Restoration

Introduction

The principal goal of restoration methods (French: *méthodes de restauration*) is to improve an image by using methods based on objective considerations.

Restoration methods need knowledge of the phenomenon that has degraded the image. This knowledge gives a mathematical model of the degradation. The degradation phenomenon is “inversed” to recover the original image: hence restoration methods are said to be an *inverse problem* (French: *problème inverse*). This approach usually involves formulating a *criterion of goodness* (French: *critère de qualité*) to help the search for an optimal estimation: a criterion is a mathematical function whose minimum corresponds to the best possible restoration.

Denoising

Denoising (French: *débruitage*) consists in reducing noise in an image.

Noise Sources

The main sources of noise in digital images are during the acquisition (quantity of photons collected too low, sensor temperature…) or during any transmission (echoes and atmospheric distortions in wireless communication). In some cases, noise is also considered to model the inaccuracies in the mathematical model of image formation, the latter being necessarily different compared to reality, like any physical model!

The noise is by nature a random phenomenon, it is modelled by a probability density which represents the intensity distribution of the noise.

In the following, we denote by y the observed (and noisy) image, b the noise and x the non-noisy image.

Additive Gaussian white noise

Additive white Gaussian noise (AWGN, in French: *bruit blanc gaussien additif*) models each pixel (m, n) of the observation y by the sum of the pixel (m, n) of the noiseless image x and of a pixel of the noise b:

$$\forall m, n \quad y(m, n) = x(m, n) + b(m, n)$$

where $b(m, n) \sim \mathcal{N}(0, \sigma^2)$.

This model is simple and facilitates calculations. It is used in most applications, including photography.

Poisson Noise

Poisson noise (also called *shot noise*, in French: *bruit de Poisson*) models the acquisition of photons on a photosite. The number of photons is random and depends on the illumination. The corresponding Poisson process has a mean equals to the illumination. The intensity of each pixel (m, n) of the observation y is:

$$\forall\, m, n \qquad y(m,n) \sim \mathcal{P}\big(x(m,n)\big).$$

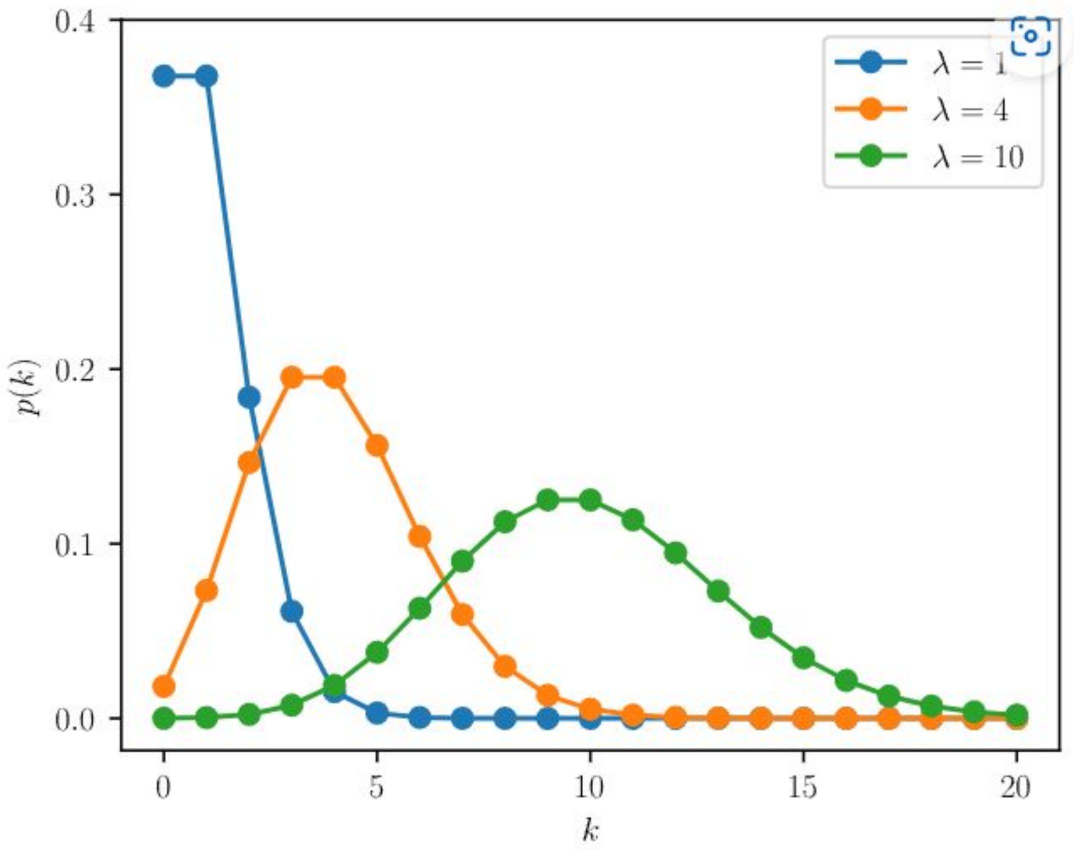
This model is used in the case of acquisitions with a low number of photons, for example in astronomy.

Fig. 73 represents the Poisson distribution for three values of parameter λ.

As a reminder,

the Poisson distribution $\mathcal{P}(\lambda)$ writes

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$



The mean and variance of the Poisson distribution are both equal λ, which depends on the number of incident photons. So, the noise b depends on the noiseless image x. Moreover, when λ increases, the Poisson distribution tends towards a Gaussian distribution, implying that AWGN becomes a good model of Poisson noise, provided that enough photons are collected.

Salt-and-pepper noise

Salt-and-pepper noise (French: *bruit poivre et sel*), also called less poetically *impulse noise*, models saturated or dead pixels (due to photosite malfunction or saturation).

$$\forall\, m, n \quad y(m,n) = \begin{cases} x_{\min} & \text{with probability } p_{\min}, \\ x_{\max} & \text{with probability } p_{\max}, \\ x(m,n) & \text{with probability } 1 - p_{\min} - p_{\max}. \end{cases}$$

where $x_{\min}$ and $x_{\max}$ are the intensity minimum and maximum.

Noise power

Fig. 74 illustrates the effect of the previous noises on the same image. One can observe that:

- for Gaussian noise, the whole image is affected in the same way by the noise,
- for Poisson noise, the lighter parts are noisier than the dark parts,
- for impulse noise, only a few pixels are modified and they are replaced by black or white pixels.

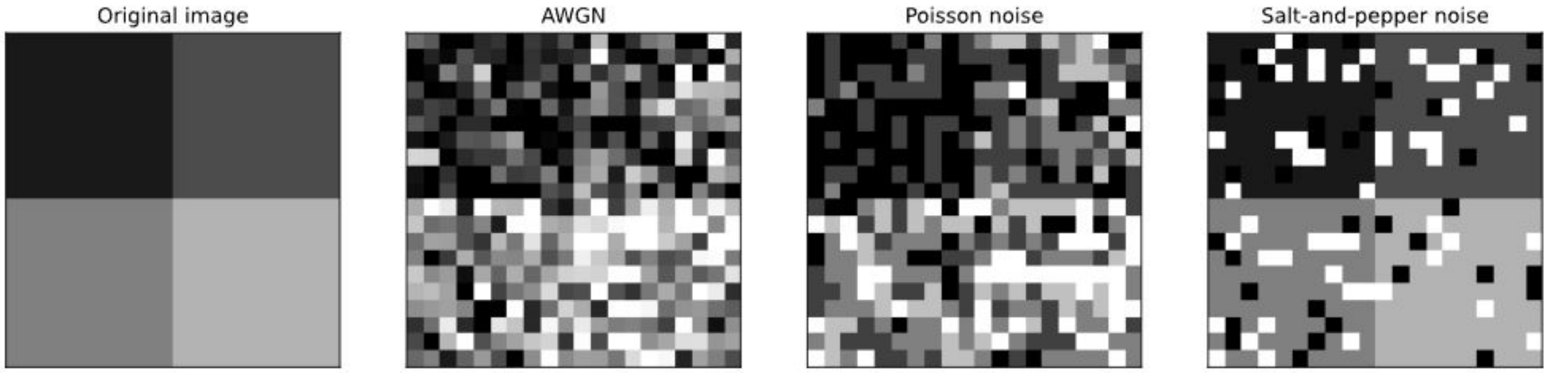


Fig. 74 Example of different types of noise (with almost the same power).

Signal-to-noise ratio (SNR, in French: RSB for *rapport signal-sur-bruit*) is a measure of the noise level. It is defined as the ratio between the power of the non-noisy image over the power of the noise, where the power of an image x is defined by:

$$P_x = \frac{1}{M \times N} \sum_{m,n} x(m,n)^2$$

Because SNR is most often expressed on a logarithmic scale (unit: decibel), it is also defined as:

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{\sum_{m,n} x(m,n)^2}{\sum_{m,n} b(m,n)^2} \right)$$

For additive noise, another measure exists: the peak signal-to-noise ratio (PSNR) is the ratio of the squared dynamics of the non-noisy image (difference between maximum intensity and minimum intensity) to the power of the noise:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\Delta x^{\;2}}{\frac{1}{M\times N}\sum_{m,n} b(m,n)^2}\right)$$

Fig. 75 represents the same image corrupted with additive white Gaussian noise, at different SNR and PSNR. As you can see, when the RSB or the PSNR increases, the noise decreases!

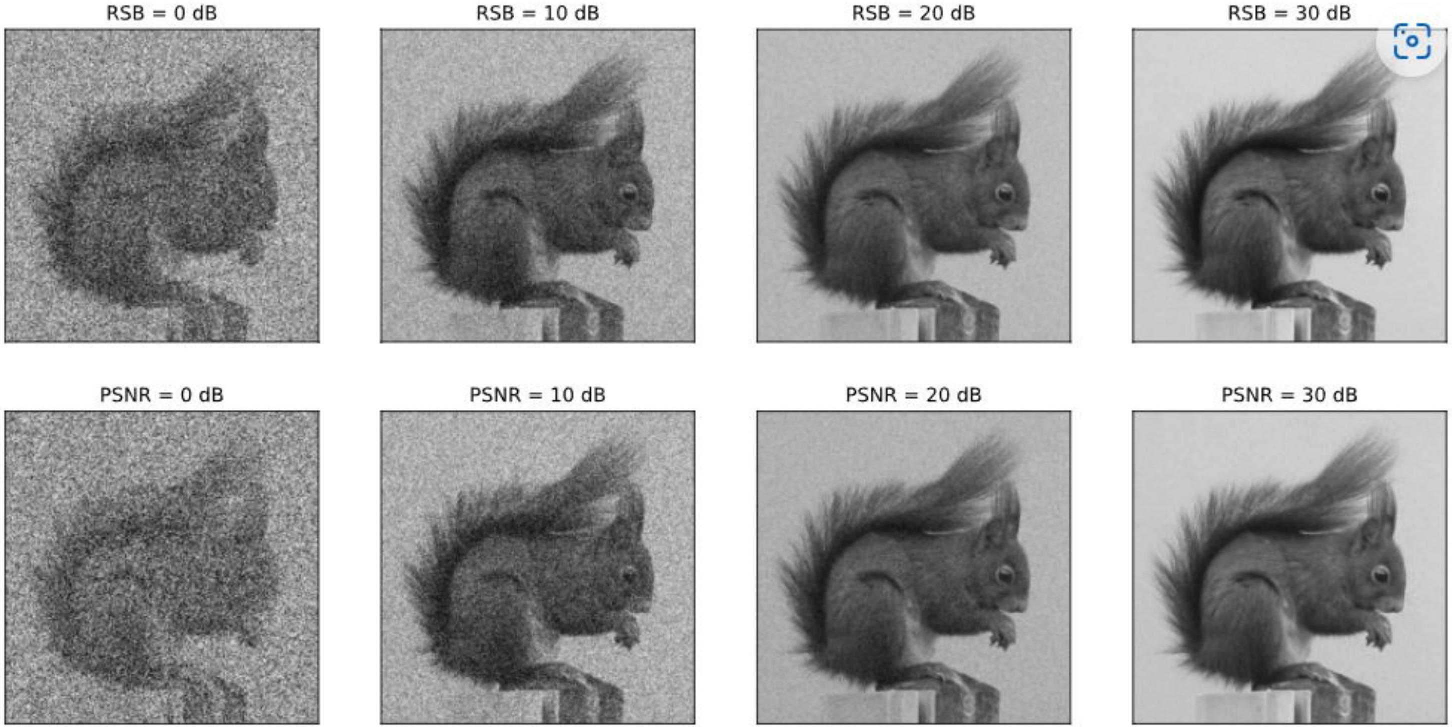


Fig. 75 Noisy image at different RSB and PSNR. #

Mean filter

The mean filter (French: *filtre moyenneur*) is a very simple denoising method. Each pixel (m, n) of the denoised image $\widehat{x}$ is the average of the pixels of the noisy image y around (m, n):

$$\forall m,\, n \quad \widehat{x}(m, n) = \frac{1}{|V_{m,n}|} \sum_{(u,v) \in V_{m,n}} y(u, v)$$

where

- $V_{m,n}$ is the neighbourhood, that is the set of pixels around (m, n);
- $|V_{m,n}|$ is the cardinality of $V_{m,n}$, that is, the number of pixels in the neighbourhood.

Fig. 76 illustrates the effect of the mean filter for different sizes of the neighbourhood. If the neighbourhood grows, then the noise decreases but at the same time the image becomes more blurry.

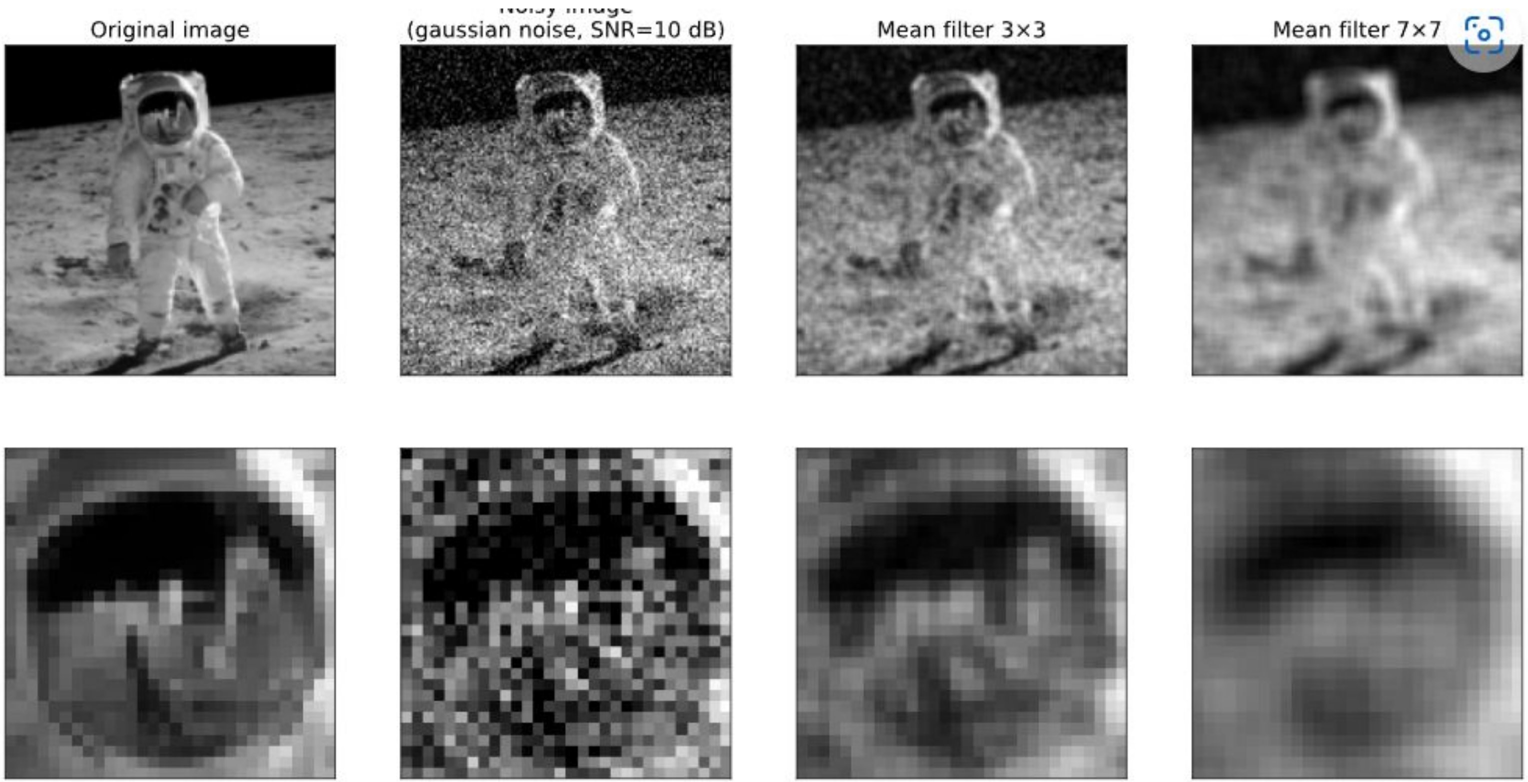


Fig. 76 Effect of the size of the mean filter.

The mean filter can be expressed with a convolution product. Indeed, consider the case where the neighbourhood is a square of size $N \times N$ pixels, then the definition of the mean filter gives:

$$\widehat{x}(m,n) = \frac{1}{N^2} \sum_{(u,v)\in V_{m,n}} y(u,v) = \sum_{u,v} g(m-u,n-v)y(u,v)$$

where

$$g(u,v) = \begin{cases} 1/N^2 & \text{if } u \in \{-\frac{N}{2},\dots,\frac{N}{2}\} \text{ and } v \in \{-\frac{N}{2},\dots,\frac{N}{2}\} \\ 0 & \text{otherwise} \end{cases}$$

This definition can be extended to any type of kernel g! For example, Fig. 77 gives the result for two different kernels.

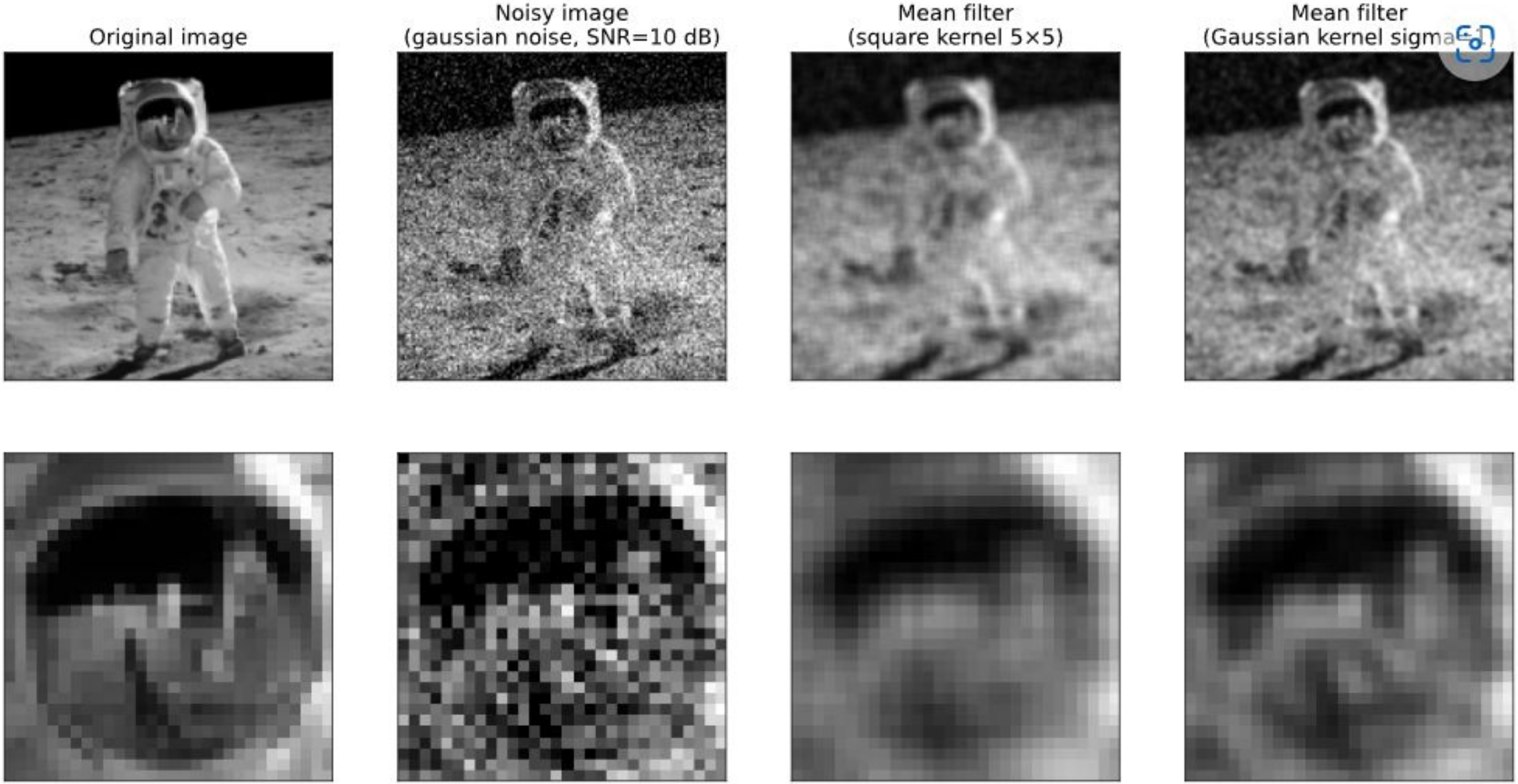


Fig. 77 Two mean filters with different kernels. #

In summary:

- the mean filter calculates the average of the pixels in a neighbourhood,
- the mean filter can be written as a convolution,
- the noise is reduced by averaging the intensities but the image is blurred.

Median filter

The median of a set of numbers is the element m of the set such that there are as many numbers smaller than m as there are numbers larger than m. For example, the median of $\{1,\ 2,\ 4,\ 8,\ 16\}$ is 4.

The median filter (French: *filtre médian*) is defined by:

$$\forall m,\ n \quad \widehat{x}(m,n) = \mathrm{median}\Big(\{y(u,v) \mid (u,v) \in V_{m,n}\}\Big)$$

The median filter is excellent for denoising an image in the case of salt-and-pepper noise because it does not blur the image, as a mean filter would do.

Despite its name, the median filter is not a filter because it does not respect the linearity property. Therefore it cannot be written as a convolution.

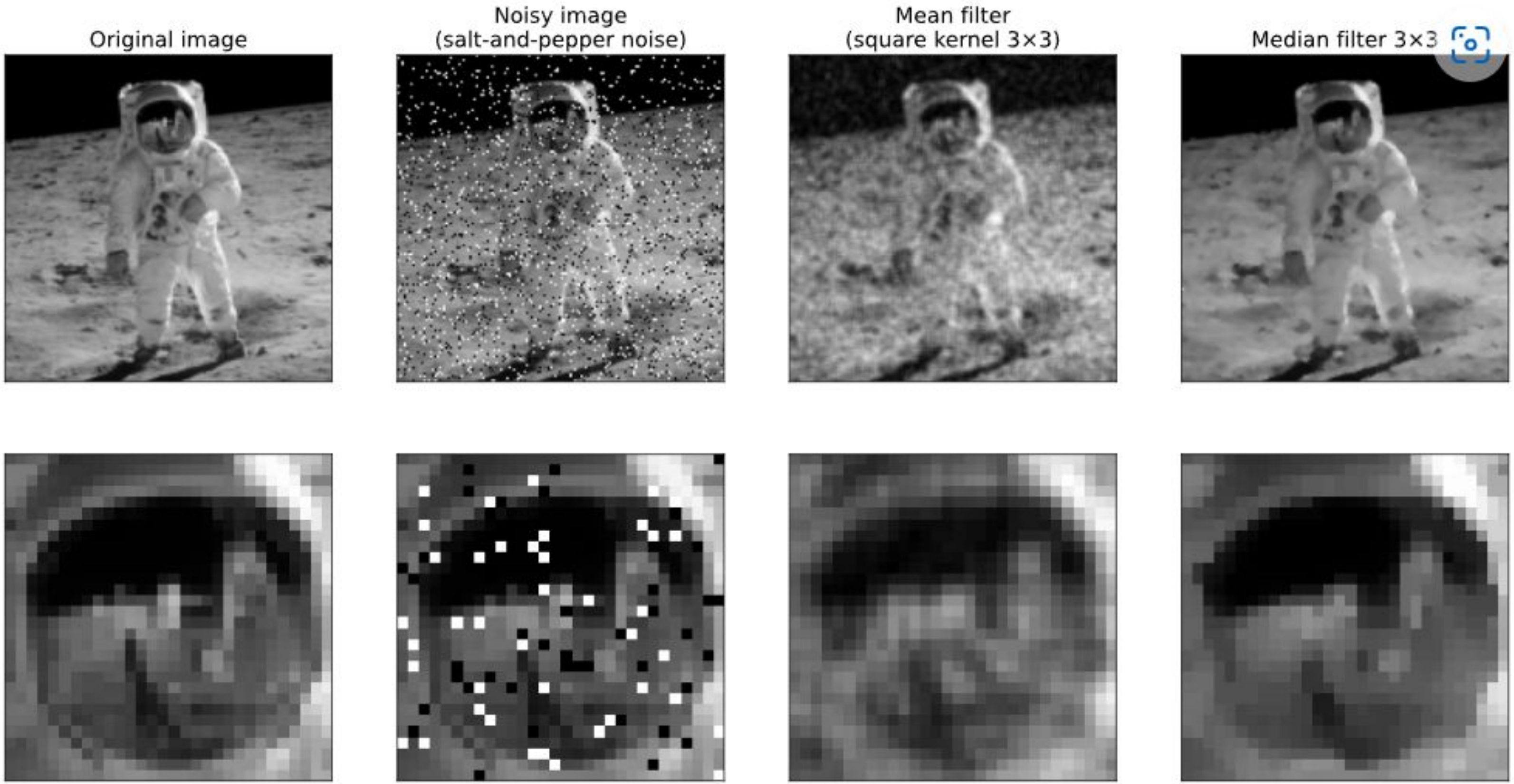


Fig. 78 Comparison between a mean filter and a median filter, on an image with salt-and-pepper noise. #

Periodic noise filtering

Periodic noises are characterized by structures in the Fourier transform. These structures can be removed in the Fourier domain by cancelling the coefficients using a mask. The denoised image is then obtained by an inverse Fourier transform.

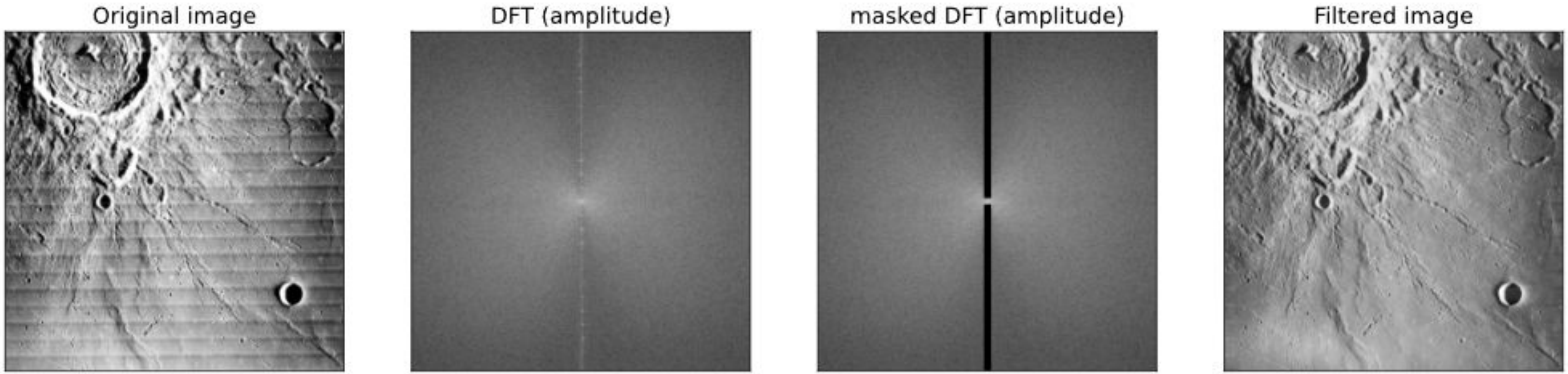


Fig. 79 Filtering a periodic noise on a photograph of the Moon: the image is cleaned of periodic image artefacts.

TV regularization

From a certain point of view, the goal of denoising is to obtain an image $\widehat{x}$ not only with small variations in intensity between pixels but also close to the observation y. TV regularization (*total variation*, French: *variation totale*) [Rudin et al. 1992, Chambolle 2004] is a denoising method that describes these two objectives by mathematical functions (the so-called “criteria”).

- The wish to have a denoised image $\widehat{x}$ close to the observation y results in a "data-fit" criterion (French: *critère d'adéquation aux données*) which measures the difference between x and y. A classic choice to measure this difference is the least-squares criterion:

$$E(x, y) = \sum_{m,n} (y(m, n) - x(m, n))^2$$

Indeed, for E to decrease, we must find an image x close to y.

- The desire to have an image with small variations in intensity results in a "regularization" criterion (French: *critère de régularisation*) which measures the difference between the neighbouring pixels of x. A simple choice is the "total variation":

$$R(x) = \sum_{m,n} |x(m+1,n) - x(m,n)| + \sum_{m,n} |x(m,n+1) - x(m,n)|$$

  So, for R to decrease, the difference between two consecutive pixels, whether they are in a row or column, must be small. This implies the image x to be nearly constant in intensity.

The goal is then to find the image x which minimizes both the data-fit and the regularization. Mathematically, one look for the image x which minimizes $E(x, y) + \lambda R(x)$, where λ is the "regularization parameter" (French: *paramètre de régularisation*) which is used to adjust the compromise between the two criteria. The value of λ is chosen by the user. Mathematically, we write:

$$\widehat{x} = \arg\min_x E(x, y) + \lambda R(x)$$

This comes to an optimization problem, and there are a large number of algorithms to minimize $E(x, y) + \lambda R(x)$. The choice and description of these algorithms are beyond the scope of the course.

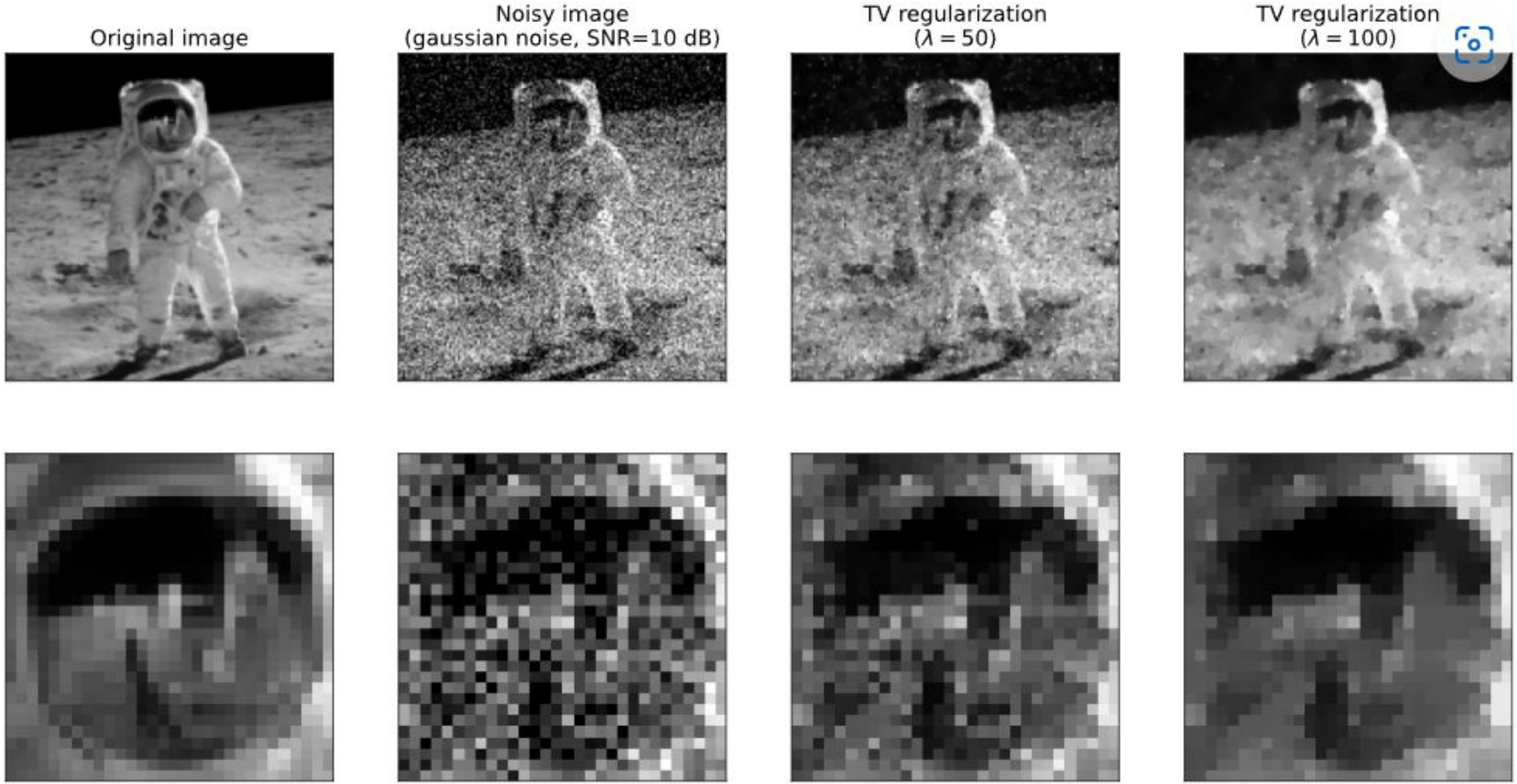


Fig. 80 Denoising with TV regularization, for two values of λ. #

Deconvolution

Usually, images acquired by a vision system suffer from degradation that can be modelled as a convolution. For example, some images present a camera shake effect (Fig. 81) or a blur due to poor focus (Fig. 82). The goal of deconvolution is to cancel the effect of a convolution.

Fig. 81 An example of motion blur (the parliament of Budapest shot by a camera).

Fig. 82 Hubble's view of Ganymede in 1996.

The degradation phenomenon is modelled as in Fig. 83: The observed image y is degraded by the convolution with a PSF h and, possibly, by a noise b (considered to be additive).

$$y = h * x + b$$

The deconvolution computes a deconvolved image $\widehat{x}$ from the observation y. We will consider only linear methods, thus deconvolution comes to filtering by g:

$$\widehat{x} = g * y$$

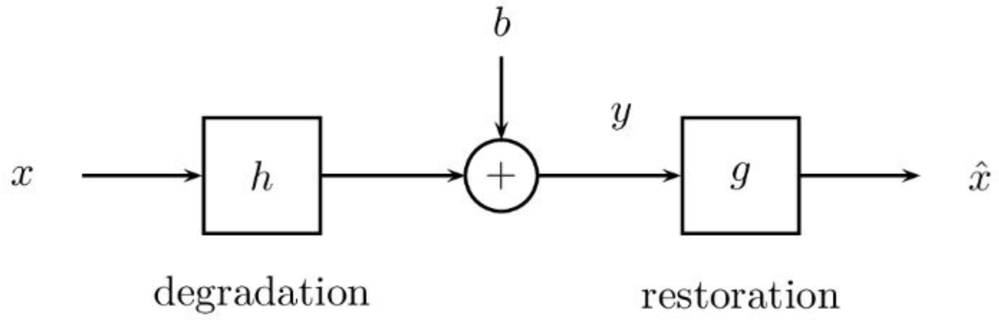


Fig. 83 Deconvolution model.

Deconvolution needs a degradation model, thus having knowledge about both h and b.

- The PSF h can be estimated by observation, *i.e.* by finding in the image some factors to estimate h. For example, a single point object in the image is h. The PSF can also be estimated by experimentation by reproducing the observation conditions in a laboratory. So, the image of a pulse gives an estimate of h. Finally, it is also possible to estimate the PSF from a mathematical model of the physics of the observation. Note also that some deconvolution methods estimate the PSF h at the same time as x: these are called blind deconvolution methods (French: *déconvolution myope*).
- Models for the noise have already been presented in chapter Denoising.

Inverse filter

The inverse filter is the simplest deconvolution method. Since the degradation is modelled $y = h * x + b$, then this equation becomes in the Fourier domain:

$$Y = HX + B$$

so we can write:

$$X = \frac{Y - B}{H}.$$

We obtain x by calculating the inverse Fourier transform of the previous expression:

$$x = \mathcal{F}^{-1}\left[\frac{Y - B}{H}\right].$$

As the noise (and therefore its spectrum B) is unknown, we can approximate the expression of x by cancelling B in the previous expression, and thus get the deconvolved image:

$$\widehat{x} = \mathcal{F}^{-1}\left[\frac{Y}{H}\right]$$

The result of the inverse filter applied on an image is given Fig. 84. The result is not usable, and yet the observed image is very little blurred with very little noise!

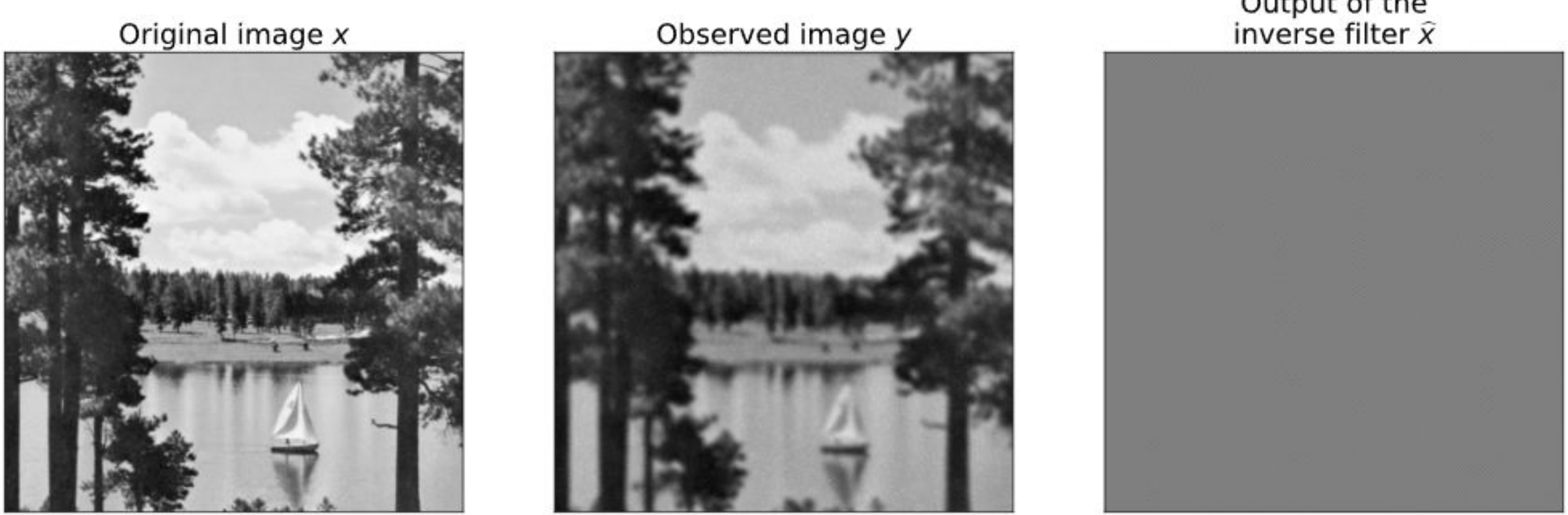


Fig. 84 Result of the inverse filter.

The catastrophic result of the inverse filter is due to the fact of having considered the noise to be zero. Indeed, according to the definition of $\widehat{x}$ and considering $Y = HX + B$, then:

$$\widehat{x} = \mathcal{F}^{-1}\left[\frac{Y}{H}\right] = \mathcal{F}^{-1}\left[X + \frac{B}{H}\right] = x + \mathcal{F}^{-1}\left[\frac{B}{H}\right]$$

Thus, the deconvolved image $\widehat{x}$ corresponds to x with an additional term which is the inverse Fourier transform of B/H. The PSF H is generally a low-pass filter, so the values of $H(m, n)$ tend towards 0 for high frequencies (m, n). Because H is in the denominator, this tends to drastically amplify the high frequencies of the noise, and then the term B/H quickly dominates X. This explains the result of Fig. 84.

One solution consists in considering only the low frequencies of Y/H. This is equivalent to truncating the result given by the inverse filter by cancelling the high frequencies before calculating the inverse Fourier transform. The result of the deconvolution is much more acceptable, as shown by Fig. 85, although the result is still far from perfect (there are many variations in intensity around objects, such as tree trunks)…

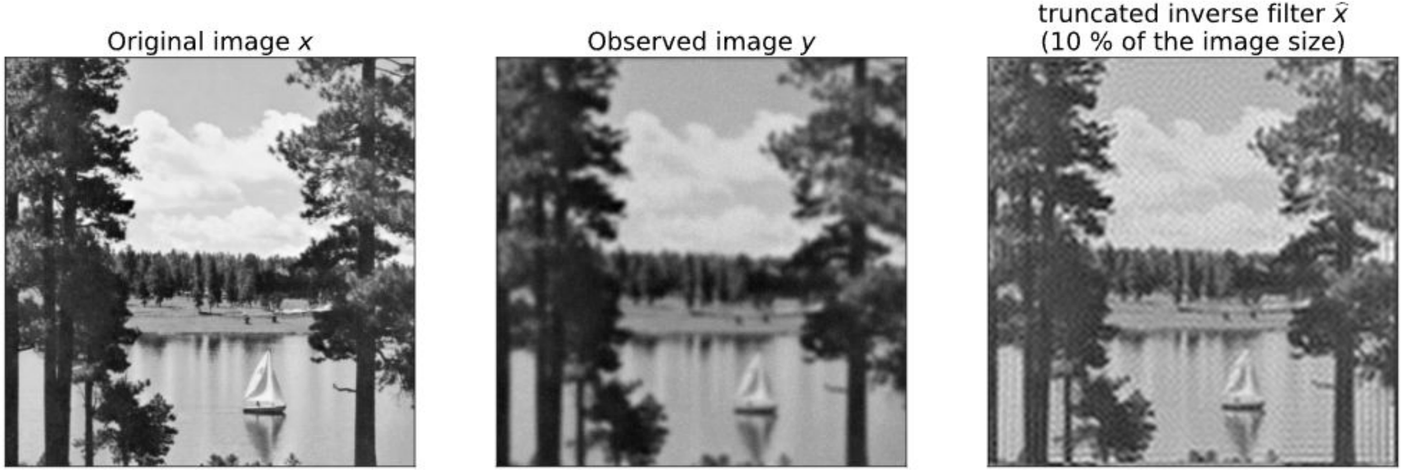


Fig. 85 Result of the truncated inverse filter with very small noise. #

Wiener Filter

Wiener filter, denoted by g (with Fourier transform G), applies to the observation y such that:

$$\widehat{x} = g * y \qquad \Leftrightarrow \qquad \widehat{X} = GY.$$

This filter is established in the statistical framework: the image x and the noise b are considered to be random variables. They are also assumed to be statistically independent. As a result, the observation y and the estimate $\widehat{x}$ are also random variables.

The calculations are done in the Fourier domain for simplicity (since convolutions become multiplications). The goal of Wiener filter is to find the image $\widehat{X} = \mathcal{F}[\widehat{x}]$ closest to $X = \mathcal{F}[x]$, in the sense of the mean squared error $\mathrm{MSE} = \mathbb{E}\left[(\widehat{X} - X)^2\right]$. Thereby :

$$
\begin{aligned}
\mathrm{MSE} &= \mathbb{E}\left[(\widehat{X} - X)^2\right] \\
&= \mathbb{E}\left[(GY - X)^2\right] \\
&= \mathbb{E}\left[\left(G(HX + B) - X\right)^2\right] \\
&= \mathbb{E}\left[\left((GH - I)X + GB\right)^2\right]
\end{aligned}
$$

where I is an image filled with 1. So:

$$\mathrm{MSE} = \mathbb{E}\Big[(GH - I)^*(GH - I)X^*X + (GH - I)^*GX^*B + G^*(GH - I)B^*X + G^*$$

where $\cdot^*$ denotes the conjugate of the variables. Since the expectation $\mathbb{E}$ is linear and only X and B are random variables, we can decompose the previous expression into four terms:

$$\begin{aligned}\mathrm{MSE} =& (GH - I)^*(GH - I)\,\mathbb{E}\big[X^*X\big] \\ &+ (GH - I)^*G\,\mathbb{E}\big[X^*B\big] \\ &+ G^*(GH - I)\,\mathbb{E}\big[B^*X\big] \\ &+ G^*G\,\mathbb{E}\big[B^*B\big].\end{aligned}$$

Since X and B are independent, then the covariances $\mathbb{E}\left[X^*B\right]$ and $\mathbb{E}\left[B^*X\right]$ are zeros. Moreover, $\mathbb{E}\left[X^*X\right]$ and $\mathbb{E}\left[B^*B\right]$ are the power spectral densities denoted as S_x and S_b (the power spectral density is the expectation of the square of the modulus of the Fourier transform). So the mean squared error simplifies into:

$$\mathrm{MSE} = (GH-1)^*(GH-1)S_x + G^*GS_b$$

We look for the filter G that minimizes the MSE, or equivalently, that cancels the derivative of MSE:

$$
\begin{aligned}
& \frac{\partial \mathrm{MSE}}{\partial G} = (GH-1)^* H S_x + G^* S_b = 0 \\
\Leftrightarrow \quad & G^* H^* H S_x - H S_x + G^* S_b = 0 \\
\Leftrightarrow \quad & G^* (H^* H S_x + S_b) = H S_x \\
\Leftrightarrow \quad & G^* = \frac{H S_x}{H^* H S_x + S_b} \\
\Leftrightarrow \quad & G = \frac{H^* S_x}{H^* H S_x + S_b} \\
\Leftrightarrow \quad & G = \frac{H^* S_x}{|H|^2 S_x + S_b}
\end{aligned}
$$

Here we are, we get the expression of the Wiener filter G! 🥳 Finally, the deconvolved image is the inverse Fourier transform of GY:

$$\widehat{x} = \mathcal{F}^{-1}\left[\frac{H^* S_x}{|H|^2 S_x + S_b} Y\right]$$

We can consider that the power spectral densities S_x and S_b are constant (for S_b, it is necessary to assume white noise). Thus, the expression of the Wiener filter can be written

$$\widehat{x} = \mathcal{F}^{-1}\left[\frac{H^*}{|H|^2 + S_b/S_x} Y\right]$$

and the term S_b/S_x is replaced by a constant K, which becomes the parameter of the method, to be set by the user.

Two remarks:

- where H vanishes (typically in high frequencies), the problem of noise increase is no longer observed as with the inverse filter, since the inverse filter tends towards 0,
- moreover, if the noise in the image is zero, then $S_b = 0$ and Wiener filter comes back to the inverse filter:

$$\widehat{x} = \mathcal{F}^{-1}\left[\frac{H^*}{|H|^2}Y\right] = \mathcal{F}^{-1}\left[\frac{Y}{H}\right]$$

The result of Wiener filter is presented Fig. 86: it is clearly much better than the inverse filter, even its truncated version!

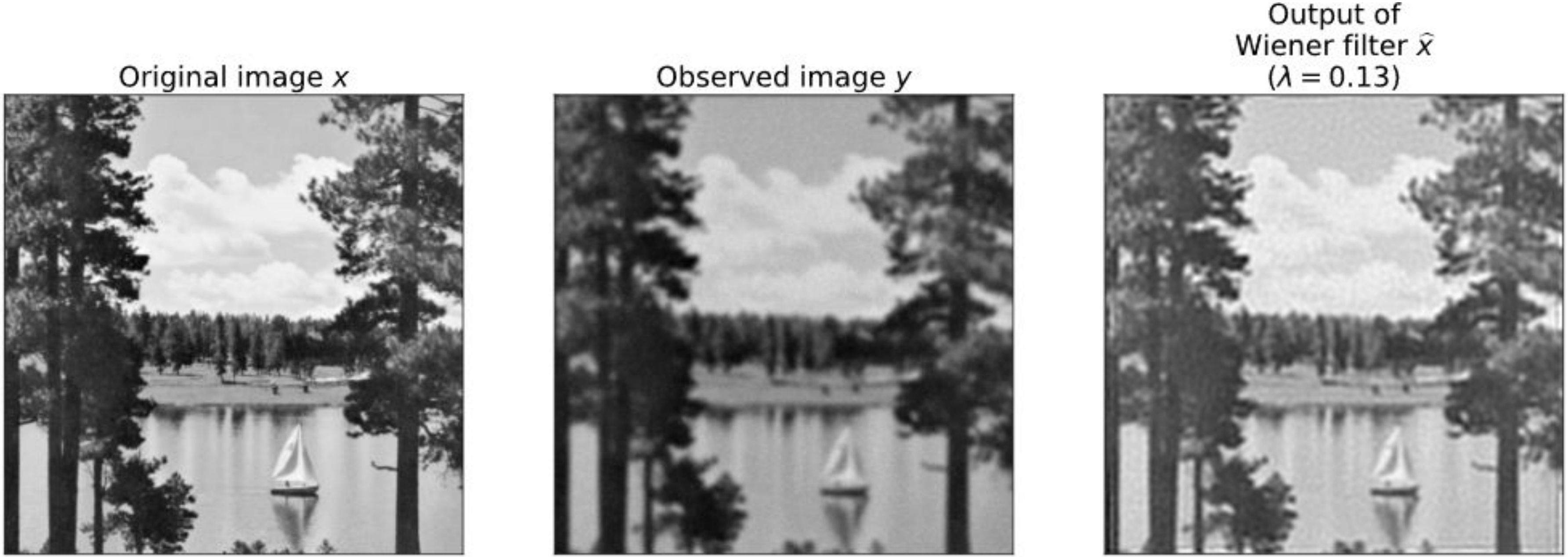


Fig. 86 Result of Wiener filter (λ is set so that the estimation is the best in terms of MSE).

References

- A. Buades, B. Coll, J.-M. Morel, “A non-local algorithm for image denoising”, CVPR, 2005.
- A. Chambolle, “An Algorithm for Total Variation Minimization and Applications”, *Journal of Mathematical Imaging and Vision*, vol. 20, p. 89–97, 2004.
- L.B. Lucy, “An iterative technique for the rectification of observed distributions”, *Astronomical Journal*, vol. 79, no 6, p. 745–754, 1974.
- W.H. Richardson, “Bayesian-based iterative method of image restoration”, *Journal of the Optical Society of America*, vol. 62, no 1, p. 55–59, 1972.
- L.I. Rudin, S. Osher, E. Fatemi, “Nonlinear total variation based noise removal algorithms”, *Physica D*, vol. 60, no. 1–4, p. 259–268, 1992.