Exercise 6 - Fourier Transforms

May 15, 2025

LTE, EPFL, Switzerland

Question 1

Describe the figure. What are the axes with the dominant frequencies? To which features do they correspond in the image?

- Values at the center of the plot are stronger, i.e. those with low frequency (in absolute value).
- There are 2 dominant axes, at approximately 45° and 135° (orthogonal to each other).
- The dominant axes correspond to the orientation of the street network which follows a rectangular grid layout.
- The first (up left to bottom right) corresponds to the edges between small streets and buildings.
- The second is less intense and corresponds to the edges between large boulevards and block of buildings.

Question 2

How do the translated and resized images differ from the Fourier transform of the original aerial image?

1. The same grid pattern is seen, with the same dominant axes.
2. The frequency range changes: the bounds are roughly $[-0.3, 0.3]$ m^{-1} instead of $[-0.6, 0.6]$ m^{-1}.
3. When we compare the same frequency range -0.3 to 0.3 m, we can see that the F transforms of the 2 resized images are the same. The translation does not influence the transform (shift property).

4. However, there are differences between the F transforms of the resized and original images. Since the images were rescaled with a factor 1/2, the magnitude is also rescaled but by a factor 1/4 (scaling property).

Question 3

Based on your understanding of the Fourier transform, explain the features you observed.

1. The grid pattern is the same because we have the same structure in the image.
2. The Fourier transform is sensitive to spatial frequency and not to the localization of features on the image. Hence translation does not affect the transform.
3. Intuitively we can reason in the following way: the image now takes up fewer pixels in the total image (4 times less: we divide by 2 in each direction), so for each frequency we have 4 times fewer pixels that can contribute to it.
4. The pixel size 'on the ground' is twice as large as before. We have averaged the information which used to be in 4 pixels into one pixel. This means that in each perpendicular direction, the size of the smallest detail which can be resolved on the ground has been increased by a factor of 2. High frequency patterns on the ground can no longer be seen, so the frequency range is reduced by a factor of 2.

Question 4

Explain how the rotated image differs, in the Fourier domain, from the original aerial image. Does the rotation affect the Fourier transform differently from the sole translation + resizing?

The F transform of the rotated image is the same as the resized and translated image but rotated to the chosen angle. This is because the directions of the principal changes in the image have changed.

Question 5

What is the effect of a low-pass filter on the image? What effect does the modification of CUTTING_FREQUENCY have on the final result?

1. The low-pass filter cuts the frequencies higher than a certain cutoff frequency and keep only the lower frequencies.
2. The output is a smoothed image where the sharp transitions are softened but where the dominant features are still visible.
3. Changing the value of CUTTING_FREQUENCY increases or reduces the area of the section of the Fourier transform that we keep in re-creating the image.
4. A smaller CUTTING_FREQUENCY results in a more blurred/smoothed image (we keep "less" of the transform output), while a larger one increases the level of details that we keep in the final product.

Question 6

What is the effect of a high-pass filter on the image? What effect does the modification of CUTTING_FREQUENCY have on the final result?

1. A high-pass filter is the complement of a low-pass filter.
2. It cuts the frequencies lower than a certain cutoff frequency and keeps only the higher frequencies.
3. This accentuates the small details of the image, which will look sharper. When low frequencies are removed, the edges are highlighted.
4. Similarly to the low-pass case, changing the value of CUTTING_FREQUENCY increases or reduces the area of the section of the Fourier transform that we keep in re-creating the image.
5. A larger CUTTING_FREQUENCY keeps only a few details in the image (with a large "inner circle", only a small part of the transform output, the one at higher frequencies, is kept for re-creating the image), while smaller values keep more information.

Question 7

What do you observe?

We get the original image. When we add the results of the low-pass and high-pass filters set with the same cutting frequency, we add complementary parts. The low-pass filter keeps all the frequencies lower than the cutting frequency whereas the high-pass filter keeps all the frequencies higher than the cutting frequency. By adding them, we retrieve the original image.

Note that if we had summed the absolute values of the two filtered transforms directly, the result would have been different from the original image. This is due to the nature of complex numbers. The triangle inequality states that, given two complex number a and b:

$$|a + b| \leq |a| + |b| \tag{1}$$

A simple explanation can be found here.

Question 8

Where do you see the signature of the stripes on the spectrum? Give an explanation.

The stripes are visible in the magnitude spectrum as small dots on the horizontal. The presence of these stripes causes a higher variability in the direction perpendicular to the stripes (short transition from light to darker pixels). The facts that there are several dots is due to the fact that the stripes are long rectangles and as such they can not be represented by a finite sum of sinusoids.

Question 9

How would the Fourier spectrum be different if the stripes on the Landsat image were more spaced out (larger distance between two stripes)?

If the stripes are more spaced out the dots will be less spaced out (still along the same "line"). The space between dots decreases because the frequency is lower (ie larger spatial wavelength).

Question 10

In your opinion, what are the best values for the three parameters (ANGLE_CENTER, ANGLE_WIDTH, MIN_FREQ)? Explain the reasoning behind your choice.

Our best estimates are:

- ANGLE_CENTER $= 165.0$
- ANGLE_WIDTH $= 20$
- MIN_FREQ $= 0.0008$

The center angle needs to be aligned to the main direction of the dots. The width must be large enough to cover the peak around the dots, without covering unaffected regions of the spectrum. If MIN_FREQ is set too large then the stripes will not be properly removed. Since the slice tapers off close to the center of the axis, a very small MIN_FREQ does not negatively impact the image.

Question 11

Even with the "optimal" set of values, the result is quite far from "perfect". What artefacts do you see? What areas are problematic?

Here the stripes we removed are replaced by gray pixels. In the dark region (bottom right) we see the removed stripes well. Depending on the chosen parameters (and especially the value of MIN_FREQ) we might have the following artifacts:

- Some grayish features where the stripes used to be, like here (if MIN_FREQ is still a bit large)
- Increase in blurriness (if MIN_FREQ is small)

Question 12

Briefly describe the histograms. Do these histograms confirm the dark appearance of the Landsat image to the eye? Can you identify signatures from certain features in the image?

The main part of the distribution is contained between pixel values of 25 and 100 rather than the full rage of possible pixel values. Since a pixel value of zero is a dark pixel, all three bands are quite dark.

Question 13

What are the correct values for the following variables: $DN_{0,min}$, $DN_{0,max}$, $DN_{n,max}$, $DN_{n,min}$? *Briefly explain your reasoning, and include the resulting enhanced image. What do you think of the result?*

- Correct values: DN0MIN = 0, DN0MAX = 195, DNNMAX = 255 and DNNMIN = 0. According to the function min and max, the min intensity DN0MIN is 0 and the max DN0MAX is 195. To correct the intensity, we want to extend the range to 0 to 255. (Optional: Why 255? Because pixels are encoded as bytes which are groups of 8-bits and the maximum value obtained with a group of 8-bits is 256 (11111111 = 256 in base 2) but since we include 0 the range of 256 values goes from 0 to 256-1.)
- Enhanced image: indeed luminosity is better, although still a bit dark (possibly due to the fact that the histograms are rather skewed, so even after stretching them so that the new maximum is 255 we still have most of the pixels with rather dark values).