Reconstrução de Imagens via Métodos Bayesianos

Gabriel Ligabô

3 de julho de 2025

1 Introdução

A reconstrução de imagens é um problema clássico que busca recuperar uma imagem original a partir de uma versão degradada. Esse tipo de problema é importante em diversos campos de estudo, um exemplo de sua utilidade é na reconstrução de imagens de tomografia computadorizada [1]. Este trabalho explora uma abordagem bayesiana para a deconvolução e redução de ruído de imagens, focando em um método que combina as vantagens das representações nos domínios da frequência e wavelet.

2 Metodologia

2.1 Definição do Problema

Queremos restaurar a imagem original $x \in \mathbb{R}^n$ a partir de uma observação degradada $y \in \mathbb{R}^m$, relacionada por um operador linear $H \in \mathbb{R}^{m \times n}$. Em geral, assume-se que $m = n = N$.

A formulação matemática do problema é:

$$y = Hx + n,$$

onde n representa ruído aditivo gaussiano com média zero e variância σ^2.

Isto é, assumimos que:

$$n \sim \mathcal{N}(0, \sigma^2 I).$$
$$p(n) = \mathcal{N}(n|0, \sigma^2 I).$$

Neste trabalho, focamos em problemas onde H modela uma convolução periódica (e.g. os pixels na borda direita são tratados como vizinhos dos pixels na borda esquerda) e espacialmente invariante ("borrão"é o mesmo em toda a imagem). Nesse caso, a matriz H é do tipo bloco-circulante e, por essa propriedade, pode ser diagonalizada pela Transformada Discreta de Fourier (DFT) bidimensional:

$$H = U^H D U$$

Isto é, estamos diagonalizando o operador de convolução H usando a Transformada Discreta de Fourier Bidimensional (2D DFT). Dessa forma, U é a matriz que representa a transformada, U^H é a transposta conjugada e também a inversa de U, e por fim D é uma matriz diagonal que contém os coeficientes DFT do operador de convolução representado por H.

2.2 Abordagem no Domínio da Frequência (FFT)

As definições feitas na seção anterior implicam que a multiplicação Hx pode ser feita no domínio de Fourier:

$$Hx = U^H D U x$$

Se H for inversível, então a estimativa de x é facilmente obtida ao fazer:

$$\hat{x} = U^H D^{-1} U y$$

No entanto, na maioria dos casos, H não é inversível ou é dita uma matriz mal condicionada [2] (i.e., a diagonal de D tem valores muito pequenos), o que leva à amplificação do ruído na estimativa anterior. Para contornar esse problema, utiliza-se um método de regularização. Uma escolha comum é o estimador de máxima verossimilhança penalizado (MPLE), ou, sob uma perspectiva Bayesiana, o estimador máximo a posteriori (MAP):

$$\hat{x} = \arg\max_x \{\log p(y|x) - pen(x)\}$$

onde $p(y|x) = \mathcal{N}(y|Hx, \sigma^2 I)$ é a função de verossimilhança e $pen(x)$ é uma função de penalidade (ou -log da priori, $p(x)$). Se a priori sobre a imagem $p(x)$ for assumida como uma gaussiana com média μ e matriz de covariância G, a estimativa MAP pode ser escrita como:

$$\hat{x} = \arg\max_{x} \{ \frac{-1}{\sigma^2} ||Hx - y||^2 - (x - \mu)^H G^{-1} (x - \mu) \} \tag{1}$$

$$= \mu + G H^H (\sigma^2 I + H G H^H)^{-1} (y - H\mu) \tag{2}$$

Para que a solução MAP na equação (2) seja computacionalmente eficiente, assume-se que a imagem original x pode ser modelada como um campo gaussiano estacionário com condições de contorno periódicas. Essa premissa implica que a matriz de covariância da priori, G, também é bloco-circulante e, portanto, também é diagonalizada pela DFT. Podemos escrever $G = U^H C U$, onde C é uma matriz diagonal.

Com isso, a equação (2) pode ser implementada de forma eficiente no domínio da frequência, resultando no conhecido filtro de Wiener:

$$\hat{x} = \mu + U^H C D^H (\sigma^2 I + D C D^H)^{-1} (Uy - DU\mu)$$

Apesar da eficiência, essa abordagem tem uma limitação fundamental: imagens do mundo real não são bem modeladas por campos gaussianos estacionários. A energia do sinal de uma imagem típica não se concentra no domínio de Fourier, o que torna difícil separar o sinal do ruído de forma eficaz. Essa limitação motiva a busca por representações mais adequadas, como o domínio Wavelet.

2.3 Abordagem no Domínio Wavelet (DWT)

A grande vantagem das transformadas wavelet é que elas fornecem representações esparsas para imagens, ou seja, a maior parte da energia da imagem é capturada por um pequeno número de coeficientes grandes, enquanto a maioria dos coeficientes é próxima de zero.

Nesta abordagem, a imagem x é re-expressa em termos de seus coeficientes wavelet θ através da Transformada Wavelet Discreta (DWT) Inversa, denotada por W:

$$x = W\theta$$

O critério de estimação MAP é então formulado no domínio wavelet:

$$\hat{\theta} = \arg\max_{\theta}\{\log p(y|\theta) - pen(\theta)\} = \arg\max_{\theta}\{-\frac{||y - HW\theta||^2}{2\sigma^2} - pen(\theta)\}$$

A função de penalidade $pen(\theta)$ é escolhida para induzir a esparsidade dos coeficientes θ.

No entanto, essa formulação apresenta um grande desafio computacional. O operador combinado HW não é, em geral, bloco-circulante, o que impede a diagonalização eficiente via FFT que era possível com H isoladamente. Além disso, HW não é uma matriz ortogonal, o que descarta o uso de regras de limiarização (thresholding) coeficiente a coeficiente, que são muito eficientes em problemas de remoção de ruído (denoising).

2.4 O Algoritmo Expectation-Maximization (EM)

O problema que queremos resolver, após entendermos os conceitos das seções anteriores, é:

$$y = HW\theta + n$$

No entanto, como já foi visto na seção anterior, esse tipo de formulação traz diversos imbróglios.

Para superar as dificuldades da abordagem wavelet, Figueiredo e Nowak [3] propõem um algoritmo de Expectation-Maximization (EM). A chave do método é introduzir uma variável latente z que desacopla a deconvolução da remoção de ruído (denoising).

O modelo anterior é decomposto em duas etapas:

$$\begin{cases} z = W\theta + \alpha n_1 \\ y = Hz + n_2 \end{cases}$$

onde $n_1 \sim \mathcal{N}(0, I)$ e $n_2 \sim \mathcal{N}(0, \sigma^2 I - \alpha^2 HH^T)$ são ruídos gaussianos independentes. O parâmetro α é escolhido tal que a matriz de covariância de n_2 seja semipositiva definida. Note que se tivéssemos acesso a z, o problema seria simplesmente remover o ruído αn_1 de z para estimar $W\theta$, um problema de denoising padrão.

O algoritmo EM itera entre dois passos para encontrar a estimativa $\hat{\theta}$:

- **E-step (Expectation):** Calcula a esperança condicional do dado completo, o que, neste caso, se resume a estimar a variável latente z dado a observação y e a estimativa atual da imagem, $\hat{x}^{(t)} = W\hat{\theta}^{(t)}$. A atualização é dada por:

$$\hat{z}^{(t)} = E[z|y, \hat{\theta}^{(t)}] = \hat{x}^{(t)} + \frac{\alpha^2}{\sigma^2} H^T (y - H\hat{x}^{(t)})$$

  Este passo é computacionalmente eficiente, pois envolve a matriz H e sua transposta H^T, que podem ser implementadas via FFTs no domínio da frequência.

- **M-step (Maximization):** Atualiza a estimativa dos coeficientes wavelet, $\hat{\theta}^{(t+1)}$. Neste trabalho, utilizaremos a **penalização** l_1, que é amplamente usada por sua capacidade de induzir esparsidade. A função de penalidade é dada por:

$$pen(\theta) = \tau ||\theta||_1 = \tau \sum_i |\theta_i|$$

  onde τ é um parâmetro de regularização positivo que controla a intensidade da esparsidade. O passo de maximização se torna:

$$\hat{\theta}^{(t+1)} = \arg\max_{\theta} \left\{ -\frac{||W\theta - \hat{z}^{(t)}||^2}{2\alpha^2} - \tau \sum_i |\theta_i| \right\}$$

  Como a matriz wavelet W é ortogonal ($||W\theta - \hat{z}^{(t)}||^2 = ||\theta - W^T\hat{z}^{(t)}||^2$), o problema se desacopla e pode ser resolvido coeficiente a coeficiente. A solução é a conhecida função *soft-thresholding*. Especificamente, cada componente $\hat{\theta}_i^{(t+1)}$ é obtido aplicando-se a seguinte regra aos coeficientes $\hat{\omega}^{(t)} = W^T\hat{z}^{(t)}$:

$$\hat{\theta}_i^{(t+1)} = \text{sgn}(\hat{\omega}_i^{(t)})(|\hat{\omega}_i^{(t)}| - \tau\alpha^2)_+$$

  onde $(\cdot)_+$ denota o operador da parte positiva, i.e., $(x)_+ = \max\{x, 0\}$, e o valor $\tau\alpha^2$ atua como o limiar (threshold) de decisão.

2.5 Esquema Geral do Algoritmo

O algoritmo combina o melhor dos dois mundos: a diagonalização do operador de convolução no domínio de Fourier e a modelagem esparsa de imagens no domínio wavelet. Cada iteração do algoritmo pode ser resumida da seguinte forma:

$$\hat{x}^{(t+1)} = \mathcal{P}\left(\hat{x}^{(t)} + \frac{\alpha^2}{\sigma^2} H^T (y - H\hat{x}^{(t)})\right)$$

onde $\mathcal{P}(\cdot)$ denota a operação de denoising no domínio wavelet: $\mathcal{P}(v) = W\mathcal{D}(W^T v)$, sendo $\mathcal{D}$ a função de encolhimento (shrinkage) aplicada aos coeficientes.

O processo é o seguinte:

1. **Inicialização:** Comece com uma estimativa inicial da imagem, $\hat{x}^{(0)}$, por exemplo, a imagem observada y ou uma estimativa do filtro de Wiener.

2. **Iteração $t+1$:**

   (a) **(E-step)** Calcule o resíduo $y - H\hat{x}^{(t)}$, aplique o operador H^T e adicione à estimativa atual $\hat{x}^{(t)}$ para obter $\hat{z}^{(t)}$. Esta etapa é realizada eficientemente com FFTs.

   (b) **(M-step)** Aplique a transformada wavelet (DWT) em $\hat{z}^{(t)}$, aplique a função de *soft-thresholding* nos coeficientes e, em seguida, aplique a transformada inversa (IDWT) para obter a nova estimativa da imagem, $\hat{x}^{(t+1)}$.

3. **Convergência:** Repita o passo 2 até que um critério de parada seja satisfeito, como a mudança relativa entre estimativas consecutivas ser menor que um limiar.

3 Resultados

Para avaliar o desempenho do algoritmo EM-Wavelet, realizamos um experimento numérico utilizando a imagem de teste "cameraman". A imagem original foi degradada por um borrão de movimento uniforme, modelado por um kernel de convolução de 9x9, e pela adição de ruído gaussiano branco, resultando em uma imagem observada com uma Relação Sinal-Ruído de Pico (PSNR) de 23.19 dB.

O algoritmo foi inicializado com uma estimativa gerada pelo filtro de Wiener (ver Apêndice A para uma explicação detalhada), que por si só já elevou o PSNR para 26.40 dB. Em seguida, o processo iterativo do EM foi executado. Os hiperparâmetros utilizados foram a wavelet 'haar' com 4 níveis de decomposição e um parâmetro de regularização $\lambda = 1$.

Figura 1: Imagem original, borrada e ruidosa e reconstruída.

O método demonstrou uma melhora progressiva a cada iteração, atingindo o melhor resultado na 25ª iteração. A imagem restaurada final alcançou um PSNR de 29.89 dB. Isso representa uma Melhoria na Relação Sinal-Ruído (SNRI) de 6.70 dB em relação à imagem degradada. A Figura 1 ilustra a qualidade visual da imagem original, da versão degradada e do resultado da restauração.

4 Conclusão

O algoritmo EM para deconvolução de imagens no domínio wavelet provou ser uma abordagem poderosa e computacionalmente viável. Ele combina a eficiência da FFT para lidar com a convolução e a capacidade das wavelets de modelar esparsamente o conteúdo de imagens. Os resultados experimentais, que mostraram um ganho de 6.70 dB em SNRI, demonstram a eficácia do método. A estrutura iterativa, que alterna entre um passo de filtragem no domínio da frequência e um passo de denoising no domínio wavelet, converge para uma solução ótima sob condições suaves, oferecendo resultados de alta qualidade na tarefa de restauração de imagens.

Referências

[1] Avinash C. Kak e Malcolm Slaney. *Principles of Computerized Tomographic Imaging*. New York: IEEE Press, 1988.

[2] Fikri Öztürk e Fikri Akdeniz. "Ill-conditioning and multicollinearity". Em: *Linear Algebra and Its Applications* 321.1-3 (2000), pp. 295–305.

[3] M.A.T. Figueiredo e R.D. Nowak. "An EM algorithm for wavelet-based image restoration". Em: *IEEE Transactions on Image Processing* 12.8 (2003), pp. 906–916. DOI: 10.1109/TIP.2003.814255.

A Filtro de Wiener

O Filtro de Wiener é um método clássico e computacionalmente eficiente para restaurar uma imagem que foi degradada por borrão e ruído. Embora o algoritmo principal deste trabalho utilize uma abordagem iterativa mais sofisticada, o Filtro de Wiener serve como um excelente ponto de partida.

A.1 Ideia Central

Imagine que uma imagem é borrada e ruidosa. Uma abordagem ingênua seria tentar reverter o borrão de forma exata. No entanto, essa "filtragem inversa"tem um efeito colateral grave: ela amplifica o ruído que já existe na imagem, muitas vezes tornando o resultado final pior do que a imagem degradada.

O Filtro de Wiener resolve este problema de forma mais inteligente. Ele não busca apenas reverter o borrão, mas sim encontrar um equilíbrio ótimo entre a nitidez da imagem e a supressão do ruído. O seu objetivo não é ser um filtro inverso perfeito, mas sim produzir a estimativa mais "limpa"possível, no sentido de minimizar o erro quadrático médio entre a imagem estimada e a original.

A.2 Estratégia: Equilíbrio no Domínio da Frequência

A implementação do filtro é mais intuitiva no domínio da frequência, para onde a imagem é levada usando a Transformada de Fourier. Neste domínio, a imagem é decomposta em suas frequências constituintes.

Para cada frequência, o filtro avalia a **Relação Sinal-Ruído (SNR)**, que é a proporção entre a energia do sinal da imagem original e a energia do ruído. Com base nessa avaliação, ele decide como agir:

- **Em frequências com SNR alta:** Onde o sinal da imagem é muito mais forte que o ruído, o filtro age de forma agressiva para reverter o borrão, pois há pouco risco de amplificar o ruído.
- **Em frequências com SNR baixa:** Onde o ruído domina o sinal, o filtro atenua ou até mesmo zera essa frequência. Isso evita que o ruído seja amplificado, mesmo que signifique não reverter completamente o borrão naquela parte específica da imagem.

Essencialmente, o filtro faz uma troca inteligente: sacrifica um pouco da nitidez nas áreas ruidosas para obter uma imagem globalmente mais limpa.

A.3 Aplicação Prática e Papel no Algoritmo

Na prática, não conhecemos a energia exata da imagem original para calcular a SNR perfeitamente. Por isso, como visto no código deste trabalho, utilizamos uma estimativa para a potência do sinal e a variância conhecida do ruído para criar um *termo de regularização*. Esse termo controla o equilíbrio do filtro.

Devido à sua rapidez (requer apenas algumas operações de FFT) e à sua capacidade de fornecer uma estimativa inicial de boa qualidade, o Filtro de Wiener é frequentemente usado como o ponto de partida para algoritmos iterativos mais complexos. Neste trabalho, ele gera a estimativa $\hat{x}^{(0)}$, que é então refinada progressivamente pelos passos E e M do algoritmo principal para alcançar um resultado superior.