Binary Shape Clustering via Zernike Moments

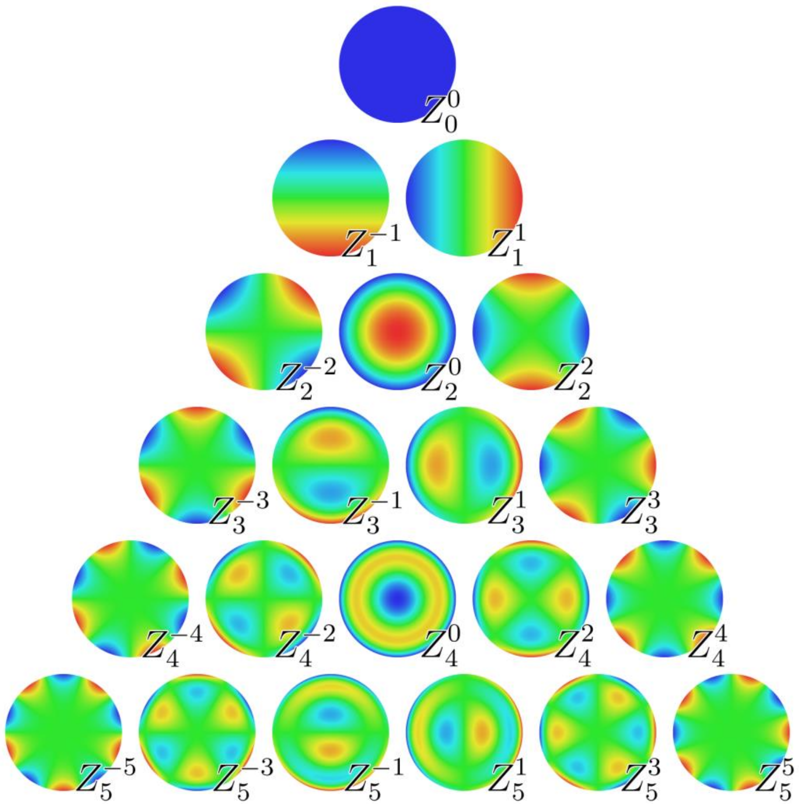
By:

Stephen Yoo

Michael Vorobyov

Moments

- In general, moments describe numeric quantities at some distance from a reference point or axis.



Regular (Cartesian) Moments

- A regular moment has the form of projection of $f(x,y)$ onto the monomial $x^p y^q$

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy$$

Problems of Regular Moments

- The basis set $x^p y^q$ is not orthogonal → The moments contain redundant information.
- As $x^p y^q$ increases rapidly as order increases, high computational precision is needed.
- Image reconstruction is very difficult.

Benefits of Regular Moments

- Simple translational and scale invariant properties

- By preprocessing an image using the regular moments we can get an image to be translational and scale invariant before running Zernike moments

Orthogonal Functions

$$\int_a^b y_m(x) y_n(x) dx = 0 \; ; \; m \neq n$$

- A set of polynomials orthogonal with respect to integration are also orthogonal with respect to summation.

Orthogonal Moments

- Moments produced using orthogonal basis sets.

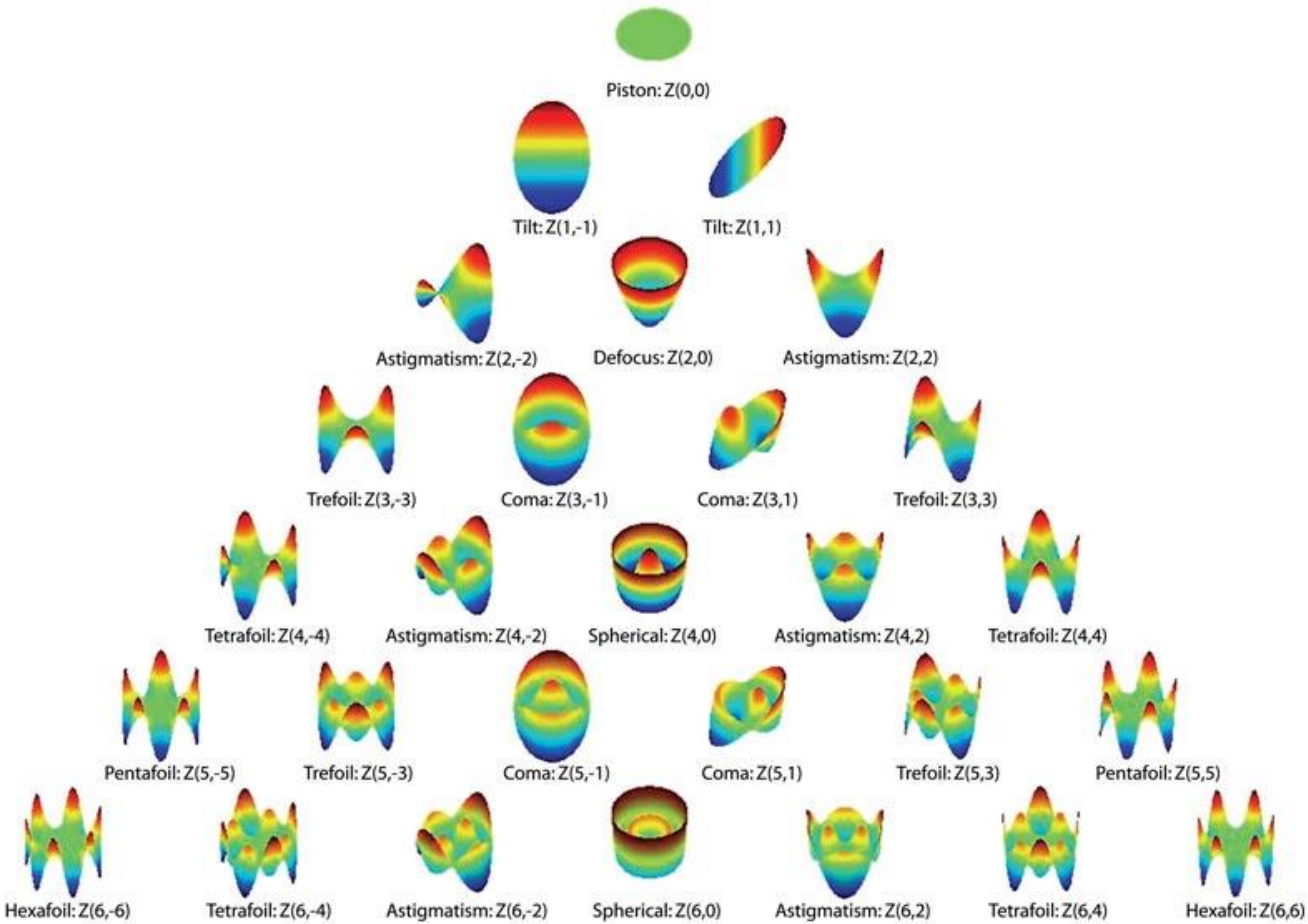
- Require lower computational precision to represent images to the same accuracy as regular moments.

Zernike Polynomials

- Set of orthogonal polynomials defined on the unit disk.

$$V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm\theta}$$

$$R_n^m(\rho) = \begin{cases} \displaystyle\sum_{l=0}^{(n-m)/2} \frac{(-1)^l\,(n-l)!}{l!\left[\frac{1}{2}(n+m)-l\right]!\left[\frac{1}{2}(n-m)-l\right]!}\,\rho^{n-2l} & \text{for } n-m \text{ even} \\ 0 & \text{for } n-m \text{ odd.} \end{cases}$$


Piston: Z(0,0)
Tilt: Z(1,-1)
Tilt: Z(1,1)
Astigmatism: Z(2,-2)
Defocus: Z(2,0)
Astigmatism: Z(2,2)
Trefoil: Z(3,-3)
Coma: Z(3,-1)
Coma: Z(3,1)
Trefoil: Z(3,3)
Tetrafoil: Z(4,-4)
Astigmatism: Z(4,-2)
Spherical: Z(4,0)
Astigmatism: Z(4,2)
Tetrafoil: Z(4,4)
Pentafoil: Z(5,-5)
Trefoil: Z(5,-3)
Coma: Z(5,-1)
Coma: Z(5,1)
Trefoil: Z(5,3)
Pentafoil: Z(5,5)
Hexafoil: Z(6,-6)
Tetrafoil: Z(6,-4)
Astigmatism: Z(6,-2)
Spherical: Z(6,0)
Astigmatism: Z(6,2)
Tetrafoil: Z(6,4)
Hexafoil: Z(6,6)

Zernike Moments

- Simply the projection of the image function onto these orthogonal basis functions.

$$A_{mn} = \frac{m+1}{\pi} \int_x \int_y f(x,y)[V_{mn}(x,y)]^* \; dx \; dy$$

$$\text{where } x^2 + y^2 \leq \; 1$$

Advantages of Zernike Moments

- Simple rotation invariance
- Higher accuracy for detailed shapes
- Orthogonal
  - Less information redundancy
  - Much better at image reconstruction (vs. normal moments)

Scale and Translational Invariance

Scale: Multiply by the scale factor raised to a certain power

$$m'_{pq} = \alpha^{1+p}\beta^{1+q}m_{pq},\ \ p \neq q$$
$$m'_{pq} = \alpha^{2+p+q}m_{pq},\ \ p = q$$

Translational: Shift image's origin to centroid (computed from normal first order moments)

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

Rotational Invariance

The magnitude of each Zernike moment is invariant under rotation.

Fig. 2. The image of character A and five rotated versions of it. From left to right rotation angles are: 0°, 30°, 60°, 150°, 180°, and 300°.

TABLE II
MAGNITUDES OF SOME OF THE ZERNIKE MOMENTS FOR ROTATED IMAGES SHOWN IN FIG. 2 AND THEIR CORRESPONDING STATISTICS

	$\|A_{20}\|$	$\|A_{22}\|$	$\|A_{31}\|$	$\|A_{33}\|$
0°	439.62	41.79	57.97	172.57
30°	436.70	40.20	63.82	171.96
60°	440.63	40.08	66.28	169.41
150°	438.53	41.55	65.47	170.83
180°	439.01	46.85	62.39	168.47
300°	438.43	39.19	65.77	170.84
μ	438.82	41.61	63.62	170.68
σ	1.32	2.74	3.12	1.53
σ/μ %	0.30	6.57	4.90	0.90

Image Reconstruction

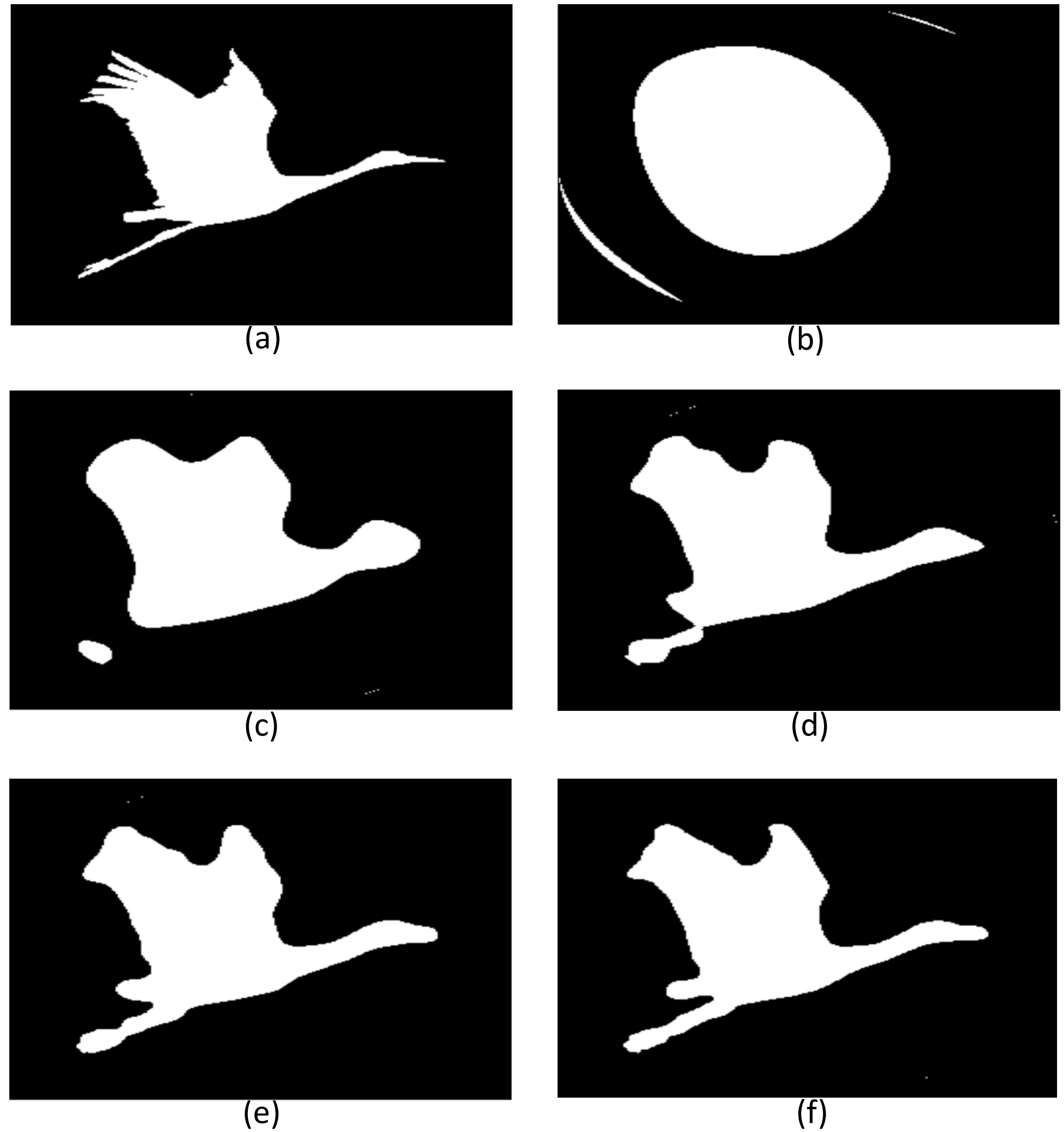
- Orthogonality enables us to determine the individual contribution of each order moment.
- Simple addition of these individual contributions reconstructs the image.

$$\hat{f}(r,\theta) = \sum_{m=0}^{N_{max}} \sum_{n} A_{mn} V_{mn}(r,\theta)$$

Image Reconstruction

Reconstruction of a crane shape via Zernike moments up to order 10k-5,

k = {1,2,3,4,5}.



(a) (b) (c) (d) (e) (f)

Determining Min. Order

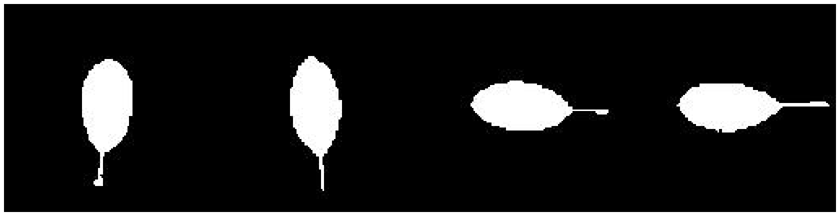
After reconstructing image up to moment i

1. Calculate the Hamming distance, $H(\hat{f}_i, f)$ which is the number of pixels that are different between $\hat{f}_i$ and f.
2. Since, in general, $H(\hat{f}_i, f)$ decreases as i increases, finding the first i for which $H(\hat{f}_i, f) < \epsilon$ will determine the minimum order to reach a predetermined accuracy.

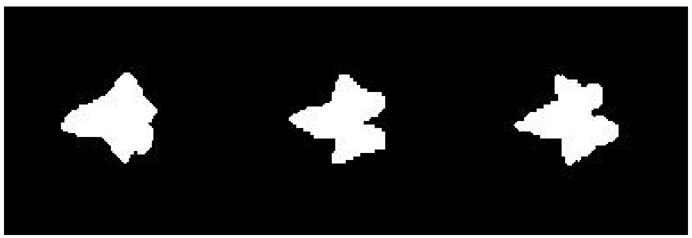
Experimentation & Results

- 40^{th} order moments on 22 binary 128 x 128 images of 7 different leaf types.
- Clustering was done by K-means, with the farthest-first approach for seeding original means.

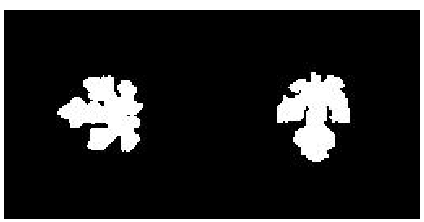
Original Clusters



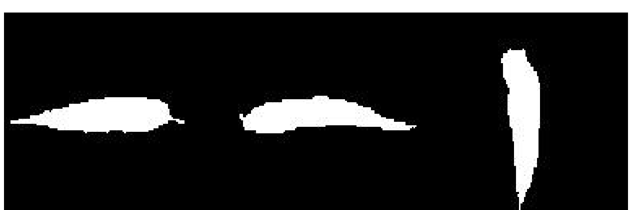
Type 1



Type 2

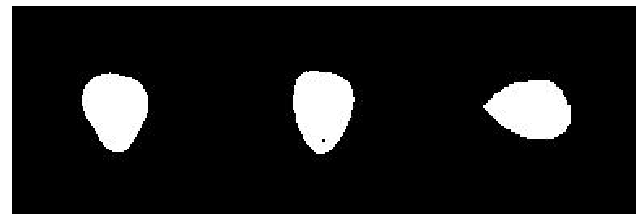


Type 3

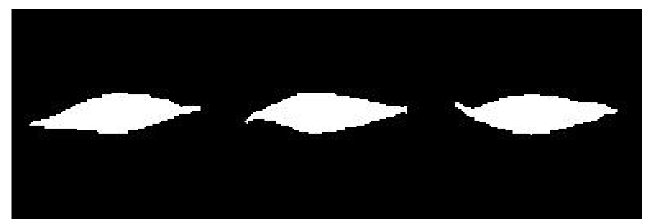


Type 4

Type 5

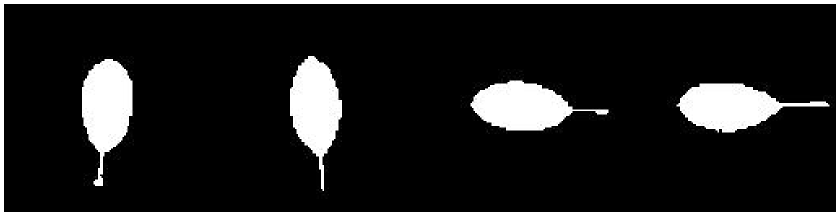


Type 6

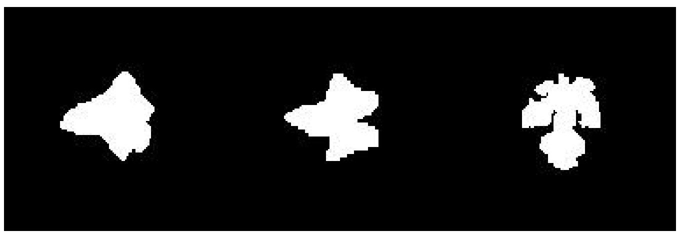


Type 7

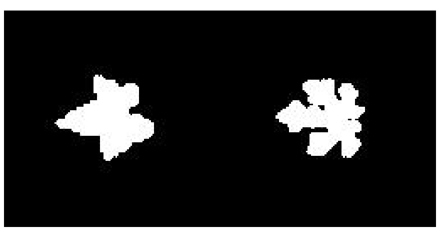
The Zernike Clusters



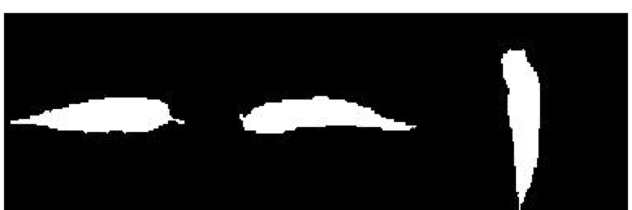
Type 1



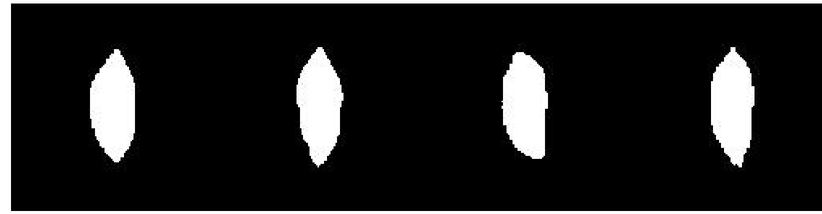
Type 2



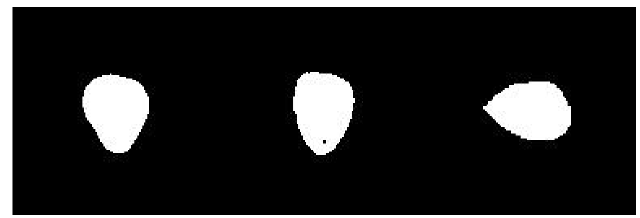
Type 3



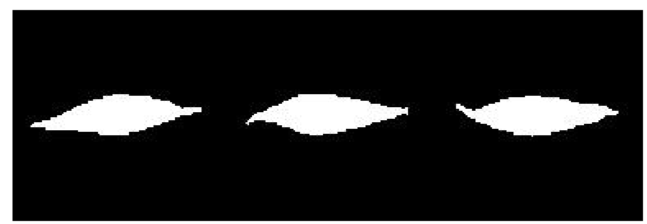
Type 4



Type 5



Type 6



Type 7

Conclusion

- Zernike moments have rotational invariance, and can be made scale and translational invariant, making them suitable for many applications.
- Zernike moments are accurate descriptors even with relatively few data points.
- Reconstruction of Zernike moments can be used to determine the amount of moments necessary to make an accurate descriptor.

Future Research

- Better seeding algorithm for K-means/ different clustering algorithm
- Apply Zernike’s to supervised classification problem
- Make hybrid descriptor which combines Zernike’s and contour curvature to capture shape and boundary information