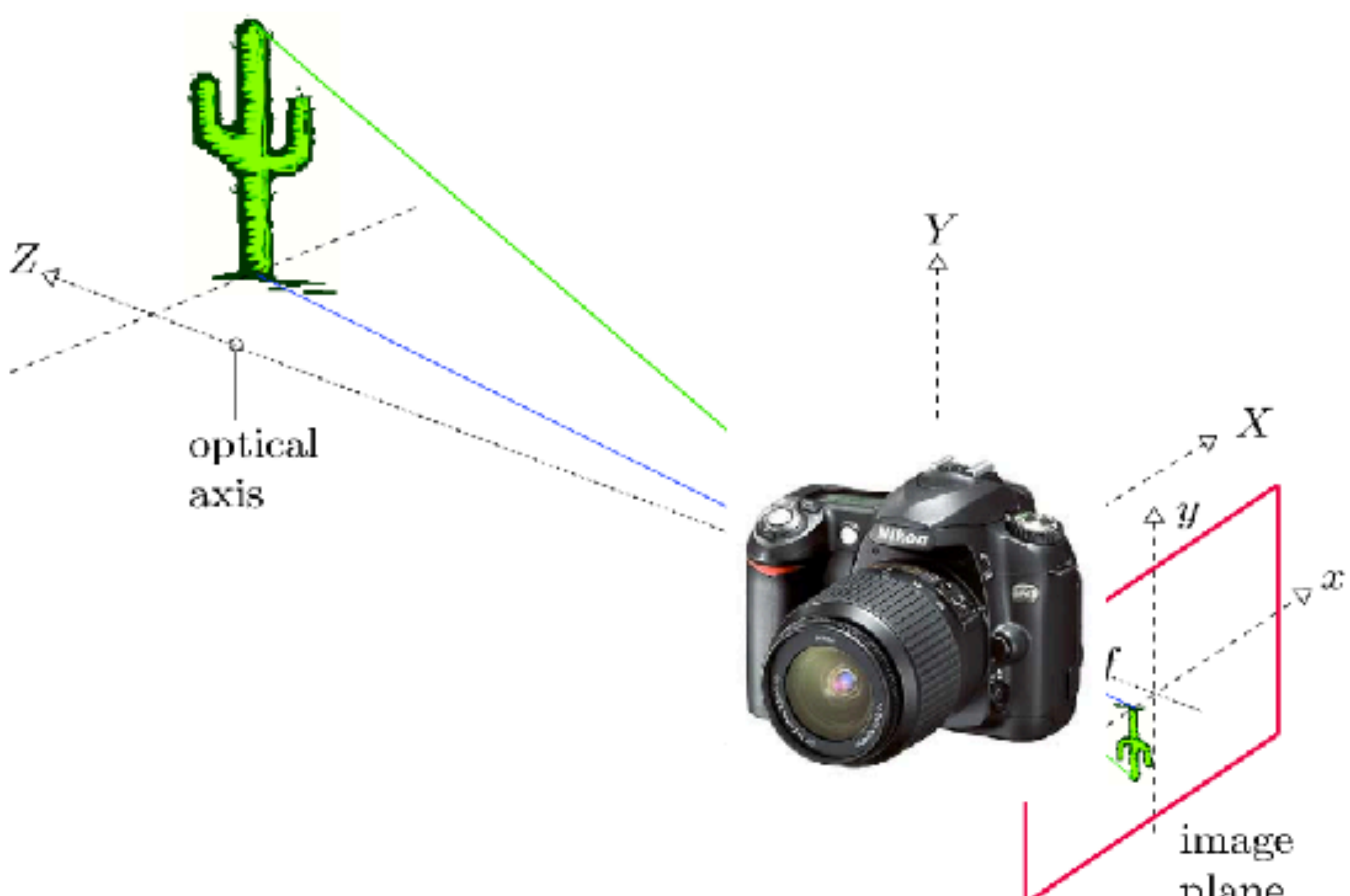
MAURICIO CERDA

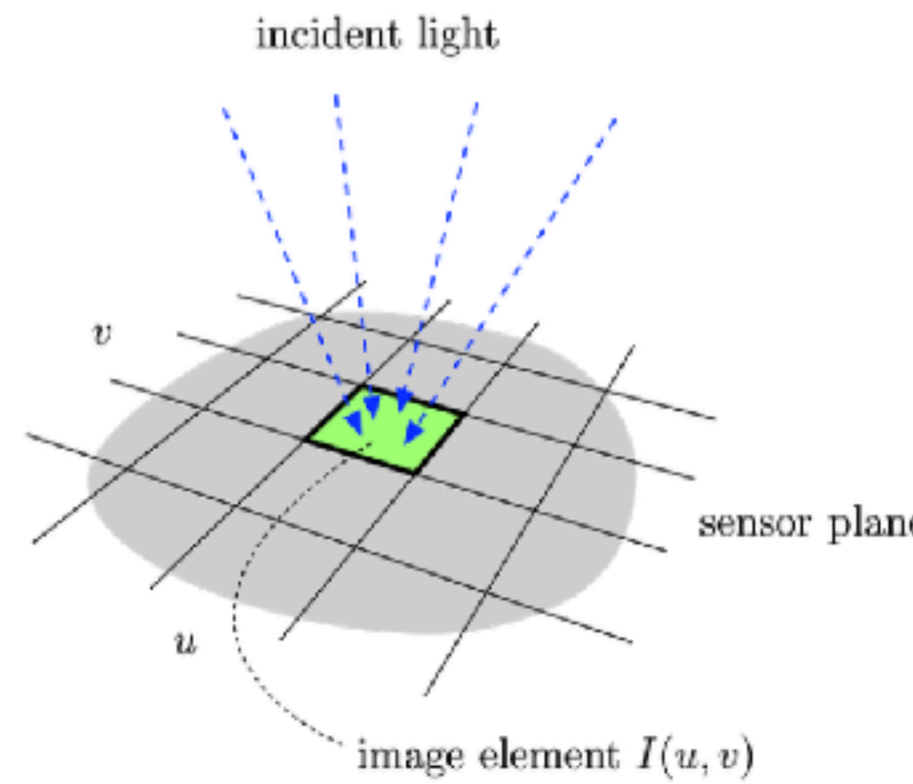
LAB “BIO-RELATED”: IMAGE PROCESSING METHODS FOR MICROSCOPY IMAGING

- La Serena, 8/24/2017 -

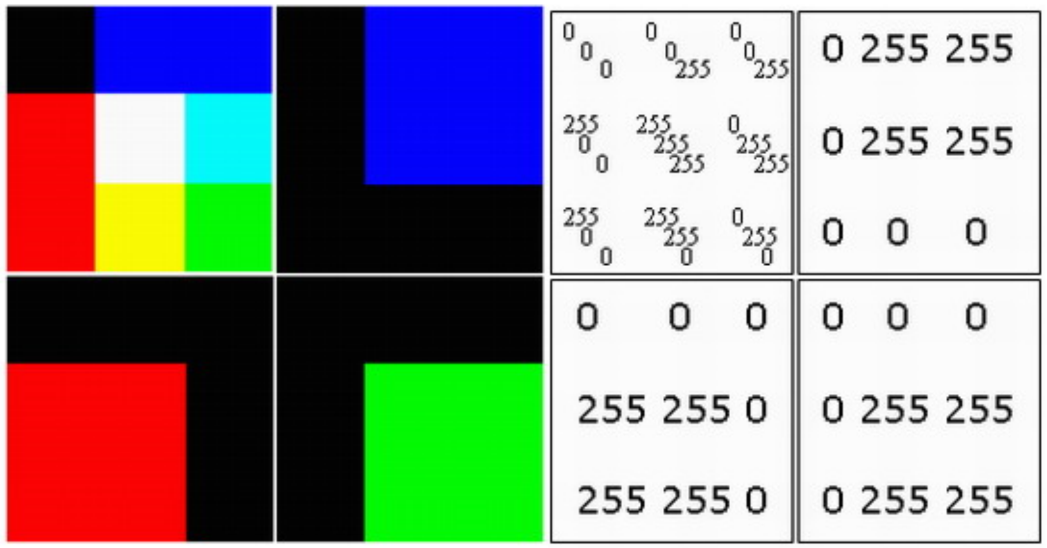
- Image processing?
- Segmentation (clustering)
- Shape description (PCA)
- Lab: challenge !

- **Image:** is an artifact that depicts visual perception, for example, a photo or a two-dimensional picture, that has a similar appearance to some subject—usually a physical object or a person, thus providing a depiction of it [*wikipedia, 2017*].





- Raster image:
    - Array or matrix of pixels with spatial coordinates I(x,y).
    - A numerical value or color per location.



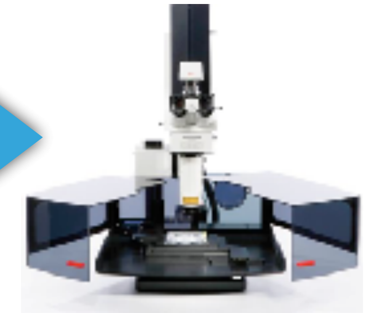
- Vectorial image:
  - Defined by basic functions (points, lines), instead of pixels.
  - To be displayed on screen needs to be rasterized (transforms to a raster image).
  - Raster or Vectorial in a paper?

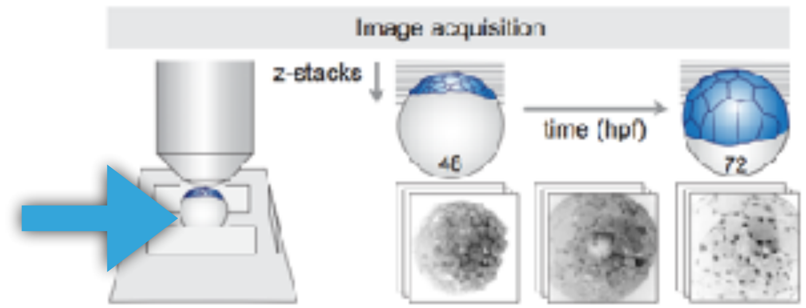
- We want to process images to understand biological systems (models):
  - Tissue development
  - Medical imaging (parasites counting)

IMAGE PROCESSING: TISSUE DEVELOPMENT

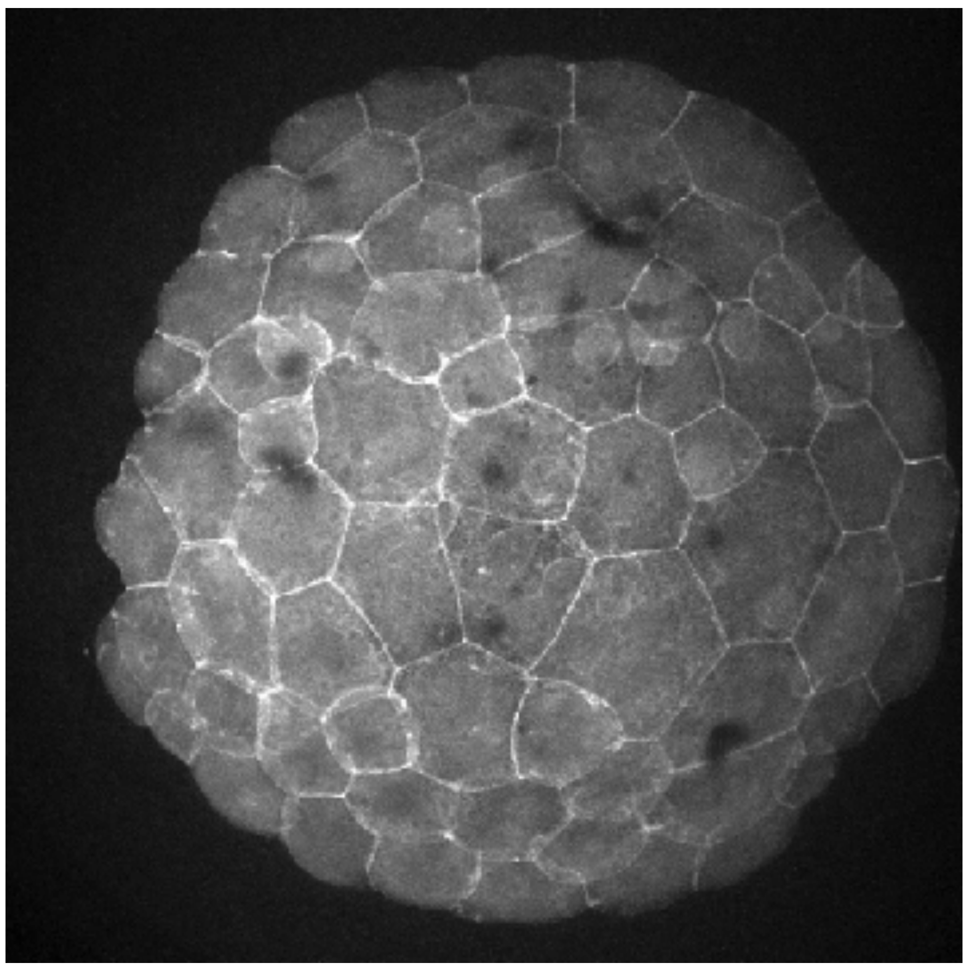
killifish (uruguay)

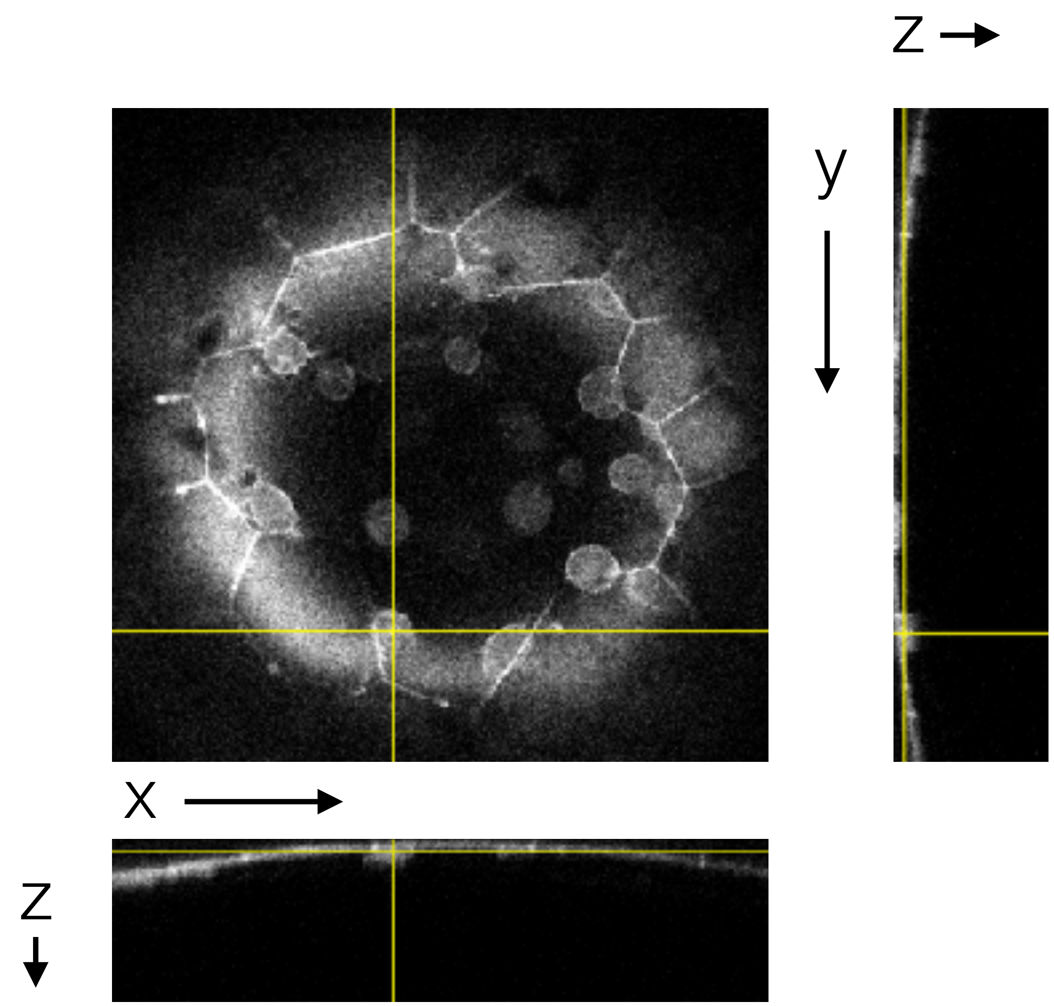
eggs





confocal microscope (U-Chile)





Austrolebias nigripinnis 48-72 HPF, membrane marked with actine EGFP.

- How cells migrate over other cells?

IMAGE PROCESSING: CHAGAS

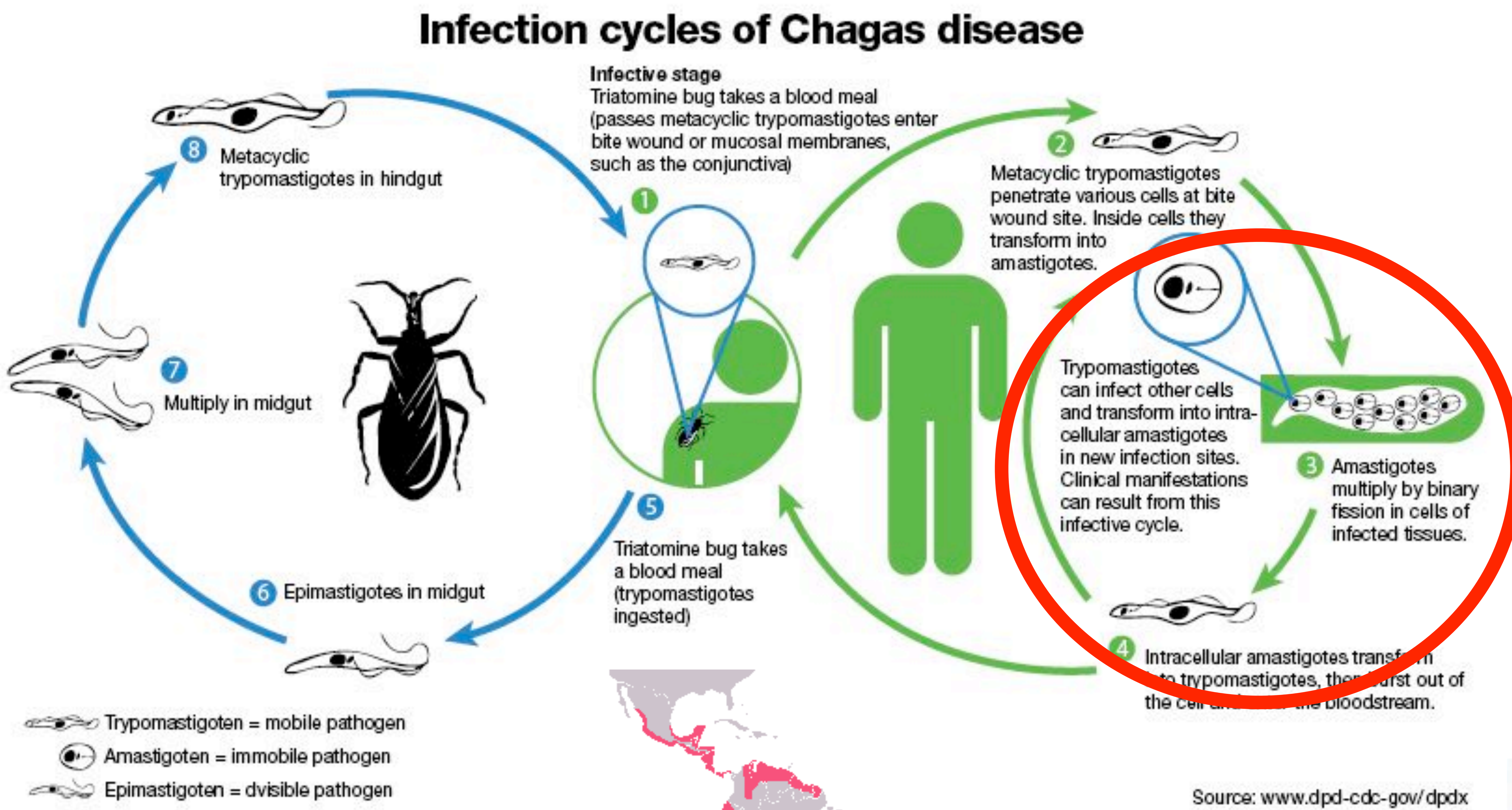
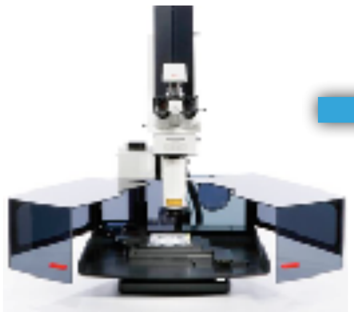


IMAGE PROCESSING: PARASITES



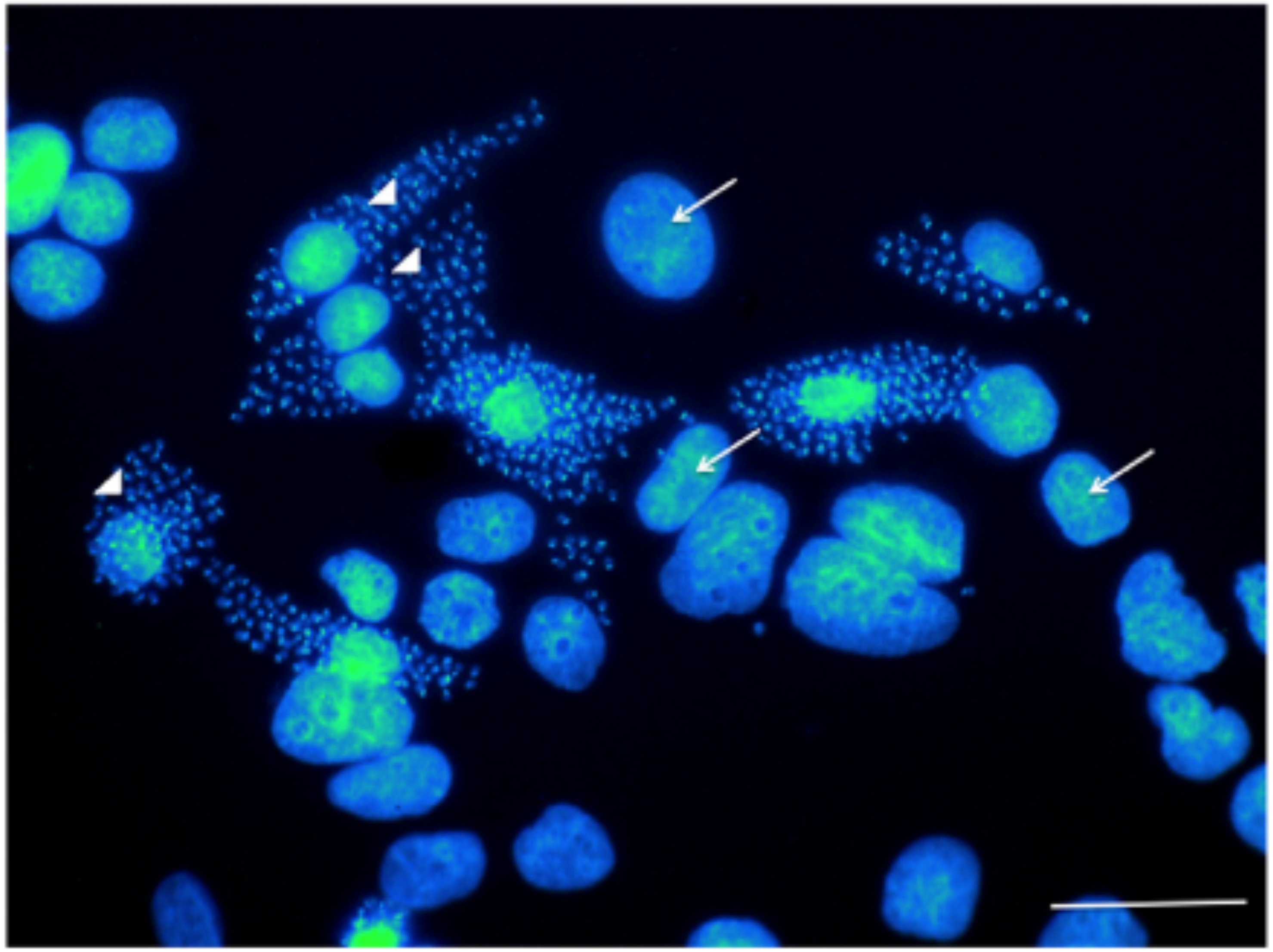
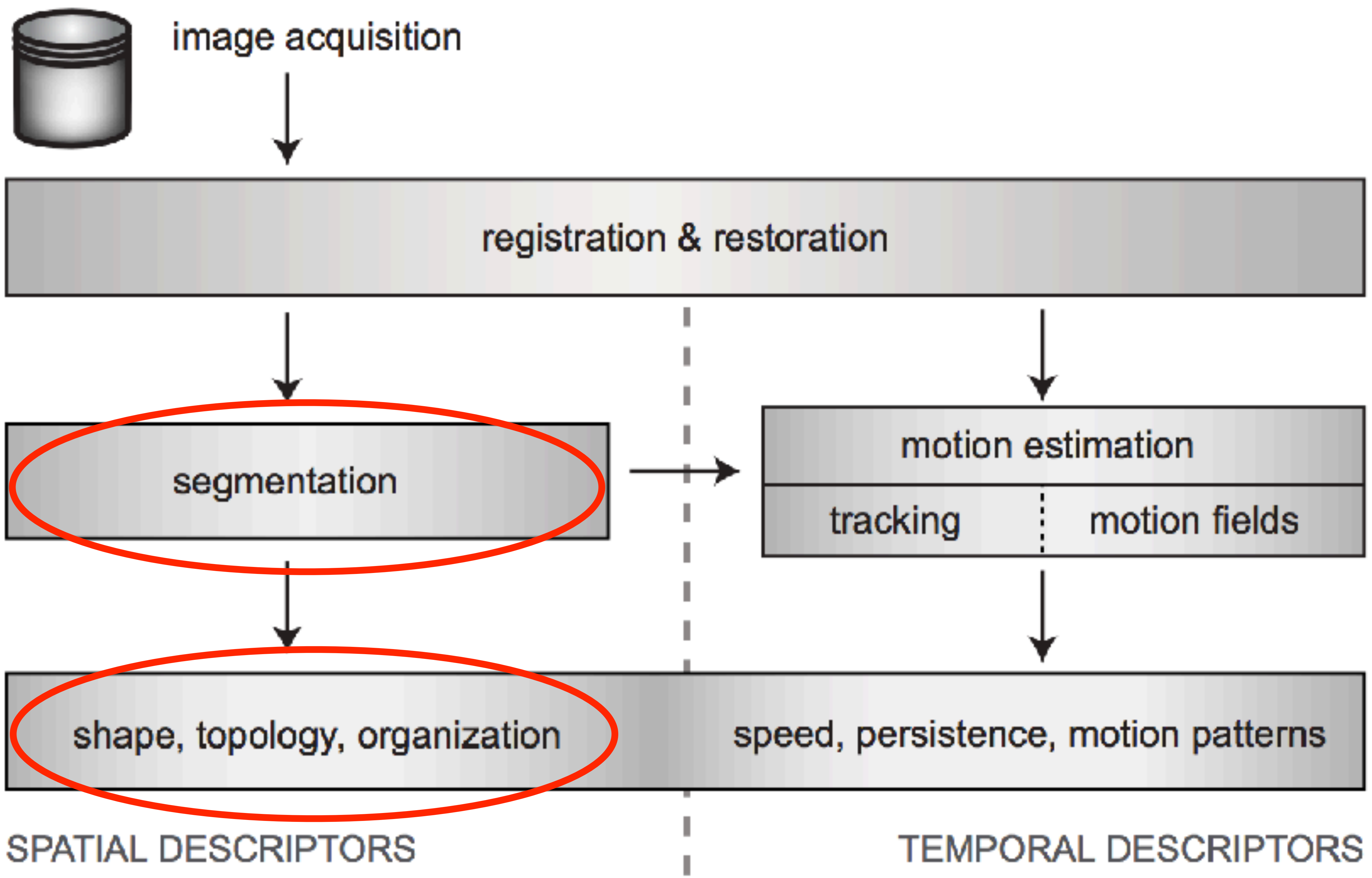


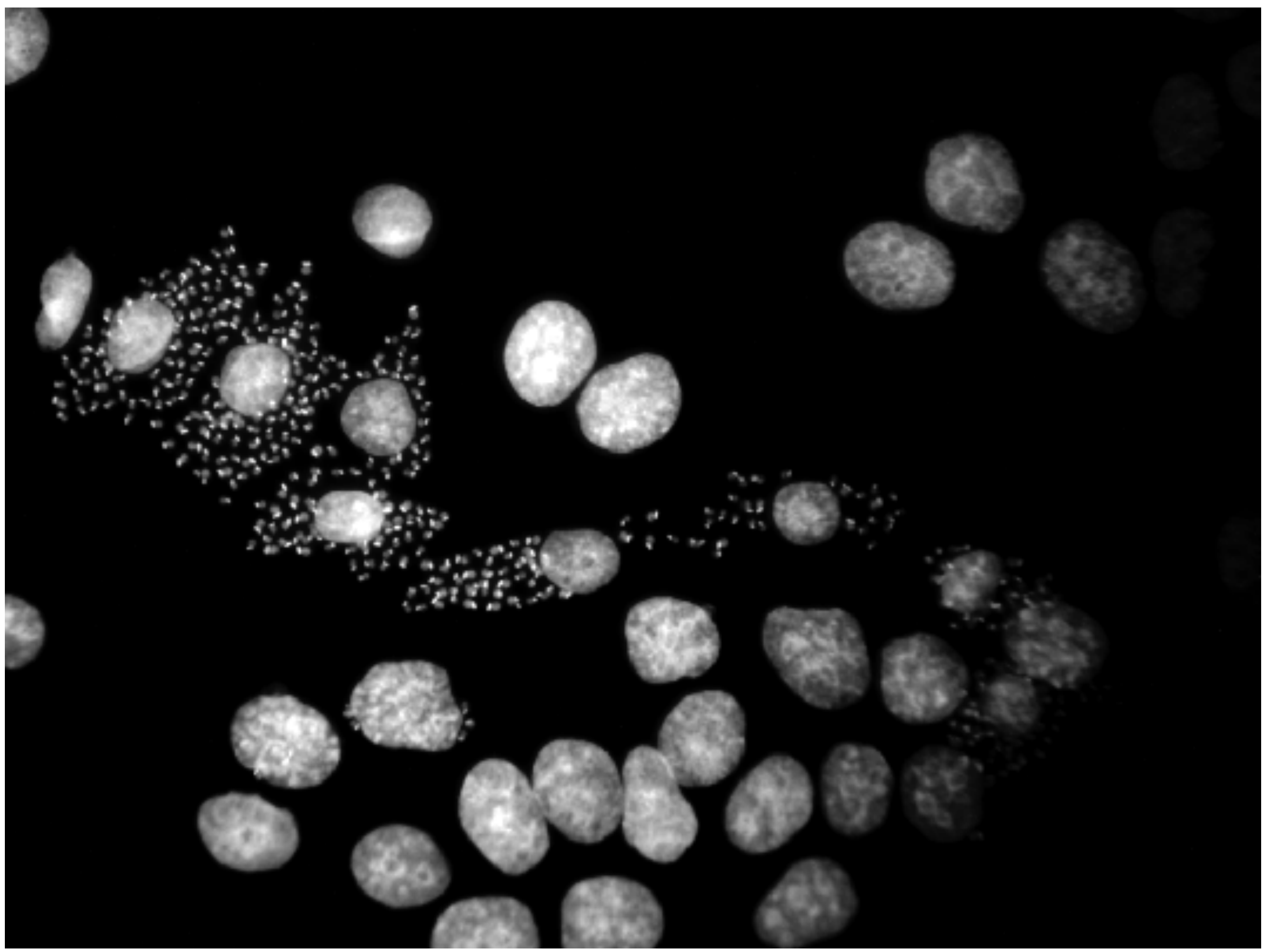
Fig. 1. Infection of BeWo cells with *T. cruzi* amastigotes. BeWo cells were challenged with *T. cruzi* Ypsilon strain trypomastigotes at a parasite:cell ratio of 1:1 for 24 h and were processed for DAPI staining after 48 h. The arrows show BeWo cell nuclei, and the arrowheads show intracellular amastigotes. Scale bar: 10 µm.

- Pregnancy?

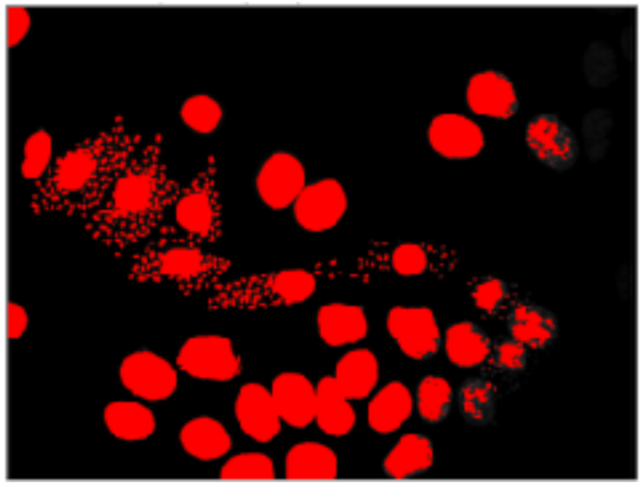
IMAGE PROCESSING: PIPELINE



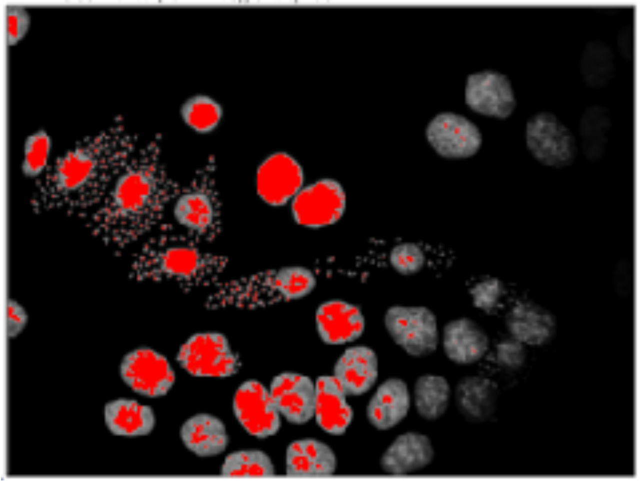
- The simplest segmentation… a manual global threshold [demo FIJI].



raw image

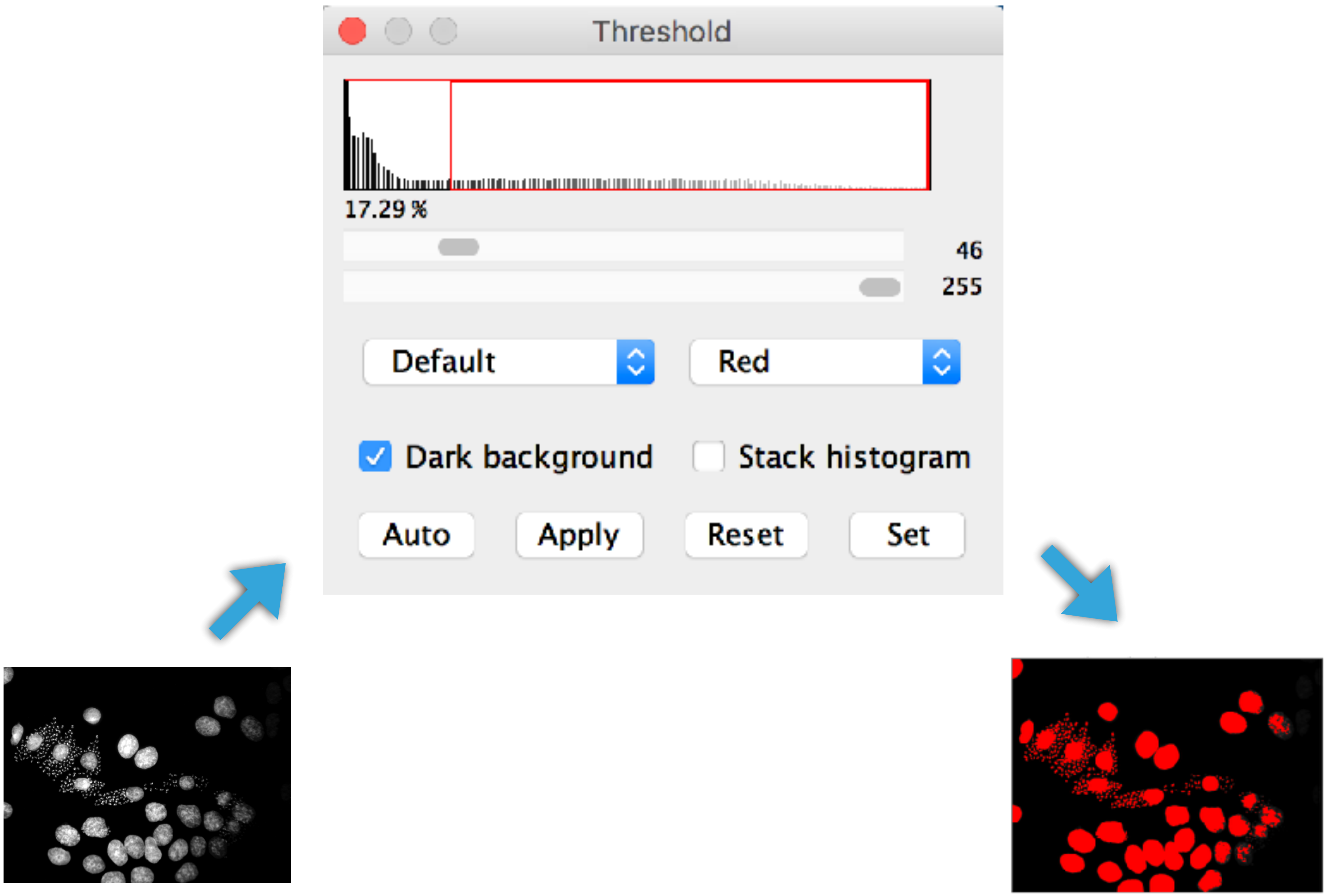


segmentation (>46)



segmentation (>158)

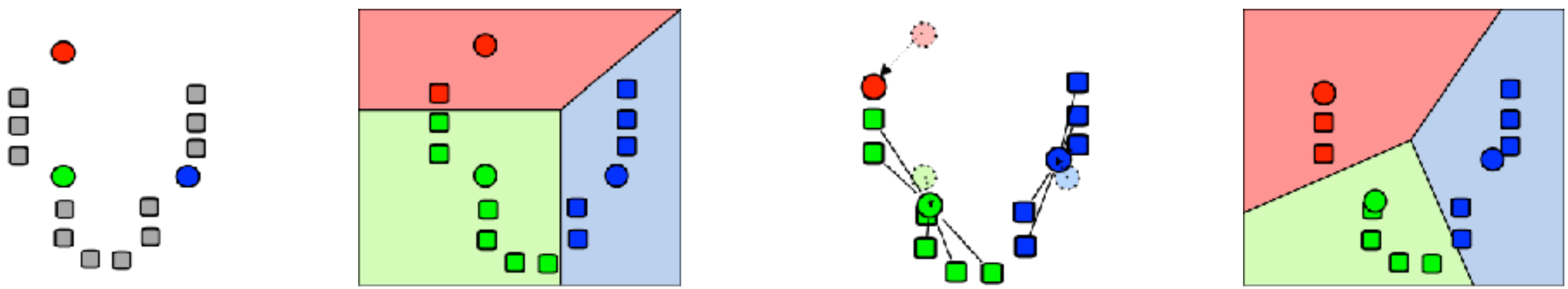
- But, it looks like ? …



- We will start with two objects: cells, and background.
- We don't have examples (!)

- This is another kind of learning problem:
  - Supervised: regression, classification
  - Unsupervised: **clustering**

- We can model it as how to discover the best k groups or clusters at a pixel level.
- K-means clustering (k=3):

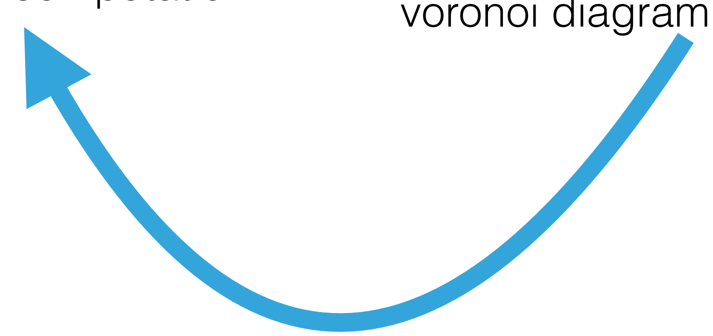


Random centroids

clusters assignation + voronoi diagram

centroids re-computation

cluster assignation + voronoi diagram

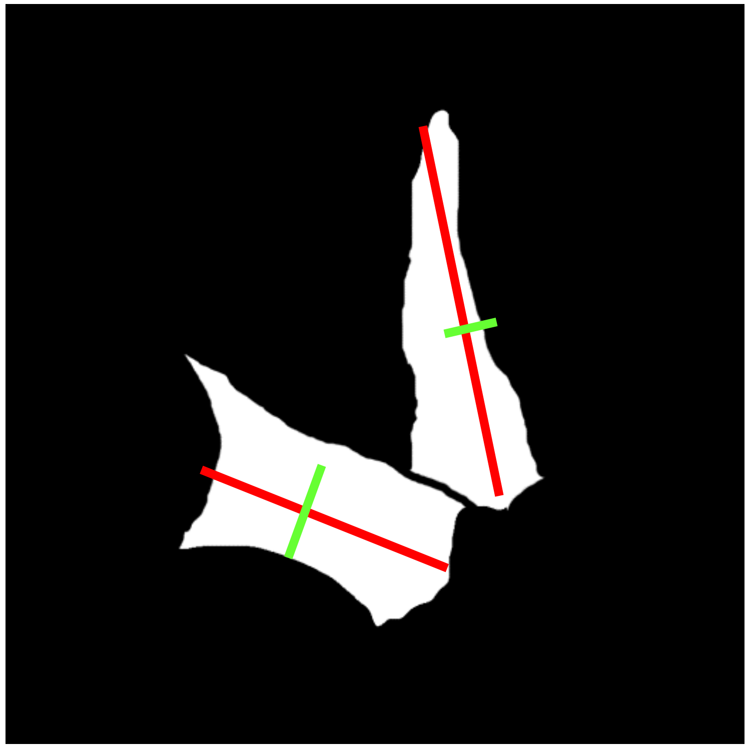


- K-Mean pseudocode [demo FIJI]:

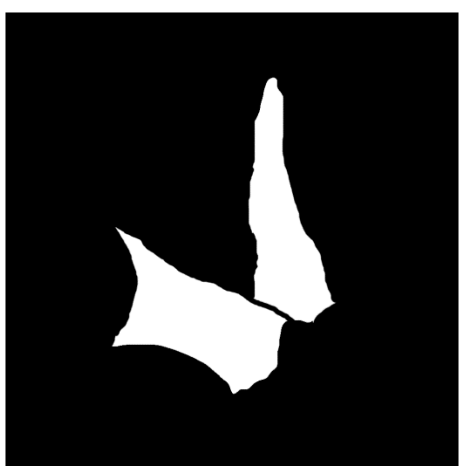
- But we can also make examples! (label data)
- In that case, segmentation may be a supervised problem. Let's try to solve it as a Random Forest problem [demo FIJI].
- How can we decide?

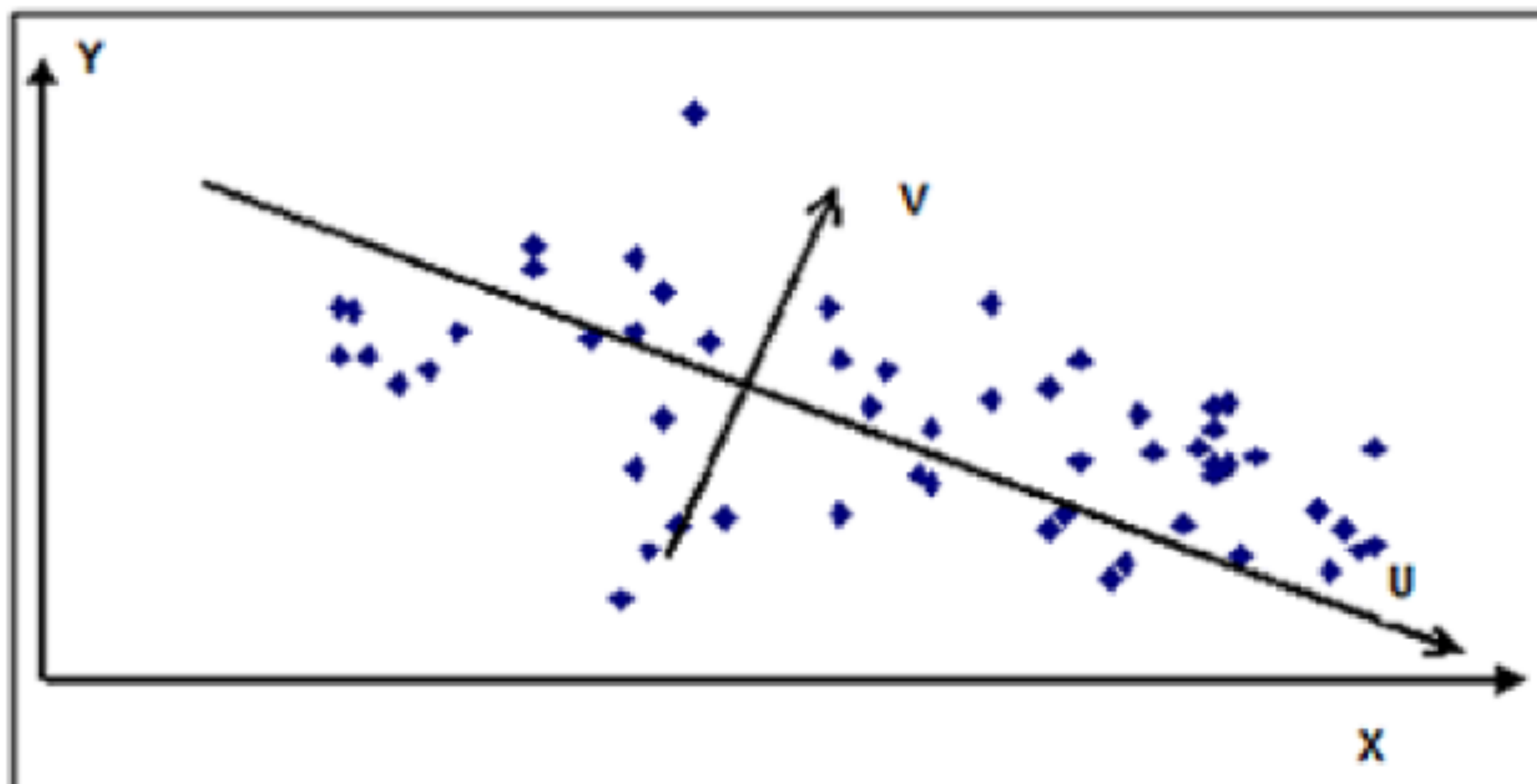
- If images are segmented, we can easily count objects.
- But, we cannot tell the difference between small and big or circle-like vs elongated cells.



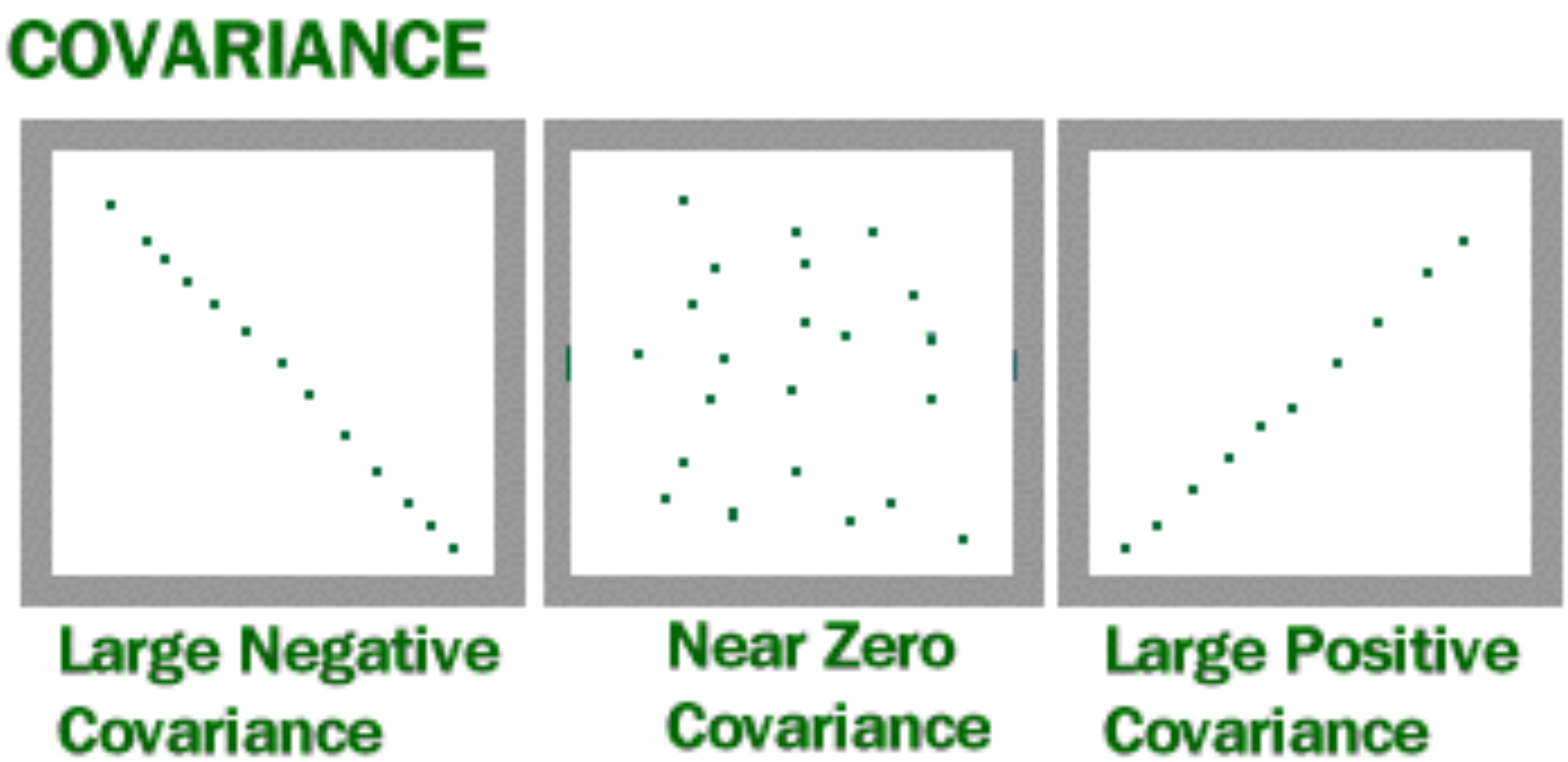


- The image as a set of pixels
- We can always find a direction to maximize variance

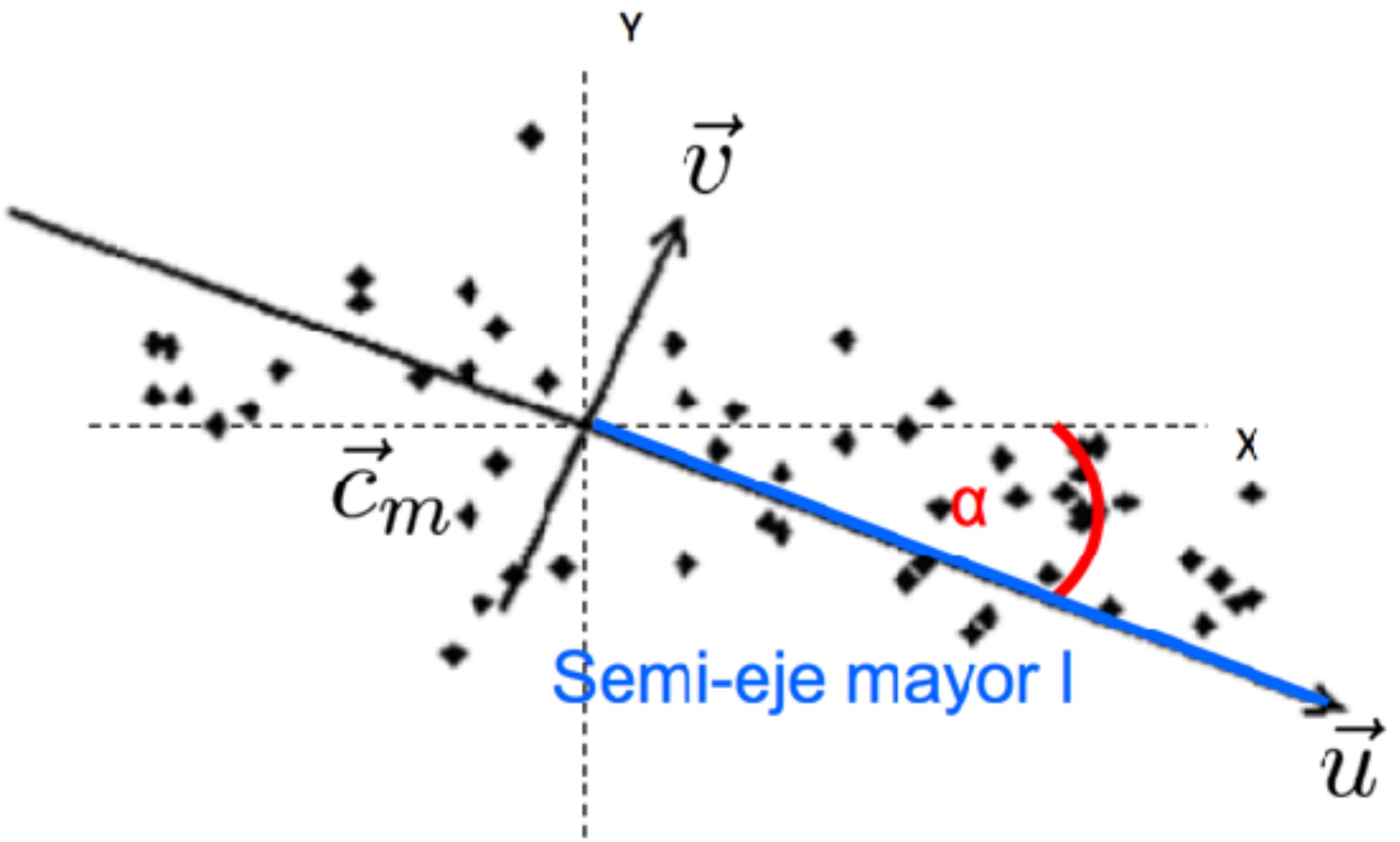




- We can always find a direction to maximize variance
- Equivalent to diagonalize covariance matrix



- We look for a rotation where covariance matrix is diagonal.



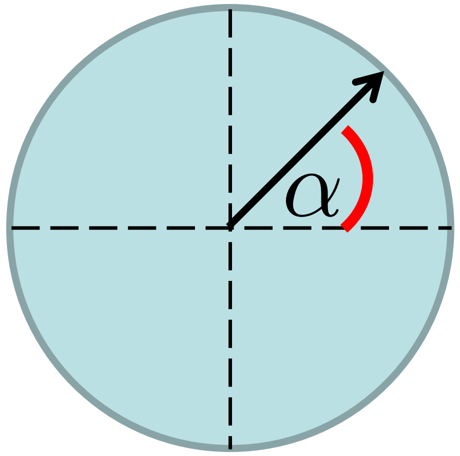
- If we call the rotation α
- Covariance matrix is diagonal (eigenvectors):

$$\sigma^2 \vec{u} = \lambda \vec{u} \qquad \sigma^2 : \text{ covariance matrix}$$

- If we assume a size 1 vector

$$\cos(\alpha)^2 + \sin(\alpha)^2 = 1$$

$$\hat{u} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$$



SHAPE DESCRIPTION: GEOMETRICAL VIEW

$$\sigma^2 \vec{u} = \lambda \vec{u}$$

$$\vec{u} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$$

$$\begin{pmatrix} (\sigma_{xx}^2 - \lambda)\cos(\alpha) + \sigma_{xy}^2 \sin(\alpha) \\ (\sigma_{xy}^2 \cos(\alpha) + (\sigma_{yy}^2 - \lambda)\cos(\alpha) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\alpha = \frac{1}{2} \arctan\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right)$$

- With 1st eigenvalue we can measure the "length" (l) of the object in its intrinsic shape.

$$l^2 = \lambda = \frac{Tr(\sigma^2)}{2} + \sqrt{\frac{T^2}{4} - det(\sigma^2)}$$

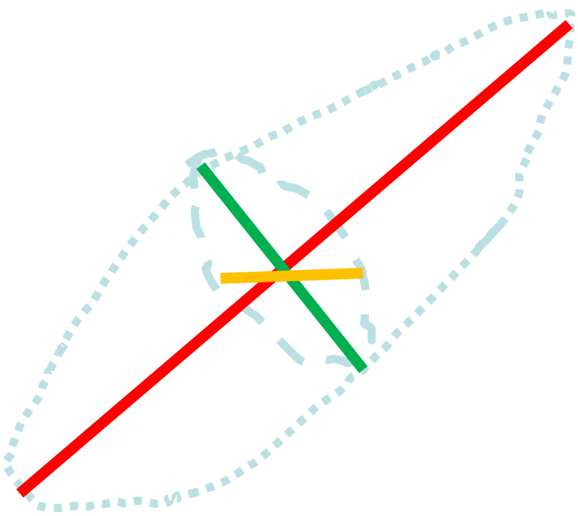
$$l^2 = \lambda = \frac{1}{2}\left(\sigma^2_{xx} + \sigma^2_{yy} + \sqrt{(\sigma^2_{xx} - \sigma^2_{yy})^2 + 4(\sigma^2_{xy})^2}\right))$$

- We can now define eccentricity as:

$$Elong = 1 - \frac{\lambda_2}{\lambda_1}$$

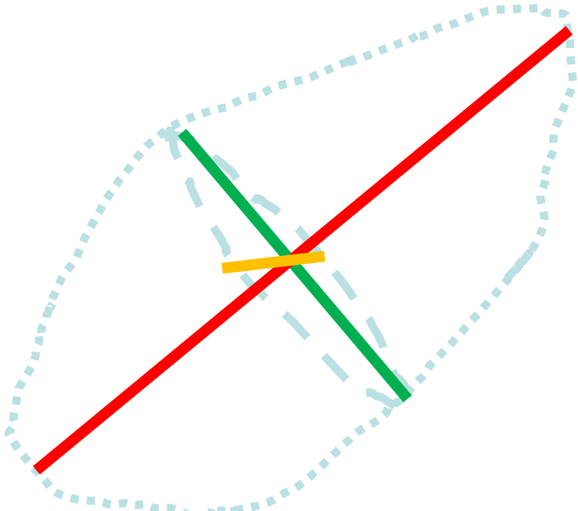
$$R.Elong = 1 - \frac{\lambda_3}{\lambda_1}$$

$$Flatness = 1 - \frac{\lambda_3}{\lambda_2}$$



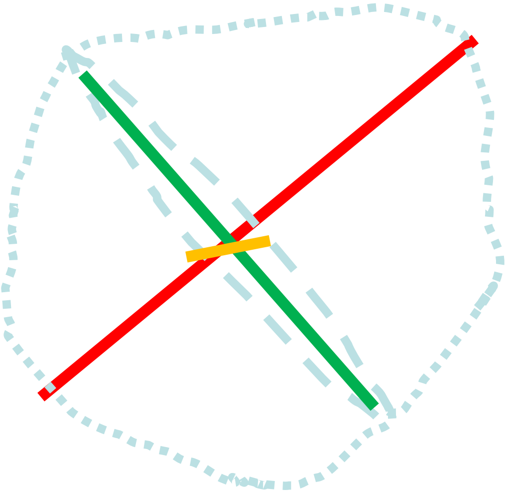
$\lambda_1 >> \lambda_2$

Elong. ~ 1



$\lambda_1 >> \lambda_3$

R. Elong. ~ 1



$\lambda_2 >> \lambda_3$

Flatn. ~ 1

- We can compute it fast in binary images with image moments.

$$m_{p,q} = \sum x^p y^q I(x,y) \qquad \mu_{p,q} = \sum (x-\overline{x})^p (y-\overline{y})^q I(x,y)$$

Order 0

$$m_{0,0} = \sum I(x,y)$$

$$area = m_{0,0}$$

Order 1

$$m_{1,0} = \sum xI(x,y)$$

$$m_{0,1} = \sum yI(x,y)$$

$$\vec{c}_m = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}}\right)$$

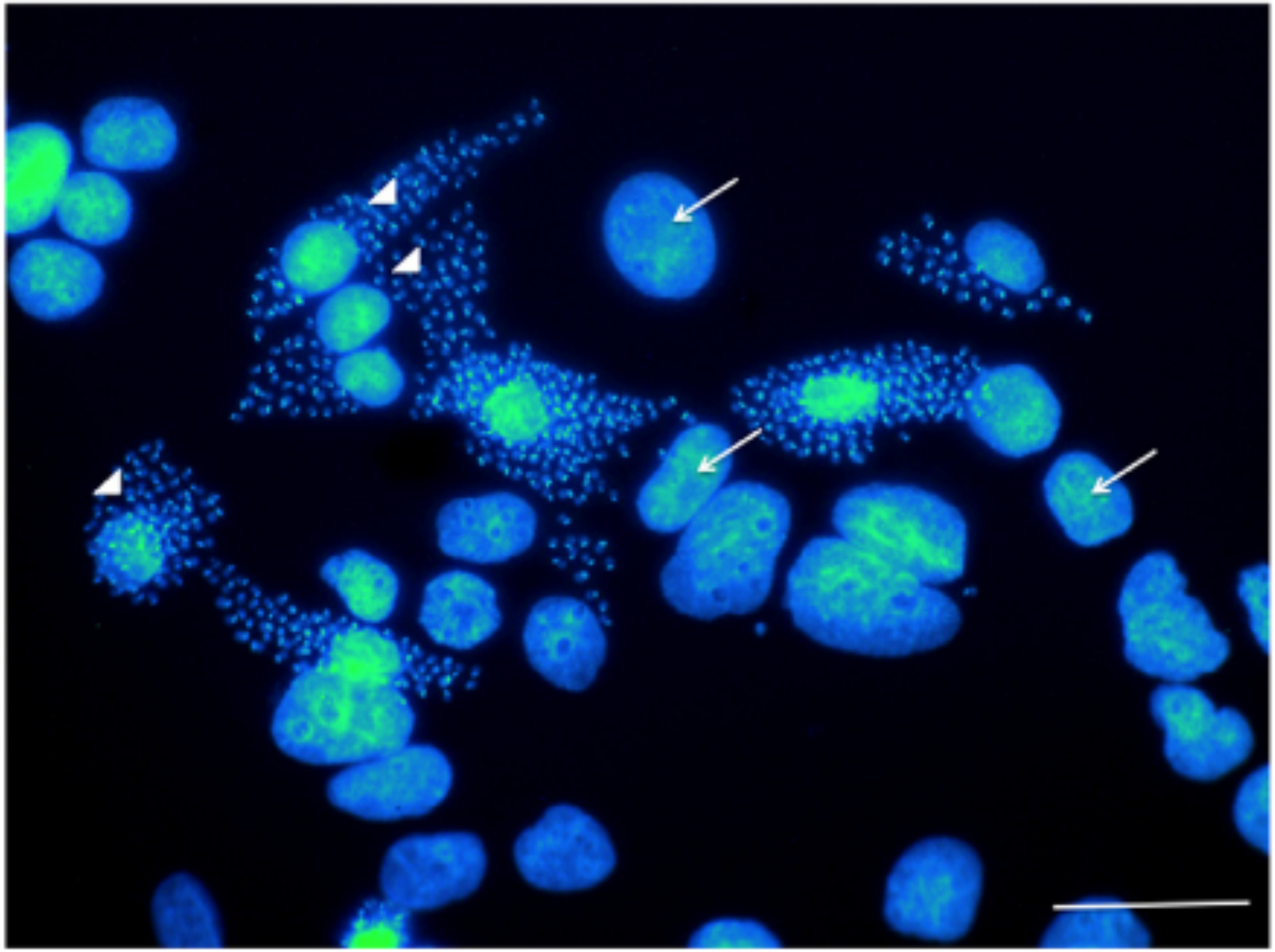
- With 2nd order moments covariance matriz is:

$$\sigma_x^2 = \frac{\sum(x-\overline{x})^2}{N} = \frac{\mu_{2,0}}{\mu_{0,0}} \qquad \sigma_{xy}^2 = \frac{\sum(x-\overline{x})(y-\overline{y})}{N} = \frac{\mu_{1,1}}{\mu_{0,0}}$$

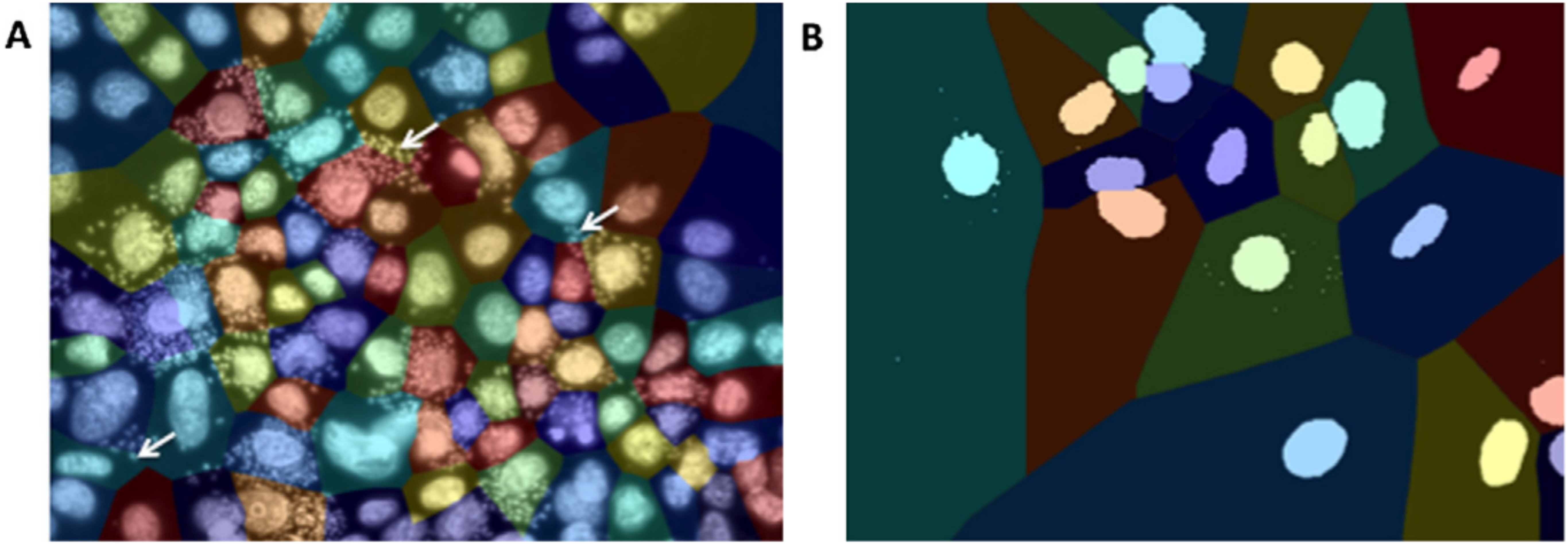
$$\sigma^2 = \frac{1}{\mu_{0,0}} \begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix}$$

- To find the rotation is known as *Principal Component Analysis* (PCA).
- Multiple applications, eg. select eigenvectors to simplify distribution (objects) to a few numbers.
- PCA in FIJI [demo].

- Challenge 1: to estimate the number of parasites, BeWo cells, and their eccentricity.



- Count the number of infected cells by assigning (eg Voronoi) parasites to BeWo cells.



Acta Trop. 2015 Mar;143:47-50. doi: 10.1016/j.actatropica.2014.12.006. Epub 2014 Dec 29.

THANKS

F-Medicine

<u>SCIAN-Lab</u>
Steffen Härtel
Mauricio Cerda
Victor Castañeda
Jorge Jara
Marilyn Gatica
Paula Llanos
Alejandro Lavados
Juan Jose Alegría
Jimena López
Jorge Mansilla
Lucas Ale
Jorge Toledo

<u>LEO-Lab</u>
Miguel Concha
German Reig
Loreto López
Eduardo Pulgar

<u>Lab Chagas</u>
Ulrike Kämmerling
Anita Liempi

Hospital Clínico
<u>U-Chile</u>

F-Engineering

<u>Computer Science</u>
Nancy Hitschfeld

<u>Physics</u>
Rodrigo Soto
Nestor Sepúlveda

GOCCHI
Grupo Oncológico Cooperativo Chileno de Investigación
Chilean Cooperative Group For Oncologic Research

Biomedical Neuroscience Institute
Chile

PIA Programa de Investigación Asociativa
Anillo: VISUAL·D

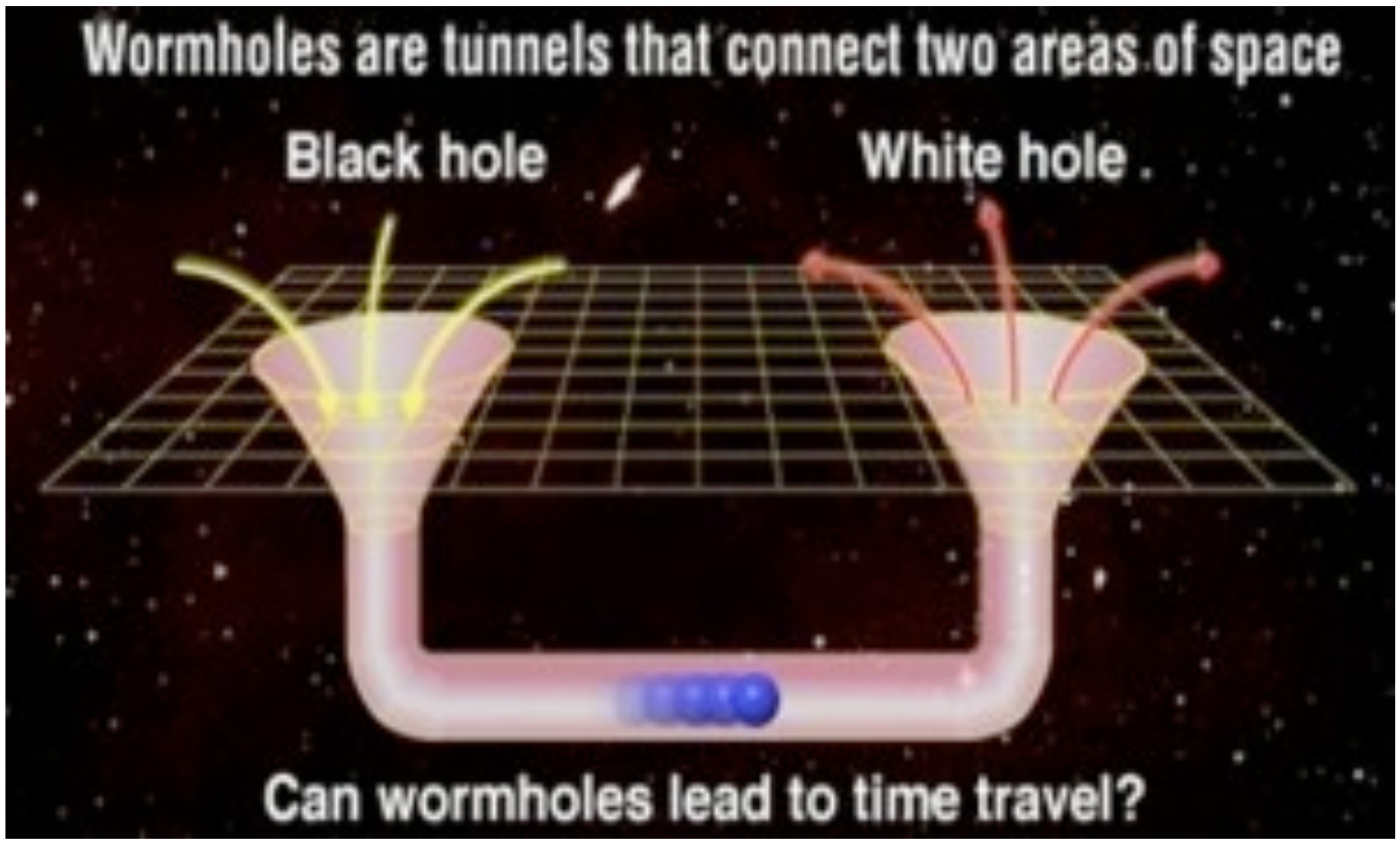
FONDECYT
Fondo Nacional de Desarrollo Científico y Tecnológico
11161033

Facultad de Medicina
Universidad de Chile

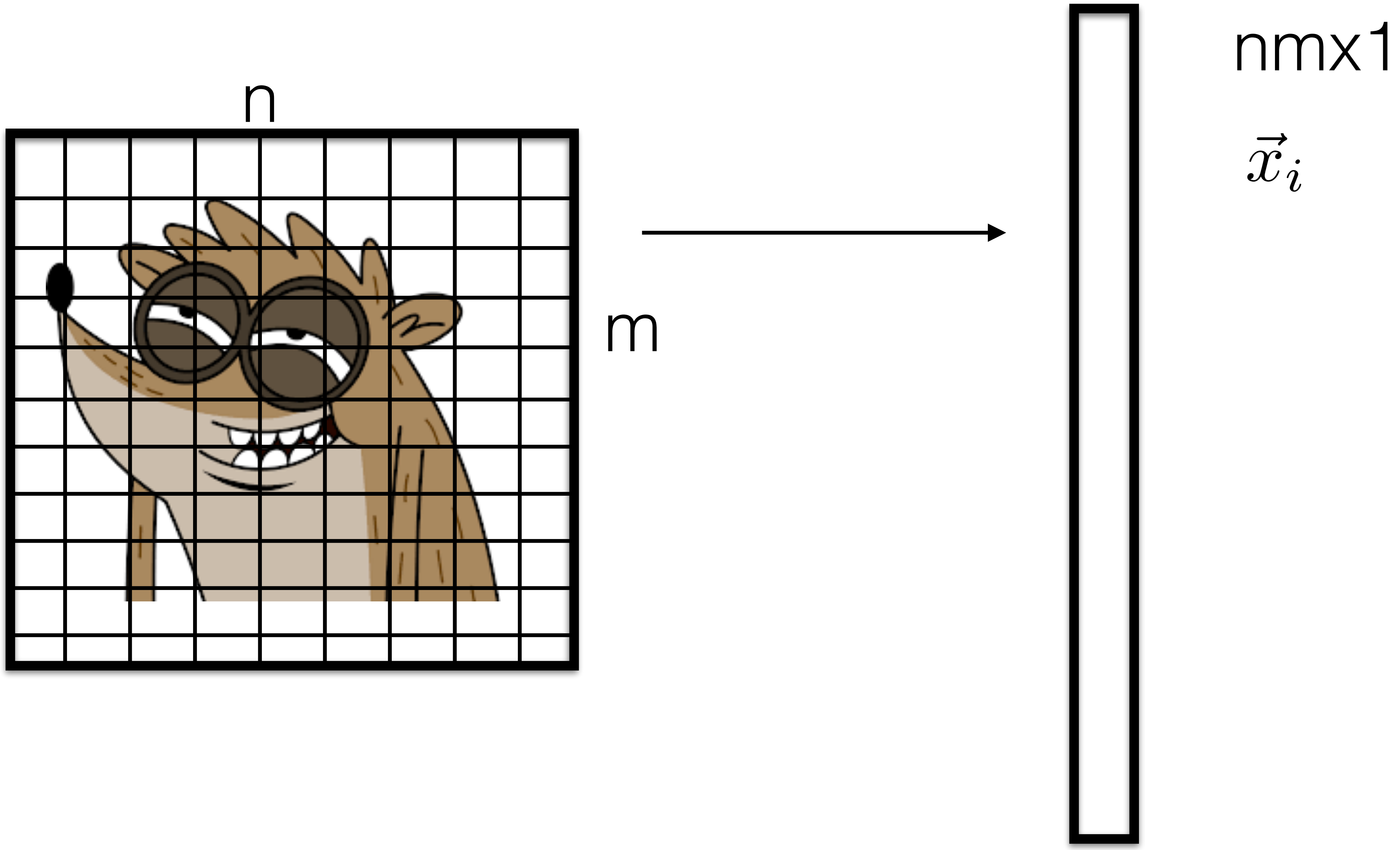
ICBM
Instituto de Ciencias Biomédicas

MAURICIO CERDA
MAURICIOCERDA@MED.UCHILE.CL
HTTP://WWW.SCIAN.CL
HTTP://WWW.BNI.CL

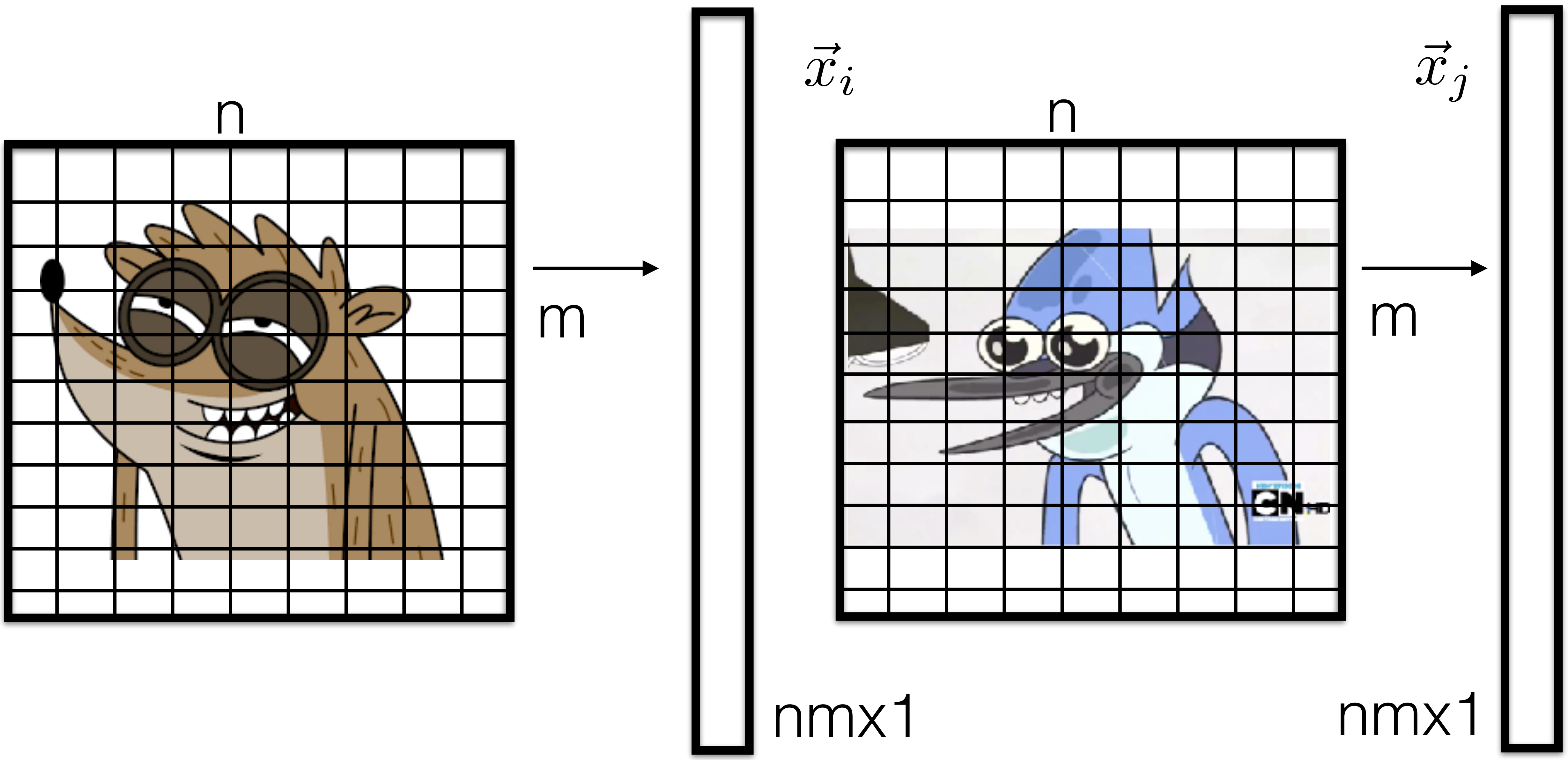
SHAPE DESCRIPTION: HIGH DIMENSION?



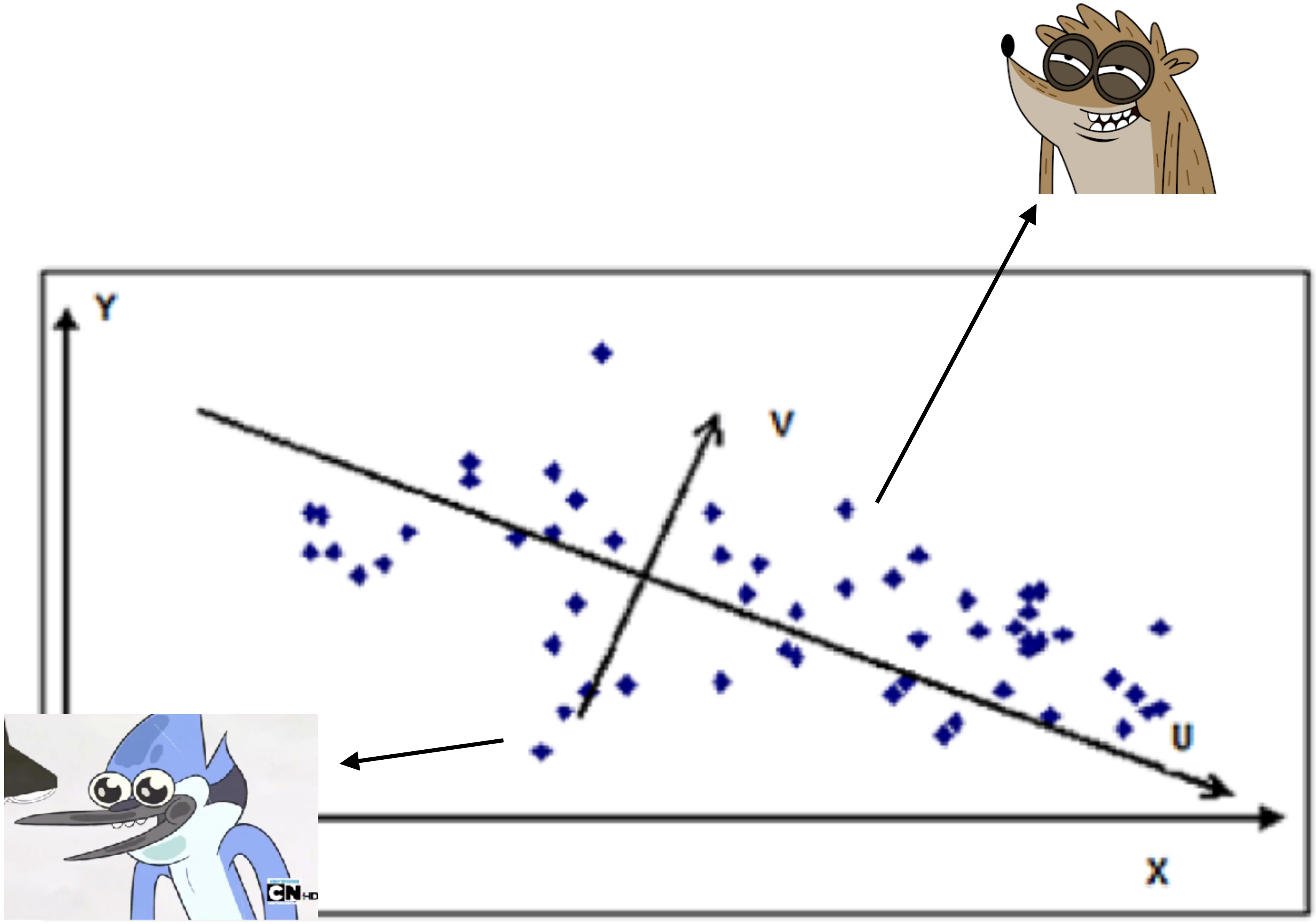
- How to understand an image in high dimension?



- For 2D images, we now have a nm size vector per image



- Now, each image is a point in your feature space.



- Now, each image is a point in our feature space.

$$\mathbf{X} = [\vec{x}_1 ... \vec{x}_i ... \vec{x}_k]$$ (column vector)

$$\mu_i = E(\vec{x}_{:,i})$$ (mean of row 1)

$$\Sigma_{ij} = cov(\vec{x}_{:,i}, \vec{x}_{:,j}) = E[(\vec{x}_{:,i} - \mu_i)(\vec{x}_{:,j} - \mu_j)]$$

$$\mathbf{\Sigma} = \begin{bmatrix} E[(\vec{x}_{:,1} - \mu_1)(\vec{x}_{:,1} - \mu_1)] & E[(\vec{x}_{:,1} - \mu_1)(\vec{x}_{:,2} - \mu_2)] & \cdots & E[(\vec{x}_{:,1} - \mu_1)(\vec{x}_{:,k} - \mu_k)] \\ \vdots & \ddots & \vdots & \\ E[(\vec{x}_{:,k} - \mu_k)(\vec{x}_{:,1} - \mu_1)] & E[(\vec{x}_{:,k} - \mu_k)(\vec{x}_{:,2} - \mu_2)] & \cdots & E[(\vec{x}_{:,k} - \mu_k)(\vec{x}_{:,k} - \mu_k)] \end{bmatrix}$$

- If $\mu_i = 0$ (centered data)

$$\mathbf{\Sigma} = \frac{1}{k}\mathbf{X}^T\mathbf{X}$$

- We can diagonalize the matriz (eg. using SVD)

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{V}^\mathbf{T}$$

Eigenvalues (diagonal matrix)

Eigenvectors

- If eigenvalues are sorted from higher to lower.
- The first eigenvector will indicate the direction that maximizes variance.
- If the input vector are size nm, how many eigenvector are in the base?

- <u>Example</u>. Face representation (*eigenfaces*) from a set of k photos form the same person.

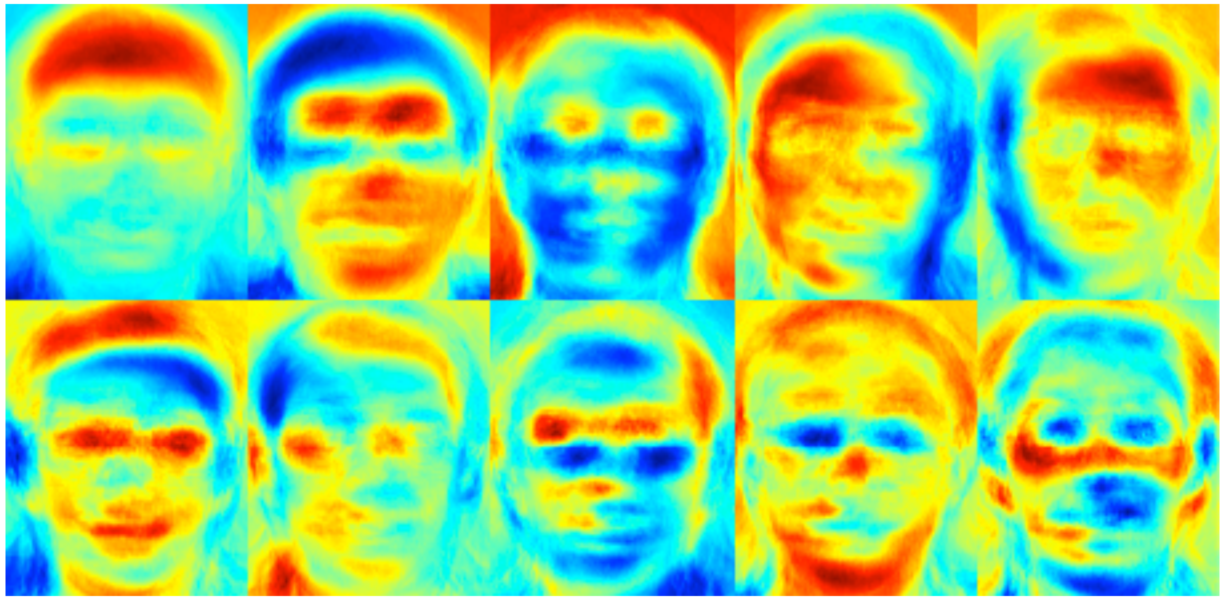

1


2

…


k

- Example. Face representation (*eigenfaces*). The first 10 eigenvectors are:



- Example. Face representation (*eigenfaces*). We can use it as a way to reduce dimensionality: