

Color image analysis and recognition using discrete orthogonal quaternion Zernike moments

Zsolt Németh and Gergely Nagy

Eötvös Loránd University, Faculty of Informatics

13th Joint Conference on Mathematics and
Computer Science (MaCS 2020)
Budapest, October 1–3, 2020

EFOP-3.6.3-VEKOP-16-2017-00001

SZÉCHENYI 2020



HUNGARIAN
GOVERNMENT

European Union
European Social
Fund



INVESTING IN YOUR FUTURE

Image moments

The term "image moment" usually refers to some numerical descriptor of an image, computed directly from pixel intensities as a discretization of a weighted integral.

A well-known example: the geometric moments

$$M_{pq}(I) = \int_0^1 \int_0^1 x^p y^q I(x, y) \, dx dy \approx \sum_i \sum_j x_i^p y_j^q I(x_i, y_j).$$

Zernike moments: Fourier-type moments w.r.t. the Zernike function system over the unit disk \mathbb{D} .

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

where $R_{n,m}(r)$ are an orthogonal system of *radial* polynomials.

Image moments for multiple channels

Classical approach:

- ▶ Conversion of the input image to a single channel (i.e. grayscale) one.
- ▶ Analyze separately for each channel, merge the results later on.

An improvement for RGB over the past decade: interpret the three-channel $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ image as a quaternion-valued function:

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

The idea is to allow the common treatment of color channels at any point over the analysis using simple but general tools.

Examples include QFMM (Fourier–Mellin), QG-CHFM (Chebysev–Fourier), QG-PJFM (Jacobi–Fourier), **QZM (Zernike)**

Quaternion Zernike moments

Generalization of Zernike functions to quaternions:

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

where μ is a pure unit quaternion (a usual choice is $\mu = \frac{i+j+k}{\sqrt{3}}$). Quaternion multiplication is generally not commutative; so we define left- and right-side moments:

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

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

Chen et al. (2012): extraction of features ($\bar{\Psi}_{n,k}^m$) from Zernike moments invariant under image translation, scaling and *rotation*.

Zernike moment invariants

Translation: obtained by translating the origin to the centroid of the image, computed from the three channel centroids $\left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right)$.

Rotation: for any $m \in \mathbb{Z}$ and $n, k \in \mathbb{N}$, the quaternions

$$\Phi_{n,k}^m = Z_{n,m}^R(f) Z_{k,-m}^L(f) = -Z_{n,m}^R(f) (Z_{k,m}^R(f))^*$$

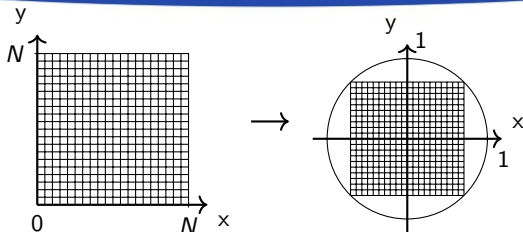
are invariant under image rotation, i.e. it is the same value for functions f and $f'(r, \theta) = f(r, \theta - \alpha)$, irrespective of α .

Scaling: certain linear combinations of Zernike moments are scaling invariant, i.e. for non-negative integers m and l , the quaternions

$$L_{m+2l,m}^R(f) = \sum_{t=0}^l \sum_{k=t}^l \left(\sqrt{|Z_{0,0}^R(f)|} \right)^{-(m+2k+2)} c_{m,l}^{t,k} Z_{m+2t,m}^R(f)$$

are invariant under image scaling.

Previous approaches to discretization



- Transformation of the image inside the unit disk:

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

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

- Now simply use the new pixel positions for discretization:

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

Novel discretization method

Problem with the previous approach: no discrete orthogonality -> losing numeric precision, low robustness to noise, worse compression ratio.

F. Schipp F. and M. Pap (2005): construction of a point system and a corresponding discrete integral for classical (complex valued) Zernike functions.

We extended this construction for their quaternion-valued counterparts.

Short recap: For a positive integer N , let $\rho_{k,N}$ denote the roots of the Legendre polynomial of degree N . With these, we can define the polar point system

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

Novel discretization method (cont.)

Now let

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

denote the Christoffel numbers w.r.t the Lagrange base polynomials $\ell_{k,N}$ over $\rho_{k,N}$. We introduce a discrete integral w.r.t. weights $w(r_{k,N}, \theta_{j,N}) = \frac{\mathcal{A}_{k,N}}{8N}$ as

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

Specifically, for QZMs we obtain the approximation

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

Discrete orthogonality

Theorem (Discrete orthogonality)

Suppose that natural numbers $n, n' \in \mathbb{N}$ and integers $m, m' \in \mathbb{Z}$ satisfy

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

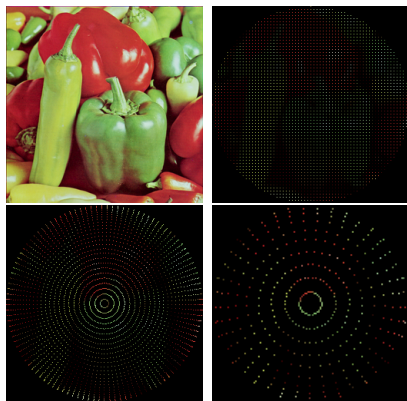
Then we have

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

Now, approximation of the moments can be performed without introducing further discretization errors.

Additionally, we also obtain the convergence of the discrete integral to the continuous one as $N \rightarrow +\infty$.

Pixel transformation



First, we linearly transform the image to the unit disk.

Second, resample the pixel data for the novel system of points:

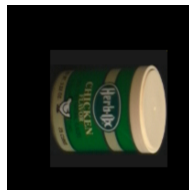
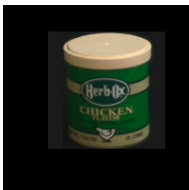
- ▶ Linear interpolation (advised for \approx the same number of points)
- ▶ Discrete integration locally over some neighbourhood (for a sparse system of points)

Tests

The previous and the novel methods were compared based on the following tests:

- ▶ Invariance test
- ▶ Image reconstruction from the moments
- ▶ Recognition of transformed, noisy images

For these, we generated a dataset of transformed images using images from the Columbia Object Image Library and the Amsterdam Library of Object Images



Invariance

The variation coefficient $\left(\frac{\sigma}{\mu}\right)$ of the modulus of low order moments over the entire transformed image set:

	$\frac{\sigma}{\mu}$
$ \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: Previous method

	$\frac{\sigma}{\mu}$
$ \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: Novel method

Image reconstruction

An image can be reconstructed using a finite number of moments by the following formula:

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

Let f be the original, and \hat{f} the reconstructed image. The mean square error of the reconstruction is given by:

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

Image reconstruction





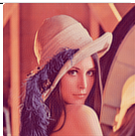





	Original	50	100	150	250
Previous					
ϵ^2		0.02659	0.1341	0.00868	0.00428
Novel					
ϵ^2		0.01611	0.00790	0.00463	0.00190

Image recognition

Goal: correctly recognize the transformed, noisy image from the original set of images.

Different types of noise used for testing:

- ▶ Gaussian noise
- ▶ Salt-and-pepper noise

Real-valued vectors are extracted from low order moments (quaternions). Classification is done by minimum Euclidean distance.

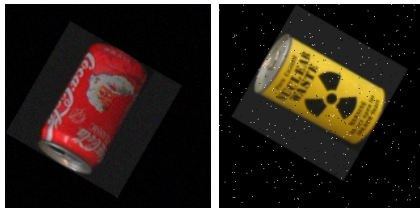


Image recognition - Gaussian noise

σ	Previous (%)	Novel – "many" points (%)	Novel – "few" points (%)
No noise	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

Image recognition - Salt-and-pepper noise

σ	Previous (%)	Novel – "many" points (%)	Novel – "few" points (%)
No noise	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

Applications

Moment-based techniques are applicable in multiple fields, some examples:

- ▶ **Watermarking:** A watermark can be placed on an image at the level of moments, thus it will be robust against different transformations and noise. The accuracy of reconstruction is important for this application.
- ▶ **Neural networks, machine learning:** Invariant moments, extracted from the images, can be part of the input vector. A well-known issue of neural networks is their lack of robustness against noise, moments can help improve this robustness.
- ▶ **Face recognition and template matching:** combined with image segmentation and possibly other feature extraction methods, moment-based techniques excel at these image processing problems.

Future improvement possibilities

- ▶ Improve existing applications based on the proposed method.
- ▶ Apply a similar construction to other moment methods based on various other function systems.
- ▶ Further generalization to 3-dimensions and application to, for example, LiDAR point clouds.

Summary

- ▶ Construction of a discrete orthogonal point system for QZMs.
- ▶ Comparisons with previous methods.
 - ▶ Invariance
 - ▶ Reconstruction
 - ▶ Image recognition
- ▶ Conclusion: the novel method improves image reconstruction capability and provides robustness against various types of noise.

Thank you for your attention!

SZÉCHENYI 



HUNGARIAN
GOVERNMENT

European Union
European Social
Fund



INVESTING IN YOUR FUTURE