Image Processing

Valentina Hubeika, Jan Černocký

DCGM FIT BUT Brno, {ihubeika,cernocky}@fit.vutbr.cz

Image signals

- grayscale picture is a projection of space to a plane, 3D $\Rightarrow$ 2D.
- video is a projection of 4D space (3 coordinates + time) to 3D (only 2 coordinates plus time).
- according to some, color is onother dimension – for a color video we work with a reduction from 5D (color, time, 3D space) to 4D (color, time, 2D projection).

Dimension reduction – this lecture is dealing with a still picture in gray-scale.

"Analog image" is a continuous picture $z(x, y)$ of space coordinates x (horizontal) and y (vertical) for $x, y \in [-\infty, +\infty]$ – normally it is called a 2D signal.

For processing we have to take the following steps:

- sample in both dimesions. Variable x is replaced by a horisontal counter of samples l. Variable y is replaced by a vertical counter of samples k. Sample $x[k, l]$ (fundamental area of an image) is called a **pixel** (picture element).

- limit the size of the picture: only L samples horizontally (L columns), and K samples verticaly (K raws). Point density is usually represented by **dpi** (dots per inch).
- sample quantization: we dispose of a limited number of quantization levels. In black-n-white pictures samples are usually represeneted by 8 bits (256 quantization levels).

Example: Lena: original and quantized, 4 quantization levels.

Spectral analysis of a 2D signal

A 2D signal can be represented by a matrix :

$$x[k,l] = \begin{bmatrix} x[0,0] & x[0,1] & \cdots & x[0,L-1] \\ x[1,0] & x[1,1] & \cdots & x[1,L-1] \\ \vdots & & & \vdots \\ x[K-1,0] & x[K-1,1] & \cdots & x[K-1,L-1] \end{bmatrix}$$

For a 2D signal, we define a 2D Fourier transform:

$$X(f,g) = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j2\pi(fk+gl)},$$

where f and g are image frequencies. In case of an analog 2D signal, image frequencies are of a unit m^{-1} (note that Hz=s^{-1}). Images we work with are sampled thus f and g are **normalized image frequencies**. The introduced equation corresponds to DTFT (Discrete Time Fourier Transform), where f and g can be of an arbitrary value.

Practically we use **two dimensional discrete Fourier transform (2D-DFT)**:

$$X[m,n] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)}.$$

Compute only for **discrete frequences**:

$$f = m\Delta f \qquad g = n\Delta g,$$

where

$$\Delta f = \frac{1}{M} \qquad \Delta g = \frac{1}{N}.$$

M, N are integer numbers, most often we use $M = K$, $N = L$.

2D-DFT can be computed separately for each dimension:

$$X[m,n] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)} = \sum_{k=0}^{K-1} e^{-j2\pi\frac{mk}{M}} \sum_{l=0}^{L-1} x[k,l] e^{-j2\pi\frac{nl}{L}},$$

thus we can compute 2D-DFT as a sequence of two 1D-DFT – first over rows and then over colums (or vice versa).

Inverse two-dimensional discrete Fourier transform (2D-IDFT):

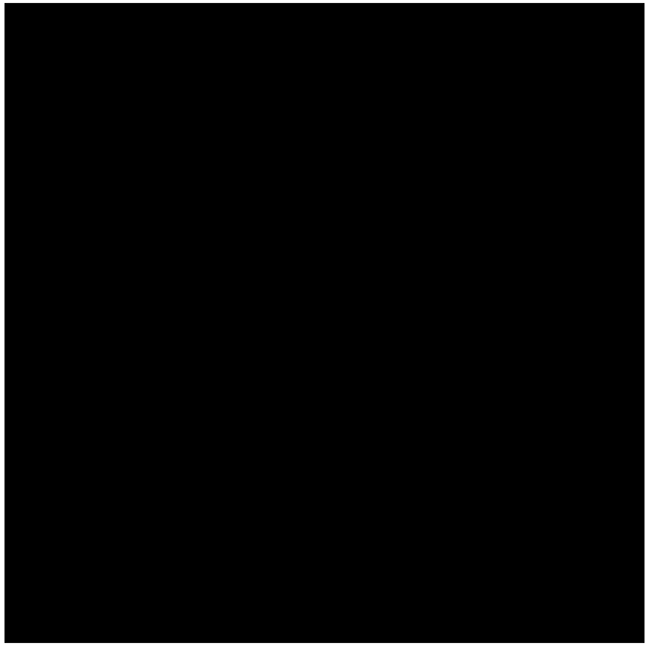
$$x[k,l] = \frac{1}{MN} \sum_{m=0}^{K-1} \sum_{n=0}^{L-1} X[m,n] e^{+j2\pi\left(\frac{mk}{M}+\frac{nl}{L}\right)}.$$

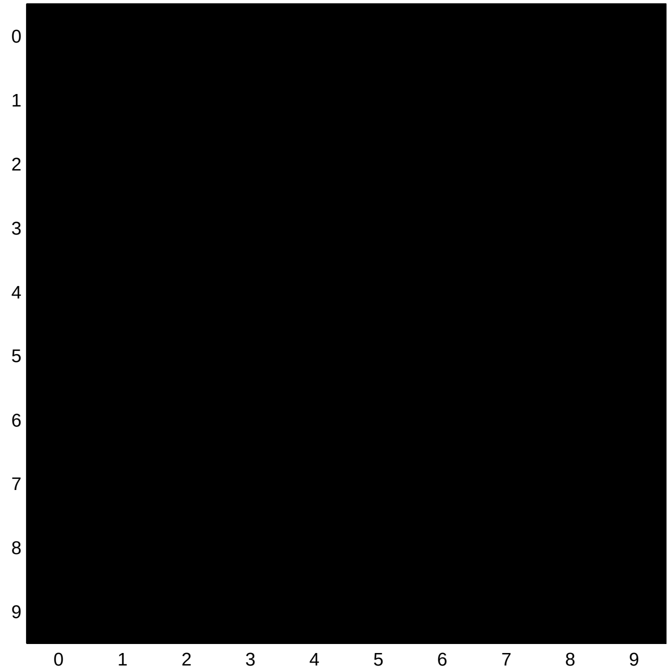
Due properties similar to DFT, matrix $X[m,n]$ is **symmetic**, thus its enough to look at the first quadrant with indeces $m = 0 \ldots \frac{M}{2} - 1$, $n = 0 \ldots \frac{N}{2} - 1$. 2D-DFT is **complex** but in the examples we will be looking only at the modules.

Image frequency – examples

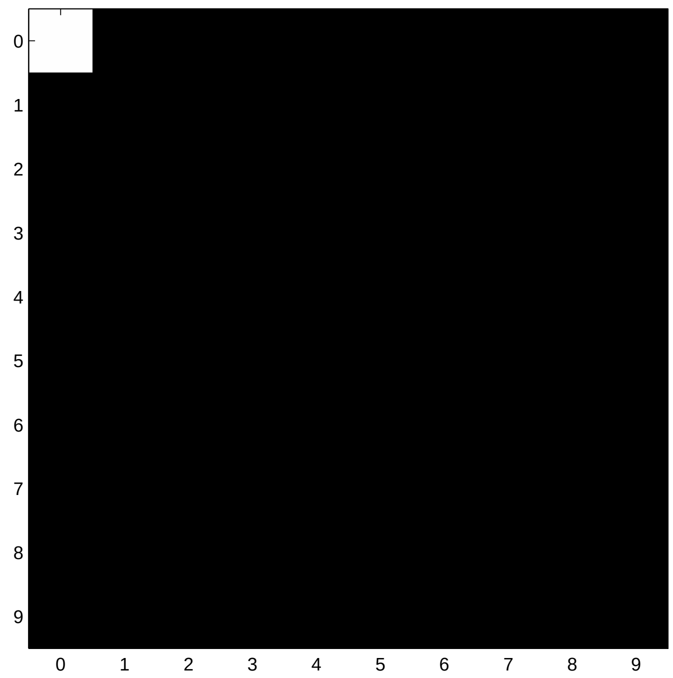
$K = L = M = N = 256.$

Black only: $x[k,l] = 0, \qquad X[m,n] = 0$

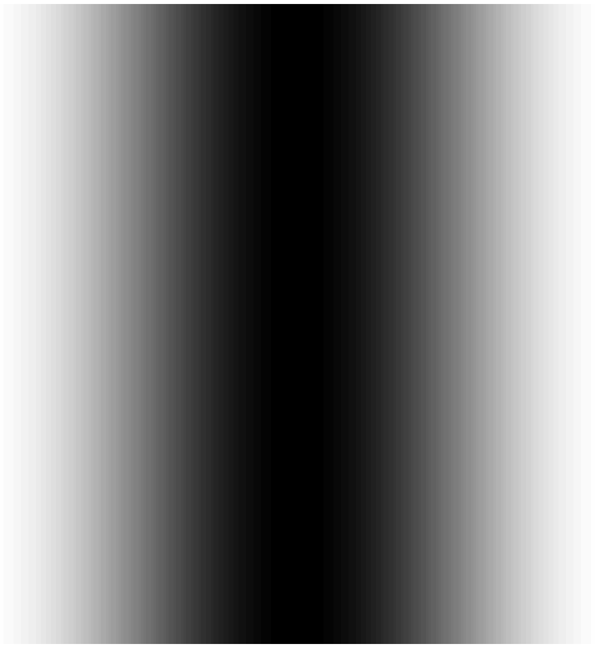


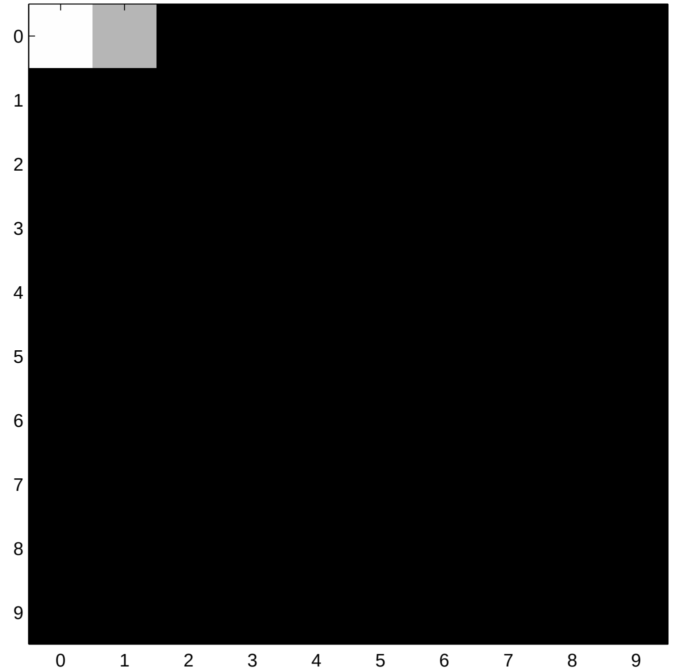


White only: $x[k,l] = 1, \qquad X[0,0] = 65536$

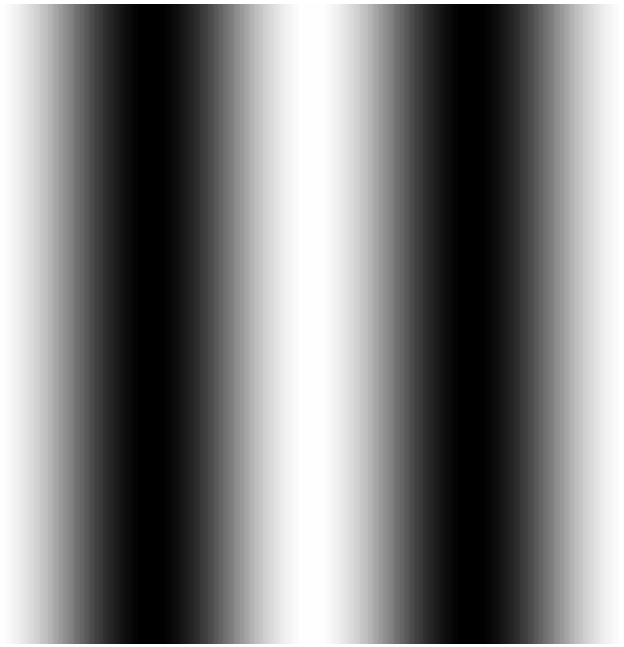


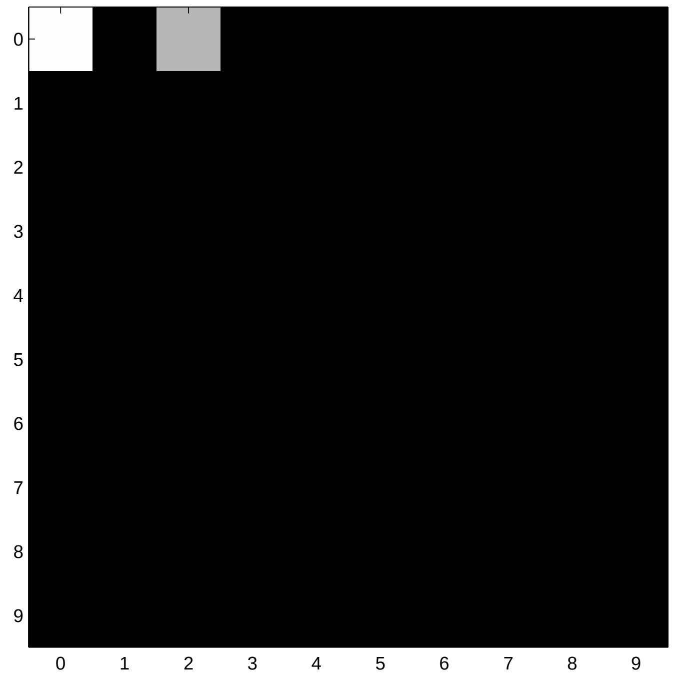
Horisontal cosine: $f = \frac{1}{L}$: $x[k,l] = \frac{1}{2} + \frac{1}{2}\cos 2\pi f l$, $X[0,0] = 32768$, $X[0,1] = 16384$,



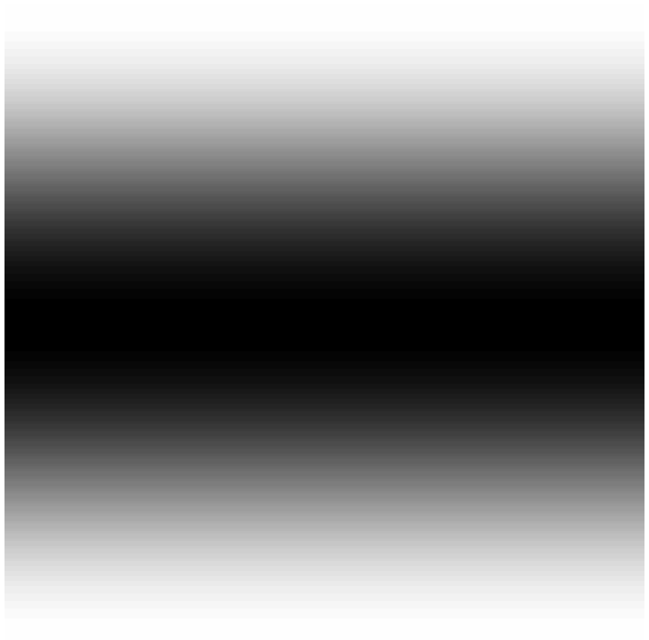


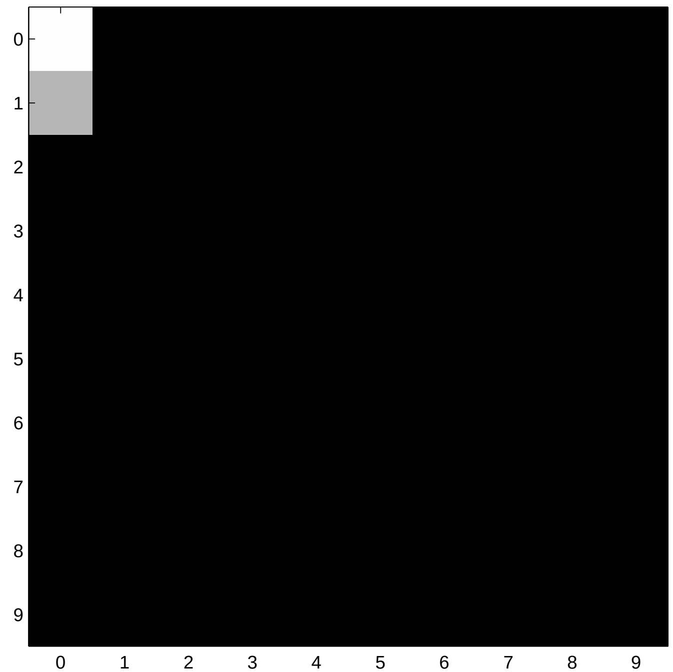
Once more: $f = \frac{2}{L}$: $x[k,l] = \frac{1}{2} + \frac{1}{2}\cos 2\pi f l$, $\quad X[0,0] = 32768, \quad X[0,2] = 16384,$



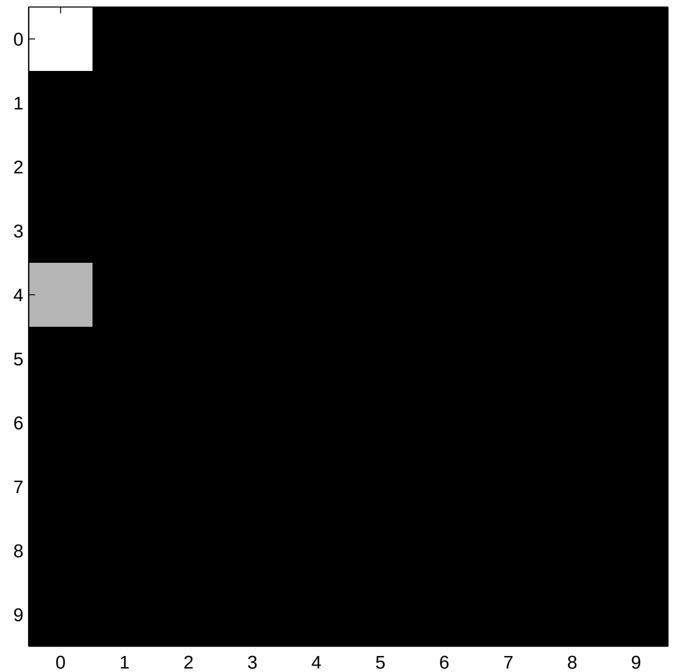


Vertical cosine: $g = \frac{1}{K}$: $x[k,l] = \frac{1}{2} + \frac{1}{2}\cos 2\pi g k$, $\quad X[0,0] = 32768, \quad X[1,0] = 16384,$

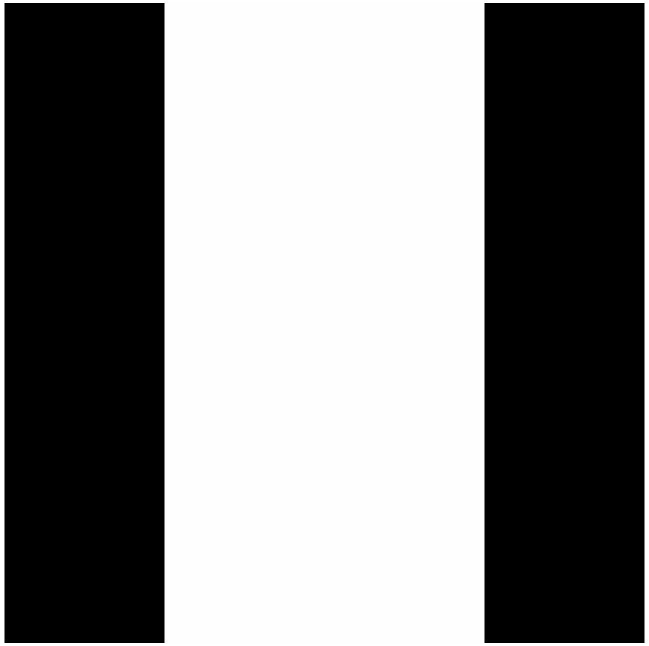


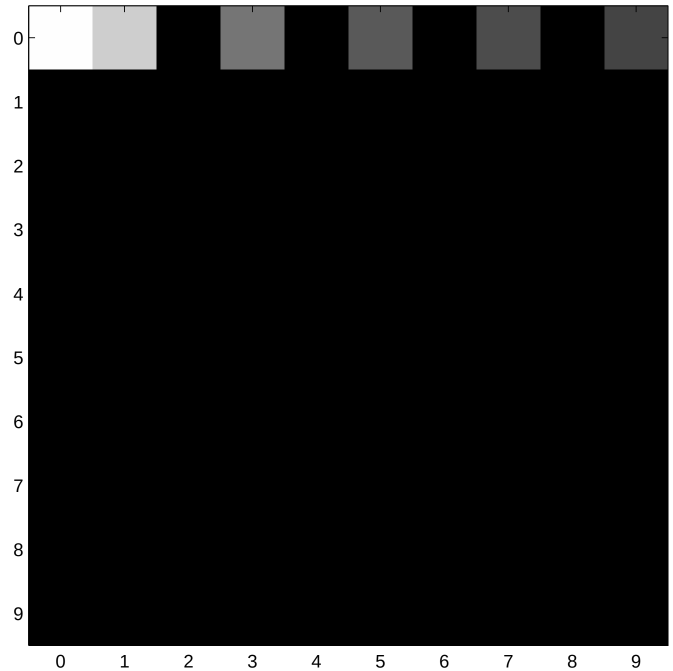


Once more: $g = \frac{4}{K}$: $x[k,l] = \frac{1}{2} + \frac{1}{2}\cos 2\pi g k$, $\quad X[0,0] = 32768, \quad X[4,0] = 16384,$

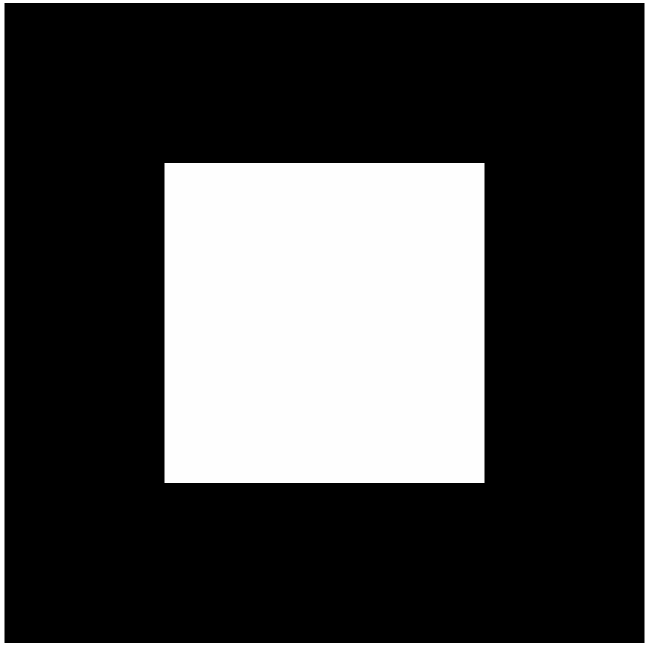


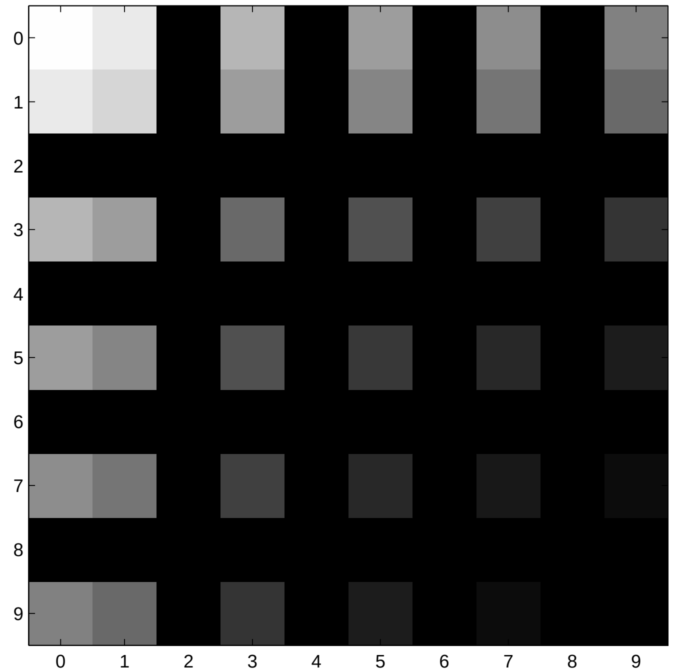
Vertical rectangle: what function defines horisontal frequencies?



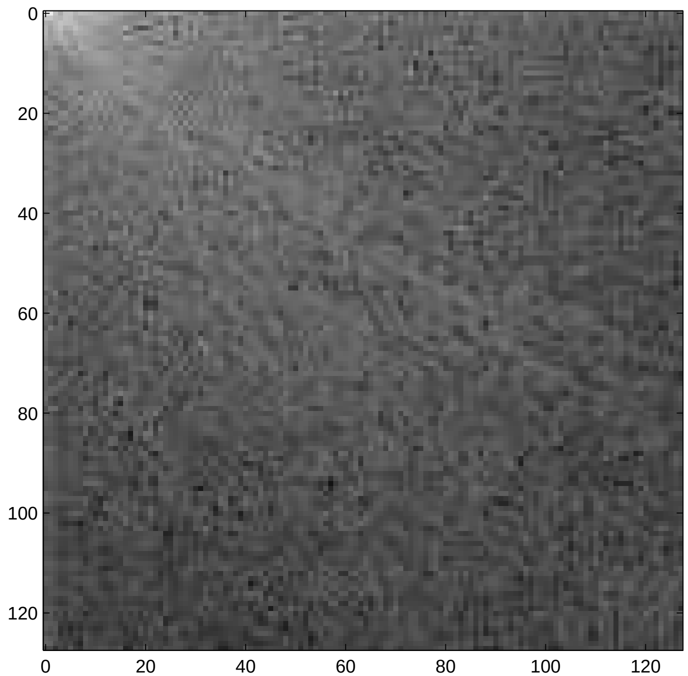


Square



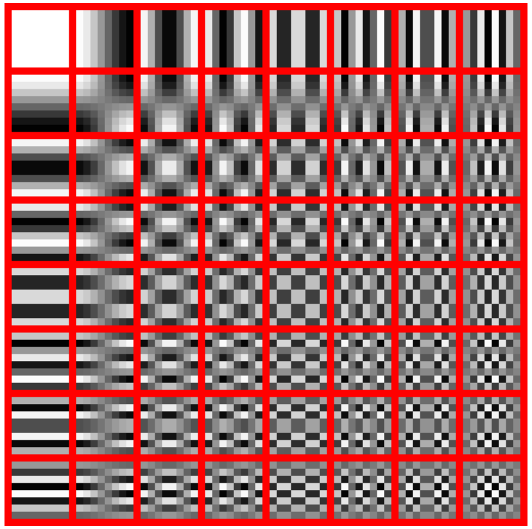


Real signal – Lena



Discrete Cosine Transform (DCT)

- analogy to DFT, real and even
- given sourse data 8x8, the transform projects it into a linear combination of the bases defined as shown in the graphic (with each step – left to right, top to bottom – there is an increase in frequency by 1/2)



Application of DFT – JPEG

(Joint Picture Encoding Group):

- DCT is aplied on “squares” of 8x8 pixel.
- sensibility correction according to the human vision (not all information that is contained in a picture a human eye can see, reduction of redundant information).
- Huffman coding (lossless)
- compromise between quality and the size.
- colors: standard RGB model or some alternative models – see courses of doc. Zemčík.
- compression rate reach up to 1:20.
- drawback: image can contain visible “square” structure (8x8-pixel blocks)

Linear filtering

for 2D signals is usually done by means of FIR filters that are defined by the impulse response:

$$h[i,j] \quad \text{for} \quad -\frac{I}{2} \leq i \leq \frac{I}{2}, \quad -\frac{J}{2} \leq j \leq \frac{J}{2}$$

The filters are most often defined by a matrix "mascs" – with size $(I+1) \times (J+1)$.

Filtering is done as following: "place" the matrix over the image to each point, corresponding values multiply together and then summed up resulting to the output point:

$$y[k,l] = x[k,l] \star h[k,l] = \sum_{i=-\frac{I}{2}}^{\frac{I}{2}} \sum_{j=-\frac{J}{2}}^{\frac{J}{2}} h[i,j]x[k-i,l-j]$$

Example 1. – low-pass filter , $h[i,j] = \mathbf{I}_{10\times 10}$. Lena image is modified by noise $\sigma = 0.1$.

Example 2. – Sobel filters for detection of vertical and horisontal edges:

$$h_v[i,j] = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \qquad h_h[i,j] = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

two detectors together: $y[k, l] = |y_v[k, l]| + |y_h[k, l]|$