Mathematical characterization of galaxies using IFS data: II - adding spatial information

Yago Ascasibar
(Universidad Autónoma de Madrid)

INAOE, 02/09/2014

Outline

1. Binning
   - Voronoi
   - HII explorer
   - BaTMAn
2. Morphological characterization

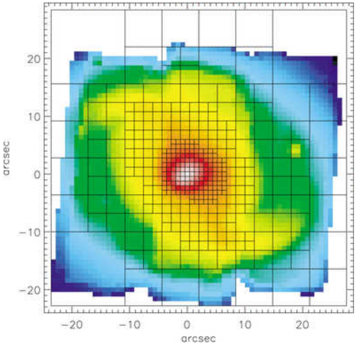
Voronoi binning

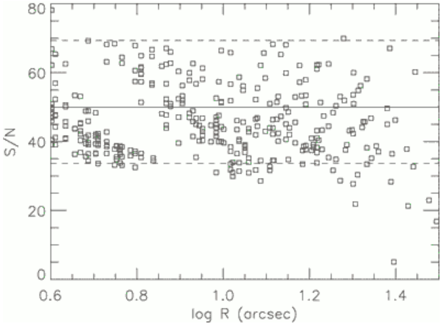
(Cappellari & Copin 2003)

Motivation

Constant signal-to-noise

$$(S/N)_b \simeq (S/N)_0$$





Motivation

Constant signal-to-noise

$$(S/N)_b \simeq (S/N)_0$$

Desired properties

- Topological tesselation
- Morphological compactness
- Uniformity

Algorithm

Weighted average

$$(S/N)_{A+B} = \frac{w_A S_A + w_B S_B}{\sqrt{w_A^2 N_A^2 + w_B^2 N_B^2}}$$

"Optimal" weighting

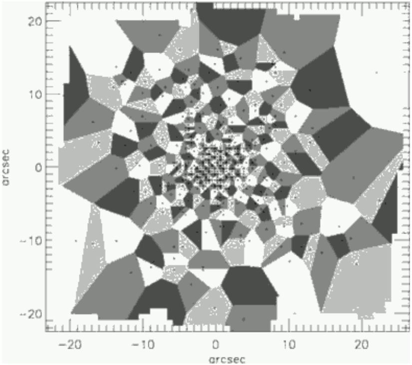
$$w_i \propto S_i/N_i^2$$

$$(S/N)_{A+B} = \sqrt{(S/N)_A^2 + (S/N)_B^2}$$

(S/N) "density"

$$\rho(\vec{x}) \equiv (S/N)^2(\vec{x})$$

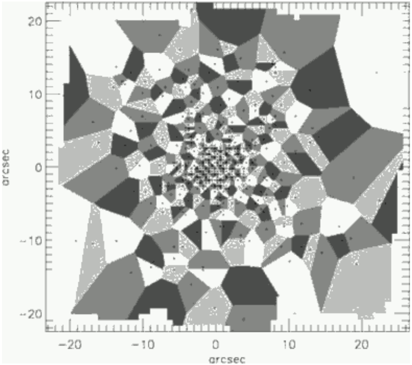
Algorithm

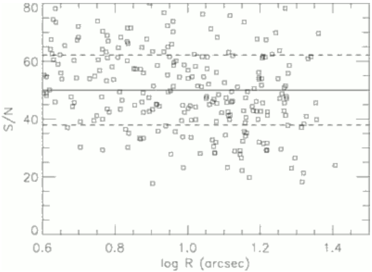


Modified k-means

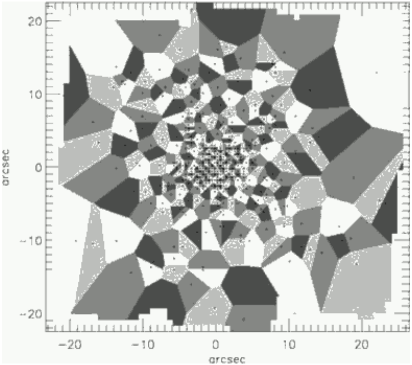
1. Select random seed
2. Tesselate
3. $\vec{x}_b = \frac{\int_b \vec{x}\,\rho^2(\vec{x})\,\mathrm{d}\vec{x}}{\int_b \rho^2(\vec{x})\,\mathrm{d}\vec{x}}$
4. Iterate to convergence

Algorithm


20
10
0
-10
-20
arcsec
-20
-10
0
10
20
arcsec


80
60
40
20
0
S/N
0.6
0.8
1.0
1.2
1.4
log R (arcsec)

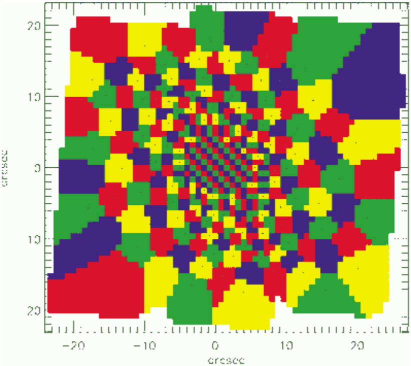
Algorithm

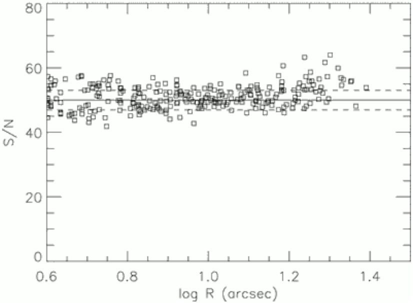


Initial seeds

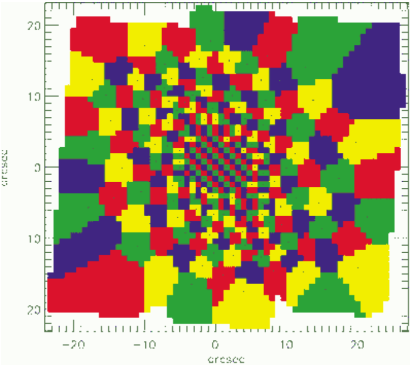
1. Start with highest S/N
2. Accrete nearest pixel while
   - adjacent
   - "roundness" $R = \frac{r_{max}}{r_{eff}} - 1 < 0.3$
   - $(S/N)_b$ approaches $(S/N)_0$
3. Mark as unsuccessful if $(S/N)_b < 0.8(S/N)_0$
4. Start new bin with nearest pixel

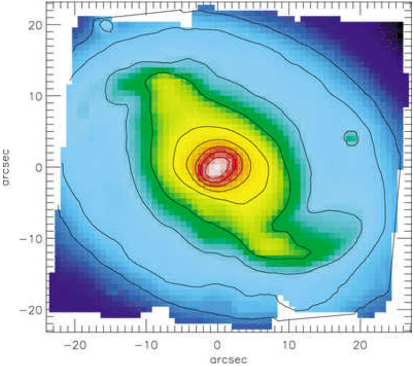
Results


20
10
0
10
20
arcsec
-20
-10
0
10
20
arcsec


80
60
40
20
0
S/N
0.6
0.8
1.0
1.2
1.4
log R (arcsec)

Results


arcsec
arcsec
20
10
0
10
20
−20
−10
0
10
20


arcsec
arcsec
20
10
0
−10
−20
−20
−10
0
10
20

HII explorer

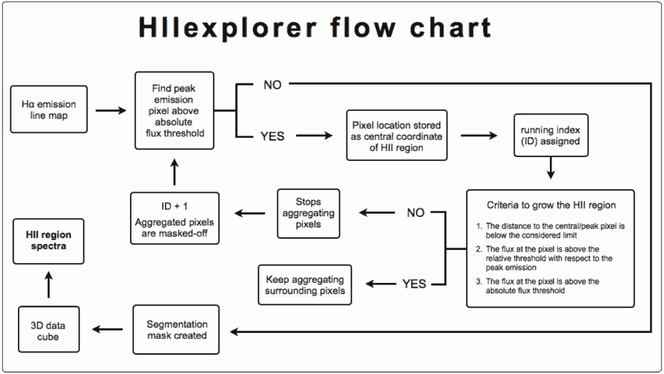
(Sánchez et al. 2012)

Algorithm

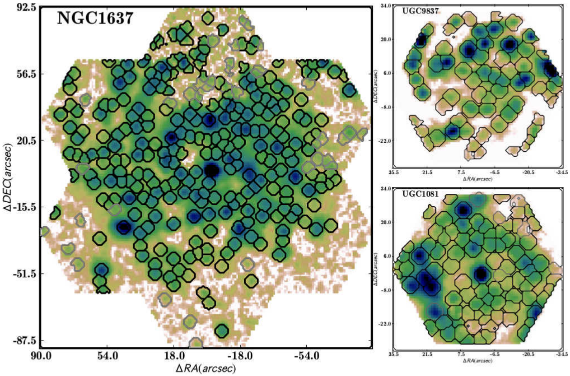
Input

- Emission-line map (Hα)
- Minimum peak intensity
- Intensity contrast
- Maximum radius
- Minimum flux

Algorithm


HIIexplorer flow chart
Hα emission line map
Find peak emission pixel above absolute flux threshold
NO
YES
Pixel location stored as central coordinate of HII region
running index (ID) assigned
Criteria to grow the HII region
1. The distance to the central/peak pixel is below the considered limit
2. The flux at the pixel is above the relative threshold with respect to the peak emission
3. The flux at the pixel is above the absolute flux threshold
NO
Stops aggregating pixels
ID + 1
Aggregated pixels are masked-off
YES
Keep aggregating surrounding pixels
HII region spectra
3D data cube
Segmentation mask created

Results


NGC1637
92.5
56.5
20.5
-15.5
-51.5
-87.5
90.0
54.0
18.0
-18.0
-54.0
ΔRA(arcsec)
ΔDEC(arcsec)
UGC9837
UGC1081

Bayeseian **T**echnique for **M**ulti-image **An**alysis

(Casado et al., in prep.)

A long-standing battle

Frequentist approach

- Likelihood
- Confidence intervals

Bayesian approach

- Likelihood
- Prior prob.
- Model evidence
- Posterior prob.

Likelihood

$\mathcal{L}_m(\vec{x}|\vec{\theta}_m)\ \mathrm{d}\vec{x}$: probability of observing the datum $\vec{x}$, given model m with parameters $\vec{\theta}_m$

Prior probability

$\pi_m(\vec{\theta}_m)\ \mathrm{d}\vec{\theta}_m$: expected probability distribution for $\vec{\theta}_m$

Model evidence

$E_m(\vec{x}) = \int \pi(\vec{\theta}_m)\, \mathcal{L}_m(\vec{x}|\vec{\theta}_m)\, \mathrm{d}\vec{\theta}_m$: probability of $\vec{x}$, given m

Posterior probability

$p(\vec{\theta}_m|\vec{x}) = \frac{\pi(\vec{\theta}_m)\, \mathcal{L}_m(\vec{x}|\vec{\theta}_m)}{E_m(\vec{x})}$: probability of $\vec{\theta}_m$, given m and $\vec{x}$

Dear Sir or Madam?

Data

The subject gave birth to twins

Likelihood

Probability of a woman having twins $\sim 10^{-5}$
Probability of an amoeba having twins ~ 1

Frequentist approach

The hypothesis that the subject is a woman can be ruled out at the 99.999% level

Bayesian approach

- Prior: $\pi_m = \pi_w = 0.5$
- $E = \pi_m \mathcal{L}_m + \pi_w \mathcal{L}_w = 5 \times 10^{-6}$
- Posterior: $p_m = 0$; $p_w = 1$

Model comparison

Data

N measurements with errors: $\{x_i, \sigma_i\}_{i=1,N}$

Model 1

- One object with actual value x
- $\pi(x)\,\mathrm{d}x = \frac{\mathrm{d}x}{L}$ if $0 < x < L$
- $\mathcal{L}_1(\{x_i\}|x) = \prod_{i=1}^N \frac{e^{-\frac{(x_i-x)^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}}$

Model 2

- Two objects, with x_A and x_B
- $\pi(x_A) = \pi(x_B) = \pi(x)$
- "labels" $\{l_i\}_{i=1,N} = A|B$
- $\mathcal{L}(\{x_i\}|x) = \prod_{i=1}^N \frac{e^{-\frac{(x_i-x_{l_i})^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}}$

$$E_1(\{x_i\}) = \int \prod_{i=1}^N \frac{e^{-\frac{(x_i-x)^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}} \frac{\mathrm{d}x}{L}$$

$$E_2(\{x_i\}) = \int \int \prod_{i=1}^N \frac{e^{-\frac{(x_i-x_{l_i})^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}} \frac{\mathrm{d}x_A}{L} \frac{\mathrm{d}x_B}{L}$$

BaTMAn

Formulation of the problem

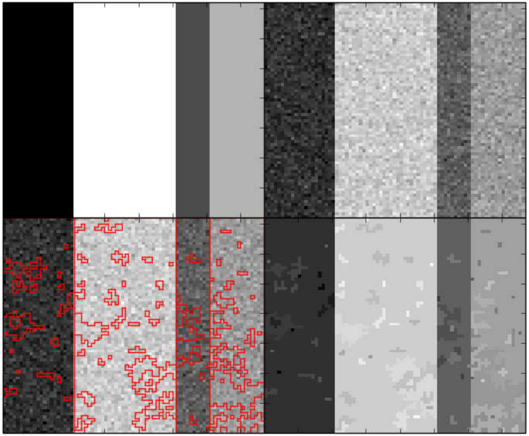
- Data: N measurements (pixels) with
  - intensity $\{x_i\}_{i=1,N}$
  - errors $\{\sigma_i\}_{i=1,N}$
- Model: B regions (bins) where $x = x_b$
  - $\pi(x_b)\,\mathrm{d}x_b \sim \frac{\mathrm{d}x_b}{\max(x_i)-\min(x_i)}$
  - "labels" $\{b_i\}_{i=1,N}$
  - $\mathcal{L}(\{x_i\}_{i=1,N}|\{x_b\}_{b=1,B}) = \prod_{i=1}^{N} \frac{e^{-\frac{(x_i-x_{b_i})^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}}$

BaTMAn

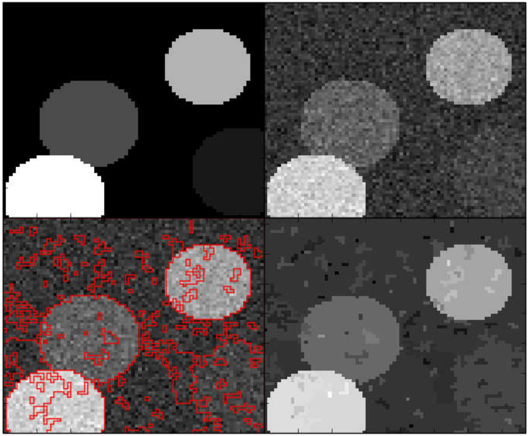
Algorithm

1. Start with $B = N$
2. While $E_{B-1} > E_B$:
    - merge the two adjacent regions that yield the highest E_{B-1}

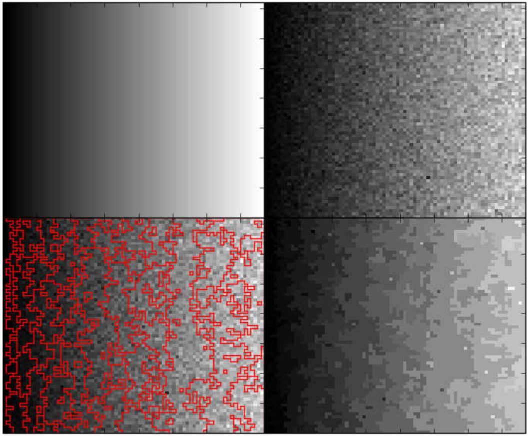
Results



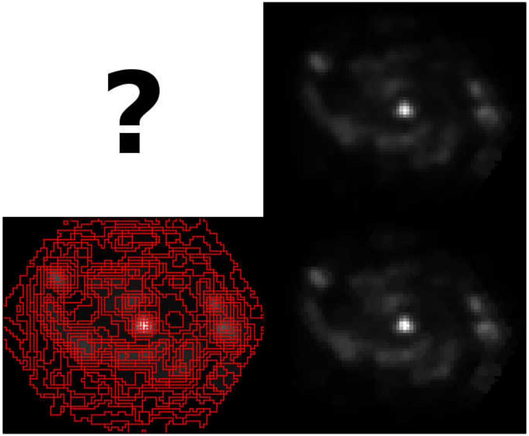
Results



Results

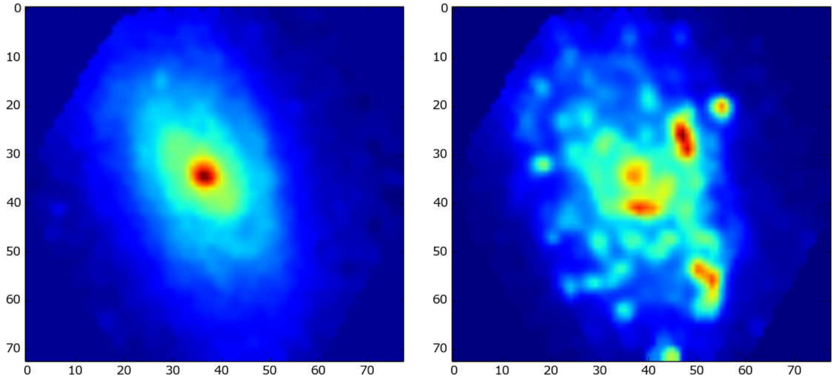


Results



Morphological characterization

Multi-wavelength morphology



Multi-wavelength morphology

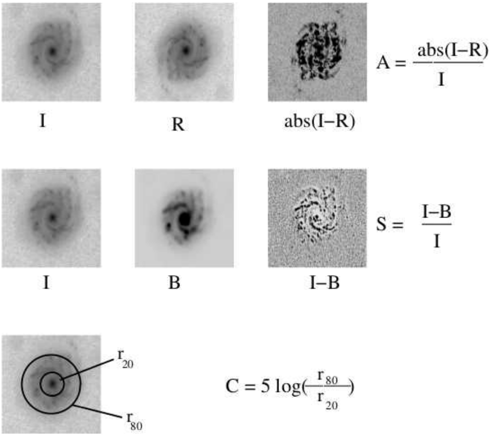
What for?

1. Classification and comparison
2. Statistics of the galaxy population
3. Input for chemical evolution models
4. Low signal-to-noise kinematics

How?

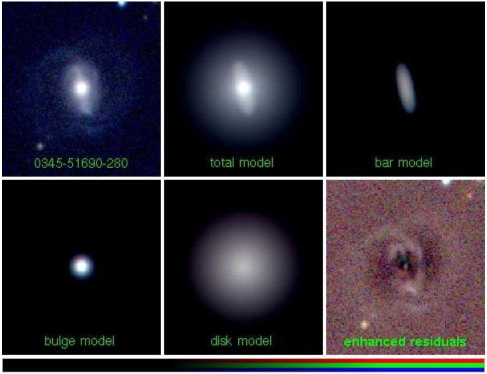
1. Statistical descriptors
2. Parametric models
3. Non-parametric models

Statistical descriptors



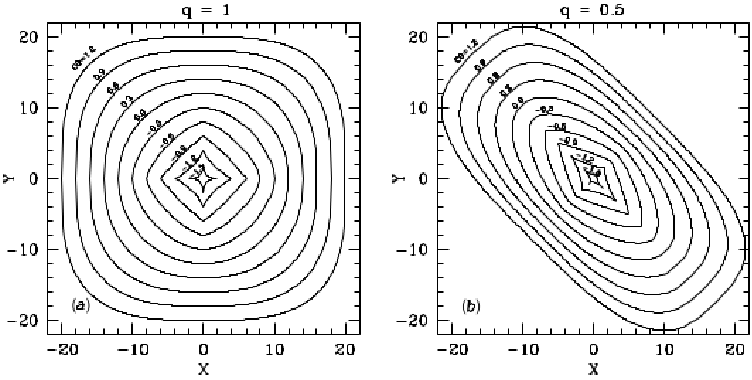
Conselice (2003)

Model-based decomposition



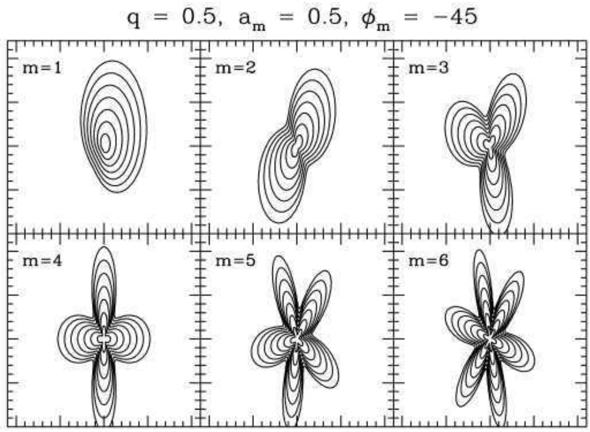
de Souza et al. (2004)

Generalized ellipsoids



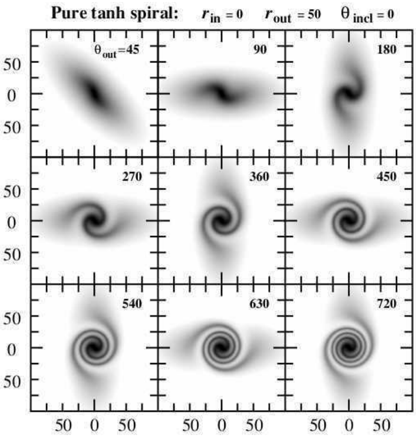
Peng et al. (2010)

Fourier modes



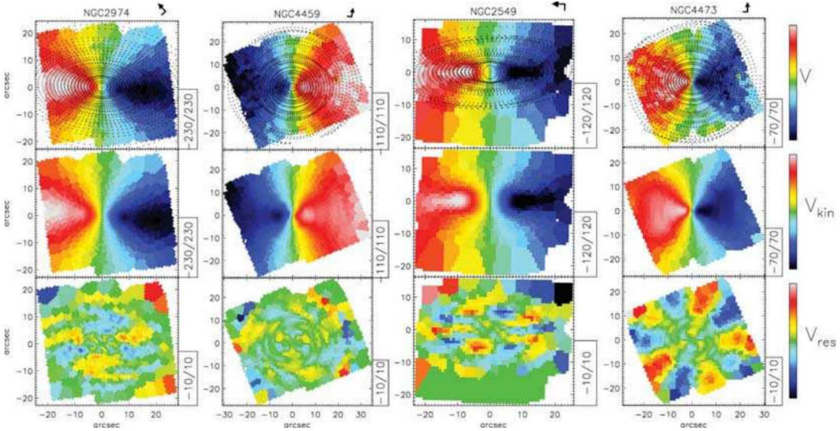
Peng et al. (2010)

Spiral modes



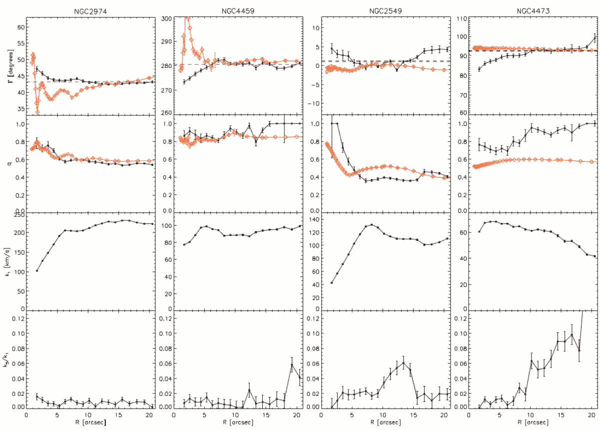
Peng et al. (2010)

Kinemetry



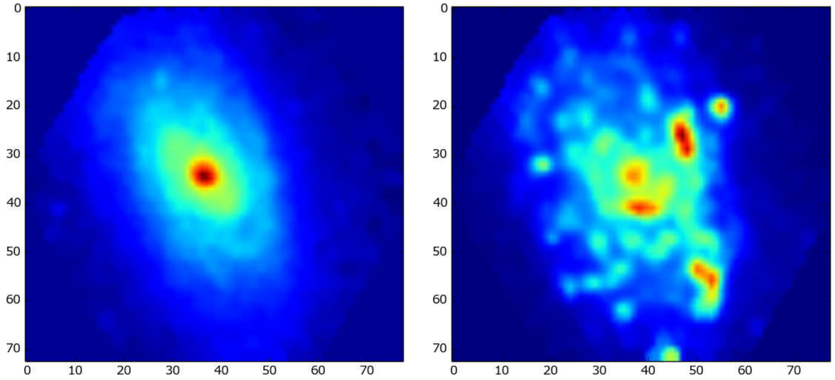
Krajnović et al. (2006)

Kinemetry



Krajnović et al. (2006)

What next?


0
10
20
30
40
50
60
70
0
10
20
30
40
50
60
70
0
10
20
30
40
50
60
70
0
10
20
30
40
50
60
70

Summary

Binning

Bayes' theorem

Morphological characterization

There is no need to re-invent the wheel...
...but there is plenty of information to recover!