HOMEWORK 5
DECONVOLUTION

CSED551 – COMPUTATIONAL PHOTOGRAPHY
FALL 2024
SUNGHYUN CHO (S.CHO@POSTECH.AC.KR)

OVERVIEW

In this homework, you will implement two deconvolution techniques: 1) wiener deconvolution, and 2) TV deconvolution. For the second one, you need to derive an energy function for a classical non-blind deconvolution method, and implement its optimization.

NON-BLIND DECONVOLUTION

Image deblurring is a long-standing image restoration problem in computer vision and image processing. Most classical image deblurring methods model blur using a convolution operation. Specifically, they model a blurred image b as:

$$b = k * l + n$$

where k is a blur kernel or a point spread function (PSF), l is a latent sharp image, and n is additive noise, which is often assumed to be white Gaussian such that $n \sim N(0, \sigma^2)$ where `white' noise means the noise variance is constant for all frequency bands as will be explained later. * is the convolution operator. Then, image deblurring becomes a deconvolution problem, in which we invert the convolution operation. If we know both b and k, the problem is called *non-blind* deconvolution. On the other hand, if we know only b, the problem is called *blind* deconvolution.

Non-blind deconvolution serves as a fundamental building block in many applications, e.g., removing motion blur, out-of-focus blur, and chromatic aberration.

INVERSE FILTERING

The most naïve approach to non-blind deconvolution is the inverse filtering. We can derive the inverse filtering by applying the Fourier transform to the blur model above. Specifically, we can obtain:

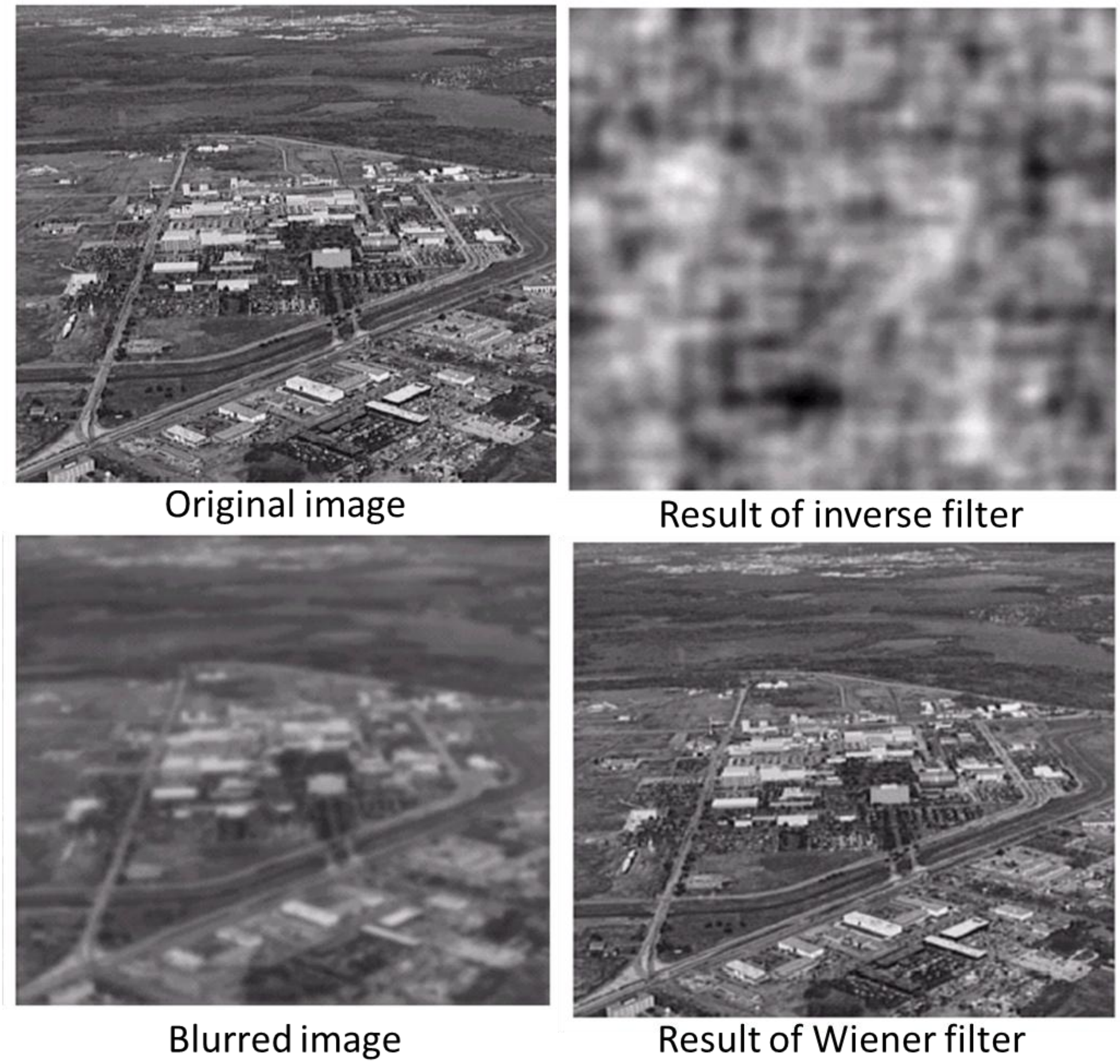
$$B = K \otimes L + N$$

where B is the Fourier transform of b, i.e., $B = \mathcal{F}(b)$ where $\mathcal{F}$ represents the forward Fourier transform function. K, L, and N are also the Fourier transforms of k, l, and n, respectively. $\otimes$ is the element-wise multiplication operator. Ignoring noise N, we can obtain L as:

$$L \approx \frac{B}{K}$$

where $\frac{B}{K}$ is the pixel-wise division between B and K in the frequency domain. Then, by computing the inverse Fourier transform, we can obtain a deconvolution result. As this technique directly applies the inverse of a blur kernel k to the input blurred image b in the frequency domain, it is called the inverse filtering.

The inverse filtering, however, cannot produce satisfying results. The blur kernel K in the frequency domain may have elements whose values are zero or close to zero, and dividing B with such elements leads to totally corrupted images with significantly amplified noise. To overcome this issue, we need proper handling of noise and regularization.



WIENER DECONVOLUTION

One classical approach to non-blind deconvolution is the wiener deconvolution. The wiener deconvolution is derived by minimizing the expected error in the frequency domain. Specifically, we want to find a deconvolution filter H in the frequency domain such that

$$\min_H E[\|L - HB\|^2]$$

where HB is the pixel-wise multiplication between H and B in the frequency domain. By solving the equation above, we can obtain a deconvolution filter H such that HB is a deconvolution result with the least expected squared error.

By replacing B in the equation above with $B = KL + N$, we can obtain:

$$\min_H E[\|L - H(KL + N)\|] = \min_H E[\|(1 - HK)L - HN\|]$$

Then, we can extend the squares and obtain:

$$\min_H E[\|(1 - HK)L - HN\|] = \min_H \{\|1 - HK\|^2 E[\|L\|^2] - 2(1 - HK)E[LN] + \|H\|^2 E[\|N\|^2]\}$$

To further continue the derivation, we assume that noise N is zero-mean Gaussian noise, which is independent to L (in practice, it's not, but many image restoration works use this assumption to simplify their problems.). Then $E[LN] = E[L]E[N]$. As N is zero-mean, $E[N] = 0$. As a result, we obtain:

$$\min_H \{\|1 - HK\|^2 E[\|L\|^2] + \|H\|^2 E[\|N\|^2]\}$$

By differentiating the equation above with respect to H, and setting it to zero, we can solve the equation for H:

$$\Rightarrow -2(1 - HK)E[\|L\|^2] + 2HE[\|N\|^2] = 0$$

$$\Rightarrow H = \frac{KE[\|L\|^2]}{K^2 E[\|L\|^2] + E[\|N\|^2]} = \frac{K^2}{K^2 + \frac{E[\|N\|^2]}{E[\|L\|^2]}} \cdot \frac{1}{K}$$

Finally, multiplying H to B, we can obtain the formulation of the wiener deconvolution:

$$l = \mathcal{F}^{-1}\left(\frac{|\mathcal{F}(k)|^2}{|\mathcal{F}(k)|^2 + \frac{1}{SNR(\omega)}} \cdot \frac{\mathcal{F}(b)}{\mathcal{F}(k)}\right)$$

where all the multiplication and division operations are pixel-wise operations. $SNR(\omega)$ is the signal-to-noise ratio at each frequency ω, which is required to be estimated. $SNR(\omega)$ is defined as:

$$SNR(\omega) = \frac{signal\ variance\ at\ \omega}{noise\ variance\ at\ \omega}$$

Natural images tend to have signal variance that scales as $\frac{1}{\omega^2}$. Noise variance tends to be a constant, i.e., $\sigma_N(\omega) = constant$ for all ω. This is called white noise. As a result, we can assume

$$SNR(\omega) = \frac{1}{c\omega^2}$$

and obtain a simplified version of the wiener deconvolution as:

$$l = \mathcal{F}^{-1}\left(\frac{|\mathcal{F}(k)|^2}{|\mathcal{F}(k)|^2 + c\omega^2} \cdot \frac{\mathcal{F}(b)}{\mathcal{F}(k)}\right)$$

where c is a constant. In practice, a more simplified version is often used, which is defined as:

$$\boldsymbol{l = \mathcal{F}^{-1}\left(\frac{|\mathcal{F}(k)|^2}{|\mathcal{F}(k)|^2 + c} \cdot \frac{\mathcal{F}(b)}{\mathcal{F}(k)}\right)} \quad \textbf{(1)}$$

By setting $c = 0$, the equation above reduces to the inverse filtering. In this homework, you need to implement this simplified version of the wiener deconvolution. You also need to discuss the effect of the parameter c with example images.

TOTAL-VARIATION (TV) DECONVOLUTION

Non-blind deconvolution is an ill-posed problem. Due to the loss of information caused by the noise n, even if we know k, there can be an infinite number of solutions for l. To resolve such ill-posedness, most methods adopt a regularization term or a prior on the latent image l. One of the most widely-used prior is the total variation, which is defined as:

$$E_{\mathrm{TV}}(l) = \|\partial_{\mathrm{x}} * l\|_1 + \left\|\partial_{\mathrm{y}} * l\right\|_1$$

where $\|f\|_1$ is the L_1 norm of an image f which is defined as:

$$\|f\|_1 = \sum_{x,y} |f(x,y)|$$

and ∂_{x} and ∂_{y} are partial derivative filters along the x- and y-axes, e.g., $\partial_{\mathrm{x}} = [0,-1,1]$ and $\partial_{\mathrm{y}} = [0,-1,1]^{\mathrm{T}}$. With the prior $E_{\mathrm{TV}}(l)$, we can define an energy function for non-blind deconvolution as:

$$E(l) = \|b - k * l\|^2 + \alpha E_{\mathrm{TV}}(l)$$

where α is a parameter to control the strength of the regularization term.

You can also formulate the non-blind deconvolution problem as a maximum-a-posteriori estimation problem, and derive the energy function above from the MAP formulation. Specifically, you can define a posterior distribution:

$$p(l|b,k) \propto p(b|l,k)p(l)$$

where $p(b|l,k)$ is a likelihood, and $p(l)$ is a prior distribution.

Optimization

The regularization term $E_{\mathrm{TV}}(l)$ in the energy function $E(l)$ is a non-quadratic term, so the optimization of $E(l)$ needs a non-linear optimization algorithm such as the iterative reweighted least-squares (IRLS) method, which is explained in the course material (see 14.global_optimization.pdf). At each iteration of the IRLS method, you solve a least-squares problem, which is formulated as:

$$\mathbf{l}^* = \arg\min_{\mathbf{l}} \left\{ \|\mathbf{b} - \mathbf{K}\mathbf{l}\|^2 + \alpha\left(\mathbf{l}^{\mathrm{T}}\mathbf{D}_x^{\mathrm{T}}\mathbf{W}_x\mathbf{D}_x\mathbf{l} + \mathbf{l}^{\mathrm{T}}\mathbf{D}_y^{\mathrm{T}}\mathbf{W}_y\mathbf{D}_y\mathbf{l}\right)\right\}$$

where $\mathbf{b}$ and $\mathbf{l}$ are vector representations of b and l, respectively. $\mathbf{K}$ is a matrix representing the convolution operation with k. The equation above can be re-written as:

$$\mathbf{l}^* = \arg\min_{\mathbf{l}} \left\{ \mathbf{l}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{l} - 2\mathbf{l}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{b} + \mathbf{b}^{\mathrm{T}}\mathbf{b} + \alpha\left(\mathbf{l}^{\mathrm{T}}\mathbf{D}_x^{\mathrm{T}}\mathbf{W}_x\mathbf{D}_x\mathbf{l} + \mathbf{l}^{\mathrm{T}}\mathbf{D}_y^{\mathrm{T}}\mathbf{W}_y\mathbf{D}_y\mathbf{l}\right)\right\}$$

As the right-hand side of this equation is a quadratic function with respect to $\mathbf{l}$, its solution can be found by setting its derivative to zero. Specifically, the derivative of the right-hand side is derived as:

$$2\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{l} - 2\mathbf{K}^{\mathrm{T}}\mathbf{b} + 2\alpha\left(\mathbf{D}_x^{\mathrm{T}}\mathbf{W}_x\mathbf{D}_x + \mathbf{D}_y^{\mathrm{T}}\mathbf{W}_y\mathbf{D}_y\right)\mathbf{l}$$

By setting the derivative to zero, we can derive the following equation:

$$\left(\mathbf{K}^{\mathrm{T}}\mathbf{K} + 2\alpha\left(\mathbf{D}_x^{\mathrm{T}}\mathbf{W}_x\mathbf{D}_x + \mathbf{D}_y^{\mathrm{T}}\mathbf{W}_y\mathbf{D}_y\right)\right)\mathbf{l} = \mathbf{K}^{\mathrm{T}}\mathbf{b}$$

Note that this equation has the same form as $\mathbf{A}\mathbf{x} = \mathbf{b}$, which means that you can find a solution $\mathbf{l}$ using a linear solver. However, the size of the matrix $\left(\mathbf{K}^{\mathrm{T}}\mathbf{K} + 2\alpha\left(\mathbf{D}_x^{\mathrm{T}}\mathbf{W}_x\mathbf{D}_x + \mathbf{D}_y^{\mathrm{T}}\mathbf{W}_y\mathbf{D}_y\right)\right)$ is huge, so it is hard to directly solve the linear system. Instead, you can use an iterative solver such as the conjugate gradient method. For more details, refer to the reference code that is provided with this homework document.

REQUIREMENTS

In this homework, you need to implement two non-blind deconvolution algorithms that remove blur from a given blurred image b and a blur kernel k. Specifically, you are required to:

- Wiener deconvolution
  - Implement the wiener deconvolution according to Equation (1).
  - Discuss the effect of the parameter c with example images.
- TV deconvolution
  - Formulate the non-blind deconvolution problem as a maximum-a-posteriori (MAP) problem, and derive the energy function $E(l)$ described above.
  - Explain the relationship between the MAP formulation, and the energy function. Specifically, you need to explain how α can be derived from your MAP formulation.
  - You need to implement an algorithm to optimize the energy function. The energy function above involves $E_{\text{TV}}(l)$, which is non-quadratic. Thus, you need to implement a non-linear optimization algorithm. In this homework, you can implement the iterative-reweighted-least-squares (IRLS) method, which was presented in our course material. See the file '14.global_optimization.pdf'.

The IRLS optimization for non-blind deconvolution needs to compute convolution operations with a blur kernel several times, which can take an excessive amount of computation time for large blur kernels. To accelerate the computation, you can use the convolution theorem.

Your report must include:

- Derivation of the energy function from the MAP formulation and discussion about their relationship
- Detailed discussion on your implementation of the wiener deconvolution and TV deconvolution
- Your results with a detailed discussion including the discussion on the parameter c of the wiener deconvolution
- Limitations of the techniques that you found

You must upload a single zip file that contains the following to the LMS:

- code/ - a directory containing all your code for this assignment
- images/ - a directory containing your input images and their results
- report.pdf – your report as a PDF file

Due: Dec. 20th, 23:59

Penalty for late submission

- 1 day: 70%
- 2 days: 30%
- 3 days: 0%

RUBRIC

- 30 pts: wiener deconvolution
- 20 pts: Energy function derivation for the TV deconvolution
- 30 pts: Implementation of the TV deconvolution
- 20 pts: Report

Your homework will be scored based on your report, and I am not going to compile or run your code. Thus, your report must include all necessary details of your implementation and results.