M2 AIC, TC5 project :Image Super-Resolution Reconstruction

Wafa Bouzouita, Dhia Eddine Youssfi

Supervized by :
M. Matthieu Kowalski

Introduction

Imaging plays a key role in many diverse areas, such as astronomy, remote sensing, microscopy or tomography, just to name few. Due to imperfections of measuring devices (optical degradations, limited size of sensors, camera shake) and instability of observed scene (object motion, air turbulence), captured images are blurred, noisy and of insufficient spatial or temporal resolution. Image restoration methods try to improve their quality.
Our aim in this project is to apply different methods of image processing, in order to reconstruct a higher image resolution than the original resolution of the used sensor. Solving this type of problems amounts to resolve a linear inverse problem. During this paper, we will follow the following plan :

1. Introducing the direct model : how to obtain a low resolution image from a super resolution image.

2. Estimating the original super resolution image by solving an inverse problem, using two techniques : Inpainting and deconvolution.

3. Generating high resolution images, with the assumption that we don't know the translations, the scaling factor..

In this project we create two notebooks, the first containing the 2 first parts and the second containing the last part. To test our code, we use the Barbara image. During this project, we used two references [1] and [2].

1 Super Resolution : direct model

In order to produce a high resolution image from several low resolution images, firstly it is necessary to define a model which relates these two images. This type of model is called the direct model of the super-resolution inverse problem and it describes how we can obtain a low resolution images from a high resolution image. This model can be summarized on fig. 1.

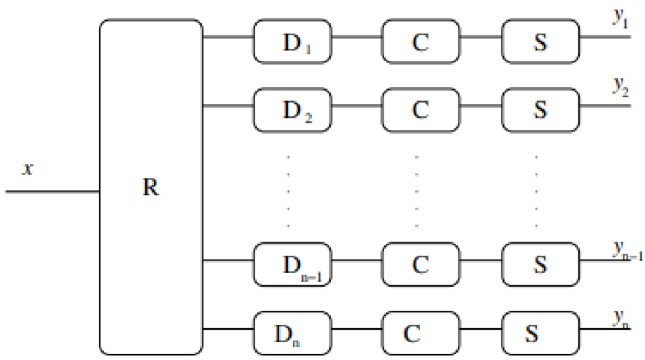

Figure 1: Direct model of a super-resolution problem.

Where

- $x \in R^N$ is the original high resolution image.
- $y_i, \quad \forall i = 1, ..., n$ is a low resolution image.
- **R** is a replication operator to create n images from one original image.
- $\mathbf{D_i}, \quad \forall i = 1, ..., n$ is a translation operator.
- **C** is the impulse response of the sensor.
- **S** is a subsampling operator to produce a low resolution image from a high resolution image.

At end, we have for all low-resolution images i:

$$y_i = \mathbf{SCD_i}x$$

1.1 Generating low resolution images y_i

In this context, from the initial image of Barbara (512×512) pixels, we have generated 4 different images, by translating every one in different directions (right,left,up,down), downsampling by a scaling factor equal to 4 and adding noise (gaussian noise). Then, we obtained the following image :

Figure 2: Low resolution image, that we obtained by applying the direct model.

evidently, we can see that the image is of bad quality, it is blured with a PSF C which is a 5 5 matrix of ones (normalized), it is noised with gaussian noise, translated to the right (4 pixels) and downsampled with a scaling factor equal to 4.

1.2 Images LR combination : construct the binary Mask

Our goal is to estimate the original high-resolution image x form its low resolution versions y_i, that is several translated then subsampled version of x. For this, it is convenient to describe the direct model by the following equation :

$$\mathbf{S^T} y_i = \mathbf{S^T SCD_i} x + \mathbf{S^T} n_i$$

Or the convolution operator C and the translation operators Di are commutative. Then we can rewrite the previous equation as :

$$\mathbf{S^T} y_i = \mathbf{S^T SD_i C} x + \mathbf{S^T} n_i$$

$$z_i = \mathbf{M_i C} x + \mathbf{M_i} e_i$$

Where $\mathbf{M_i} = \mathbf{S^T SCD_i}$ is a binary mask of the size of the original high-resolution image and $e_i = \mathbf{S^T} n_i$ is the additive error due to noise. Now, in order to take into account all the data observed in all the z_i, one can create one binary mask of the size of the original high-resolution image using the following procedure:

$$z = \frac{\sum_i (\mathbf{M_i C} x + e_i)}{\sum_i M_i} = \mathbf{M}(c + e)$$

Where the division is element-wise, and only on non-zero values. M is then still a binary mask. and e is a noise but not necessarily white anymore. Then, the direct model of our problem involve a deconvolution and an inpainting direct problems.
In our project, we deal with matrix,

$$M,\ x,\ c,\ ,e.. \in R^{h,w},$$

Where h and w are respectively the number of row and column of original barbara image. To inverse the direct problem $z = M(c + e)$, we are interested firstly in finding the binary mask $\mathbf{M}$. In fact, this mask can be constructed by considering that the pixels which can be determined by the low resolution images are equal to 1, otherwise they are equal to zero.

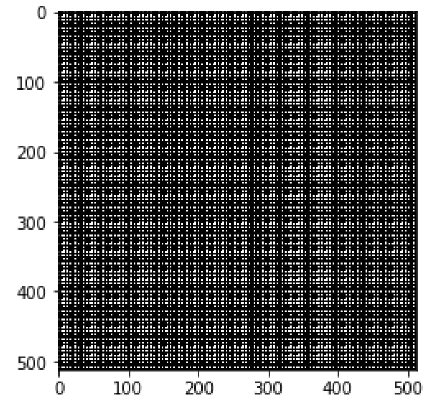

Figure 3: Mask M.

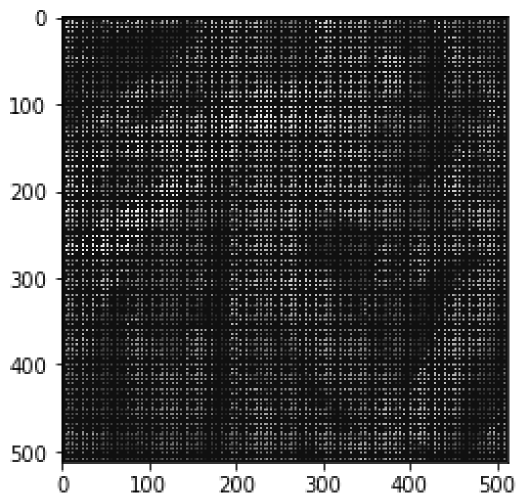

Figure 4: Combined low resolution images.

We can see from the Figure 4., that the image that we noted by z (combination of low resolution images), contains many holes. In fact, when we computed the percentage of pixels derived from the low resolution images, we found only 18.701%. This is very challenging!

2 Super Resolution : inverse problem

2.1 Inpainting

At present, we have identified the Mask M and reconstructed the barabara image with the original size. Then, the necessary tools to inverse the direct problem are ready. In order to fill automatically the missing parts in the reconstructed image (Figure 4.), we need to use a method called Inpainting. This method consists in finding an estimation of the original signal c (convolved signal) by minimizing a regularized loss function, that could be written as following equation :

$$L(c) = \frac{1}{2}||z - Mc||_2^2 + \lambda R(c)$$

Where :

- L is a data fidelity term, this term links the observation z and the original signal c through the system M. It aims to model the additive noise n.
- R is a regularizer, or prior. This prior aims to model the knowledge of the original signal c.
- λ is a hyperparameter that controls the tradeoff between L and R.

Then, the estimation of c is given by

$$\hat{c} = argmin_c L(c)$$

To inpaint the damaged image c, we have used as a regularizer term, the variational energies : Sobolev and total variation. Thus, we implemented the proximal gradient descent algorithm adapted to this two regularizer term, to minimize the regularized Loss function and then estimate the signal c.

2.1.1 Sobolev Inpainting

This method is interested in minimizing the sobolev norm of the image c :

$$R(c) = ||\nabla c||^2$$

Choice of hyperparameter λ: To choose the adequate hyperparameter λ, we tested the SNR between the original image and the inpainted image for different values of λ. We find that $\lambda = 0.08$ give the best result.

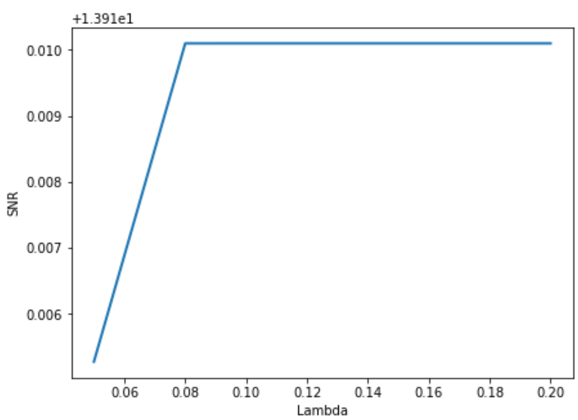

Figure 5: SNR in terms of the values of lambda.

Algorithm output : we fix *lambda* to 0.08, and we visualize the images obtained by the algorithm in terms of iterations.

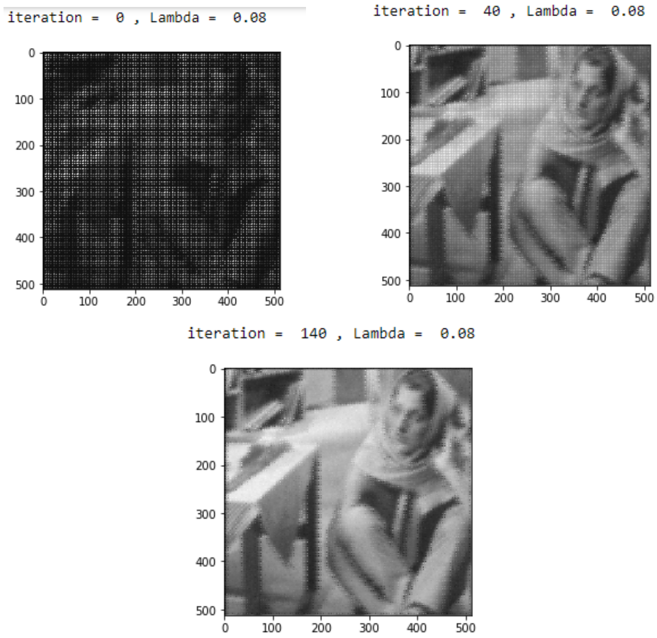

Figure 6: Results of the Sobolev inpainting.

2.1.2 TV Inpainting

In this method we replace the previous regularize term by the TV norm, that tends to better reconstruct edges.

$$R(c) = J_\epsilon(c) = \sum_i \sqrt{||\nabla c(i)||^2 + \epsilon}$$

Choice of hyperparameter λ: to have convergence of the algorithm, the step size λ should satisfy $\lambda < \frac{1}{4}$. Then, we fixed in our code, $\lambda = \frac{0.9\epsilon}{4}$ and we have interested changing only the value of ϵ (hyperparameter). We find that $\epsilon = 0.1$ give the best result, $SNR = 13.97$.

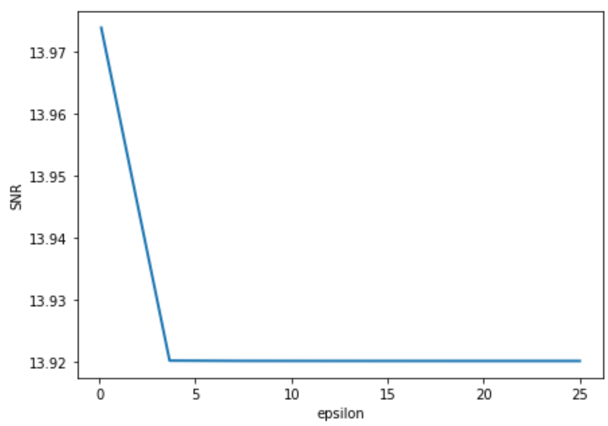

Figure 7: SNR in terms of the values of epsilon.

Algorithm output : we fix *epsilon* to 0.1, and we visualize the images obtained by the algorithm in terms of iterations.

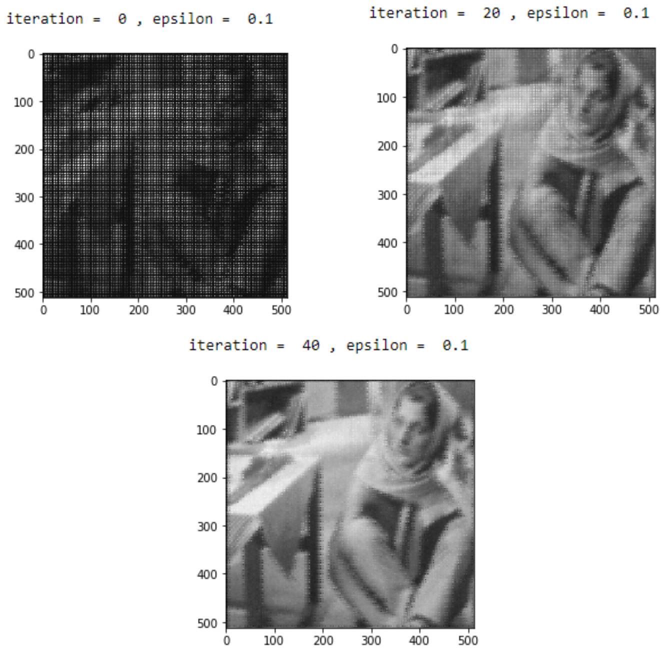

Figure 8: Results of the TV inpainting.

2.2 Deconvolution

In the previous section, we performed the inpainting, then we estimated c. Or $c = Cx$. Where x is the original signal we want to determine.
In the direct problem, we have applied to the original barbara image a box blur filter which is a 5 5 matrix of ones (normalized). It is a form of low-pass ("blurring"). Now, our goal is deconvolve the blurred image c by solving an inverse problem which consist in the inversion of the convolution matrix. To do that, we used the method of the deconvolution with L2 Regularization.
Choice of hyperparameter λ: To choose the adequate hyperparameter λ, we tested the SNR between the original image and the deconvolved image for different values of $Lambda$. we obtain the following graph :

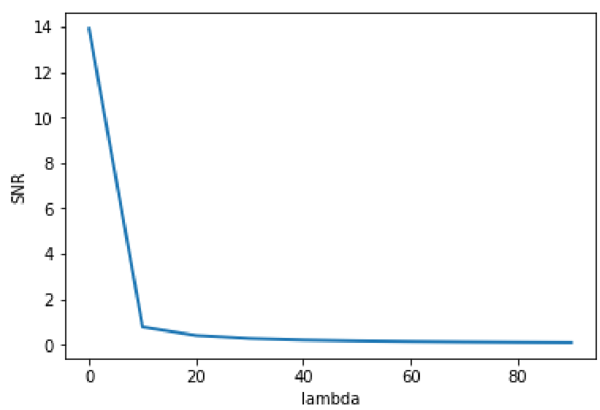

Figure 9: SNR in terms of the values of lambda.

We find that $\lambda = 0.002$ gives the best result, $SNR = 13.92$. In this case,the result obtained without the deconvolution is a little bit better !

3 From low resolution images to high resolution image

In this part we want to generate High resolution images from low resolution images. We suppose that we don't have an idea about the translations of LR images.

3.1 Translation estimation

In the case where we don't know the translations of the low resolution images, we are interested to estimate these translations. We supposed that the translations are very small: only few pixels. Then we applied a technique consist of comparing the correlation coefficient between the original image and few translated image (in different directions). Thus, we chose the ones that maximize the correlation coefficient. The estimated translations that we found by this technique wasn't 100% correct but closer to the right answer and most of the time is the same.

3.2 Generation of high resolution images

By using the translation estimation and choosing an up scaling factor, we can construct the mask M. Then, we can solve the inverse problem and determine the high resolution image. We found then for different up scaling factor the following images :

- Reconstructed image with up scaling factor $= 5$, with filled hole 12.0%.

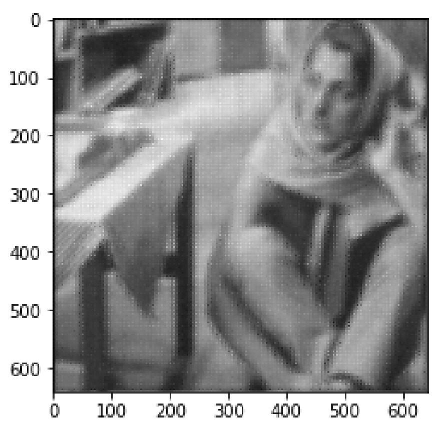

Figure 10: High resolution image.

- Reconstructed image with up scaling factor $= 4$, with filled hole 18.75% and SNR $= 14.174$.

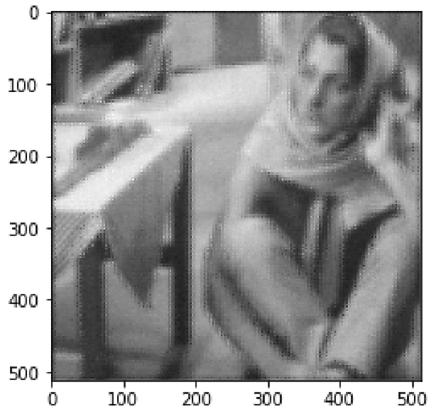

Figure 11: High resolution image.

- Reconstructed image with up scaling factor : 3, with filled hole 33.333%.

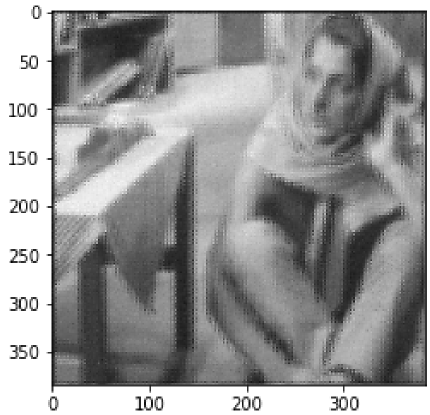

Figure 12: High resolution image.

Even if the estimated translation in not perfectly correct, we obtained a good result, SNR of the image obtained by the original upscaling factor 4 is of 14.17 db.

References

[1] `http://hebergement.u-psud.fr/mkowalski/AIC/data/M2_AIC_SuperResolution.pdf`.

[2] `https://www.numerical-tours.com/`.