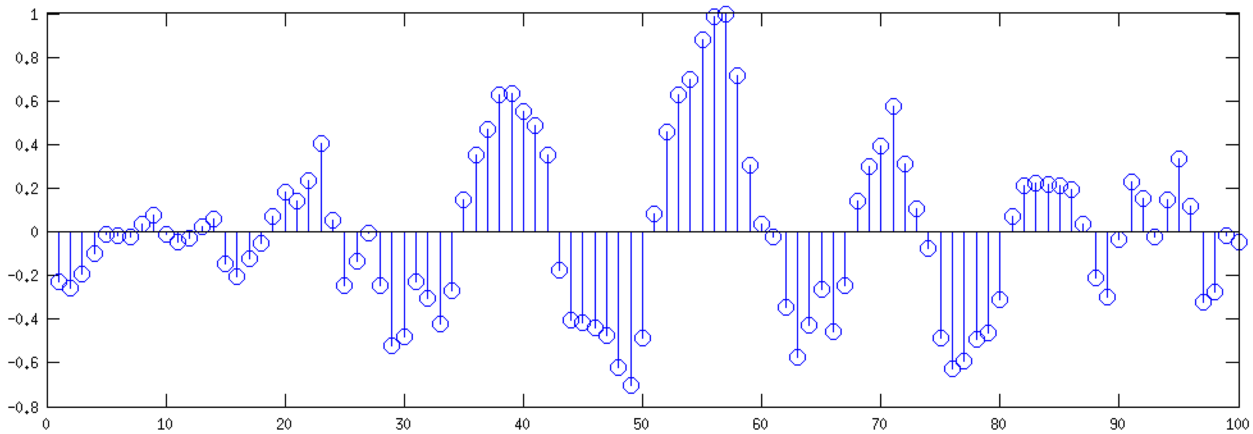
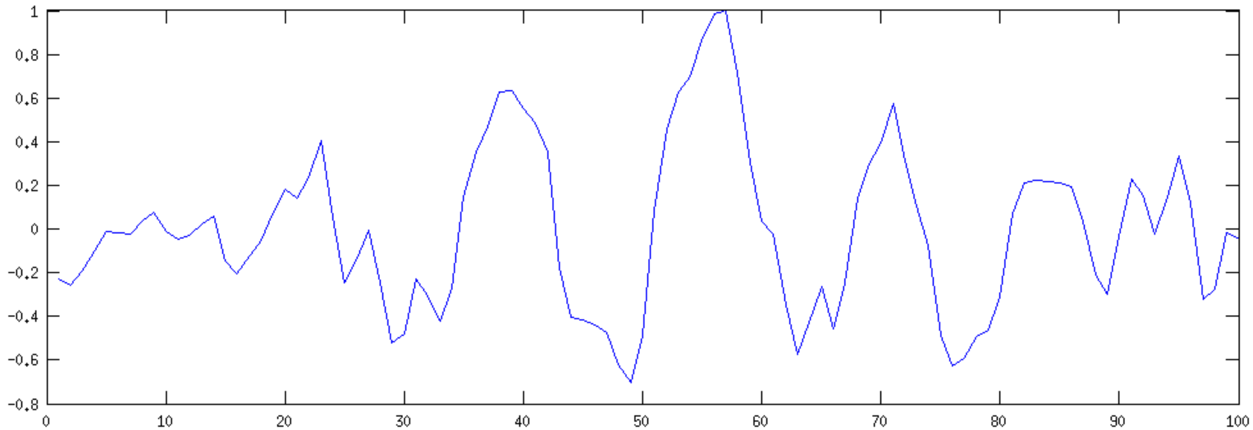
Zpracování obrazů

Honza Černocký, ÚPGM

1D signál

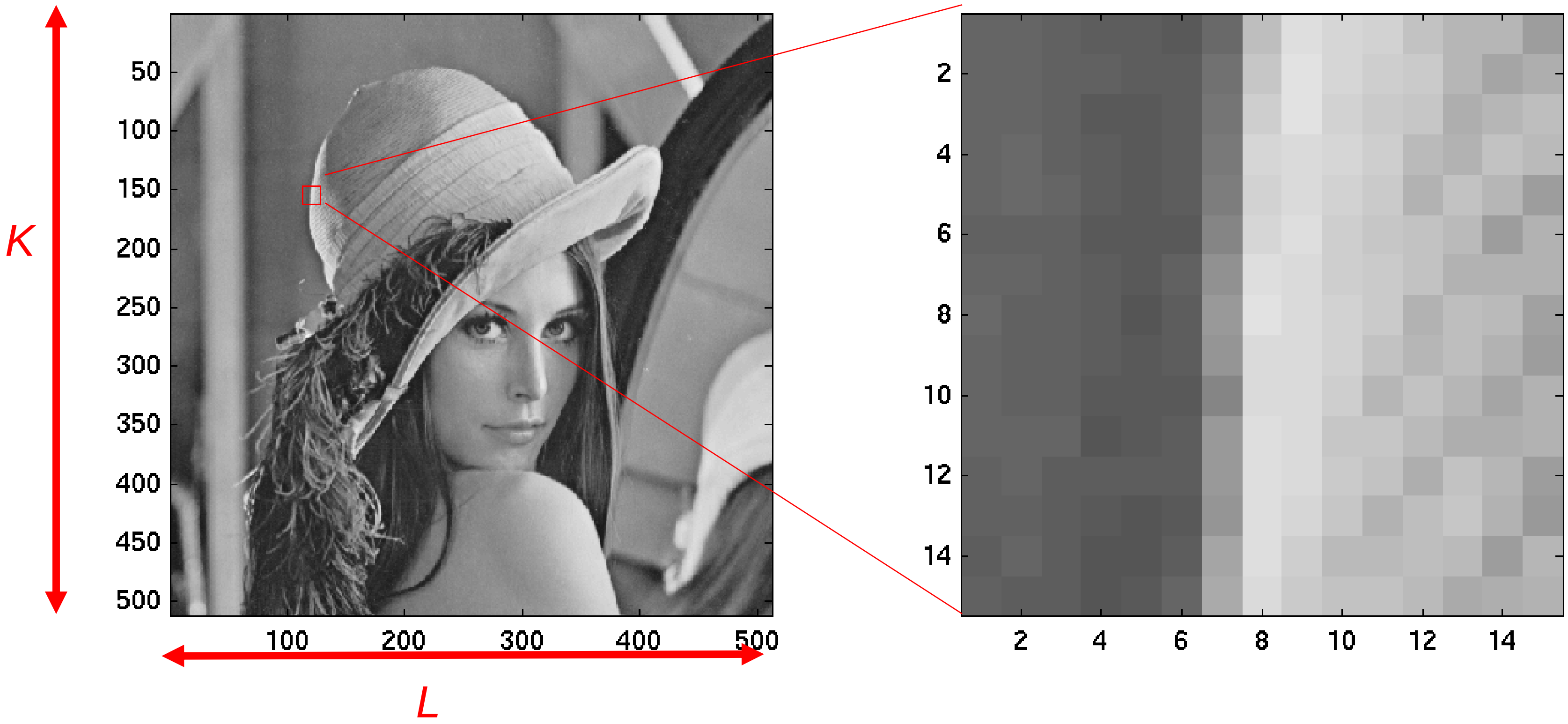

1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
10
20
30
40
50
60
70
80
90
100


1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
0
10
20
30
40
50
60
70
80
90
100

Obrázky

- 2D – šedotónový obrázek (grayscale)
- Několikrát 2D – barevné foto
- 3D – lékařské zobrazování, vektorová grafika, point-clouds (hloubková mapa, Kinect)
- Video – další dimenze je čas

Obrázek



K
L
50
100
150
200
250
300
350
400
450
500
100
200
300
400
500
2
4
6
8
10
12
14
2
4
6
8
10
12
14

Matematicky

$x[k, l]$

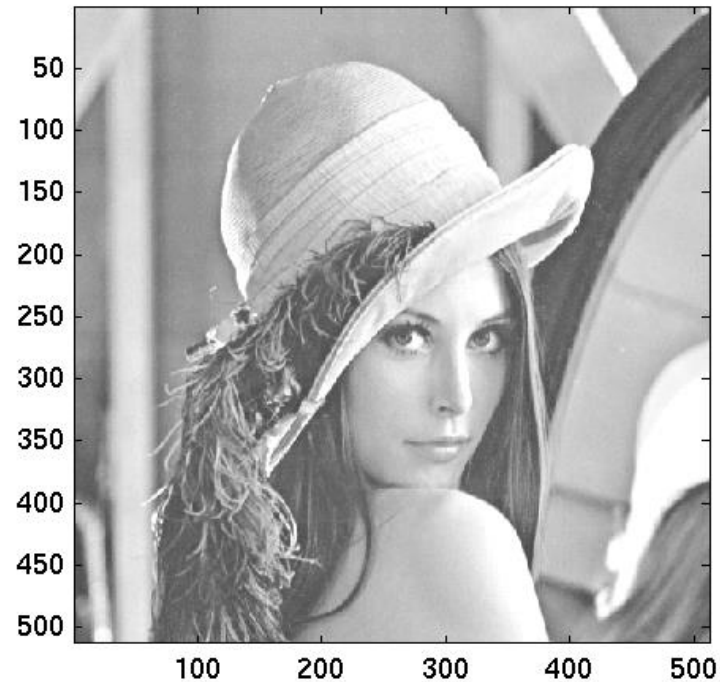
$$x[k,l] = \begin{bmatrix} x[0,0] & x[0,1] & \cdots & x[0,L-1] \\ x[1,0] & x[1,1] & \cdots & x[1,L-1] \\ \vdots & & & \vdots \\ x[K-1,0] & x[K-1,1] & \cdots & x[K-1,L-1] \end{bmatrix}$$

Representace vzorků

- Běžně kvantovány
- 8/16 bitů
- 0…255 atd
- Pro výpočty lépe [0…1]
- A jako **floats**

Operace nad pixely – jas plus

$$y[k,l] = x[k,l] + const.$$

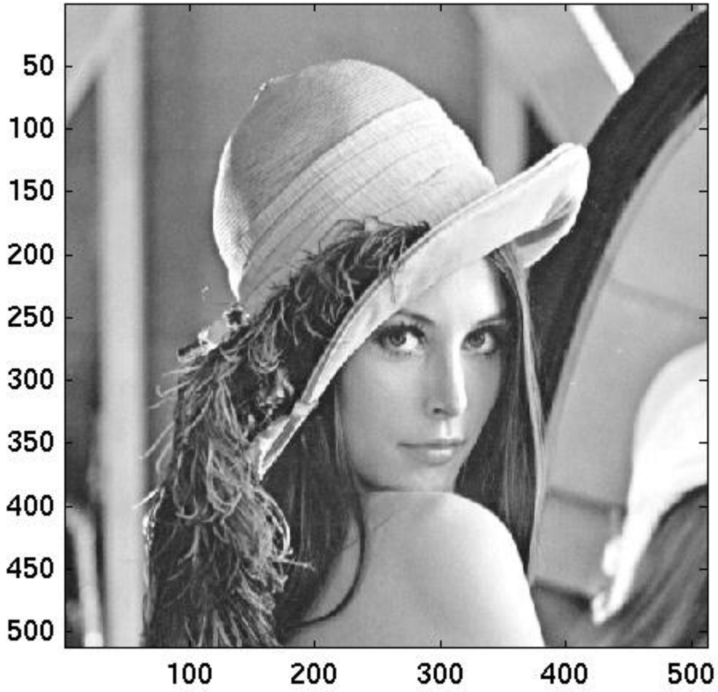


Operace nad pixely – jas minus

$$y[k,l] = x[k,l] + const.$$

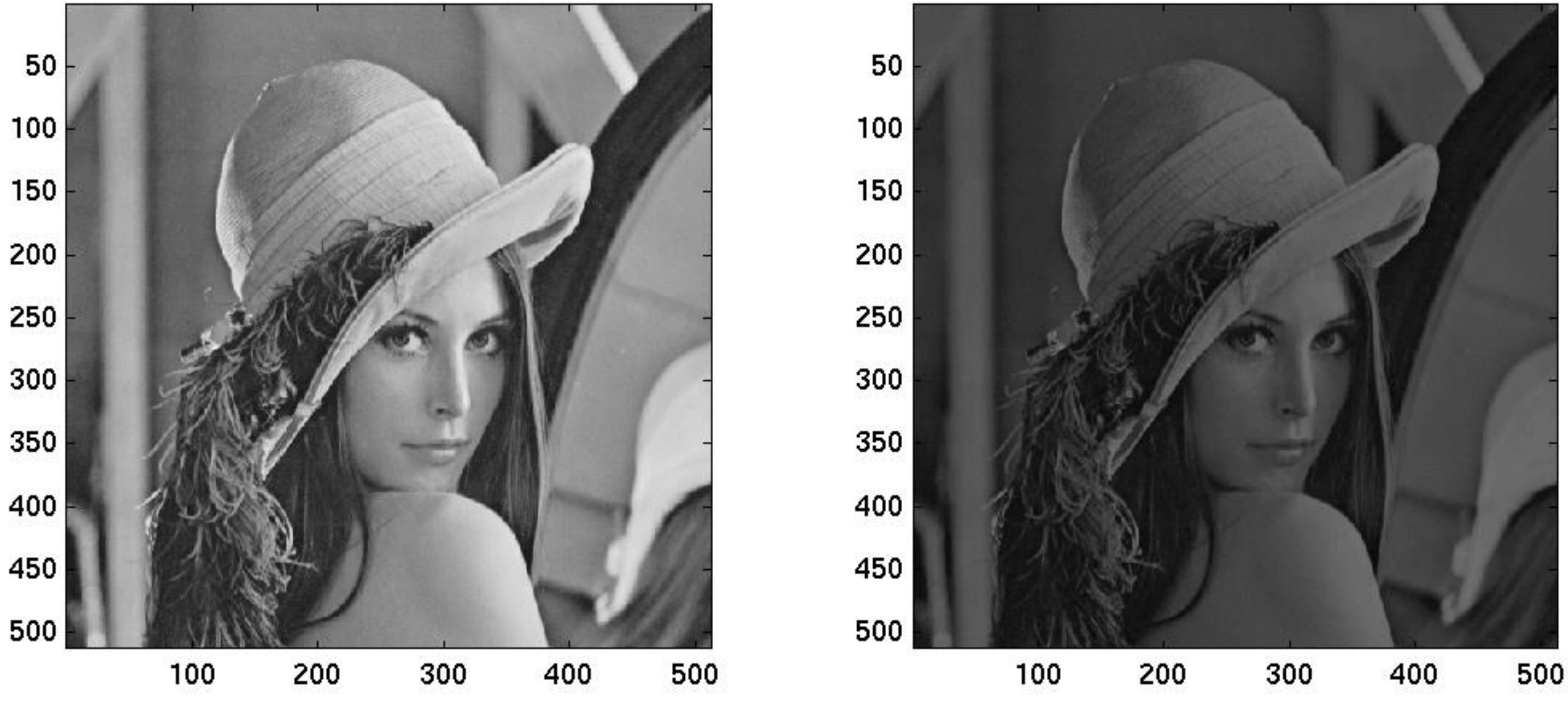
Kontrast plus

$$y[k,l] = x[k,l] \times const.$$



Kontrast minus

$$y[k,l] = x[k,l] \times const.$$



Co se špatnými hodnotami ?

x[k,l] < 0 nebo *x[k,l] > 1*

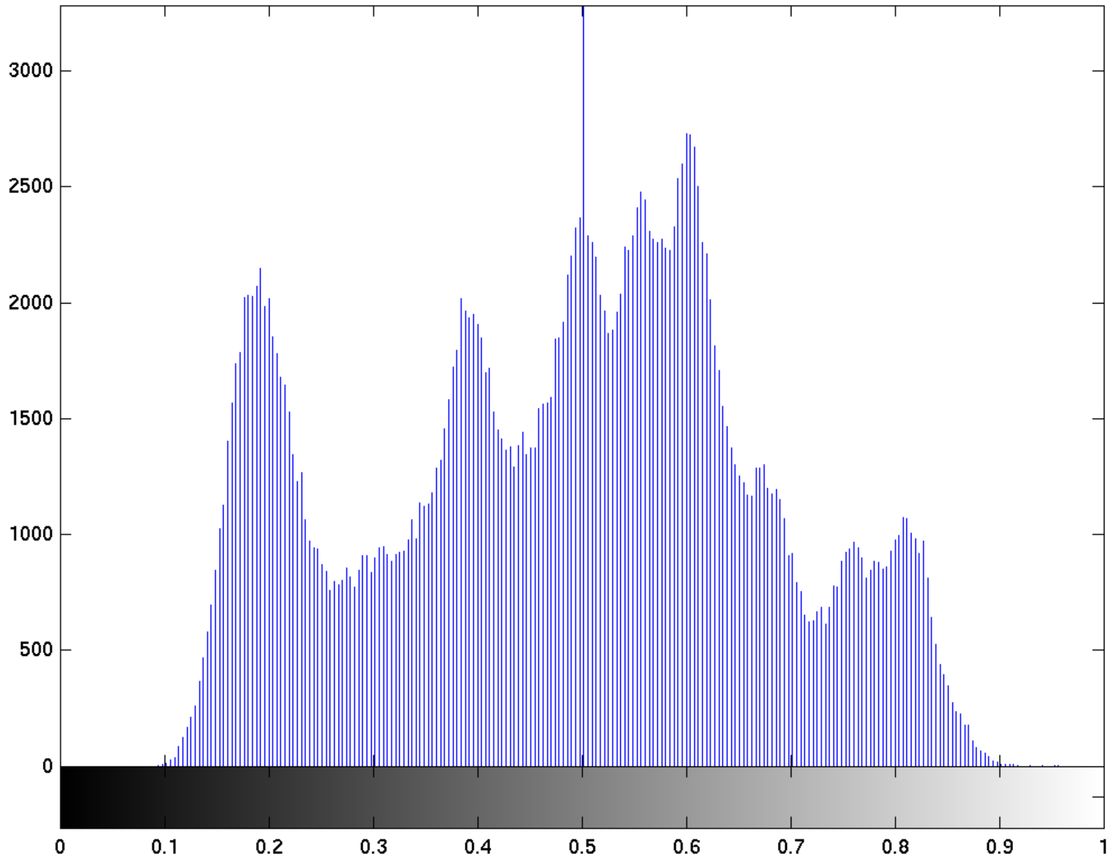
- Nechat …
- Klipovat:

$$y[k,l] = 0, \quad \text{pokud} \;\; y[k,l] < 0$$

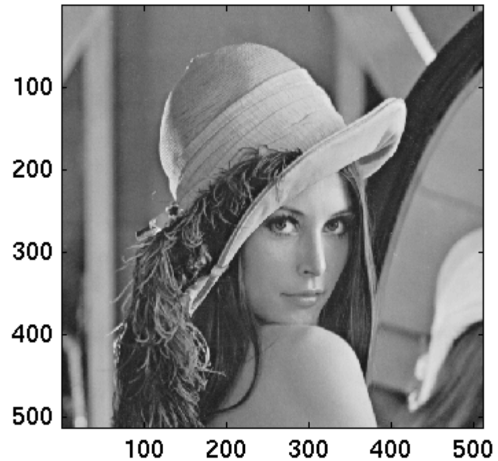
$$y[k,l] = 1, \quad \text{pokud} \;\; y[k,l] > 1$$

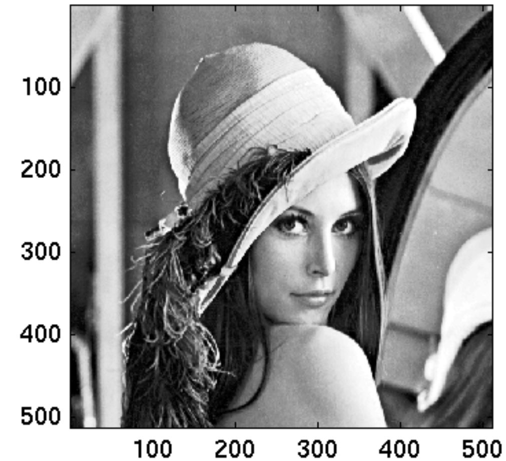
Využití statistik hodnot

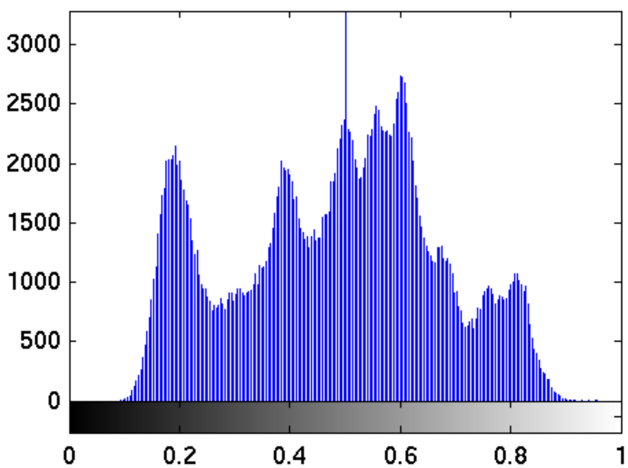
- Histogram

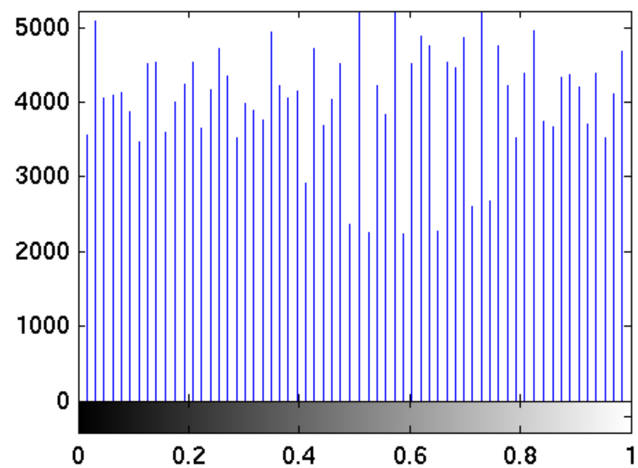


Ekvalizace histogramu

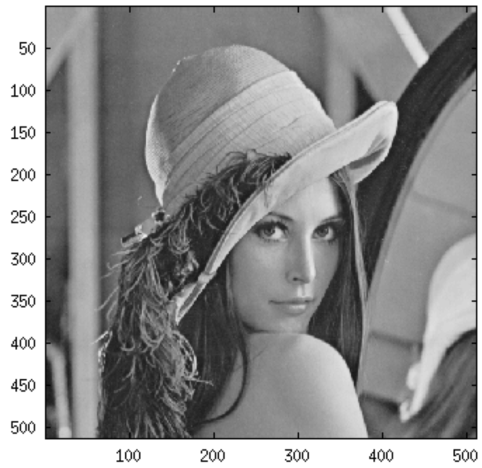




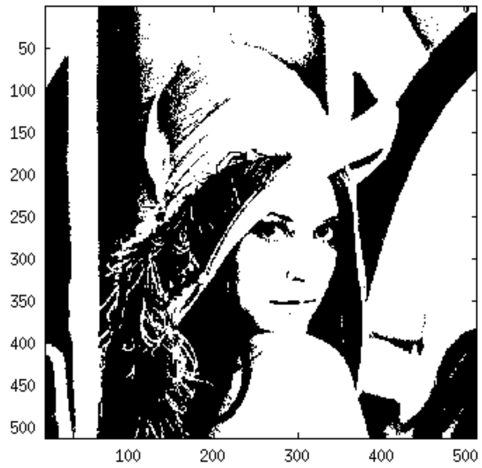




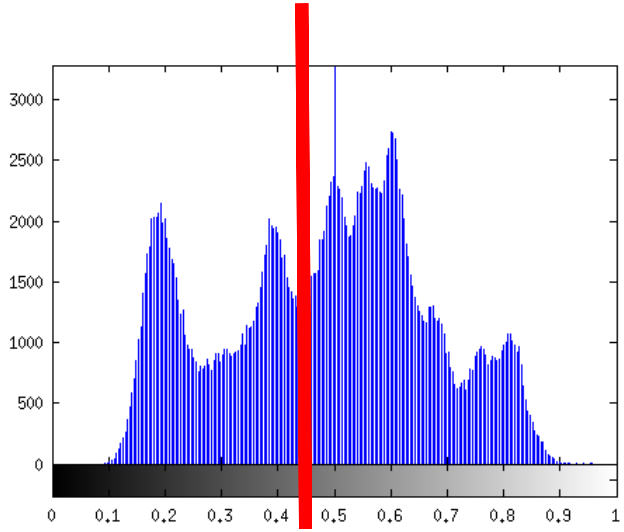
Prahování



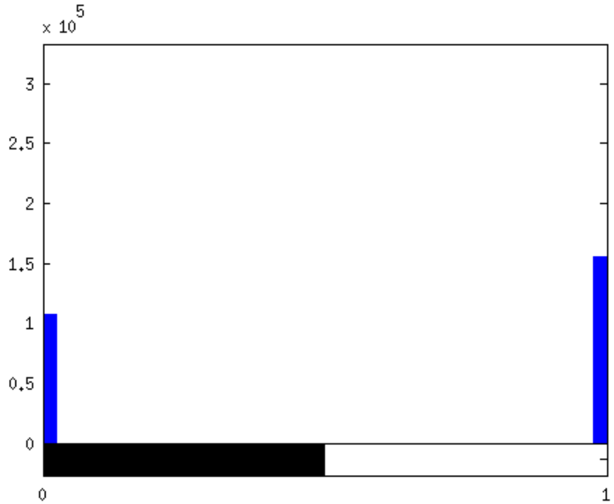
50
100
150
200
250
300
350
400
450
500
100
200
300
400
500



50
100
150
200
250
300
350
400
450
500
100
200
300
400
500



3000
2500
2000
1500
1000
500
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



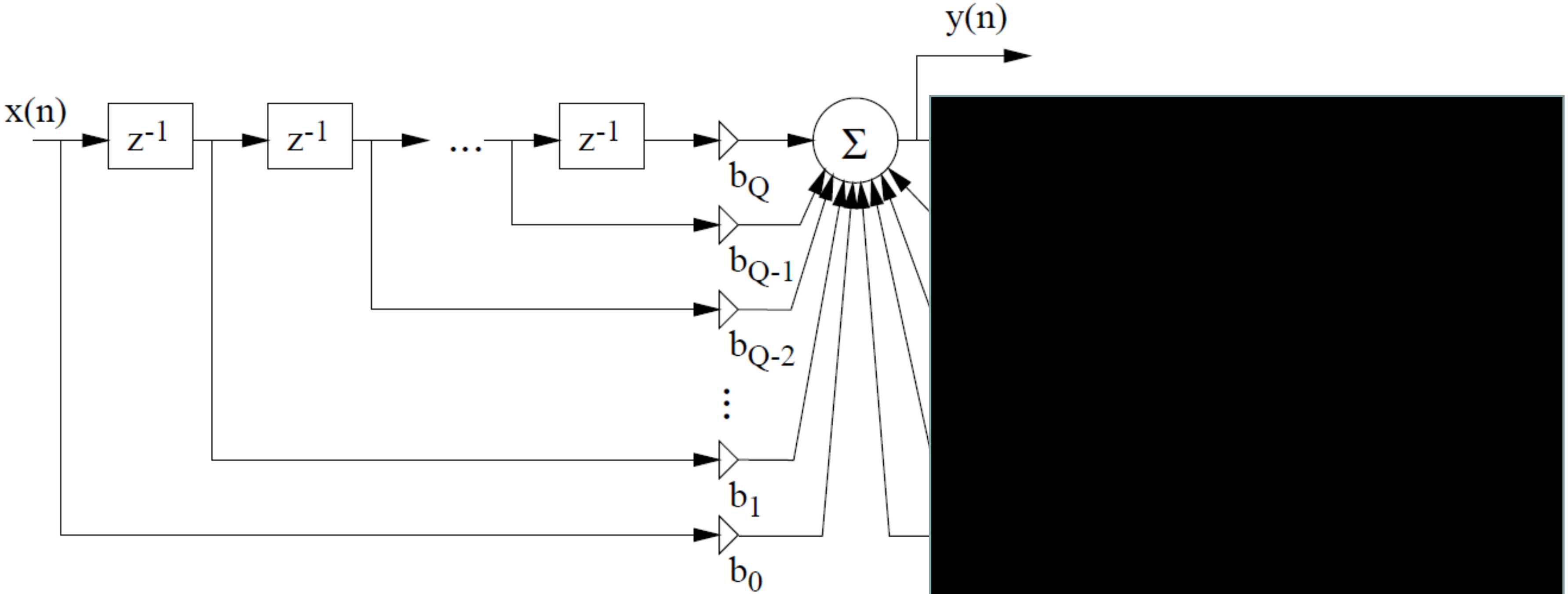
x 10^5
3
2.5
2
1.5
1
0.5
0
0
1

Filtrace

Úkol ?

Získat **nový signál** s požadovanými vlastnostmi.

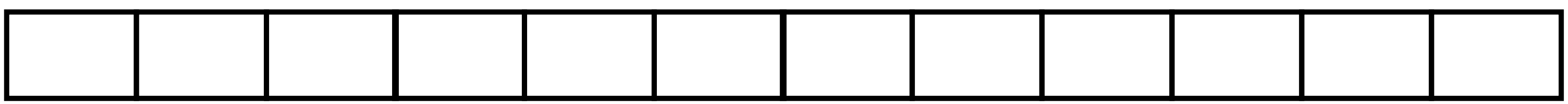
Filtrace – opakování 1D

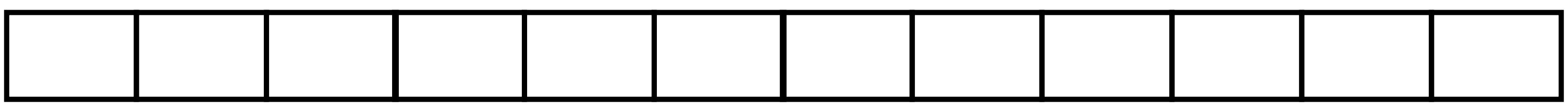


[redacted]

Filtrace – opakování 1D

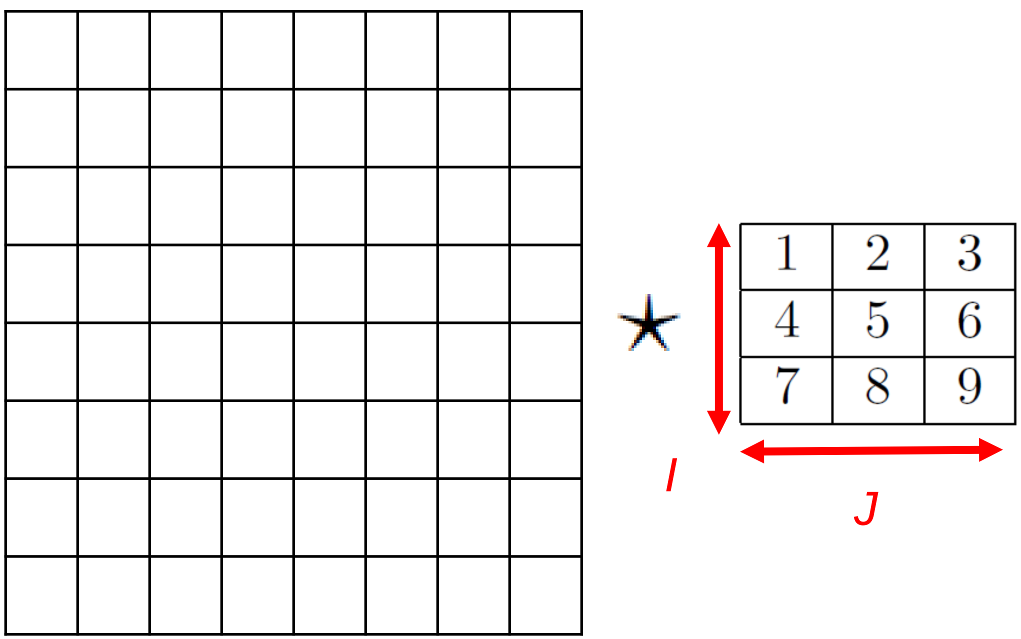
$$y[n] = x[n] \star h[n] = \sum_{k=0}^{Q} h[k]x[n-k]$$



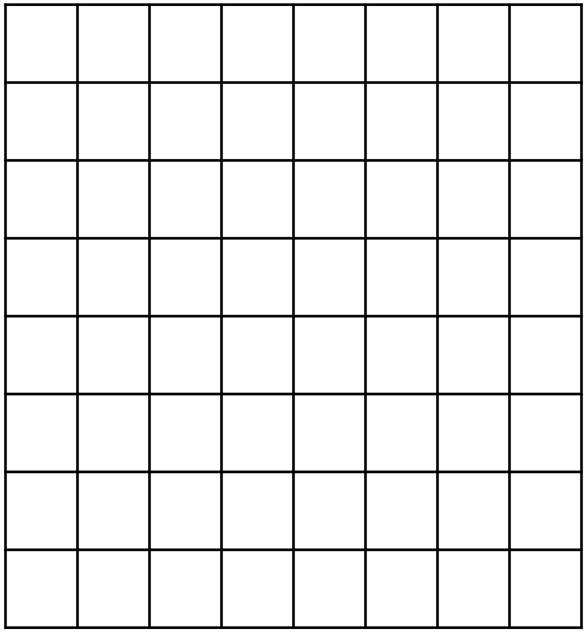


2D filtr

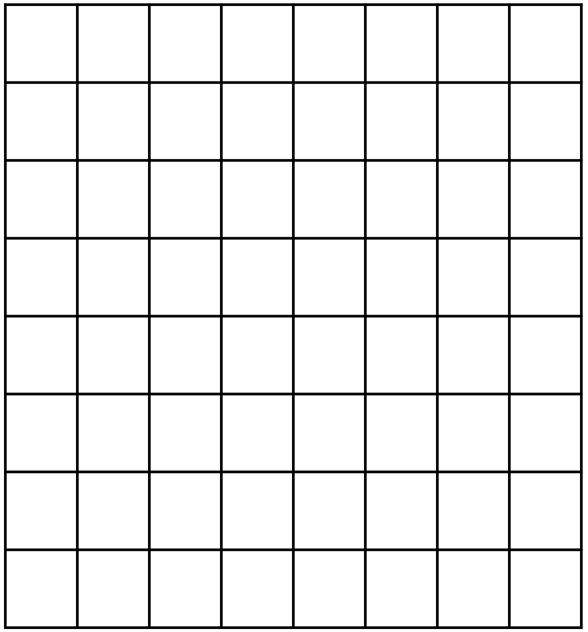
$$y[k,l] = x[k,l] \star h[k,l] = \sum_{m=-\frac{I-1}{2}}^{\frac{I-1}{2}} \sum_{n=-\frac{J-1}{2}}^{\frac{J-1}{2}} h[m,n]x[k-m,l-n]$$



Filtrace …



9	8	7
6	5	4
3	2	1



Okraj ? Třeba nuly …

Koeficienty *h[k,l]*

Říkáme jim

- Koeficienty
- Maska
- Konvoluční jádro …

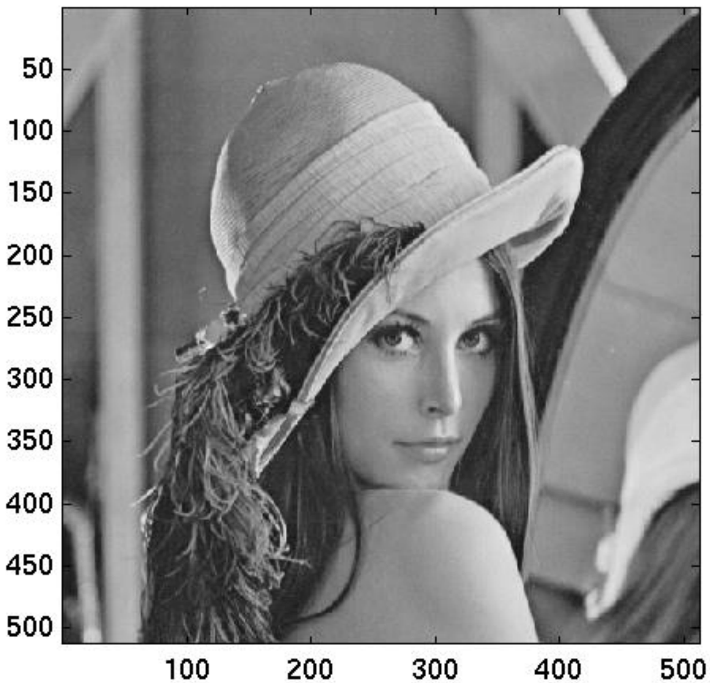
Chceme po nich

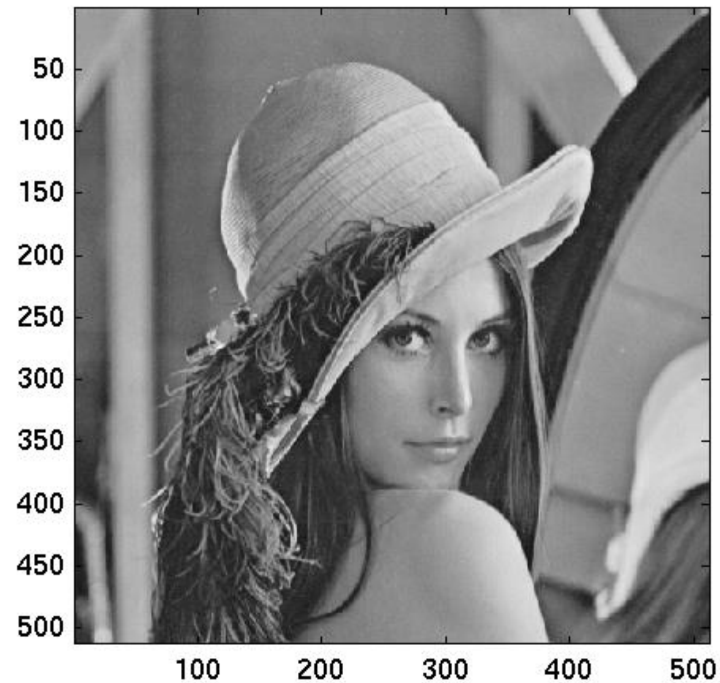
- Aby měly nějaký smysl …
- Aby neměnily dynamiku signálu

$$\sum_k \sum_l |h[k,l]| = 1$$

Drát

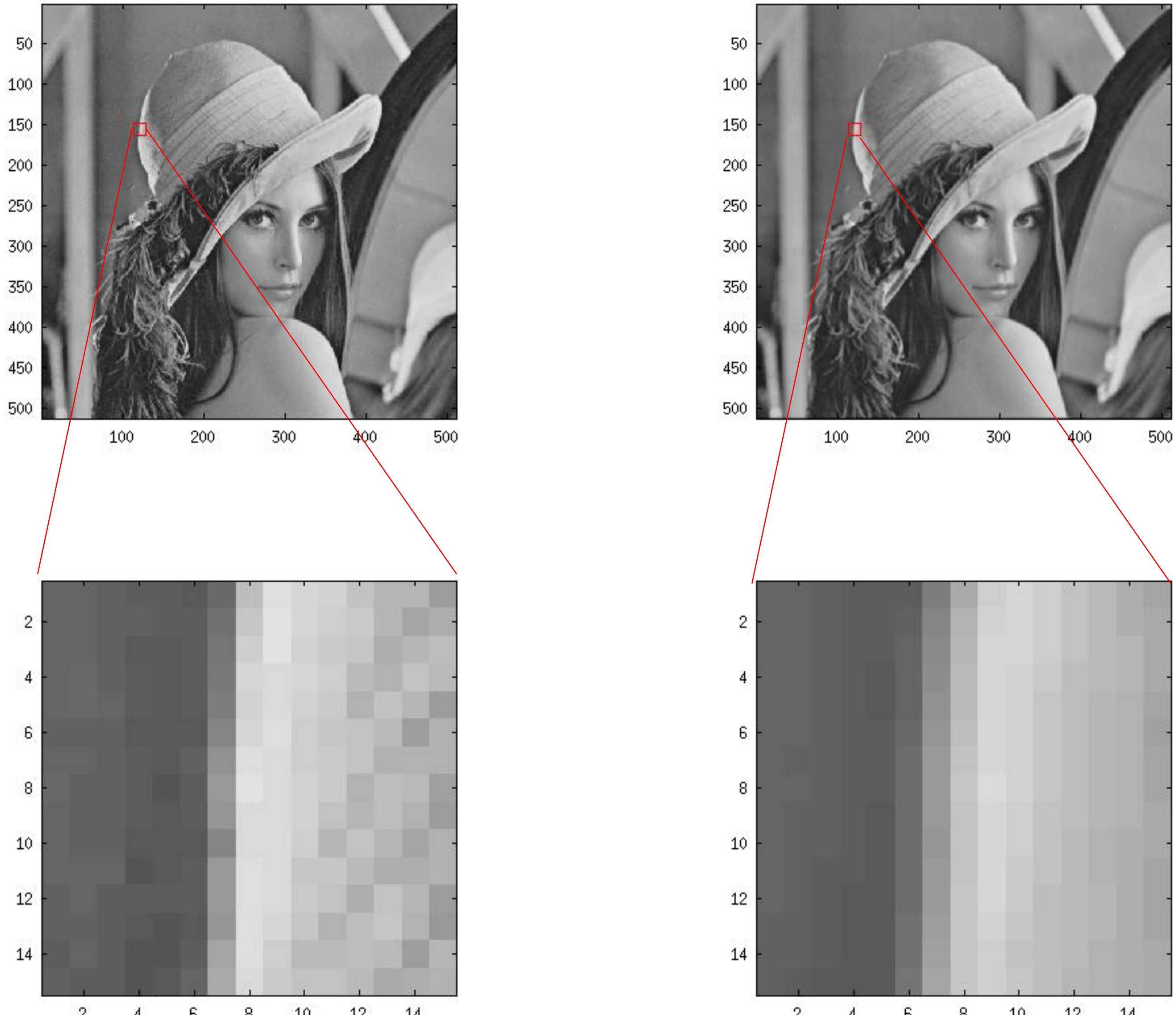
$$h[k,l] = [1]$$


50
100
150
200
250
300
350
400
450
500
100
200
300
400
500


50
100
150
200
250
300
350
400
450
500
100
200
300
400
500

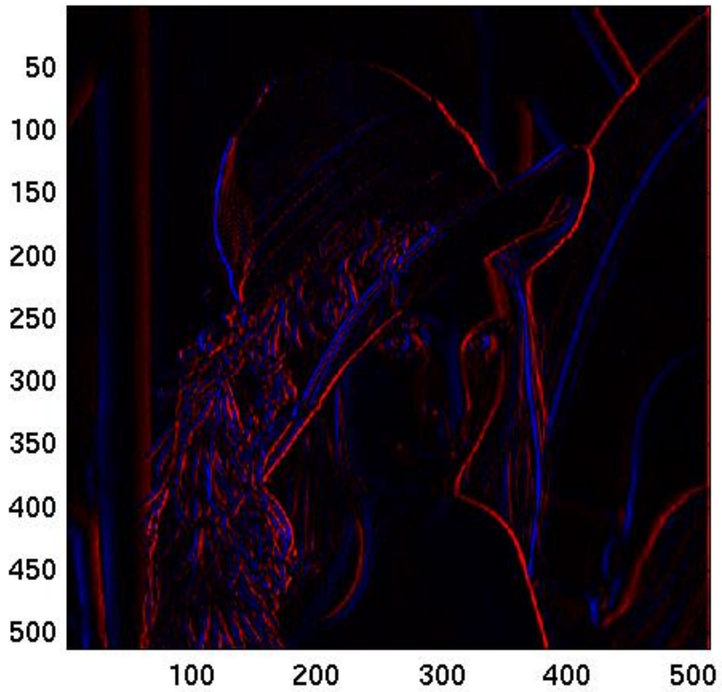
Vyhlazování

$$h[k,l] = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



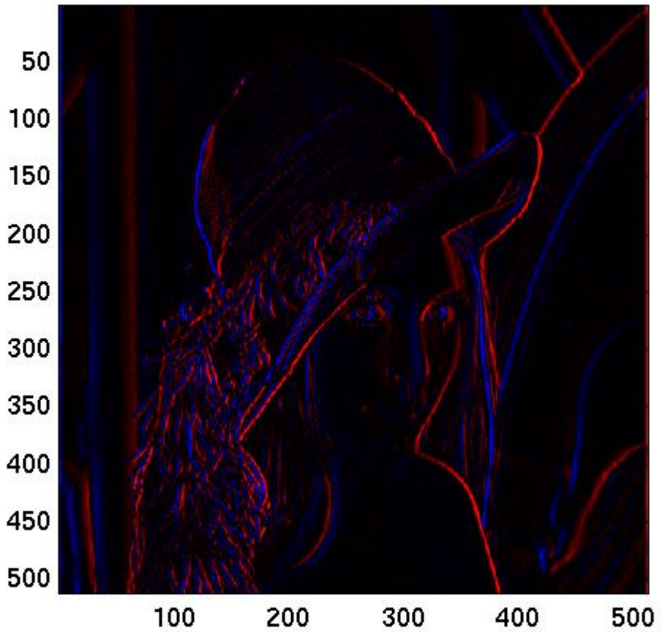
Detektor vertikálních hran

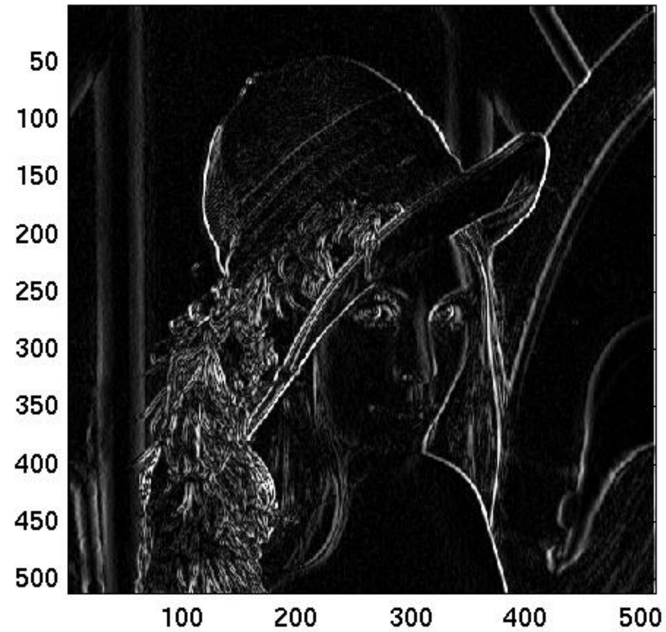
$$h_v[k,l] = \frac{1}{8}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$



Co se zápornými vzorky ?

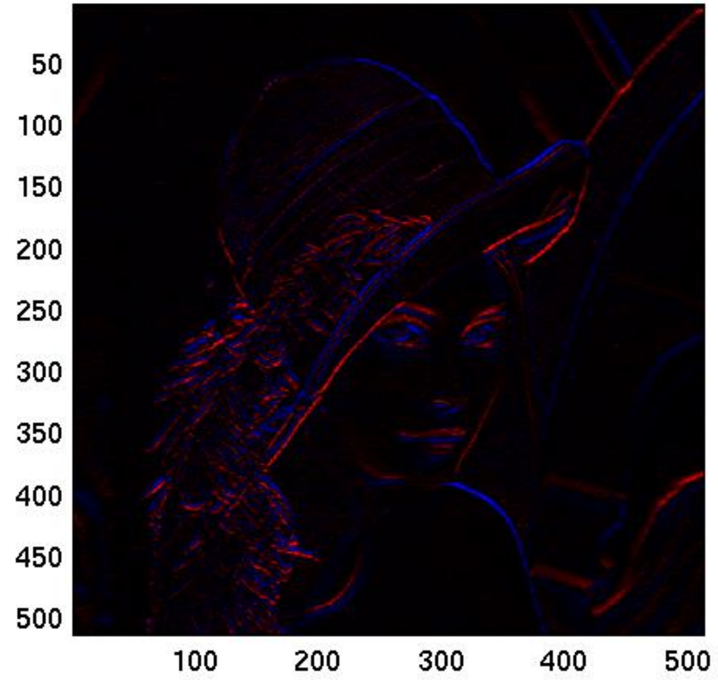
- Můžeme nechat tak pro další výpočty …
- Nebo vymyslet, jak je převést do [0…1] pro zobrazování – např. absolutní hodnota





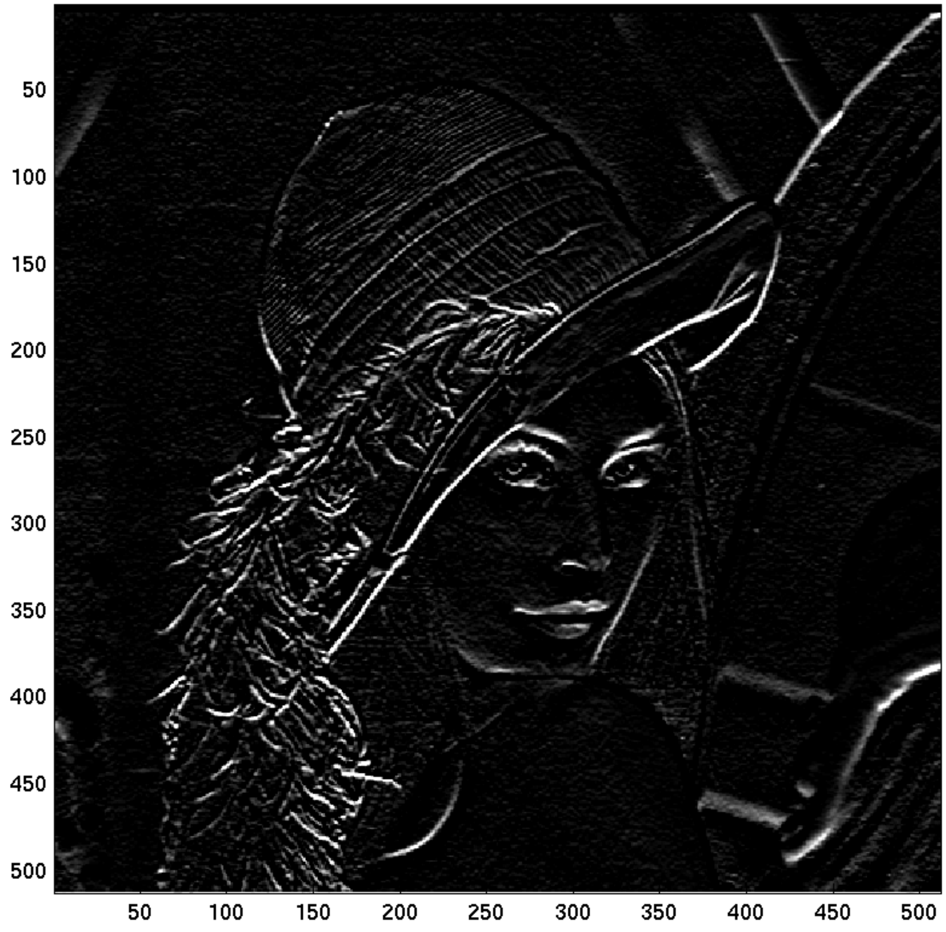
Detektor horizontálních hran

$$h_h[k,l] = \frac{1}{8}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$



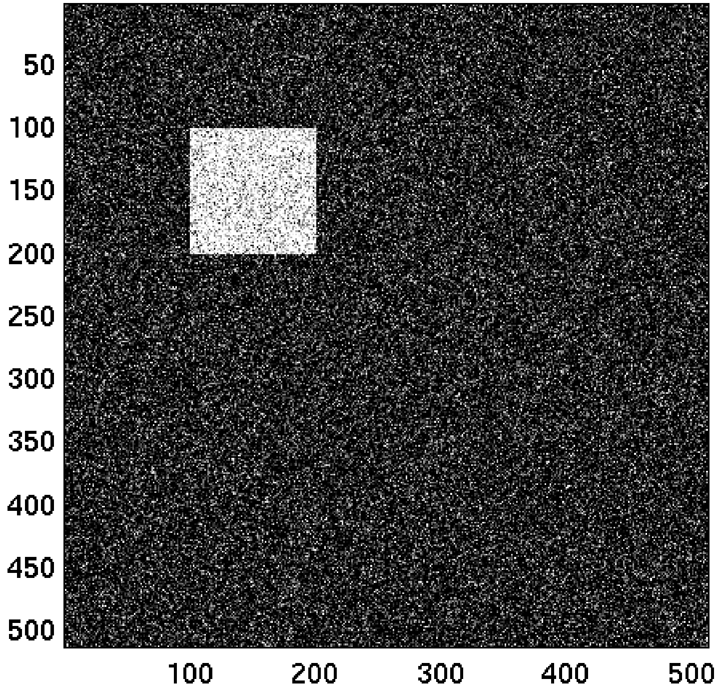
Obojí dohromady …

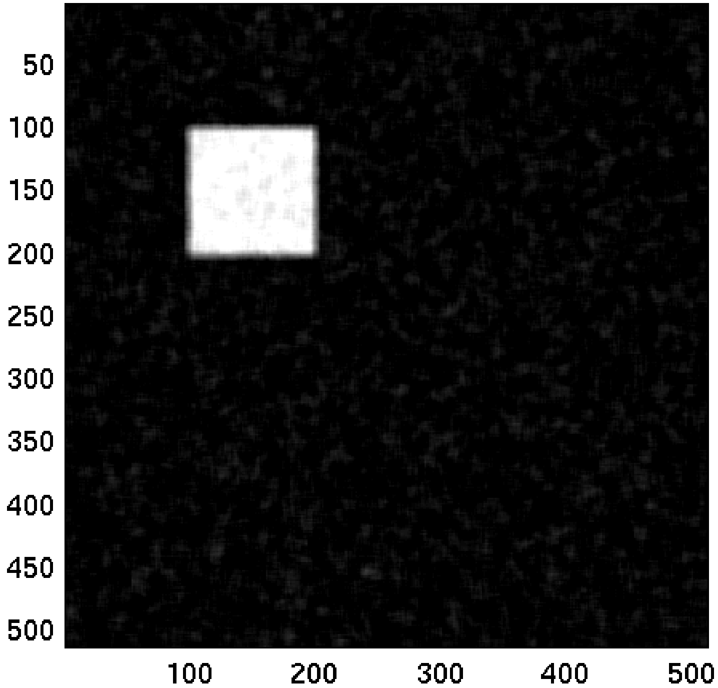
$$y[k,l] = |y_v[k,l]| + |y_h[k,l]|$$



Odšumování …

- Maska 9x9, hodnoty 1/81 … **zoom noise_low_pass.png**

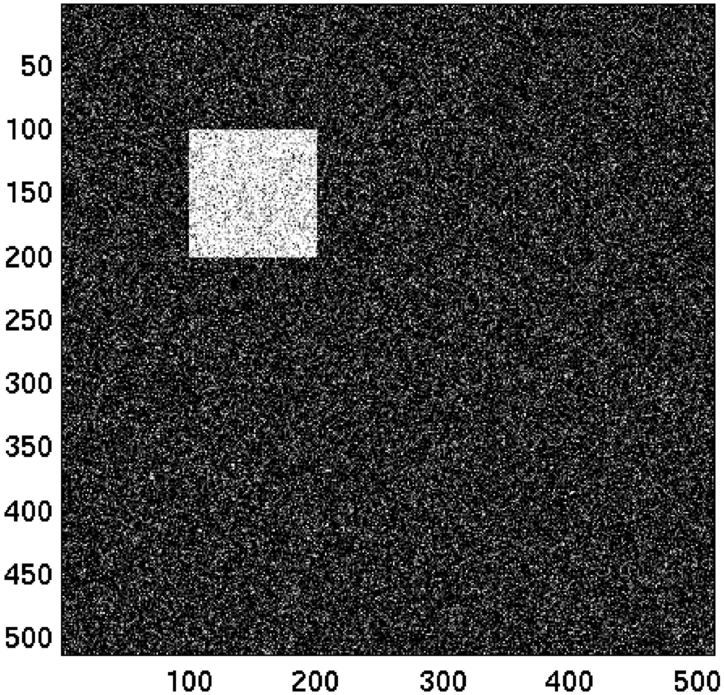


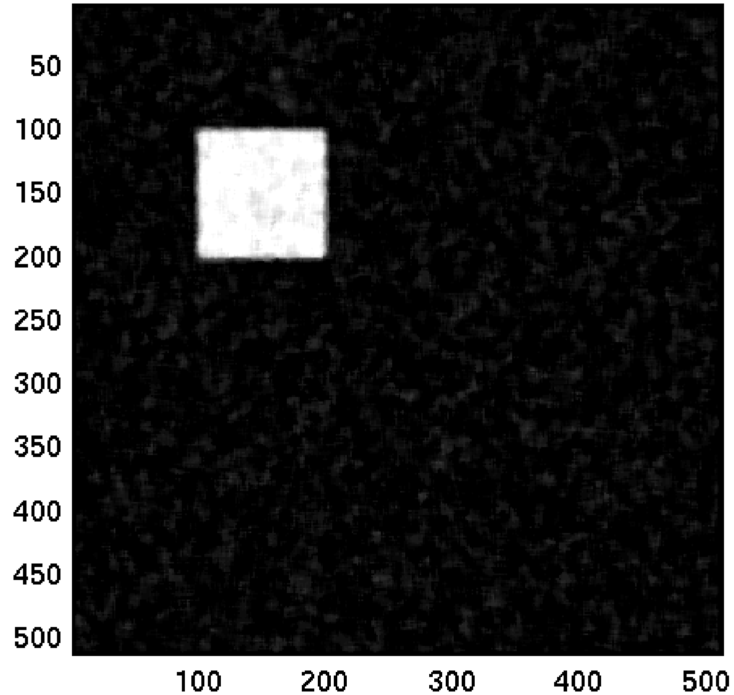


Odšumování II – mediánový filtr

$$y[k,l] = \text{median}_{k=-\frac{I-1}{2}\dots\frac{I-1}{2},\quad l=-\frac{J-1}{2}\dots\frac{J-1}{2}} x[k,l]$$

- **zoom noise_median.png**





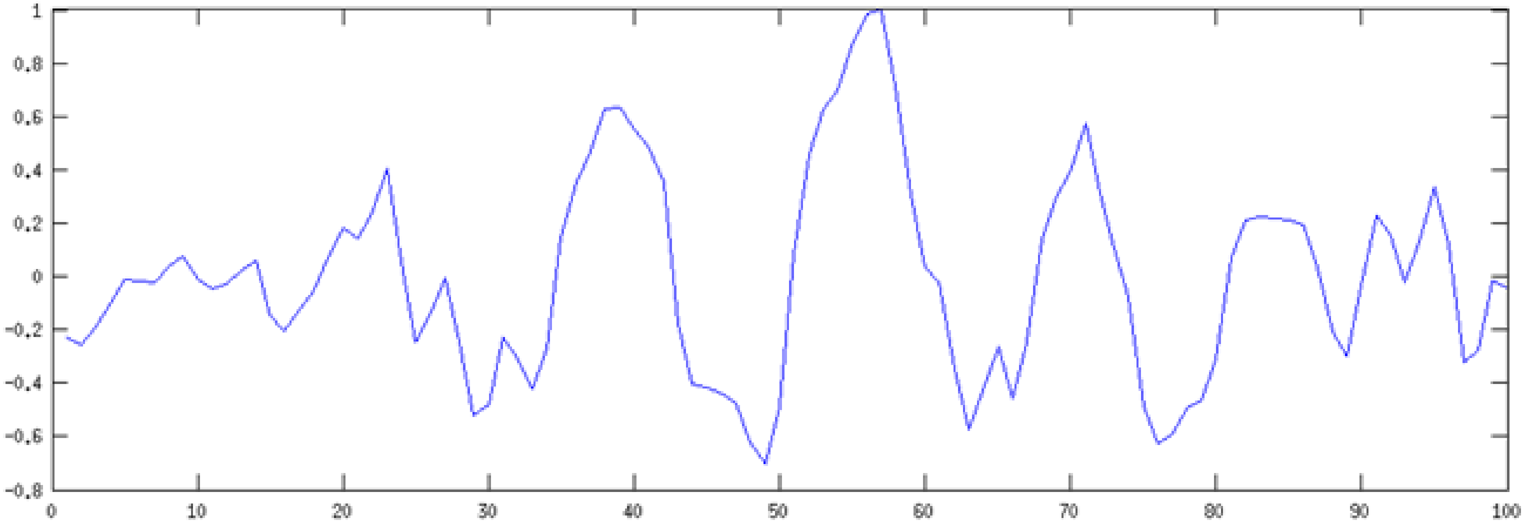
Spektrální analýza

- Úkol:

  Zjistit, co je v signálu na jakých frekvencích

- Proč ?
  - Vizualisace,
  - Výpočet parametrů (think of Facebook)
  - Filtrace (převod do spektrální oblasti a násobení tam může být efektivnější)
  - Kódování (think of JPEG)

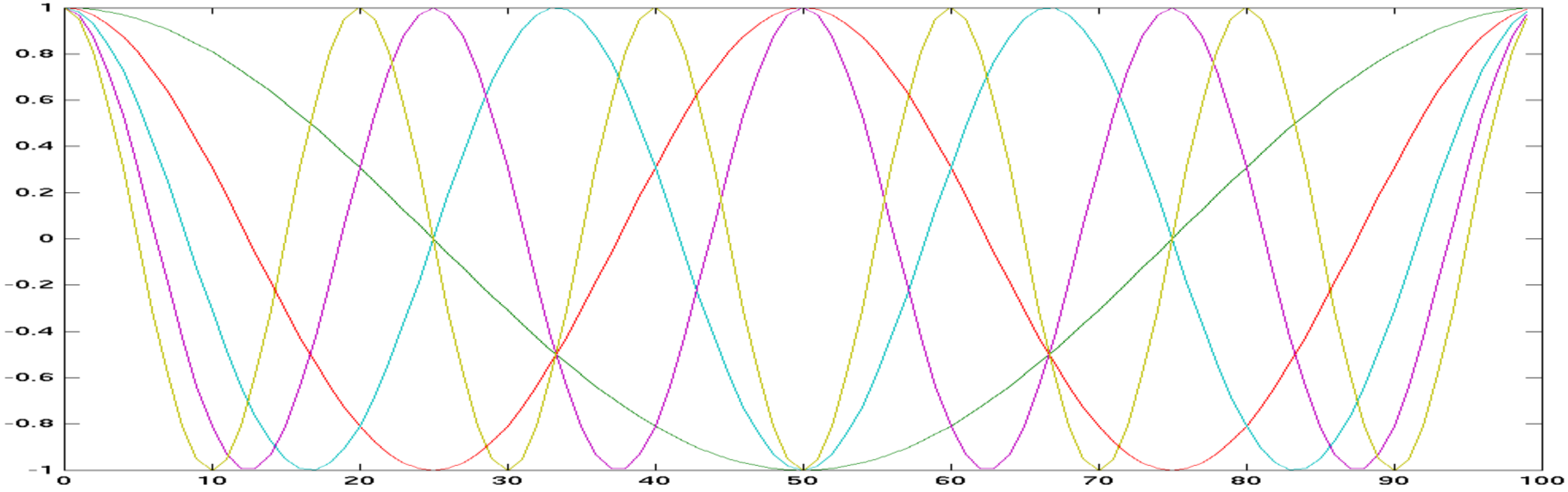
Opakování – 1D



- Co je v něm ?

$$c = \sum_{n=0}^{N-1} x[n]a[n]$$

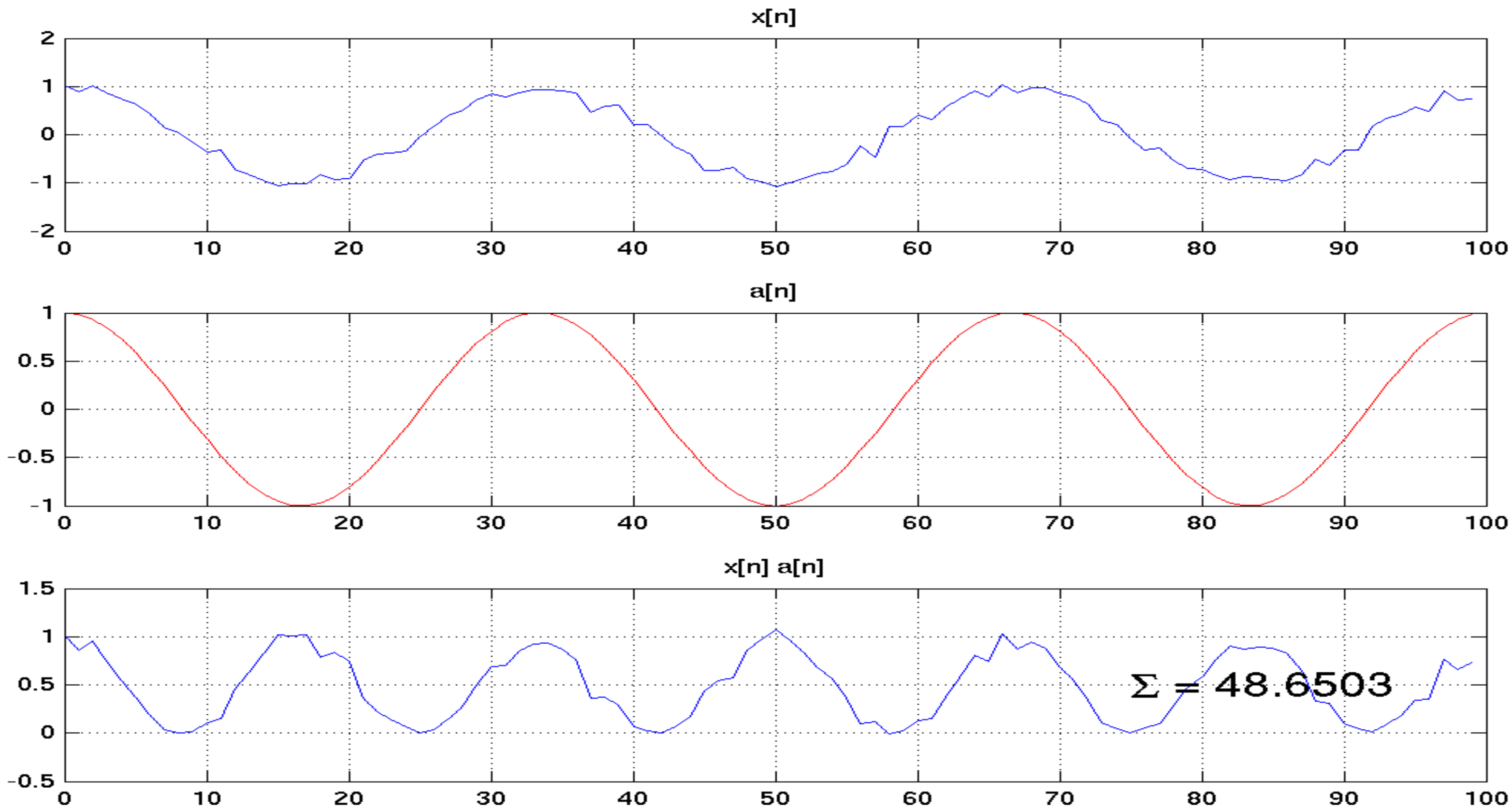
Opakování – 1D cosinové báze



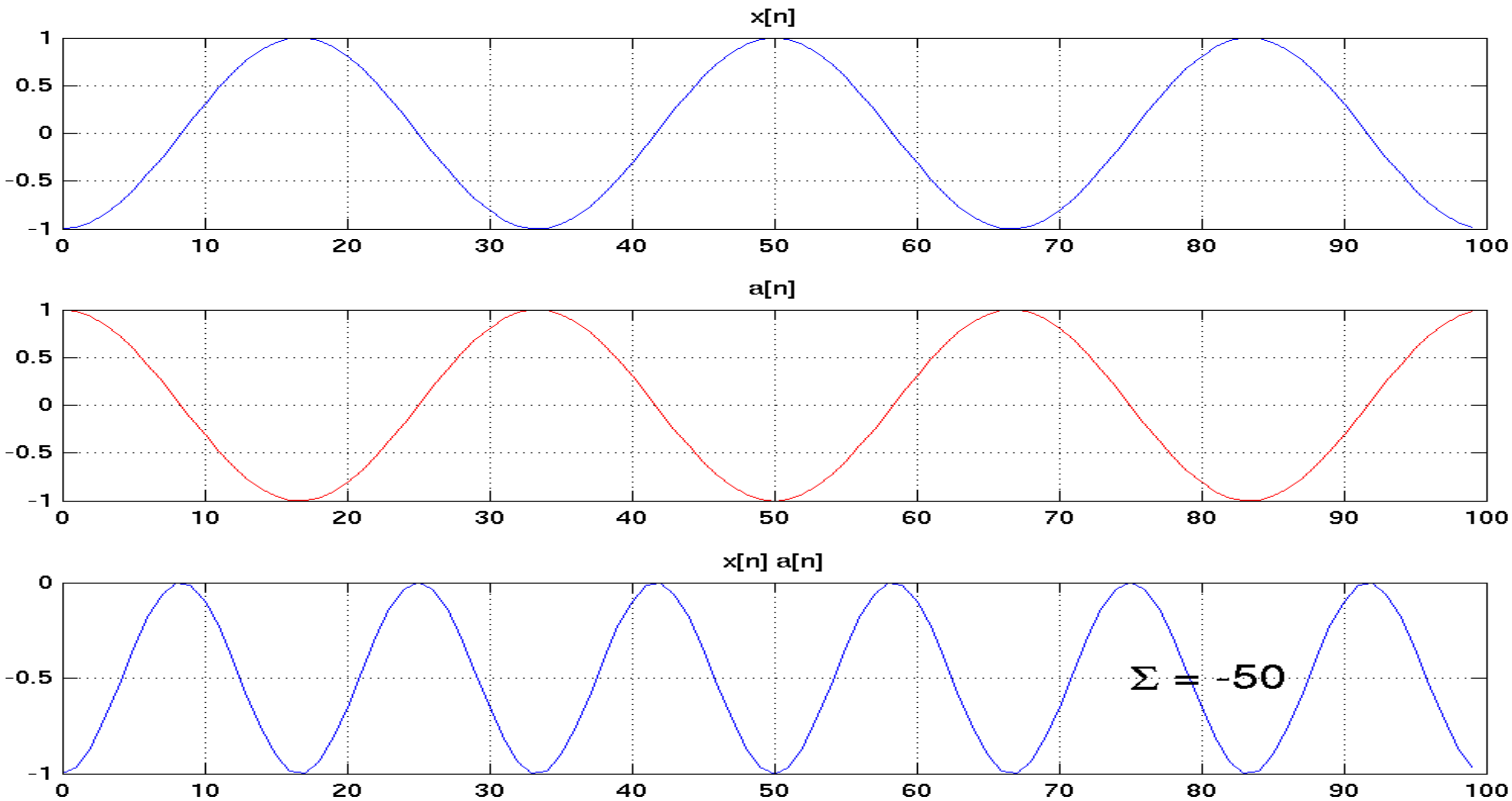
$$\cos(2\pi \frac{k}{N} n)$$

Problém s fází – „vlna“ signálu může začínat jinde než v nule …

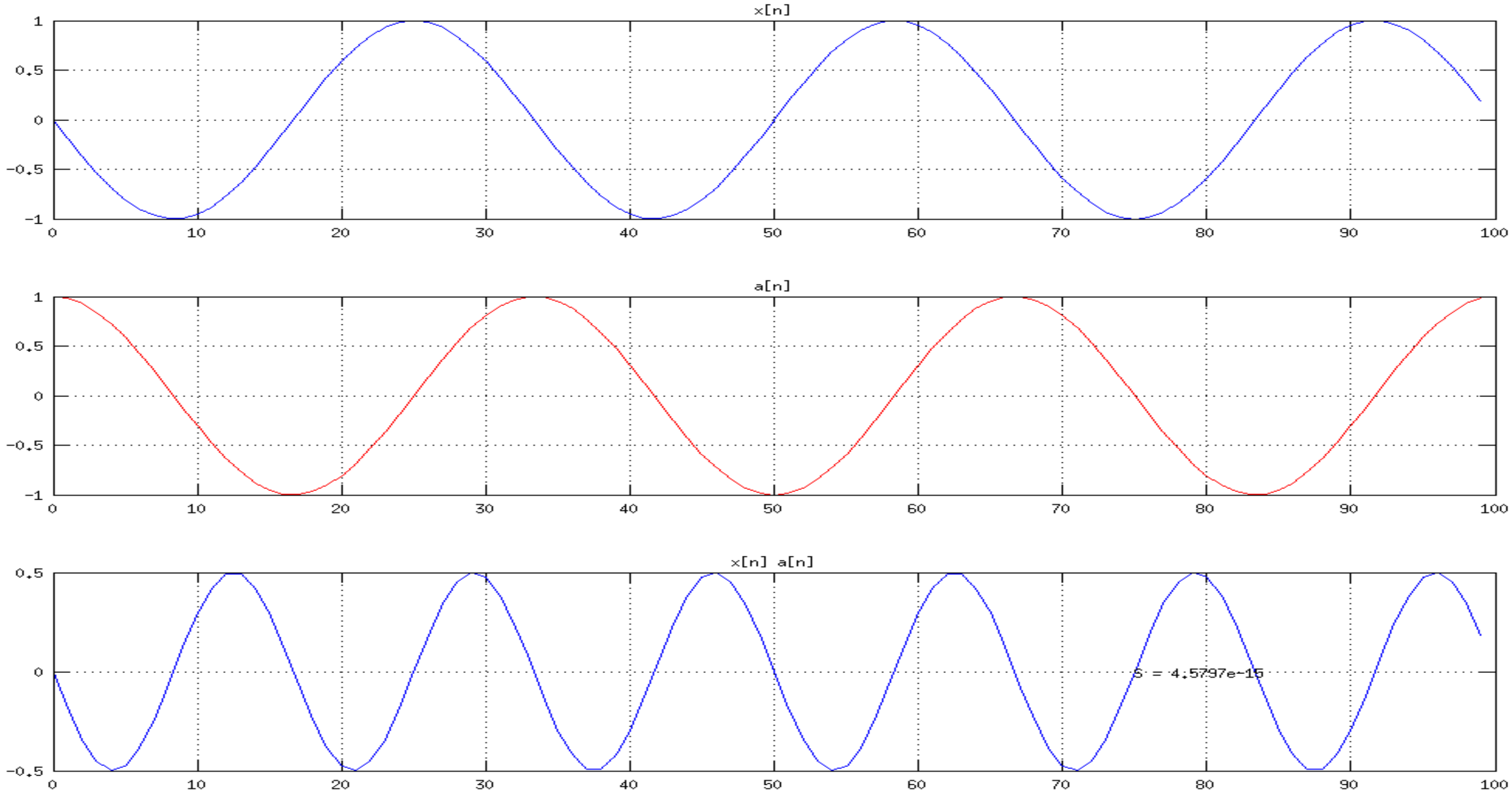
Tady dobře …



Tady také dobře …

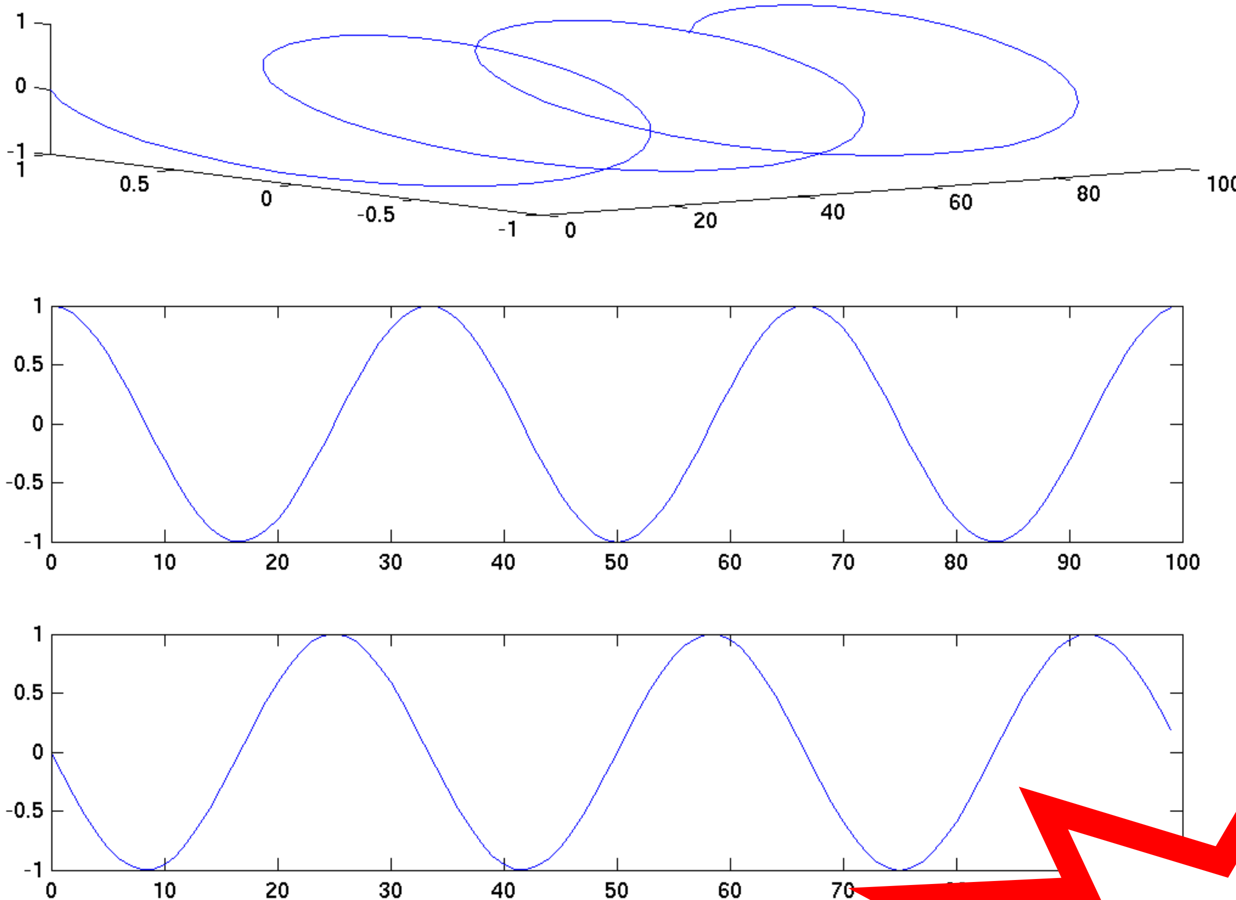


Aj aj ☹



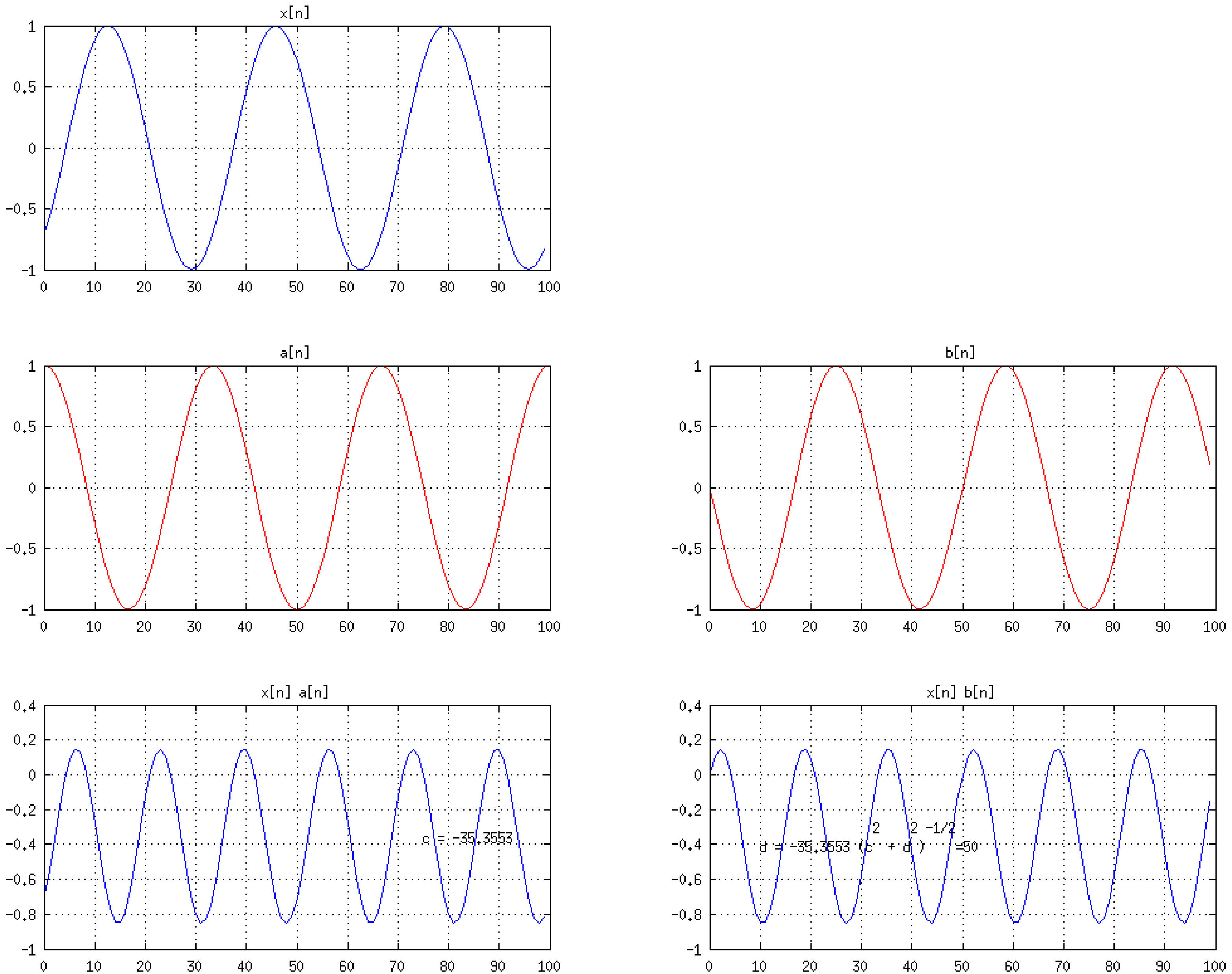
- Jak to že nula, když $\sin(x) = \cos(x - \frac{\pi}{2})$

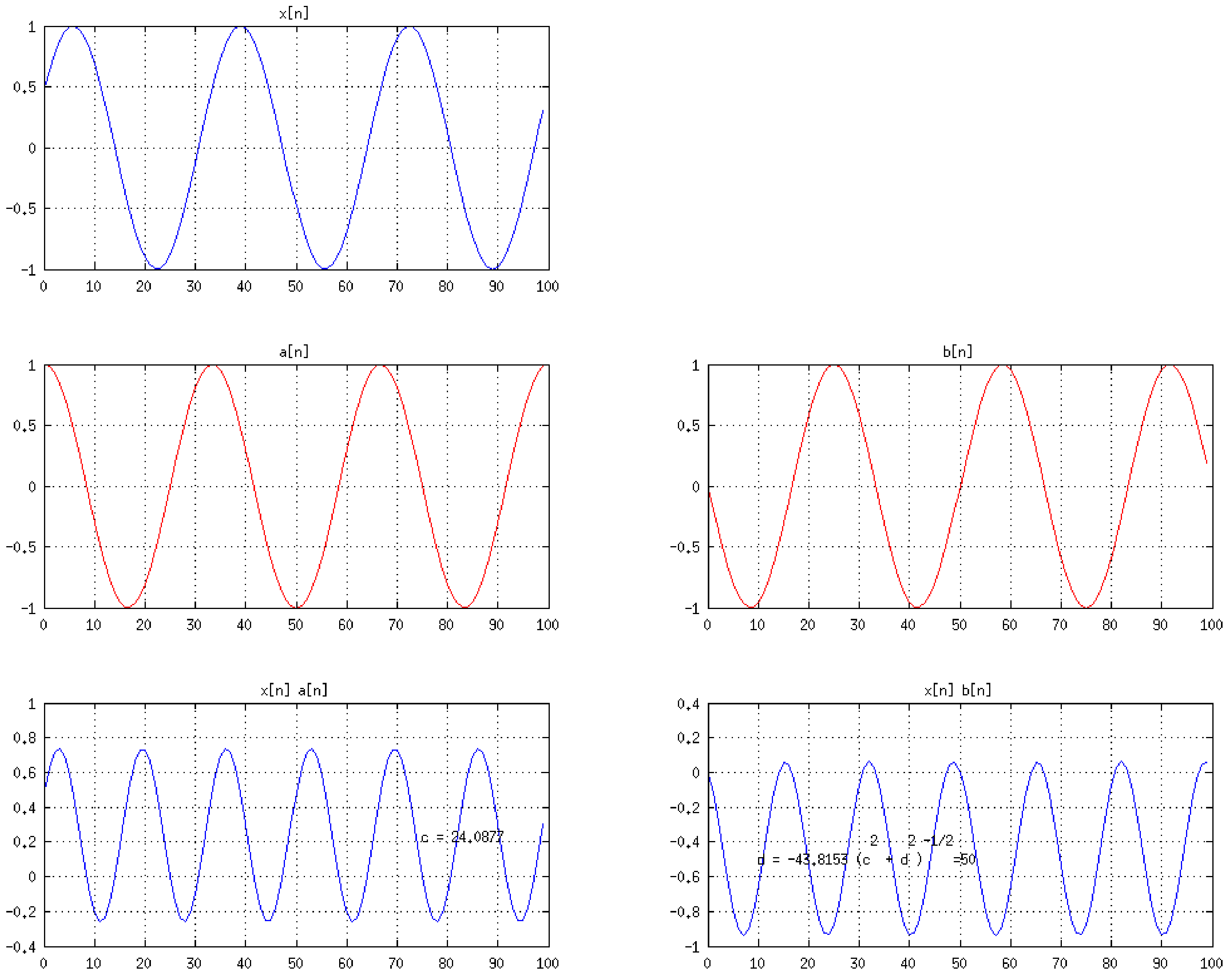
Komplexní exponenciály …



$$e^{jx} = \cos(x) + j\sin(x)$$

$$e^{-j2\pi\frac{k}{N}n} = \cos(2\pi\frac{k}{N}n) - j\sin(2\pi\frac{k}{N}n)$$


x[n]
a[n]
b[n]
x[n] a[n]
c = -35.3553
x[n] b[n]
d = -35.3553 (c^2 + d^2)^-1/2 =50


x[n]
a[n]
b[n]
x[n] a[n]
x[n] b[n]
c = 24.0877
2 2 -1/2
d = -43.8153 (c + d) =50

1D ultimate result - DFT

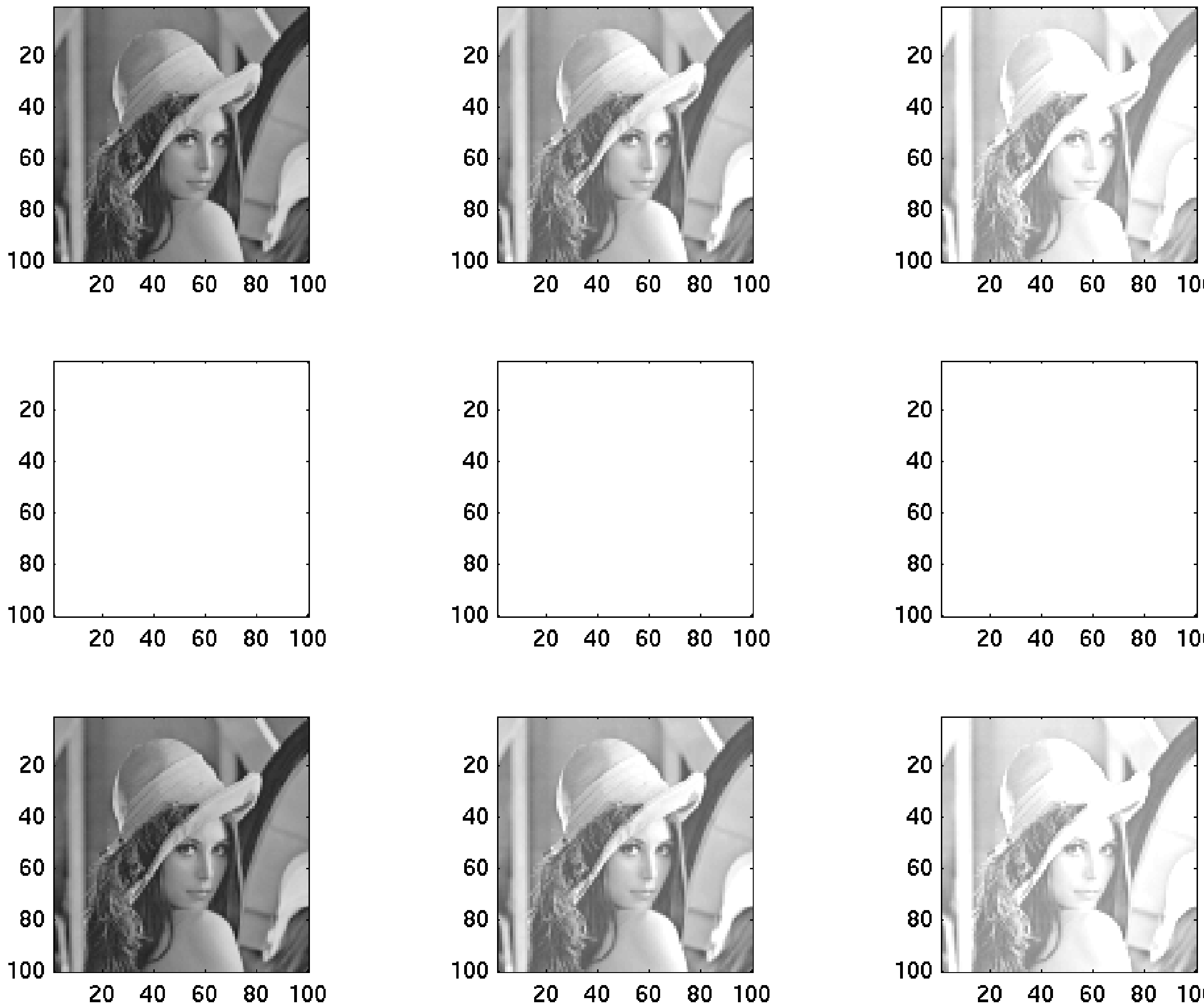
$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{k}{N} n}, \quad k = 0 \ldots N-1$$

Teď konečně 2D

- Korelace = určování podobnosti = projekce do bází …
- **To už jsme jednou viděli, je to furt to samý … ANO, JE !**

$$c = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] a[k,l]$$

Analyzační signál = d.c. $a[k,l] = 1$

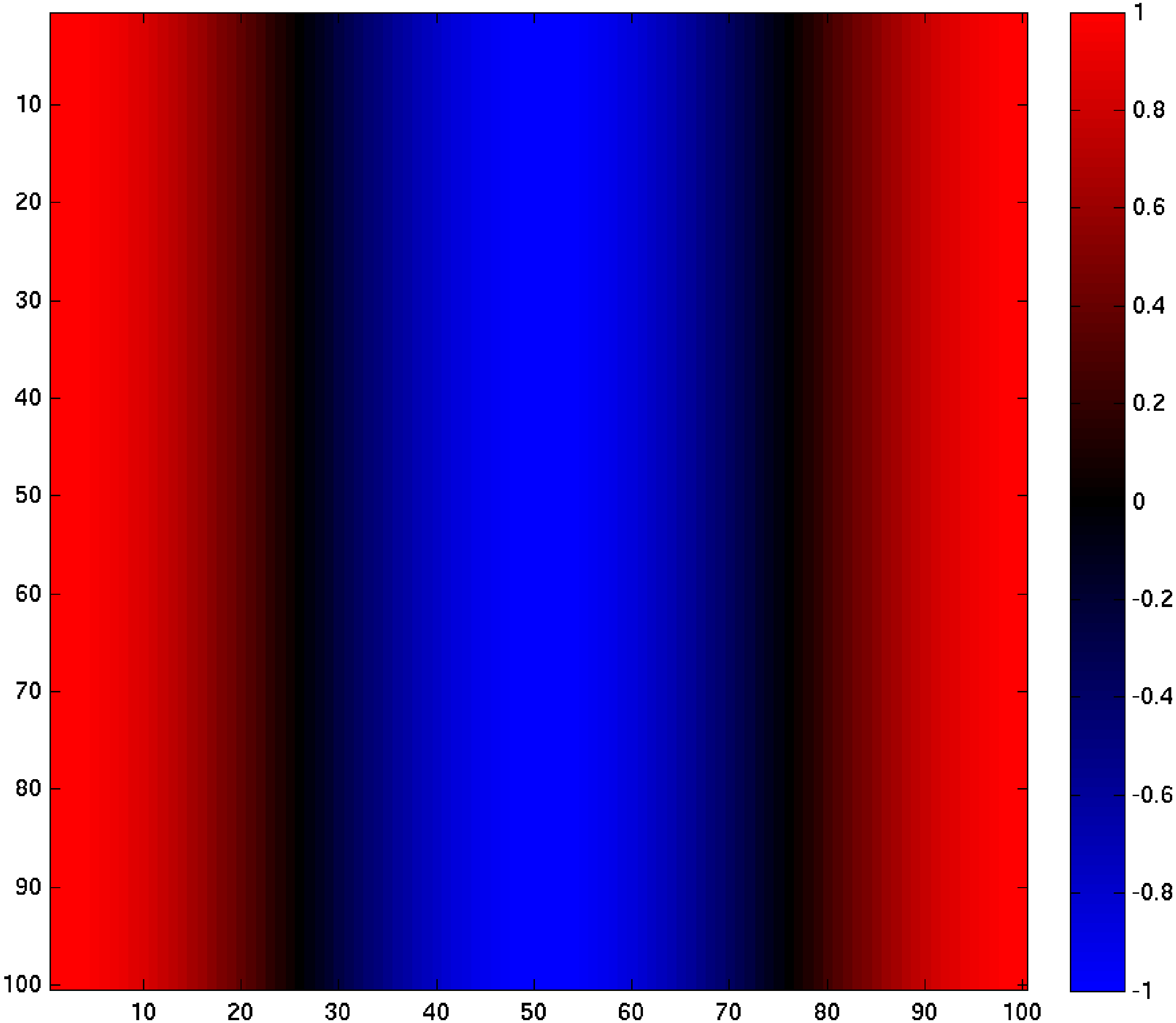


c=4.8287e+03 c=6.8287e+03 c=8.8287e+03

Analyzační signál – horizontální cosinusovka

$$a[k,l] = \cos(2\pi \frac{1}{100} l)$$

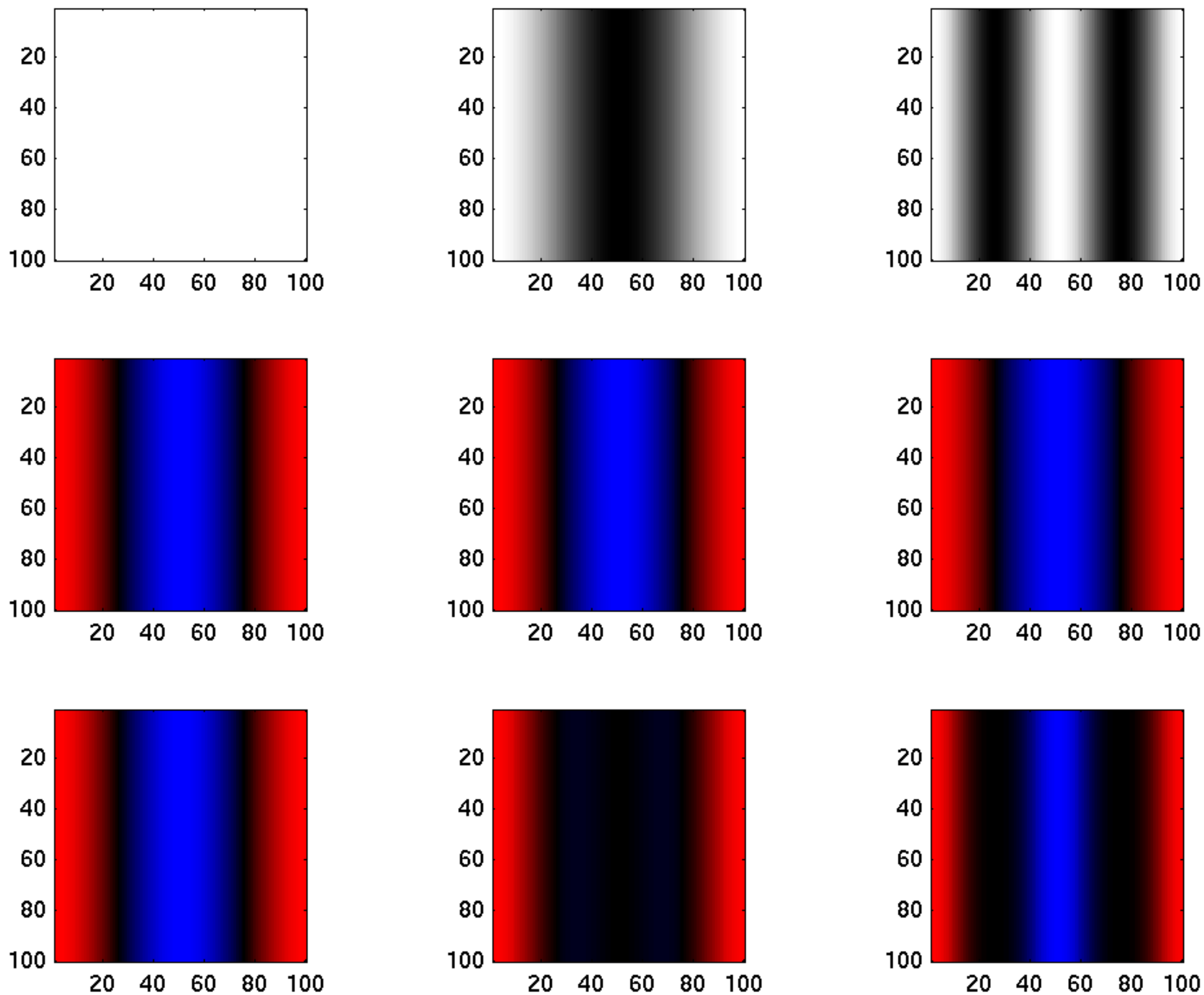
- Mění se pouze v jednom směru.
- Budeme potřebovat zobrazovat i záporné hodnoty



Geeks: jak se toto dělá v Matlabu ??

Analýza pro

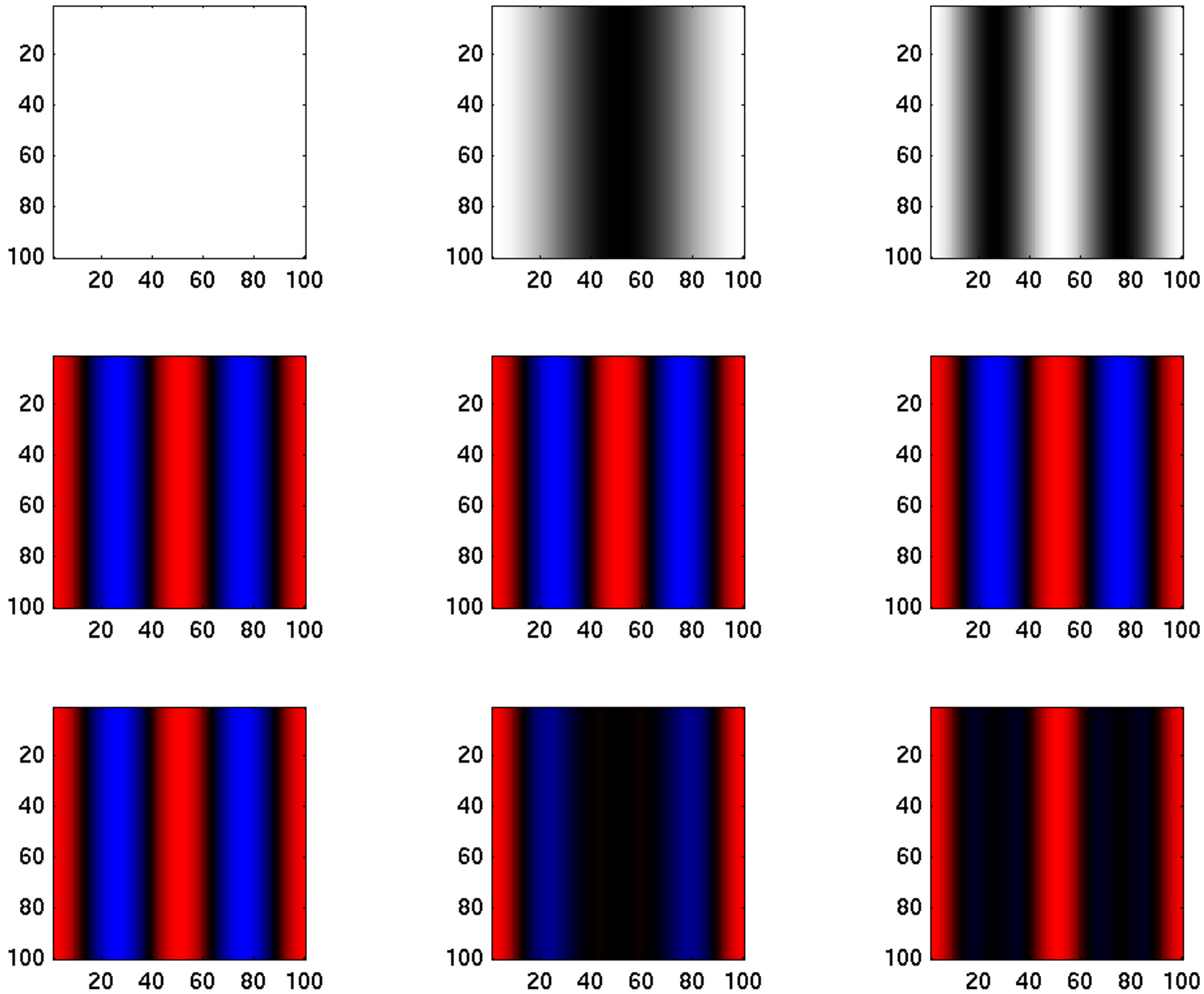
$$a[k,l] = \cos(2\pi \frac{1}{100} l)$$



c=4.5475e-13 c=2500 c=3.4106e-13

2 x rychlejší horizontální cos

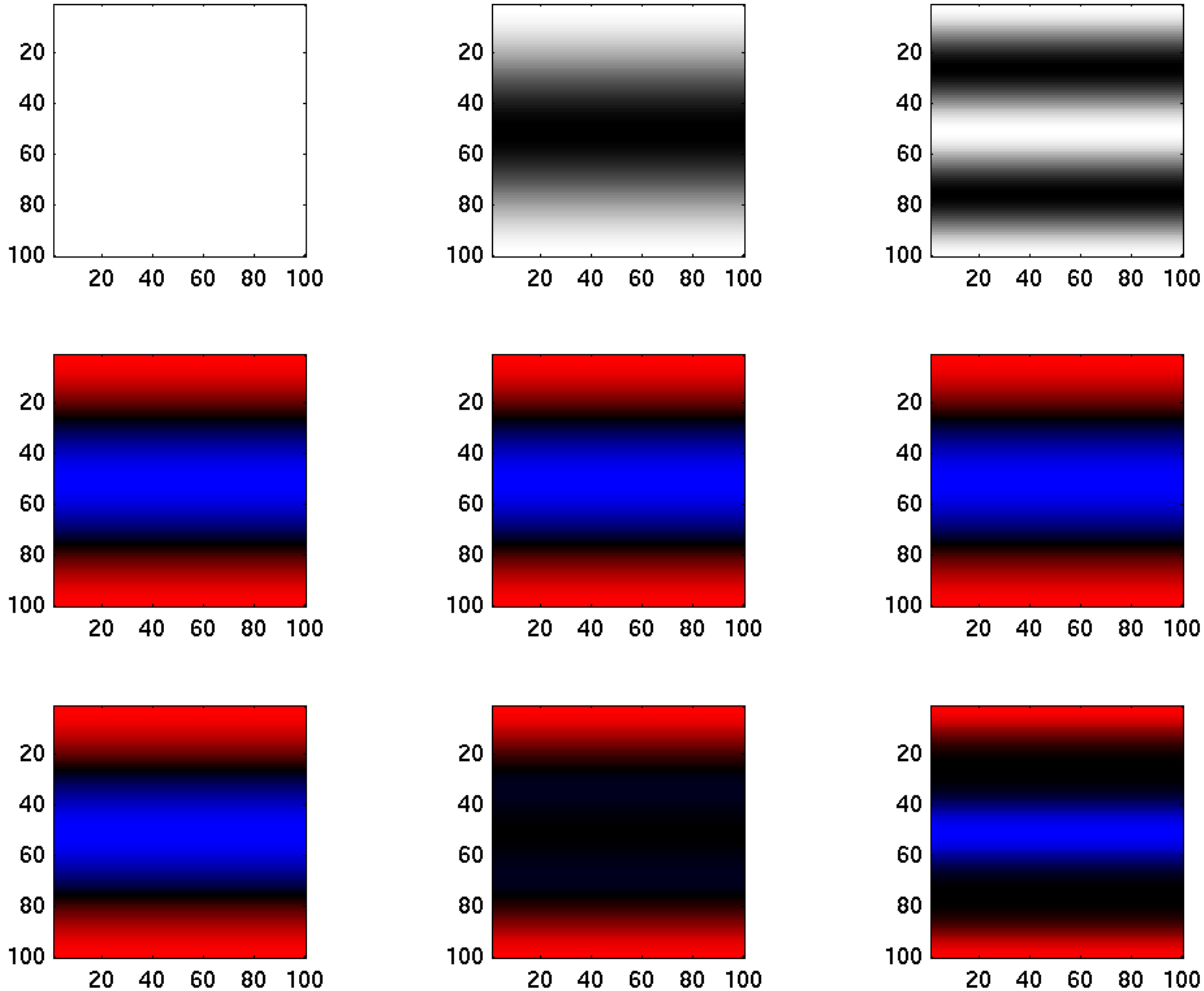
$$a[k,l] = \cos(2\pi \frac{2}{100} l)$$



c=2.9843e-13 c=1.5064e-12 c=2.5000e+03

Vertikální cos

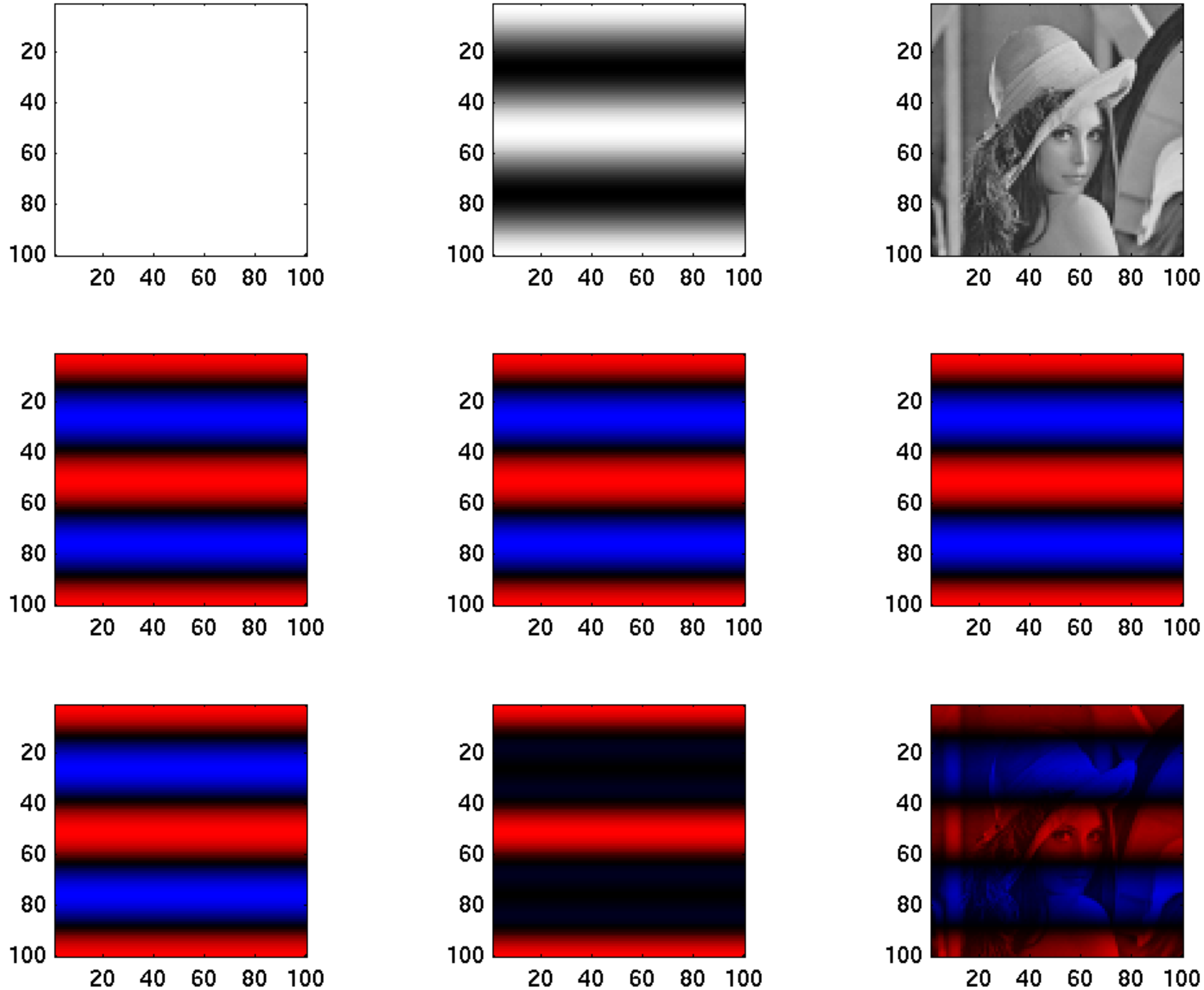
$$a[k,l] = \cos(2\pi \frac{1}{100} k)$$



c= -7.7716e-14 c=2.5000e+03 c=6.9944e-13

2 x rychlejší vertikální cos

$$a[k,l] = \cos(2\pi \frac{2}{100} k)$$

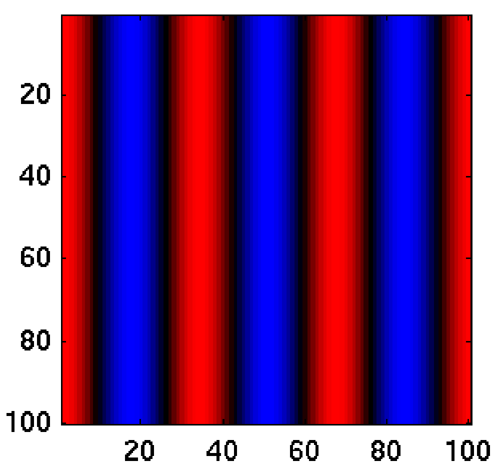


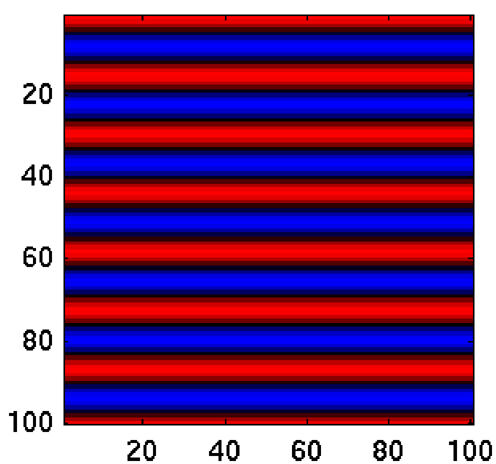
c= 4.9960e-13 c=2.5000e+03 c=-65.4694

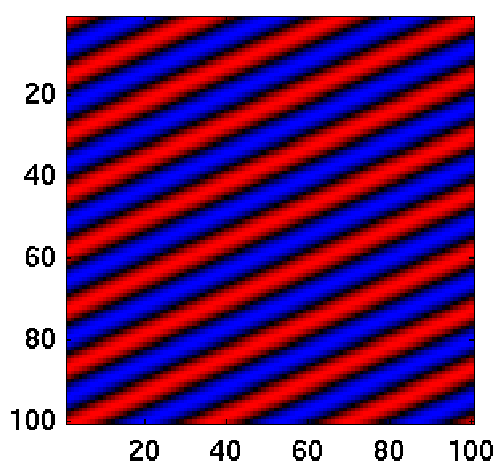
Mix obou směrů …

$$a[k,l] = \cos(2\pi \frac{3}{100} l) \qquad a[k,l] = \cos(2\pi \frac{7}{100} k)$$

$$a[k,l] = \cos\left[2\pi(\frac{7}{100}k + \frac{3}{100}l)\right]$$







Zobecnění

$$X[m,n] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] \cos\left[2\pi(\frac{m}{K}k + \frac{n}{L}l)\right]$$

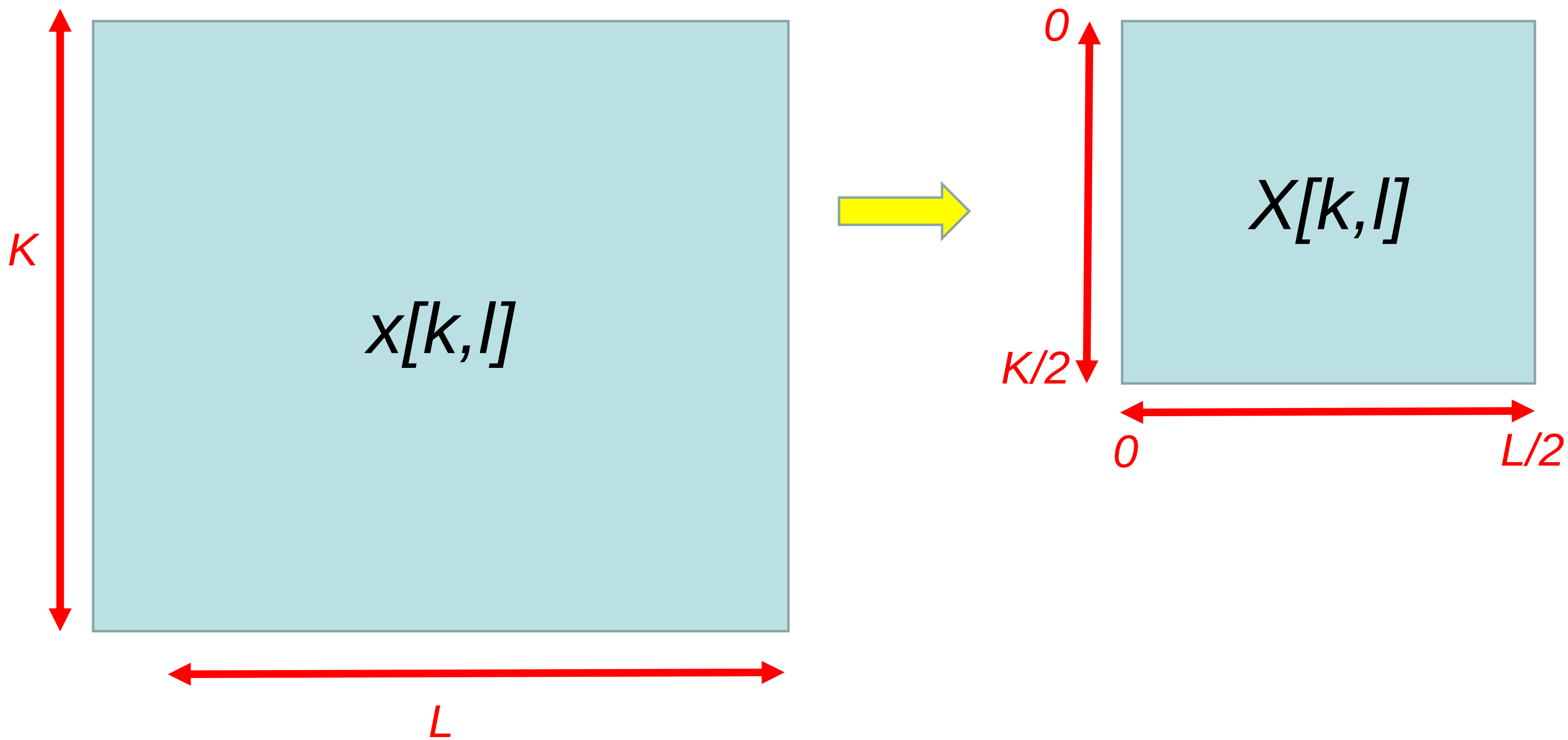
- *m/K* – svislá frekvence
- *n/L* – vodorovná frekvence

Rozumné rozsahy frekvencí ?

- *m/K* a *n/L*

  0 … ½ => OK

  > ½ => hm…



Ještě o frekvencích …

- 1D

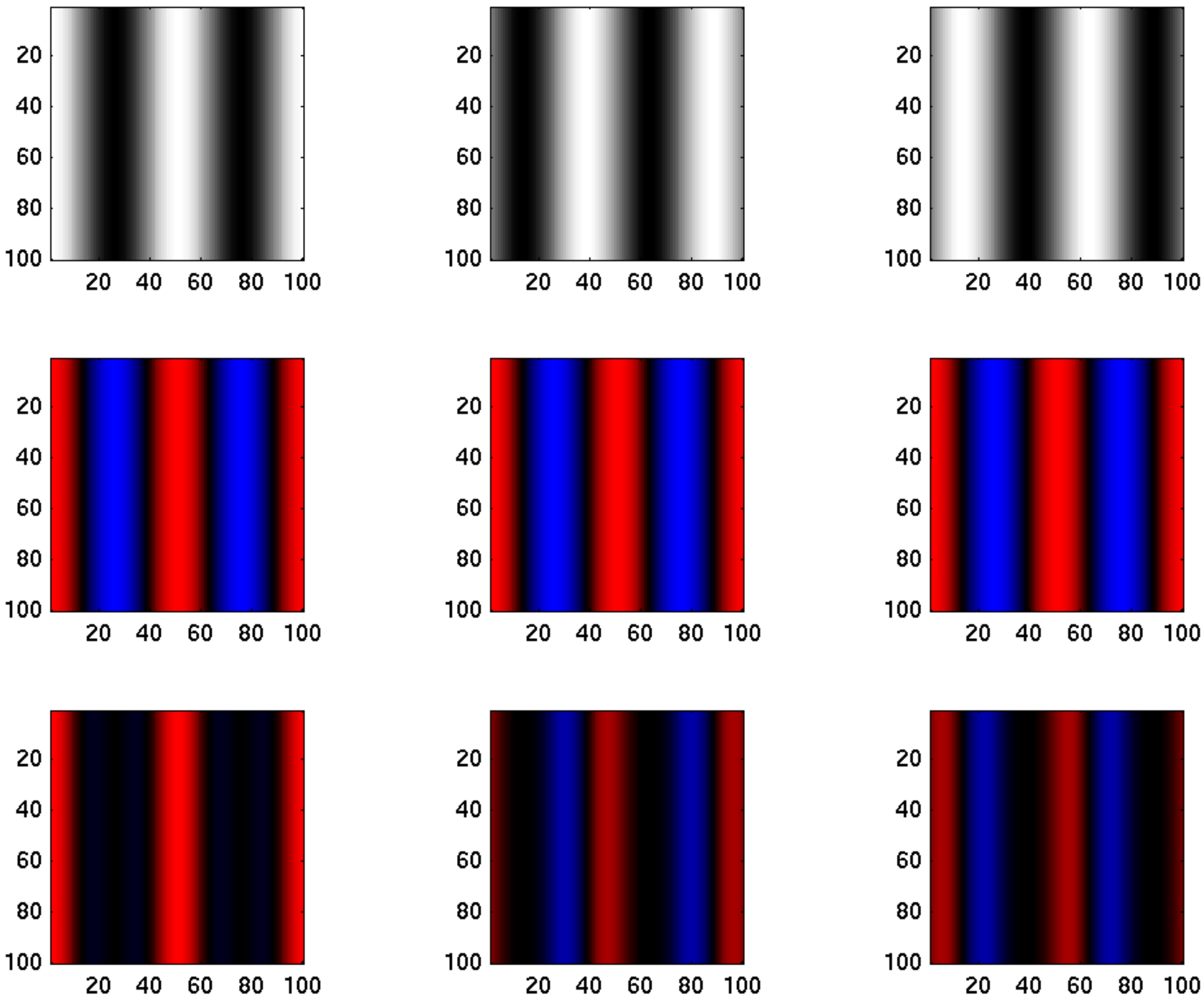
$$Hz = \frac{1}{s} \qquad F_s = \frac{\#samples}{s} \qquad f_{norm} = \frac{f_{skut}}{F_s} \qquad f_{skut} = \frac{k}{N} F_s$$

- 2D

$$dpi = \frac{1}{inch} \qquad F_s = \frac{\#pixels}{inch} \qquad f_{norm} = \frac{f_{skut}}{F_s}$$

$$f_{skut,vert} = \frac{m}{K} F_s, \quad f_{skut,horiz} = \frac{n}{L} F_s$$

Fáze – opět problém ☹



c=2.5000e+03 c=9.9476e-13 c=-5.6843e-14

Řešení – komplexní exponenciály (opět)

$$X[m,n] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j\left[2\pi\left(\frac{m}{K}k + \frac{n}{L}l\right)\right]}$$

Opakování, co je co …

- *k,l* jsou počitadla pixelů (vstup)
- *m,n* jsou počitadla frekvencí (výsledek)
- *m/K* je normalizovaná vertikální frekvence
- *n/L* je normalizovaná horizontální frekvence

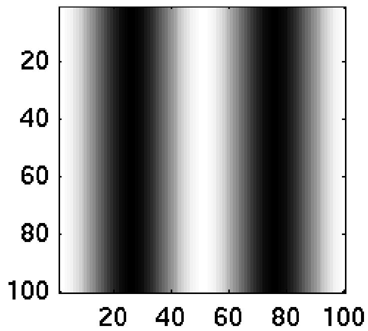
Jak si to představit I.

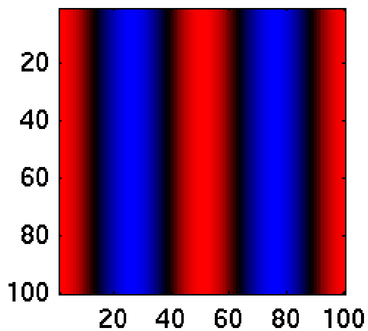
$$X[0,1] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j\left[2\pi\left(\frac{0}{K}k + \frac{1}{L}l\right)\right]} = \ldots$$

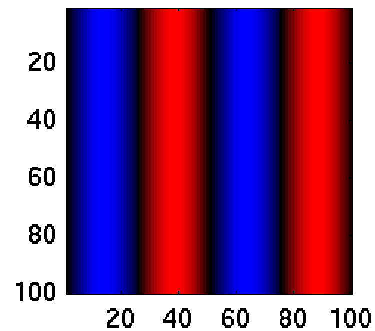
Jak si to představit II.

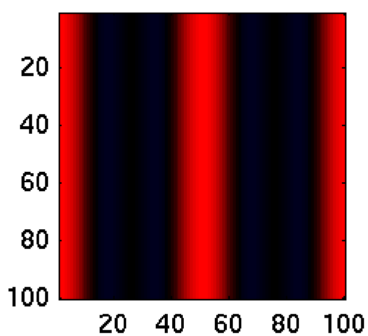
$$X[3,0] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x[k,l] e^{-j\left[2\pi\left(\frac{3}{K}k + \frac{0}{L}l\right)\right]} = \dots$$

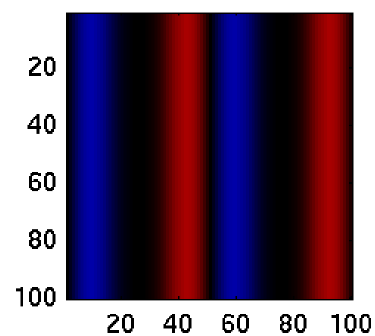
Korelace s *x[n]*









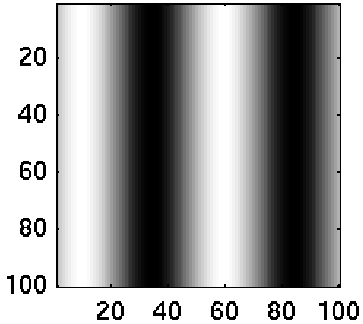


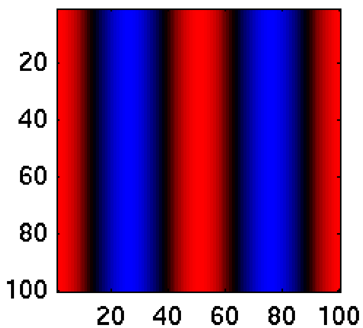
Re(c) = 2500

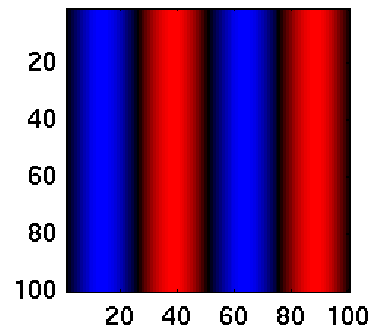
Im(c) = 0

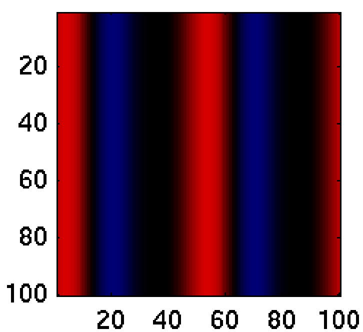
|c| = 2500

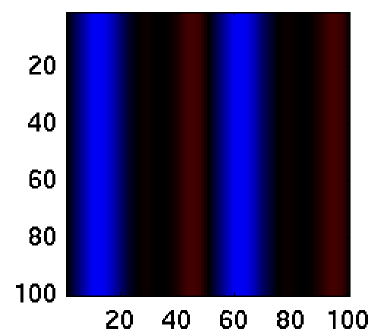
Korelace s *x[n]*









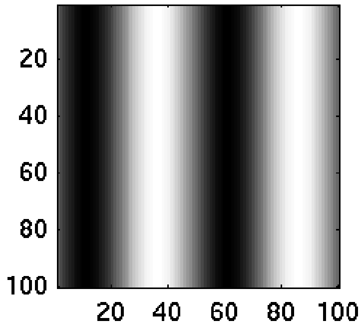


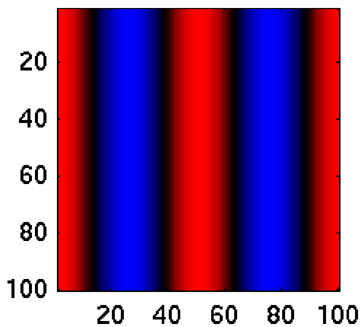
Re(c) = 1250

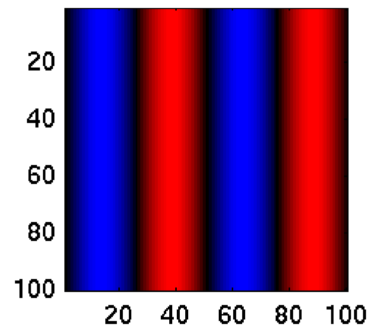
Im(c) = -2165

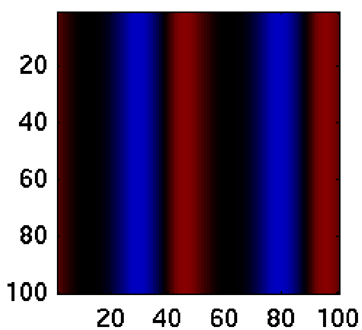
|c| = 2500

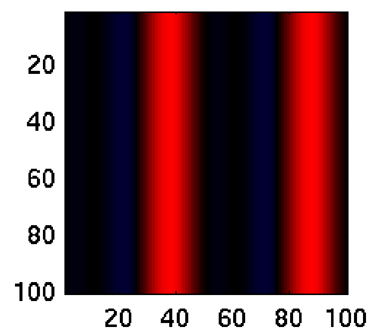
Korelace s *x[n]*









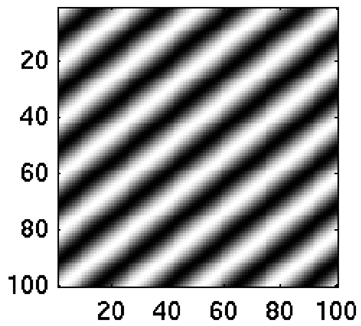


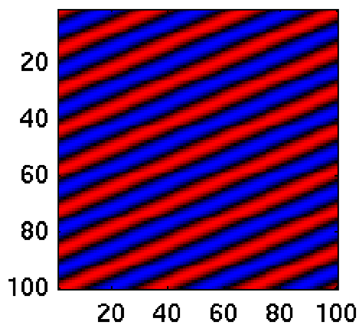
Re(c) = -772

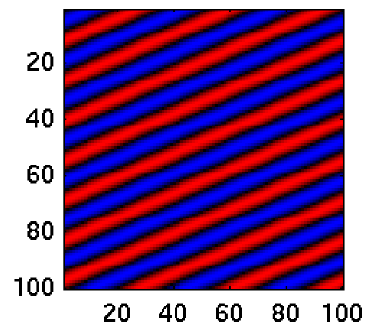
Im(c) = 2378

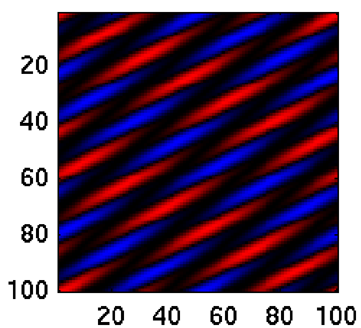
|c| = 2500

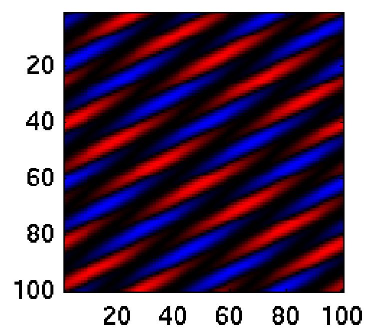
Změna v obou směrech…*X[7,3]*







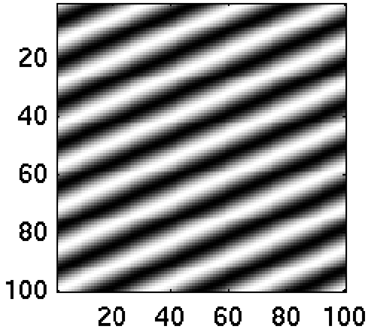


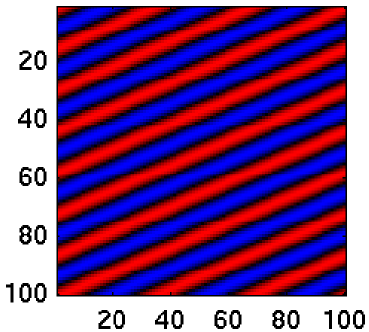


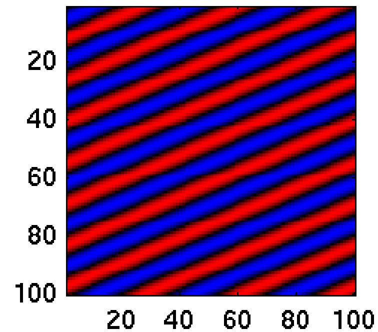
Re(c) = 0 Im(c) = 0

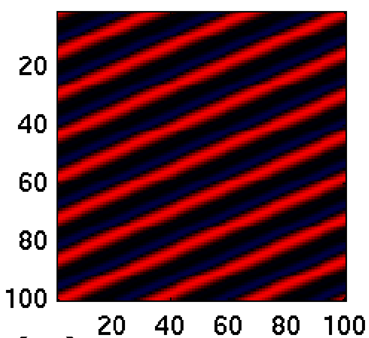
|c| = 0 … nepodobný

Změna v obou směrech…*X[7,3]*

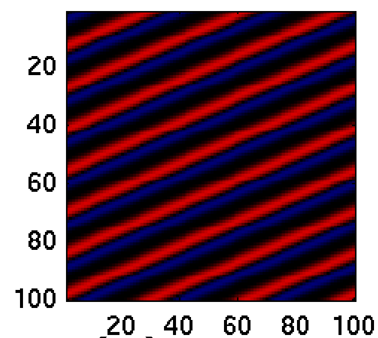








Re(c) = 2165



Im(c) = 1250

|c| = 2500 … podobný

2D DFT pomocí 2 x 1D DFT

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

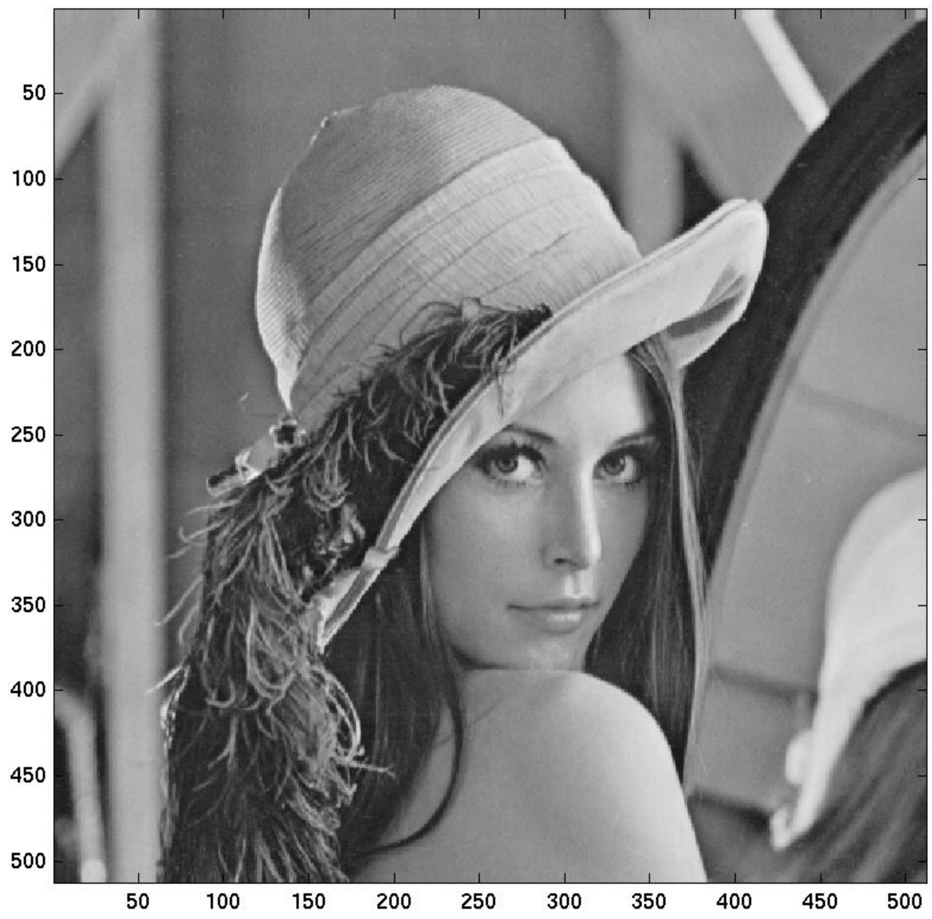
$$= \sum_{k=0}^{K-1} e^{-j2\pi\frac{mk}{M}} \sum_{l=0}^{L-1} x[k,l] e^{-j2\pi\frac{nl}{L}}, \quad \ldots \quad \text{nebo naopak}$$

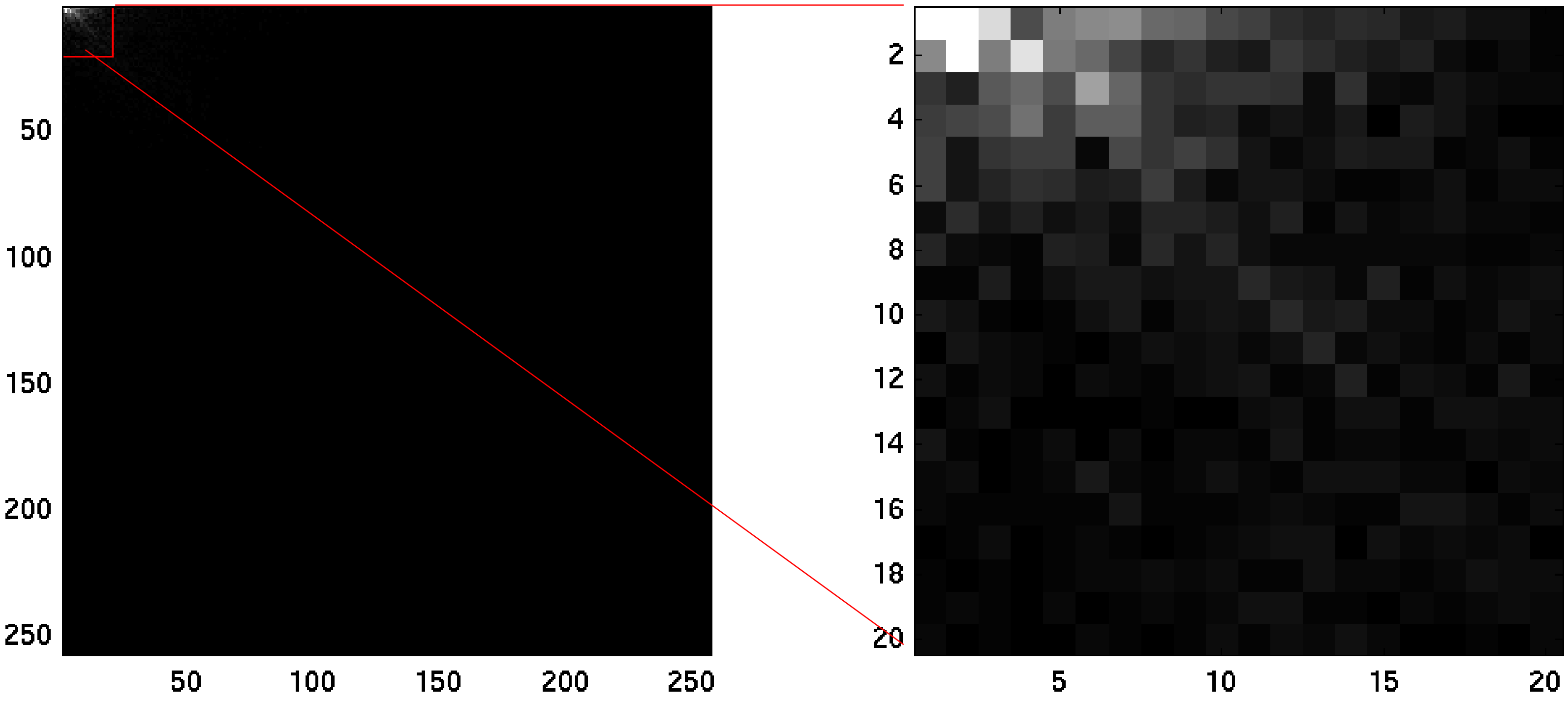
- takže

$$2DDFT\{x[k,l]\} =$$

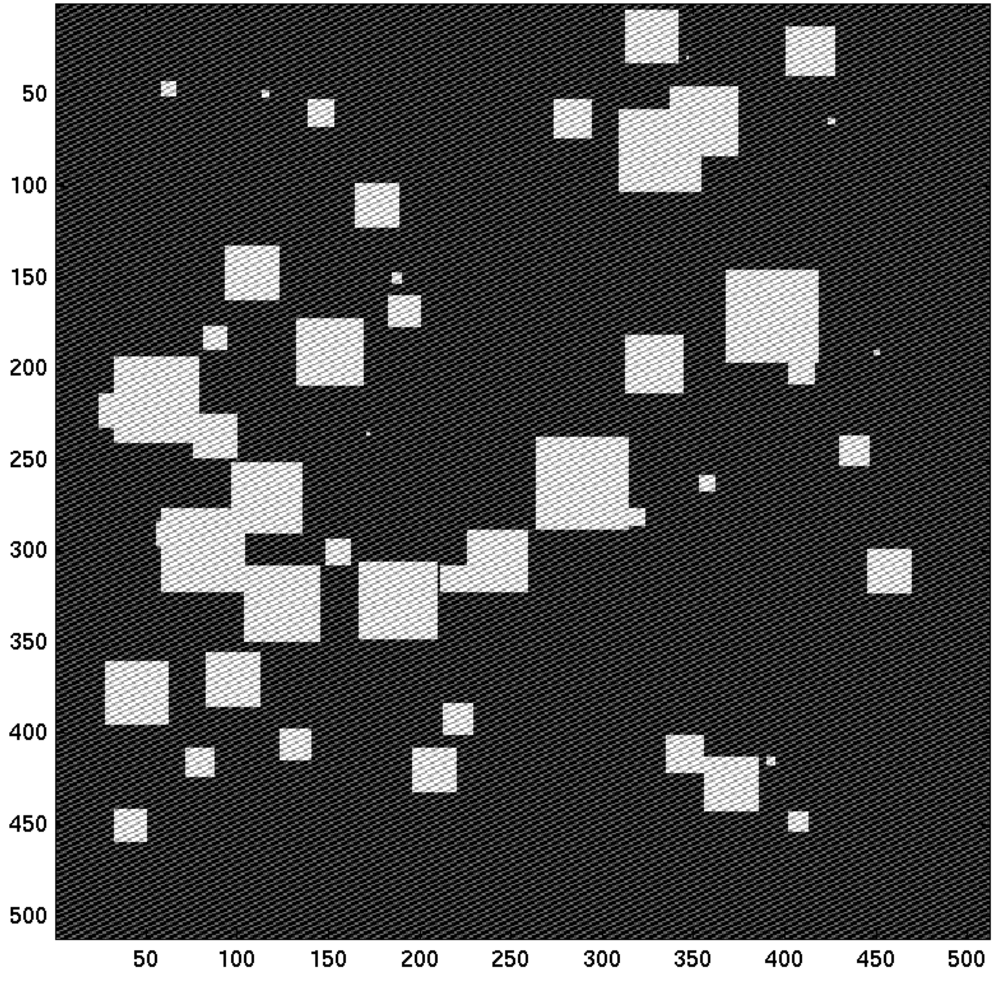
$$= 1DDFT_{sloupce}\{1DDFT_{radky} x[k,l]\} \quad \ldots \quad \text{nebo naopak}$$

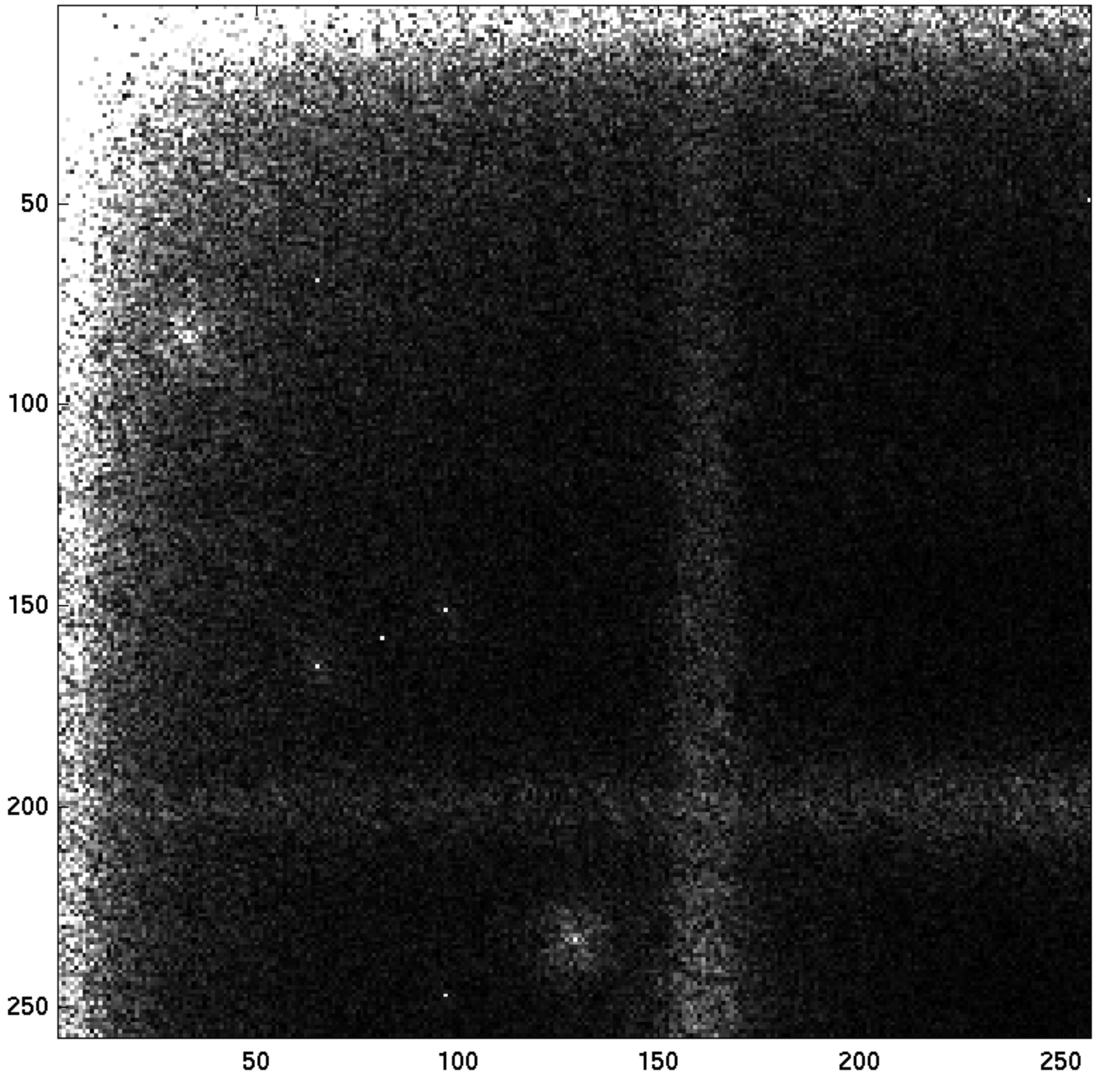
2D DFT pro skutečný signál



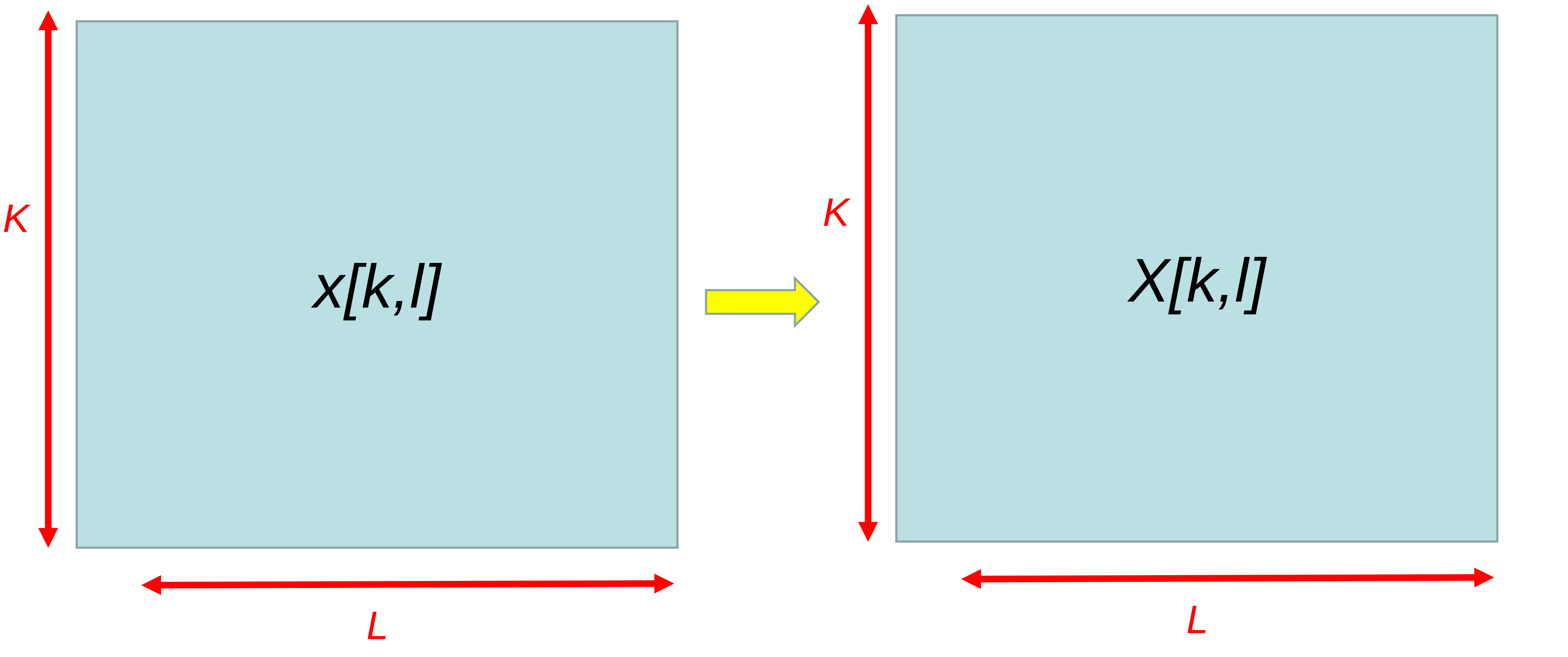

50
100
150
200
250
50
100
150
200
250
2
4
6
8
10
12
14
16
18
20
5
10
15
20

Něco s více vyššími frekvencemi …

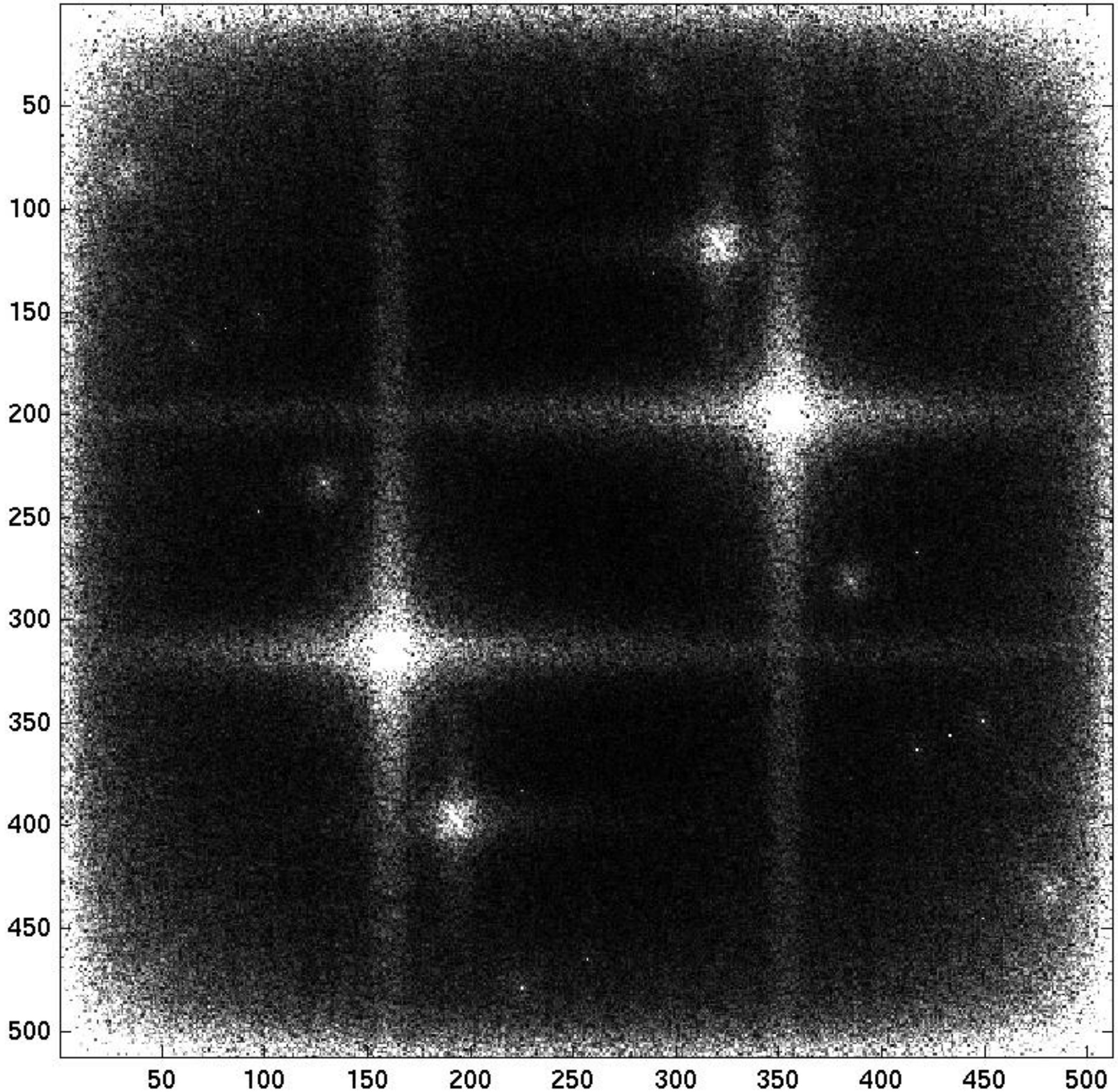



50
100
150
200
250
50
100
150
200
250

DFT produkuje *K x L* bodů !

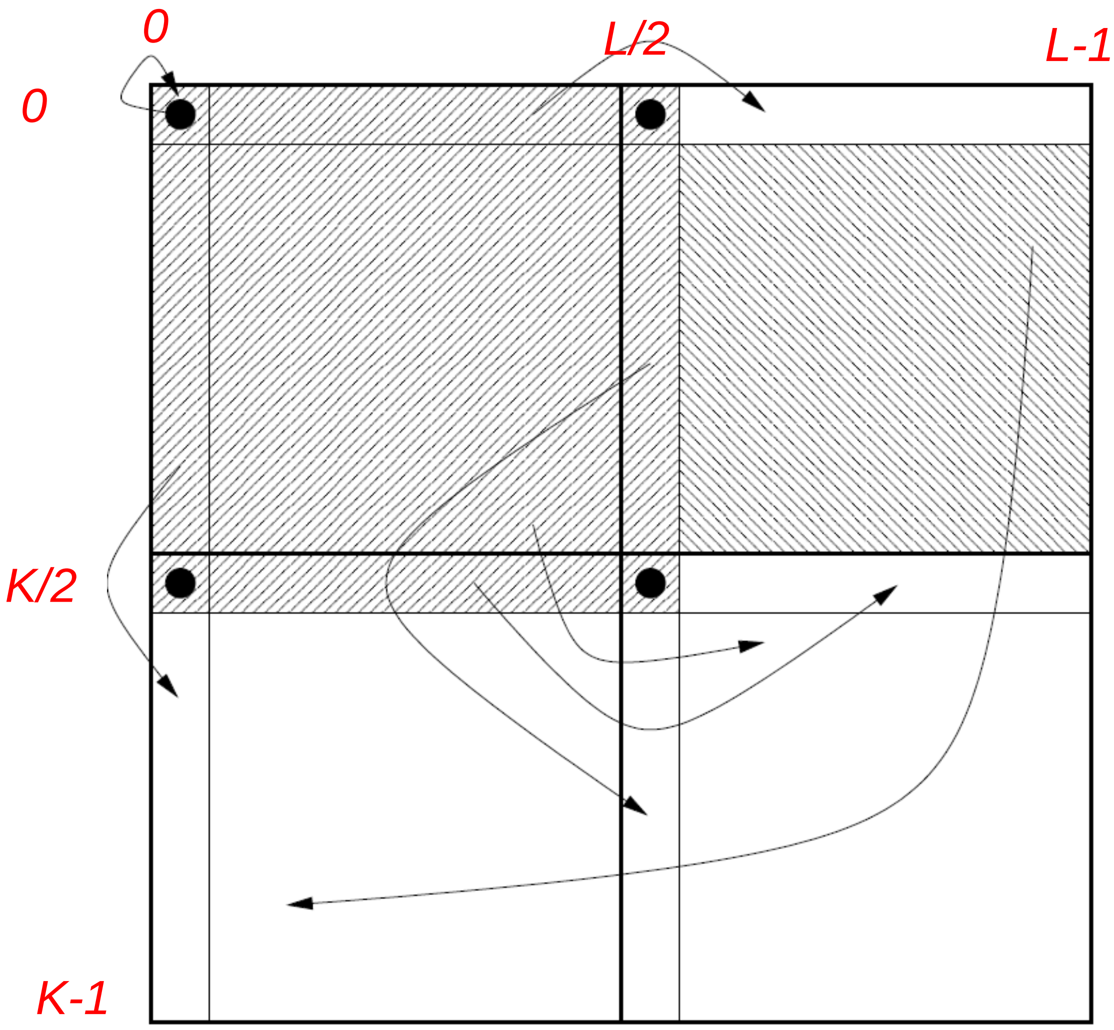


Celý výsledek *k=0…K-1, l=0…L-1*



Symetrie ?

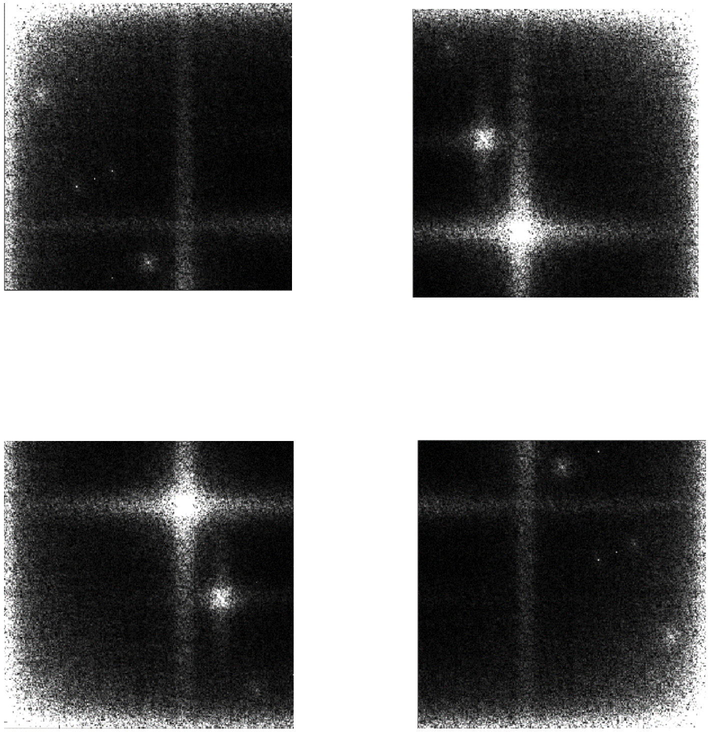
$$X[m,n] = X^{\star}[K-m, L-n]$$

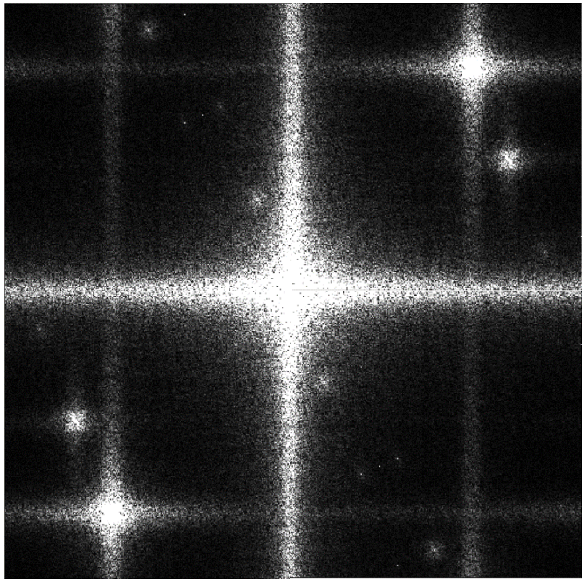


Zdroj: http://www-personal.umich.edu/~hyunjinp/notes/n-dft.pdf

Přeskládání …

- Nízké frekvence uprostřed

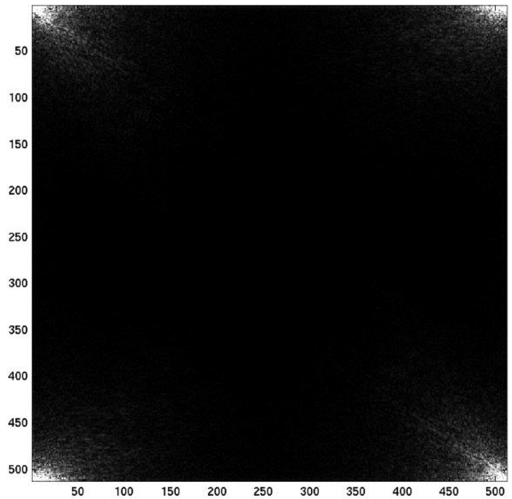




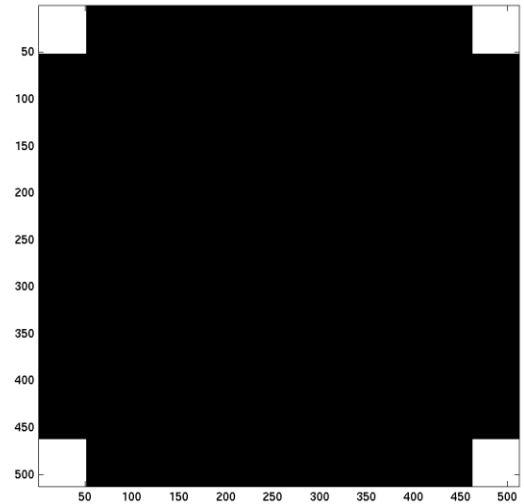
Zpětná 2D DFT

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

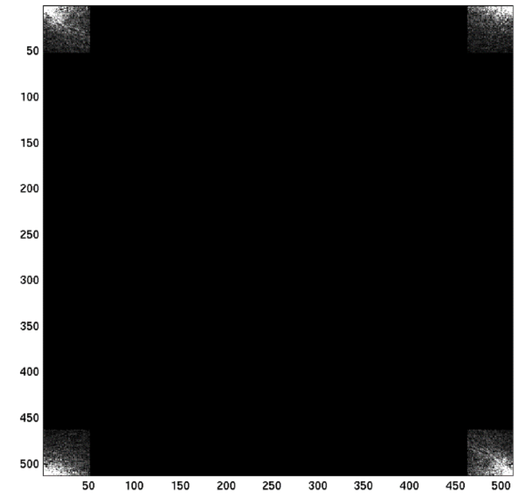
Hraní s frekvencemi …

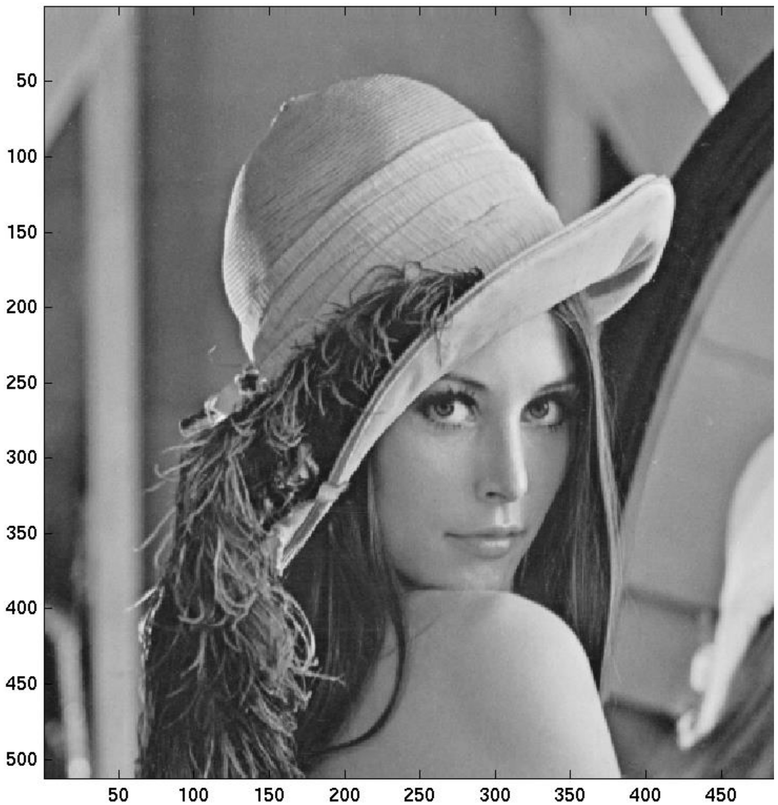


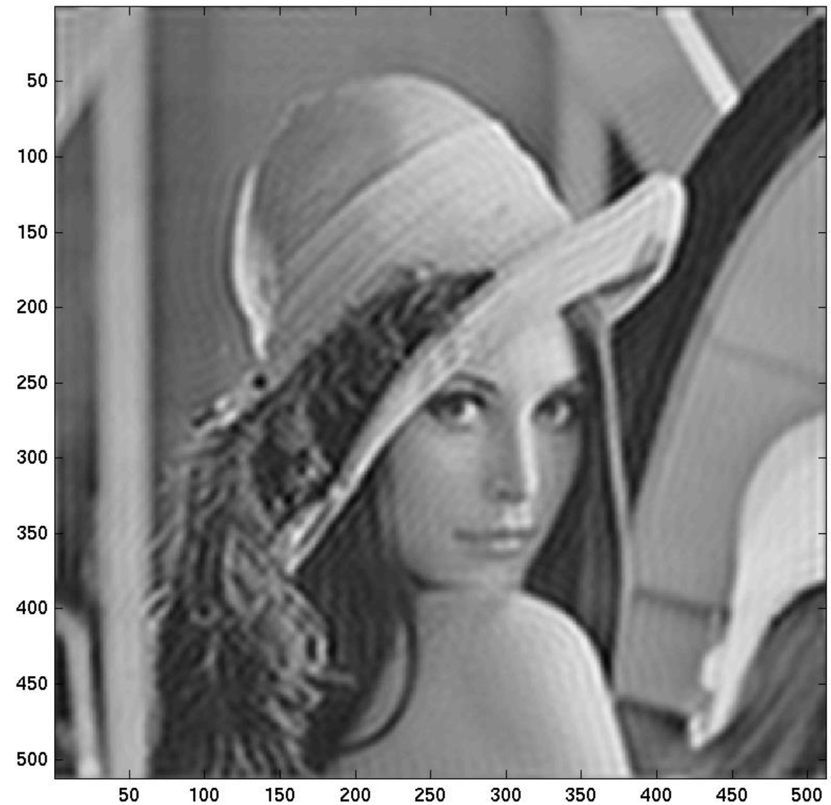
×



=




50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450


50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450 500

DCT

Proč ?

- Komplexní koeficienty jsou otravné
- Symetrie v obrázku jsou otravné
- Pro obrázek $K \times L$ chceme $K \times L$ reálných čísel.

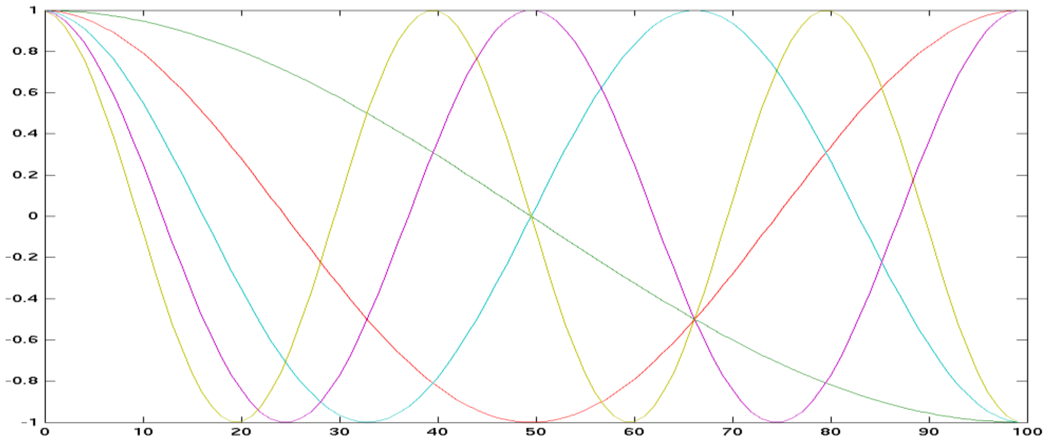
1D DCT

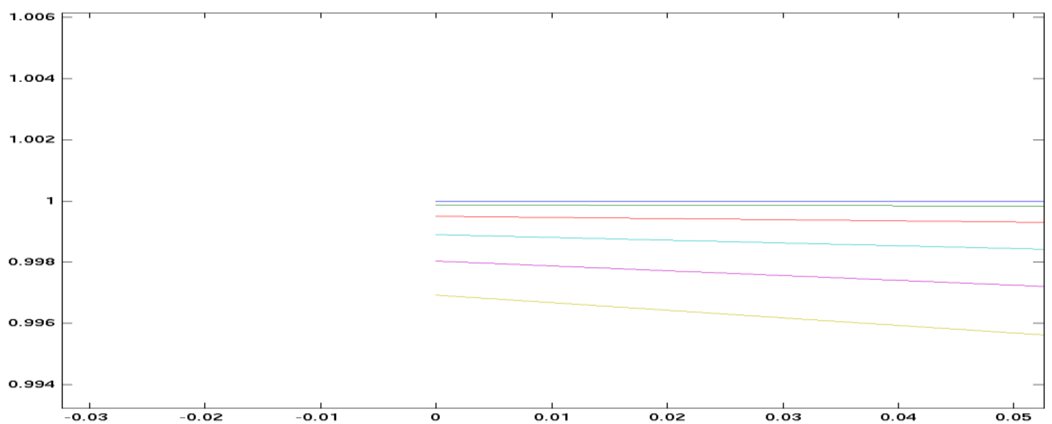
$$X[k] = \sum_{n=0}^{N-1} x[n] \cos\left[\frac{\pi}{n}(n+\frac{1}{2})k\right]$$

+ normalisace koeficientů …

$$X[0] \ldots \times \sqrt{\frac{1}{N}}$$

$$X[1 \ldots N-1] \ldots \times \sqrt{\frac{2}{N}}$$

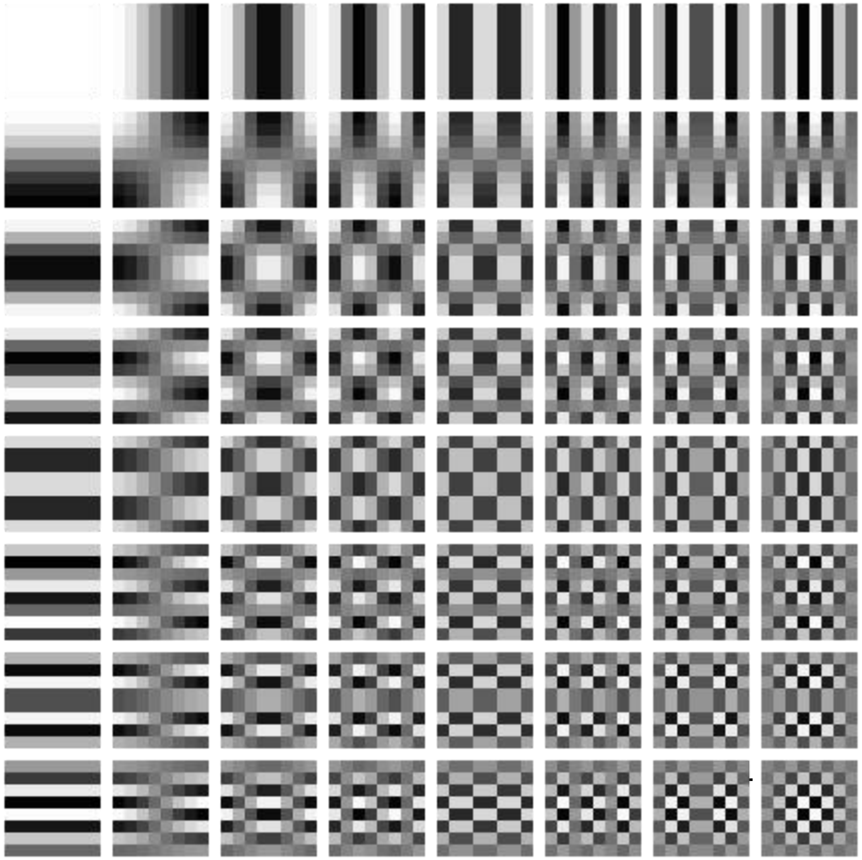




2D DCT báze

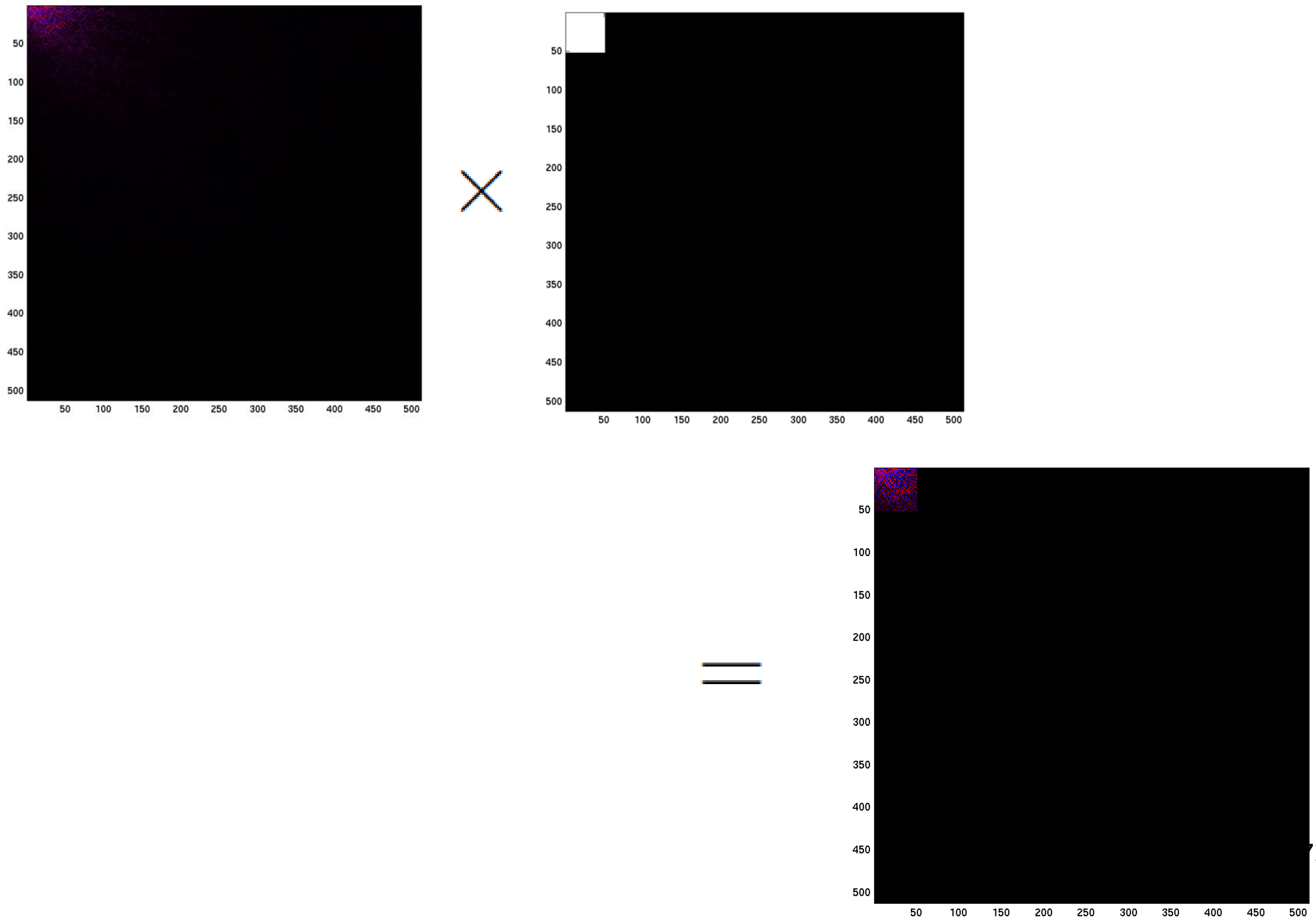
$$X_{k_1,k_2} = \sum_{n_1=0}^{N_1-1} \left(\sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_2}\left(n_2+\frac{1}{2}\right)k_2\right]\right) \cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right]$$

$$= \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right] \cos\left[\frac{\pi}{N_2}\left(n_2+\frac{1}{2}\right)k_2\right].$$

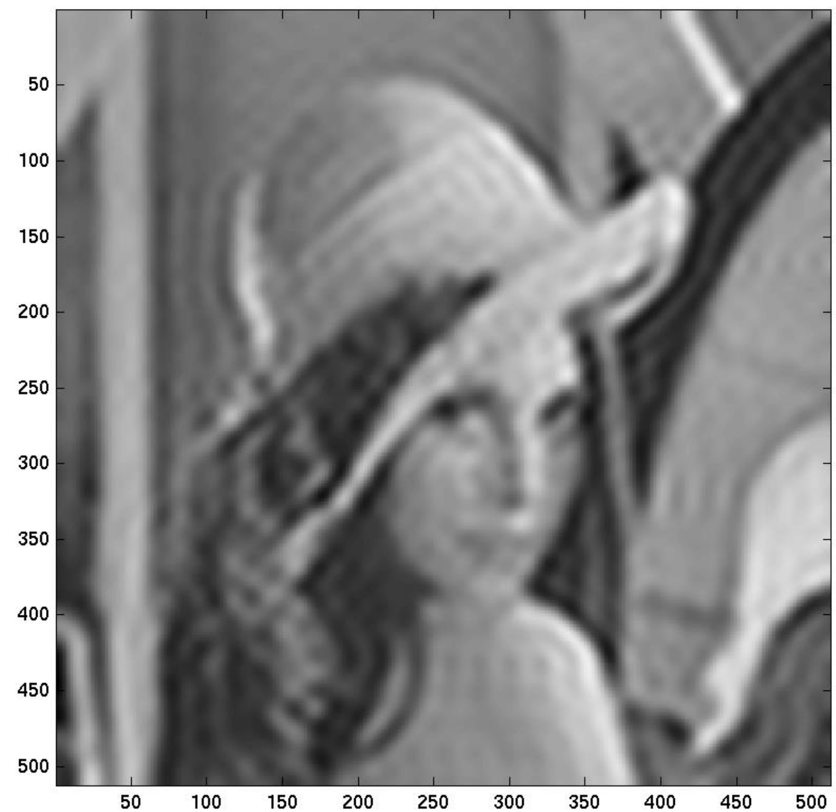


Zdroj:
https://en.wikipedia.org/wiki/Discrete_cosine_transform#Multidimensional_DCTs

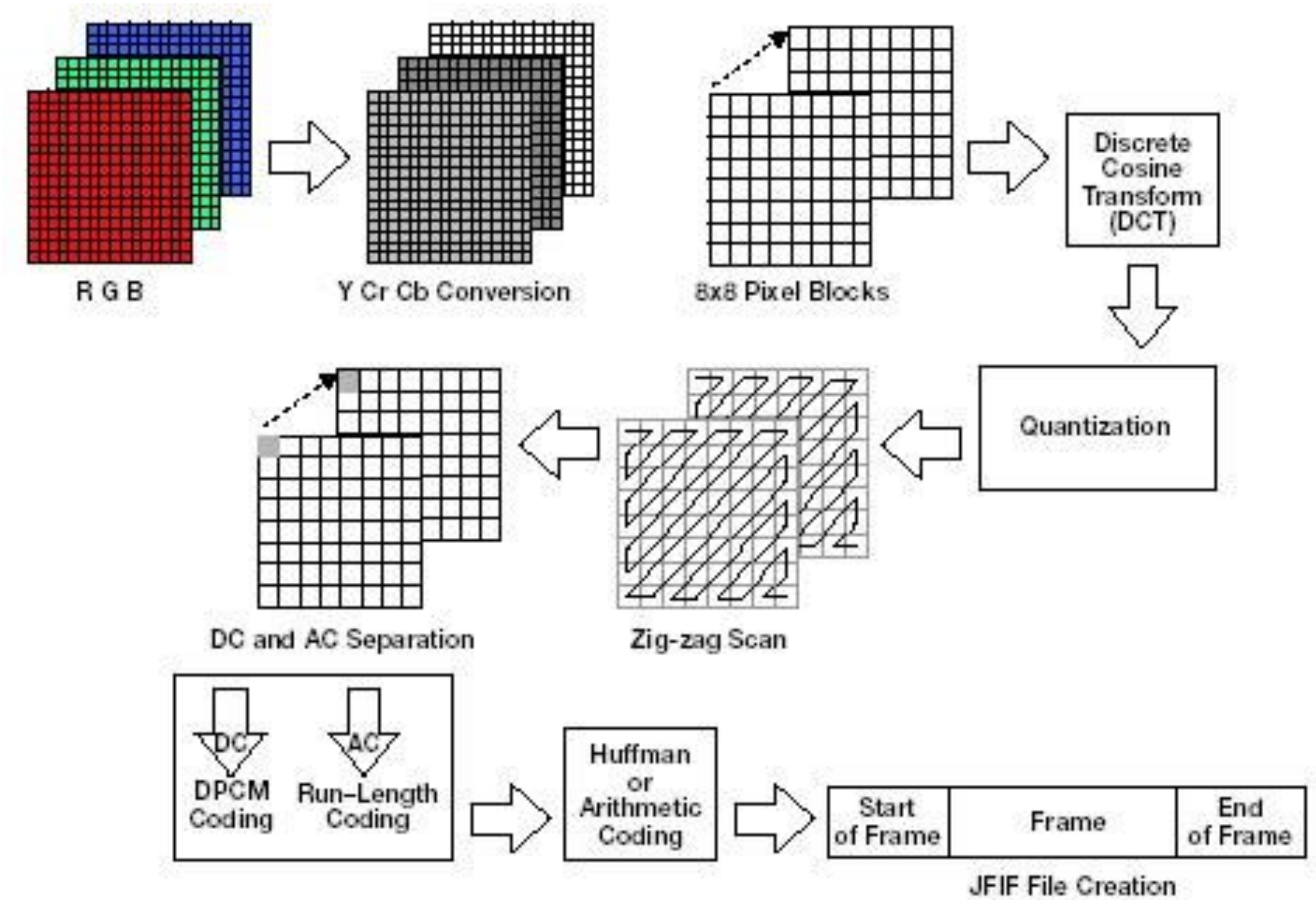
Filtrování v DCT




50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450 500


50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450 500

JPEG



- Zdroj: http://www.eetimes.com/document.asp?doc_id=1225736
- Více v grafických předmětech a 2. cvičení.

SUMMARY

- Obraz je 2D signál
- Filtrace
  - Konvoluce, analogie s 1D FIR filtry
  - IIR v obrázcích nepotkáte
- Jezdím s maskou po obrázku, co pod ni padne, vynásobím s koeficienty a sčítám.

SUMMARY II.

- Frekvenční analýza 2D
  - Stejně jako u 1D – projekce do bází
  - Obrazové frekvence mají svůj význam.
  - Cos báze jsou dobré cvičení, ale nestačí
- 2D DFT
  - Projekce/podobnost/korelace s komplexními exponenciálami
  - Koeficienty jsou komplexní, modul a argument
  - Ve výsledné DFT matici je spousta symetrií
  - Ve spektru se dá filtrovat.

SUMMARY III.

- DCT
  - 2x „pomalejší“ báze než DFT
  - Trochu složitější definice
  - Produkuje reálné koeficienty, nízké frekvence (jen!) na začátku
  - Využití v JPEG

TO BE DONE

- Určení frekvenční charakteristiky 2D filtru
- Jak je to přesně se symetriemi 2D DFT ?
- Jak je přesně definována 1D a 2D DCT a proč je tam ten půlvzorkový posuv ?
- Jak je to s barvami (barevné modely, atd).
- Jak funguje rozpoznávání xichtů na FB ?
  – Hint: https://research.facebook.com/researchers/684639631606527/yann-lecun/

The END