

# Robusztus kvaternió Zernike momentumok és alkalmazásuk színes képek elemzésére és felismerésére

Nagy Gergely

Eötvös Loránd Tudományegyetem, Informatikai Kar

Tudományos Diákköri Konferencia

Budapest, 2020. december 16.

Témavezető: Németh Zsolt

EFOP-3.6.3-VEKOP-16-2017-00001

**SZÉCHENYI** 2020



HUNGARIAN  
GOVERNMENT

European Union  
European Social  
Fund



INVESTING IN YOUR FUTURE

# Momentumok

Általában valamilyen numerikus, leíró érték a pixel intenzitások alapján, például geometriai momentumok:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y).$$

**Zernike momentumok:** az egységkörön definiált, ortogonális Zernike függvények alapján

$$Z_{n,m}(f) = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} f(r, \theta) R_{n,m}(r) e^{-im\theta} dr d\theta,$$

ahol  $R_{n,m}(r)$  az ortogonális, sugárirányú polinomok.

**Momentumok invariánsok:** forgatás, skálázás és eltolás invariáns értékek konstruálhatók.

# Momentumok alkalmazása színes képekre

Hagyományos megoldások:

- ▶ A kép szürkeárnyalatossá alakítása.
- ▶ A színcsatornák elemzése külön-külön.

Újabb módszer az utóbbi évtizedben:

Értelmezzük az  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  képet kvaternió értékű függvényként:

$$f(x, y) = \mathbf{i}f_R(x, y) + \mathbf{j}f_G(x, y) + \mathbf{k}f_B(x, y)$$

Különböző momentumok általánosítása kvaterniókra:

QFMM (Fourier-Mellin), QG-CHFM (Csebisev-Fourier), QG-PJFM (Jacobi-Fourier), QBFM (Bessel-Fourier),  
**QZM (Zernike)**

# Kvaternió Zernike momentumok

Zernike függvények általánosítása kvaterniókra:

$$\Phi_{n,m}(r, \theta) = R_{n,m}(r)e^{-\mu m\theta},$$

ahol  $\mu$  egység hosszú, tiszta kvaternió (általában  $\mu = \frac{\mathbf{i}+\mathbf{j}+\mathbf{k}}{\sqrt{3}}$ ).

Mivel a kvaterniók szorzása nem kommutatív, így jobb- és baloldali momentumok is definiálhatók:

$$Z_{n,m}^R(f) = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} f(r, \theta) \Phi_{n,m}(r, \theta) dr d\theta,$$

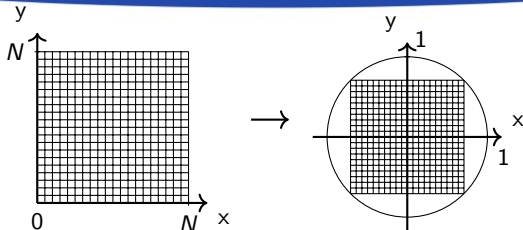
$$Z_{n,m}^L(f) = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} \Phi_{n,m}(r, \theta) f(r, \theta) dr d\theta.$$

# Invariánsok konstrukciója

Chen et al. Quaternion zernike moments and their invariants for color image analysis and object recognition. *Signal Processing*, 2012.

Forgatás, skálázás és eltolás kombinált invariánsok ( $\overline{\Psi}_{n,k}^m$ )  
konstruálása.

# Korábbi diszkretizációs módszer



- Kép transzformálása az egységkörbe:

$$r_{x,y} = \sqrt{(c_1x + c_2)^2 + (c_1y + c_2)^2}, \quad \theta_{x,y} = \tan^{-1} \left( \frac{c_1y + c_2}{c_1x + c_2} \right),$$

ahol  $c_1 = \frac{\sqrt{2}}{N-1}$  és  $c_2 = -\frac{1}{\sqrt{2}}$ .

- Ekkor a pixelek helyét alappontoknak választva:

$$Z_{n,m}^R(f) \approx \frac{2(n+1)}{\pi(N-1)^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \Phi_{n,m}(r_{x,y}, \theta_{x,y}).$$

# Új diszkretizációs módszer

Probléma a korábbi diszkretizációval: nincs diszkrét ortogonalitás.

Schipp F. és Pap M.: diszkrét ortogonális pontrendszer konstrukciója a klasszikus (komplex értékű) Zernike függvényekhez.

Ennek az ötletnek a kvaternió értékű Zernike függvényekre való kiterjesztése megfelelő pontrendszert ad.

Legyen  $N$  pozitív egész és  $\rho_{k,N}$  az  $N$ -edik Legendre polinom gyökei, ekkor a pontrendszer:

$$(r_{k,N}, \theta_{j,N}) = \left( \sqrt{\frac{1 + \rho_{k,N}}{2}}, \frac{2\pi j}{4N} \right), \quad (k = 1, \dots, N, j = 1, \dots, 4N).$$

# Új diszkretizációs módszer (folyt.)

Legyen

$$\mathcal{A}_{k,N} = \int_{-1}^1 \ell_{k,N}(x) dx, \quad (k = 1, \dots, N),$$

ahol  $\ell_{k,N}$  a Lagrange interpolációs alappolinomok a  $\rho_{k,N}$  pontokon.  
Ekkor a  $w(r_{k,N}, \theta_{j,N}) = \frac{\mathcal{A}_{k,N}}{8N}$  súlyokkal véve az integrálközelítést:

$$\frac{1}{\pi} \int_0^1 \int_0^{2\pi} f(r, \theta) d\theta dr \approx \int_{X_N} f = \sum_{k=1}^N \sum_{j=1}^{4N} f(r_{k,N}, \theta_{j,N}) \frac{\mathcal{A}_{k,N}}{8N}.$$

Azaz a QZM-ek közelíthetők a következő módon:

$$Z_{n,m}^R(f) \approx (n+1) \sum_{k=1}^N \sum_{j=1}^{4N} f(r_{k,N}, \theta_{j,N}) \Phi_{n,m}(r_{k,N}, \theta_{j,N}) \frac{\mathcal{A}_{k,N}}{8N}.$$



# Diszkrét ortogonalitás

## Tétel (Diszkrét ortogonalitás)

*Legyen  $n, n' \in \mathbb{N}$  természetes számok,  $m, m' \in \mathbb{Z}$  egészek, úgy, hogy teljesül*

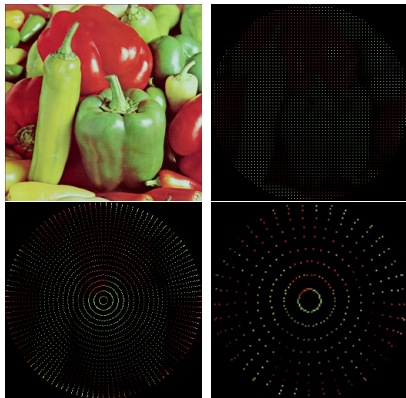
$$\frac{n + n'}{2} + \min(|m|, |m'|) < 2N.$$

*Ekkor*

$$(n + 1) \int_{X_N} \Phi_{n,m} \Phi_{n',m'}^* = \delta_{n,n'} \delta_{m,m'}.$$

Így a momentumok diszkrétizációs hiba nélkül előállíthatók és a képek visszaállítási pontossága és módszer hibatűrése is javult.

# Képpontok becslése



A képet lineárisan  
transzformáljuk az egységgörré.

A függvény értékének becslése a  
pontokban:

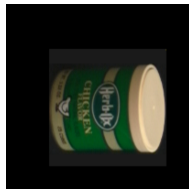
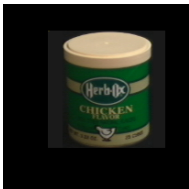
- ▶ Lineáris interpoláció (sok pont esetén)
- ▶ Diszkrét integrálás a pontok környezetében (kevés pont esetén)

# Gyakorlati tesztek

A következő tesztek alapján hasonlítottuk össze a régi és az új módszert:

- ▶ Invariancia teszt
- ▶ Kép visszaállítása momentumokból
- ▶ Transzformált, zajos képek felismerése

A tesztek során a Columbia Object Image Library és az Amsterdam Library of Object Images képeiből generált transzformált képeket használtuk.



# Invariancia

Alacsony rendű momentumok modulusának, az összes transzformált képekre vett variációs koefficiense  $\left(\frac{\sigma}{\mu}\right)$ :

	$\frac{\sigma}{\mu}$
$ \bar{\Psi}_{1,1}^1 $	3.73%
$ \bar{\Psi}_{2,0}^0 $	0.028%
$ \bar{\Psi}_{2,2}^0 $	0.057%
$ \bar{\Psi}_{2,2}^2 $	6.87%
$ \bar{\Psi}_{3,1}^1 $	3.71%
$ \bar{\Psi}_{3,3}^1 $	3.69%
$ \bar{\Psi}_{3,3}^3 $	9.40%

Table: Régi módszer

	$\frac{\sigma}{\mu}$
$ \bar{\Psi}_{1,1}^1 $	3.72%
$ \bar{\Psi}_{2,0}^0 $	0.028%
$ \bar{\Psi}_{2,2}^0 $	0.056%
$ \bar{\Psi}_{2,2}^2 $	6.82%
$ \bar{\Psi}_{3,1}^1 $	3.70%
$ \bar{\Psi}_{3,3}^1 $	3.68%
$ \bar{\Psi}_{3,3}^3 $	9.32%

Table: Új módszer

# Képek visszaállítása




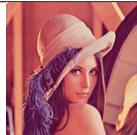
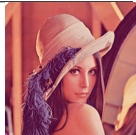








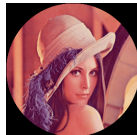

Egy kép rekonstruálható véges számú momentumot használva a következő képlet szerint:

$$f(x, y) \approx \sum_{n=0}^M \sum_{m=-n}^n Z_{n,m}^R(f) \Phi_{n,m}^*(r_{x,y}, \theta_{x,y}).$$

Ha  $f$  az eredeti,  $\hat{f}$  a visszaállított kép, akkor a hiba (mean square error):

$$\epsilon^2 = \frac{\sum_{x=1}^N \sum_{y=1}^N |f(x, y) - \hat{f}(x, y)|^2}{\sum_{x=1}^N \sum_{y=1}^N |f(x, y)|^2}.$$

# Képek visszaállítása

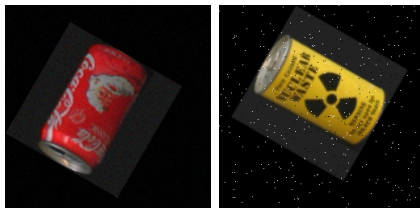
	Eredeti	50	150	250	350
QZM – Régi					
$\epsilon^2$		0.02659	0.00868	0.00428	0.00279
QZM – Új					
$\epsilon^2$		0.01611	0.00461	0.00164	0.00066
QRHFM					
$\epsilon^2$		0.01088	0.00351	0.00150	0.00092

# Képfelismerés

Cél: a transzformált, zajos képet felismerni az eredeti képek közül.  
Különböző nagyságú és típusú zaj:

- ▶ Gauss zaj
- ▶ Só-bors zaj

Alacsony rendű invariáns momentumok (kvaterniók)  
komponenseiből kinyert valós értékek vektorként kezelése.  
Osztályozás legkisebb euklideszi távolság alapján.



# Képfelismerés - Gauss zaj

$\sigma$	Régi (%)	Új – "sok" pont (%)	Új – "kevés" pont (%)
Nincs zaj	99.06	99.15	98.21
1	98.98	99.49	98.81
2	98.98	99.74	98.81
3	98.55	99.83	98.04
5	95.15	99.49	94.64
7	95.15	98.72	91.67
9	76.87	98.47	89.20
40	52.89	88.52	51.87
50	48.21	84.10	45.07
60	41.58	85.80	39.12



# Képfelismerés - Só-bors zaj

p	Régi (%)	Új – "sok" pont (%)	Új – "kevés" pont (%)
Nincs zaj	99.06	99.15	98.21
0.2%	99.66	99.32	94.98
0.4%	99.91	99.74	99.15
0.6%	99.91	99.91	99.40
1%	98.98	99.91	99.66
2%	99.66	93.96	99.74
3%	99.40	99.40	96.34
5%	97.87	94.90	97.87
10%	99.91	93.03	98.72
15%	99.91	93.20	97.87

# Forgatás invariáns Zernike és Fourier momentumok

<b>Zaj típusa</b>		<b>Zernike (%)</b>	<b>Fourier (%)</b>
Nincs zaj		100.00	100.00
Gauss	80	98.12	97.52
	90	99.01	97.02
	100	98.91	96.63
	110	97.62	96.92
	120	96.23	96.63
Só-bors	10%	100.00	100.00
	15%	100.00	100.00
	20%	100.00	99.70
	25%	100.00	100.00
	30%	100.00	100.00

A momentumok sok területen alkalmazhatók, néhány példa:

- ▶ **Vízjelek:** a képre a momentumok szintjén helyeznek vízjelet, így ellenálló lesz különböző transzformációknak és zajnak. Ehhez fontos a rekonstrukciós pontosság.
- ▶ **Neurális hálók, gépi tanulás:** A képekből kinyert invariáns momentumok lehet a bemeneti vektor része. A neurális hálók ismert hiányossága a zajra való érzékenység, ezen segíthetnek a momentumok.
- ▶ **Orvosi/optikai alkalmazások:** lencsék leképezési hibáinak azonosítására, szaruhártya vizsgálatára, stb.

# További lehetőségek

- ▶ Meglévő alkalmazások továbbfejlesztése az új módszer szerint.
- ▶ Más függvényrendszeren alapuló momentumokra hasonló konstrukció megadása.
- ▶ További általánosítás 3-dimenzióra és alkalmazás például LiDAR pontfelhőkre.

- ▶ Diszkrét ortogonális pontrendszer konstruálása QZM-hez.
- ▶ Összehasonlítás a korábbi módszerrel.
  - ▶ Invariancia
  - ▶ Rekonstrukció
  - ▶ Képfelismerés
- ▶ Jelentősen jobb rekonstrukciós képesség és robusztusság különféle zajokkal szemben.

# KÖSZÖNÖM A FIGYELMET!

**SZÉCHENYI** 



HUNGARIAN  
GOVERNMENT

European Union  
European Social  
Fund



**INVESTING IN YOUR FUTURE**