

David Campos Anchieta

Restauração de Imagens de Placas de Veículos Usando Super-resolução

Belém - PA

2018

David Campos Anchieta

Restauração de Imagens de Placas de Veículos Usando Super-resolução

PREÂMBULO

Universidade Federal do Pará

Instituto de Tecnologia

Faculdade de Engenharia da computação e Telecomunicações

Orientador: Prof. Dr. Ronaldo de Freitas Zampolo

Belém - PA

2018

David Campos Anchieta

Restauração de Imagens de Placas de Veículos Usando Super-resolução

PREÂMBULO

Trabalho aprovado. Belém - PA, 24 de novembro de 2012:

Prof. Dr. Ronaldo de Freitas Zampolo
Orientador

Professor
Convidado 1

Professor
Convidado 2

Belém - PA
2018

*Além disso, filho meu, sê avisado.
De fazer muitos livros não há fim;
e o muito estudar é enfado da carne.
Eclesiastes 12:12*

Resumo

Abstract

This is the english abstract.

Keywords: latex. abntex. text editoration.

Lista de ilustrações

Figura 1 – Ilustração demonstrado as transformações sofrida por uma imagem durante o processo de aquisição.	13
Figura 2 – Visualização das transformações de subamostragem, deslocamento linear e rotação.	14
Figura 3 – Exemplo de gráfico de uma função de espalhamento de ponto com $\gamma = 2$ centrada em $[0, 0]$	15
Figura 4 – Exemplo de convolução de uma imagem com uma função de espalhamento de ponto. Neste caso, a imagem da direita é resultado da convolução da imagem da esquerda com uma função de espalhamento de ponto gaussiana com $\gamma = 10$	15
Figura 5 – Imagens de alta resolução utilizadas nos testes.	23
Figura 6 – Conjunto das imagens geradas a partir da imagem de alta resolução. A diferença sutil entre as imagens é causada pelas transformações de deslocamento e rotação.	24
Figura 7 – Conjunto das imagens geradas a partir da imagem de alta resolução. A diferença sutil entre as imagens é causada pelas transformações de deslocamento e rotação.	24
Figura 8 – Progressão dos valores da avaliação da função de verossimilhança e do erro dos parâmetros estimados a cada iteração do algoritmo de otimização.	26
Figura 9 – Comparação dos resultados obtidos ao final da execução do algoritmo de otimização com os parâmetros verdadeiros.	27
Figura 10 – Comparação do resultado obtido através do método de super-resolução Bayesiana com a simples interpolação da imagem.	28

Lista de abreviaturas e siglas

SR	Super-resolução
HR	<i>High resolution</i>
LR	<i>Low resolution</i>
PDF	<i>Probability density function</i>

Lista de símbolos

\mathbf{x}	Vetor com os dados da imagem de alta resolução.
\mathbf{y}_k	Vetor da imagem de alta resolução de número k.
K	Número de imagens de baixa resolução.
N	Número de pontos na imagem de alta resolução.
M	Número de pontos na imagem de baixa resolução.
$\mathbf{W}^{(k)}$	Matriz de degradação da imagem de número k.
γ	Largura da função de espalhamento de ponto.
θ_k	Ângulo de rotação da imagem de número k.
\mathbf{s}_k	Vetor com deslocamento bidimensional da imagem de baixa resolução de número k.
\mathbf{Z}_x	Matriz de covariância da distribuição a priori de \mathbf{x} .
Σ	Matriz de covariância da distribuição a posteriori de \mathbf{x} condicionada aos parâmetros de degradação e conjunto de todos os $\mathbf{y}^{(k)}$.
$\boldsymbol{\mu}$	Vetor média da distribuição a posteriori de \mathbf{x} condicionada aos parâmetros de degradação e conjunto de todos os $\mathbf{y}^{(k)}$.
β	Inverso da variância do ruído.
$ \cdot $	Determinante de matriz.
$\ \cdot\ $	Norma de um vetor.

Sumário

1	INTRODUÇÃO	10
	Introdução	10
2	RESTAURAÇÃO DE IMAGENS USANDO SUPER-RESOLUÇÃO	12
2.1	Introdução	12
2.2	Modelo de Observação	12
2.3	Apreciação dos Métodos de Super-resolução	15
2.3.1	Métodos no domínio da frequência.	16
2.3.2	Iterative Back Projections	17
2.3.3	Métodos de Aprendizado	17
2.3.4	Métodos Probabilísticos	18
3	SUPER RESOLUÇÃO BAYESIANA	20
3.1	Introdução	20
3.2	Máximo a posteriori	20
3.3	Equacionamento do Método de Super-resolução Bayesiana	20
4	RESULTADOS	23
4.1	Introdução	23
4.2	Geração do conjunto de imagens de teste	23
4.3	Estimação dos parâmetros de degradação	25
4.4	Estimação da imagem de alta resolução	27
5	CONCLUSÃO	29
	Referências	30

1 Introdução

A disponibilidade de imagens de alta resolução é uma vantagem para a execução de qualquer trabalho que envolva a captação e interpretação de informações através de imagens digitais. Como exemplo disso podemos citar as imagens médicas, nas quais quanto melhor a qualidade da imagem, mais fácil é diferenciar as estruturas retratadas com o objetivo de diagnosticar doenças. Também há o caso dos sistemas de vigilância, nos quais quanto melhor a resolução espacial da imagem, mais viável se torna a identificação de veículos e criminosos.

Hoje em dia, com as frequentes ameaças à segurança das pessoas em diversos lugares do mundo, os cidadãos sentem a necessidade de investir na própria segurança privada para proteger a si mesmos e seus patrimônios. Por causa dessa necessidade, houve uma popularização de sistemas de vigilância monitoramento. As câmeras de segurança já são tão onipresentes em grandes cidades que é praticamente impossível que um cidadão comum saia à rua e não tenha a sua imagem gravada por pelo menos uma delas.

No entanto, apesar da presença geral dos sistemas de vigilância, a maioria dos equipamentos utilizados é de baixa qualidade, o que se reflete nas imagens captadas por eles. A consequência disso é que muitas vezes as câmeras de monitoramento registram cenas de crimes, mas as imagens deterioradas nem sempre ajudam a trazer à luz informações importantes, como o rosto de um criminoso ou a placa de um carro, por exemplo.

Não obstante, mesmo com necessidade evidente de se melhorar a qualidade das imagens geradas por sistemas de vigilância, também há o dilema constante entre aumentar a qualidade das imagens registradas e o custo de implementar manter o sistema de monitoramento funcionando. Quanto melhor a qualidade da imagem, maior é o custo do equipamento e mais capacidade de armazenamento será necessária para guardar uma quantidade útil de todas as imagens gravadas.

Face ao dilema das imagens geradas por sistemas de vigilância, a melhor solução a curto prazo seria aprimorar as imagens que estão sendo geradas pelas câmeras de vigilância que já temos hoje. Esse aprimoramento seria feito através de restauração de imagens digitais.

A demanda pelo desenvolvimento de técnicas de restauração de imagens ganhou força no fim da década de 1950, quando a sua principal aplicação era na área de imagens astronômicas. Naquele contexto, era necessário restaurar as primeiras imagens recebidas de sondas e satélites artificiais que sofriam degradações próprias do ambiente onde esses objetos trabalhavam.

Após tantos anos, a restauração de imagens digitais continua tendo sua importância na astronomia, mas também é uma ferramenta para outras áreas como medicina, mídia e ciências forenses [1].

Este trabalho descreve a implementação e teste de um método de restauração de imagens de placas de veículos usando super-resolução bayesiana. O método utilizado foi apresentado por Michael Tipping em [10]. Em resumo, o algoritmo apresentado define um modelo de observação no qual as degradações sofridas pela imagem durante o processo de aquisição são parametrizadas. Então são usadas inferências bayesianas para estimar os parâmetros de degradação a partir das imagens de baixa resolução, para então estimar a imagem de alta resolução a partir destes parâmetros.

Esta solução foi escolhida devido à adequação do algoritmo à aplicação desejada, que é a restauração de uma imagem de placa de veículo a partir de várias imagens degradadas, como os quadros de um vídeo de câmera de vigilância. Também vale destacar a relativa simplicidade de implementação do método escolhido, o qual também possibilita ajustes afim de reduzir seu custo computacional.

2 Restauração de Imagens Usando Super-resolução

2.1 Introdução

Entende-se por Super-resolução o processo de obtenção de uma ou mais imagens de alta resolução a partir de uma ou mais imagens de baixa resolução. Técnicas de super-resolução são pesquisadas desde os anos 1970 e têm despertado um grande interesse de estudo nas últimas décadas. As aplicações de técnicas de SR são diversas e incluem aprimoramentos de imagens médicas, de imagens de rosto, de texto e impressões digitais [6].

Vale esclarecer que, neste trabalho, sempre que a palavra *resolução* é usada, ela se refere à resolução espacial de uma imagem, a qual mede a densidade de pontos por unidade de área em uma imagem [11].

Este capítulo fará uma breve síntese do histórico das técnicas de super-resolução e seus principais exemplares. A seção 2.2 descreve o modelo de observação usado pelos principais trabalhos no tema. A seção 2.3 cita e descreve resumidamente as principais técnicas e grupos de métodos de restauração de imagens via super-resolução.

2.2 Modelo de Observação

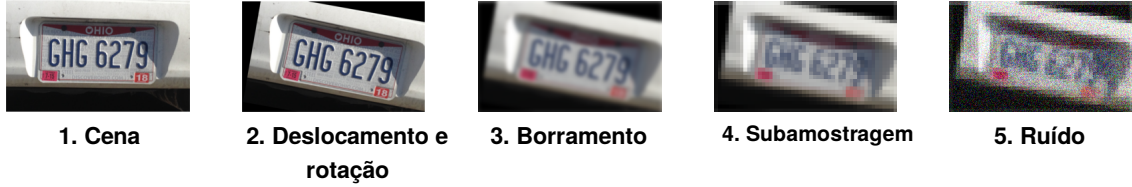
Definir um modelo de observação que relacione a imagem de alta resolução às imagens de baixa resolução observadas é fundamental para o desenvolvimento de qualquer técnica de super-resolução. Um modelo de observação permite parametrizar e quantificar as transformações às quais a imagem é submetida durante o processo de aquisição e é através dele que os principais métodos de super-resolução inferem as características da imagem de alta resolução.

O modelo de observação deste trabalho simula cinco tipos de alteração aplicadas por sistemas de aquisição de imagens, a saber: rotação, deslocamento, espalhamento de ponto, subamostragem e ruído. Este modelo de observação é descrito tanto no artigo de Tipping [10] quanto em outros trabalhos relacionados [8, 2].

As alterações de deslocamento linear e rotação tentam simular um possível movimento aleatório da câmera em relação ao objeto ou vice-versa. A transformação de espalhamento de ponto simula um efeito de desfoque causado por um desajuste na definição de foco da câmera. A subamostragem representa a perda de resolução espacial

aplicada à imagem captada por um sensor CMOS ou CCD de densidade menor do que a necessária para registrar os dados da imagem sem que ocorra *aliasing*. A figura 1 ilustra o efeito destas cinco transformações em uma imagem.

Figura 1 – Ilustração demonstrado as transformações sofrida por uma imagem durante o processo de aquisição.



Fonte: O autor.

Pela conveniência do modelo de observação, a imagem de alta resolução de dimensões $m \times n$ é representada por um vetor coluna $\mathbf{x} = [x_1, x_2, \dots, x_{N-1}, x_N]^T$ onde $N = m \times n$. Ou seja, os valores dos pixels da imagem são rearranjados em um vetor de comprimento N .

Sabendo disso, a relação entre uma imagem de alta resolução \mathbf{x} e uma imagem degradada $\mathbf{y}^{(k)}$ é resumida em (2.1). Onde $k = 1, 2, \dots, K$; sendo K o número total de quadros obtidos a partir da cena de alta resolução \mathbf{x} .

$$\mathbf{y}^{(k)} = \mathbf{W}^{(k)}\mathbf{x} + \mathbf{n}^{(k)} \quad (2.1)$$

O ruído, representado por \mathbf{n} é um vetor de variáveis aleatórias Gaussianas independentes de média zero e variância $1/\beta$.

A matriz de sistema \mathbf{W} aplica as transformações de rotação, deslocamento, subamostragem e espalhamento de ponto ao ser multiplicada pela imagem de alta resolução \mathbf{x} . Pelas propriedades da multiplicação de matrizes, é natural que as dimensões da matriz de sistema sejam $M \times N$; onde M é o número de pixels da imagem resultante. Também se espera que $N \gg M$.

A matriz \mathbf{W} depende de três parâmetros de transformação: a largura da função de espalhamento de ponto γ , um ângulo de rotação θ e um vetor de deslocamento linear \mathbf{s} . Os elementos da matriz são dados por (2.2) com a função de espalhamento de ponto descrita em (2.3) enquanto os parágrafos seguintes demonstram como as demais transformações são incorporadas à matriz de sistema.

$$W_{ji}^{(k)} = \widetilde{W}_{ji}^{(k)} / \sum_{i'} \widetilde{W}_{ji}^{(k)} \quad (2.2)$$

$$\widetilde{W}_{ji}^{(k)} = \exp \left\{ -\frac{\|\mathbf{v}_i - \mathbf{u}_j^{(k)}\|^2}{\gamma^2} \right\} \quad (2.3)$$

Os vetores $\mathbf{u}_j^{(k)}$ representam o centro da função de espalhamento de ponto e é descrito em (2.4). Caso não houvesse espalhamento de ponto, as coordenadas $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_M\}$ seriam os pontos de amostragem da imagem de baixa resolução na imagem de alta resolução. A Figura 2 mostra como se dá o processo de subamostragem e outras transformações bidimensionais.

$$\mathbf{u}_j^{(k)} = \mathbf{R}^{(k)}(\mathbf{v}_j - \bar{\mathbf{v}}) + \bar{\mathbf{v}} + \mathbf{s}_k \quad (2.4)$$

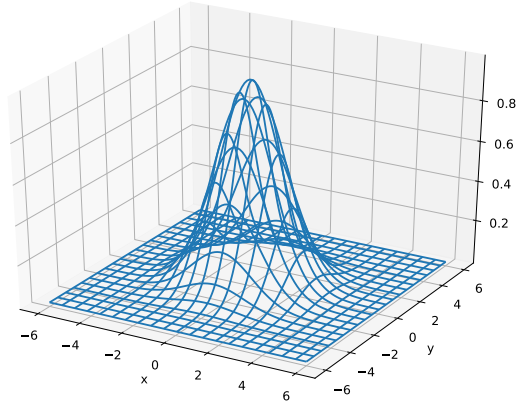
$$\mathbf{R}^{(k)} = \begin{pmatrix} \cos \theta_k & \sin \theta_k \\ -\sin \theta_k & \cos \theta_k \end{pmatrix} \quad (2.5)$$

Figura 2 – Visualização das transformações de subamostragem, deslocamento linear e rotação.



Fonte: O autor.

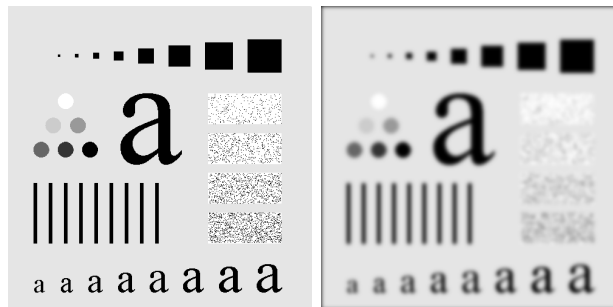
Figura 3 – Exemplo de gráfico de uma função de espalhamento de ponto com $\gamma = 2$ centrada em $[0, 0]$.



Fonte: O autor.

A transformação de espalhamento de ponto ou borramento consiste em realizar uma convolução da imagem com uma função gaussiana bidimensional como a mostrada na figura 3. Visualmente, o efeito dessa convolução é equivalente ao de um desfoque de lente, como demonstrado na figura 4

Figura 4 – Exemplo de convolução de uma imagem com uma função de espalhamento de ponto. Neste caso, a imagem da direita é resultado da convolução da imagem da esquerda com uma função de espalhamento de ponto gaussiana com $\gamma = 10$.



Fonte: O autor.

2.3 Apreciação dos Métodos de Super-resolução

Esta seção descreve resumidamente alguns dos métodos de super-resolução mais importantes

A maioria dos algoritmos de super-resolução desenvolvidos faz o processamento de imagem no domínio espacial. Os tipos de abordagens usadas são bastante diversos,

sendo que a mais frequente é a probabilística, na qual o método usado neste trabalho está incluído [6].

2.3.1 Métodos no domínio da frequência.

Os primeiros métodos de super-resolução desenvolvidos faziam o processamento das imagens no domínio da frequência. Sendo que os primeiros trabalhos com essa abordagem foram os de Gerchberg [4] e Santis e Gori [3].

O primeiro método de super-resolução no domínio da frequência a ter múltiplas imagens como entrada foi desenvolvido por Tsai e Huang em 1994 [6]. Este algoritmo foi desenvolvido para processar as imagens geradas pelo satélite Landsat 4. Este aparelho gerava várias imagens semelhantes, porém deslocadas, da mesma área na terra.

Podemos considerar que cada imagem g_k gerada pelo satélite seja uma versão deslocada de uma cena contínua f , podemos considerar que

$$g_k(m, n) = f(m + \Delta_{m_k}, n + \Delta_{n_k}). \quad (2.6)$$

Pela propriedade de deslocamento da transformada de Fourier, a transformada de Fourier contínua das imagens geradas pelo satélite, em relação à pode ser dada por

$$F_{g_k}(m, n) = \exp[i2\pi(\Delta_{m_k}m + \Delta_{n_k}n)]F_f(m, n). \quad (2.7)$$

Tendo isso, a transformada discreta de Fourier de g_k é dada por

$$G_k(m, n) = \frac{1}{T_m T_n} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} F_{g_k} \left(\frac{m}{MT_m} + p_1 \frac{1}{T_m}, \frac{n}{NT_n} + p_2 \frac{1}{T_n} \right). \quad (2.8)$$

Segundo Nasrollahi e Moeslund [6] o problema de super-resolução se resolve encontrando a transformada contínua de Fourier (\mathbf{F}_f) que soluciona (2.9). Isso pode ser resolvido usando o Método dos Mínimos Quadrados.

$$\mathbf{G} = \Phi \mathbf{F}_f \quad (2.9)$$

A abordagem dos algoritmos de super-resolução deste tipo tinha a vantagem da simplicidade em sua teoria. No entanto, as técnicas de super-resolução no domínio da frequência tinham limitações com respeito ao tipo de degradação sofrida pela imagem além da dificuldade de aplicar o conhecimento a priori da imagem de alta resolução no domínio espacial para regularização. Estes são os principais motivos pelos quais a maioria dos algoritmos de super-resolução desenvolvidos até hoje faz o processamento das imagens no domínio espacial [7].

2.3.2 Iterative Back Projections

O método *Iterative Back Projection* (IBP) foi proposto pela primeira vez por Irani e Peleg [5]. O método IBP se caracteriza por ser um método iterativo de super-resolução para restaurar uma imagem de alta resolução a partir de várias imagens de baixa resolução. Por ser iterativo, o método executa uma rotina por diversas iterações com o objetivo de se aproximar mais da solução correta a cada execução.

O método parte de um modelo de observação como o apresentado na seção 2.2 no qual uma imagem de baixa resolução $\mathbf{y}^{(k)}$ é dada por

$$\mathbf{y}^{(k)} = \mathbf{W}^{(k)} \mathbf{x}. \quad (2.10)$$

Onde \mathbf{x} é a imagem de alta resolução a ser encontrada e $\mathbf{W}^{(k)}$ é a matriz que aplica as transformações do modelo de observação.

Se for feito um palpite tanto para a matriz de sistema quanto para a imagem a ser estimada, podemos dizer que o erro do palpite é dado por 2.11. Feito isso, a melhor estimativa para a imagem de alta resolução \mathbf{x} seria a que minimiza o erro na equação 2.12. Essa estimativa é alcançada através de métodos iterativos como o de mínimos quadrados.

$$\epsilon_j^{(k)} = \mathbf{y}^{(k)} - \mathbf{W}_j^{(k)} \mathbf{x}_j \quad (2.11)$$

$$\mathbf{e}_j = \sum_{k=1}^N \sqrt{\|\mathbf{y}^{(k)} - \mathbf{W}_j^{(k)} \mathbf{x}_j\|_2^2} \quad (2.12)$$

O método *Iterative Back Projections* tem a vantagem de ser intuitivo e de fácil compreensão. No entanto, ele tem problemas por não ter uma solução única, além de não permitir que se use com facilidade o conhecimento a priori acerca da imagem de alta resolução [6, 7].

2.3.3 Métodos de Aprendizado

Os métodos de super-resolução baseados em aprendizado usam redes neurais para restaurar imagens estáticas. Para que isso seja possível, deve haver uma etapa de aprendizado onde exemplos de imagens de uma classe específica (impressões digitais, rostos, texto) são usados junto com seus equivalentes de baixa resolução para que o algoritmo reconheça o relacionamento entre elas. Esse conhecimento forma a componente a priori na reconstrução das imagens.

2.3.4 Métodos Probabilísticos

Consideremos o modelo de observação em (2.1), onde o vetor \mathbf{n} representa o ruído adicionado às imagens de baixa resolução. Sabendo que cada elemento \mathbf{n}_i é aleatório seguindo uma distribuição gaussiana de média 0 e variância $1/\beta$, podemos estabelecer que a função densidade de probabilidade de \mathbf{n} é representada em (2.13).

$$p(\mathbf{n}^{(k)}) = \left(\frac{\beta}{2\pi}\right)^{M/2} \exp\left\{-\frac{\beta}{2}\|\mathbf{n}^{(k)}\|^2\right\} \quad (2.13)$$

Ainda baseado em (2.1), a relação entre uma estimativa da imagem de alta resolução $\hat{\mathbf{x}}$ e a imagem de baixa resolução observada é:

$$\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}} = \mathbf{n}^{(k)} \quad (2.14)$$

Então é possível dizer que a probabilidade de $\mathbf{y}^{(k)}$ ocorrer baseado em uma estimativa da imagem de alta resolução $\hat{\mathbf{x}}$ é dada pela distribuição em (2.15).

$$p(\mathbf{y}^{(k)}|\hat{\mathbf{x}}, \mathbf{W}^{(k)}) = \left(\frac{\beta}{2\pi}\right)^{M/2} \exp\left\{-\frac{\beta}{2}\|\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}}\|^2\right\} \quad (2.15)$$

O logaritmo da função de verossimilhança de $\mathbf{y}^{(k)}$ é expresso em

$$L\{\mathbf{y}^{(k)}\} = -\|\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}}\|^2. \quad (2.16)$$

Teoricamente, a melhor estimativa ($\hat{\mathbf{x}}$) da imagem de alta resolução é a que maximiza a função de verossimilhança acima, como é demonstrado em (2.17).

$$\mathbf{y}_{ML}^{(k)} = \arg \max_{\hat{\mathbf{x}}} \left[-\|\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}}\|^2\right] \quad (2.17)$$

Para obter o ponto de máximo da função, podemos calcular seu diferencial, igualá-lo a zero, e isolar o $\hat{\mathbf{x}}$ na equação como é feito abaixo.

$$\frac{d}{d\hat{\mathbf{x}}} L\{\mathbf{y}^{(k)}\} = -2(\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}})\mathbf{W}^{(k)T} \quad (2.18)$$

$$(\mathbf{y}^{(k)} - \mathbf{W}^{(k)}\hat{\mathbf{x}})\mathbf{W}^{(k)T} = 0 \quad (2.19)$$

$$\hat{\mathbf{x}} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{y}^{(k)} \quad (2.20)$$

O método de máxima verossimilhança é sensível a distúrbios causados por ruídos ou imprecisões na estimação dos parâmetros de aquisição. Além disso, se o número de imagens não for suficiente, pode haver mais de uma solução possível. Esse problema deve

ser contornado adicionando conhecimento *a priori* da imagem. Esse conhecimento ajuda a restringir o número de respostas possíveis [6].

Uma das formas de adicionar conhecimento *a priori* ao método é transformá-lo em uma abordagem de máximo *a posteriori*. O método de Super-resolução Bayesiana estudado neste trabalho é um exemplo de método probabilístico que usa inferência Bayesiana para incorporar conhecimento *a priori* acerca da imagem ao modelo.

3 Super Resolução Bayesiana

3.1 Introdução

O método de super resolução desenvolvido no trabalho de Tipping e Bishop [10] é um método probabilístico para a obtenção de uma imagem de alta resolução a partir de um conjunto de imagens de baixa resolução de uma mesma cena. Esse conjunto de imagens de baixa resolução poderiam ser retirados de quadros de um vídeo produzido por uma câmera de segurança que fica parada em relação ao objeto observado, por exemplo.

Pode-se dizer que a abordagem usada neste método é uma expansão do método de máxima verossimilhança descrito na seção 2.3.4. O método de super-resolução Bayesiana adiciona uma densidade de probabilidade como conhecimento a priori da imagem para restringir a quantidade de soluções possíveis do problema, contornando assim um dos problemas da abordagem de máxima verossimilhança. A adição de uma densidade de probabilidade *a priori*, converte a abordagem de máxima verossimilhança em uma abordagem de máximo *a posteriori* (MAP).

se vale do método de inferência Bayesiana para obter uma imagem de alta resolução a partir de várias imagens de baixa resolução. Através de um método de inferência Bayesiana, é possível obter uma estimativa de valores desconhecidos (como a imagem de alta resolução) usando um conhecimento a priori que se tem desses dados e informações que já estão disponíveis (no caso, as imagens de baixa resolução) [9].

Sabendo disso, há a necessidade de formular uma distribuição a priori que represente as características estatísticas da imagem que pretendemos estimar a partir dos dados. A distribuição de probabilidade a priori escolhida para a imagem é uma gaussiana como descrita a seguir:

3.2 Máximo a posteriori

3.3 Equacionamento do Método de Super-resolução Bayesiana

A distribuição de probabilidade *a priori* $p(\mathbf{x})$ é uma Gaussiana de média $\mathbf{0}$ e matriz de covariância \mathbf{Z}_x como definido em (3.2).

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|0, \mathbf{Z}_x) \quad (3.1)$$

$$Z_x(i, j) = A \exp \left\{ -\frac{\|\mathbf{v}_i - \mathbf{v}_j\|^2}{r^2} \right\} \quad (3.2)$$

Onde A é uma constante que indica a *força* da distribuição e r determina o quanto cada ponto da imagem de alta resolução depende dos pontos vizinhos. Expandindo a função densidade de probabilidade $p(\mathbf{x})$ temos:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{Z}_x|}} \exp -\frac{1}{2} \mathbf{x}^T \mathbf{Z}_x^{-1} \mathbf{x} \quad (3.3)$$

Este tipo de distribuição não é o ideal, visto que se definirmos $r = 1$, estaríamos dizendo que a maioria dos valores dos pontos da imagem de alta resolução estão contidos no intervalo $[-1, 1]$, o que não condiz com a realidade do que se espera de uma imagem digital. No entanto, o uso desta distribuição a priori será mantido no desenvolvimento deste trabalho pela conveniência proporcionada pela função de probabilidade Gaussiana, que facilita algumas manipulações – como aplicação de logaritmos naturais – utilizadas a seguir.

Ciente da relação entre a imagem original e as imagens de baixa resolução, podemos escrever a probabilidade a posteriori das imagens de baixa resolução como:

$$p(\mathbf{y}^{(k)} | \mathbf{x}, \mathbf{s}_k, \theta_k, \gamma) = \left(\frac{\beta}{2\pi} \right)^{M/2} \exp \left\{ -\frac{\beta}{2} \|\mathbf{y}^{(k)} - \mathbf{W}^{(k)} \mathbf{x}\|^2 \right\}. \quad (3.4)$$

Note que esta distribuição é condicionada na imagem de alta resolução e nos parâmetros de degradação, dados que são desconhecidos em um situação real.

Podemos usar o Teorema de Bayes para isolar a imagem original \mathbf{x} , ao que obtemos:

$$p(\mathbf{x} | \{\mathbf{y}^{(k)}, \mathbf{s}_k, \theta_k\}, \gamma) = \frac{p(\mathbf{x}) \prod_{k=1}^K p(\mathbf{y}^{(k)} | \mathbf{x}, \mathbf{s}_k, \theta_k, \gamma)}{p(\{\mathbf{y}^{(k)}\} | \{\mathbf{s}_k, \theta_k\}, \gamma)} \quad (3.5)$$

$$= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (3.6)$$

$$\boldsymbol{\Sigma} = \left[\mathbf{Z}_x^{-1} + \beta \left(\sum_{k=1}^K \mathbf{W}^{(k)T} \mathbf{W}^{(k)} \right) \right]^{-1} \quad (3.7)$$

$$\boldsymbol{\mu} = \beta \boldsymbol{\Sigma} \left(\sum_{k=1}^K \mathbf{W}^{(k)T} \mathbf{y}^{(k)} \right). \quad (3.8)$$

Esta função de probabilidade já poderia ser utilizada para inferir a imagem de alta resolução. Para isso, bastaria encontrar o valor de \mathbf{x} que maximiza o numerador em (3.5).

entretanto, isso exigiria saber os valores dos parâmetros de degradação $(\gamma, \theta_k, \theta_k)$. Como esses parâmetros não estão disponíveis, eles também terão de ser estimados a partir das imagens de baixa resolução \mathbf{y}_k .

A solução proposta no artigo é marginalizar \mathbf{x} em (3.4) para que ele não seja mais uma variável na função. O resultado disto é a seguinte distribuição:

$$\log p(\mathbf{y}|\{\mathbf{s}_k, \theta_k\}, \gamma) = -\frac{1}{2} \left[\beta \sum_{k=1}^K \|\mathbf{y}^{(k)} - \mathbf{W}^{(k)} \boldsymbol{\mu}\|^2 + \boldsymbol{\mu}^T \mathbf{Z}_x \boldsymbol{\mu} + \log |\mathbf{Z}_x| - \log |\Sigma| - KM \log \beta \right]. \quad (3.9)$$

Onde \mathbf{y} é o conjunto de todas as imagens de baixa resolução \mathbf{y}_k .

Em uma situação ideal, a melhor estimativa dos valores dos parâmetros de degradação é a que maximiza a função de verossimilhança em (3.9). Tendo a estimativa desses valores, podemos então usá-los em (3.5) para procurar pela imagem de alta resolução \mathbf{x} que maximiza aquela função de verossimilhança.

Um fato que se constatou empiricamente na tentativa de calcular os valores da função de verossimilhança é que as probabilidades são muito pequenas ao ponto de estarem fora do intervalo de precisão de uma variável do tipo ponto flutuante. Os valores também variam muito pouco ao longo do espaço vetorial de \mathbf{x} o que dificultaria o uso de métodos iterativos para encontrar o ponto de máximo da função.

Para resolver isso, aplicou-se o logaritmo natural à função. Isso resolveu o problema da escala e da variação. Além disso eliminou os exponenciais e os produtos, facilitando o cálculo do gradiente e Jacobiano da função.

$$L(\mathbf{x}) = -\frac{1}{2} \left\{ \sum_{k=1}^K \|\mathbf{y}^{(k)} - \mathbf{W}^{(k)} \mathbf{x}\|^2 + \mathbf{x}^T \mathbf{Z}_x^{-1} \mathbf{x} - KM \ln(\beta) - \ln |\mathbf{Z}_x| \right\} \quad (3.10)$$

$$L'(\mathbf{x}) = \sum_{k=1}^K (\mathbf{y}^{(k)} - \mathbf{W}^{(k)} \mathbf{x})(\mathbf{W}^{(k)})^T - \frac{1}{2}(\mathbf{Z}_x^{-1} + (\mathbf{Z}_x^{-1})^T) \mathbf{x} \quad (3.11)$$

$$L''(\mathbf{x}) = -\sum_{k=1}^K (\mathbf{W}^{(k)})^T \mathbf{W}^{(k)} - \frac{1}{2}(\mathbf{Z}_x^{-1} - (\mathbf{Z}_x^{-1})^T)^T \quad (3.12)$$

4 Resultados

4.1 Introdução

Para avaliar a eficácia do método de super-resolução bayesiana, foi elaborado um teste que simula um sistema de aquisição de imagem, gerando as imagens de baixa resolução a serem restauradas. Na sequência, as imagens de baixa resolução foram submetidas ao método de super-resolução como descrito capítulo 3.

Este capítulo descreve o processo de implementação e teste do método de super-resolução bayesiana. A seção 4.2 apresenta o processo que gerou as imagens de teste simulando as transformações aplicadas pelo sistema de aquisição. A seção 4.3 relata como se deu o processo de restauração da imagem de alta resolução a partir do conjunto de imagens de baixa resolução.

4.2 Geração do conjunto de imagens de teste

O método de super-resolução Bayesiana foi testado com dois conjuntos de imagens de baixa resolução geradas artificialmente a partir de duas imagens de alta resolução distintas. O modelo de observação exposto na seção 2.2 foi usado para gerar dois conjuntos de 20 imagens de baixa resolução a partir das imagens na figura 5.

Figura 5 – Imagens de alta resolução utilizadas nos testes.



(a) Exemplar 1: “*Virginia*”



(b) Exemplar 2: “*Ohio*”

Fonte: O autor.

Para os testes feitos neste trabalho, ambas as imagens de alta resolução foram redimensionadas de forma que as dimensões das imagens de baixa resolução são iguais a 1/4 das dimensões da imagem HR. Na prática, há 16 pontos nas imagens de alta resolução para cada ponto nas imagens de baixa resolução.

Sobre os parâmetros de degradação escolhidos. A largura da função de espalhamento de ponto (γ) escolhida foi 2. Para cada imagem, o ângulo de rotação foi escolhido aleatoriamente de uma distribuição uniforme entre -4 e 4 graus. A mesma seleção aleatória foi feita para o deslocamento linear da imagem; a quantidade de pontos de deslocamento foi retirada de uma distribuição uniforme e contínua entre -2 e 2 em cada um dos eixos. A tabela 1 resume os parâmetros utilizados na degradação das imagens.

Figura 6 – Conjunto das imagens geradas a partir da imagem de alta resolução. A diferença sutil entre as imagens é causada pelas transformações de deslocamento e rotação.



Fonte: O autor.

Figura 7 – Conjunto das imagens geradas a partir da imagem de alta resolução. A diferença sutil entre as imagens é causada pelas transformações de deslocamento e rotação.



Fonte: O autor.

Tabela 1 – Parâmetros de degradação usados para gerar o conjunto de imagens de teste.

Ângulos de rotação (θ_k)	$\sim \mathcal{U}(-4, 4)$
Deslocamento (\mathbf{s}_k)	$\sim \mathcal{U}(-2, 2)$
Largura da PSF (γ)	2
Dimensões da imagem de alta resolução	139×104
Dimensões das imagens de baixa resolução	35×26

Fonte: O autor.

O processo de geração das imagens de baixa resolução foi todo desenvolvido usando a linguagem Python, beneficiando-se dos módulos Numpy e Pillow. Os parâmetros de degradação bem como as dimensões da imagem original foram salvos em arquivos de valores separados por vírgula (csv) junto com as imagens geradas. Esses valores foram usados para avaliar quantitativamente a efetividade do algoritmo de otimização em inferir os valores corretos dos parâmetros.

4.3 Estimação dos parâmetros de degradação

A primeira etapa no processo de restauração da imagem é estimar os parâmetros que definiram as transformações que degradaram a imagem durante o processo simulado de captura. Como já elucidado no capítulo 3, encontrar os valores dos parâmetros de degradação é um problema de otimização. Teoricamente, os valores verdadeiros de \mathbf{s}_k , θ_k e γ são os que maximizam a função de verossimilhança em (3.9).

O artigo no qual este trabalho se baseia relata que o método de otimização usado para estimar tanto os parâmetros de degradação quanto da imagem foi o de gradientes conjugados [10]. Como o método de super-resolução deste trabalho foi todo escrito em Python, não foi possível utilizar o mesmo método de implementação utilizado no artigo. Em vez disso, foi usado um Método de Newton Truncado presente no módulo *Scipy*.

O processo de otimização dos parâmetros se deu da seguinte forma: Os parâmetros de rotação e deslocamento foram definidos como zero; a largura da função de espalhamento de ponto inicial foi 4.

Primeiramente, somente os valores de deslocamento são otimizados enquanto os valores dos ângulos de rotação permanecem fixos. Após isso, é realizada a otimização dos ângulos de rotação junto com os deslocamentos lineares. Em uma terceira etapa, a largura da função de espalhamento de ponto (γ) é otimizada concomitantemente com os demais parâmetros. Essa abordagem é usada no artigo original e os testes realizados para este trabalho constatarem que a otimização dos parâmetros de deslocamento se aproxima mais rápido da solução correta quando esses valores são estimados isoladamente. A estimação dos ângulos também funciona melhor quando os valores de deslocamento já estão otimizados.

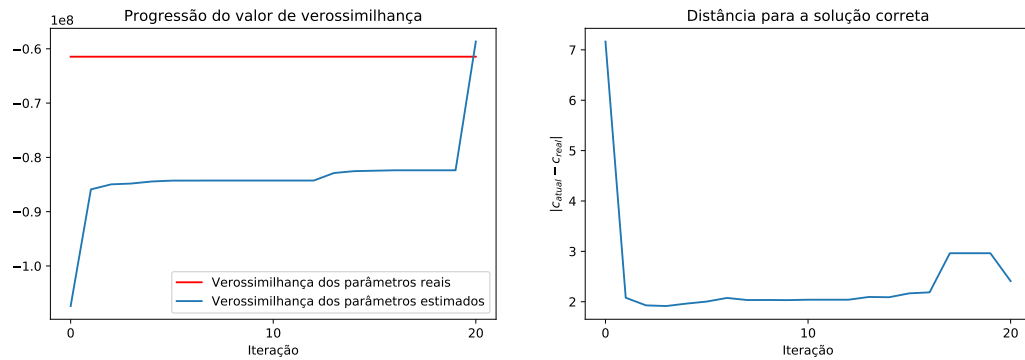
Um fato que merece destaque é o de que, como a função em (3.9) não é uma função vetorial do tipo $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^1$, não foi possível passá-la diretamente ao método de otimização. Por isso, foi necessário empacotar a função de forma que ela recebesse os parâmetros de entrada (\mathbf{s}_k , θ_k e γ) reorganizados na forma de um único vetor e retornasse somente um escalar na saída. Isso também implicou na exigência de calcular numericamente o gradiente da função, necessário na execução do método de otimização.

A função de otimização com algoritmo de Newton truncado presente na biblioteca *Scipy* permite que sejam definidos limites para restringir o intervalo de busca no espaço vetorial onde a estimação é feita. Isso é especialmente útil, pois permite usarmos um conhecimento a priori acerca dos parâmetros para impedir que a otimização retorne resultados inviáveis que, por ventura, tenham um alto valor de verossimilhança. Os limites definidos para os parâmetros foram de -4 a 4 no caso dos ângulos de rotação, -2 a 2 para ambas as direções dos deslocamentos e 1 a 7 para a largura da função de espalhamento de ponto.

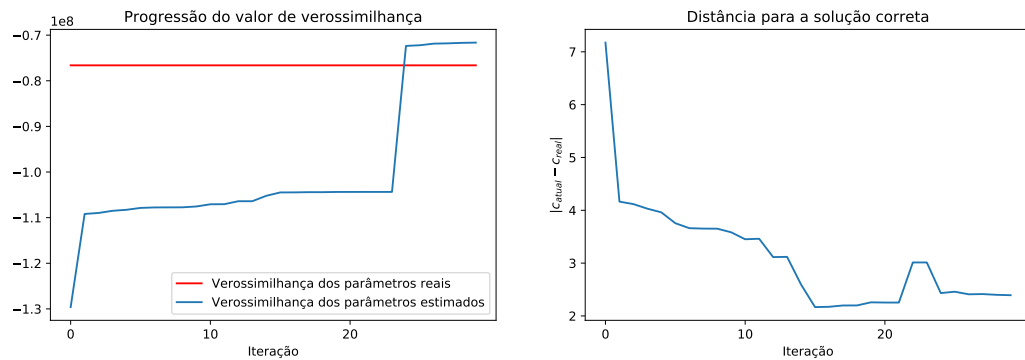
Para reduzir o tempo gasto na execução da otimização dos parâmetros, o número máximo de avaliações da função de verossimilhança por etapa foi restringido. Na primeira etapa, o número máximo de avaliações da função de verossimilhança foi de 70. Em ambas as etapas subsequentes este limite foi de 50 avaliações. Esses limites não interferem na qualidade do resultado pois, mesmo que eles sejam bem maiores, o valor das avaliações da função estabiliza após poucas iterações. Isso será demonstrado mais adiante.

Durante a execução do Método de Newton Truncado, foram registrados o valor da função de verossimilhança e o erro da solução atual ao fim de cada iteração. O erro calculado é igual $\|\mathbf{c}_{atual} - \mathbf{c}_{real}\|$ onde \mathbf{c}_{atual} é o vetor com os parâmetros obtidos ao fim da iteração e \mathbf{c}_{real} é o vetor com os parâmetros verdadeiros usado para a comparação. Os resultados desta análise são ilustrados na figura 8.

Figura 8 – Progressão dos valores da avaliação da função de verossimilhança e do erro dos parâmetros estimados a cada iteração do algoritmo de otimização.



(a) Resultados obtidos com a estimação do Exemplar 1: “Virginia”



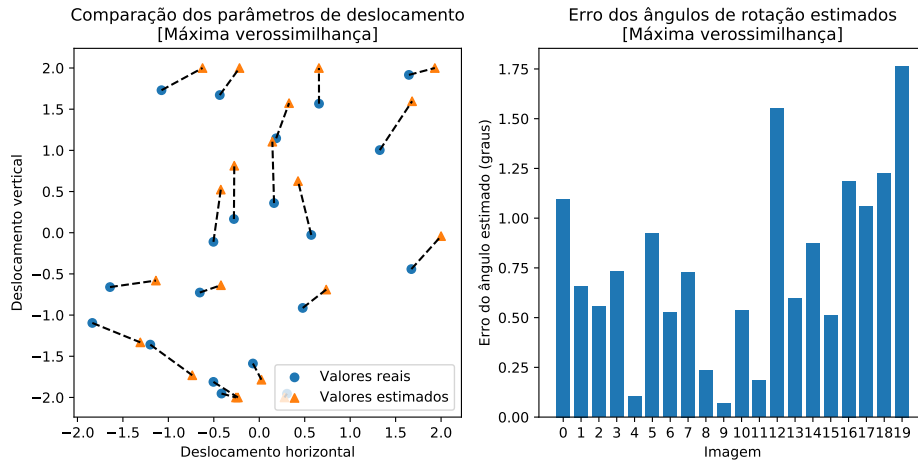
(b) Resultados obtidos com a estimação do Exemplar 2: “Ohio”

Fonte: O autor

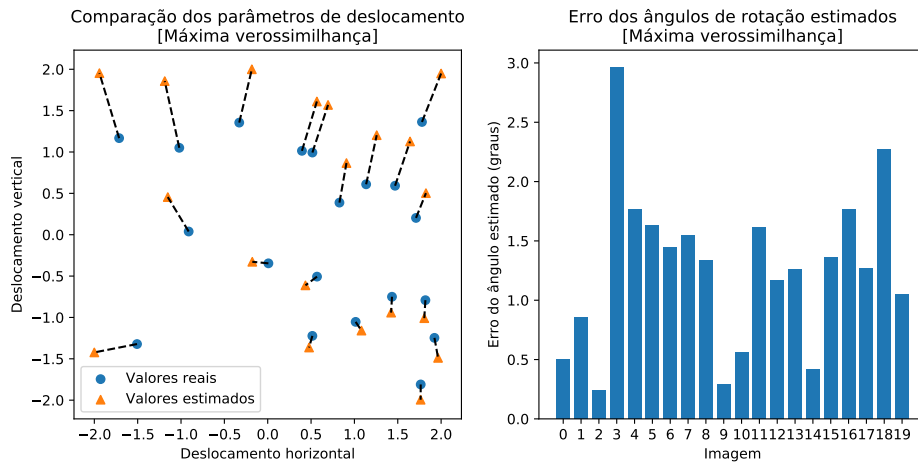
Ao final da execução do algoritmo de otimização, os os valores obtidos foram colhidos e comparados com os parâmetros de transformação reais. O resultado dessa comparação está retratado na figura 9. A estimativa da largura da função de espalhamento de ponto foi de 1 para ambas as imagens testadas. Esse erro pode ser observado também

nos gráficos da figura 8 onde há um aumento acentuado tanto no valor da função de verossimilhança quanto no erro da solução encontrada. Esse erro ocorre no

Figura 9 – Comparação dos resultados obtidos ao final da execução do algoritmo de otimização com os parâmetros verdadeiros.



(a) Resultados obtidos com a estimação do Exemplar 1: “*Virginia*”



(b) Resultados obtidos com a estimação do Exemplar 2: “*Ohio*”

Fonte: O autor

4.4 Estimação da imagem de alta resolução

Após a estimação dos parâmetros de degradação, os valores otimizados deve ser passados para a função de verossimilhança da imagem em (3.10) para que seja estimada a imagem de alta resolução em si.

O algoritmo de Newton truncado também foi usado para estimar a imagem de acordo com as imagens observadas e os parâmetros de degradação. Ao contrário

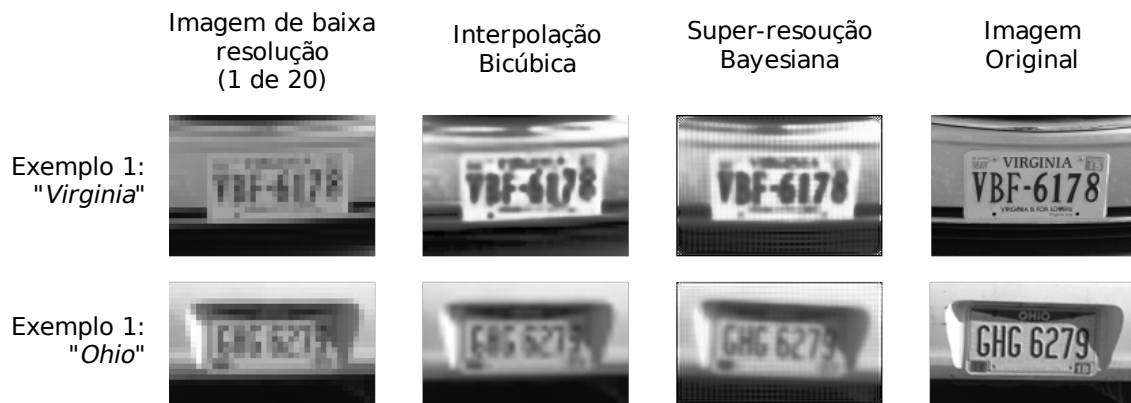
do que aconteceu na etapa anterior, não foi necessário calcular o gradiente da função numericamente para esta etapa. Como a função de verossimilhança em (3.10) é uma função vetorial do tipo $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^1$, então foi relativamente simples calcular seu gradiente mostrado na equação (3.11).

A possibilidade de se calcular o gradiente da função analiticamente em vez de numericamente teve como consequência um ganho considerável de desempenho com redução no tempo de execução do algoritmo. Esse ganho de desempenho foi bastante importante devido à quantidade de dados a serem estimados, que chegam a 15600 pontos no caso da imagem de maior resolução.

O processo de otimização da imagem teve como resultado uma imagem de baixo contraste que precisou ter seu histograma equalizado para

A figura mostra as imagens obtidas através do método de super-resolução Bayesiana e a compara com uma imagem que foi redimensionada usando interpolação bicúbica; que é um método bastante utilizado nos programas de edição de imagens.

Figura 10 – Comparação do resultado obtido através do método de super-resolução Bayesiana com a simples interpolação da imagem.



Fonte: O autor.

5 Conclusão

Referências

- [1] Mark R Banham and Aggelos K Katsaggelos. Digital image restoration. *IEEE signal processing magazine*, 14(2):24–41, 1997.
- [2] D. P. Capel. *Image Mosaicing and Super-resolution*. PhD thesis, University of Oxford, 2001.
- [3] P De Santis and F Gori. On an iterative method for super-resolution. *Journal of Modern Optics*, 22(8):691–695, 1975.
- [4] R.W. Gerchberg. Super-resolution through error energy reduction. *Optica Acta: International Journal of Optics*, 21(9):709–720, 1974.
- [5] Michal Irani and Shmuel Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, 53(3):231 – 239, 1991.
- [6] Kamal Nasrollahi and Thomas B Moeslund. Super-resolution: a comprehensive survey. *Machine vision and applications*, 25(6):1423–1468, 2014.
- [7] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, 20(3):21–36, 2003.
- [8] Lyndsey C Pickup, David P Capel, Stephen J Roberts, and Andrew Zisserman. Bayesian image super-resolution, continued. In *Advances in Neural Information Processing Systems*, pages 1089–1096, 2007.
- [9] C. Therrien and M. Tummala. *Probability and Random Processes for Electrical and Computer Engineers, Second Edition*. Taylor & Francis, 2011.
- [10] Michael E Tipping and Christopher M Bishop. Bayesian image super-resolution. In *Advances in neural information processing systems*, pages 1303–1310, 2003.
- [11] Marcelo Victor Wüst Zibetti. *Super-resolução simultânea para sequência de imagens*. PhD thesis, Universidade Federal de Santa Catarina, 2007.