritmo ICM	35
4.1	Considerações Iniciais	35
4.2	Teoria de Decisão Bayesiana	36
4.3	O Estimador MAP	38
4.4	O Algoritmo ICM	41
4.5	Considerações Finais	42
5	Método Proposto	43
5.1	Considerações Iniciais	43
5.2	Modelo de Formação das Imagens	44
5.3	Registro das Imagens	46
5.4	Formulação Bayesiana do Problema	50
5.5	Modelo de Potts-Strauss	50
5.6	O Algoritmo ICM Adaptativo a Descontinuidades	51
5.7	Considerações Finais	55
6	Resultados	56
6.1	Considerações Iniciais	56
6.2	Descrição da Simulação	56
6.3	Análise dos Resultados	59
6.4	Estudo de Caso	65
6.5	Considerações Finais	66
7	Conclusão e Trabalhos Futuros	67
7.1	Considerações Finais	67
7.2	Contribuições	68
7.3	Trabalhos Futuros	69
7.3.1	Uma Ferramenta para Super Resolução	69
7.3.2	Projeto HERON	70
	Referências Bibliográficas	72

Lista de Figuras

2.1	Imagem “Cameraman”	8
2.2	Ilustração a respeito da relação entre a área das células fotoelétricas e a densidade de células no sensor.	8
2.3	Demonstração da presença de deslocamentos <i>sub-pixel</i> entre as imagens de baixa resolução observadas.	10
2.4	Ilustração da abordagem clássica da Reconstrução por Super Resolução.	11
2.5	Modelo de formação das imagens de baixa resolução a partir da cena original (PARK et al., 2003).	12
3.1	Sistemas de vizinhança de (a) primeira ordem e de (b) segunda ordem.	24
3.2	Sistemas de vizinhança de primeira a sexta ordem.	24
3.3	Cliques referentes aos sistemas de vizinhança de primeira ordem – de (a) a (c) – e de segunda ordem – de (a) a (j).	25
3.4	Abordagem utilizada por DERIN e ELLIOT (1987), associando um parâmetro a cada tipo de clique	26
3.5	Formas qualitativas das funções de interação adaptativa mostradas na Tabela (3.1).	34
3.6	Detecção de descontinuidades na imagem Lena utilizando uma função de interação adaptativa	34
4.1	Possível paralelismo do algoritmo ICM.	42
5.1	Processo de subamostragem provocando os deslocamentos <i>sub-pixel</i> entre as imagens (HOMEM et al., 2006).	45
5.2	Imagem de alta resolução (HOMEM et al., 2006).	46
5.3	Imagens resultantes da sub-amostragem da imagem de alta resolução mostrada na Figura (5.2) (HOMEM et al., 2006).	46
5.4	Estimação dos deslocamentos <i>sub-pixel</i> entre as imagens de baixa resolução, dada uma imagem de referência (HOMEM et al., 2006).	48
5.5	Ilustração do Registro aplicado às imagens de baixa resolução, permitindo que elas sejam alinhadas e que uma primeira estimativa de alta resolução seja adquirida (HOMEM et al., 2006).	48

5.6	Ilustração utilizando a imagem Lena, do Registro aplicado às imagens de baixa resolução.	49
5.7	Situação em que existem deslocamentos equidistantes e uniformemente espaçados entre as imagens de baixa resolução	49
5.8	Situação em que os deslocamentos presentes entre as imagens de baixa resolução estão dispostos de forma irregular	49
5.9	Forma qualitativa da função que detecta as descontinuidades mostrada na Equação (5.13).	54
5.10	Ilustração do cálculo da probabilidade <i>a posteriori</i> local, utilizando a informação da função que detecta as descontinuidades.	54
6.1	Imagem “Lena” de dimensão 512x512, usada para gerar as várias imagens de baixa resolução.	57
6.2	Ilustração mostrando o processo de sub-amostragem da imagem de alta resolução, dando origem a 16 imagens de baixa resolução.	58
6.3	Imagens de baixa resolução de dimensão 128x128, geradas utilizando a imagem “Lena” borrada e ruidosa.	58
6.4	Imagem de dimensão 512x512, resultante do registro e interpolação das 16 imagens de baixa resolução simuladas.	59
6.5	Imagem resultante da aplicação do algoritmo ICM puro, sem a imposição das descontinuidades.	60
6.6	ICM adaptativo a descontinuidade aplicado à imagem Lena registrada.	60
6.7	Comparação do ICM adaptativo a descontinuidade com a Interpolação Bilinear aplicada a uma das imagens de baixa resolução.	61
6.8	Recorte da imagem resultante da décima iteração do algoritmo proposto por IRANI e PELEG (1991).	63
6.9	<i>Frames</i> , de dimensão 128x128 <i>pixels</i> , utilizados para avaliar a aplicação do método proposto em uma situação real.	65
6.10	Imagens resultantes da décima iteração do algoritmo proposto por Irani e Peleg, e do algoritmo ICM adaptativo a descontinuidades.	66
7.1	Imagens de ressonância magnética de um corte transversal do trato vocal humano, cedidas pelo Instituto de Engenharia Eletrônica e Telemática de Aveiro.	71
7.2	Imagens de ressonância magnética de um corte oblíquo do trato vocal humano, cedidas pelo Instituto de Engenharia Eletrônica e Telemática de Aveiro.	71

Lista de Tabelas

3.1	Quatro possibilidades de funções de interação adaptativa (LI, 1995b).	33
6.1	Índice de qualidade <i>Universal Image Quality Index</i> aplicado às imagens com relação à imagem original.	63
6.2	Índice de qualidade <i>Universal Image Quality Index</i> aplicado às imagens com relação à imagem borrada.	64
6.3	NMSE calculado para as imagens com relação à imagem original.	64
6.4	NMSE calculado para as imagens com relação à imagem borrada.	65

Resumo

A Reconstrução por Super Resolução consiste em, utilizando várias imagens de baixa resolução da mesma cena, gerar uma aproximação da cena original, que possua resolução espacial mais alta que a presente em qualquer uma das imagens observadas. Para isso, tais imagens devem possuir algum tipo de deslocamento da ordem *sub-pixel* uma em relação às demais, de forma que não exista apenas a mesma informação replicada em todas as imagens. Assim, a incerteza inerente a tais deslocamentos pode ser usada como informação adicional no aumento de resolução.

Nesse contexto, esta pesquisa propõe uma abordagem Bayesiana do problema, utilizando Campos Aleatórios de Markov e o Modelo de Potts-Strauss na caracterização das imagens. Isso torna possível a imposição de informações espaciais de contexto conhecidas *a priori* da imagem de alta resolução a ser estimada. A estimativa de Máximo *a Posteriori* (MAP) de alta resolução é encontrada por meio do algoritmo *Iterated Conditional Modes* (ICM) adaptativo a descontinuidades.

Dessa forma, utilizando como estimativa inicial de alta resolução a imagem resultante do registro e interpolação das imagens de baixa resolução observadas, foi possível reconstruir imagens de maior resolução que respeitassem as descontinuidades inicialmente presentes, e que apresentassem maior riqueza de detalhes.

Abstract

Super resolution image reconstruction consists in using a set of low resolution images from the same scene to generate a high resolution estimate of the original scene. For that purpose, all the observed low resolution images need to have sub-pixel displacements among each other. In this way, there is more than just the same information replicated in each image and then the uncertainty inherent to the displacements can be used as additional information to increase the spatial resolution. This master's thesis proposes a Bayesian approach for the super resolution reconstruction problem using Markov Random Fields and the Potts-Straus model for the image characterization. Therefore, it is possible to incorporate previously known context spatial information about the high resolution image to be estimated. Moreover, a discontinuity adaptive ICM algorithm was used to estimate the *maximum a posteriori* solution. Using an initial high resolution estimate constructed from the registration and interpolation of all the observations made it possible to reconstruct an image that respected the initially presented discontinuities. We also observed that the resulted high resolution image hold finer details when compared to the initial estimation.

Introdução

1.1 Contextualização

A Reconstrução por Super Resolução, também chamada na literatura de Realce de Resolução (*Resolution Enhancement*), consiste em, utilizando várias imagens de baixa resolução da mesma cena, gerar uma aproximação da cena original, que possua resolução espacial mais alta que a presente em qualquer uma das imagens observadas. As imagens de baixa resolução utilizadas possuem algum tipo de transformação geométrica (translações, rotações, transformações perspectivas, mudança de escala, etc.), uma em relação às demais.

Os deslocamentos presentes nas imagens observadas devem ser da ordem *sub-pixel*, de forma que existam informações novas em cada uma das imagens, e não apenas a mesma informação replicada em todas elas. Se os deslocamentos forem inteiros, a Reconstrução por Super Resolução se resumirá a uma simples interpolação, não recuperando informações das altas frequências, ou seja, os detalhes perdidos no processo de aquisição das imagens.

Imagens possuindo tais características podem ser adquiridas por: múltiplas câmeras; sistemas de imageamento que possuam controle de deslocamento; movimento do objeto de interesse; sistemas de imageamento vibratórios; na utilização de *frames* de um vídeo; dentre outras maneiras.

Adquiridas tais imagens, o aumento de resolução é conseguido com o processamento dessas. Sendo assim, os meios de aquisição já utilizados não precisam ser substituídos por meios

mais potentes e mais caros, o que faz das técnicas de Reconstrução por Super Resolução uma solução bastante interessante para aplicações que necessitam de maior detalhamento nas imagens utilizadas.

Desde o trabalho pioneiro de TSAI e HUANG (1984), surgiram várias abordagens do problema da Reconstrução por Super Resolução. A abordagem clássica divide o processo em três etapas bem definidas: Registro, Interpolação e Restauração (PARK et al., 2003). O registro é aplicado para se identificar os deslocamentos presentes nas imagens de baixa resolução. A interpolação visa o aumento da resolução, com base nos deslocamentos identificados. Por fim, a restauração atenua as degradações inseridas nas imagens durante o processo de aquisição.

Existem propostas que tratam o processo no domínio do espaço ou da frequência, de forma direta ou iterativa, sendo que as três etapas podem ocorrer em ordem diferente da adotada pela abordagem clássica, podendo inclusive se fundir em apenas um passo ou se desdobrar em mais passos.

De forma generalizada, a reconstrução da imagem de alta resolução é um problema mal-condicionado, já que a solução pode não existir ou podem existir várias soluções. Para que o problema se torne bem condicionado, se faz necessária a imposição de algum tipo de conhecimento *a priori* a respeito da solução. Na maior parte das abordagens, isso é feito assumindo a hipótese de suavidade da imagem (TIKHONOV e ARSENIN, 1977).

Para se trabalhar com a imposição da hipótese de suavidade da imagem, as abordagens estatísticas, em comparação com abordagens determinísticas, oferecem maior facilidade para se encontrar um modelo coerente para o processo de formação das imagens de baixa resolução e para descrever a imagem de alta resolução que se quer estimar. Dessa forma, decidindo-se por uma abordagem estatística, freqüentemente adota-se a modelagem Bayesiana do problema, sendo que a imposição da suavidade é feita pela distribuição de probabilidade conhecida *a priori* da imagem de alta resolução.

Entretanto, na maioria dos casos não se possui a distribuição de probabilidade da imagem a ser estimada. Segundo DUBES e JAIN (1993), os modelos de imagens que consideram informações de contexto não precisam ser altamente precisos. Por essa razão, os Campos Aleatórios de Markov (MRF - *Markov Random Fields*) vêm sendo cada vez mais utilizados. No problema da Reconstrução por Super Resolução, MRFs permitem a imposição de informações espaciais de contexto conhecidas *a priori*, a respeito da imagem de alta resolução a ser reconstruída.

Quando se está utilizando MRFs na caracterização das imagens, a imposição da característica *a priori* de suavidade na imagem pode provocar uma super suavização, desrespeitando as discontinuidades inicialmente presentes nessa imagem. LI (1995b) comenta que abordagens adaptativas a discontinuidades (DA - *Discontinuity Adaptive*), principalmente quando se está

utilizando MRFs, são bastante adotadas desde a introdução do conceito *Line Process* pelo trabalho de GEMAN e GEMAN (1984). Por meio de uma análise da equação de Euler associada à minimização da energia em MRFs e em modelos de regularização, o autor mostrou que as peculiaridades entre as diferentes abordagens estão apenas na maneira como ocorre a interação entre pontos vizinhos, segundo o sistema de vizinhança do modelo adotado.

Nesse contexto, é preciso trabalhar duas etapas na estimação da imagem de alta resolução: a detecção das descontinuidades e como proceder quando tais descontinuidades são detectadas.

De forma geral, a detecção das descontinuidades da imagem deve ocorrer de forma que as mudanças de intensidade correspondentes a ruído, não sejam erroneamente consideradas como descontinuidades. O método deve ser capaz de fazer a distinção entre ruído e descontinuidade. LI (1995b) apresenta algumas funções capazes que fazer tal distinção.

Os procedimentos para a estimação da imagem, quando se encontra uma descontinuidade, podem ser os mais diversos possíveis, sendo que, na grande maioria, as descontinuidades são impostas na função energia da distribuição de Gibbs que caracteriza o MRF. Um problema quando se utiliza a distribuição de Gibbs na caracterização do MRF, é que o cálculo da energia considera a imagem como um todo, o que pode acarretar em um maior custo computacional, já que tal energia é calculada praticamente a cada *pixel* da imagem.

Um caso específico da distribuição de Gibbs é o Modelo de Potts-Strauss. Ele pode ser adotado para a caracterização da imagem de alta resolução. Esse modelo permite que se trabalhe considerando apenas uma pequena localidade a cada *pixel*, e, dessa forma, o processo de estimação se torna mais rápido.

Mesmo utilizando o Modelo de Potts-Strauss na caracterização da imagem de alta resolução, para se encontrar a estimativa MAP, considerando as descontinuidades detectadas, uma busca exaustiva dentre todas as configurações possíveis do MRF é uma alternativa computacionalmente intratável devido à dimensão do conjunto de configurações possíveis que o campo aleatório pode assumir. Para solucionar esse problema, são utilizados métodos de otimização (KIRKPATRICK et al., 1988).

Entretanto, mesmo que métodos de otimização, como o clássico *Simulated Annealing* proposto por KIRKPATRICK et al. (1988), sejam utilizados, o tempo dispendido na busca ainda é um problema. Uma alternativa é o algoritmo determinístico proposto por BESAG (1986), denominado *Iterated Conditional Modes* (ICM), tem sido bastante utilizado na literatura como forma de se encontrar uma solução para problemas de otimização, contando com baixo custo computacional. Sua peculiaridade está na convergência para um máximo local. Apesar de não garantir a convergência para o máximo global, se bem inicializado, esse algoritmo garante bons resultados em poucas iterações.

1.2 Motivação

A resolução espacial de uma imagem digital está ligada ao número de *pixels* que essa imagem possui. Este é um dos principais fatores responsáveis pelo grau de detalhamento da imagem. Tal detalhamento é desejado, senão exigido, na grande maioria das aplicações que manipulam imagens digitais.

Apenas para exemplificar a necessidade de se encontrar soluções para o problema do aumento de resolução, observa-se a crescente demanda por vídeos de alta resolução e câmeras digitais capazes de adquirir imagens de resolução cada vez maior, além da digitalização do sistema de televisão que também têm exigido esforços nesse contexto.

Além disso, diagnósticos baseados na análise de imagens médicas, poderiam possuir maior grau de confiança se as imagens em questão possuíssem alto grau de detalhamento. Nesse contexto, um diagnóstico errôneo poderia, no pior caso, culminar na morte do paciente. Dessa forma, trata-se de uma aplicação crítica.

Outras duas aplicações, não tão críticas, mas em que se pode observar a necessidade de maior detalhamento nas imagens, são o reconhecimento de objetos em aplicações de visão computacional, e os métodos de reconhecimento de padrões aplicados a imagens de sensoriamento remoto. Tais aplicações possuiriam resultados bem mais precisos se os *frames* dos respectivos sistemas de visão possuíssem alta resolução.

Em todos os casos citados, maior resolução poderia ser adquirida aumentando a capacidade dos meios de aquisição das imagens digitais, porém em muitas aplicações existe um limite imposto pela tecnologia existente, e tal limite já foi praticamente atingido (TOM et al., 2001). Além disso, os meios de aquisição já utilizados teriam que ser descartados e substituídos por meios mais potentes e conseqüentemente mais caros.

Assim, a máxima resolução adquirida e o preço dos dispositivos de maior capacidade estimularam o aparecimento de abordagens que, utilizando técnicas de Processamento de Imagens, permitissem o aumento de resolução das imagens. Nesse contexto, a Reconstrução por Super Resolução se mostra uma boa alternativa.

1.3 Objetivo

O objetivo desse trabalho é propor uma abordagem de Reconstrução por Super Resolução que visa baixo custo computacional, por meio de soluções sub-ótimas. Isso é feito adotando uma abordagem estatística do problema, com caracterização das imagens por MRFs. O modelo de

Potts-Strauss é utilizado como alternativa à Distribuição de Gibbs, e para se encontrar a estimativa MAP da imagem de alta resolução, utiliza-se o algoritmo ICM adaptativo a descontinuidades.

Primeiramente o método é exemplificado por meio de uma situação controlada onde imagens de baixa resolução possuindo deslocamentos *sub-pixel* são geradas. Tais imagens são registradas e interpoladas, gerando uma primeira estimativa da imagem de alta resolução. Essa imagem, já de alta resolução em relação às demais, é utilizada como entrada para o algoritmo ICM adaptativo a descontinuidades.

Por fim, resultados utilizando *frames* de um vídeo, são gerados e analisados, comprovando a eficiência do método.

1.4 Organização

No Capítulo 2 é apresentada uma revisão bibliográfica a respeito da Reconstrução por Super Resolução, comentando as abordagens existentes na literatura, e as etapas inerentes do processo de estimação da imagem de alta resolução. O Capítulo 3 detalha e discute os Campos Aleatórios de Markov e o Modelo de Potts-Strauss, adotados na caracterização das imagens de alta resolução. Como adotou-se uma abordagem probabilística do problema, a Teoria de Decisão Bayesiana e o algoritmo ICM são detalhados no Capítulo 4. No Capítulo 5, o método proposto é descrito, assim como as abordagens e algoritmos utilizados. No Capítulo 6 são discutidos os resultados alcançados em uma situação controlada em que imagens possuindo deslocamentos *sub-pixel* uma em relação às demais foram geradas. No Capítulo 7 é feita uma discussão quanto às expectativas e resultados alcançados e, finalmente, são discutidas as contribuições do trabalho e possíveis trabalhos futuros.

Reconstrução por Super Resolução

2.1 Considerações Iniciais

Também chamada Realce de Resolução (*Resolution Enhancement*), a Reconstrução por Super Resolução consiste em, utilizando várias imagens de baixa resolução da mesma cena, que possuem algum tipo de transformação geométrica uma em relação às demais, gerar uma imagem de resolução mais alta que a presente em qualquer uma das imagens observadas.

Segundo KANG e CHAUDHURI (2003), o termo Super Resolução foi originalmente usado em ótica, se referindo a algoritmos que operam em uma única imagem a fim de extrapolar seu espectro de frequências além do limite de difração. Essa abordagem é chamada na literatura de Restauração por Super Resolução (*Super Resolution Restoration*) e, assim como a Reconstrução por Super Resolução, preocupa-se com a recuperação das informações das altas frequências, perdidas ou corrompidas durante o processo de aquisição das imagens.

RUBERT et al. (2004) afirmam que o termo Super Resolução também é utilizado para denominar algoritmos baseados em aprendizagem, que também trabalham com o aumento de resolução espacial. Detalhes não existentes na imagem observada são extraídos de um conjunto de imagens de treinamento (FREEMAN et al., 2002), (HERTZMANN et al., 2001). Apesar dessa abordagem possuir como objetivo o aumento de resolução, assim como a Reconstrução por Super Resolução, ela é aplicável somente quando o conjunto de imagens de interesse possui características conhecidas que serão armazenadas durante o treinamento do método.

Visto que o objetivo final da Reconstrução por Super Resolução é o aumento de resolução espacial, é interessante que se defina Resolução Espacial de uma imagem digital.

2.2 Resolução Espacial da Imagem

Segundo GONZALEZ e WOODS (2001), resolução espacial de uma imagem digital é seu menor detalhe discernível. Ainda segundo os autores a amostragem durante o processo de aquisição da imagem é o principal fator que determina sua resolução espacial, pois é nessa etapa que a densidade de *pixels* da imagem é determinada.

A aquisição da imagem é feita captando a energia luminosa refletida ou transmitida pelas superfícies dos objetos. Desde a década de 70, dispositivos CCD (*Charge-Coupled Device*) são amplamente utilizados na aquisição de imagens digitais. Um CCD consiste em células sensíveis a luz, normalmente dispostas na forma de uma matriz (PARK et al., 2003). Cada uma dessas células atribui uma quantidade numérica à intensidade de energia luminosa captada, processo denominado Quantização da imagem. A discretização espacial, amostrando a cena original nas coordenadas x e y , determina a densidade de *pixels* da imagem.

Dessa forma, a imagem digital pode ser definida como uma função bidimensional discreta $f(x, y)$, composta por um conjunto finito de elementos, chamado *suporte* da imagem. Cada elemento dessa função – denominado *pixel* – possui uma localização no suporte da imagem (coordenadas x e y).

Quando CCDs estão sendo utilizados na aquisição, a dimensão da matriz de sensores determina os limites da amostragem. Supondo que tal matriz possui $M \times N$ células fotoelétricas, a imagem $f(x, y)$ poderia ser representada pela matriz dada pela Equação (2.1).

$$f(x, y) = \begin{pmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N - 1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M - 1, 0) & f(M - 1, 1) & \dots & f(M - 1, N - 1) \end{pmatrix} \quad (2.1)$$

Sendo a resolução espacial o número de *pixels* por unidade de área, ela pode ser relacionada à dimensão da matriz $f(x, y)$. Mantendo as devidas proporções, essa dimensão é fator determinante da qualidade ou nível de detalhamento da imagem. Quanto maior a resolução e maior o número de *pixels* por unidade de área, maior a dimensão de $f(x, y)$, e maior o detalhamento na imagem, o que pode ser observado na Figura (2.1).

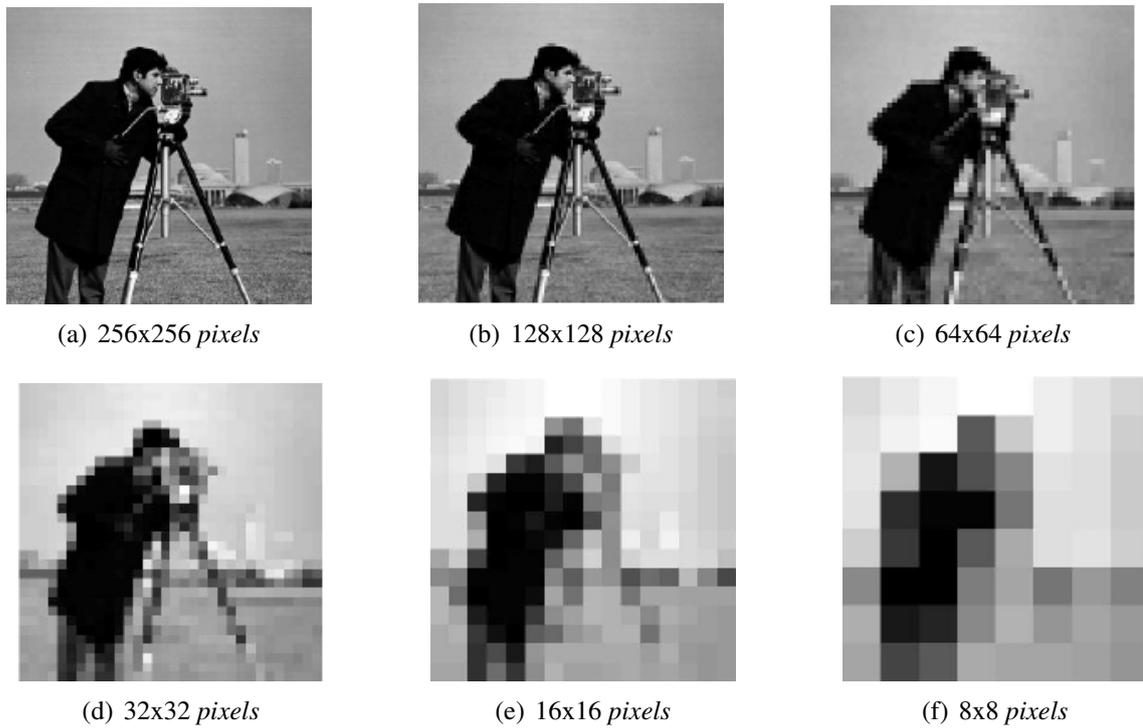


Figura 2.1: Imagem “Cameraman”

Na grande maioria das aplicações baseadas em imagens digitais, alto grau de detalhamento é desejado, senão exigido. Assim, poder-se-ia aumentar a capacidade dos meios de aquisição, de forma que as imagens possuíssem maior resolução espacial.

Em sistemas de aquisição de imagens baseados em CCDs, aumenta-se a resolução das imagens reduzindo a área de cada célula fotoelétrica, aumentando, assim, a densidade de células na matriz de CCDs, como pode ser visto na Figura (2.2). Porém, existe um limite imposto pela tecnologia existente.

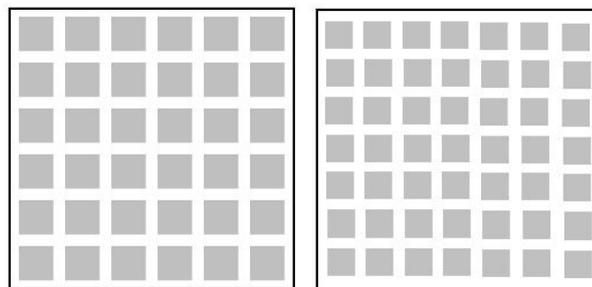


Figura 2.2: Ilustração a respeito da relação entre a área das células fotoelétricas e a densidade de células no sensor.

A presença de *shot noise* nas imagens adquiridas, característica inerente a sistemas de aquisição baseados em CCDs, é a responsável por esse limite (TOM et al., 2001). Ao reduzir a área de cada CCD, o número de *photons* que atingem o CCD diminui proporcionalmente a sua área. Porém, o número de *photons* correspondente ao ruído praticamente não se altera, o que faz com que a proporção entre *photons* relativos ao ruído e aqueles relativos ao sinal, se torne inaceitável.

Segundo TOM et al. (2001), o limite no tamanho de cada CCD é aproximadamente $50\mu m^2$, e a tecnologia CCD atual está quase atingindo esse limite.

Além das limitações impostas pela tecnologia existente, aumentar a capacidade dos meios de aquisição significa substituir os meios sendo utilizados por meios mais potentes. Em algumas situações essa substituição pode ser financeiramente inviável, já que meios mais potentes são invariavelmente mais caros.

Outra situação em que a substituição dos meios de aquisição não é interessante configura-se quando o imageamento ocorre em condições longe de serem ideais – por exemplo em sensoria-mento remoto. Nesses casos, as imagens podem ser adquiridas à noite, sob neblina, e/ou com o céu nublado. O investimento para substituir os meios de aquisição em tais situações seria inútil, já que não resolveria o problema das irregularidades presentes na imagem.

A máxima resolução adquirida, aliada à inviabilidade de investimento financeiro para a compra de meios de aquisição mais potentes, estimularam o aparecimento de métodos que, utilizando técnicas de Processamento de Imagens, permitissem o aumento de resolução das imagens já adquiridas. A Reconstrução por Super Resolução tem sido uma abordagem bastante pesquisada, justamente por propor uma solução para esse problema.

2.3 Reconstrução por Super Resolução

Segundo TOM et al. (2001), para que seja possível obter uma imagem de alta resolução a partir de várias imagens de baixa resolução da mesma cena, é preciso que as imagens observadas estejam sub-amostradas (*aliasing*) e possuam deslocamentos *sub-pixel*, uma em relação às demais.

Imagens possuindo deslocamentos *sub-pixel* estão deslocadas por parcelas menores do que o tamanho do *pixel* de baixa resolução, como pode ser visto na Figura (2.3). Isso é necessário para que existam informações adicionais em cada uma das imagens observadas, já que com deslocamentos inteiros ter-se-ia apenas a mesma informação replicada em todas as imagens de entrada. Nesse caso, a Reconstrução por Super Resolução se resumiria a uma simples interpolação.

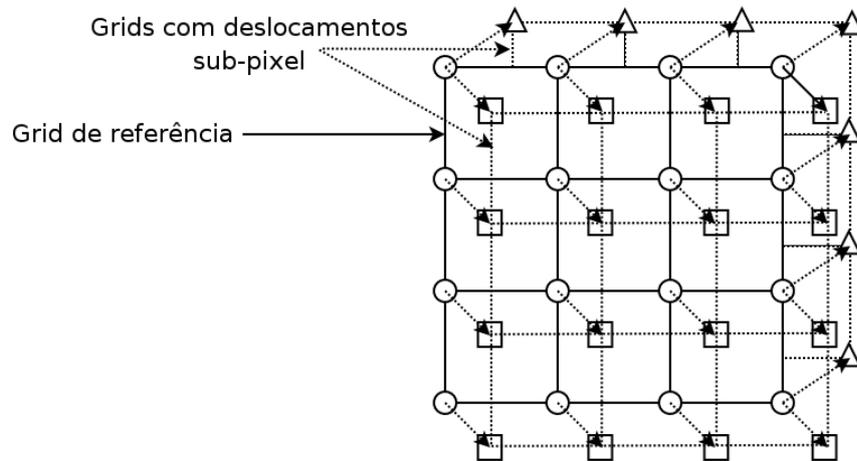


Figura 2.3: Demonstração da presença de deslocamentos *sub-pixel* entre as imagens de baixa resolução observadas.

Existem várias metodologias empregadas no processo de Reconstrução por Super Resolução. A abordagem mais intuitiva – considerada clássica – divide o processo em três tarefas bem definidas (PARK et al., 2003):

1. Primeiramente os deslocamentos *sub-pixel* são estimados, tomando uma das imagens observadas como referência. Isso é feito aplicando técnicas de Registro de Imagens nas imagens duas a duas, tomando sempre a imagem de referência como uma delas.
2. Estimados os deslocamentos, as imagens são alinhadas e encontra-se uma primeira estimativa de alta-resolução por meio de uma técnica de Interpolação.
3. Por fim, técnicas de Restauração de Imagens são utilizadas para atenuar as degradações invariavelmente inseridas durante o processo de aquisição e manipulação das imagens.

A abordagem clássica, ilustrada na Figura (2.4), se baseia no modelo de formação das imagens, mostrado na Figura (2.5). Esse modelo descreve o processo de formação das imagens de baixa resolução da seguinte forma:

- As imagens, ainda de alta resolução, podem receber borrimento devido às seguintes situações: câmera fora de foco, movimento relativo entre a câmera e o objeto, turbulência atmosférica etc.;
- Cada uma dessas imagens sofre algum tipo de transformação geométrica (*warping*), o que provoca os deslocamentos *sub-pixel*: translação, rotação, transformação perspectiva etc.;

- As imagens são sub-amostradas (decimação), dando origem às várias imagens de baixa resolução borradas e possuindo deslocamentos *sub-pixel*;
- Por fim, é inserido ruído nas imagens de baixa resolução, o que pode ocorrer devido à manipulação e armazenamento das imagens (na Figura (2.5) só está sendo considerado ruído aditivo).



Figura 2.4: Ilustração da abordagem clássica da Reconstrução por Super Resolução.

Assim, a abordagem clássica aplica o processo inverso ao modelo de formação das imagens mostrado na Figura (2.5). Nesse sentido, as degradações são modeladas matematicamente, a fim de encontrar uma aproximação da cena original aplicando o processo inverso.

Dessa forma, a modelagem matemática da formação das imagens de baixa resolução é definida a seguir.

2.4 Modelo de Formação das Imagens

De acordo com o modelo clássico de formação das imagens de baixa resolução (Figura (2.5)), a modelagem matemática da formação dessas imagens pode ser dada pela Equação (2.2), sendo f a cena original a ser estimada de dimensão $M_1 \times M_2$. Considerando que existem q imagens de baixa resolução, g_k é a k -ésima imagem de baixa resolução de dimensão $N_1 \times N_2$, $k = 1, 2, \dots, q$ ($M_1 \geq N_1$ e $M_2 \geq N_2$).

$$g_k = D_k W_k H_k f + r \quad (2.2)$$

Em notação lexicográfica, a imagem de alta resolução pode ser dada por $f = [f_1, f_2, \dots, f_M]$, $M = M_1 M_2$, e uma das imagens de baixa resolução por $g_k = [g_{k,1}, g_{k,2}, \dots, g_{k,N}]$, para $k = 1, 2, \dots, q$, onde $N = N_1 N_2$.

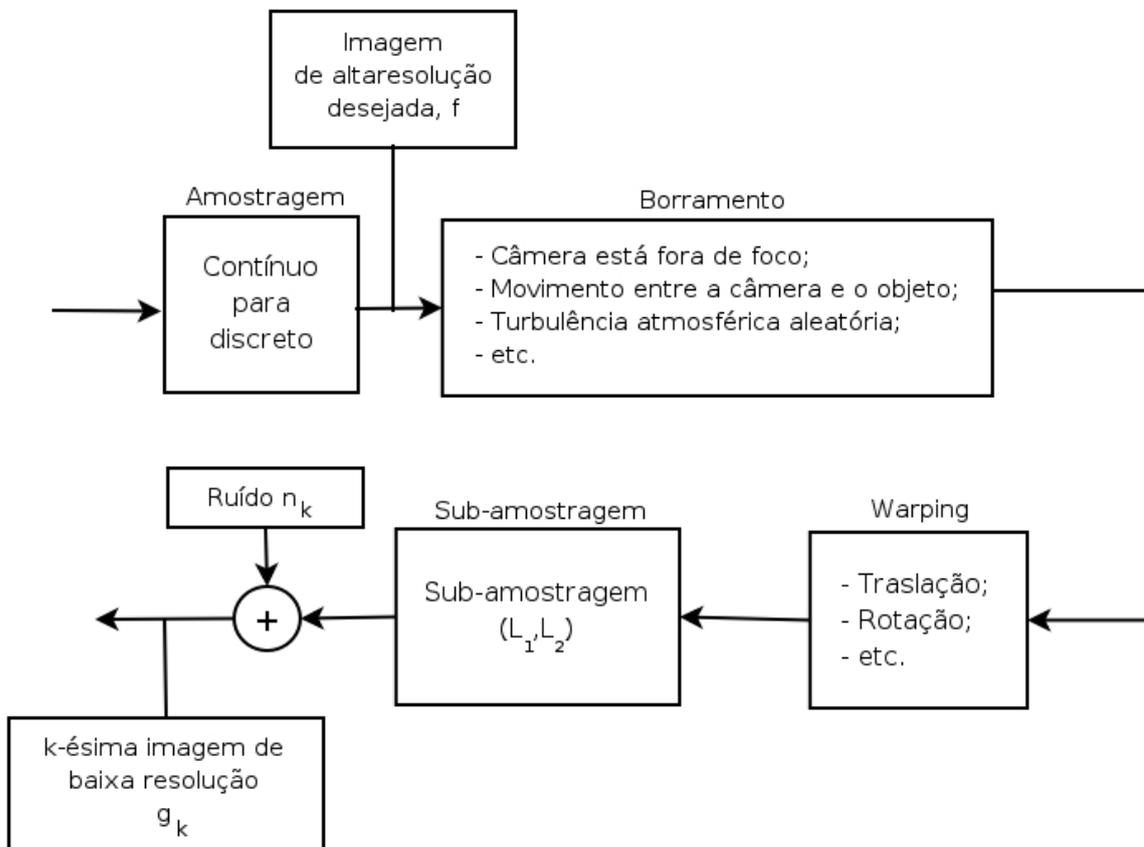


Figura 2.5: Modelo de formação das imagens de baixa resolução a partir da cena original (PARK et al., 2003).

O borramento é modelado pela Função de Espalhamento Pontual (PSF - *Point Spread Function*) do sistema de aquisição das imagens, representada pela matriz H_k na Equação (2.2), de dimensão $M_1 M_2 \times M_1 M_2$. Considerando que todas as q imagens de baixa resolução foram captadas pelo mesmo sistema de aquisição, H_k é idêntica para todo $k = 1, 2, \dots, q$, portanto podemos considerar $H_k = H$, qualquer que seja $k = 1, 2, \dots, q$.

A imagem borrada, ainda de alta resolução, sofre algum tipo de transformação geométrica (*warping*) diferente para cada uma das imagens de baixa resolução. É nesse momento que os deslocamentos *sub-pixel* são aplicados às imagens, e eles devem ser diferentes para cada uma delas, pois, caso contrário, teria-se apenas a mesma informação replicada em todas as imagens. Na Equação (2.2), os deslocamentos são modelados pela matriz W_k , de dimensão $M_1 M_2 \times M_1 M_2$, $k = 1, 2, \dots, q$.

Cada uma das q imagens, ainda de alta resolução, borradas e deslocadas, sofre o processo de sub-amostragem (*downsampling*) devido à resolução espacial máxima permitida pelo sensor. Nesse momento, são geradas as q imagens de baixa resolução.

Novamente considerando que todas as imagens foram captadas pelo mesmo sensor, tem-se que o operador de sub-amostragem é o mesmo para todas as q imagens observadas. Assim, $D_k = D$ para todo $k = 1, 2, \dots, q$, e a dimensão das imagens após a sub-amostragem é $N_1 \times N_2$, $N_1 \leq M_1$ e $N_2 \leq M_2$.

Assumindo que $M_1 = L_1 N_1$ e $M_2 = L_2 N_2$, tem-se que L_1 e L_2 são os fatores de sub-amostragem na horizontal e na vertical, respectivamente. Assim, o operador de sub-amostragem, representado por D , é responsável pela diminuição de resolução pelos fatores L_1 e L_2 na horizontal e na vertical, respectivamente.

Supondo que cada *pixel* tenha dimensão $T_1 \times T_2$, cada uma das imagens de baixa resolução estará deslocada em relação à imagem de referência por um valor proporcional a $T_1/L_1 \times T_2/L_2$. Isso garante que os deslocamentos não sejam inteiros, e sim *sub-pixel*. Se as imagens de baixa resolução estivessem deslocadas por valores proporcionais a $T_1 \times T_2$, os deslocamentos seriam inteiros, e o método de Reconstrução por Super Resolução ineficaz (CORTIJO et al., 2003).

Por fim, as imagens já de baixa resolução, borradas e que sofreram deslocamentos, sofrem a ação de ruído aditivo, inerente ao processo de manipulação e armazenamento das imagens. Na Equação (2.2), o ruído aditivo é representado pela adição do termo r . Esse modelo está considerando apenas a presença e ruído aditivo, porém ele é facilmente estendido para outras formas de ruído.

Dadas as considerações discutidas, a modelagem matemática do processo de formação das imagens de baixa resolução pode, portanto, ser dada pela Equação (2.3).

$$g_k = DW_k Hf + r \quad (2.3)$$

Apesar da modelagem clássica dividir o processo de Reconstrução por Super Resolução em três etapas bem definidas, desde o trabalho pioneiro de TSAI e HUANG (1984), as abordagens existentes tratam o processo no domínio do espaço ou da frequência, de forma direta ou iterativa, sendo que as três etapas podem ocorrer em ordem diferente da colocada, podendo se fundir em apenas um passo ou mesmo se desdobrar em mais passos.

2.5 Abordagens e Aplicações Existentes

A seguir é apresentada a abordagem pioneira de TSAI e HUANG (1984), além de trabalhos publicados mais recentemente.

2.5.1 Trabalhos Pioneiros

TSAI e HUANG (1984) publicaram o primeiro trabalho utilizando várias imagens da mesma cena, a fim de gerar uma imagem de resolução espacial mais alta do que qualquer uma das imagens observadas. Foram utilizadas imagens de um satélite Landsat adquiridas a cada 18 dias, sendo que essas invariavelmente possuíam alguns deslocamentos, uma em relação às demais.

Os autores utilizaram uma abordagem no domínio da frequência com base na propriedade de deslocamento no espaço da Transformada de Fourier. Essa propriedade afirma que o deslocamento de uma função no domínio do espaço é equivalente à multiplicação no domínio de Fourier, como mostra a Equação (2.4), sendo $f(x, y)$ a cena original contínua.

$$f(x + \delta_{k,1}, y + \delta_{k,2}) \longleftrightarrow F(u, v)e^{j2\pi(\delta_{k,1}u + \delta_{k,2}v)} \quad (2.4)$$

Considerou-se que existiam p frames disponíveis representados por $\{f_{ij}^k\}_{i=0, \dots, M-1, j=0, \dots, N-1}$, sendo f_{ij}^k a (i, j) -ésima amostra do k -ésimo frame. Existiam, portanto, um total de MN amostras em cada frame, representadas como mostra a Equação (2.5), sendo $\delta_{k,1}$ e $\delta_{k,2}$ os deslocamentos nas direções horizontal e vertical do k -ésimo frame com relação ao frame adotado como referência, e T_1 e T_2 os intervalos de amostragem na horizontal e na vertical.

$$f_{ij}^k = f(iT_1 + \delta_{k,1}, jT_2 + \delta_{k,2}) \quad (2.5)$$

Os autores afirmaram que o k -ésimo frame poderia ser visto como a amostragem da cena original, deslocada por $\delta_{k,1}$ na horizontal e $\delta_{k,2}$ na vertical, consideração que pode ser representada por $f_k(x, y)$ como mostra a Equação (2.6).

$$f_k(x, y) = f(x + \delta_{k,1}, y + \delta_{k,2}) \quad (2.6)$$

Nesse sentido, pela propriedade de deslocamento no tempo da Transformada de Fourier (Equação (2.4)), a Transformada Contínua de Fourier de $f_k(x, y)$, representada por $F_k(u, v)$, pode ser dada pela Equação (2.7), sendo $F(u, v)$ a Transformada Contínua de Fourier da cena original $f(x, y)$.

$$F_k(u, v) = e^{j2\pi(\delta_{k,1}u + \delta_{k,2}v)} F(u, v) \quad (2.7)$$

Com base no Teorema da Amostragem e na relação de *aliasing* entre a Transformada Discreta de Fourier e a Transformada Contínua de Fourier, a (m, n) -ésima componente de frequência do k -ésimo frame pode ser representada como mostra a Equação (2.8), sendo $w_{si} = T_i^{-1}$, $i = 1, 2$.

$$F_{mn}^k = \frac{1}{T_1 T_2} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} F_k \left(\frac{m}{MT_1} + mw_{s1}, \frac{n}{NT_2} + nw_{s2} \right) \quad (2.8)$$

Aplicando a Transformada Discreta de Fourier no k -ésimo *frame*, F_{mn}^k também pode ser representado como mostra a Equação (2.9).

$$F_{mn}^k = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f_{ij}^k \exp \left[-j2\pi \left(\frac{mi}{M} + \frac{nj}{N} \right) \right], m = 0, \dots, M-1, n = 0, \dots, N-1 \quad (2.9)$$

Para concluir seu raciocínio, os autores consideram um dado L_i , tal que $F(w_1, w_2) = 0$ para $|w_i| \geq L_i w_{si}$, $i = 1, 2$. Combinando as Equações (2.7) e (2.8), chega-se à Equação (2.10), sendo G_{mn} um vetor de dimensão $px1$ com k -ésimo elemento F_{mn}^k . F_{mn} é um vetor de dimensão $4L_1 L_2 \times 1$, cujo i -ésimo elemento é dado pela Equação (2.11), e ϕ_{mn} uma matriz de dimensão $px4L_1 L_2$, cujo (k, l) -ésimo elemento é dado pela Equação (2.12).

$$G_{mn} = \phi_{mn} F_{mn} \quad (2.10)$$

$$[F_{mn}]_i = F \left(\frac{m}{MT_1} + rw_{s1}, \frac{n}{NT_2} + sw_{s2} \right), r = \text{imod}(2L_1) - L_1, s = \lfloor L_i/2L_1 \rfloor - L_2 \quad (2.11)$$

$$[\phi_{mn}]_{kl} = \exp \left\{ j2\pi \left[\delta_{k1} \left(\frac{m}{MT_1} - \frac{r}{T_1} \right) + \delta_{k2} \left(\frac{n}{NT_2} - \frac{s}{T_2} \right) \right] \right\} \quad (2.12)$$

Resolvendo a Equação (2.10) para F_{mn} , para todo mn , encontram-se $4L_1 L_2 MN$ amostras de $F(u, v)$. Aplicando-se a Transformada Inversa de Fourier nessas amostras, a resolução é aumentada em $4L_1 L_2$.

Assim, TSAI e HUANG (1984) utilizaram a propriedade de deslocamento no espaço da Transformada de Fourier para, utilizando as várias imagens deslocadas, encontrar mais componentes de frequência, e conseqüentemente aumentar a resolução. Porém, os autores não consideraram ruído e borramento em sua abordagem. Trabalhos posteriores, como o de KIM et al. (1990) e o de KIM e SU (1993) estenderam o trabalho de TSAI e HUANG (1984) para o caso em que existia borramento e ruído nas imagens de baixa resolução, e para o caso em que existiam diferentes tipos de borramento em cada uma das imagens de baixa resolução.

2.5.2 Outras Abordagens

Após o trabalho pioneiro de TSAI e HUANG (1984), surgiram as mais diversas abordagens para resolver o problema de aumento de resolução quando são utilizadas várias imagens da mesma cena. Tais abordagens sempre colocam a contraposição entre métodos no domínio da frequência ou no domínio do espaço, determinísticos ou estatísticos, iterativos ou diretos, mostrando suas vantagens e desvantagens.

Uma abordagem bastante interessante, que trata o processo de forma iterativa foi proposta por STARK e OSKOUI (1989). Os autores utilizaram um algoritmo iterativo baseado em Projeções em Conjuntos Convexos (POCS - *Projection Onto Convex Sets*) na reconstrução da imagem de alta resolução. Esse algoritmo permite a imposição de dados conhecidos *a priori* no processo de reconstrução. Informações a respeito do meio de aquisição e do processo de formação das imagens, além de informações previamente conhecidas a respeito do conjunto de soluções, são impostas na forma de restrições modeladas como conjuntos convexos fechados. A convergência do método para a intersecção de todos esses conjuntos faz com que a estimativa encontrada obedeça todas as restrições impostas. Posteriormente, TEKALP et al. (1992) estenderam o problema para a presença de ruído, implementando também a estimação dos deslocamentos presentes entre as imagens.

IRANI e PELEG (1991) também utilizaram uma abordagem iterativa, mas, dessa vez, similar ao algoritmo de retro-projeção empregado na reconstrução de projeções para a formação de imagens tomográficas. Os autores modelaram cada *pixel* de baixa resolução como a projeção de uma dada região na imagem de alta resolução a ser estimada. O método considerou translações e rotações entre as imagens de baixa resolução, e alcançou resultados bastante satisfatórios. Basicamente, foi utilizada uma primeira estimativa da imagem de alta resolução para simular as imagens de baixa resolução, dado o modelo de formação dessas imagens. A estimativa de alta resolução é então atualizada com base no erro entre as imagens simuladas e observadas. O trabalho de IRANI e PELEG (1991) é considerado um marco no contexto da Reconstrução por Super Resolução.

KOMATSU et al. (1993), também por meio de uma solução iterativa, consideraram imagens adquiridas por múltiplas câmeras, com aberturas diferentes. Por abertura entende-se o diâmetro da abertura da lente, que controla a quantidade de luz que atinge o sensor. Os autores exploram as diferentes aberturas para conseguir melhores resultados na fusão das imagens. A reconstrução utilizou a Iteração da Landweber (BIEMOND et al., 1990).

TOM e KATSAGGELOS (1995) utilizaram o algoritmo EM (*Expectation Maximization*) para encontrar uma estimativa por máxima verossimilhança (ML - *Maximum Likelihood*). Os autores implementaram as três etapas da abordagem clássica em um único passo, afirmando

que o Registro, Interpolação e Restauração são interdependentes e por isso não deveriam ser resolvidos separadamente.

HARDIE et al. (1997) apresentaram um *framework* Bayesiano para estimar simultaneamente os parâmetros do registro dos *frames* de baixa resolução e da estimativa de alta resolução. Os parâmetros foram iterativamente atualizados por meio de um procedimento de otimização.

van VLIET e HENDRIKS (1999) discutiram um método de complexidade moderada, para aumento da resolução de imagens de infra-vermelho adquiridas por um sistema vibratório que proporcionou deslocamentos *sub-pixel* aleatórios. Os métodos de registro CZP (*Cross-correlation with Zero-Padding*), CPF (*Cross-correlation with phase fitting*) e MTS (*First Order Taylor Series*), foram testados a fim de se estimar os deslocamentos entre as imagens observadas, e seus resultados comparados. Em seguida, estimados os deslocamentos, os métodos de interpolação LPS (*Least Squares Plane Fitting*), NC (*Normalized Convolution*), ER (*Exact Reconstruction*) e IT (*Iterative Reconstruction*), foram aplicados na fusão das imagens registradas.

TOM et al. (2001), através de comparações entre métodos aplicados na estimação dos deslocamentos e entre métodos de fusão das imagens, procuraram encontrar combinações que chegassem a bons resultados, além de discriminar características que influenciavam cada etapa da reconstrução. Os autores trabalharam com imagens monocromáticas e coloridas, sendo que no caso das imagens coloridas os canais de cores foram explorados para aumentar a precisão das estimação dos deslocamentos.

ELAD e HEL-OR (2001) consideraram o caso em que as imagens de baixa resolução possuíam apenas deslocamentos translacionais, borramento invariante no espaço, idêntico para todas as imagens, e ruído branco. Os autores utilizaram um método direto, preservando a otimalidade da solução de Máxima Verossimilhança. Além disso, foram discutidas as vantagens do método quando comparado a outras abordagens existentes.

NGUYEN et al. (2001) aplicaram um método de gradiente conjugado para a Super Resolução regularizada. Foi utilizado o método de regularização de Tikhonov e o método Generalizado de Validação Cruzada (GVC - *Generalized Cross-Validation*) foi utilizado para calcular os parâmetros da regularização.

A aplicação mais comum da Reconstrução por Super Resolução é encontrar uma única imagem de alta resolução a partir de um conjunto de imagens de baixa resolução da mesma cena, entretanto alguns trabalhos como os de ALTUNBASAK e PATTI (2000), SEGALL et al. (2002), PARK et al. (2003), GUNTURK et al. (2002) e SEGALL et al. (2004), utilizaram a Reconstrução por Super Resolução na recuperação de *frames* de alta resolução a partir de vídeos comprimidos.

CORTIJO et al. (2003) encontraram uma solução de Máximo *a Posteriori* (MAP - *Maximum a Posteriori*), utilizando um método de autoregressão simultânea (SAR - *Simultaneous Autoregression*) para encontrar um modelo *a priori* da imagem original. Os autores consideraram borramento, ruído e deslocamentos presentes nas imagens de baixa resolução como previamente conhecidos.

Abordagens estatísticas como a adotada por CORTIJO et al. (2003) apresentam maior facilidade na modelagem do processo de formação das imagens de baixa resolução, além de permitir que se utilize informações conhecidas *a priori* no processo de reconstrução. Porém, na maioria dos casos não se possui um modelo da imagem a ser estimada.

Segundo a premissa de que, a menos que a imagem se resuma a ruído aleatório, *pixels* vizinhos possuem valores próximos, exceto no caso de informações de borda, pode-se utilizar Campos Aleatórios de Markov na representação das imagens. Se adotados no processo de Reconstrução por Super Resolução, eles são capazes de impor informações de contexto a respeito da imagem de alta resolução.

A seguir são apresentados alguns trabalhos que alcançaram resultados bastante interessantes no contexto da Reconstrução por Super Resolução ao utilizarem Campos Aleatórios de Markov na modelagem das imagens.

2.5.3 Utilização de Informações Contextuais

Os trabalhos pioneiros de BESAG (1974) e GEMAN e GEMAN (1984) impulsionaram o uso de Campos Aleatórios de Markov na modelagem de imagens digitais. No contexto da Reconstrução por Super Resolução, vários trabalhos encontraram nessa modelagem maior facilidade na imposição de informações e restrições conhecidas *a priori* sobre as imagens e seu processo de formação, chegando, assim, a melhores resultados.

BORMAN e STEVENSON (1998) afirmam que os Campos Aleatórios de Markov estão praticamente se tornando padrão na modelagem de imagens, por se tratarem de uma maneira intuitiva, realística e conveniente de se modelar características típicas de uma imagem. Essa afirmação é justificada por se tratar basicamente de um modelo de interação entre vizinhos próximos. Decidindo por um comportamento coerente do modelo, informações conhecidas *a priori* podem ser definidas na forma de restrições que explorem essa interação, como a restrição de suavidade comumente imposta quando se trabalha com imagens (TIKHONOV e ARSENIN, 1977). BORMAN e STEVENSON (1998) apresentaram uma revisão dos métodos propostos até então, discutindo os rumos que as pesquisas em Reconstrução por Super Resolução poderiam tomar.

RAJAN e CHAUDHURI (2001) adotaram uma abordagem Bayesiana do problema da Reconstrução por Super Resolução para o caso em que as imagens observadas não possuíam deslocamentos entre si, porém cada uma foi adquirida sob condições de iluminação diferentes. Dessa forma, a etapa de Registro das imagens não foi necessária. Os autores utilizaram Campos Aleatórios de Markov para modelar as imagens e, para acelerar o processo de estimação, utilizaram o método de otimização *Simulated Annealing*. A imposição da hipótese de suavidade na imagem diminui a alta sensibilidade ao ruído, característica inerente a uma estimativa de Máxima Verossimilhança, na qual não existe imposição de um modelo *a priori* no processo de estimação.

Em um trabalho posterior (RAJAN e CHAUDHURI, 2002), os autores aplicaram essa abordagem para o caso em que ainda não existiam deslocamentos entre as imagens observadas, mas essas possuíam diferentes tipos de borramento. Nesse caso, o algoritmo proposto tratou simultaneamente a etapa de restauração e fusão das imagens de baixa resolução, sendo que uma minimização por Gradiente Descendente (GD - *Gradient Descent*) acelerou o processo.

JOSHI et al. (2004) propuseram uma técnica de Reconstrução por Super Resolução tendo como observações imagens da cena original com diferentes níveis de *zoom*, sendo tais níveis desconhecidos. A imagem de alta resolução foi modelada como um Campo Aleatório de Markov decidindo-se pela estimativa MAP. Não existiam deslocamentos entre as imagens de baixa resolução, e, dessa forma, o registro não foi implementado. Os autores afirmaram que o Campo Aleatório de Markov serviu como uma restrição contextual para regularizar a solução. Entretanto, foi preciso preservar de alguma forma as discontinuidades da imagem, já que *pixels* com mudanças significativas de intensidade corregam muita informação a respeito da imagem. Foi utilizado o método de otimização *Simulated Annealing* para acelerar o processo de estimação.

SROUBEK e FLUSSER (2006) geraram uma imagem de alta resolução a partir de duas ou mais imagens de baixa resolução da mesma cena adquiridas pelo mesmo sensor. Os autores implementaram uma técnica de Deconvolução Cega em Multicanais (MBD - *Multichannel Blind Deconvolution*), utilizando as informações presentes em todas as imagens de baixa resolução. A distribuição *a priori* da imagem de alta resolução foi modelada como um Campo Aleatório de Markov, e decidiu-se por uma estimativa MAP, utilizando o método de otimização Minimização Alternada.

JOSHI e CHAUDHURI (2006), na reconstrução de uma imagem de alta resolução a partir de imagens de baixa resolução adquiridas posicionando a fonte de luz em diferentes posições, levaram em consideração a estrutura tridimensional da cena original. Os autores utilizaram Campos Aleatórios de Markov para modelar a dependência contextual entre as intensidades dos *pixels*, a profundidade dos objetos, e outras características espacialmente correlacionadas.

Todas essas características foram modeladas separadamente, o que permitiu que a forma tridimensional fosse preservada.

Segundo STEPHENSON e TSUHAN (2006), aumento de resolução utilizando imagens de baixa resolução pode ser alcançado por meio de métodos de interpolação, Super Resolução utilizando múltiplos *frames* de um vídeo, e Super Resolução baseada em exemplos. Os autores afirmam que para o caso de imagens que características conhecidas e bem comportadas, como imagens de faces, a Super Resolução baseada em exemplos se mostra mais interessante. Nesse contexto, os autores utilizaram Campos Aleatórios de Markov e *Hallucinations* para inferir as informações de alta frequência da imagem de uma face, buscando exemplares do mesmo domínio dessa imagem. Pela flexibilidade oferecida pelos Campos Aleatórios de Markov, o método implementado possuía comportamento adaptativo a cada região da imagem.

HUMBLLOT e MOHAMMAD-DJAFARI (2006) adotaram o modelo clássico de formação das imagens de baixa resolução, onde a cena original de alta resolução recebe borramento, sofre os deslocamentos, é sub-amostrada, e finalmente, recebe ruído aditivo. Os autores implementaram um método não supervisionado, caracterizando as imagens pelo modelo Potts-Markov, em que era possível estimar de forma conjunta a classificação dos *labels*, os parâmetros da distribuição *a priori* e o modelo que caracteriza o ruído presente na imagem. Segundo os autores, o modelo de Potts-Markov permitiu a inserção da correlação entre *pixels* vizinhos no método.

SURESH e RAJAGOPALAN (2007) utilizaram um regularizador adaptativo a descontinuidades, no qual o grau de interação entre os *pixels* era ajustado de forma adaptativa a fim de preservar descontinuidades na imagem. Os autores caracterizaram as imagens como Campos Aleatórios de Markov, utilizando a Distribuição de Gibbs. A imposição das descontinuidades foi imposta na função energia da Distribuição de Gibbs, e a estimativa MAP de alta resolução foi encontrada utilizando o algoritmo ICM.

2.6 Considerações Finais

Campos Aleatórios de Markov têm-se mostrado uma ferramenta bastante interessante na modelagem de imagens por sua flexibilidade e facilidade na incorporação de características espaciais previamente conhecidas a respeito das imagens. Como visto, muitos trabalhos na literatura utilizaram Campos Aleatórios na solução do problema da Reconstrução por Super Resolução, chegando a resultados muito interessantes.

No próximo capítulo são discutidos Campos Aleatórios de Markov, a Distribuição de Gibbs e o Modelo de Potts-Strauss.

Campos Aleatórios de Markov e o Modelo de Potts-Strauss

3.1 Considerações Iniciais

Como visto no capítulo anterior, muitas abordagens do problema da Reconstrução por Super Resolução têm adotado Campos Aleatórios de Markov (MRF - *Markov Random Fields*) na modelagem das imagens. Os trabalhos pioneiros de BESAG (1974) e GEMAN e GEMAN (1984) são considerados marcos no Processamento de Imagens por terem impulsionado várias pesquisas na utilização de MRFs para a modelagem de imagens digitais.

Dentre outras razões, uma que torna os MRFs modelos tão interessantes é que na grande maioria dos problemas de estimação de imagens não se possui um modelo exato da imagem a ser estimada. Segundo DUBES e JAIN (1993), modelos considerando informações contextuais não precisam ser altamente precisos, e os MRFs modelam basicamente como ocorre a interação entre vizinhos próximos, dado um sistema de vizinhança.

Outra descoberta importante, que também incentivou a utilização de MRFs na modelagem de imagens, foi o trabalho de HAMMERSLEY e CLIFFORD (1971) que provou a equivalência entre os Campos Markovianos e a Distribuição de Gibbs. Essa equivalência deixou a modelagem das imagens menos complexa.

Dando continuidade ao trabalho de HAMMERSLEY e CLIFFORD (1971), foram desenvolvidas especializações da Distribuição de Gibbs, em muitas situações mais adequadas ao contexto do Processamento de Imagens. Uma dessas especializações é o modelo de Potts-Strauss, modelo utilizado nesta pesquisa.

Neste capítulo será dado um embasamento teórico a respeito de MRFs, a Distribuição de Gibbs e o Modelo de Potts-Strauss, necessário para o entendimento desta pesquisa.

3.2 Campos Aleatórios de Markov

Seja f uma imagem digital de dimensão $M \times N$. Em notação lexicográfica, f é um vetor de dimensão $1 \times MN$.

Chama-se *suporte* da imagem, denotado por S , os MN índices correspondentes aos *pixels* da imagem. Cada índice $s \in S$ é chamado *site*.

Ainda, a imagem pode ser vista como um vetor de variáveis aleatórias, indexadas pelos *sites* $s \in S$. O valor assumido por um *pixel* de índice s , denotado por f_s , $s \in S$, pertence a um conjunto finito enumerável G_s , chamado *espaço amostral* do *pixel* correspondente ao site s . Tratando-se de uma imagem formada por 256 níveis de cinza, o espaço amostral G_s , $s \in S$, é idêntico para todo s e igual a $\{0, 1, \dots, 255\}$. Apesar de na grande parte das aplicações os espaços amostrais G_s , $s \in S$, serem idênticos para todo s , isso não necessariamente precisa ocorrer.

Uma configuração possível da imagem, denotada por F_s , é o evento de todos os *sites* do suporte assumirem um dado valor. F_s é representada pelo Equação (3.1).

$$F_s = f_s, s \in S \quad (3.1)$$

O conjunto de todas as configurações possíveis, $F_s = f_s, s \in S$, é chamado *espaço amostral de configurações possíveis* F . Considerando as distribuições de probabilidade de F , denotadas por $\{P(f)_{f \in F}\}$, se $P > 0$ e $\sum_{s \in S} P(f) = 1$, pode-se afirmar que P é um *campo aleatório*. Trata-se da distribuição de probabilidade conjunta, imposta a um conjunto de variáveis aleatórias (DUBES e JAIN, 1993).

Campos aleatórios são processos estocásticos, ou funções aleatórias, cujos valores são mapeados em uma região do espaço (BESAG, 1974). Portanto, uma imagem F_s é interpretada como a realização de um campo aleatório.

Probabilidades condicionais como a definida na Equação (3.2), sendo $S \setminus \{a\} = S - \{a\}$, são as chamadas *características locais* e especificam a probabilidade do *pixel* indexado pelo site a , $a \in S$, assumir o valor f_a , dada a configuração no restante da imagem $f_{S \setminus \{a\}}$.

$$P(F_a = f_a | F_{S \setminus \{a\}} = f_{S \setminus \{a\}}), a \in S \quad (3.2)$$

Considerando que em uma imagem o valor de um *pixel* é altamente dependente dos valores dos *pixels* vizinhos a ele, a menos que se trate apenas de ruído aleatório, as características locais dependem apenas de um pequeno número de *pixels* espacialmente próximos.

Os MRFs são especificados em termos de probabilidades condicionais associadas a sistemas de vizinhança espacial. Um sistema de vizinhança espacial η_S é definido como um conjunto de subconjuntos de S , como mostra a Equação (3.3), sendo η_a o conjunto de vizinhos do *site* a .

$$\eta_S = \{\eta_a, a \in S\} \quad (3.3)$$

Para todo $a \in S$ tem-se que:

- $a \notin \eta_a$;
- e $b \in \eta_a$ se, e somente se $a \in \eta_b$.

Nessas condições, η_a é chamado *vizinhança de a*, e seus elementos são chamados *vizinhos de a*.

Um campo aleatório é dito Campo Aleatório de Markov se a probabilidade condicional mostrada na Equação (3.2), obedece a Equação (3.4), ou seja, o valor assumido por um *pixel* depende apenas de seus vizinhos, dado o sistema de vizinhança. Essa é a chamada *propriedade de Markov*.

$$P(F_a = f_a | F_{S \setminus \{a\}} = f_{S \setminus \{a\}}) = P(F_a = f_a | F_b = f_b, b \in \eta_a), \forall a \in S \quad (3.4)$$

Visto isso, o sistema de vizinhança espacial deve ser definido formalmente para que a interdependência entre os *pixels* seja especificada. Para um sistema de vizinhança definido como $\eta_a = \emptyset$, para todo $a \in S$, não existe dependência entre os *pixels* da imagem. Para o extremo oposto, $\eta_a = S \setminus \{a\}$, cada *pixel* depende de todos os demais.

Com base em tais constatações, todo campo aleatório pode ser considerado um MRF, basta apenas que se especifique o sistema de vizinhança de forma conveniente (WINKLER, 2006). Mesmo assim, na maioria das aplicações são definidos sistemas que englobam apenas os *pixels* mais próximos.

Os sistemas de vizinhança mais utilizados são o de primeira e de segunda ordem, mostrados na Figura (3.1). Outros sistemas comumente utilizados são mostrados na Figura (3.2).



Figura 3.1: Sistemas de vizinhança de (a) primeira ordem e de (b) segunda ordem.

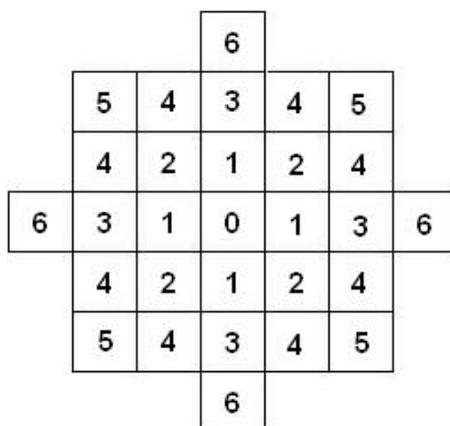


Figura 3.2: Sistemas de vizinhança de primeira a sexta ordem.

O teorema provado por HAMMERSLEY e CLIFFORD (1971) afirma que F é um MRF no suporte S , estabelecido o sistema de vizinhança η_S , se, e somente se sua distribuição de probabilidades P , for uma Distribuição de Gibbs em S com relação a η_S (LI, 1995a). Como a representação Gibbsiana é mais usual, o uso dos MRFs se tornou mais comum após esse teorema.

3.3 Distribuição de Gibbs

Um conjunto de variáveis aleatórias é dito um Campo Aleatório de Gibbs, com suporte S e com relação ao sistema de vizinhança η_S , se suas configurações obedecem à Distribuição de Gibbs dada pela Equação (3.5). Z , dado pelo Equação (3.6), é uma constante de normalização chamada *função de partição*.

$$P(f) = \frac{1}{Z} e^{-U(f)} \tag{3.5}$$

$$Z = \sum_{f \in F} e^{-U(f)} \quad (3.6)$$

Estabelecido o sistema de vizinhanças, todos os subconjuntos de S , cujos elementos são vizinhos entre si, são chamados *cliques*, denominados por $C(\eta_S)$. Se a vizinhança estipulada for de primeira ordem – Figura (3.1(a)) – os cliques serão os mostrados na Figura (3.3(a) – (c)). Da mesma forma, para o sistema de vizinhança de segunda ordem – Figura (3.1(b)) – os cliques serão os mostrados na Figura (3.3(a) – (j)).

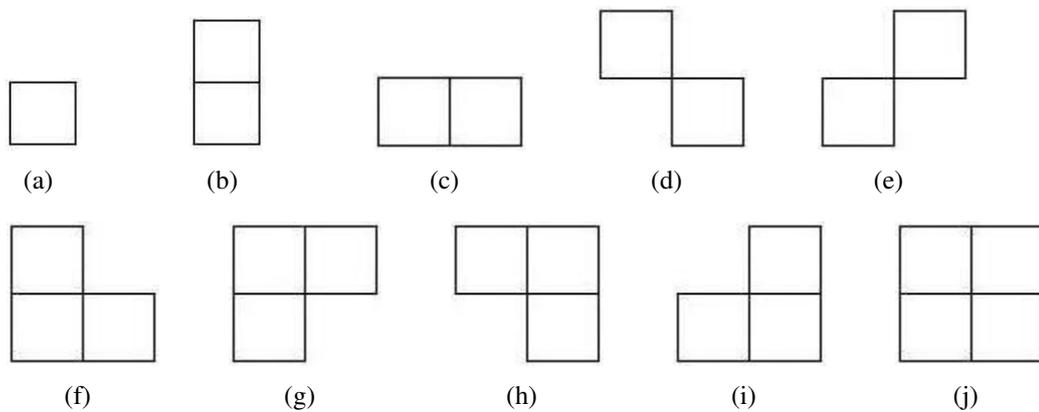


Figura 3.3: Cliques referentes aos sistemas de vizinhança de primeira ordem – de (a) a (c) – e de segunda ordem – de (a) a (j).

Na Equação (3.5), a função $U(f)$, chamada *função energia*, é definida pela Equação (3.7). $V_c(f)$ é o potencial associado ao clique $c \in C(\eta_S)$, e depende, portanto, apenas dos valores dos *pixels* pertencentes ao clique c . Em comparação às características locais do MRF, a função potencial é uma forma mais conveniente e mais natural de modelar a imagem.

$$U(f) = \sum_{c \in C(\eta_S)} V_c(f) \quad (3.7)$$

DERIN e ELLIOT (1987), utilizando o sistema de vizinhança de segunda ordem, associaram um parâmetro a cada tipo de clique, exceto ao clique formado apenas pelo *pixel* de interesse, como mostra a Figura (3.4). Assim, os potenciais foram definidos como mostra a Equação (3.8), sendo ζ o parâmetro associado ao clique.

$$V_c(f) = \begin{cases} -\zeta & \text{se todas as classes em } c \text{ são iguais} \\ \zeta & \text{caso contrário} \end{cases} \quad (3.8)$$

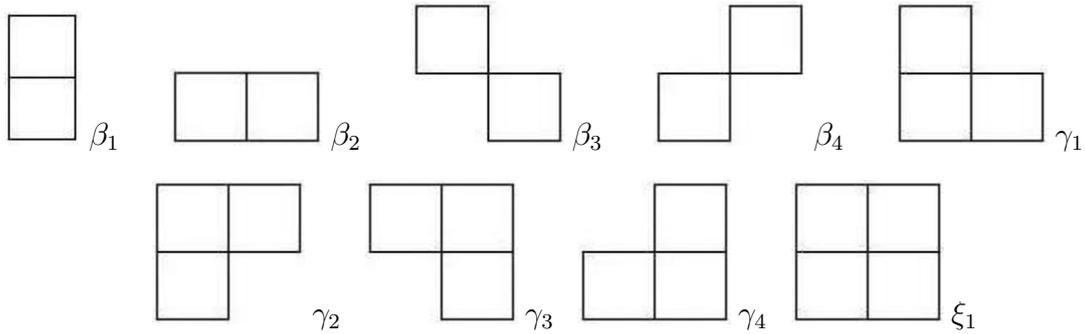


Figura 3.4: Abordagem utilizada por DERIN e ELLIOT (1987), associando um parâmetro a cada tipo de clique

Nesse contexto, poder-se-ia atribuir valores aos parâmetros de forma que configurações do MRF bastante homogêneas estivessem relacionadas a altas probabilidades, o contrário ocorrendo com imagens bastante irregulares. A abordagem utilizada por DERIN e ELLIOT (1987) mostra a facilidade de se impor alguma característica da imagem através de uma modelagem que utiliza a Distribuição de Gibbs para descrever a imagem. Em contra partida, ela também mostra a importância de um modelo coerente, já que características não desejadas poderiam ser erroneamente impostas à imagem.

A função energia $U(f)$ da distribuição de Gibbs, onde é modelada a interação entre *pixels* vizinhos, considera a imagem como um todo. Em muitas situações isso pode ser um inconveniente por acarretar um custo computacional muito alto para a aplicação. Existem modelos que são especializações da Distribuição de Gibbs, bastante utilizados na caracterização de MRFs.

O Modelo de Potts-Strauss, considera a interação entre *pixels* vizinhos localmente, o que, em aplicações que utilizam imagens de resolução muito alta, por exemplo, pode ser mais interessante do que tratar a imagem como um todo a cada cálculo da função energia. Esse modelo é utilizado nos casos em que a dimensão do espaço amostral das variáveis aleatórias que compõem o campo, é maior do que dois. No caso em que a dimensão do espaço amostral é dois, utiliza-se o Modelo de Ising (CIPRA, 1987), considerado uma especialização do Modelo de Potts.

3.4 Modelo de Potts-Strauss

Seja S o suporte finito da imagem, $F = \prod_{s \in S} F_s$ o espaço de configurações possíveis do MRF, e β um número real, segundo o Modelo de Potts-Strauss, a distribuição de probabilidade de F é dada pela Equação (3.9). $\Pi_A(x)$ é a função indicadora do conjunto A e Z_β a chamada

função de partição do Modelo de Potts-Strauss, dadas pelas Equações (3.10) e (3.11), respectivamente.

$$P(F = f) = \frac{1}{Z_\beta} \exp \left(\beta \sum_{\|s-t\|=1} \Pi_{\{f_s\}}(f_t) \right) \quad (3.9)$$

$$\Pi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{caso contrário} \end{cases} \quad (3.10)$$

$$Z_\beta = \sum_{f \in F} \exp \left\{ \sum_{\|s-t\|=1} \Pi_{\{f_s\}}(f_t) \right\} \quad (3.11)$$

A probabilidade conjunta, $P(F = f)$, pode também ser especificada pelo conjunto de todas as distribuições condicionais, como mostra a Equação (3.12), sendo $\#A$ o número de elementos do conjunto A .

$$P(F_s = f_s | F_t = f_t) = \exp \{ \beta \# \{t \in \eta_s : f_s = f_t\} \} \forall s \in S \quad (3.12)$$

β é chamado parâmetro de atratividade, sendo que para $\beta < 0$ tem-se um modelo repulsivo, para $\beta > 0$ um modelo atrativo, e para $\beta = 0$ um modelo independente. O Modelo de Potts-Strauss apresenta um comportamento chamado *transição de fase*, que depende do valor de um parâmetro, denominado *parâmetro crítico* (FRERY, 1993).

Segundo BESAG (1974), as características locais do MRF determinam unicamente a distribuição conjunta $P(F = f)$. A transição de fase ocorre quando existe mais de uma probabilidade em F , dada a família de características locais.

FRERY (1993) apresenta o fenômeno da transição de fase para o modelo de Ising, e esclarece que o valor de β controla esse fenômeno, impedindo ou não que $P(F = f)$ seja unicamente identificada pelas características locais. O autor afirma que existe $0 < \beta_c < +\infty$ tal que a distribuição apresenta o fenômeno da transição de fase apenas se $\beta > \beta_c$.

Nesse contexto, um dos maiores problemas é estimar o valor do parâmetro β , dada uma ocorrência de F . Estimá-lo pelo critério de Máxima Verossimilhança se torna computacionalmente inviável, já que a dependência do modelo com relação a esse parâmetro acontece de forma muito complexa.

A fim de obter um estimador computacionalmente viável, FRERY (1993) provou a consistência do estimador por Máxima Pseudoverossimilhança.

3.4.1 Estimação por Máxima Pseudoverossimilhança

Seja $\hat{\beta}$ a estimativa de Máxima Pseudoverossimilhança do parâmetro de atratividade β , que determina a unicidade da distribuição de Potts-Strauss. Essa estimativa é calculada resolvendo a Equação (3.13), sendo $\vartheta_s(t)$ dado pela Equação (3.14).

$$\sum_{s \in S} \left[\vartheta_s(f_s) - \frac{\sum_{x \in F_s} \vartheta_s(x) \exp(\hat{\beta} \vartheta_s(x))}{\sum_{x \in F_s} \exp(\hat{\beta} \vartheta_s(x))} \right] = 0 \quad (3.13)$$

$$\vartheta_s(t) = \# \{u \in \eta_s : f_u = t\} \quad (3.14)$$

FRERY (1993) desenvolve a equação de Máxima Pseudoverossimilhança para os casos em que o espaço amostral possui dois, três e quatro ou mais elementos, sendo que em todos os casos F_s é idêntico para todo s . O autor considerou o sistema de vizinhança de primeira ordem (Figura (3.1(a))) para todos os casos.

Quando o espaço amostral possui apenas dois elementos, $F_s = \{0, 1\}$ para todo s , $\hat{\beta}$ é encontrado resolvendo a Equação (3.15), C_k^l dado pela Equação (3.16).

$$2(C_4^1 - C_0^1 + C_0^0 + C_4^1 + C_3^1) - C_1^1 + C_1^0 - \frac{2(C_0^0 + C_4^1 + C_4^0) \exp(4\hat{\beta})}{1 + \exp(4\hat{\beta})} - \frac{(C_1^0 + C_3^1 + C_3^0) \exp(2\hat{\beta})}{1 + \exp(2\hat{\beta})} = 0 \quad (3.15)$$

$$C_k^l = \# \left\{ s \in S : f_s = l, \sum_{t \in \eta_s} f_t = k \right\} \quad (3.16)$$

Desenvolvendo a Equação (3.13) para o caso em que o espaço amostral possui três elementos, $F_s = \{0, 1, 2\}$ para todo s , e chamando os *sites* vizinhos de s como $\eta_s = \{t, u, v, w\}$, $\hat{\beta}$ é encontrado resolvendo a Equação (3.17), V^2 , V^3 e V^4 dados pela Equação (3.18).

$$\sum_{s \in S} \vartheta_s(f_s) - \frac{4V^4 \exp(4\hat{\beta})}{1 + \exp(4\hat{\beta})} - V^3 \frac{3 \exp(3\hat{\beta}) + \exp(\hat{\beta})}{\exp(3\hat{\beta}) + \exp(\hat{\beta})} - 2V^2 = 0 \quad (3.17)$$

$$\begin{aligned} V^2 &= \# \{f_{\eta_s} \subset S : f_t = f_u \neq f_v = f_w\} \\ V^3 &= \# \{f_{\eta_s} \subset S : f_t = f_u = f_v \neq f_w\} \\ V^4 &= \# \{f_{\eta_s} \subset S : f_t = f_u = f_v = f_w\} \end{aligned} \quad (3.18)$$

Por fim, FRERY (1993) desenvolveu a Equação (3.13) para o caso em que o espaço amostral possuía K elementos, $K \geq 4$, $F_s = \{0, 1, 2, \dots, k-1\}$, novamente chamando os sites vizinhos de s por $\eta_s = \{t, u, v, w\}$. Assim, a equação de Máxima Pseudoverossimilhança se reduz à Equação (3.19), com K^0 , K^1 , K^2 , K^3 e K^4 dados pela Equação (3.20).

$$\begin{aligned} \sum_{s \in S} \vartheta_s(f_s) - 4K^4 \frac{\exp(4\hat{\beta})}{\exp(4\hat{\beta})+K} - K^3 \frac{3\exp(3\hat{\beta})+\exp(\hat{\beta})}{\exp(3\hat{\beta})+\exp(\hat{\beta})+K-1} - \\ - K^2 \frac{4\exp(2\hat{\beta})}{2\exp(2\hat{\beta})+K-1} - K^1 \frac{2(\exp(2\hat{\beta})+\exp(\hat{\beta}))}{\exp(2\hat{\beta})+2\exp(\hat{\beta})+K-2} - \\ - K^0 \frac{4\exp(\hat{\beta})}{4\exp(\hat{\beta})+K-3} = 0 \end{aligned} \quad (3.19)$$

$$\begin{aligned} K^0 &= \#\{f_{\eta_s} \subset S : f_i \neq f_j, \forall i \neq j\} \\ K^1 &= \#\{f_{\eta_s} \subset S : f_t = f_u \neq f_v \neq f_w \neq f_t\} \\ K^2 &= \#\{f_{\eta_s} \subset S : f_t = f_u \neq f_v = f_w\} \\ K^3 &= \#\{f_{\eta_s} \subset S : f_t = f_u = f_v \neq f_w\} \\ K^4 &= \#\{f_{\eta_s} \subset S : f_t = f_u = f_v = f_w\} \end{aligned} \quad (3.20)$$

Desenvolvendo da equação de Máxima Pseudoverossimilhança para vizinhança de segunda ordem (Figura 3.1(b)) e espaço amostral com 256 elementos – os 256 níveis de cinza – chega-se à Equação (3.21). Denominando os oito vizinhos do site s como $\eta_s = \{a, b, c, d, e, f, g, h\}$, tem-se k_i , N_i e D_i mostrados nas Equações (3.22), (3.23) e (3.24), respectivamente.

$$\sum_{s \in S} \vartheta_s(f_s) - \sum_{i=0}^{21} k_i \frac{N_i}{D_i} \quad (3.21)$$

A Equação (3.22) representa todas as configurações possíveis entre os oito vizinhos do *pixel* de interesse. Em k_0 temos os oito vizinhos pertencentes à mesma classe, em k_1 sete vizinhos pertencentes a uma classe, e um pertencente a uma classe diferente dos demais, em k_2 seis vizinhos pertencentes a uma classe, e os outros dois pertencentes a uma classe diferente dos demais vizinhos, porém iguais entre si, e assim por diante.

No contexto da Reconstrução por Super Resolução, assim como no contexto da Restauração de Imagens, em geral se possui observações degradadas da cena original e algum conhecimento a respeito do processo de formação dessas observações e da cena original e ser estimada. Entretanto o conhecimento a respeito da cena original não é exatamente preciso, e por isso os MRFs são modelos bastante adotados para sua caracterização.

Adotar MRFs significa adotar a suposição de que *pixels* próximos possuem classes próximas, ou seja, existe imposição da hipótese de suavidade na imagem. Porém, essa imposição

$$\begin{aligned}
k_0 &= \#\{f_{\eta_s} \subset S : f_i = f_j, \forall i \neq j\} \\
k_1 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_h, \forall i, j \in \{a, b, c, d, e, f, g\}\} \\
k_2 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_g = f_h, \forall i, j \in \{a, b, c, d, e, f\}\} \\
k_3 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_g \neq f_h, \forall i, j \in \{a, b, c, d, e, f\}\} \\
k_4 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_f = f_g = f_h, \forall i, j \in \{a, b, c, d, e\}\} \\
k_5 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_f = f_g \neq f_h, \forall i, j \in \{a, b, c, d, e\}, f_h \neq f_i\} \\
k_6 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_f \neq f_g \neq f_h, \forall i, j \in \{a, b, c, d, e\}, f_g \neq f_i, f_h \neq f_i, f_h \neq f_f\} \\
k_7 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_k = f_l, \forall i, j \in \{a, b, c, d\}, \forall k, l \in \{e, f, g, h\}\} \\
k_8 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_k = f_l \neq f_h, \forall i, j \in \{a, b, c, d\}, \forall k, l \in \{e, f, g\}, f_h \neq f_i\} \\
k_9 &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_e = f_f \neq f_g = f_h, \forall i, j \in \{a, b, c, d\}, f_h \neq f_i\} \\
k_{10} &= \#\{f_{\eta_s} \subset S : f_i = f_j \neq f_e = f_f \neq f_g \neq f_h, \forall i, j \in \{a, b, c, d\}, f_g \neq f_i, f_h \neq f_i, f_h \neq f_e\} \\
k_{11} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c = f_d \neq f_i \forall i \in \{e, f, g, h\}\} \\
k_{12} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c \neq f_d = f_e = f_f \neq f_g = f_h, f_g \neq f_a\} \\
k_{13} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c \neq f_d = f_e = f_f \neq f_g \neq f_h, f_g \neq f_a, f_a \neq f_h \neq f_d\} \\
k_{14} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c \neq f_d = f_e \neq f_f = f_g \neq f_h, f_f \neq f_a \neq f_h \neq f_d\} \\
k_{15} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c \neq f_d = f_e \neq f_f \neq f_g \neq f_h, f_a \neq f_f \neq f_h, f_a \neq f_h \neq f_d \neq f_g\} \\
k_{16} &= \#\{f_{\eta_s} \subset S : f_a = f_b = f_c \neq f_i, \forall i \in \{d, e, f, g, h\}\} \\
k_{17} &= \#\{f_{\eta_s} \subset S : f_a = f_b \neq f_c = f_d \neq f_e = f_f \neq f_g = f_h, f_a \neq f_e, f_c \neq f_g\} \\
k_{18} &= \#\{f_{\eta_s} \subset S : f_a = f_b \neq f_c = f_d \neq f_e = f_f \neq f_g \neq f_h, f_a \neq f_e \neq f_h, f_a \neq f_g \neq f_c, f_a \neq f_h \neq f_c\} \\
k_{19} &= \#\{f_{\eta_s} \subset S : f_a = f_b \neq f_c = f_d \neq f_i, f_i \neq f_a, \forall i \in \{e, f, g, h\}\} \\
k_{20} &= \#\{f_{\eta_s} \subset S : f_a = f_b \neq f_i, \forall i \in \{c, d, e, f, g, h\}\} \\
k_{21} &= \#\{f_{\eta_s} \subset S : f_i \neq f_j, \forall i, j \in \{a, b, c, d, e, f, g, h\}\}
\end{aligned} \tag{3.22}$$

pode causar uma suavização indevida em toda a imagem, e descontinuidades presentes na cena original são desconsideradas.

O que se faz geralmente é adotar modelos que sejam adaptativos a descontinuidades, de forma a não suavizá-las. A seguir é discutido como esse processo ocorre.

3.5 Adaptação a Descontinuidades

Desde o conceito *Line Process* introduzido pelo trabalho de GEMAN e GEMAN (1984), modelos adaptativos a descontinuidade, principalmente quando se está utilizando MRFs na caracterização das imagens, são bastante adotados (LI, 1995a).

Bordas e descontinuidades são mudanças abruptas de níveis de cinza, presentes na imagem. Entretanto, o ruído também provoca tais mudanças abruptas. Os modelos adaptativos a descontinuidades devem ser capazes de diferenciar entre essas duas possibilidades.

Por meio de uma análise dos mecanismos de suavização em termos da Equação de Euler, LI (1995a) mostrou que a principal diferença entre os modelos adaptativos a descontinuidade

$$\begin{aligned}
N_0 &= 8 \exp \left(8\hat{\beta} \right) \\
N_1 &= 7 \exp \left(7\hat{\beta} \right) + \exp \left(\hat{\beta} \right) \\
N_2 &= 6 \exp \left(6\hat{\beta} \right) + 2 \exp \left(2\hat{\beta} \right) \\
N_3 &= 6 \exp \left(6\hat{\beta} \right) + 2 \exp \left(\hat{\beta} \right) \\
N_4 &= 5 \exp \left(5\hat{\beta} \right) + 3 \exp \left(3\hat{\beta} \right) \\
N_5 &= 5 \exp \left(5\hat{\beta} \right) + 2 \exp \left(2\hat{\beta} \right) + \exp \left(\hat{\beta} \right) \\
N_6 &= 5 \exp \left(5\hat{\beta} \right) + 3 \exp \left(\hat{\beta} \right) \\
N_7 &= 8 \exp \left(4\hat{\beta} \right) \\
N_8 &= 4 \exp \left(4\hat{\beta} \right) + 3 \exp \left(3\hat{\beta} \right) + \exp \left(\hat{\beta} \right) \\
N_9 &= 4 \exp \left(4\hat{\beta} \right) + 4 \exp \left(2\hat{\beta} \right) \\
N_{10} &= 4 \exp \left(4\hat{\beta} \right) + 2 \exp \left(2\hat{\beta} \right) + 2 \exp \left(\hat{\beta} \right) \\
N_{11} &= 4 \exp \left(4\hat{\beta} \right) + 4 \exp \left(\hat{\beta} \right) \\
N_{12} &= 6 \exp \left(3\hat{\beta} \right) + 2 \exp \left(2\hat{\beta} \right) \\
N_{13} &= 6 \exp \left(3\hat{\beta} \right) + 2 \exp \left(\hat{\beta} \right) \\
N_{14} &= 3 \exp \left(3\hat{\beta} \right) + 4 \exp \left(2\hat{\beta} \right) + \exp \left(\hat{\beta} \right) \\
N_{15} &= 3 \exp \left(3\hat{\beta} \right) + 2 \exp \left(2\hat{\beta} \right) + 3 \exp \left(\hat{\beta} \right) \\
N_{16} &= 3 \exp \left(3\hat{\beta} \right) + 5 \exp \left(\hat{\beta} \right) \\
N_{17} &= 8 \exp \left(2\hat{\beta} \right) \\
N_{18} &= 6 \exp \left(2\hat{\beta} \right) + 2 \exp \left(\hat{\beta} \right) \\
N_{19} &= 4 \exp \left(2\hat{\beta} \right) + 4 \exp \left(\hat{\beta} \right) \\
N_{20} &= 2 \exp \left(2\hat{\beta} \right) + 6 \exp \left(\hat{\beta} \right) \\
N_{21} &= 8 \exp \left(\hat{\beta} \right)
\end{aligned} \tag{3.23}$$

propostos é a forma com que a interação entre vizinhos ocorre. De maneira geral tais modelos agem de forma que quando uma descontinuidade é encontrada, a interação entre vizinhos diminui.

Nesse sentido, adota-se uma função responsável pela detecção das descontinuidades $g(x)$, denominada *função de interação adaptativa*. Se tal função tomar a forma puramente quadrática (Equação (3.25)), intuitivamente adotada na detecção de bordas, quanto mais irregular o valor do *pixel* em relação a seus vizinhos, maior o resultado da função e maior a suavização. Con-

$$\begin{aligned}
D_0 &= \exp\left(8\hat{\beta}\right) + (256 - 1) \\
D_1 &= \exp\left(7\hat{\beta}\right) + \exp\left(\hat{\beta}\right) + (256 - 2) \\
D_2 &= \exp\left(6\hat{\beta}\right) + \exp\left(2\hat{\beta}\right) + (256 - 2) \\
D_3 &= \exp\left(6\hat{\beta}\right) + 2\exp\left(\hat{\beta}\right) + (256 - 3) \\
D_4 &= \exp\left(5\hat{\beta}\right) + \exp\left(3\hat{\beta}\right) + (256 - 2) \\
D_5 &= \exp\left(5\hat{\beta}\right) + \exp\left(2\hat{\beta}\right) + \exp\left(\hat{\beta}\right) + (256 - 3) \\
D_6 &= \exp\left(5\hat{\beta}\right) + 3\exp\left(\hat{\beta}\right) + (256 - 4) \\
D_7 &= 2\exp\left(4\hat{\beta}\right) + (256 - 2) \\
D_8 &= \exp\left(4\hat{\beta}\right) + \exp\left(3\hat{\beta}\right) + \exp\left(\hat{\beta}\right) + (256 - 3) \\
D_9 &= \exp\left(4\hat{\beta}\right) + 2\exp\left(2\hat{\beta}\right) + (256 - 3) \\
D_{10} &= \exp\left(4\hat{\beta}\right) + \exp\left(2\hat{\beta}\right) + 2\exp\left(\hat{\beta}\right) + (256 - 4) \\
D_{11} &= \exp\left(4\hat{\beta}\right) + 4\exp\left(\hat{\beta}\right) + (256 - 5) \\
D_{12} &= 2\exp\left(3\hat{\beta}\right) + \exp\left(2\hat{\beta}\right) + (256 - 3) \\
D_{13} &= 2\exp\left(3\hat{\beta}\right) + 2\exp\left(\hat{\beta}\right) + (256 - 4) \\
D_{14} &= \exp\left(3\hat{\beta}\right) + 2\exp\left(2\hat{\beta}\right) + \exp\left(\hat{\beta}\right) + (256 - 4) \\
D_{15} &= \exp\left(3\hat{\beta}\right) + \exp\left(2\hat{\beta}\right) + 3\exp\left(\hat{\beta}\right) + (256 - 5) \\
D_{16} &= \exp\left(3\hat{\beta}\right) + 5\exp\left(\hat{\beta}\right) + (256 - 6) \\
D_{17} &= 4\exp\left(2\hat{\beta}\right) + (256 - 4) \\
D_{18} &= 3\exp\left(2\hat{\beta}\right) + 2\exp\left(\hat{\beta}\right) + (256 - 5) \\
D_{19} &= 2\exp\left(2\hat{\beta}\right) + 4\exp\left(\hat{\beta}\right) + (256 - 6) \\
D_{20} &= \exp\left(2\hat{\beta}\right) + 6\exp\left(\hat{\beta}\right) + (256 - 7) \\
D_{21} &= 8\exp\left(\hat{\beta}\right) + (256 - 8)
\end{aligned} \tag{3.24}$$

seqüentemente, haverá uma suavização indevida na imagem e as discontinuidades não serão respeitadas.

$$g(x) = x^2 \tag{3.25}$$

O conceito *Line Process* age de forma a “desligar” a suavização quando a derivada do sinal excede um dado limite. Primeiramente as discontinuidades são encontradas e uma matriz de 0s e 1s é mantida, 1 indicando a presença de uma discontinuidade e 0 indicando a ausência. Ao

trabalhar com a hipótese de suavidade, quando uma descontinuidade é encontrada – presença do valor 1 na matriz – essa hipótese é desconsiderada.

Nesse contexto, as descontinuidades encontradas são tratadas exatamente da mesma forma, não considerando sua intensidade. Além disso, dependendo do limite imposto, descontinuidades de menor intensidade podem ser desrespeitadas ou ruído pode ser considerado uma descontinuidade. O ideal seria que o comportamento da função fosse específico para cada caso individual.

Segundo LI (1995b), uma condição necessária a qualquer função de interação adaptativa $g(x)$, é a dada pela Equação (3.26), onde x é a diferença entre *pixels* vizinhos, $C \in [0, \infty)$ uma constante e $h(x)$ a função que modela a maneira como se dá a interação entre *pixels* vizinhos.

$$\lim_{x \rightarrow \infty} |g'(x)| = \lim_{x \rightarrow \infty} |2xh(x)| = C \quad (3.26)$$

É preciso que essa função cresça monotonicamente juntamente com x em um certo intervalo dentro do qual as mudanças de níveis de cinza são consideradas ruído e por isso devem ser suavizadas. Fora desse intervalo a suavização deve ir decaindo à medida que as mudanças de níveis de cinza aumentam, proporcionando uma suavização proporcional à intensidade da descontinuidade.

Quatro funções que possuem esse comportamento nos respectivos intervalos são mostradas na Tabela (3.1). Todas essas funções são escritas em função de um dado parâmetro γ , o qual flexibiliza os intervalos.

Tabela 3.1: Quatro possibilidades de funções de interação adaptativa (LI, 1995b).

$h_\gamma(x) = e^{-\frac{x^2}{\gamma}}$	$g_\gamma(x) = -\gamma e^{-\frac{x^2}{\gamma}}$	$B_\gamma = (-\sqrt{\frac{\gamma}{2}}, \sqrt{\frac{\gamma}{2}})$
$h_\gamma(x) = \frac{1}{[1+\frac{x^2}{\gamma}]^2}$	$g_\gamma(x) = -\frac{\gamma}{1+\frac{x^2}{\gamma}}$	$B_\gamma = (-\sqrt{\frac{\gamma}{3}}, \sqrt{\frac{\gamma}{3}})$
$h_\gamma(x) = \frac{1}{1+\frac{x^2}{\gamma}}$	$g_\gamma(x) = \gamma \log(1 + \frac{x^2}{\gamma})$	$B_\gamma = (-\sqrt{\gamma}, \sqrt{\gamma})$
$h_\gamma(x) = \frac{1}{1+\frac{ x }{\gamma}}$	$g_\gamma(x) = \gamma x - \gamma^2 \ln(1 + \frac{ x }{\gamma})$	$B_\gamma = (-\infty, +\infty)$

Na Tabela (3.1), B_γ é o intervalo dentro do qual a função correspondente cresce monotonicamente juntamente com x . A Figura (3.5) mostra as formas qualitativas dessas funções.

A Figura (3.6) mostra as descontinuidades detectadas na imagem Lena borrada e ruidosa, quando se utiliza a primeira função mostrada na Tabela (3.1).

Detectadas as descontinuidades por meio de uma função de interação adaptativa, em geral essa informação é imposta no processo de Restauração ou Reconstrução de forma a prevenir a

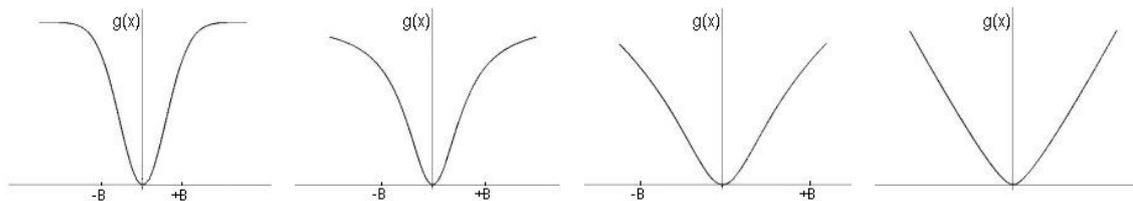


Figura 3.5: Formas qualitativas das funções de interação adaptativa mostradas na Tabela (3.1).



Figura 3.6: Detecção de descontinuidades na imagem Lena utilizando uma função de interação adaptativa

suavização da imagem. Os resultados quando a adaptação a descontinuidade é aplicada são bem mais satisfatórios do que quando ela não é aplicada, o que fica bastante evidente no trabalho de BHATT e DESAI (1994).

3.6 Considerações Finais

Como visto anteriormente, a escolha por um modelo realístico para descrever a imagem é um passo crítico no processo de estimação. Modelagens utilizando MRFs e a Distribuição de Gibbs, ou uma especialização dessa como o Modelo de Potts-Strauss, permitem a introdução de informações espaciais de contexto, não necessitando de alta precisão.

A Teoria de Decisão Bayesiana é uma forma de estipular regras de decisão a fim de se encontrar a melhor aproximação da cena original, utilizando para isso informações previamente conhecidas a respeito da imagem e do modelo de formação da imagem.

No próximo capítulo são discutidos o estimador MAP e o Algoritmo ICM, aplicáveis ao caso em que se trabalha com uma abordagem Bayesiana do problema e se adota MRFs como modelo *a priori* da imagem.

Inferência Estatística e o Algoritmo ICM

4.1 Considerações Iniciais

A grande maioria dos problemas existentes no Processamento de Imagens – Segmentação, Restauração, Reconstrução, Classificação, etc. – consistem fundamentalmente em problemas de tomada de decisão. Decidir a qual classe pertence um dado padrão ou qual a real intensidade de um *pixel* exige primeiramente que exista um modelo descritivo dos dados de entrada (DUBES e JAIN, 1993).

Nesse sentido, comparando-se com abordagens determinísticas, as abordagens estatísticas oferecem maior facilidade para se encontrar um modelo coerente para o processo de formação das imagens. A Teoria de Decisão Bayesiana oferece meios para que regras de decisão sejam estipuladas na estimação da cena original.

Devido à presença de ruído, esse processo de estimação normalmente é mal-condicionado. Para tratar o problema do mal-condicionamento, faz-se a imposição de informações conhecidas *a priori* da imagem a ser estimada. Nesse sentido, pela Teoria de Decisão Bayesiana, procura-se maximizar a distribuição de probabilidade *a posteriori*, adotando um estimador MAP.

Geralmente estimadores MAP são lentos e computacionalmente caros, já que se busca uma configuração dentre as milhares possíveis. Para otimizar essa busca, são utilizados métodos de otimização dentre os quais está o Algoritmo ICM.

A seguir são discutidos alguns conceitos dentro da Teoria de Decisão Bayesiana que serão úteis no entendimento desta pesquisa.

4.2 Teoria de Decisão Bayesiana

A Teoria de Decisão Bayesiana é uma abordagem fundamentalmente estatística quando o problema é modelado em termos probabilísticos e as distribuições relevantes conhecidas (DUDA e HART, 1973). No contexto do Processamento de Imagens tal abordagem foi originalmente utilizada apenas em problemas de Segmentação de Imagens, posteriormente se mostrando útil na modelagem de problemas de Restauração e Reconstrução de Imagens.

O modelo probabilístico é especificado em termos de uma distribuição de probabilidade, sendo que não é possível explicitar ou definir um resultado particular de antemão. Dessa forma, esse modelo é utilizado quando se tem um grande número de variáveis – que não podem ser controladas – influenciando o resultado (MONTGOMERY e RUNGER, 2003).

Suponha que em um experimento, um dentre k eventos possa ocorrer, sendo os eventos denominados $w_i, i = 1, \dots, k$. Por exemplo, no lançamento de um dado, as seis faces representam os possíveis eventos. O conjunto de todos os eventos possíveis é chamado *espaço amostral* e será denominado por Ω .

Nesse contexto, a probabilidade *a priori* do j -ésimo evento, $P(w_j)$, reflete o pré-conhecimento de quão provável o evento w_j pode ocorrer, antes que ele realmente ocorra (DUDA e HART, 1973).

Em problemas modelados utilizando uma abordagem estatística, uma regra de decisão é definida de forma a indicar a ação a ser tomada para todas as possíveis observações. Se a regra de decisão fosse decidir pelo evento de maior probabilidade *a priori*, a decisão seria sempre a mesma e pelo evento mais provável.

Para que isso não ocorra, além da probabilidade conhecida *a priori* na grande maioria dos problemas conta-se também com algumas amostras ou observações que ajudam a complementar a regra de decisão. Assim, a decisão depende também do “estado da natureza”.

Tal dependência é representada pela probabilidade condicional $P(x|w_i)$, de um evento x ocorrer, dado que w_i ocorreu. Essa probabilidade é também chamada *verossimilhança* de x com relação a w_i .

O *Teorema de Bayes* relaciona a probabilidade *a priori* e a probabilidade condicional, representando uma forma de modificar a decisão com base nas observações. A Equação (4.1) mostra a *Regra de Bayes* que implementa o Teorema de Bayes, onde $P(x)$ é dado pela Equação (4.2).

$$P(w_j|x) = \frac{P(x|w_j) P(w_j)}{P(x)} \quad (4.1)$$

$$P(x) = \sum_{i=1}^k P(x|w_i) P(w_i) \quad (4.2)$$

Chamando os eventos de *classes*, suponha que existe um conjunto de k ações $A = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ sendo que tomar a ação α_i significa decidir pela classe w_i .

Para cada ação existe uma perda por se tomar a ação errada, $\lambda(\alpha_i|w_j)$ sendo a perda pela decisão por α_i sendo que a classe observada foi w_j . Assim, o *risco condicional* inerente à ação α_i dado que um evento x ocorreu é dado pela soma das probabilidades condicionais de cada um dos possíveis eventos dado que x ocorreu, multiplicadas pela perda ocasionada pela ação α_i (Equação (4.3)).

$$R(\alpha_i|x) = \sum_{j=1}^k \lambda(\alpha_i|w_j) P(w_j|x) \quad (4.3)$$

Para o caso em que existem múltiplas classes $k \geq 2$, a regra de decisão é definida por meio de k *funções discriminantes*. Como o próprio nome diz, funções discriminantes servem para discriminar entre as possibilidades existentes. Decide-se pela ação de maior função discriminante $g_i(x) > g_j(x)$ para todo $j \neq i$.

A regra de decisão poderia ser definida como “decide-se pela ação que implica em menor risco condicional”. Nesse caso, a função discriminante seria definida como mostra a Equação (4.4).

$$g_i(x) = -R(\alpha_i|x) \quad (4.4)$$

Supondo que não haja perda pela ação correta, $\lambda(\alpha_i|w_i) = 0$ para $i = 1, 2, \dots, k$, e que a perda por uma ação incorreta seja 1, $\lambda(\alpha_i|w_j) = 1$ para todo $i \neq j$, tem-se que o risco condicional pode ser dado pela Equação (4.5).

$$R(\alpha_i|x) = \sum_{j=1, j \neq i}^k P(w_j|x) = 1 - P(w_i|x) \quad (4.5)$$

Nesse caso a função discriminante poderia ser definida como mostra a Equação (4.6) e decidir-se-ia pela classe de maior probabilidade *a posteriori*, o chamado estimador de Máximo a Posteriori (MAP - *Maximum a Posteriori*).

$$g_i(x) = P(w_i|x) \quad (4.6)$$

Pela Regra de Bayes (Equação (4.1)) a função discriminante dada pela Equação (4.6) poderia ser escrita como mostra a Equação (4.7).

$$g_i(x) = \frac{P(x|w_i) P(w_i)}{P(x)} \quad (4.7)$$

Finalmente, como $P(x)$ independe da classe observada, esse termo pode ser desconsiderado. Assim, o estimador MAP pode ser definido como mostra a Equação (4.8).

$$g_i(x) = P(x|w_i) P(w_i) \quad (4.8)$$

Para casos em que a probabilidade *a priori* das classes não é conhecida, ou mesmo quando todas as classes são equiprováveis e têm-se apenas as probabilidades condicionais, a função discriminante pode ser definida como mostra a Equação (4.9). O estimador que utiliza tal função discriminante é chamado estimador de Máxima Verossimilhança (ML - *Maximum Likelihood*).

$$g_i(x) = P(x|w_i) \quad (4.9)$$

Na maioria das situações o estimador MAP é preferível, justamente por permitir a inserção de informações a respeito da distribuição *a priori* das classes no processo de estimação. A seguir o estimador MAP para o contexto do Processamento de Imagens será discutido.

4.3 O Estimador MAP

A distribuição de probabilidade *a priori* deve refletir o pré-conhecimento a respeito da imagem. Como visto anteriormente, imagens podem ser modeladas como realizações de campos aleatórios indexados pelos respectivos suportes. Mais especificamente, devido ao pré-conhecimento que *pixels* próximos possuem intensidades parecidas, as imagens podem ser modeladas como realizações de MRFs representando a interdependência entre os *pixels* de acordo com o sistema de vizinhança adotado.

Seja f uma imagem formada por 256 níveis de cinza, de dimensão $1 \times MN$ – notação lexicográfica – e suporte $S = \{s : 1 \leq s \leq MN\}$, seus *pixels* são modelados como variáveis aleatórias que podem assumir um dentre os 256 níveis de cinza.

Dado o sistema de vizinhança $\eta_S = \{\eta_a, a \in S\}$, onde $\eta_a \subseteq S$ é o conjunto de vizinhos do *pixel* de índice a , a imagem é modelada como a realização de um MRF cujas características locais determinam unicamente sua distribuição conjunta $P(f)$ (BESAG, 1974). Assim, tem-se que a probabilidade *a priori* $P(f)$ pode ser caracterizada por uma Distribuição de Gibbs, dada a função energia $U(f)$ e os respectivos potenciais associados aos cliques $V_c(f)$ (Equação (3.7)).

Na maior parte das situações, além do conhecimento *a priori* a respeito da imagem, tem-se também observações degradadas da cena original e algum conhecimento a respeito do processo de degradação. Segundo o modelo de formação das imagens discutido no Capítulo 2, $g_k = DW_k Hf + r$ é a k -ésima observação degradada e de baixa resolução, $k = 1, \dots, q$, sendo r ruído Gaussiano aditivo com média μ e variância σ^2 , H a PSF considerada idêntica para todas as observações, W_k os deslocamentos aplicados à k -ésima observação, e D o operador de sub-amostragem também considerado idêntico para todas as observações.

Por essa representação da formação das imagens, pode-se afirmar que $r = g_k - DW_k Hf$. Assim, a probabilidade condicional da k -ésima observação dada a imagem original $P(g_k|f, \sigma^2)$, é proporcional à Gaussiana representada pela Equação (4.10) (CORTIJO et al., 2003).

$$P(g_k|f, \sigma^2) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{MN}{2}} \exp \left(-\frac{1}{2\sigma^2} \|g - DW_k Hf\|^2 \right) \quad (4.10)$$

Como existem q observações degradadas e de baixa resolução, as quais são consideradas eventos independentes, a distribuição conjunta $P(g|f, \sigma^2)$ é dada pelo produto mostrado na Equação (4.11), proporcional à Equação (4.12) (CORTIJO et al., 2003).

$$P(g|f, \sigma^2) = \prod_{i=1}^q P(g_i|f, \sigma^2) \quad (4.11)$$

$$P(g|f, \sigma^2) = \prod_{i=1}^q \frac{1}{(2\pi\sigma^2)^{\frac{MN}{2}}} \exp \left[-\frac{1}{2} \sum_{i=1}^q \frac{1}{\sigma^2} \|g_i - DW_i Hf\|^2 \right] \quad (4.12)$$

Nesse contexto o estimador MAP (Equação (4.8)) é dado pela Equação (4.13).

$$P(f|g) = P(g|f, \sigma^2) P(f) \quad (4.13)$$

Para encontrar a estimativa de máxima probabilidade *a posteriori*, a busca exaustiva dentre todas as configurações possíveis é uma alternativa computacionalmente inviável. Cada *pixel* da

imagem pode assumir um dentre os 256 níveis de cinza e, supondo que se trata de uma imagem de dimensão 128x128, existem 256^{16384} configurações possíveis. Assim, os algoritmos utilizando estimadores MAP, apesar da facilidade na modelagem do problema, tendem a convergir muito lentamente.

A fim de solucionar o problema do custo computacional inerente a estimadores MAP, foram propostas várias abordagens que utilizam heurísticas no processo de estimação. KIRKPATRICK et al. (1988) propuseram uma das primeiras tentativas de otimizar funções objetivo relacionadas a MRFs, o *Simulated Annealing*. Segundo os autores as duas estratégias a serem adotadas para otimizar o processo de estimação se resumem na divisão do processo em partes mais simples e a atualização iterativa. Dividir o problema em sub-problemas menores e independentes, faz com que seja possível chegar a um solução com menor dificuldade.

Durante a atualização iterativa o processo é iniciado com uma configuração inicial e uma operação padrão que é aplicada a todas as partes do sistema. Essa operação padrão visa modificar a configuração corrente de forma a maximizar a função objetivo. Esse processo é realizado várias vezes até que nenhuma atualização seja feita.

GEMAN e GEMAN (1984) desenvolveram um algoritmo de restauração denominado *Stochastic Relaxation* (SR), inspirado no algoritmo *Metropolis* (METROPOLIS et al., 1953), o qual foi desenvolvido com o intuito de estudar as propriedades de equilíbrio de sistemas constituídos por um grande número de componentes interagindo.

O *Stochastic Relaxation* utiliza o *Simulated Annealing* de KIRKPATRICK et al. (1988) aliado a um amostrador também proposto por GEMAN e GEMAN (1984) denominado *Gibbs Sampler*. Esse algoritmo é considerado clássico e é muito utilizado como critério de comparação para métodos de otimização propostos (BERTHOD et al., 1996).

Entretanto, apesar desse algoritmo propor a otimização na busca de uma estimativa MAP com garantia de convergência, ele é de difícil implementação e a convergência para uma aproximação da solução desejada geralmente é lenta devido à dimensão do espaço de configurações possíveis do MRF.

O algoritmo proposto por BESAG (1986), denominado *Iterated Conditional Modes* (ICM) surgiu com o intuito de suprir as deficiências do *Stochastic Relaxation*. Trata-se de uma abordagem determinística que não garante convergência para o máximo global, mas que se devidamente inicializado garante bons resultados em poucas iterações. O ICM é um algoritmo iterativo que foi desenvolvido de forma a ser computacionalmente simples e não sofrer a presença de características em larga escala inerente a MRFs.

A seguir o algoritmo ICM é descrito detalhadamente.

4.4 O Algoritmo ICM

Os dois principais problemas dos algoritmos que utilizam estimadores MAP e também do algoritmo *Stochastic Relaxation* proposto por GEMAN e GEMAN (1984) são:

- Os algoritmos geralmente são lentos devido à dimensão do espaço de configurações possíveis do MRF;
- As características locais do campo aleatório em conjunto com a suposição que *pixels* vizinhos possuem classes próximas levam à presença de características em larga escala – formação de grandes áreas homogêneas ou de mesma classe.

Como dito anteriormente, o ICM foi desenvolvido por BESAG (1986) a fim de suprir esses dois problemas, além de ser computacionalmente simples.

Supondo uma estimativa inicial \hat{f} da cena original f , pretende-se atualizar o valor do *pixel* de índice i , \hat{f}_i de forma a encontrar uma estimativa MAP. Uma escolha plausível é o valor que maximiza a probabilidade condicional de f_i dado o vetor formado pelas observações y , e a configuração atual no restante da imagem $f_{S \setminus i}$. Assim, \hat{f}_i maximiza $P(f_i | y, f_{S \setminus i})$, que pelo Teorema de Bayes é dada pela Equação (4.14), sendo f_{η_i} os valores dos vizinhos do *pixel* de índice i .

$$P(f_i | y, f_{S \setminus i}) = P(y_i | f_i) P(f_i | f_{\eta_i}) \quad (4.14)$$

Nesse caso a implementação é bastante simplificada já que apenas as configurações do *pixel* e de seus vizinhos são analisadas. Eventualmente o ICM irá convergir para um máximo local já que a probabilidade *a posteriori* dada pela Equação (4.15) não irá decair em momento algum. Isso é garantido pelo fato de que $P(f_i | y, f_{S \setminus i})$ é maximizado para todo i .

$$P(f | y) = \prod_{i=1}^{MN} P(f_i | y, \hat{f}_{S \setminus i}) \quad (4.15)$$

O ICM não garante convergência para o máximo global da função, mas, se bem inicializado, ele garante bons resultados em poucas iterações. BESAG (1986) afirmou que a convergência do algoritmo se dá de forma relativamente rápida, sendo que praticamente nenhuma atualização é feita a partir do sexto ciclo. Sua modelagem é extremamente simples já que são feitos apenas cálculos locais e pode-se trabalhar com qualquer modelo adotado na caracterização do MRF.

Flexibilidade e rapidez são características que tornam o ICM tão interessante para implementações que manipulam grandes quantidades de dados.

O algoritmo ICM utiliza as seguintes entradas:

- Número de classes – níveis de cinza;
- Estimativa inicial da cena original;
- Número máximo de modificações a cada iteração, escolhido de forma empírica;
- Número máximo de iterações para, caso o algoritmo não alcance a convergência, o processo seja interrompido.

Esse algoritmo foi originalmente desenvolvido para processamento seqüencial, entretanto como a mesma operação é realizada para todos os *pixels*, bastando apenas que cada *pixel* transmita seu valor para seus vizinhos, existe a possibilidade de implementá-lo de forma paralela. Os *pixels* poderiam ser visitados como mostra a Figura (4.1), sendo que os números representam a ordem em que essa visita ocorre e os diferentes níveis de cinza o possível paralelismo.

1	10	2	11	3	12
19	28	20	29	21	30
4	13	5	14	6	15
22	31	23	32	24	33
7	16	8	17	9	18
25	34	26	35	27	36

Figura 4.1: Possível paralelismo do algoritmo ICM.

4.5 Considerações Finais

Devido à facilidade de implementação e versatilidade com relação ao modelo adotado na caracterização das imagens, o algoritmo ICM se mostra uma boa alternativa quando é preciso encontrar uma estimativa MAP da imagem original.

Juntamente com o Modelo de Potts-Strauss e de forma adaptativa à descontinuidades, o algoritmo pode ser aplicado ao contexto da Reconstrução por Super Resolução, o que é mostrado no próximo capítulo.

Método Proposto

5.1 Considerações Iniciais

Nesta pesquisa foi considerado o problema de aumento de resolução a partir de um conjunto de imagens de baixa resolução borradas e ruidosas. Além disso, essas imagens possuíam deslocamentos *sub-pixel* uma em relação às demais.

Nesse contexto, as diferentes informações presentes em cada uma das imagens observadas são utilizadas na busca de uma estimativa de resolução maior do que a presente em qualquer uma das imagens de entrada. Na literatura, essa abordagem é denominada Reconstrução por Super Resolução ou Realce de Resolução.

Adotando um *framework* Bayesiano, esta pesquisa propõe um algoritmo para a Reconstrução por Super Resolução que caracteriza as imagens como MRFs utilizando o modelo de Potts-Strauss como função densidade de probabilidade *a priori* da imagem de alta resolução a ser estimada.

Dadas as imagens de baixa resolução da mesma cena – que possuem deslocamentos *sub-pixel* uma em relação às demais – os deslocamentos são estimados por meio de uma técnica de Registro. Uma primeira estimativa de alta resolução é encontrada interpolando as imagens registradas de acordo com os deslocamentos estimados. Por fim, o algoritmo ICM adaptativo a discontinuidades é utilizado para melhorar a qualidade dessa primeira estimativa de alta resolução.

Neste capítulo o método proposto é discutido de forma detalhada.

5.2 Modelo de Formação das Imagens

Como discutido no Capítulo 2, considera-se que existem q imagens de baixa resolução da mesma cena, cada qual podendo ser vista como uma matriz de dimensão $N_1 \times N_2$. Da mesma forma a cena original a ser estimada pode ser vista como uma matriz de dimensão $M_1 \times M_2$, sendo que $N_1 \leq M_1$ e $N_2 \leq M_2$. O modelo de formação das imagens de baixa resolução discutido no Capítulo 2 é dado pela Equação (5.1).

$$g_k = DW_k Hf + r \quad (5.1)$$

Utilizando notação lexicográfica, g_k é o vetor de dimensão $1 \times N_1 N_2$ contendo os *pixels* da k -ésima imagem de baixa resolução, $k = 1, \dots, q$. Tais imagens são adquiridas por meio do borramento, *warping* e sub-amostragem da cena original de alta resolução, cujos *pixels* são dados pelo vetor f de dimensão $1 \times M_1 M_2$. D , de dimensão $M_1 M_2 \times N_1 N_2$, modela o operador de *downsampling* ou sub-amostragem da imagem de alta resolução para a grade de baixa resolução. W_k , de dimensão $M_1 M_2 \times M_1 M_2$, modela os deslocamentos *sub-pixel* presentes na k -ésima imagem de baixa resolução, sendo que esse operador é diferente para cada uma das imagens observadas. H , de dimensão $M_1 M_2 \times M_1 M_2$, modela a PSF inerente ao sensor que pode ser considerada a mesma para todas as imagens adquiridas. E, por fim, r modela ruído aditivo provocado pela manipulação e armazenamento das imagens.

Em contrapartida a essa modelagem, os operadores de sub-amostragem D , e o que modela os deslocamentos *sub-pixel* entre as imagens W_k , podem ser vistos de forma conjunta, já que na maioria das situações o processo de amostragem da cena original invariavelmente provoca os deslocamentos *sub-pixel* presentes nas imagens observadas. Assim, a versão de baixa resolução da cena original pode ser dada pela Equação (5.2).

$$d = Df \quad (5.2)$$

d , de dimensão $1 \times N_1 N_2$, é o vetor contendo os *pixels* de baixa resolução dados pela Equação (5.2), sendo f_m os componentes do vetor de alta resolução f .

$$d_n = \sum_m \delta_{n,m} \cdot f_m \quad (5.3)$$

De acordo com a Equação (5.3), os *pixels* de baixa resolução são interpretados como uma soma ponderada dos *pixels* de alta resolução, sendo $\delta_{n,m}$ os pesos dessa soma, ou seja, os elementos do operador de sub-amostragem D . Dessa forma, a sub-amostragem da imagem, que inerentemente provoca os deslocamentos *sub-pixel*, ocorre como mostra a Figura (5.1).

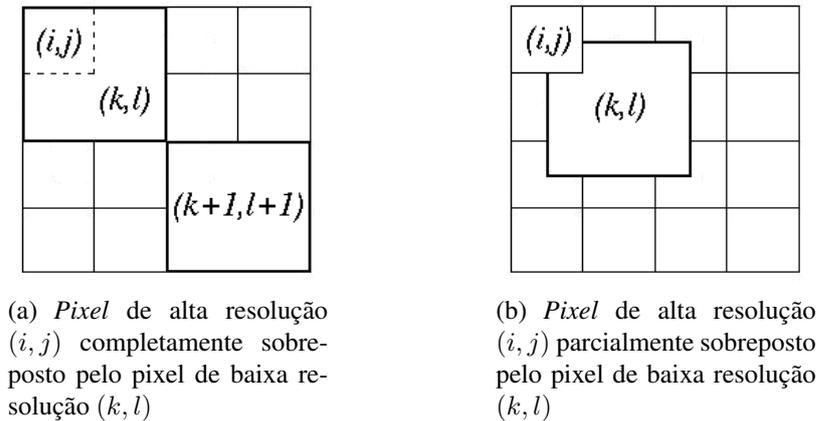


Figura 5.1: Processo de subamostragem provocando os deslocamentos *sub-pixel* entre as imagens (HOMEM et al., 2006).

Na Figura (5.1(a)), a grade de baixa resolução está perfeitamente alinhada à grade de alta resolução, e, por isso, o *pixel* de alta resolução (i, j) está completamente sobreposto pelo *pixel* de baixa resolução (k, l) . Para esse *pixel* de baixa resolução, na soma ponderada, o peso relativo aos *pixels* de alta resolução sobrepostos seria 1, e o peso dos demais 0, já que os sobrepostos contribuem completamente para aquele *pixel* de baixa resolução, e os demais não contribuem em nada.

No caso da Figura (5.1(b)), a grade de baixa resolução está um pouco deslocada para baixo e para a direita. Dessa forma, o *pixel* de alta resolução (i, j) possui apenas $1/4$ de sua área sobreposta pelo *pixel* de baixa resolução (k, l) . Além dele, outros oito *pixels* de alta resolução estão parcialmente ou completamente sobrepostos pelo *pixel* de baixa resolução (k, l) . Os quatro *pixels* que possuem $1/2$ de sua área sobreposta possuiriam peso 0.5, os quatro *pixels* que possuem $1/4$ de sua área sobreposta possuiriam peso 0.25, e o *pixel* de alta resolução completamente sobreposto possuiria peso 1.

Assim, o processo de formação das imagens de baixa resolução modelando a sub-amostragem juntamente com os deslocamentos *sub-pixel*, pode ser dado pela Equação (5.4).

$$g_k = D_k H f + r \tag{5.4}$$

Encontrado o modelo de formação das imagens, para que seja possível estimar a imagem de alta resolução é necessário primeiramente estimar os deslocamentos presentes entre as imagens de baixa resolução por meio de uma técnica de Registro de Imagens.

5.3 Registro das Imagens

De acordo com a modelagem do operador de sub-amostragem discutido na seção anterior, temos que a formação das imagens de baixa resolução a partir da cena original de alta resolução se dá como mostram as Figuras (5.2) e (5.3).

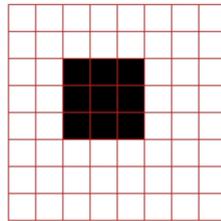


Figura 5.2: Imagem de alta resolução (HOMEM et al., 2006).

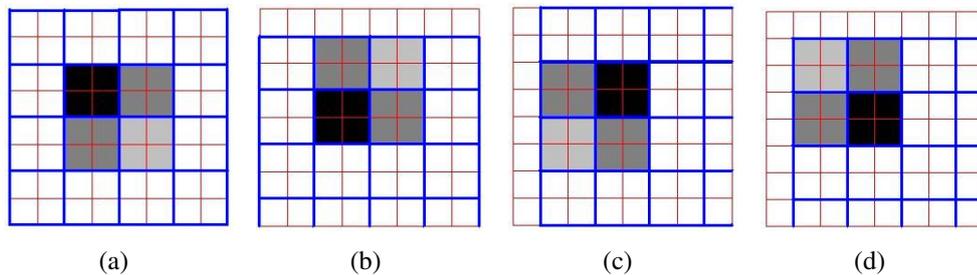


Figura 5.3: Imagens resultantes da sub-amostragem da imagem de alta resolução mostrada na Figura (5.2) (HOMEM et al., 2006).

A imagem de alta resolução mostrada da Figura (5.2) é sub-amostrada nas quatro situações mostradas nas Figuras (5.3(a) – (d)). Na Figura (5.3(a)), a grade de baixa resolução, mostrada em azul, está perfeitamente alinhada à grade de alta resolução, mostrada em vermelho. Tem-se que quatro *pixels* de alta resolução possuindo nível de cinza preto estão completamente sobrepostos por um *pixel* de baixa resolução. Dessa forma, o *pixel* de baixa resolução também recebe esse nível de cinza, já que na soma ponderada dos *pixels* de alta resolução, os *pixels* possuem peso 1 por estarem completamente sobrepostos.

No caso dos *pixels* de baixa resolução à direita e abaixo desse *pixel* que recebeu nível de cinza preto, como dois *pixels* com nível de cinza preto e dois *pixels* com nível de cinza branco

estão sobrepostos, tem-se que o *pixel* de baixa resolução adquire um valor intermediário entre esses dois níveis de cinza, ou seja, uma tonalidade de cinza intermediário entre branco e preto.

Para o *pixel* de baixa resolução na diagonal inferior daquele que recebeu nível de cinza preto, como ele está sobrepondo apenas um *pixel* de alta resolução preto e quatro brancos, ele recebe uma tonalidade de cinza mais claro, ou seja, mais próximo de branco do que de preto.

Nas Figuras (5.3(b) – (d)), a grade de baixa resolução está deslocada uma unidade para baixo, uma unidade para a direita, e uma unidade para baixo e para a direita, respectivamente. Da mesma forma que na Figura (5.3(a)), os *pixels* de baixa resolução adquirem um dado nível de cinza, de acordo com a soma ponderada dos *pixels* de alta resolução sendo sobrepostos.

Adquiridas as imagens de baixa resolução da forma explicada, os deslocamentos entre elas são identificados em relação a uma imagem de referência, a qual é escolhida entre as imagens observadas. Dessa forma, as imagens são comparadas duas a duas, cada uma comparada à imagem de referência. Os deslocamentos x_0 e y_0 são estimados minimizando a função de similaridade mostrada na Equação (5.5) (HOMEM et al., 2006). g_0 é a imagem de referência, e cada uma das imagens é comparada a essa imagem de referência.

$$s(x_0, y_0) = \sum_{x \in X} \sum_{y \in Y} [g_0(x, y) - g_1(x - x_0, y - y_0)]^2 \quad (5.5)$$

O resultado da otimização da Equação (5.5) é dado pelas Equações (5.6) e (5.7), sendo g_{1x} e g_{1y} as derivadas de primeiro ordem de g_1 com relação a x e com relação a y , respectivamente.

$$x_0^* = \frac{\sum_x \sum_y [(g_0 - g_1)(x, y) - y_0^* g_{1y}(x, y)] g_{1x}(x, y)}{\sum_x \sum_y g_{1x}^2(x, y)} \quad (5.6)$$

$$y_0^* = \frac{\sum_x \sum_y [(g_0 - g_1)(x, y) - x_0^* g_{1x}(x, y)] g_{1y}(x, y)}{\sum_x \sum_y g_{1y}^2(x, y)} \quad (5.7)$$

De acordo com a Equação (5.5), os deslocamentos são encontrados minimizando a soma das diferenças entre os *pixels* da imagem de referência g_0 , e de uma das imagens de baixa resolução g_1 . Essa soma de diferenças é feita várias vezes, deslocando a imagem comparada à de referência por vários deslocamentos x_0 e y_0 , até que a função de similaridade seja minimizada. Esse processo é mostrado na Figura (5.4).

Toma-se a Figura (5.4(a)) como imagem de referência, sendo a área de interesse mostrada em verde. Para se estimar os deslocamentos dessa imagem com relação à imagem mostrada na Figura (5.4(b)), primeiramente calcula-se $s(x_1, y_1)$ para $x_1 = 0$ e $y_1 = 0$. Calculando $s(x_2, y_2)$ para $x_2 = 1$ e $y_2 = 1$, ou seja, a imagem deslocada para baixo e para a direita como mostra a

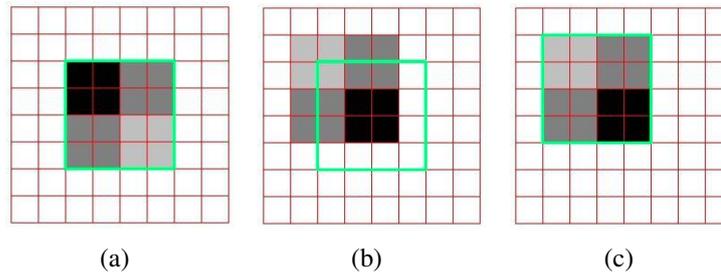


Figura 5.4: Estimação dos deslocamentos *sub-pixel* entre as imagens de baixa resolução, dada uma imagem de referência (HOMEM et al., 2006).

Figura (5.4(c)), tem-se que $s(x_2, y_2) < s(x_1, y_1)$. Dessa forma, x_2 e y_2 são os deslocamentos que minimizam a função de similaridade.

Com o mesmo procedimento feito para cada uma das imagens de baixa resolução, os deslocamentos presentes entre cada uma delas e a imagem de referência são estimados. Feito isso, as imagens de baixa resolução podem ser alinhadas e, por meio de um método de interpolação, uma primeira estimativa de alta resolução pode ser adquirida como ilustrado nas Figuras (5.5) e (5.6).

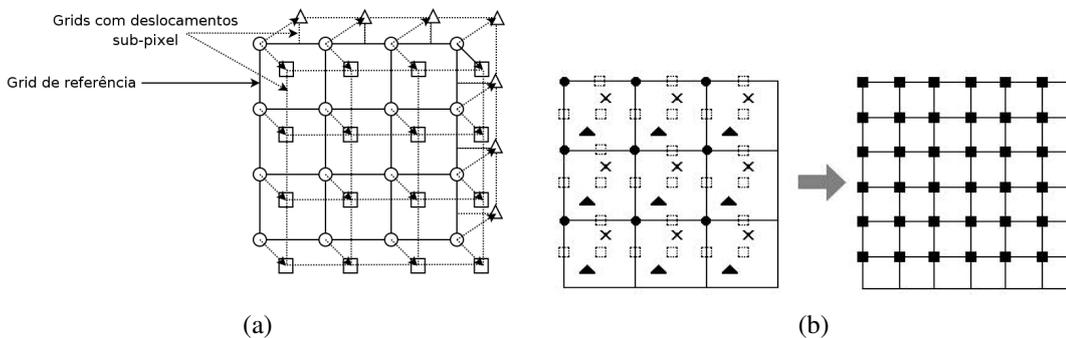


Figura 5.5: Ilustração do Registro aplicado às imagens de baixa resolução, permitindo que elas sejam alinhadas e que uma primeira estimativa de alta resolução seja adquirida (HOMEM et al., 2006).

Estimados os deslocamentos e alinhadas as imagens, caso tais deslocamentos sejam equidistantes e uniformemente espaçados, não será necessária a aplicação de um método interpolatório para se chegar a uma estimativa inicial de alta resolução. Nesse caso, os *pixels* de baixa resolução já estarão perfeitamente posicionados na grade de alta resolução. Essa situação é ilustrada na Figura (5.7), onde quatro imagens de baixa resolução, cujos *pixels* são representados por quadrados, círculos, estrelas e pentágonos, respectivamente, foram alinhadas de acordo com os deslocamentos estimados na etapa de registro.



Figura 5.6: Ilustração utilizando a imagem Lena, do Registro aplicado às imagens de baixa resolução.

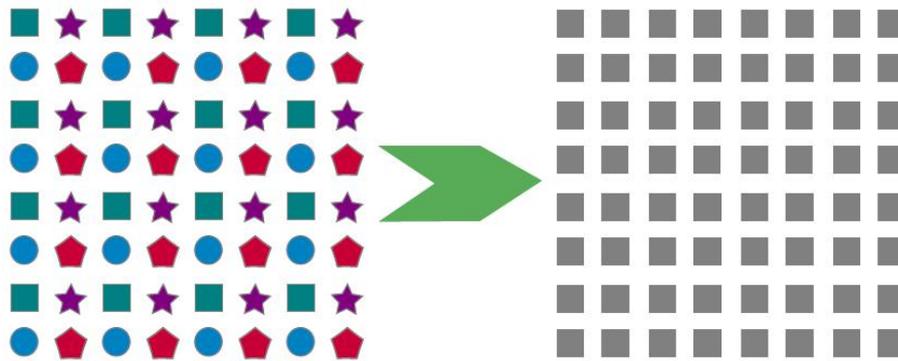


Figura 5.7: Situação em que existem deslocamentos equidistantes e uniformemente espaçados entre as imagens de baixa resolução

A Figura (5.8) ilustra a situação em que os deslocamentos estimados estão dispostos de forma irregular. Dessa forma, é necessário que se aplique algum procedimento interpolatório nos *pixels* de baixa resolução, para se chegar à grade de alta resolução. Esse procedimento pode ser qualquer um existente na literatura.

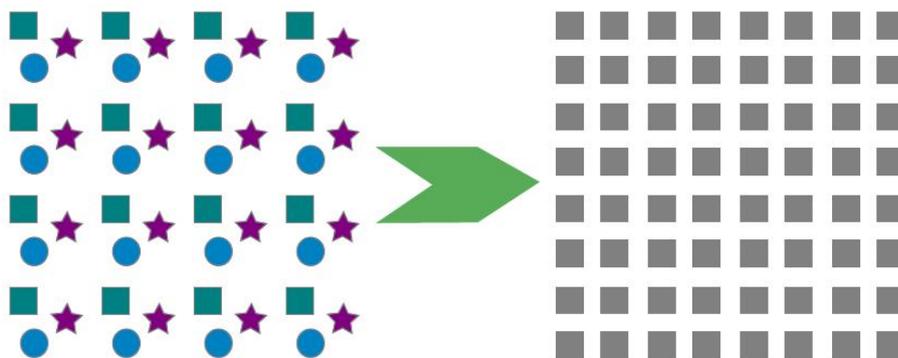


Figura 5.8: Situação em que os deslocamentos presentes entre as imagens de baixa resolução estão dispostos de forma irregular

5.4 Formulação Bayesiana do Problema

De forma geral, a Reconstrução por Super Resolução é um problema mal-condicionado já que a solução pode não existir ou podem existir várias soluções. Como discutido no Capítulo 4, para tratar o problema do mal-condicionamento faz-se a imposição de informações conhecidas *a priori* a respeito da imagem a ser estimada, e do processo de formação das imagens observadas.

Abordagens estatísticas, em comparação com abordagens determinísticas, oferecem maior facilidade para se encontrar um modelo coerente para o processo de formação das imagens de baixa resolução e para descrever a imagem de alta resolução que se quer estimar. Dessa forma, decidindo-se por uma abordagem estatística, a imposição de informações *a priori* é feita pela distribuição de probabilidade da imagem a ser estimada.

Segundo a abordagem Bayesiana do problema, pretende-se encontrar uma estimativa \hat{f} da cena original de alta resolução f , dadas as observações de baixa resolução g_k , $k = 1, 2, \dots, q$, e dada a distribuição de probabilidade *a priori* de f . Decide-se pela estimativa de maior probabilidade *a posteriori*, a chamada estimativa de *Máximo a Posteriori* (MAP), como mostra a Equação (5.8), onde g é formado pelas q imagens de baixa resolução observadas.

$$\hat{f} = \arg \max_f \{p(f|g)\} \quad (5.8)$$

Pela Regra de Bayes, tem-se que a estimativa de alta resolução é encontrada pela Equação (5.9), sendo $p(f)$ a distribuição de probabilidade *a priori* da imagem de alta resolução.

$$\hat{f} = \arg \max_f \{p(g|f)p(f)\} \quad (5.9)$$

Como discutido no Capítulo 3, a maior dificuldade nessa abordagem é que na maioria das situações não se possui a distribuição de probabilidade da imagem a ser estimada. Nesse caso, pode-se adotar modelos contextuais como MRFs, já que esses não exigem grande precisão na caracterização das imagens.

Modelos como a distribuição de Gibbs ou especializações dessa, como o modelo de Potts-Strauss, são utilizados na caracterização de MRFs. Esta pesquisa adotou o modelo de Potts-Strauss discutido a seguir.

5.5 Modelo de Potts-Strauss

Especificado o sistema de vizinhança η , o modelo de Potts-Strauss pode ser definido pelo conjunto de todas as distribuições condicionais dadas pela Equação (5.10), sendo $\beta \in \Re$

o chamado parâmetro de atração ou repulsão se for positivo ou negativo, respectivamente (MASCARENHAS e FRERY, 1994).

$$p(f_i | \hat{f}_{\eta_i}) \sim e^{\beta \cdot \#\{t \in \eta_i | f_i = f_t\}} \quad (5.10)$$

Dessa forma, adotado o modelo de Potts-Strauss, a distribuição de probabilidade *a priori* da imagem de alta resolução é dada pelo conjunto das probabilidades condicionais de cada *pixel* f_i , dados seus vizinhos f_{η_i} .

O parâmetro β é estimado pela Equação de Máxima Pseudoverossimilhança como explicado no Capítulo 3, sendo que esta pesquisa adotou o sistema de vizinhança de segundo grau (Figura (3.1(b))). Para esse caso, a Equação de Máxima Pseudoverossimilhança toma a forma mostrada na Equação (3.21), sendo o número de classes igual a 256 – os 256 níveis de cinza.

Definida a distribuição *a priori*, para encontrar a estimativa MAP da imagem de alta resolução, esta pesquisa utilizou o algoritmo ICM discutido a seguir.

5.6 O Algoritmo ICM Adaptativo a Descontinuidades

Como visto no Capítulo 4, os métodos que implementam estimadores MAP, apesar da facilidade na modelagem do problema, tendem a convergir muito lentamente. Nesse sentido, o algoritmo *Iterated Conditional Modes* proposto por BESAG (1986), se mostra uma alternativa computacionalmente simples para solucionar esse problema.

Trata-se de uma abordagem determinística que não garante convergência para o máximo global, mas, se bem inicializada, garante bons resultados em poucas iterações. Nesta pesquisa o algoritmo ICM foi utilizado com o intuito de melhorar a estimativa de alta resolução conseguida pelo processo de Registro e Interpolação das imagens de baixa resolução observadas.

Essa melhora é alcançada maximizando as probabilidades condicionais locais $p(f_i | g, \hat{f}_{\eta_i})$ sequencialmente, ou seja, maximizando a probabilidade de cada *pixel* de alta resolução f_i , dados os *pixels* de baixa resolução g , e a vizinhança desse *pixel* de alta resolução \hat{f}_{η_i} , sendo η_i os vizinhos do *pixel* i , estabelecido o sistema de vizinhança do MRF, η .

Dessa forma, segundo a Regra de Bayes a probabilidade *a posteriori* do *pixel* de alta resolução, dados os *pixels* de baixa resolução e os vizinhos desse *pixel* de alta resolução $p(f_i | g, \hat{f}_{\eta_i})$, pode ser aproximada como mostra a Equação (5.11).

$$p(f_i | g, \hat{f}_{\eta_i}) \sim p(g_i | f_i) \cdot p(f_i | \hat{f}_{\eta_i}) \quad (5.11)$$

Como discutido na seção anterior, adotado o Modelo de Potts-Strauss como distribuição de probabilidade *a priori* da imagem de alta resolução, a função densidade de probabilidade de cada *pixel* de alta resolução f_i , dados seus vizinhos \hat{f}_{η_i} , $p(f_i|\hat{f}_{\eta_i})$, é dada pela Equação (5.10).

Nesta pesquisa adotou-se a função densidade de probabilidade condicional dos *pixels* de baixa resolução dado o *pixel* de alta resolução $p(g_i|f_i)$, como mostra a Equação (5.12), sendo σ^2 a variância, $m_i = \frac{1}{C} \left(\sum_{j \in \eta_i} f_j \right) + \frac{1}{C} f_i$, e $C = \#\eta_i + 1$.

$$p(g_i|f_i) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(g_i - m_i)^2}{2\sigma^2}} \quad (5.12)$$

Visto isso, o algoritmo ICM utiliza uma estimativa inicial da imagem de alta resolução, para, a cada iteração, atualizar essa estimativa maximizando $p(f_i|g, \hat{f}_{\eta_i})$ para todo $i \in S$, sendo S o suporte da imagem. A estimativa inicial adotada foi a imagem de alta resolução resultante do processo de Registro e Interpolação das imagens de baixa resolução, como explicado anteriormente.

Assim, o algoritmo ICM ocorre da seguinte forma:

1. Inicializa-se \hat{f} com a estimativa inicial de alta resolução encontrada ao se registrar e interpolar as imagens de baixa resolução observadas;
2. Para todo $i \in S$, atualiza-se \hat{f}_i de forma a maximizar

$$p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$$

3. Repete-se o item (2) por um número máximo de iterações T_{iter} ou até que nenhuma alteração seja feita.

Devido à natureza dos MRFs e do Modelo de Potts-Strauss, o algoritmo ICM provoca uma suavização indevida na imagem, já que o valor de f_i que maximiza a probabilidade *a posteriori* $p(f_i|g, \hat{f}_{\eta_i})$, sempre será aquele que é mais freqüente entre seus vizinhos \hat{f}_{η_i} . Assim, como áreas mais homogêneas são favorecidas, as descontinuidades presentes na imagem são desrespeitadas.

Nesse contexto, seria mais interessante que o algoritmo adaptasse seu comportamento para que quando uma descontinuidade fosse encontrada a suavização diminuísse, ou fosse desativada. Segundo LI (1995a), desde o trabalho pioneiro de GEMAN e GEMAN (1984), esse é um comportamento bastante adotado quando se caracteriza as imagens como MRFs.

Para que esse comportamento adaptativo seja possível, primeiramente é preciso detectar as descontinuidades presentes na imagem. Como discutido no Capítulo 3, isso é feito utilizando funções capazes de diferenciar entre mudanças de níveis de cinza que são causadas

pela presença de ruído, e mudanças causadas por descontinuidades da imagem. Além disso, quando uma descontinuidade é encontrada, tais funções também devem ser capazes de avaliar sua intensidade.

Como mostrado no trabalho de LI (1995a), funções possuindo esse comportamento crescem dentro de um certo intervalo, dentro do qual as mudanças de nível de cinza são consideradas ruído e por isso a suavização deve ocorrer, e decrescem fora desse intervalo, quando as mudanças de nível de cinza ocorrem devido a descontinuidades que devem ser respeitadas.

Nesse sentido, a função adotada por esta pesquisa é mostrada na Equação (5.13), sendo x a diferença entre f_i e cada um de seus vizinhos f_{η_i} . A forma qualitativa dessa função pode ser vista na Figura (5.9), mostrando que a função é convexa no intervalo $(-B, +B)$, e não convexa fora.

$$g_\gamma(x) = \gamma - \frac{\gamma}{1 + \frac{x^2}{\gamma}} \quad (5.13)$$

Adotando o sistema de vizinhança de segundo grau, a Equação (5.13) toma a forma da Equação (5.14).

$$\begin{aligned} g_\gamma(x) = & 8 * \gamma - \frac{\gamma}{1 + \frac{[x(i,j)-x(i-1,j-1)]^2}{\gamma}} - \frac{\gamma}{1 + \frac{[x(i,j)-x(i-1,j)]^2}{\gamma}} \\ & - \frac{\gamma}{1 + \frac{[x(i,j)-x(i-1,j+1)]^2}{\gamma}} - \frac{\gamma}{1 + \frac{[x(i,j)-x(i,j-1)]^2}{\gamma}} - \frac{\gamma}{1 + \frac{[x(i,j)-x(i,j+1)]^2}{\gamma}} \\ & - \frac{\gamma}{1 + \frac{[x(i,j)-x(i+1,j-1)]^2}{\gamma}} - \frac{\gamma}{1 + \frac{[x(i,j)-x(i+1,j)]^2}{\gamma}} - \frac{\gamma}{1 + \frac{[x(i,j)-x(i+1,j+1)]^2}{\gamma}} \end{aligned} \quad (5.14)$$

Detectadas as descontinuidades, essas são impostas quando o produto $p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$ é maximizado. Como já comentado, o valor que maximiza esse produto com certeza será aquele mais freqüente entre os vizinhos do *pixel* de interesse. Dessa forma, para maximizar esse produto, bastaria calculá-lo para o valor de cada um dos oito vizinhos de i .

A forma adotada para a imposição das descontinuidades se inicia calculando a função $g_\gamma(x)$, mostrada na Equação (5.14), para todos os *pixels* da estimativa de alta resolução. Possuindo tais valores para cada um dos *pixels* de alta resolução, o cálculo de $p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$ só será feito para os vizinhos que possuírem $g_\gamma(x)$ próximo ao valor dessa função para o *pixel* de interesse. Esse processo é ilustrado na Figura (5.10).

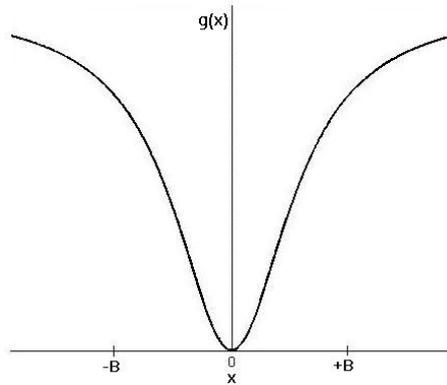


Figura 5.9: Forma qualitativa da função que detecta as descontinuidades mostrada na Equação (5.13).

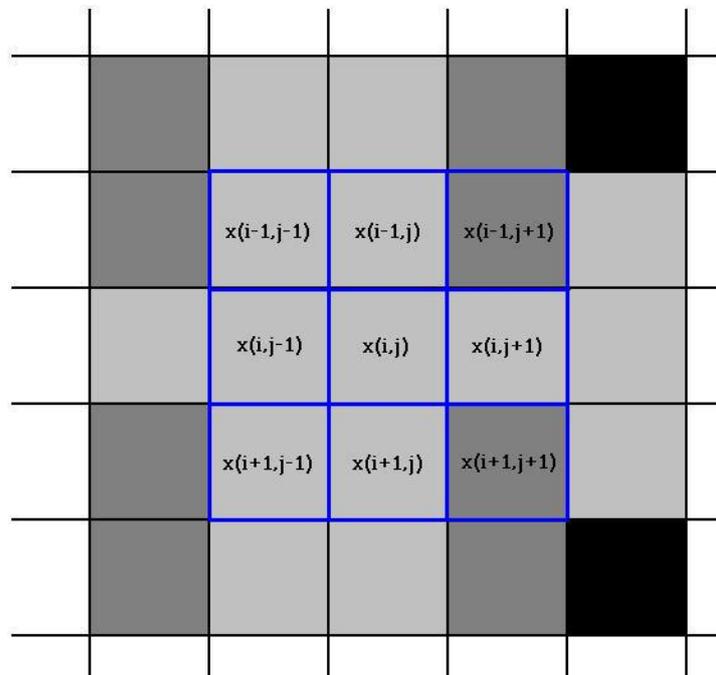


Figura 5.10: Ilustração do cálculo da probabilidade *a posteriori* local, utilizando a informação da função que detecta as descontinuidades.

Na Figura (5.10), o valor da função $g_\gamma(x)$ para $x(i, j)$ é bem próximo ou igual ao valor dessa função calculada para $x(i - 1, j - 1)$, $x(i - 1, j)$, $x(i, j - 1)$, $x(i, j + 1)$, $x(i + 1, j - 1)$ e $x(i + 1, j)$. Entretanto, esse valor será um pouco distante do valor da função calculada para $x(i - 1, j + 1)$ e $x(i + 1, j + 1)$. Assim, $p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$ só é calculado para $x(i - 1, j - 1)$, $x(i - 1, j)$, $x(i, j - 1)$, $x(i, j + 1)$, $x(i + 1, j - 1)$ e $x(i + 1, j)$, sendo que o valor do *pixel* de interesse $x(i, j)$ é substituído pelo valor que maximiza esse produto.

A cada iteração do algoritmo ICM, o mesmo processo é repetido para todos os *pixels* da estimativa de alta resolução. Visto isso, o algoritmo ICM adaptativo a descontinuidades ocorre da seguinte forma:

1. Inicializa-se \hat{f} com a estimativa inicial de alta resolução encontrada ao se registrar e interpolar as imagens de baixa resolução observadas;
2. Calcula-se $g_\gamma(x)$ para todos os *pixels* de \hat{f} ;
3. Assume-se um intervalo para o qual o valor de $g_\gamma(x)$ será considerado próximo ao valor dessa função para o *pixel* de interesse;
4. Para todos os *pixels* da estimativa de alta resolução, atualiza-se \hat{f}_i de forma a maximizar

$$p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$$

sendo que esse produto é calculado apenas para os vizinhos que possuem $g_\gamma(x)$ dentro do intervalo estipulado;

5. Repete-se os itens de (2) a (4) por um número máximo de iterações T_{iter} ou até que nenhuma alteração seja feita.

Aplicado de forma adaptativa a descontinuidade como explicado, o algoritmo ICM consegue melhorar a estimativa inicial de alta resolução com baixo custo computacional, já que o número de operações é reduzido. Isso acontece pelo fato dos cálculos serem feitos localmente para cada *pixel* considerando apenas os vizinhos desse *pixel*. Além disso, como são considerados apenas os vizinhos que possuem valores da função que detecta descontinuidades próximos ao valor dessa função para o *pixel* de interesse, $p(g_i|f_i) \cdot p(f_i|\hat{f}_{\eta_i})$ é calculada no máximo oito vezes para cada *pixel*.

5.7 Considerações Finais

O algoritmo ICM adaptativo a descontinuidades é aplicado a fim de melhorar a estimativa inicial de alta resolução, preservando as descontinuidades inicialmente presentes.

No próximo capítulo são discutidos os resultados alcançados em uma situação controlada, quando imagens de baixa resolução foram geradas, registradas e interpoladas. Além disso, o método proposto foi aplicado a uma situação real, utilizando *frames* de um vídeo. Utilizou-se o algoritmo ICM adaptativo a descontinuidades e o algoritmo ICM puro a fim de comparar e discutir a influência da adaptação a descontinuidades nas imagens resultantes.

Resultados

6.1 Considerações Iniciais

Como discutido no capítulo anterior, o algoritmo ICM adaptativo a descontinuidades é aplicado a fim de melhorar a estimativa de alta resolução preservando as descontinuidades inicialmente detectadas. Esta pesquisa adotou como estimativa inicial para o algoritmo, a imagem de alta resolução resultante do processo de registro e interpolação das imagens observadas.

O método foi aplicado em uma situação controlada, onde imagens possuindo deslocamentos *sub-pixel* uma em relação às demais foram geradas. Além disso, foram utilizados *frames* de um vídeo em um estudo de caso, a fim de gerar uma imagem de alta resolução aplicando o método proposto. A seguir a situação controlada e o estudo de caso são detalhados, discutindo e comparando os resultados com a aplicação do algoritmo ICM puro.

6.2 Descrição da Simulação

Primeiramente uma imagem foi utilizada para gerar as várias imagens de baixa resolução da mesma cena, que foram registradas e interpoladas a fim de gerar a primeira estimativa de alta resolução. Dessa forma, utilizou-se a imagem “Lena” mostrada na Figura (6.1(a)). A

essa imagem, considerada aqui como a cena original, foi aplicado borramento por uma janela uniforme 3x3 e ruído Gaussiano a 40dB, resultando na imagem mostrada na Figura (6.1(b)).



Figura 6.1: Imagem “Lena” de dimensão 512x512, usada para gerar as várias imagens de baixa resolução.

De acordo com o modelo de formação das imagens discutido no Capítulo 5, os deslocamentos *sub-pixel* entre as imagens de baixa resolução foram gerados durante o processo de sub-amostragem. Isso foi feito como mostra a Figura (6.2).

Supondo que o tamanho do *pixel* de alta resolução seja $L \times L$, assumiu-se que cada *pixel* de baixa resolução possui tamanho $4L \times 4L$. Dessa forma, a primeira imagem de baixa resolução foi adquirida alinhando perfeitamente a grade de baixa resolução com a grade de alta resolução, como mostra a primeira imagem da Figura (6.2), sendo a grade de alta resolução mostrada em vermelho e a grade de baixa resolução mostrada em azul. O valor de cada *pixel* de baixa resolução é encontrado segundo a Equação (6.1), como explicado no Capítulo 5, onde d_n é o *pixel* de baixa resolução e $\delta_{n,m}$ o peso relacionando o *pixel* de alta resolução f_m ao *pixel* de baixa resolução d_n .

$$d_n = \sum_m \delta_{n,m} \cdot f_m \quad (6.1)$$

Foram geradas dezesseis imagens de baixa resolução da mesma forma, porém deslocando a grade de baixa resolução, como mostram as demais imagens da Figura (6.2). As dezesseis imagens geradas são mostradas na Figura (6.3).

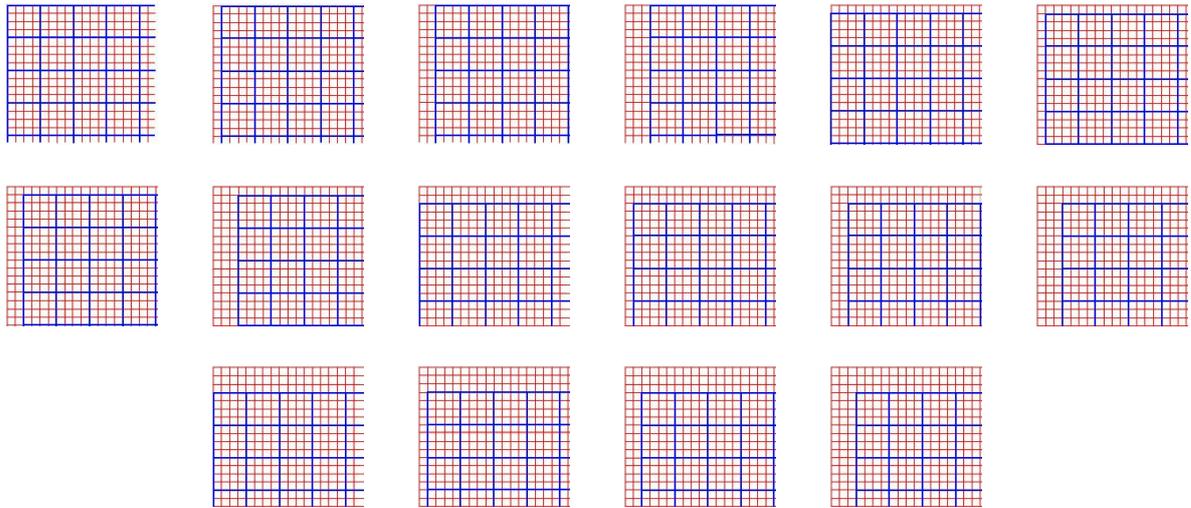


Figura 6.2: Ilustração mostrando o processo de sub-amostragem da imagem de alta resolução, dando origem a 16 imagens de baixa resolução.



Figura 6.3: Imagens de baixa resolução de dimensão 128x128, geradas utilizando a imagem “Lena” borrada e ruidosa.

Geradas as imagens de baixa resolução, essas foram registradas como explicado no Capítulo 5, chegando-se à primeira estimativa de alta resolução mostrada na Figura (6.4), aumentando a resolução por um fator 4. Como os deslocamentos presentes entre as imagens de baixa

resolução simuladas são equidistantes e uniformemente espaçados, não foi necessário aplicar nenhum procedimento interpolatório nos *pixels* de baixa resolução para se chegar à grade de alta resolução. Eles já estavam perfeitamente dispostos na grade de alta resolução – Figura (5.7).

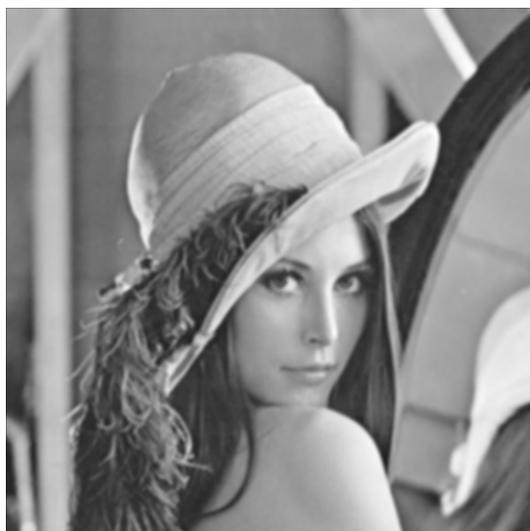


Figura 6.4: Imagem de dimensão 512x512, resultante do registro e interpolação das 16 imagens de baixa resolução simuladas.

6.3 Análise dos Resultados

A imagem resultante do registro e interpolação das 16 imagens de baixa resolução simuladas foi utilizada como estimativa inicial de alta resolução do algoritmo ICM puro, onde as descontinuidades detectadas não foram impostas no processo. A Figura (6.5) mostra um recorte da imagem registrada – Figura (6.5(a)) – ao lado de um recorte da imagem resultante da décima iteração do algoritmo ICM puro – Figura (6.5(b)). Nota-se que, devido à característica do algoritmo de considerar a classe mais freqüente entre os vizinhos do *pixel* como a de maior probabilidade, as descontinuidades de menor intensidade foram desrespeitadas. Dessa forma surgiram regiões homogêneas na imagem, dando a impressão de que essa possuía um número insuficiente de níveis de cinza. Esse problema é denominado na literatura como *Falsos Contornos*.

Na aplicação do algoritmo ICM adaptativo a descontinuidades, primeiramente utilizou-se a função dada pela Equação (5.14) para detectar as descontinuidades, com valor inicial de $\gamma = 300$ decaindo a cada iteração por um fator de 0,95. Isso foi feito pelo fato de que a função mostrada na Equação (5.14) não é convexa, e conseqüentemente a função custo também não é convexa. Para aliviar o efeito de mínimos locais, torna-se a função custo aproximadamente con-

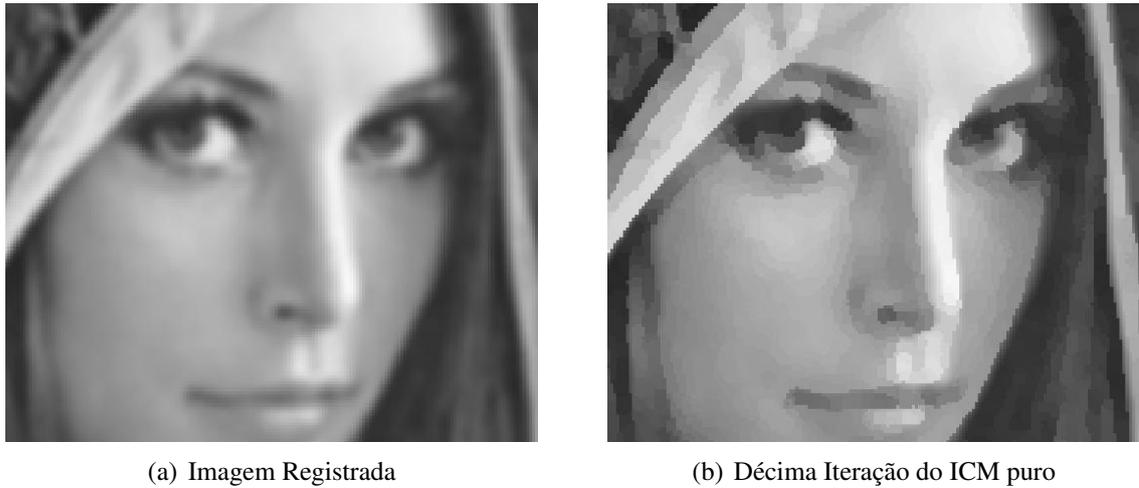


Figura 6.5: Imagem resultante da aplicação do algoritmo ICM puro, sem a imposição das descontinuidades.

vexa inicialmente escolhendo um valor arbitrariamente alto para γ . À medida que as iterações acontecem, o efeito de suavização decai, preservando as descontinuidades na imagem. Além disso, assumiu-se, de maneira experimental, o valor 0,01 para a variância σ^2 da distribuição de probabilidade condicional dos *pixels* de baixa resolução, dado o *pixel* de alta resolução $p(g_i|f_i)$ – Equação (5.12).

A Figura (6.6) mostra um recorte da imagem registrada – Figura (6.6(a)) – ao lado de um recorte da imagem resultante da décima iteração do algoritmo ICM adaptativo a descontinuidades – Figura (6.6(b)). Como pode ser visto, a formação das áreas homogêneas não ocorreu e as descontinuidades foram preservadas.

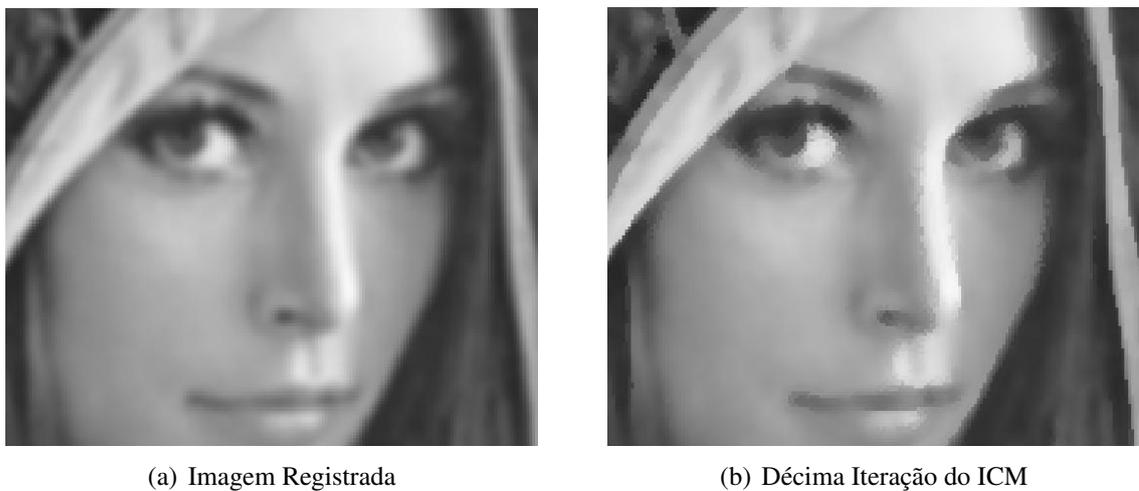


Figura 6.6: ICM adaptativo a descontinuidade aplicado à imagem Lena registrada.

A Figura (6.7) mostra um recorte da imagem resultante da décima iteração do ICM adaptativo a descontinuidades – Figura (6.7(a)) – ao lado de um recorte da imagem resultante da aplicação da interpolação bilinear à imagem de baixa resolução adotada como imagem de referência – Figura (6.7(b)). Nota-se que a imagem resultante da aplicação do ICM adaptativo a descontinuidades possui detalhes muito mais nítidos.



Figura 6.7: Comparação do ICM adaptativo a descontinuidade com a Interpolação Bilinear aplicada a uma das imagens de baixa resolução.

Para avaliar numericamente os resultados alcançados, foi utilizado o índice de qualidade *Universal Image Quality Index*, proposto por WANG e BOVIK (2002). Ao contrário dos métodos de avaliação tradicionais que normalmente utilizam o erro médio quadrático (MSE - *Mean Squared Error*) para avaliar as imagens, esse índice de qualidade modela qualquer distorção como uma combinação de três fatores – perda de correlação, distorção de luminância e distorção de contraste. Os autores mostraram algumas imagens com MSEs praticamente idênticos, e que possuíam distorções distintas. Tais imagens foram ordenadas de forma empírica, apenas considerando sua qualidade visual. Ao aplicar o índice de qualidade proposto, a mesma ordenação foi alcançada, mostrando que essa poderia ser uma forma mais interessante de avaliar a qualidade da imagem.

Além disso, o MSE utilizado como forma de avaliação, favorece a suavidade presente nas imagens. Acredita-se que no contexto do aumento de resolução por meio de técnicas de Reconstrução por Super Resolução, favorecer a suavidade seria semelhante a favorecer a perda de detalhamento inerente à perda de resolução das imagens observadas.

Seja $x = \{x_i | i = 1, 2, \dots, N\}$ a imagem original e $y = \{y_i | i = 1, 2, \dots, N\}$ a imagem a ser avaliada, o *Universal Image Quality Index* é dado pela Equação (6.2), sendo \bar{x} , \bar{y} , σ_x^2 , σ_y^2 e σ_{xy} dados pelas Equações (6.3) a (6.7).

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2) [(\bar{x})^2 + (\bar{y})^2]} \quad (6.2)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (6.3)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (6.4)$$

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (6.5)$$

$$\sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2 \quad (6.6)$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (6.7)$$

Q assume valores no intervalo $[-1, 1]$, sendo igual a 1 apenas quando $x_i = y_i$ para todo $i = 1, 2, \dots, N$, e igual a -1 quando $y_i = 2\bar{x} - x_i$ para todo $i = 1, 2, \dots, N$. Reescrevendo a Equação (6.2) de forma a esclarecer a avaliação por meio dos três fatores citados, chega-se à Equação (6.8).

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (6.8)$$

O primeiro termo da Equação (6.8) é dado pelo coeficiente de correlação que mede o grau de correlação linear entre x e y , assumindo valores no intervalo $[-1, 1]$. O segundo termo mede o quão próxima a luminância média está entre x e y , assumindo valores entre $[0, 1]$. O terceiro termo aponta o grau de similaridade no contraste das imagens, assumindo valores no intervalo $[0, 1]$.

A Tabela (6.1) mostra os valores de Q calculados para as imagens registrada – Figura (6.5(a)) –, resultante da décima iteração do ICM puro – Figura (6.5(b)) –, resultante da décima iteração do ICM adaptativo a descontinuidades – Figura (6.6(b)) –, e resultante da aplicação da interpolação bilinear na imagem de baixa resolução considerada como imagem de referência – Figura (6.7(b)). Além disso, o índice de qualidade é calculado para a imagem resultante da décima iteração do algoritmo proposto por IRANI e PELEG (1991), já que trata-se de um dos

trabalhos utilizados como referência entre os métodos propostos no contexto da Reconstrução por Super Resolução.

Tabela 6.1: Índice de qualidade *Universal Image Quality Index* aplicado às imagens com relação à imagem original.

Imagem Registrada	0.9449
ICM puro	0.9108
ICM adaptativo a descontinuidade	0.9686
Interpolação Bilinear	0.9458
Irani-Peleg	0.9904

Os valores apresentados pela Tabela (6.1) estão de acordo com a simples inspeção visual das imagens, classificando a imagem resultante da aplicação do algoritmo ICM adaptativo a descontinuidades como a de melhor qualidade quando comparado às imagens registrada, resultante da interpolação, e a resultante da aplicação do ICM puro. Além disso, devido à formação das áreas homogêneas inerentes ao algoritmo ICM puro, a imagem resultante de sua aplicação é avaliada como de pior qualidade.

A Figura (6.8), mostra o recorte da imagem resultante da aplicação da décima iteração do método proposto por IRANI e PELEG (1991). Apesar dessa imagem ter apresentado o melhor índice de qualidade, esse índice foi calculado com relação à imagem original mostrada na Figura (6.1(a)). Como os algoritmos não aplicaram nenhum método para retirar o borramento presente nas imagens observadas, o ideal seria calcular tal índice com relação à imagem borrada.



Figura 6.8: Recorte da imagem resultante da décima iteração do algoritmo proposto por IRANI e PELEG (1991).

A Tabela (6.2) mostra os valores de Q comparando as imagens com a imagem borrada – Figura (6.1(b)). Nota-se que o algoritmo ICM adaptativo a descontinuidades apresentou um

valor bastante próximo daquele apresentado pelo algoritmo proposto por IRANI e PELEG (1991).

Tabela 6.2: Índice de qualidade *Universal Image Quality Index* aplicado às imagens com relação à imagem borrada.

Imagem Registrada	0.9660
ICM puro	0.9331
ICM adaptativo a descontinuidade	0.9875
Interpolação Bilinear	0.9668
Irani-Peleg	0.9897

A fim de comparar os valores alcançados pelo *Universal Image Quality Index*, aplicou-se o erro médio quadrático normalizado (NMSE - *Normalized Mean Square Error*) (JAIN, 1988) nas imagens – Equação (6.9).

$$NMSE = \frac{\|x - y\|^2}{\|x\|^2} = \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N x_i^2} \quad (6.9)$$

A Tabela (6.3) mostra os valores do NMSE comparando as imagens com a imagem original – Figura (6.1(a)). Assim como os resultados da aplicação do índice de qualidade proposto por WANG e BOVIK (2002), o NMSE é menor para a imagem resultante da aplicação do algoritmo ICM adaptativo a descontinuidades quando comparado às outras imagens, exceto com relação à imagem resultante da aplicação do método proposto por IRANI e PELEG (1991).

Tabela 6.3: NMSE calculado para as imagens com relação à imagem original.

Imagem Registrada	0.1557
ICM puro	0.1818
ICM adaptativo a descontinuidade	0.1332
Interpolação Bilinear	0.1569
Irani-Peleg	0.0501

Os valores do NMSE com relação à imagem borrada, são mostrados na Tabela (6.4). Tais valores estão de acordo com aqueles apresentados pelo índice de qualidade, sendo a imagem resultante do algoritmo ICM adaptativo a descontinuidades bem próxima da resultante do algoritmo proposto por IRANI e PELEG (1991).

Apesar de apresentar melhores resultados tanto do índice de qualidade, quanto do NMSE, a imagem resultante do método proposto por IRANI e PELEG (1991) está um pouco mais escura quando comparada à imagem borrada. O mesmo não acontece com a imagem resultante do algoritmo ICM adaptativo a descontinuidades.

Tabela 6.4: NMSE calculado para as imagens com relação à imagem borrada.

Imagem Registrada	0.1439
ICM puro	0.1709
ICM adaptativo a descontinuidade	0.0577
Interpolação Bilinear	0.1445
Irani-Peleg	0.0209

Além disso, o algoritmo ICM, assim como foi feita a imposição das descontinuidades detectadas, permite que outras informações previamente conhecidas a respeito da cena original sejam impostas no processo de estimação. Essa imposição não é possível no algoritmo proposto por IRANI e PELEG (1991).

6.4 Estudo de Caso

A fim de aplicar o método proposto em uma situação real, foram utilizados *frames* de um vídeo adquirido com uma câmera Sony Cyber-shot DSC-S90 a 30 quadros por segundo, com resolução de 640x480.

A Figura (6.9) mostra os quatro *frames*, de dimensão 128x128 *pixels*, utilizados na aplicação do método proposto. Essas imagens foram registradas chegando à primeira estimativa de alta resolução utilizada como entrada para o algoritmo ICM adaptativo a descontinuidades. Assim como na simulação, não foi utilizado nenhum procedimento interpolatório para se chegar à primeira estimativa de alta resolução. Os *frames* foram tomados de forma que os deslocamentos entre eles fossem equidistantes e uniformemente espaçados – situação ilustrada na Figura (5.7).

**Figura 6.9:** *Frames*, de dimensão 128x128 *pixels*, utilizados para avaliar a aplicação do método proposto em uma situação real.

A Figura (6.10) mostra as imagens resultantes da décima iteração do algoritmo proposto por IRANI e PELEG (1991) e do algoritmo ICM adaptativo a descontinuidades. Em ambos os casos a resolução foi aumentada por um fator de dois, chegando a imagens de dimensão

256x256 *pixels*. Novamente, apesar da imagem resultante do algoritmo de IRANI e PELEG (1991) possuir aparentemente maior resolução, a imagem se mostra mais escura que a alcançada pelo ICM adaptativo a descontinuidades.

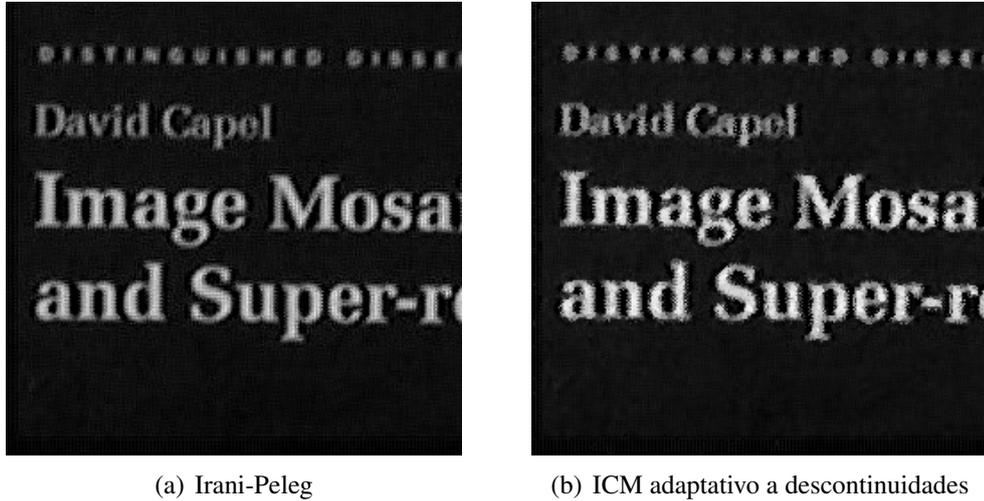


Figura 6.10: Imagens resultantes da décima iteração do algoritmo proposto por Irani e Peleg, e do algoritmo ICM adaptativo a descontinuidades.

6.5 Considerações Finais

Ao se utilizar a imagem resultante do registro e interpolação das imagens de baixa resolução observadas como estimativa inicial de alta resolução no algoritmo ICM adaptativo a descontinuidades, foi possível reconstruir uma imagem de alta resolução que respeitasse as descontinuidades inicialmente presentes, e que apresentasse maior riqueza de detalhes.

No próximo capítulo é feita uma discussão quanto às expectativas e resultados alcançados e são discutidas as contribuições do trabalho e possíveis trabalhos futuros.

Conclusão e Trabalhos Futuros

7.1 Considerações Finais

A abordagem proposta apresentou baixo custo computacional, com relação ao número de operações realizadas para cada *pixel* da imagem a cada iteração do algoritmo. A maioria das abordagens que utilizam MRFs na caracterização das imagens, modelam tais imagens com a distribuição de Gibbs. Quando isso acontece, a cada *pixel* da imagem calcula-se a função energia considerando a imagem como um todo, realizando um número elevado de operações para cada *pixel*.

Quando se utiliza o modelo de Potts-Strauss na modelagem da imagem de alta resolução, cada *pixel* considera apenas o valor de seus vizinhos, dado o sistema de vizinhança adotado. Dessa forma, a menos que o sistema de vizinhança seja $\eta_a = S - \{a\}$, ou seja, todos os *pixels* da imagem menos o *pixel* de interesse, o número de operações executadas para cada *pixel* será bastante reduzido.

Além disso, o algoritmo ICM converge para um máximo local em poucas iterações. Na simulação descrita no Capítulo 6, nenhuma modificação é feita a partir da décima iteração.

Entretanto, como se pode notar, o sistema de vizinhança adotado é fator determinante do número de operações executadas pelo método. Assim, a escolha desse deve ser feita de forma a melhor balancear o número de operações e a perda de informações de contexto. Para a

abordagem adotada, o sistema de vizinhança de segundo grau mostrou-se empiricamente a melhor opção.

Da mesma forma, a detecção das descontinuidades também é dependente do sistema de vizinhança adotado. Ao se utilizar uma função responsável por essa detecção, também é preciso balancear o número de operações e a perda de informações de contexto. Para a função adotada por essa pesquisa – Equação (5.14) – o sistema de vizinhança de segundo grau também se mostrou uma boa opção.

Com relação à qualidade da imagem reconstruída, essa é dependente do número de imagens observadas e dos deslocamentos estimados entre elas. Isso acontece porque o algoritmo ICM é dependente da imagem de alta resolução que ele utiliza como estimativa inicial. Caso existam poucas imagens disponíveis, a imagem resultante do registro e interpolação de tais imagens provavelmente não será uma boa estimativa de alta resolução, e o método não chegará a bons resultados. O mesmo acontece se o registro aplicado às imagens observadas não encontra estimativas dos deslocamentos que sejam próximas à realidade.

Sendo a qualidade da imagem reconstruída, dependente da estimativa inicial de alta resolução utilizada, poder-se-ia pensar em melhorar tal estimativa. A abordagem proposta não aplicou nenhum método de restauração nas imagens observadas a fim de atenuar as degradações invariavelmente inseridas durante o processo de aquisição e manipulação dessas imagens.

Aplicado sem englobar a etapa de restauração das imagens, o método proposto é independente das degradações existentes, que podem variar dependendo das condições em que as imagens foram adquiridas. Entretanto, implementar essa etapa seria natural, bastando apenas adotar o método de restauração mais adequado para as imagens com que se está trabalhando.

Para avaliar numericamente a qualidade da estimativa de alta resolução alcançada, considerou-se o índice de qualidade *Universal Image Quality Index* mais adequado que o MSE. Isso se deve ao fato do MSE não ser capaz de captar adequadamente a correlação entre as imagens original e a que está sendo avaliada.

A seguir são discutidas as contribuições e possibilidades de continuidade desta pesquisa.

7.2 Contribuições

Os resultados deste trabalho foram publicados no *II Workshop de Visão Computacional (WVC2006)* que aconteceu de 16 a 18 de outubro de 2006 na Escola de Engenharia de São Carlos da Universidade de São Paulo (USP/EESC), sob o título “*High-Resolution Image Reconstruction with the ICM Algorithm*” (HOMEM et al., 2006).

Recentemente outro trabalho foi aceito para publicação no *IEEE International Conference on Image Processing* edição de 2007 (ICIP2007) a ser realizado nos dias 16 a 19 de setembro de 2007 na cidade de San Antonio, Texas, EUA, sob o título “*Super-Resolution Image Reconstruction using the ICM Algorithm*” (MARTINS et al., 2007).

Outros artigos discutindo a influência da adaptação a descontinuidades no algoritmo ICM estão sendo elaborados, sendo que um deles será submetido ao *Brazilian Symposium on Computer Graphics and Image Processing* edição 2007 (SIBGRAPI2007), caso seja finalizado em tempo.

As duas publicações citadas se encontram em anexo a esta dissertação.

7.3 Trabalhos Futuros

Existem inúmeras aplicações que necessitam de maior resolução nas imagens de interesse, e inúmeras possibilidades que dispõem de várias imagens da mesma cena com algum tipo de incerteza entre elas. Nesse contexto, o método proposto, ou seu aprimoramento, poderia ser explorado de forma a se adequar da melhor maneira possível. Para esse aprimoramento, um primeiro passo seria tornar o método automático como explicado a seguir.

7.3.1 Uma Ferramenta para Super Resolução

Todas as etapas empregadas no processo de Reconstrução por Super Resolução utilizando o algoritmo ICM adaptativo a descontinuidades foram implementadas separadamente. Como explicado, primeiro as imagens observadas foram registradas a fim de estimar os deslocamentos *sub-pixel* presentes entre elas. Depois essas imagens foram interpoladas de acordo com os deslocamentos identificados. E, por fim, o algoritmo ICM adaptativo a descontinuidades foi aplicado a fim de melhorar a estimativa de alta resolução.

Seria conveniente que todas essas etapas fossem aplicadas de forma automática, por meio de uma ferramenta que recebesse como entrada um vídeo, ou as várias imagens da mesma cena, e retornasse uma ou várias imagens de resolução mais alta do que as imagens de entrada.

Dessa forma, pretende-se estudar a melhor maneira de automatizar todo o processo de Reconstrução por Super-Resolução, com a possibilidade do usuário manipular alguns parâmetros para maior flexibilidade da ferramenta.

7.3.2 Projeto HERON

O Linha 1, inserido no Grupo de Ciências da Linguagem do CLC (Centro de Línguas e Culturas) da Universidade de Aveiro, Portugal, é um grupo de trabalho que tem como principal objetivo o estudo da variação linguística do Português em falantes nativos e estrangeiros. O grupo está envolvido em projetos em colaboração com o IEETA (Instituto de Engenharia Electrónica e de Telemática de Aveiro), sendo um deles o projeto HERON – Um ambiente computacional para investigação em síntese articulatória do Português.

A síntese articulatória gera o sinal de voz através da modelagem das características físicas, anatômicas e fisiológicas do aparelho humano produtor de voz. A grande diferença desse sistema para outros, é que nesta técnica se modela diretamente o aparelho produtor de voz em lugar de se modelar o sinal ou as suas características acústicas.

Dentre outras, uma das tarefas desempenhadas pelo sintetizador da Universidade de Aveiro é a análise de imagens de Ressonância Magnética do trato vocal para obtenção de novos dados anatômicos, necessários para o desenvolvimento e avaliação dos modelos articulatórios. Essa é uma forma de adquirir informação dos articuladores do trato vocal em tempo real, já que em geral o trato vocal não é facilmente acessível, as estruturas biológicas envolvidas são complexas, e os articuladores se movem relativamente rápido (AVILA-GARCÍA et al., 2004).

Dessa forma, várias imagens são obtidas durante a emissão da fala de uma palavra, sendo que o momento exato da aquisição da imagem não é fixado. Essa incerteza abre a possibilidade de se utilizar imagens de diferentes aquisições para se estimar imagens de resolução mais alta que as observadas, por meio dos métodos de Reconstrução por Super Resolução. Isso é possível porque a informação fornecida pelas diversas aquisições não é redundante.

Além das imagens adquiridas durante a fala de uma palavra, existe a disponibilidade de imagens de fonemas isolados, o que pode facilitar a tarefa de se segmentar uma seqüência de imagens a ser objeto do processo de Super Resolução.

As Figuras (7.1) e (7.2) mostram algumas das imagens de ressonância magnética do corte transversal e oblíquo, respectivamente, do trato vocal humano, cedidas pelo Instituto de Engenharia Electrónica e Telemática de Aveiro.

O método desenvolvido por esta pesquisa será aprimorado para aumentar a resolução de tais imagens em parceria com o IEETA, no processo de doutorado pelo Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Carlos (PPGBiotec UFSCar) sob orientação do Prof. Dr. Nelson Delfino D'Ávila Mascarenhas e co-orientação do Prof. Dr. Cláudio Alberto Torres Suazo.

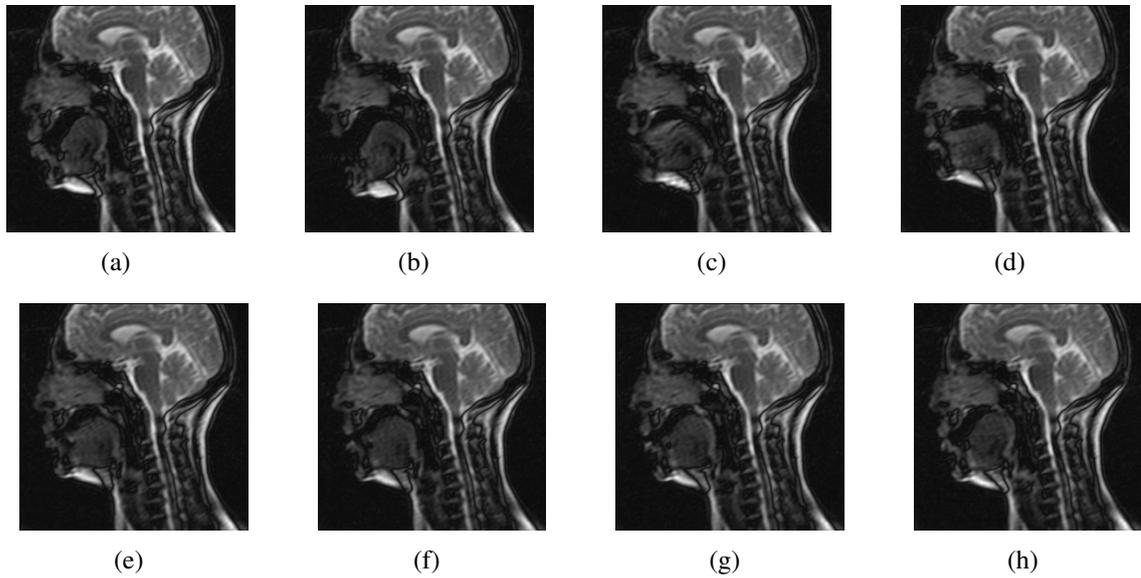


Figura 7.1: Imagens de ressonância magnética de um corte transversal do trato vocal humano, cedidas pelo Instituto de Engenharia Eletrônica e Telemática de Aveiro.

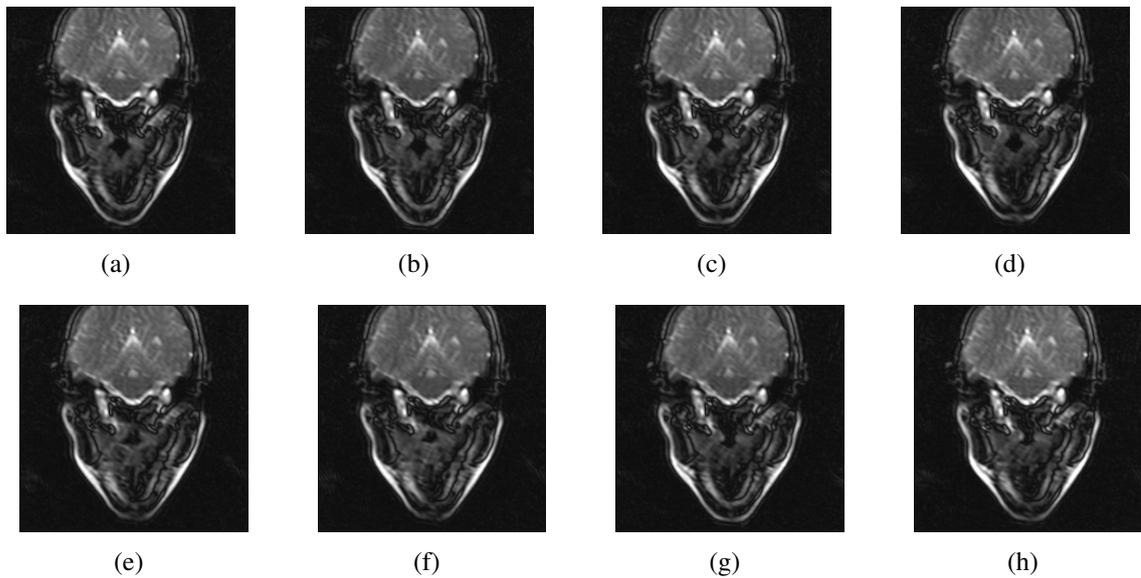


Figura 7.2: Imagens de ressonância magnética de um corte oblíquo do trato vocal humano, cedidas pelo Instituto de Engenharia Eletrônica e Telemática de Aveiro.

Referências Bibliográficas

- ALTUNBASAK, Y.; PATTI, A. J. A Maximum a Posteriori estimator for high resolution video reconstruction from MPEG video. *Proceedings of 2000 IEEE International Conference of Image Processing*, p. 649–652, 2000.
- AVILA-GARCÍA, M. S.; CARTER, J. N.; DAMPER, R. I. Extracting Tongue Shape Dynamics from Magnetic Resonance Image Sequences. *Transactions on Engineering, Computing and Technology*, v. 2, p. 288–291, 2004.
- BERTHOD, M.; KATO, Z.; YU, S.; ZERUBIA, J. Bayesian image classification using Markov random fields. *Image Vision Comput.*, v. 14, n. 4, p. 285–295, 1996.
- BESAG, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, v. 36, n. 2, p. 192–236, 1974.
- BESAG, J. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, v. 48, p. 259–302, series B, 1986.
- BHATT, M. R.; DESAI, U. B. Robust image restoration algorithm using Markov random field model. *CVGIP: Graph. Models Image Process.*, v. 56, n. 1, p. 61–74, 1994.
- BIEMOND, J.; LAGENDIJK, R. L.; MERSEREAU, R. S. Iterative Methods for Image Deblurring. *Proceedings of the IEEE*, v. 78, n. 5, 1990.
- BORMAN, J. L.; STEVENSON, R. *Spatial resolution enhancement of low-resolution image sequences - A comprehensive review with directions for future research*. Relatório Técnico, University of Notre Dame, Notre Dame, IN, 1998.
- CIPRA, B. A. An introduction to the Ising model. *American Mathematical Monthly*, v. 94, n. 10, p. 937–959, 1987.
- CORTIJO, F.; VILLENA, S.; MOLINA, R.; KATSAGGELOS, A. Bayesian super-resolution of text image sequences from low resolution observations. *Proceedings of the Seventh International Symposium on Signal Processing and Its Applications*, p. 421–424, 2003.

- DERIN, H.; ELLIOT, H. Modelling and Segmentation of noisy and texture images using Gibbs Random Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 9, n. 1, p. 39–55, 1987.
- DUBES, R. C.; JAIN, A. K. Random field models in image analysis. *Advances in Applied Statistics*, v. 16, p. 121–154, 1993.
- DUDA, R. O.; HART, P. E. *Pattern Classification and Scene Analysis*. John Wiley & Sons Inc, 1973.
- ELAD, M.; HEL-OR, Y. Fast Super-Resolution Reconstruction Algorithm for Pure Translational Motion and Common Space-Invariant Blur. *IEEE Trans. on Image Processing*, v. 10, n. 8, p. 1187–1193, 2001.
- FREEMAN, W.; JONES, T. R.; PASZTOR, E. C. Example-based super-resolution. *IEEE Computer Graphics and Applications*, v. 22, n. 2, p. 56–65, 2002.
- FRERY, A. C. *Algumas ferramentas estatísticas na síntese, processamento e análise de imagens de radar de abertura sintética*. Tese de Doutorado, INPE - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, 1993.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 6, n. 6, p. 721–741, 1984.
- GONZALEZ, A. A.; WOODS, R. E. *Digital Image Processing*. Addison-Wesley Longman Publishing, 2001.
- GUNTURK, B. K.; ALTUNBASAK, Y.; MERSEREAU, R. M. Multiframe resolution-enhancement methods for compressed video. *IEEE Signal Processing Letters*, p. 170–174, 2002.
- HAMMERSLEY, J. M.; CLIFFORD, P. Markov field on finite graphs and lattices, 1971.
- HARDIE, R. C.; BARNARD, K. J.; ARMSTRONG, E. A. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, v. 6, n. 12, p. 1621–1633, 1997.
- HERTZMANN, A.; JACOBS, C. E.; CURLESS, B.; SALESIN, D. H. Image Analogies. *SIGGRAPH'01: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive techniques*, p. 327–340, new York, NY, USA: ACM Press, 2001.
- HOMEM, M. R. P.; MARTINS, A. L. D.; MASCARENHAS, N. D. A. High-Resolution Image Reconstruction with the ICM Algorithm. In: *WVC2006 - II Workshop de Visão Computacional*, São Carlos, 2006.
- HUMBLLOT, F.; MOHAMMAD-DJAFARI, A. Super-Resolution Using Hidden Markov Model and Bayesian Detection Estimation Framework. *EURASIP Journal on Applied Signal Processing*, v. 2006, p. 1–16, 2006.

- IRANI, M.; PELEG, S. Improving resolution by image registration. *CVGIP: Graph. Models Image Process.*, v. 53, n. 3, p. 231–239, 1991.
- JAIN, A. K. *Fundamentals of Digital Image Processing*. Prentice Hall, 1988.
- JOSHI, M. V.; CHAUDHURI, S. Simultaneous estimation of super-resolved depth map and intensity field using photometric cue. *Computer Vision and Image Understanding*, v. 101, p. 31–44, 2006.
- JOSHI, M. V.; CHAUDHURI, S.; PANUGANTI, R. Super-resolution imaging: use of zoom as a cue. *Image and Vision Computing*, v. 22, n. 14, p. 1185–1196, 2004.
- KANG, M. G.; CHAUDHURI, S. Super-resolution image reconstruction. *IEEE Signal Processing Magazine*, v. 20, n. 3, p. 19–20, 2003.
- KIM, S. P.; BOSE, N. K.; VALENZUELA, H. M. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Trans. on Acoustics, Speech and Signal Processing*, v. 18, p. 1013–1027, 1990.
- KIM, S. P.; SU, W. Y. Recursive high resolution reconstruction of blurred multiframe images. *IEEE Trans. on Image Processing*, v. 2, p. 534–539, 1993.
- KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. *Neurocomputing: foundations of research*, p. 551–567, Cambridge, MA, USA: MIT Press, 1988.
- KOMATSU, T.; IGARASHI, T.; AIZAWA, K.; SAITO, T. Very high resolution imaging scheme with multiple different-aperture cameras. *Signal Processing: Image Communication*, v. 5, p. 511–526, 1993.
- LI, S. Z. *Markov random field modeling in computer vision*. London, UK: Springer-Verlag, 1995a.
- LI, S. Z. On discontinuity-adaptive smoothness priors in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. PAM-17, n. 6, p. 576–586, 1995b.
- MARTINS, A. L. D.; HOMEM, M. R. P.; MASCARENHAS, N. D. A. Super-Resolution Image Reconstruction using the ICM Algorithm. In: *Proceedings of the 2007 IEEE International Conference on Image Processing*, San Antonio, TX, (aceito para publicação), 2007.
- MASCARENHAS, N. D. A.; FRERY, A. C. SAR image filtering with the ICM algorithm. *International Geoscience and Remote Sensing Symposium*, p. 2185–2187, 1994.
- METROPOLIS, N.; ROSENBLUTH, A.; TELLER, A.; TELLER, E. Equation of state calculations by fast computing machines. *J. of Chem. Physics*, v. 21, p. 1087–1092, 1953.
- MONTGOMERY, D. C.; RUNGER, G. C. *Estatística Aplicada e Probabilidade para Engenheiros*. 2003.

- NGUYEN, N.; MILANFAR, P.; GOLUB, G. A computationally efficient super-resolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, v. 10, p. 573–583, 2001.
- PARK, S. C.; PARK, M. K.; KANG, M. G. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, v. 20, n. 3, p. 21–36, 2003.
- RAJAN, D.; CHAUDHURI, S. Generation of super-resolution images from blurred observation using Markov Random Fields. *IEEE International Conference on Acoustics Speech and Signal Processing*, p. 1837–1840, 2001.
- RAJAN, D.; CHAUDHURI, S. Data fusion techniques for super-resolution imaging. *Information Fusion*, v. 3, n. 1, p. 25–38, 2002.
- RUBERT, C.; FONSECA, L.; VELHO, L. Super-resolução de imagens de sensores remotos usando a transformada de cor YUV para a estimação das altas frequências. IV Workshop dos Cursos da Computação Aplicada do INPE - IV WORCAP, 2004.
- SEGALL, C. A.; KATSAGGELOS, A. K.; MOLINA, R.; MATEOS, J. Bayesian resolution enhancement of compressed video. *IEEE Transactions on Image Processing*, v. 13, n. 7, p. 898–911, 2004.
- SEGALL, C. A.; MOLINA, R.; KATSAGGELOS, A. K. Reconstruction of high-resolution image frames from a sequence of low-resolution and compressed observations. *Proceedings of 2002 IEEE International Conference of Acoustics, Speech and Signal Processing*, p. 1701–1704, 2002.
- SROUBEK, F.; FLUSSER, J. Resolution enhancement via probabilistic deconvolution of multiple degraded images. *Pattern Recognition Letters*, v. 27, n. 4, p. 287–293, 2006.
- STARK, H.; OSKOU, P. High-resolution image recovery from image-plane arrays, using convex projections. *Journal of Optical Society America A*, v. 6, n. 11, p. 1715–1726, 1989.
- STEPHENSON, T. A.; TSUHAN, C. Adaptive Markov Random Fields for Example-Based Super-Resolution of Faces. *EURASIP Journal on Applied Signal Processing*, v. 2006, p. 1–11, 2006.
- SURESH, K.; RAJAGOPALAN, A. A Robust and Computationally Efficient Super-resolution Algorithm. *Journal of the Optical Society of America A*, v. 24, n. 4, p. 984–992, 2007.
- TEKALP, A.; OZKAN, M.; SEZAN, M. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. *IEEE Int. Conf. Acoust., Speech, and Signal Process.*, v. III, p. 169–172, 1992.
- TIKHONOV, A. N.; ARSENIN, V. A. *Solutions of ill-posed problems*. Washington: Winston & Sons, 1977.

- TOM, B.; KATSAGGELOS, A. K. Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, p. 2539–2542, 1995.
- TOM, B. C.; GALATSANOS, N. P.; KATSAGGELOS, A. K. *Super-Resolution Imaging*, cáp. 4 Kluwer Academic Publishers, p. 73–105, 2001.
- TSAI, R. Y.; HUANG, T. S. Multi-frame image restoration and registration. *Advances in Computer Vision and Image Processing*, p. 317–339, 1984.
- VLIET, L. J.; HENDRIKS, C. L. L. Improving spatial resolution in exchange of temporal resolution in aliased image sequences. *SCIA'99, Proc. 11th Scandinavian Conference on Image Analysis*, p. 493–499, 1999.
- WANG, Z.; BOVIK, A. C. A Universal Image Quality Index. *IEEE Signal Processing Letters*, v. 9, n. 3, p. 81–84, 2002.
- WINKLER, G. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (Stochastic Modelling and Applied Probability)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.