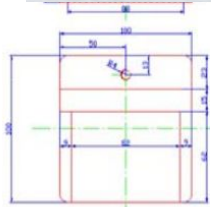
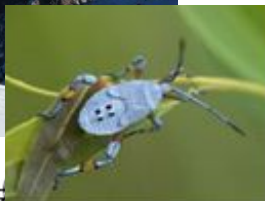


MLRF Lecture XX

J. Chazalon, LRE/EPITA, 2023

Image descriptors: Overview

Different sizes and contents \Rightarrow Different kind of descriptors

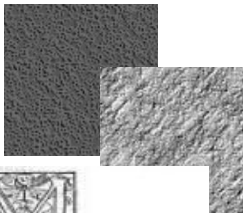


Large images, many parts

$\approx 500 \times 500$ px and more



Complex
small images



La parole se feroit la fondue dedans
enoy, de le poule cette sans mousser. Si se
ra'auis, rendus les bonnettes fainebrres, de
les -detaiens -offardes à la mousser
de ce grand Cheneier François, en fou-

Textured
areas

Office for

Words



Logos



Local image patches,
lines, etc.



Isolated
symbols and
letters

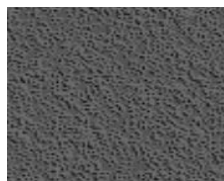
$\approx 20 \times 20$ px and less

Texture descriptors

Lecture XX part YY

What are they useful for?

To describe (then match, group, classify...) relatively large and regular images.



Vector 1



Vector 2



La parole se feroit la foudre dedans
roy, & le poultre fait sans mesure. Si se
reusent rendus les hommes en foudres, &
les dernières offrandes à la renommée
de ce grand Chevalier François, de foud-

Vector 3

A tentative taxonomy (of a selection of approaches)

Statistical

- GLCM (Grey Level Co-occurrence Matrix)
- Fractal dimension

Frequency-based

- Fourier transforms
- Difference-of-Gaussian filter
- Gabor filters
- Wavelets

Model-based

- Markov Random Fields
- Convolutional Neural Networks

Statistical approaches

GLCM (Grey Level Co-occurrence Matrix)

From an image patch of grayscale image (usually 16 levels), compute the matrix

$$C_{\Delta x, \Delta y}(i, j) = \sum_{x=1}^n \sum_{y=1}^m \begin{cases} 1, & \text{if } I(x, y) = i \text{ and } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{otherwise} \end{cases}$$

Where, for each cell (i,j) we add 1 when the pixel $I(x,y)$ has value i and the pixel $I(x+\Delta x, y+\Delta y)$ has value j .

Example, for $\Delta x = \Delta y = 1$:

Image

0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
0	0	1	0	1

GLCM

$$C = \frac{1}{16} \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

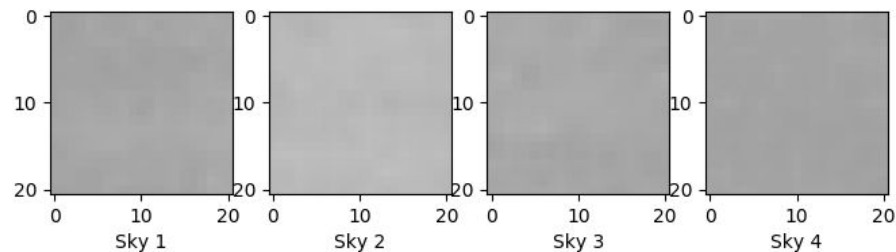
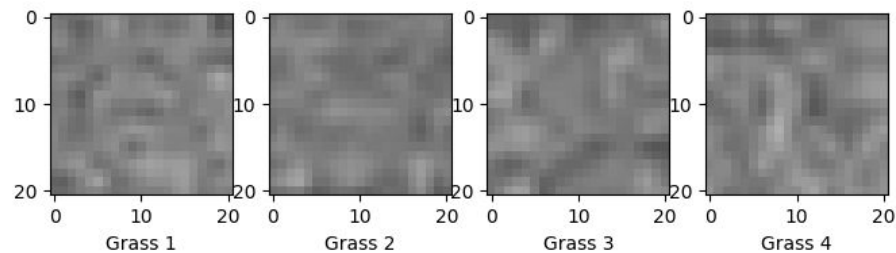
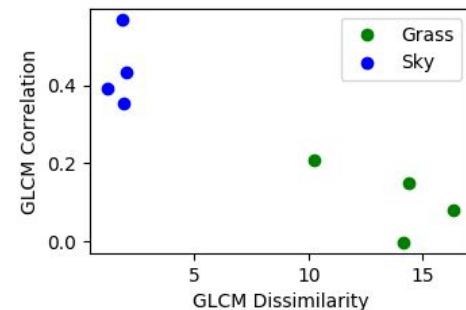
Using GLCM

We usually compute statistics on such matrix like:

- Contrast $\sum_{i,j=0}^{levels-1} P_{i,j} (i - j)^2$
- Dissimilarity $\sum_{i,j=0}^{levels-1} P_{i,j} |i - j|$
- Homogeneity $\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1 + (i - j)^2}$
- ASM $\sum_{i,j=0}^{levels-1} P_{i,j}^2$
- Energy \sqrt{ASM}
- Correlation $\sum_{i,j=0}^{levels-1} P_{i,j} \left[\frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$



Original Image



Fractal dimension

“How a smaller version of myself is equal to myself?”

Method of Range:

- Take 2 windows of size $l_1=9$ and $l_2=5$, centered on the same pixel.
- Compute the brightness range r_1 for window 1 and r_2 for window 2.
- Estimate the fractal dimension as
$$D = \frac{r_1 - r_2}{\ln l_1 - \ln l_2}$$

D: ratio between the difference of the ranges in each window and the proportion of the length of each window, in log scale.

Input image

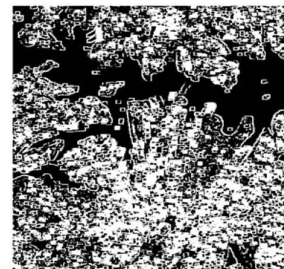
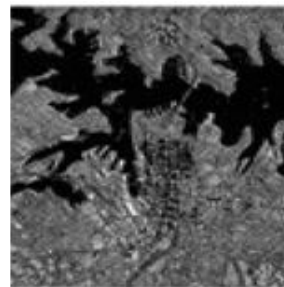
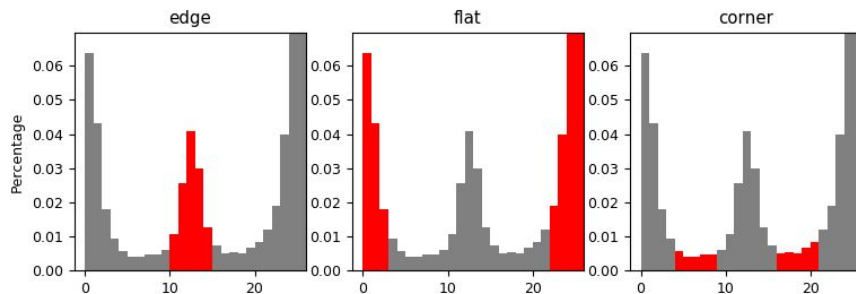
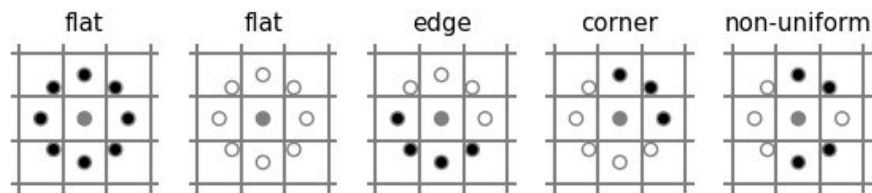


Image showing places where D is above a global threshold.

Local Binary Patterns

LBP looks at points surrounding a central point and tests whether the surrounding points are greater than or less than the central point (i.e. gives a binary result).



Frequency-based approaches

Fourier transform

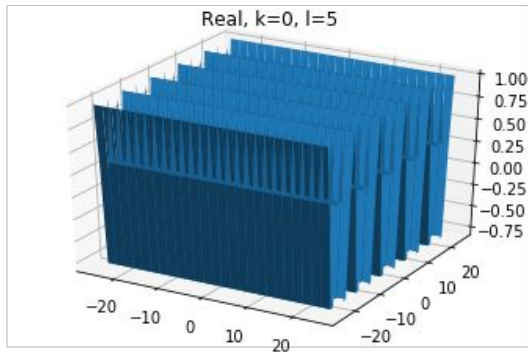
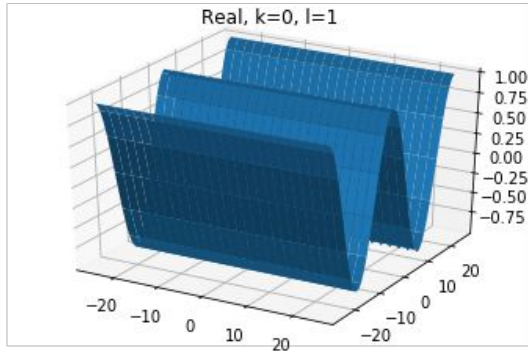
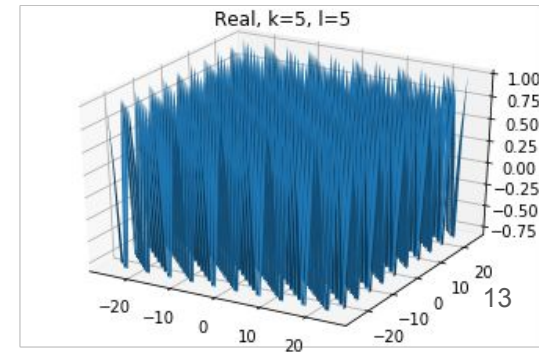
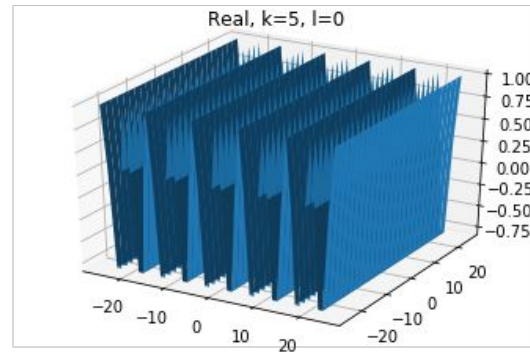
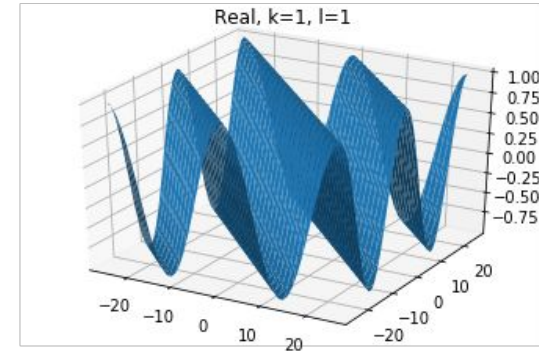
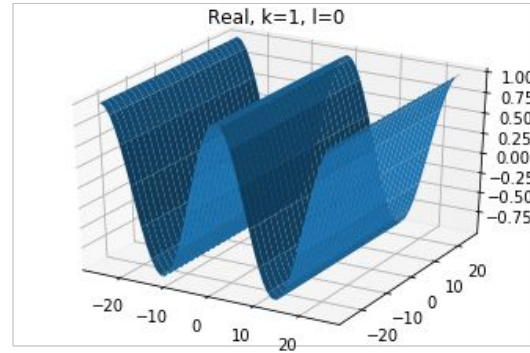
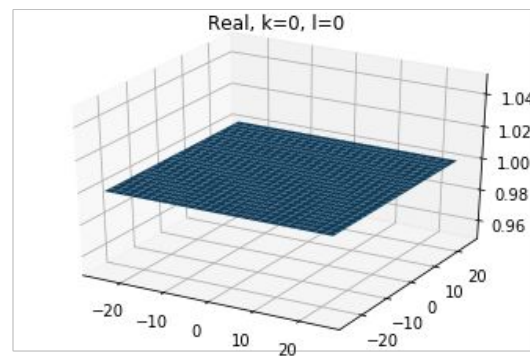
For each possible frequency, sum pixel contributions from original image.

$$F(k, l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-i2\pi(\frac{ki}{N} + \frac{lj}{N})}$$

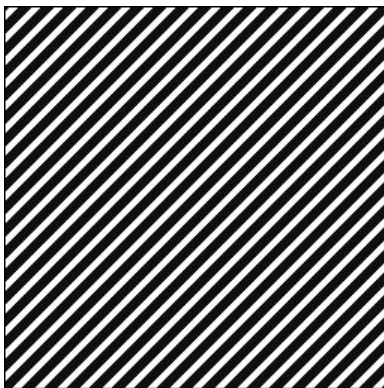
The exponent term can be viewed as the filter for the target frequency over the spatial image.

Next: values of the exponent term (only) for various k, l .

*Values of the exponent
term (only) for various k, l .*



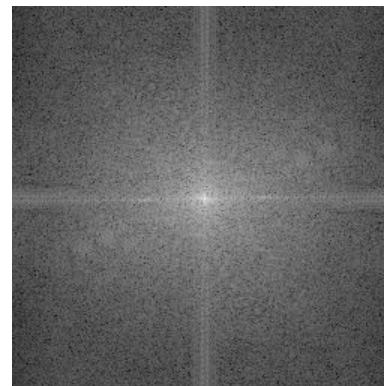
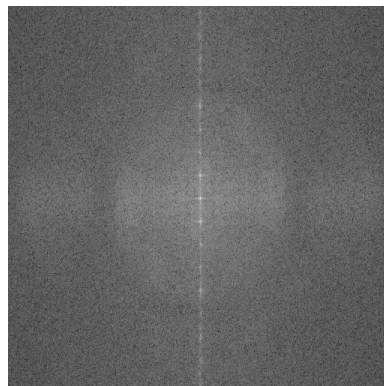
Fourier transform



Sonnet for Lena

O dear Lena, your beauty is no vast
It is hard sometimes to describe it fast.
I thought the entire world I would impress
If only your portrait I could compress.
Alas! First when I tried to use VQ
I found that your cheeks belong to only you.
Your silky hair contains a thousand lines
Hard to match with sums of discrete cosines.
And for your lips, sensual and tactual
Thirteen Crays found not the proper fractal.
And while those setbacks are all quite severe
I might have fixed them with hacks here or there
But when filters took sparkle from your eyes
I said, 'Damn all this. I'll just digitize.'

Thomas CohnAurel



Difference-of-Gaussian filter

Take a image.



Blur it.



Take the difference.

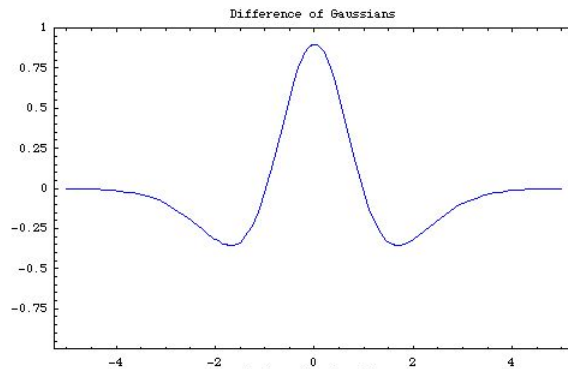


Difference-of-Gaussian filter

It is a band-pass filter.

$$\Gamma_{\sigma, K\sigma}(x, y) = I * \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} - I * \frac{1}{2\pi K^2\sigma^2} e^{-(x^2+y^2)/(2K^2\sigma^2)}$$

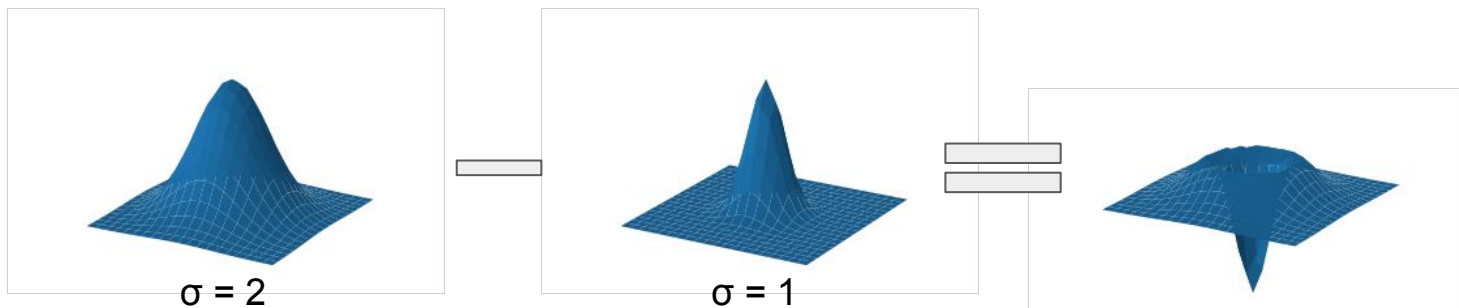
$$\Gamma_{\sigma, K\sigma}(x, y) = I * \left(\frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} - \frac{1}{2\pi K^2\sigma^2} e^{-(x^2+y^2)/(2K^2\sigma^2)} \right)$$



Difference-of-Gaussian filter

Intuition

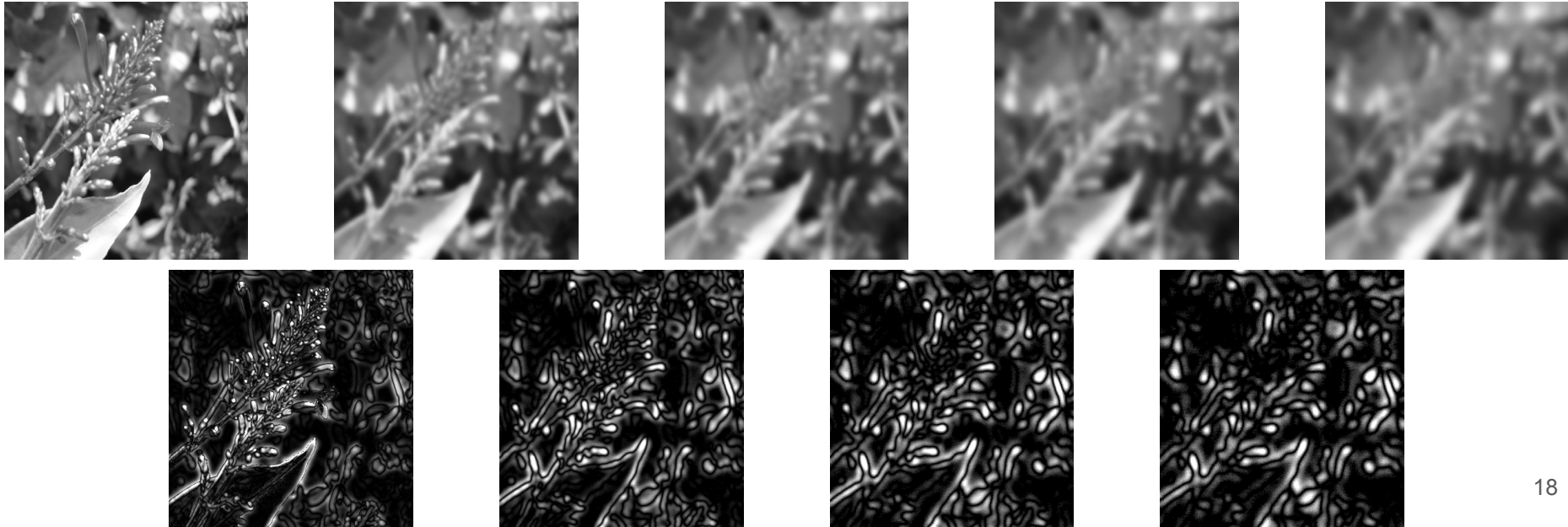
- Gaussian (g) is a low pass filter
- $(g * I)$ low frequency components
- $I - (g * I)$ high frequency components
- $g(\sigma_1) * I - g(\sigma_2) * I \Leftarrow$ Components in between these frequencies
- $g(\sigma_1) * I - g(\sigma_2) * I = [g(\sigma_1) - g(\sigma_2)] * I$



Difference-of-Gaussian filter

Many applications.

Indicates the “size” of the “stable” region around a pixel at a given freq. band.



Gabor filters

Gabor filters allow to select both a frequency band and an orientation.

$$G_c[i, j] = B e^{-\frac{(i^2 + j^2)}{2\sigma^2}} \cos(2\pi f(i \cos \theta + j \sin \theta))$$

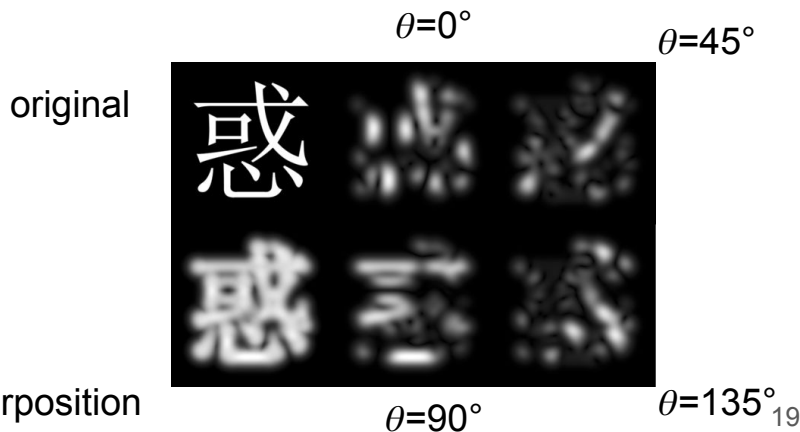
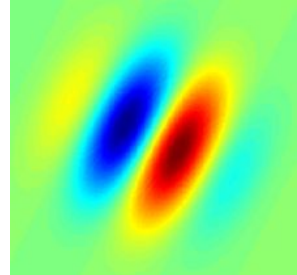
$$G_s[i, j] = C e^{-\frac{(i^2 + j^2)}{2\sigma^2}} \sin(2\pi f(i \cos \theta + j \sin \theta))$$

B and C: scaling factors

f : frequency selection

θ : angle selection

σ : size of the image region being analysed

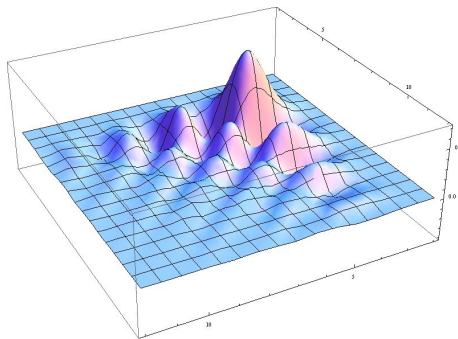


Wavelets

Much like Gabor filters, with potentially more complex patterns.

It turns the image into a grid of coefficients based on an orthogonal basis of small finite waves, or “wavelets”.

(Used in JPEG-2000.)



Learning-based approaches

Markov Random Fields

Convolutional Neural Networks

Learn to produce a high response for some texture samples / patches.

The filter bank is not orthogonal in general, but rather overcomplete, ie the original signal can be recovered using a small subset of filters.

Highly tunable, but a good random sampling over the possible patch patterns can gives good results too.

(Not covered in this course.)