5. Geometriai transzformációk

Kató Zoltán

Képfeldolgozás és Számítógépes Grafika tanszék
SZTE
(http://www.inf.u-szeged.hu/~kato/teaching/)

Kép transzformációk típusai

- **Kép értékkészletének (*radiometriai* információ) átalakítása:**

$$J(i,j) = f(I,(i,j))$$

- **Kép értelmezési tartományának *geometriai* transzformációja (warping):**

$$J(i,j) = I(t_i(i,j), t_j(i,j))$$

- **Mind az értékkészlet mind pedig az értelmezési tartomány átalakítása:**

$$J(i,j) = f(I,(t_i(i,j), t_j(i,j)))$$

Egyszerű geometriai transzformációk

- **Nagyítás / kicsinyítés**
  - Izotróp
  - Anizotróp

- **Forgatás**

Forgatás

- **RxC kép forgatása 0° < θ ≤ 90° szöggel a középpont körül**

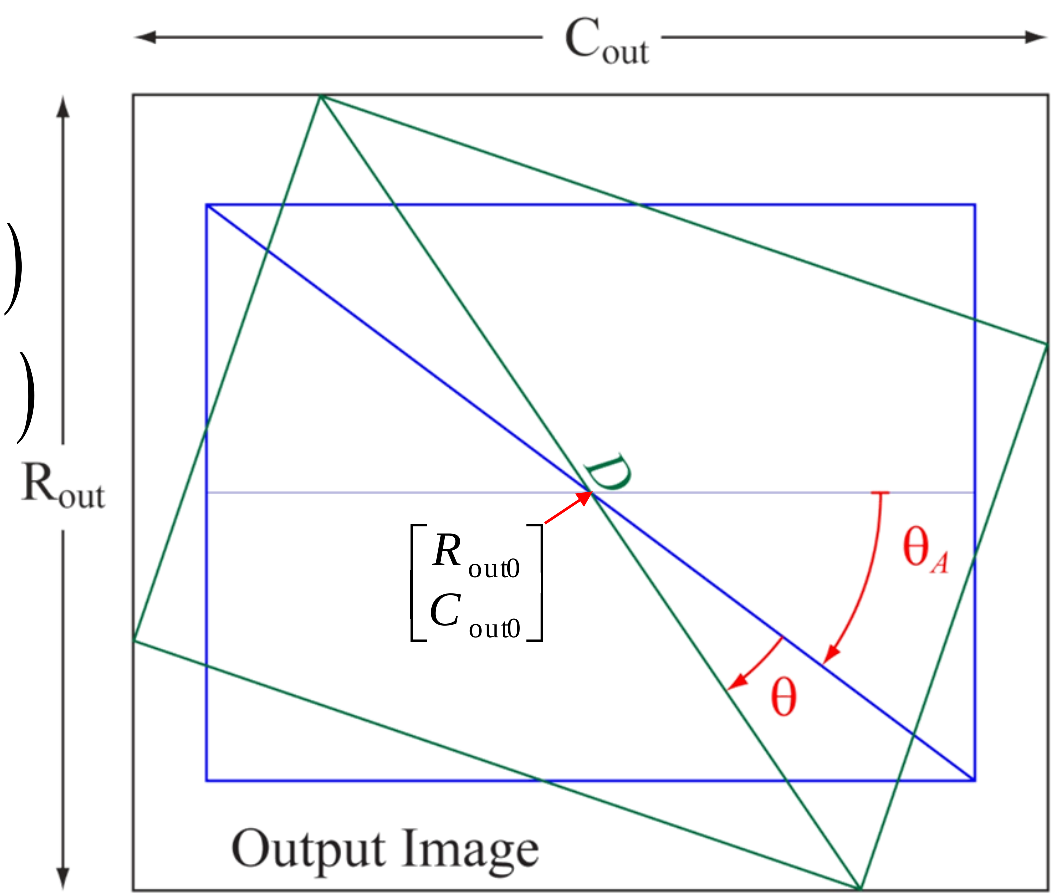
  - **Új képméret:**

$$R_{out} = \mathrm{round}\left(\left|D\sin(\theta+\theta_A)\right|\right)$$

$$C_{out} = \mathrm{round}\left(\left|D\cos(\theta-\theta_A)\right|\right)$$

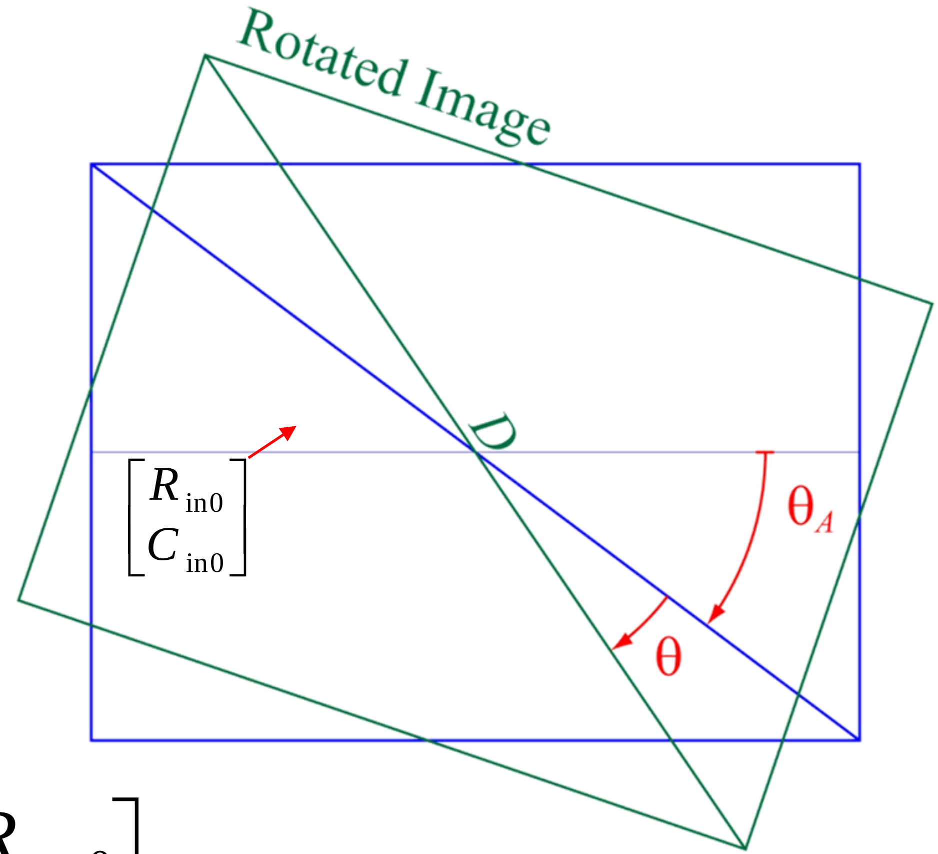
  - **Új középpont:**

$$(R_{out0}, C_{out0}) = \left(\left\lfloor \tfrac{1}{2} R_{out} \right\rfloor + 1, \left\lfloor \tfrac{1}{2} C_{out} \right\rfloor + 1\right).$$



Forgatás

- **RxC kép forgatása 0° < θ ≤ 90° szöggel a középpont körül**



- **Forgatási mátrix**

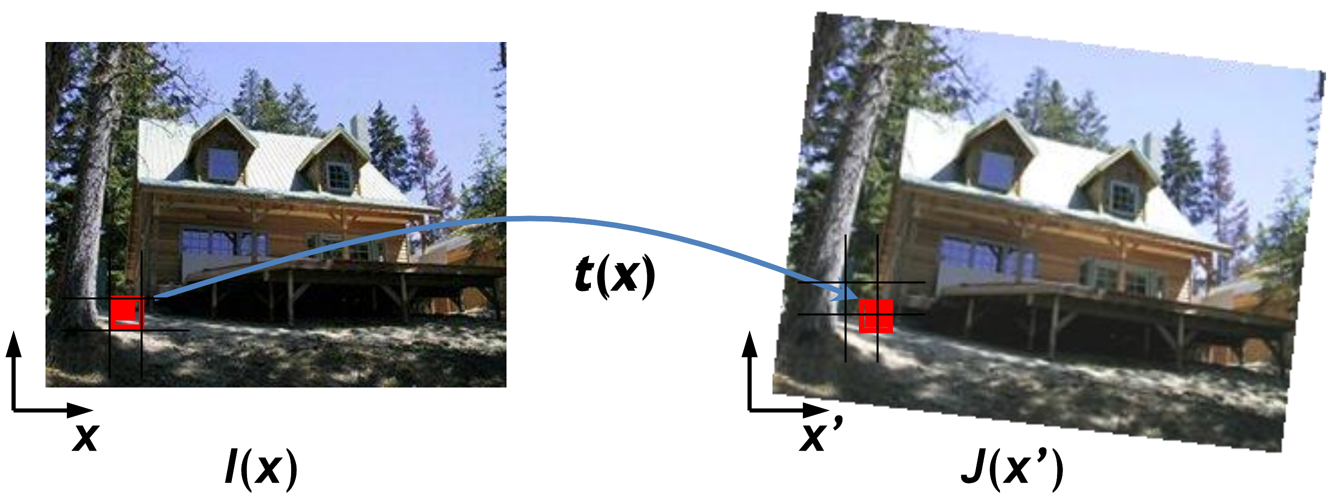
$$P(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

- **Output koordináták előállítása**

$$\begin{bmatrix} r \\ c \end{bmatrix} = P(\theta) \begin{bmatrix} r_{\mathrm{in}} - R_{\mathrm{in0}} \\ c_{\mathrm{in}} - C_{\mathrm{in0}} \end{bmatrix} + \begin{bmatrix} R_{\mathrm{out0}} \\ R_{\mathrm{out0}} \end{bmatrix}$$

Geometriai transzformációk

- Adott koordináta transzformáció $x' = t(x)$ és $I(x)$ kép esetén hogyan állítjuk elő a $J(x') = I(t(x))$ képet?
- $J(x')$ generálása tipikusan az inverz transzformáció alapján történik ➔ J minden pixele garantáltan kitöltésre kerül
- Mi történik, ha x' őse ($x=t^{-1}(x')$) I-ben nem egész pixel-koordinátára mutat?



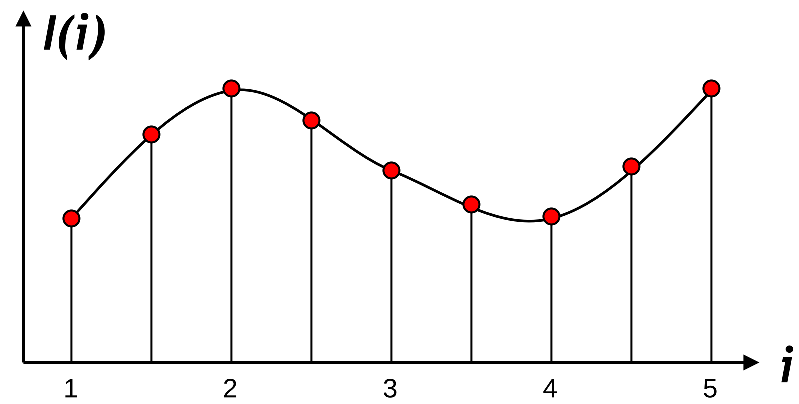
Szeliski

Újra mintavételezés

- **Emlékeztető: egy *I(i,j)* digitális kép az eredeti folytonos *f(x,y)* kép mintavételezésével és kvantálásával áll elő:**
  - ***I* egy folytonos függvény diszkrét pontszerű mintáiból áll elő.**

$$I(i,j) = \mathrm{quantize}(f(xd, yd))$$

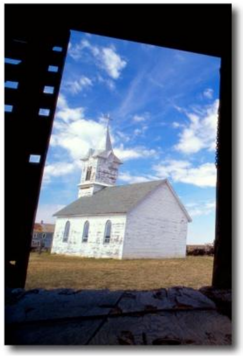
  - **Ha rekonstruálni tudnánk az eredeti *f* függvényt, akkor tetszőleg új felbontást elő tudnánk állítani**



- **Ha ismerjük egy *f* függvény értékeit diszkrét pontokban, akkor *f* közbülső értékeit <u>interpolációval</u> állíthatjuk elő**

Kató Zoltán: Digitális Képfeldolgozás (Tehetséggondozó program)

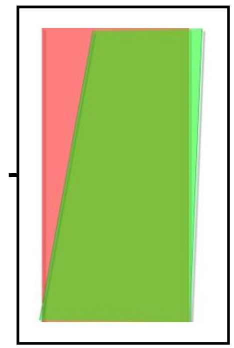
Geometriai transzformáció végrehajtása



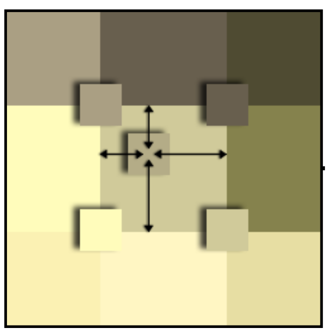
Input kép

↓

Transzformáló függvény



↓

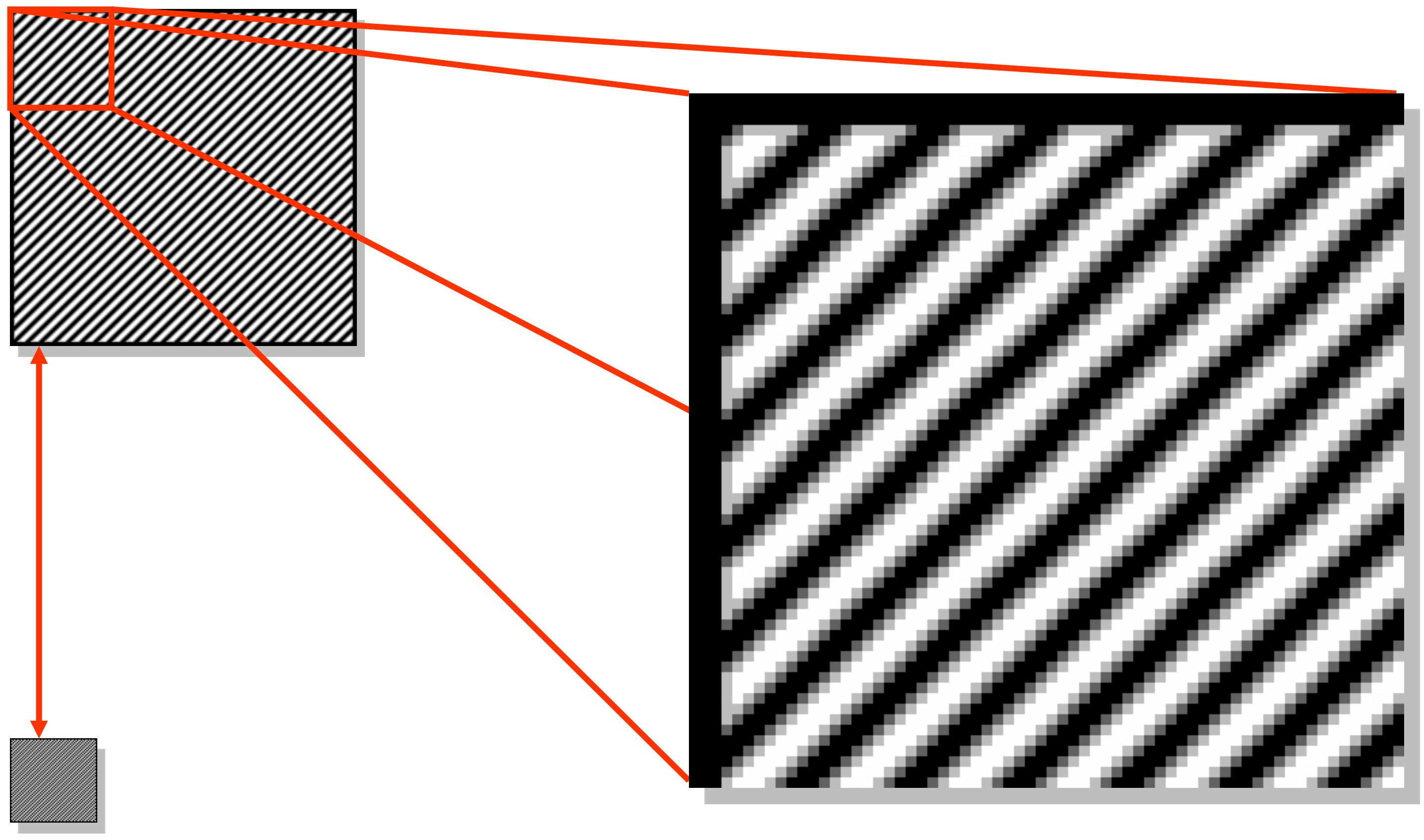


Interpolációs függvény

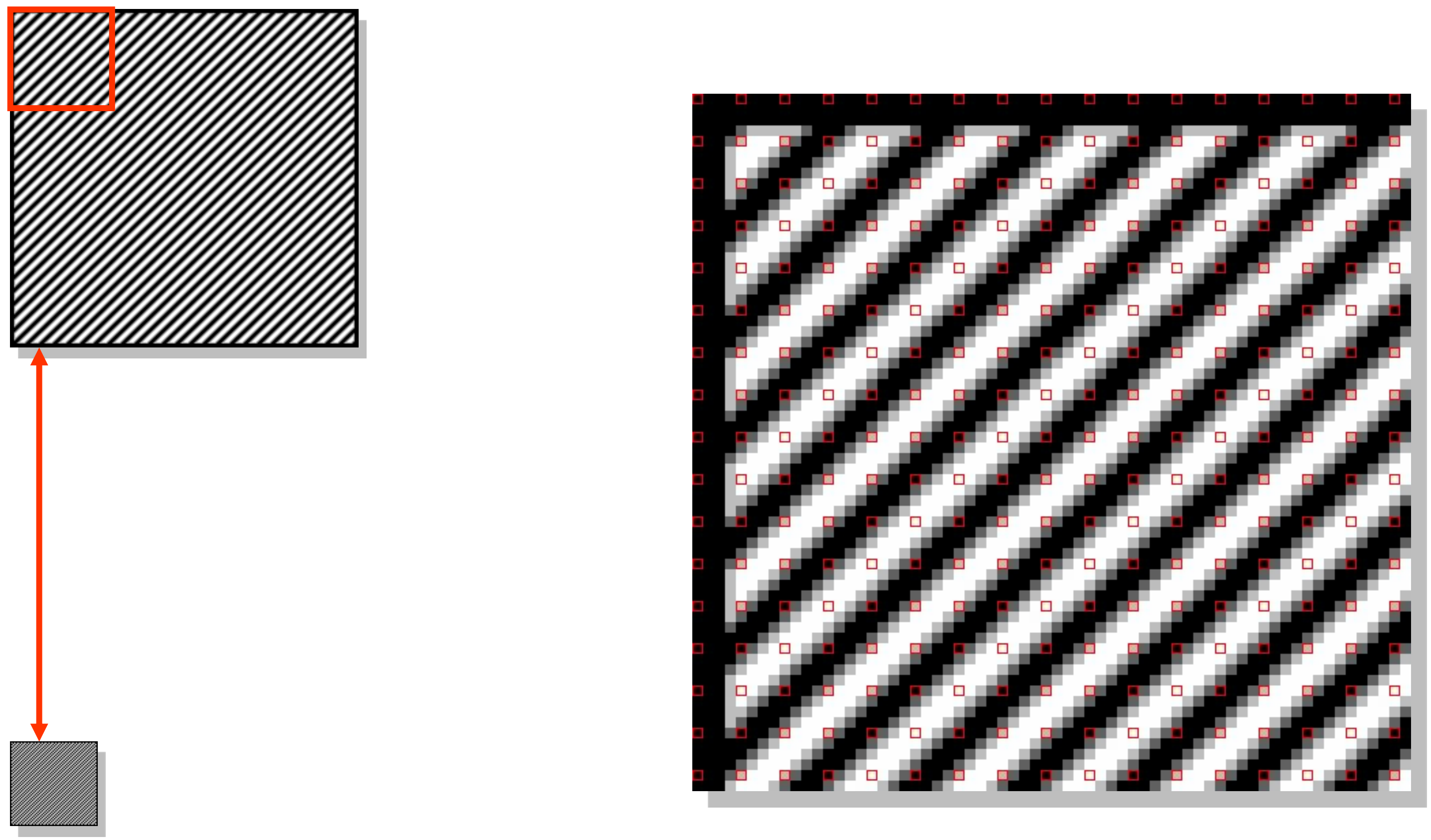
↓

Output kép

Egyszerű példa: Kép negyedére kicsinyítése

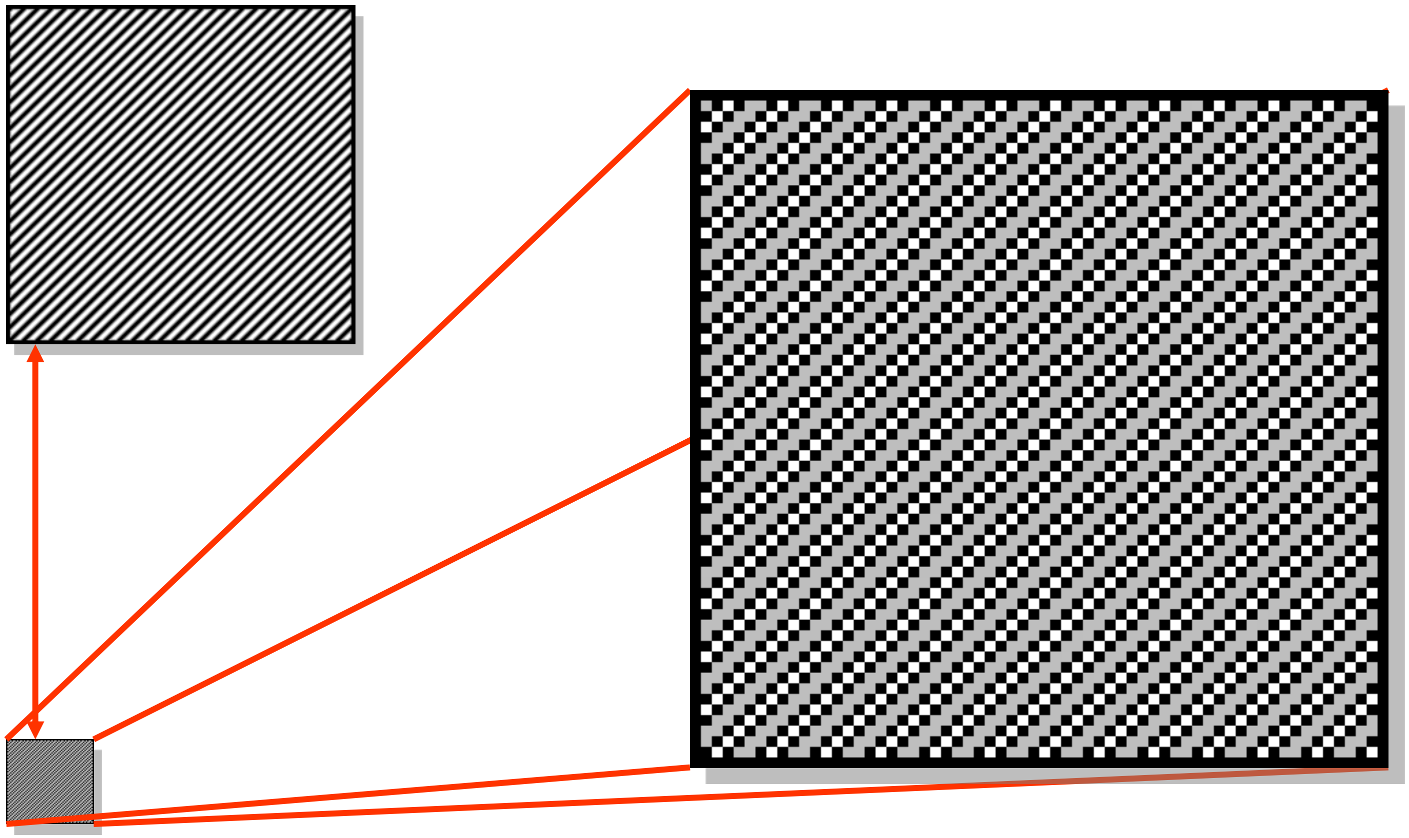


Egyszerű példa: Kép negyedére kicsinyítése



Minden 4. sor minden 4. pixelét vegyük ki és másoljuk a kicsinyített képbe.

Egyszerű példa: Kép negyedére kicsinyítése



Mi történt?

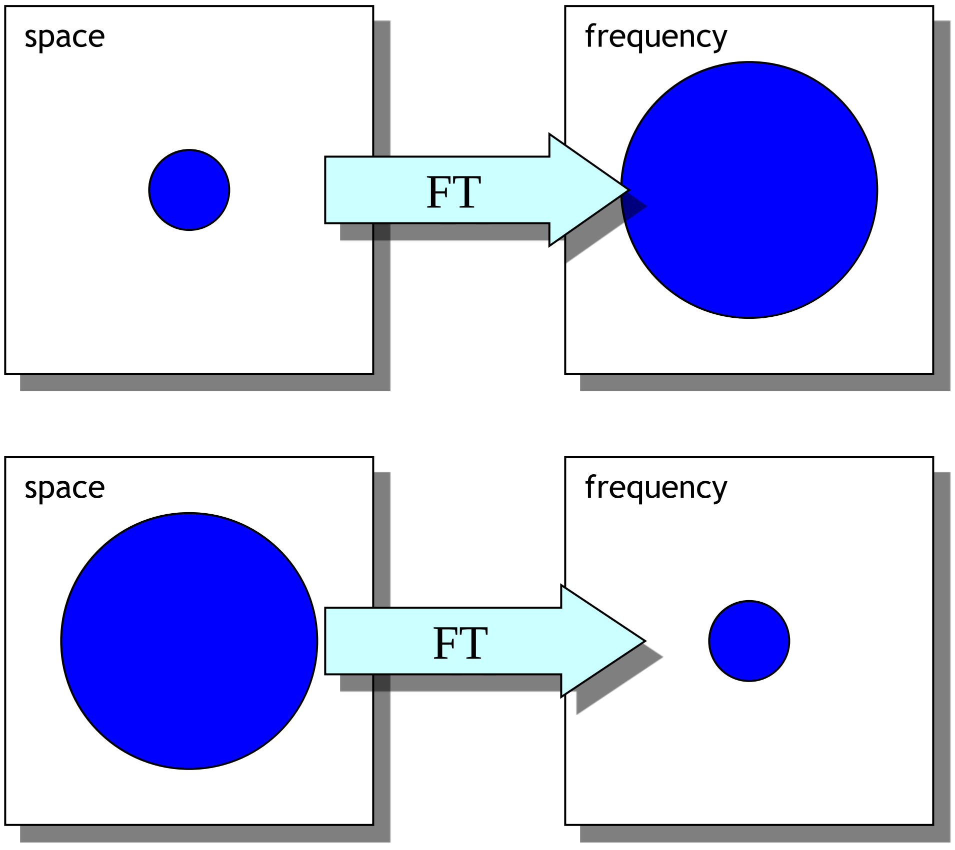
A MINTAVÉTELEZÉS KORLÁTAI

Méret és frekvencia közötti összefüggés

- **Ha *ΔxΔy* egy objektum kiterjedése a képtérben és *ΔuΔv* a Fourier térben, akkor közöttük az alábbi összefüggés áll fent:**

$$\Delta x \, \Delta y \cdot \Delta u \, \Delta v \geq \frac{1}{16\pi^2}$$

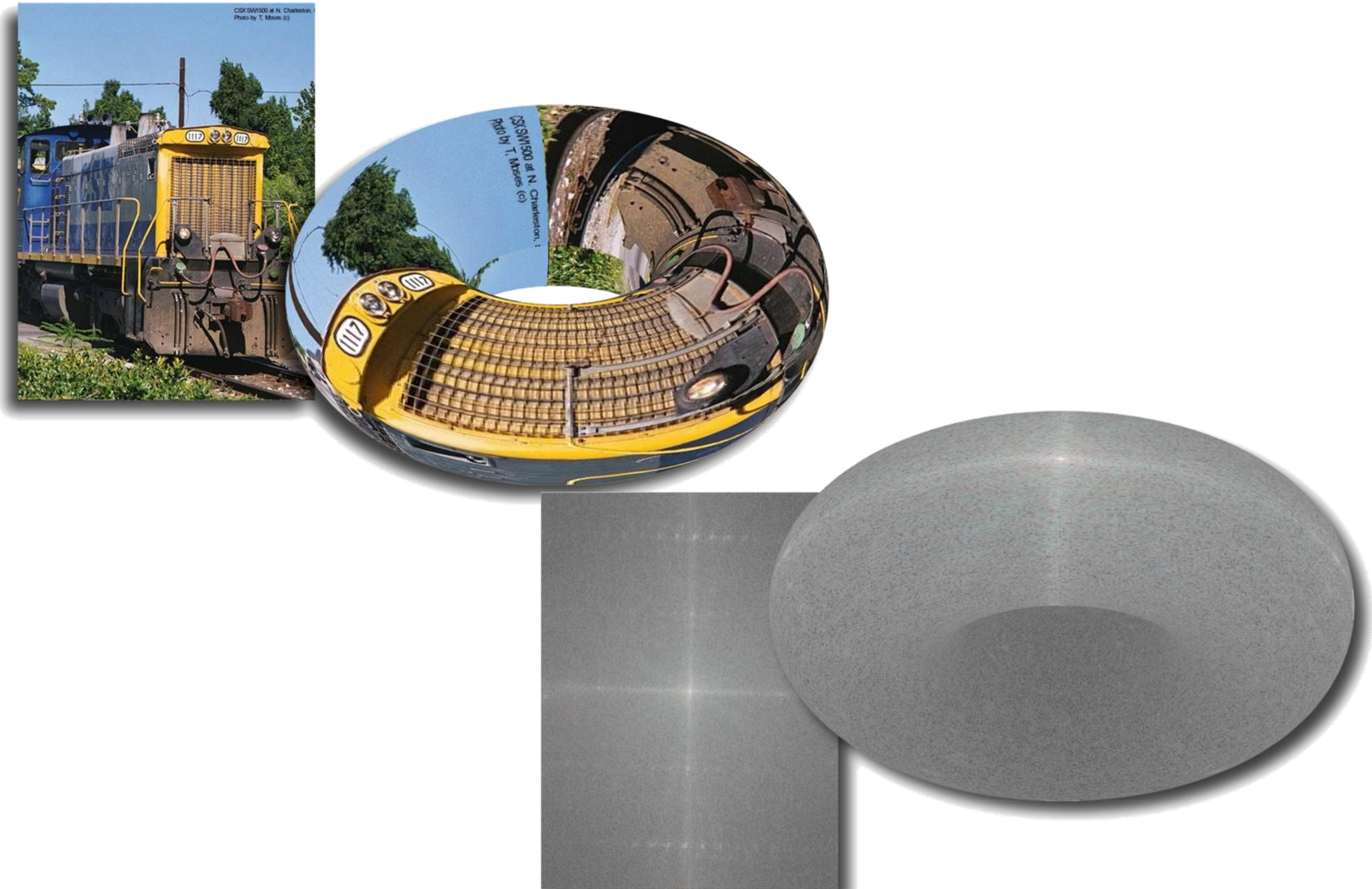
- **Egy kis méretű képelem nagy kiterjedésű lesz a Fourier térben és fordítva.**



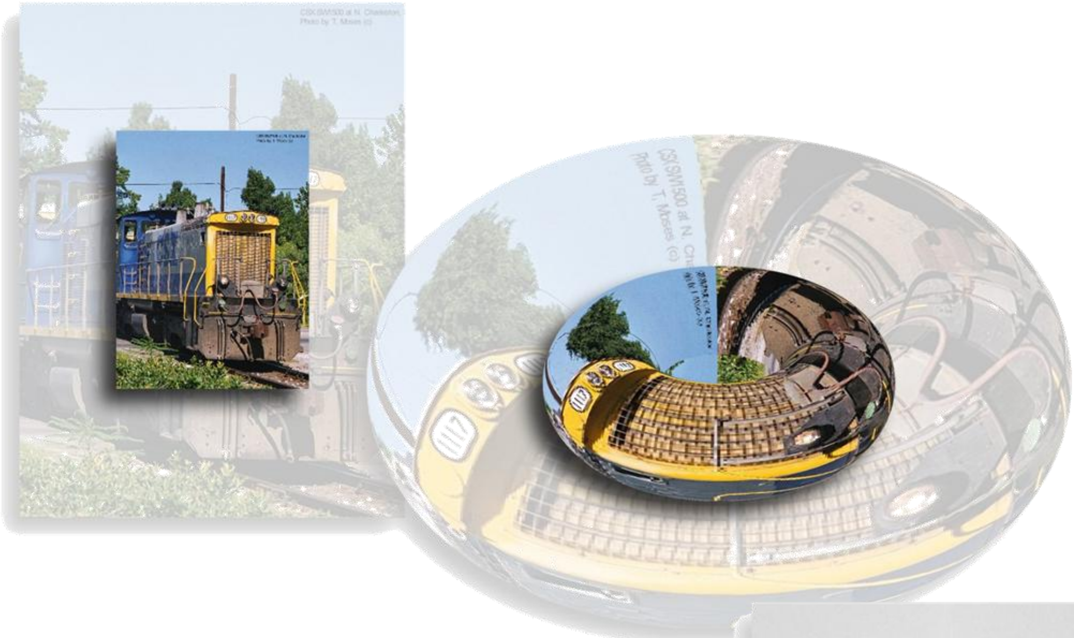
Képméret csökkentés - alulmintavételezés

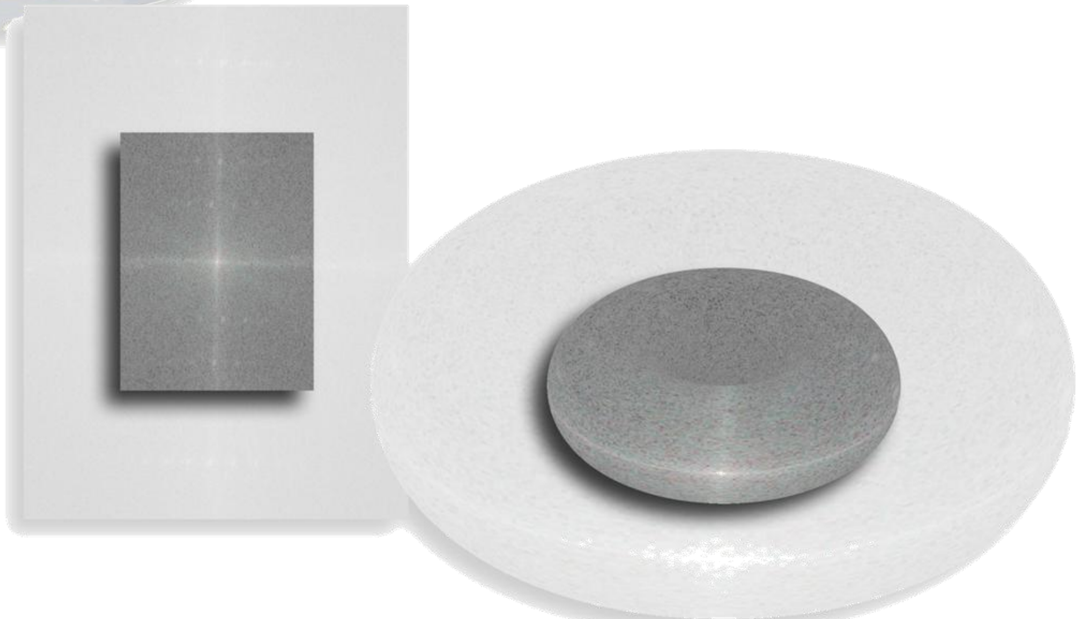
- Egy $R \times C$ méretű I kép n-ed részre kicsinyítése egy $R/n \times C/n$ méretű J kép.
- Egy J kép diszkrét Fourier transzformáltja a J-vel megegyező méretű.
- A képtér és FT jellemzők mérete fordítottan arányos ➔ J-nek $n\, R/n \times n\, C/n = R \times C$ méretűnek kellene lennie.
- A fentiek egyszerre csak úgy teljesülhetnek, ha $DFT(J)$ átfed önmagával.

Kép és DFT tóruszon értelmezett



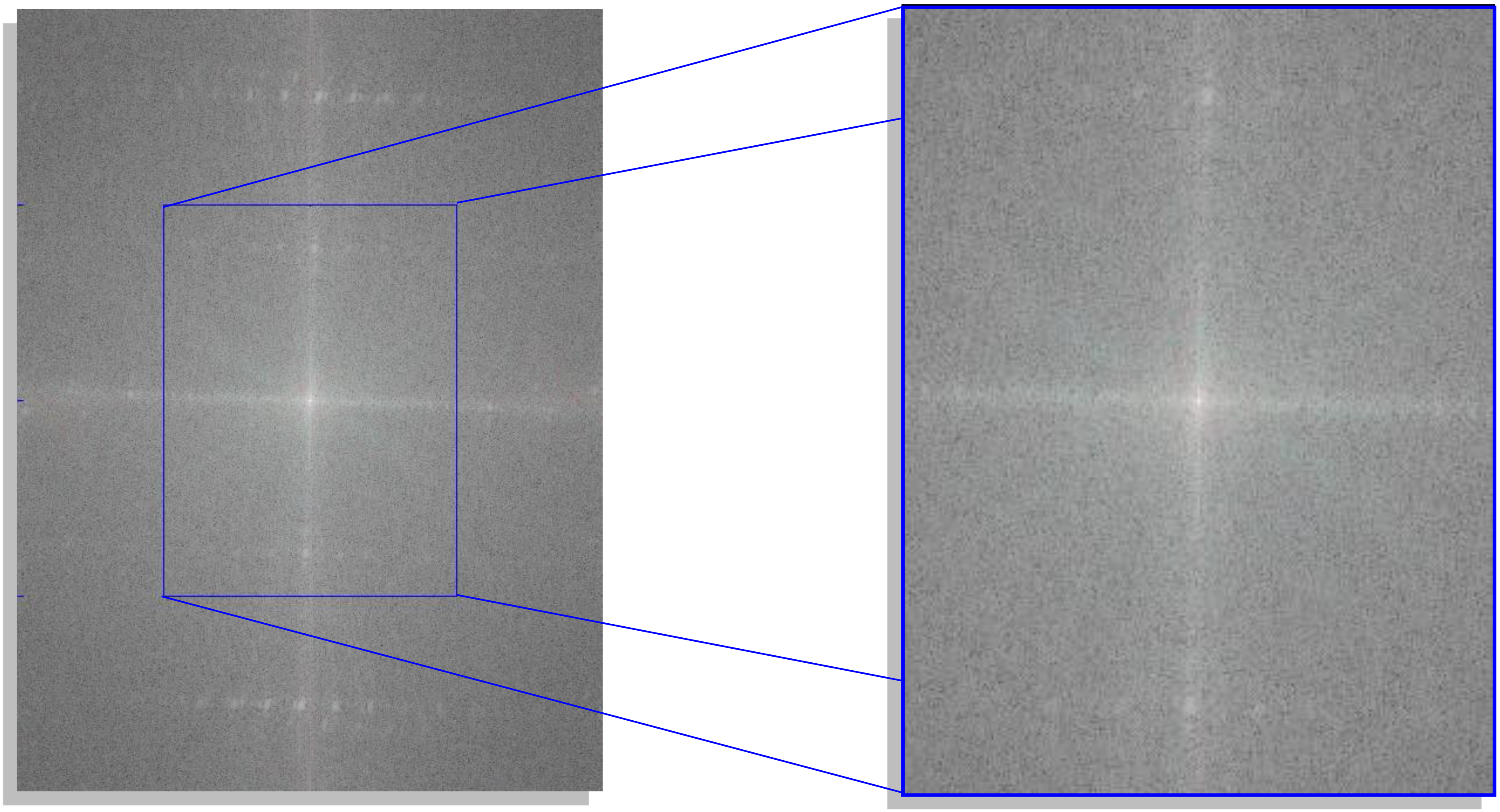
Felére kicsinyített kép





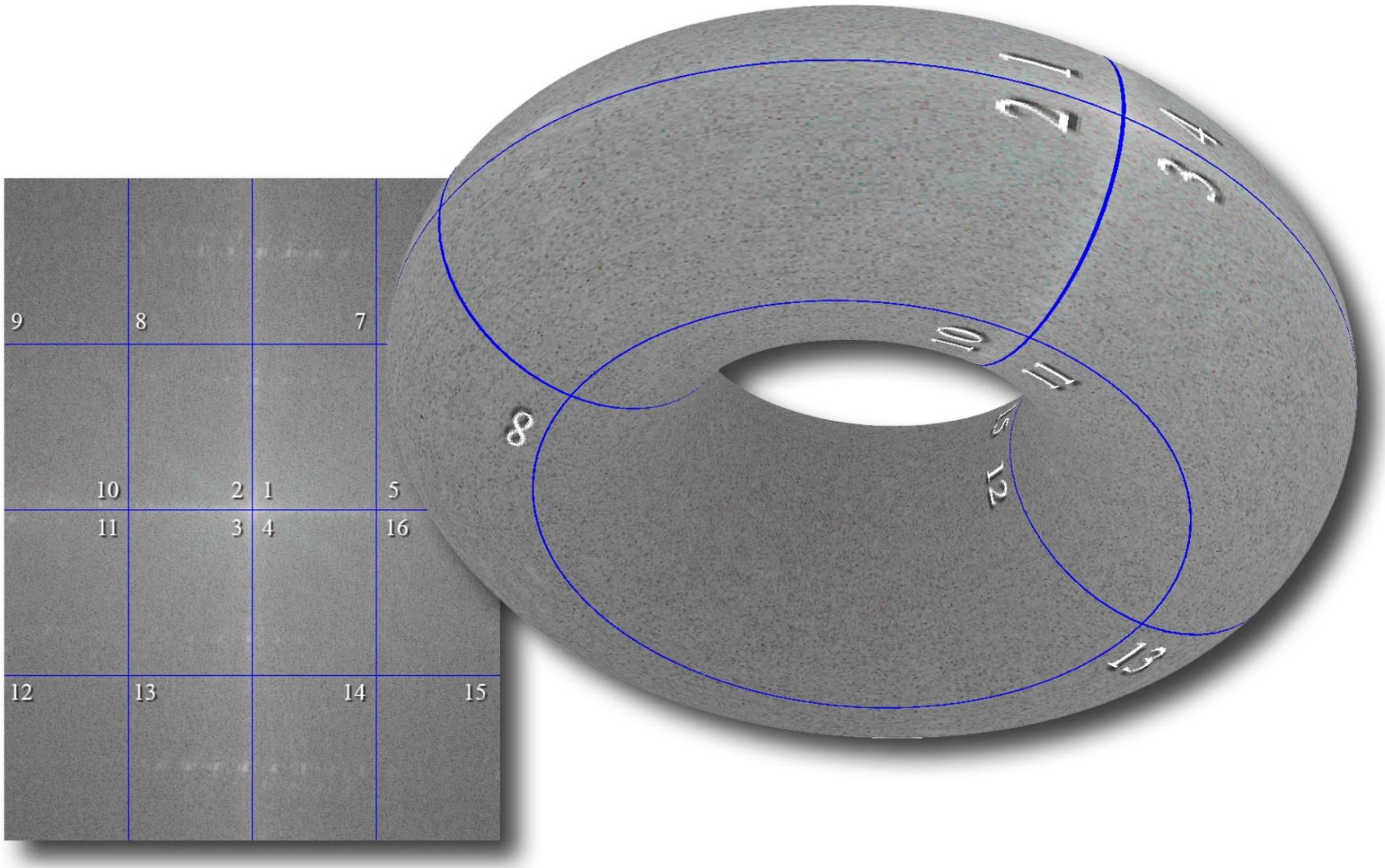
Ideális DFT - felezés

- **A felére kicsinyített kép ideális DFT-je a középső (alacsony frekvenciájú) régió lenne**



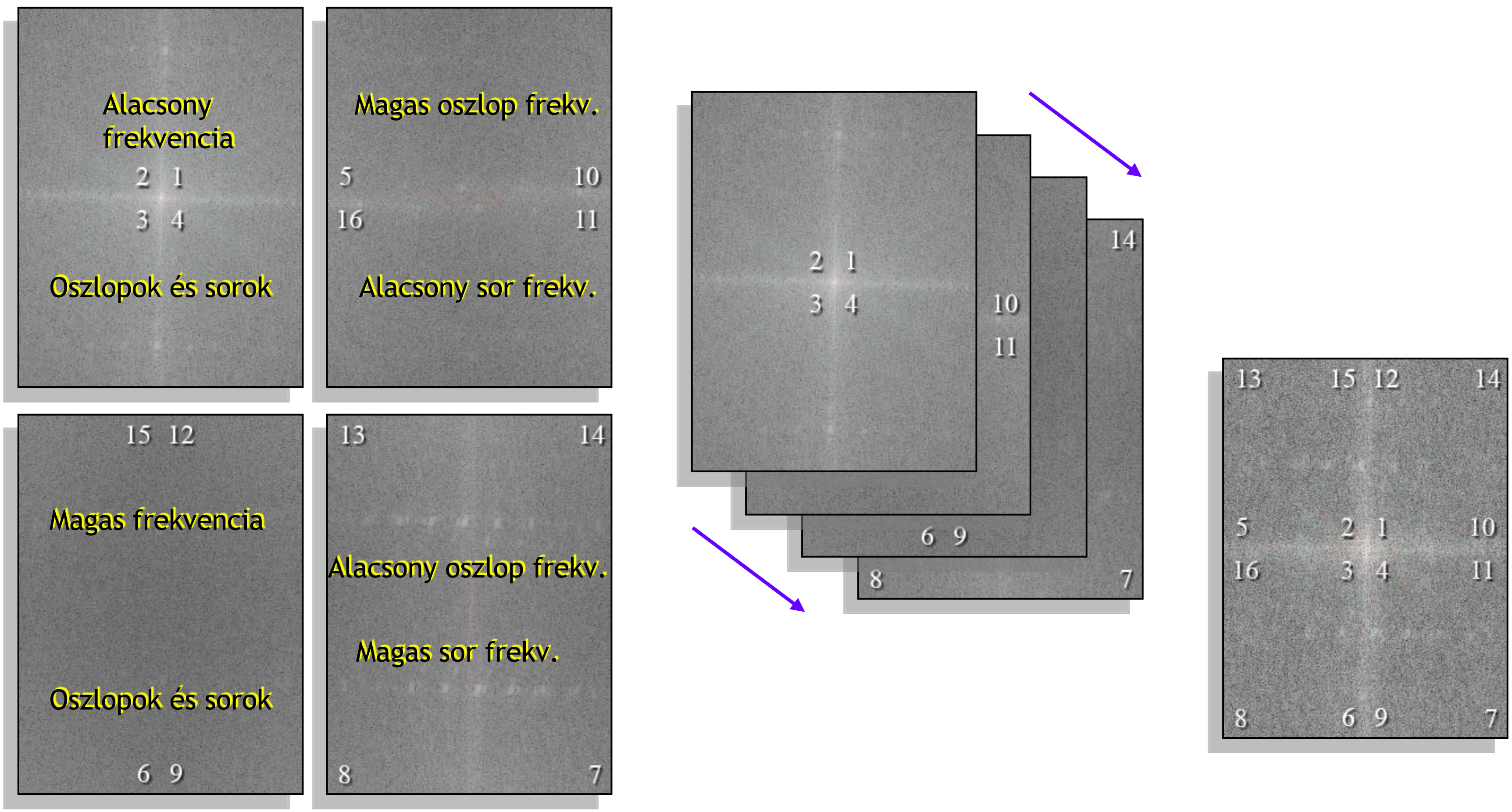
Valós DFT - felezés

- A valóságban az eredeti kép DFT-je 4 részre osztódik, amelyek az eredeti gyűrűn folytonosan helyezkednek el



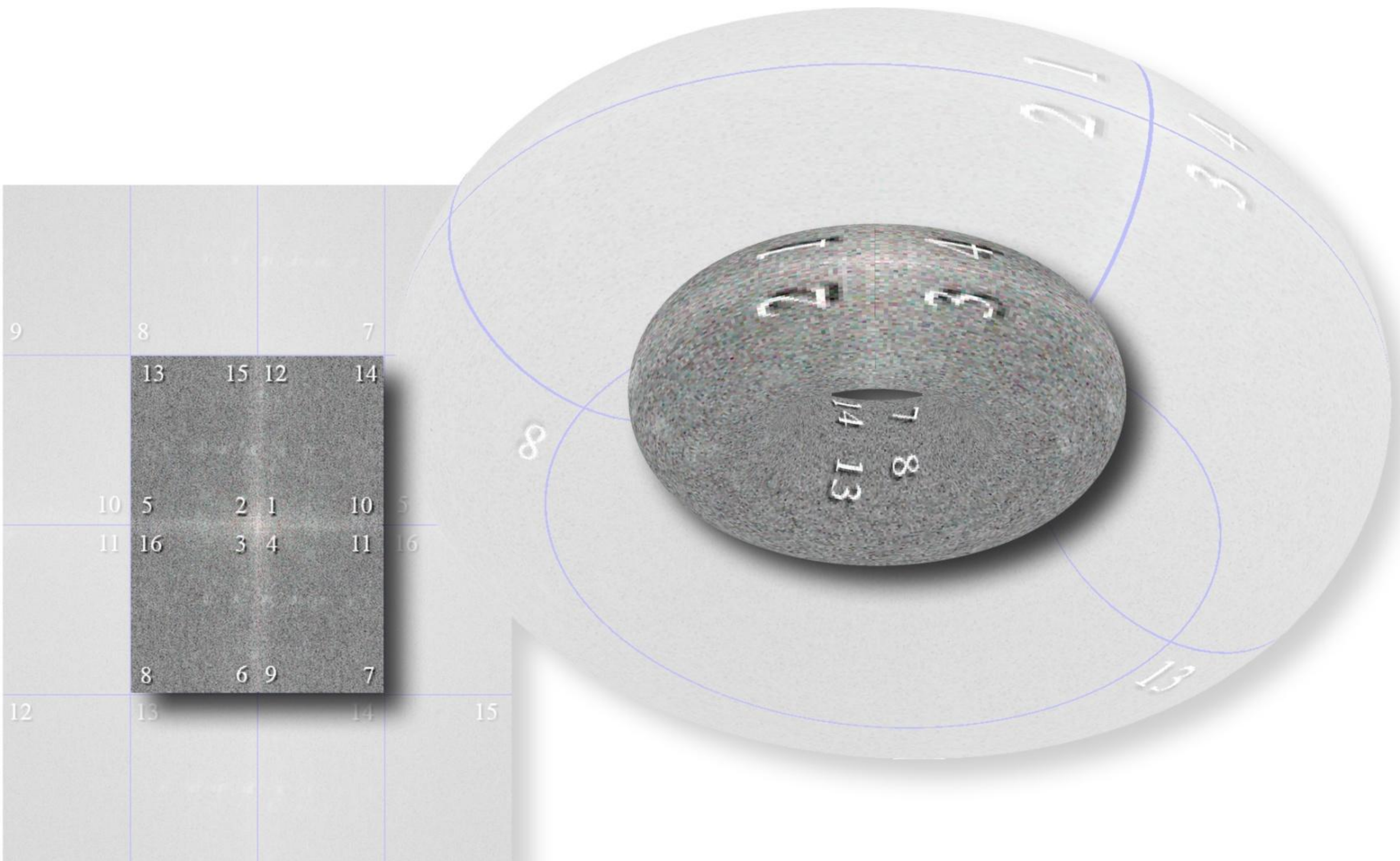
Valós DFT - felezés

- **A kicsinyített képen a 4 rész átlapolódik**



Valós DFT - felezés

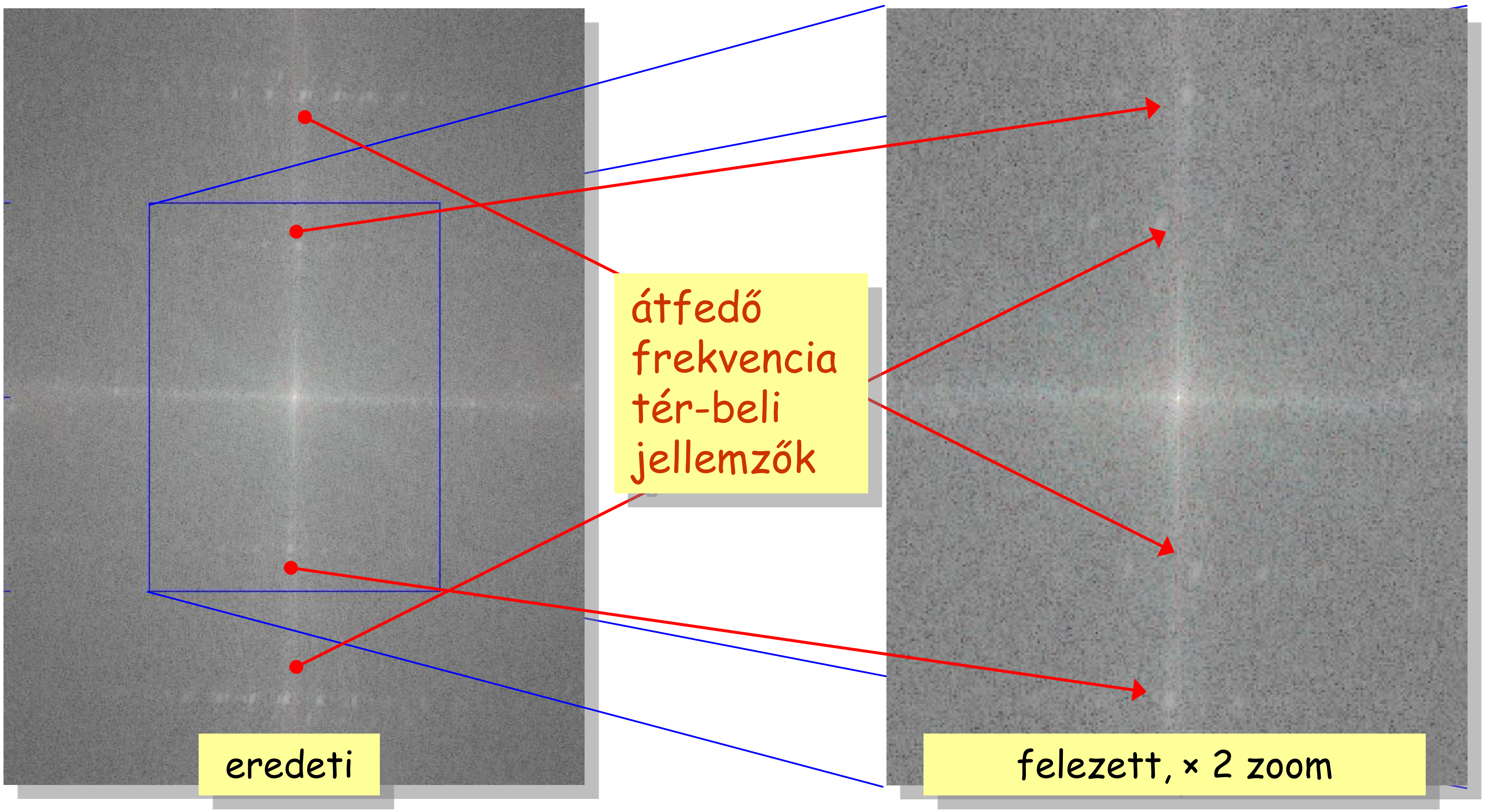
- **A 4 rész megfelel 1-1 tórusznak, amelyek átfednek**



Kató Zoltán: Digitális Képfeldolgozás (Tehetséggondozó program)

Spektrális átfedés

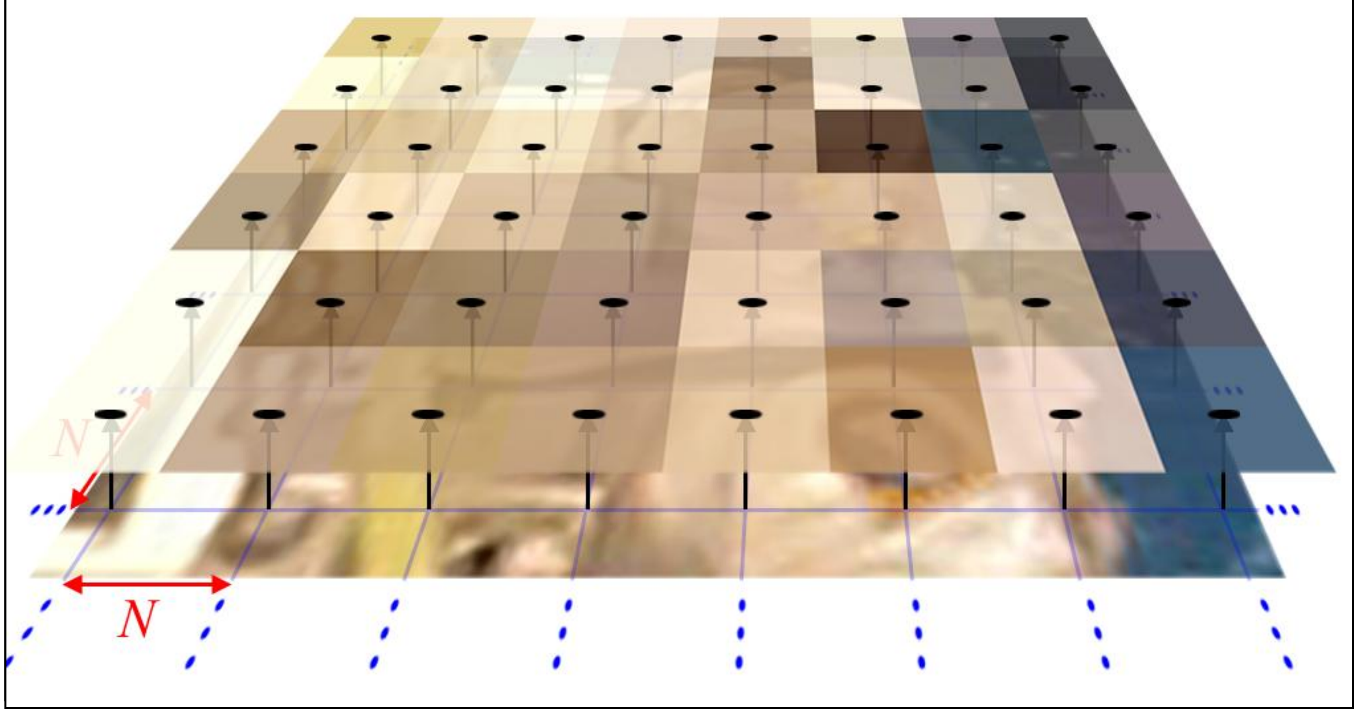
- **At eredeti és felezett kép közötti spektrális átfedés**



Egy kép mintavételezése

- **Egy kép mintavételezése a *samp()* mintavételező függvénnyel való szorzással történik**
  - **A mintavételező függvény nem más, mint egy négyzetrács mentén elhelyezett impulzusfüggvények összessége**

$$\mathrm{samp}_N\{I\}(r,c) = I(r,c)\cdot\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\delta(r-jN)\cdot\delta(c-kN)$$



Impulzus függvény

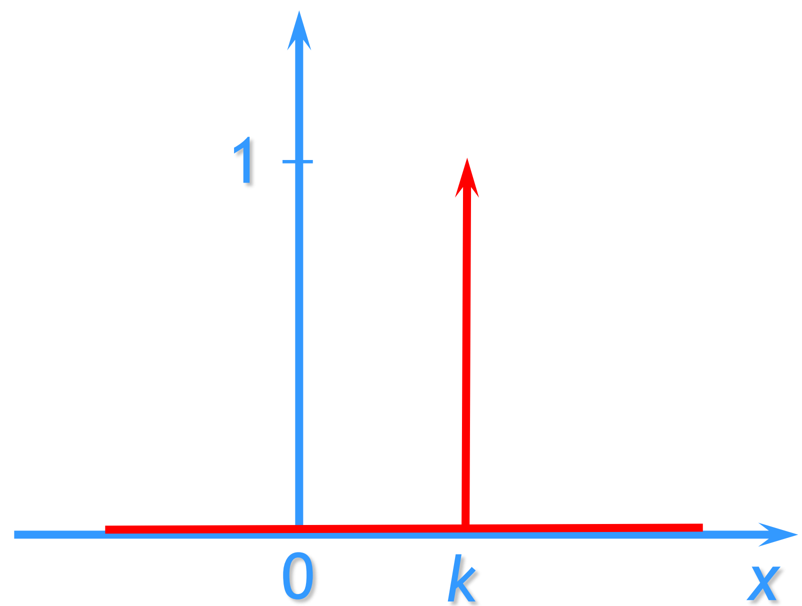
- **A Dirac δ-függvény**

$$\delta(x,y)=\begin{cases}\infty & \text{ha } (x,y)=(0,0)\\ 0 & \text{különben}\end{cases}, \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(x,y)\,dx\,dy=1$$

  - **A Dirac függvény szeparálható:** $\delta(x,y)=\delta(x)\cdot\delta(y)$

  - **A mérésben használt változata (impulzus függvény):**

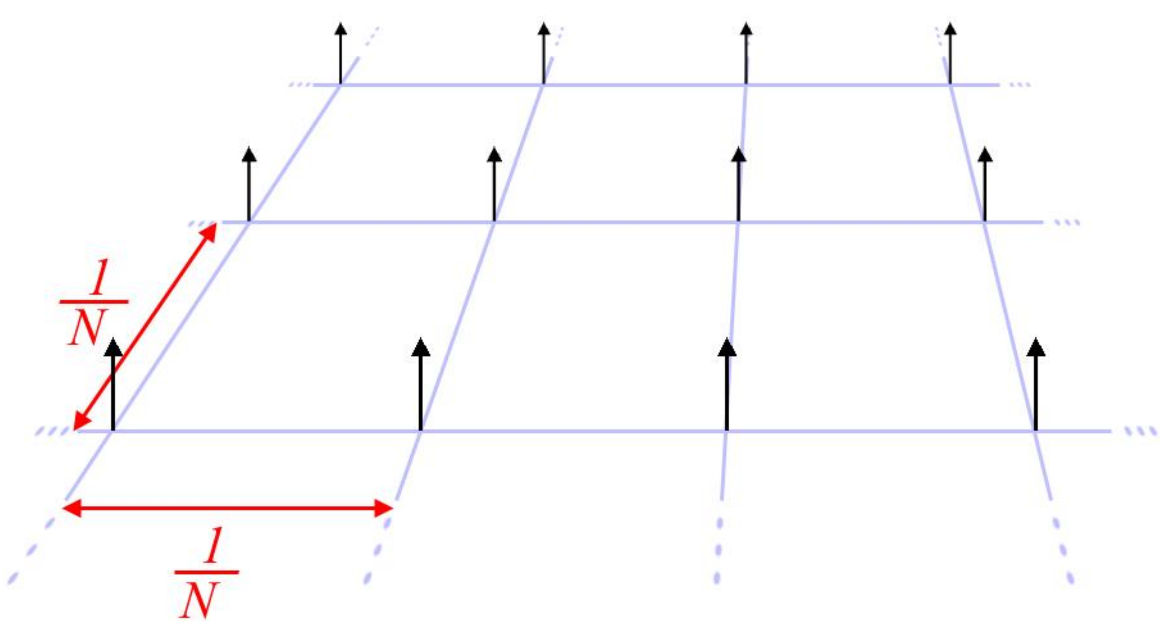
$$\delta(x-k)=\begin{cases}1 & \text{ha } x=k\\ 0 & \text{különben}\end{cases}$$



A mintavételező függvény FT-ja

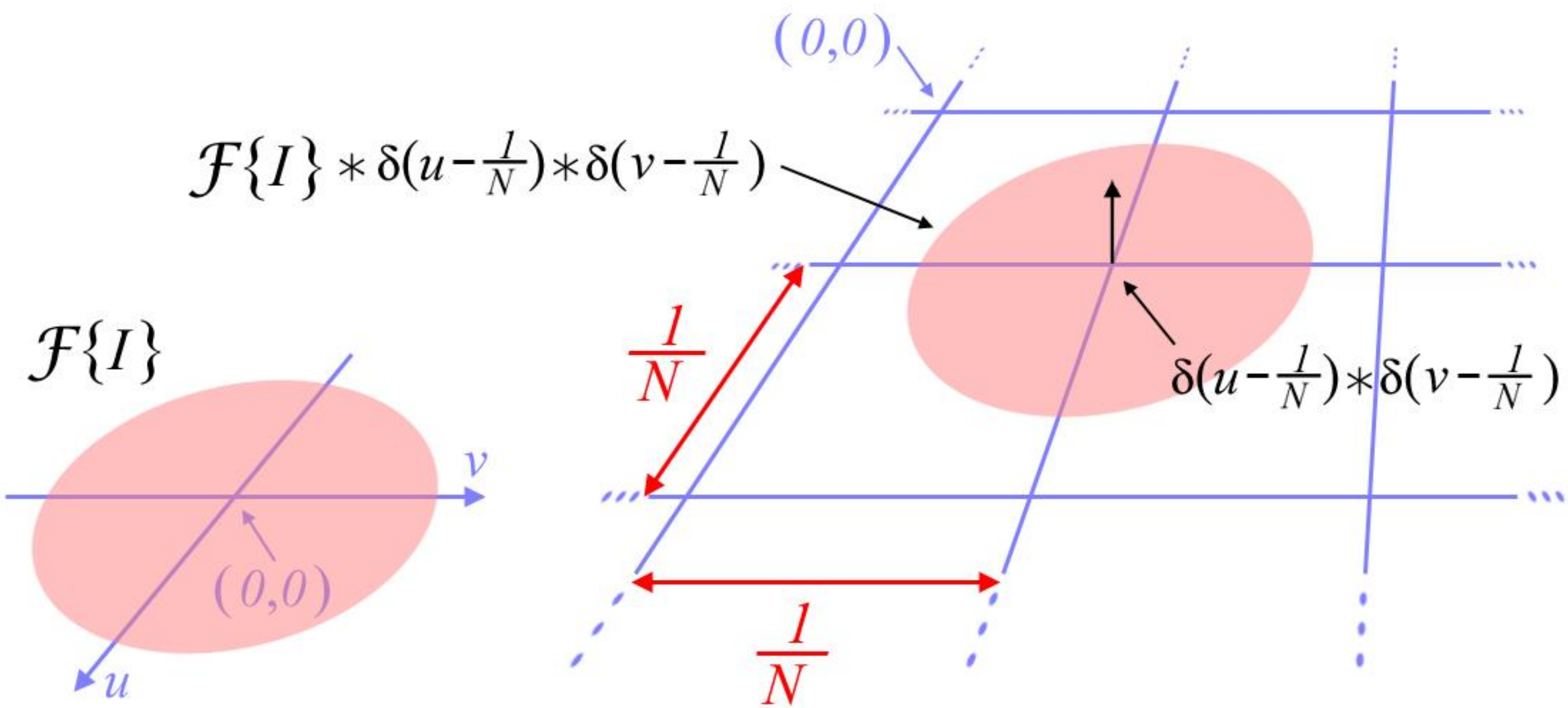
- **Maga is egy mintavételező függvény, de az impulzusok 1/N távolságra lesznek egymástól**
- **A képtérben vett szorzás (mintavételezés) a Fourier térben konvolúció lesz!**

$$\text{samp}_{1/N}(u,v) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \delta\left(u-\frac{j}{N}\right)\delta\left(v-\frac{k}{N}\right)$$



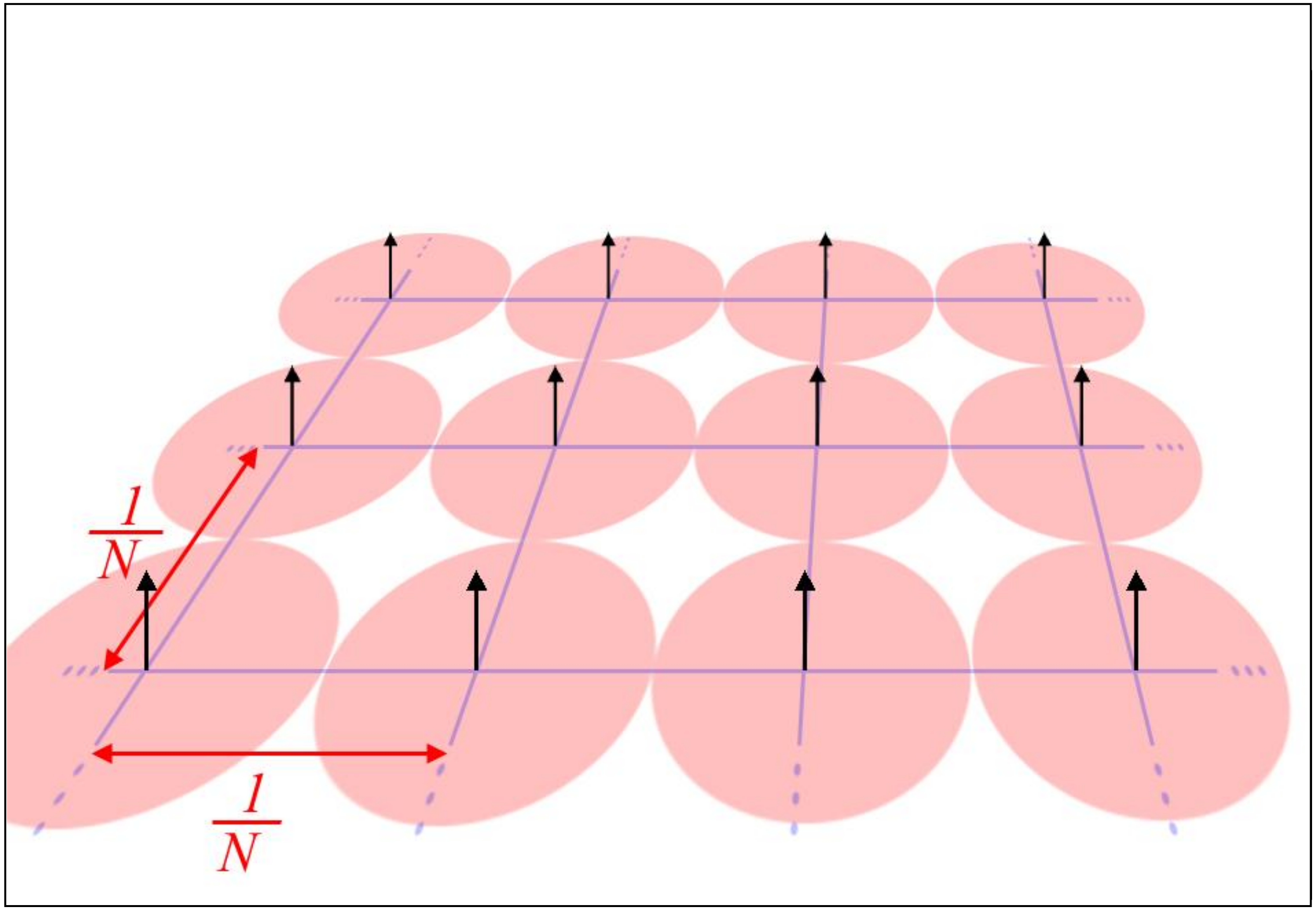
Konvolúció az impulzusfüggvénnyel

- **Az eredmény az eredeti függvény eltolása lesz az impulzusfüggvény pozíciójába**



Egy mintavételezett kép FT-ja

- **A kép FT-ja ismétlődik 1/N intervallumonként**
  - Ha a Fourier tartomány rádiusza ≥1/(2N), akkor átfedés lesz ➔ spektrum átfedés, Moiré hatás



Nyquist frekvencia

T.f.h. az $f(x)$ függvény **sávhatárolt**, vagyis a Fourier transzformáltja $F(X)$ nem tartalmaz w-nél nagyobb frekvenciákat, azaz:

$$F(X)=0, \text{ ha } |X|>w \,.$$

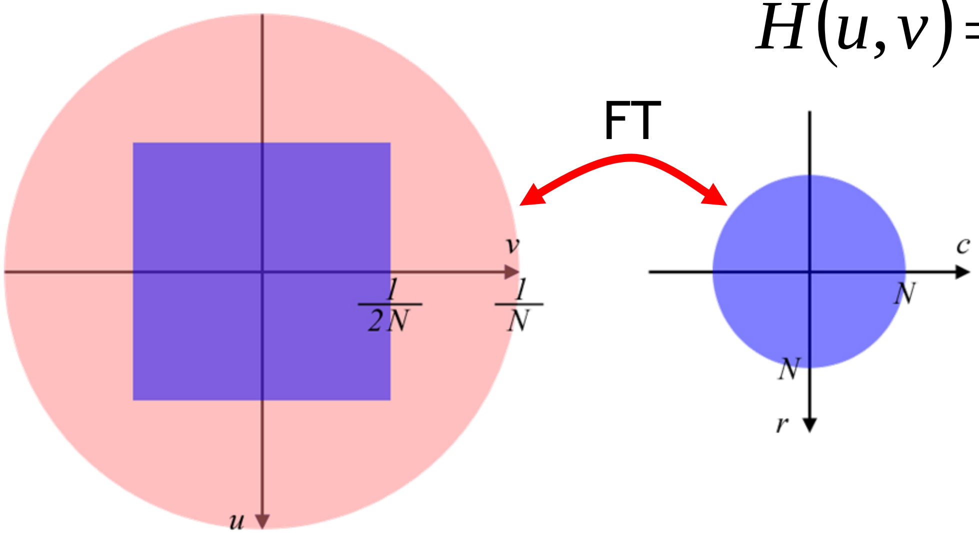
A w-t ekkor **Nyquist**-frekvenciának nevezzük, ha

$$F(-w) \neq 0 \text{ vagy } F(w) \neq 0 \,.$$

Sávhatárolt függvény

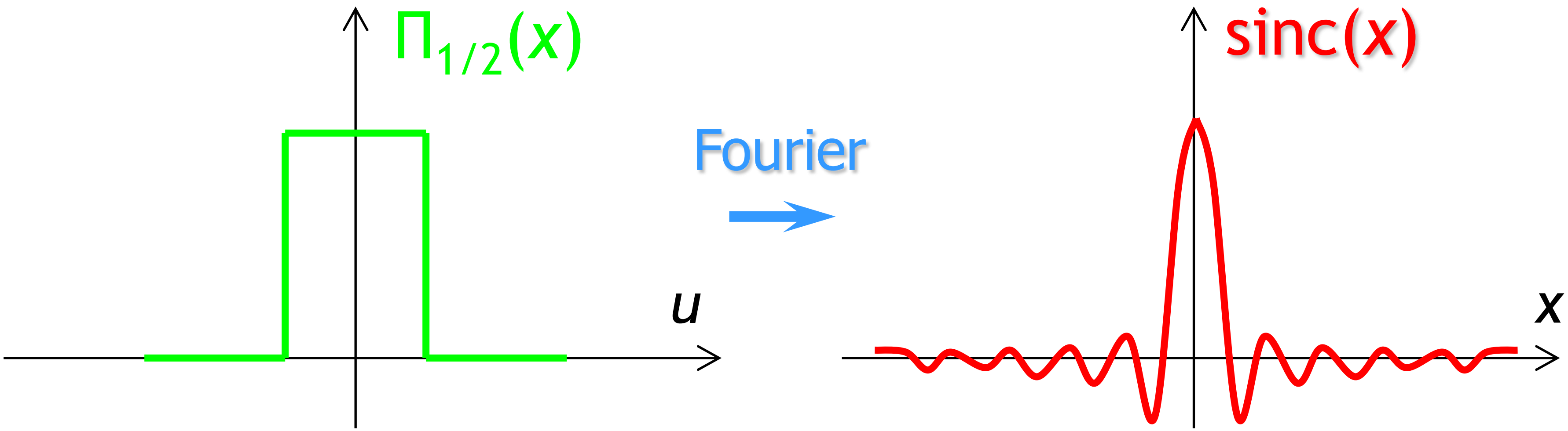
- **Mintavételezés előtt biztosítani kell, hogy a kép nem tartalmaz *1/2N*-nél nagyobb frekvenciákat:**
  - **Szorzás az alábbi dobozfüggvénnyel:**

$$H(u,v)=0 \text{ ahol } u>\frac{1}{2N} \text{ vagy } v>\frac{1}{2N}$$



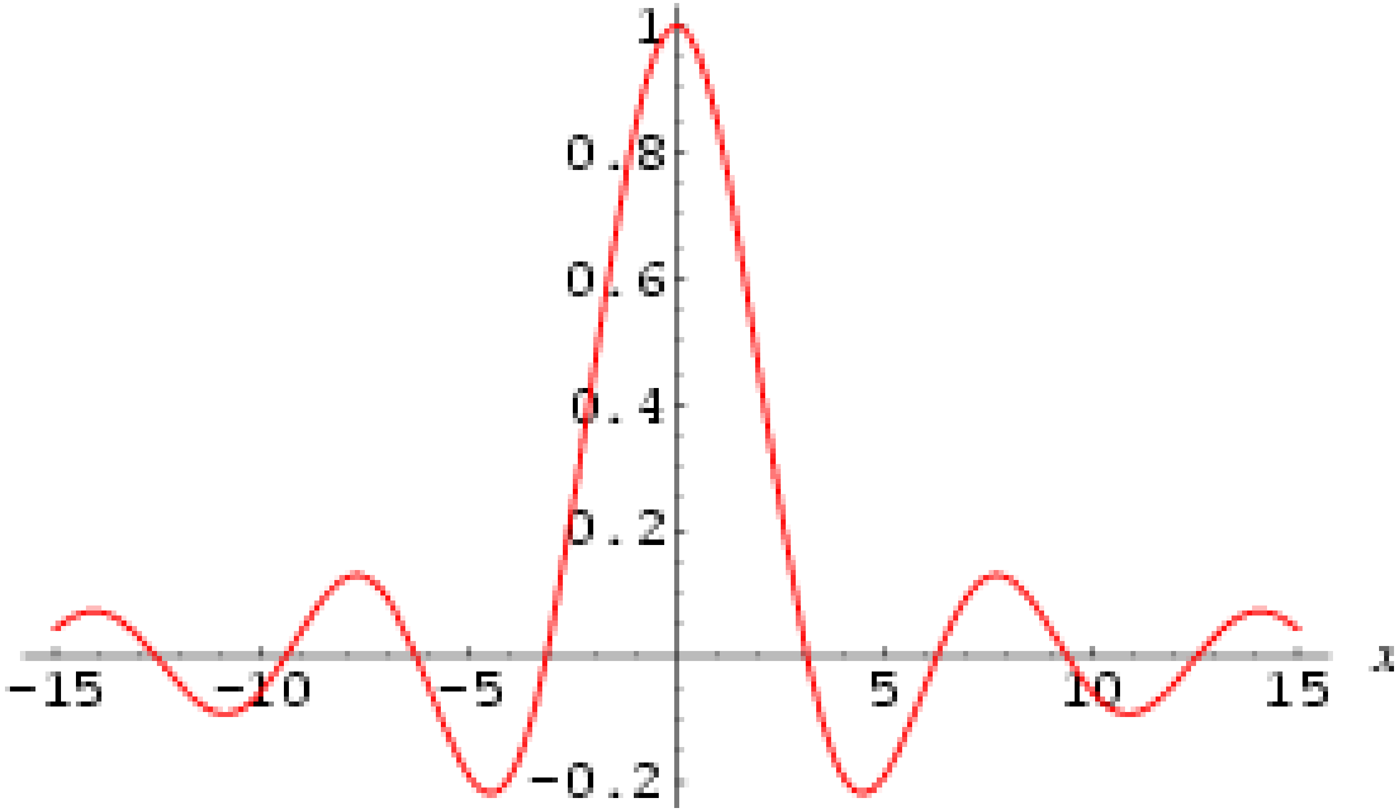
  - **Ez ekvivalens a képtérben egy 2N nagyságú szűrővel**
  - **Tehát ha felezni akarjuk a képet, akkor N=2 és így egy legalább 4x4 méretű simításra van szükség (pl. 5x5 méretű konvolúciós maszkkal simítjuk).**

A *sinc* és a doboz függvény kapcsolata



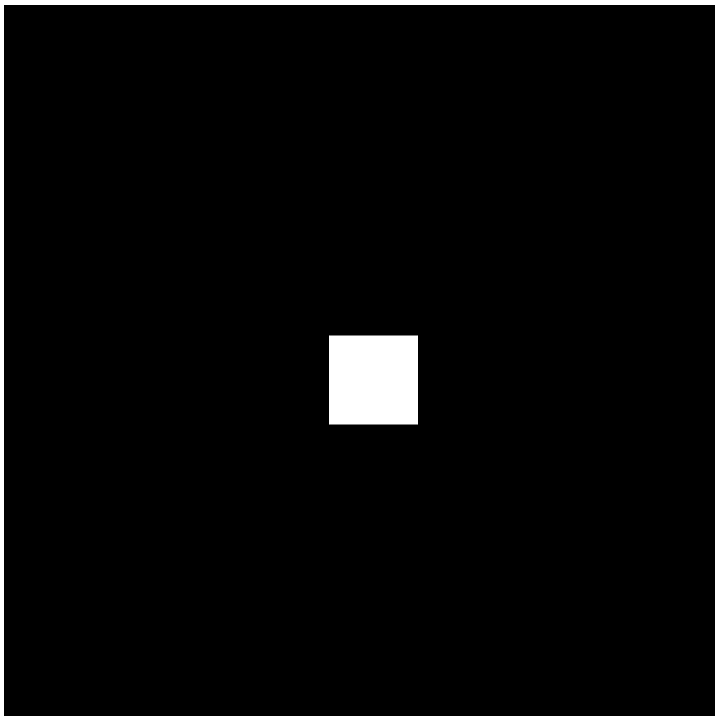
$$F[\Pi_{1/2}(x)](k) = \int_{-\infty}^{\infty} e^{-2\pi ikx} \cdot \Pi_{1/2}(x)\, dx = \mathrm{sinc}(k \cdot \pi)$$

A *sinc* függvény

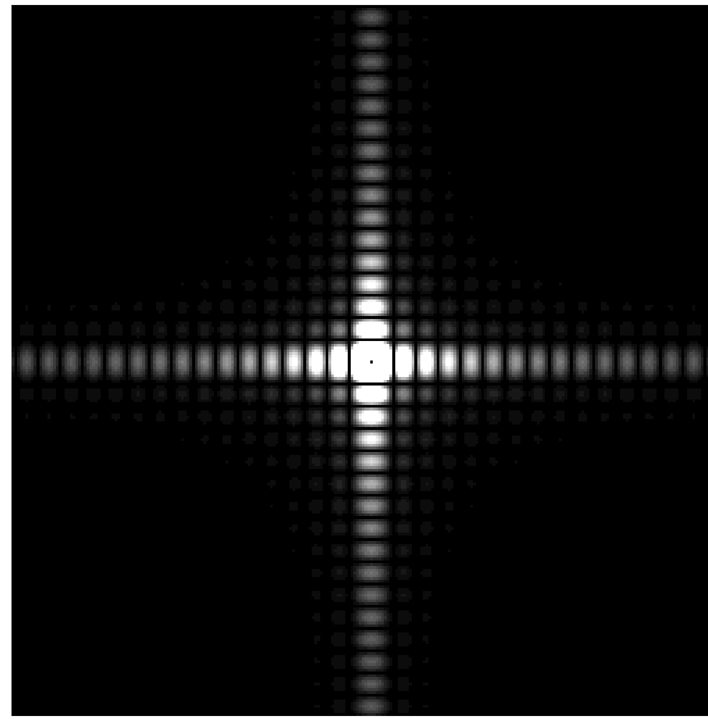


$$\mathrm{sinc}(x) = \begin{cases} 1 & ,\text{ha } x = 0 \\ \dfrac{\sin x}{x} & ,\text{ha } x \neq 0 \end{cases} = \prod_{k=1}^{\infty} \cos\left(\frac{x}{2^k}\right)$$

A *sinc* és a doboz függvény kapcsolata



Fourier →



2D doboz függvény

2D sinc függvény

Whittaker-Shannon interpolációs formulája

Ha az $f(x)$ függvény sávhatárolt (a Fourier transzformáltja nem tartalmaz w-nél nagyobb frekvenciákat), akkor:

$$f(x) = \sum_{n=-\infty}^{\infty} \underbrace{f(n \cdot \frac{1}{2w})}_{\text{diszkrét pontok}} \cdot \frac{\sin(\pi(2wx - n))}{\pi(2wx - n)}$$

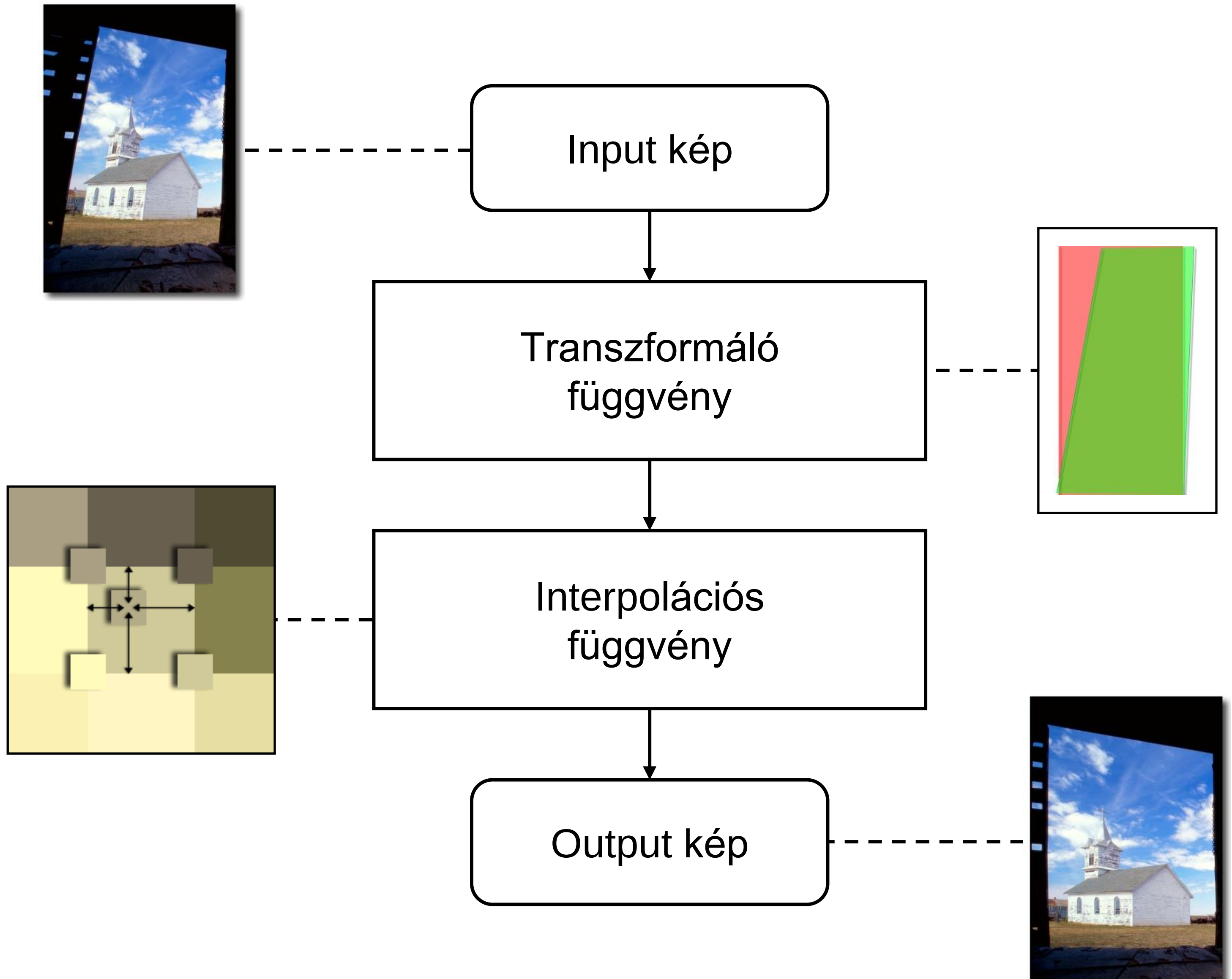
Whittaker-Shannon mintavételezési tétele

A 2D folytonos, w-nél nagyobb frekvenciákat nem tartalmazó (sávhatárolt) képfüggvény akkor és csakis akkor állítható vissza a mintavételezettjéből, ha a Δx és a Δy mintavételezési léptékekre teljesül:

$$\Delta x \leq \frac{1}{2w}\ , \quad \Delta y \leq \frac{1}{2w}\ .$$

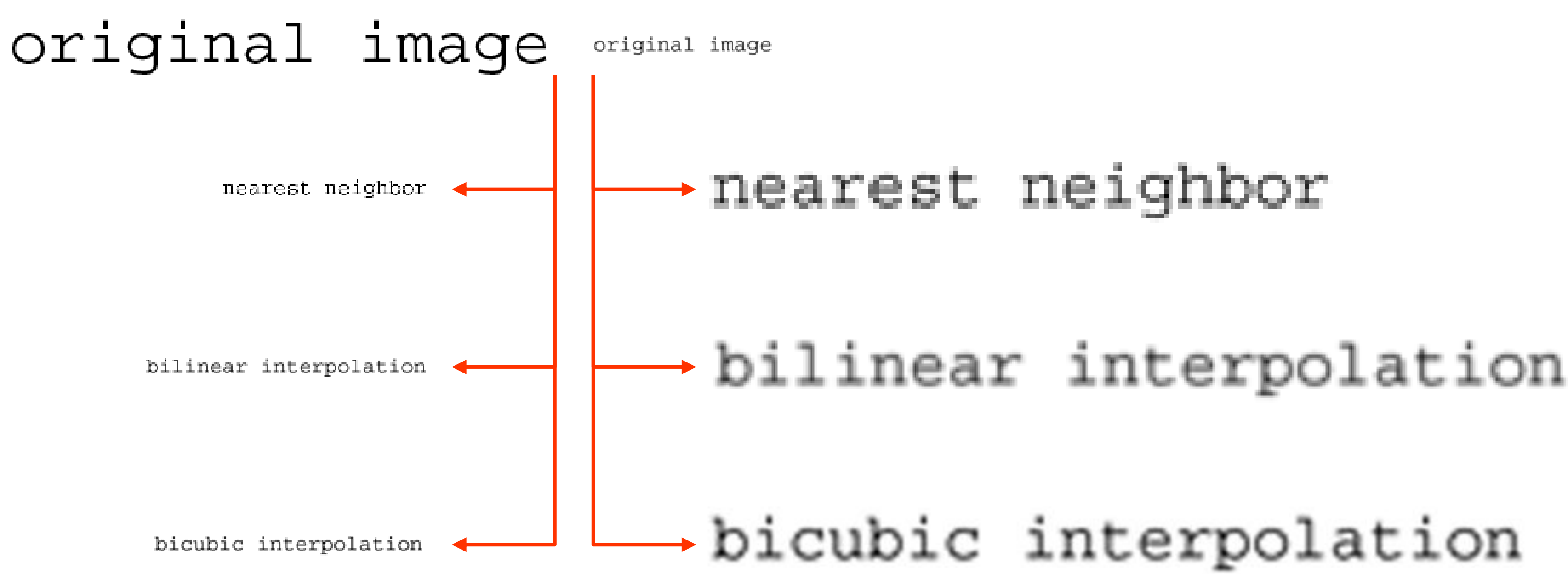
INTERPOLÁCIÓS TECHNIKÁK

Geometriai transzformáció végrehajtása



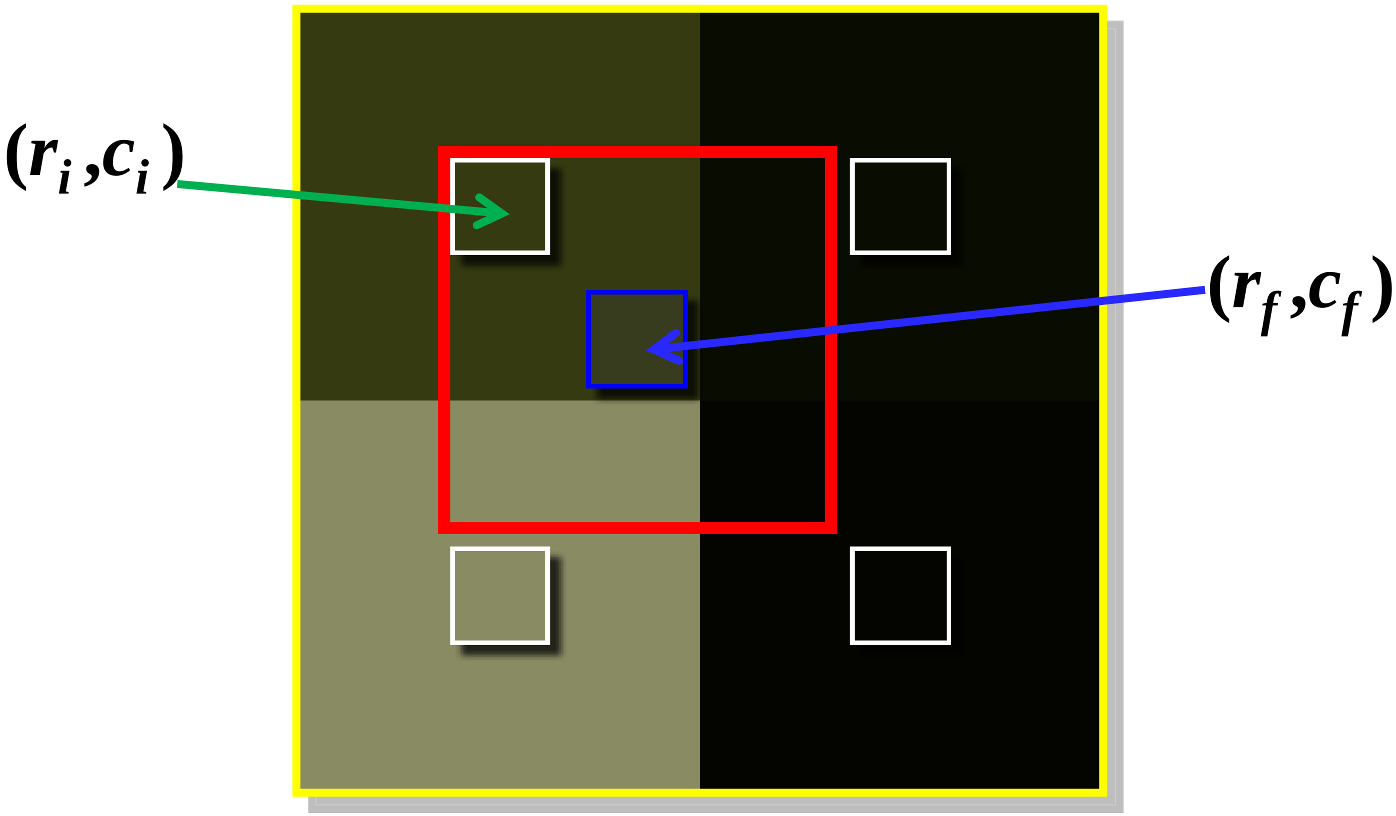
Kató Zoltán: Digitális Képfeldolgozás (Tehetséggondozó program)

Interpolációs technikák



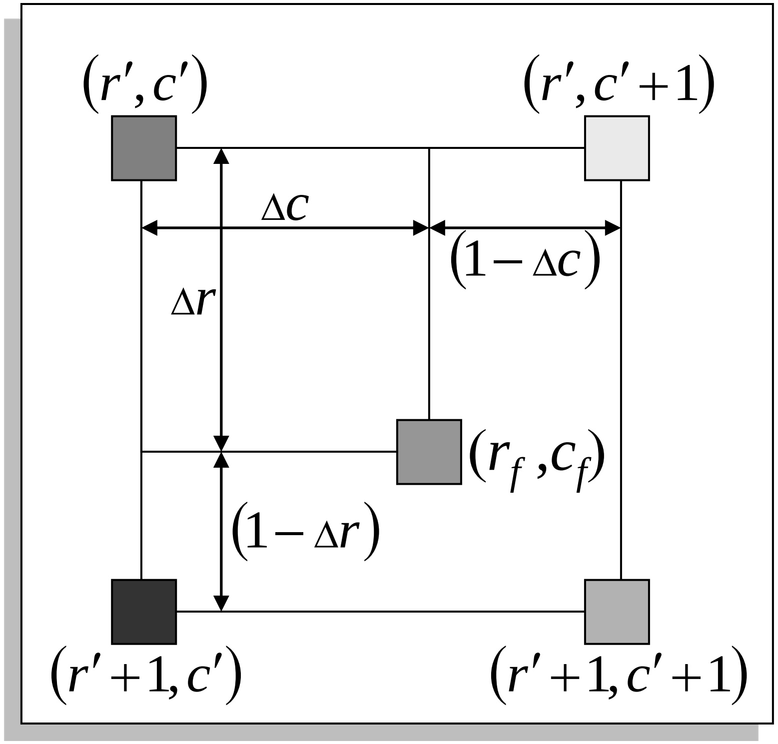
Legközelebbi szomszéd (NN)

- **A legközelebbi szomszéd (Nearest Neighbor – NN) interpoláció nem más, mint a subpixeles koordináták kerekítése:**
  - $J(r,c) = I(r_i, c_i)$ **ahol** $(r_i, c_i) = \text{round}(r_f, c_f)$



Bilineáris

- **A bilineáris interpoláció nem más, mint a subpixeles (r_f, c_f) koordinátához legközelebb eső 4 szomszédos pixelérték súlyozott átlaga:**



$$
\begin{aligned}
J(r,c) = {} & I(r',c')\cdot(1-\Delta r)\cdot(1-\Delta c) \\
& + I(r'+1,c')\cdot\Delta r\cdot(1-\Delta c) \\
& + I(r',c'+1)\cdot(1-\Delta r)\cdot\Delta c \\
& + I(r'+1,c'+1)\cdot\Delta r\cdot\Delta c.
\end{aligned}
$$

  - **Ahol**

$$r' = \lfloor r_f \rfloor;\ c' = \lfloor c_f \rfloor$$

  - **és**

$$\Delta r = r_f - r';\ \Delta c = c_f - c'$$

Kató Zoltán: Digitális Képfeldolgozás (Tehetséggondozó program)

Bicubic

- **A bicubic interpoláció a subpixeles (r_f, c_f) koordinátához legközelebb eső 4 szomszédos pixelértéket illetve azok parciális deriváltjait is használja**
  - **Ezért 4x4 szomszédság szükséges** $(r', c') = (\lfloor r_f \rfloor, \lfloor c_f \rfloor)$

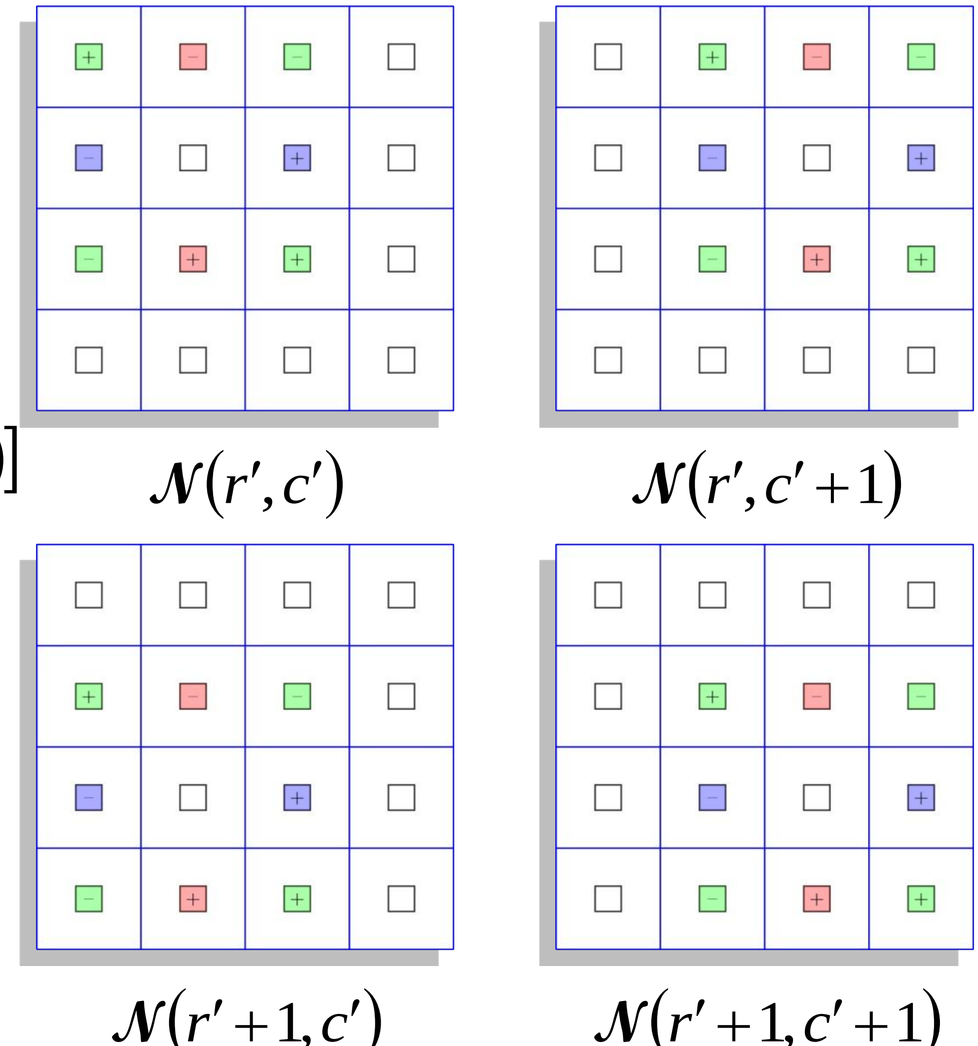
$$\left\{\left\{ \frac{\partial}{\partial c'} \mathrm{I}(r'+i, c'+j), \frac{\partial}{\partial r'} \mathrm{I}(r'+i, c'+j), \frac{\partial}{\partial c' r'} \mathrm{I}(r'+i, c'+j) \right\}_{j=0}^{1} \right\}_{i=0}^{1}$$

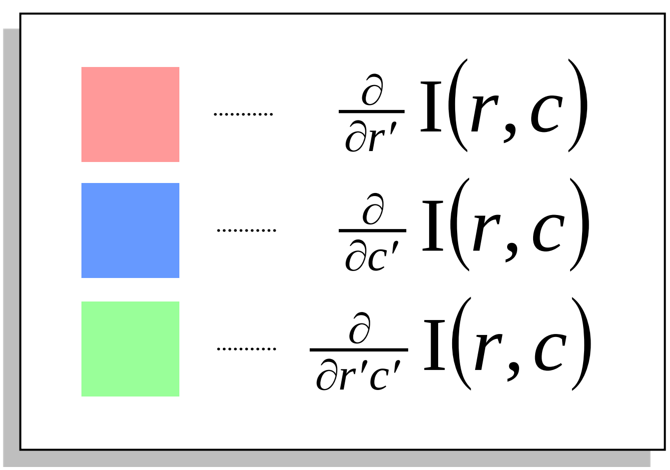
$(r'$-$1,c'$-$1)$	$(r'$-$1,c')$	$(r'$-$1,c'+1)$	$(r'$-$1,c'+2)$
$(r',c'$-$1)$	(r',c')	$(r',c'+1)$	$(r',c'+2)$
$(r'+1,c'$-$1)$	$(r'+1,c')$	$(r'+1,c'+1)$	$(r'+1,c'+2)$
$(r'+2,c'$-$1)$	$(r'+2,c')$	$(r'+2,c'+1)$	$(r'+2,c'+2)$

Bicubic

- $\mathcal{N}(r',c')$ az (r',c') 8-pixeles szomszédsága

$$
\begin{aligned}
\frac{\partial}{\partial r'} \mathrm{I}(r',c') &= (r',c') \text{ diff. átlaga} \\
&= \tfrac{1}{2}\Big[\big(\mathrm{I}(r'+1,c') - \mathrm{I}(r',c')\big) \\
&\quad + \big(\mathrm{I}(r',c') - \mathrm{I}(r'-1,c')\big)\Big] \\
&= \tfrac{1}{2}\Big[\mathrm{I}(r'+1,c') - \mathrm{I}(r'-1,c')\Big]
\end{aligned}
$$





Bicubic

- **A parciális deriváltak az alábbi harmadfokú polinomok szorzatának összegében kombinálódnak**

$$J(r,c) = \sum_{m=-1}^{2} \sum_{n=-1}^{2} I(r'+m, c'+n) P(\Delta r - m) P(n - \Delta c)$$

- **ahol**

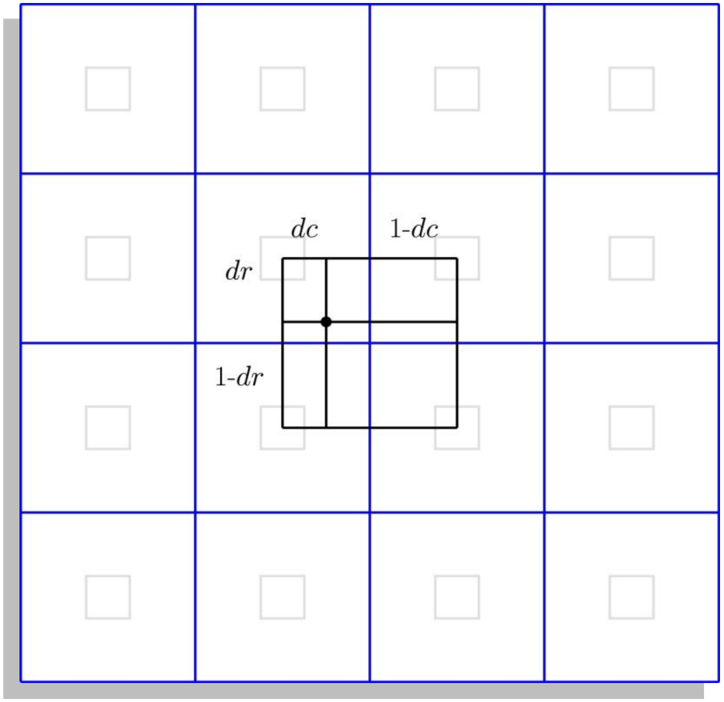
$$P(x) = \tfrac{1}{6}\left[Q(x+2)^3 - 4Q(x+1)^3 - 6Q(x)^3 - 4Q(x-1)^3 \right]$$

- **és**

$$(r', c') = (\lfloor r_f \rfloor, \lfloor c_f \rfloor)$$

$$(\Delta r, \Delta c) = (r_f - r', c_f - c')$$

$$Q(x) = \begin{cases} x & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$



Példa: Kicsinyítés 3/7 arányban

eredeti

nearest neighbor

bilinear

bicubic

Példa: Kicsinyítés 3/7 arányban - eredeti

eredeti

Példa: Kicsinyítés 3/7 arányban - NN

nearest neighbor

Példa: Kicsinyítés 3/7 arányban - bilinear

bilinear

Példa: Kicsinyítés 3/7 arányban - bicubic

bicubic

Példa: Kicsinyítés 3/7 arányban - eredeti

eredeti

Felhasznált anyagok

- Palágyi Kálmán: Digitális Képfeldolgozás
  /pub/Digitalis_kepfeldolgozas
- Trevor Darrell: C280, Computer Vision
  http://www.cs.cmu.edu/afs/cs/academic/class/15385-s06/lectures/ppts/
- Richard Alan Peters: EECE/CS 253 Image Processing
  http://www.archive.org/details/Lectures_on_Image_Processing
- További források az egyes diákon megjelölve