

Fuzzy Segmentation and Astronomical Images Interpretation

Pierre Dhérété, Laurent Wendling, Jacky Desachy

Université Paul Sabatier - IRIT
118, route de Narbonne, 31062 Toulouse Cedex FRANCE

Abstract. This paper presents a fast segmentation method using fuzzy sets theory applied to astronomical images interpretation. It consists in separation in two classes (« stars » and « other objects ») with human-like decision. This separation constitutes the first step in astronomical images analysis, before spectrographic analysis and high level interpretation. A fuzzy segmentation isolates regions relative to objects and evaluates shape parameters for each region. Then, combined criteria are computed on the obtained fuzzy regions. The decision is founded on rules using these combined criteria. For each object we obtain the location, the orientation and the membership degrees to each class, then we build two images coding the membership degrees for each pixel to each class. With this method the use of fuzzy concepts in the first step allows to use fuzzy reasoning all along the classification process.

1. Introduction

The new large telescopes, with very sensitive and wide spectrum instruments, permit to explore regions of sky invisible until nowadays. They make possible deep space study and allow new horizons to astronomical research, mainly cosmology.

Remote sensing founded on deep sky images deals with different problems. The high number of objects excludes systematic study of each object. The bad signal/noise ratio makes the analysis difficult. So, it is very important to use fast classification as automatic as possible. The goal is to build catalogs, which regroup the main characteristics of objects, allowing to select them for a specific study.

A fast classification is only an approach, the result is tainted with errors. So, it is very important to take notice of these errors and to specify the reliance of the classification for each object. The notion of uncertainty needed by this type of classification may be expressed with fuzzy logic formalism. Thus, we can associate to each object the possibility to belong more or less to a specific class.

2. Astronomical images

An astronomical image, after an instrumental deconvolution, contains different object types, stars (of our galaxy), galaxies, gravitational bows, globular clusters, galaxy clusters, cosmic rays, etc.

A large field image contains a large amount of small objects. Generally the resolution does not allow to distinguish structure or details. Sometimes the sky background is not homogeneous or flat and may modify the object morphology.

Objects can be in interaction and belong to a group (e.g. clusters). In the neighborhood of clusters, a recognition process has to determine if objects belong to the cluster or not, and if they are before or behind it.

A visual analysis can take a long time and is only the first step in object analysis. Specific astronomical programs uses only a few object categories. So, before performing an analysis, we need to classify the objects. to select interesting objects.

3. Classification

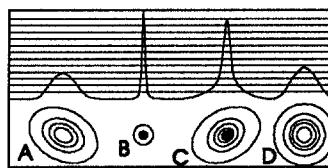
After sky background correction and cosmic rays removal, classifications are generally performed step by step. The first step consists in separation of two classes, stars and other objets. Then the classification needs spectrographic criteria, redshift, red/blue ratio, etc., to separate kinds of stars and other objects. At last, we need more complex criteria (distribution, relationships, shapes) requiring high level concepts.

Current classifications are subjective. An astronomer can attribute different types to same object, so a category is always a fuzzy concept. To obtain fuzzy decisions it becomes interesting to use a fuzzy reasoning all along the classification process. Each step provides fuzzy data to the next step, allowing revision of former decisions and thus may provide a final result closer to reality.

Most of known automated astronomical images classifiers do not use fuzzy concepts, specifically in the first step. They are often founded on FFT and wavelets transforms. They use pattern matching with models or comparison with predetermined values. The parameters are conditioned by images and the adjustments are generally manual.

4. Selection of criteria

To separate stars and other objects (Galaxy) an astronomer uses rules and criteria to take a decision. The criteria are subjective and not precise. They are fuzzy concepts.



Shapes and levels of astronomical objects

The first criterion is the **ellipticity** : A star is circular (B) ; A galaxy is generally elliptic (A) (C) ; A galaxy may be circular (D) (like a star)

⇒ we need a second criterion,

The **sharpness** : A star is very sharp (B) ; A galaxy is generally stretched (A) (D) ; A galaxy with intense nucleus has a high sharpness (C).

5. Segmentation

After sky background rejection, segmentation is performed using a set of level-cuts. Because of high dynamic of images the successive values of α follow an exponential law by decade (1..9, 10..90, 100..900, etc.), to limit the number of level-cuts and to concentrate many cuts in image low levels.

For each level-cut we obtain a binary image, after applying a closing process to cancel the noise, we extract the contour for each region.

A single object profile is gaussian, see above figure, for a same object the successive level-cuts give contours included one into another. Several close objects may define only one region at low level α , and more with increasing α . The superimposed objects give a similar case.

So we create a tree describing inclusions of contours, then we regroup contours for each elementary object. The low level contour is the support for the object.

6. Fuzzy parameters

We need ellipticity and sharpness to classify objects. But we need to locate each object and its orientation too. We process these two parameters as fuzzy quantities.

6.1 Moments, barycenter, sharpness

Barycenter, ellipticity and orientation are deducted from moments of order 0 through order 2 [3]. For each object we have one support and higher level-cuts. For each level-cut A_α we compute the six moments, the barycenter and the sharpness.

Expressions of moments are :

$$M_{00} = \sum_x \sum_y I(x,y) ; \quad M_{10} = \sum_x \sum_y x \cdot I(x,y) \quad \text{and} \quad M_{01} = \sum_x \sum_y y \cdot I(x,y)$$

$$\text{and} \quad M_{jk} = \sum_x \sum_y (x - B_x)^j \cdot (y - B_y)^k \cdot I(x,y) \quad \text{for } M_{02}, M_{11} \text{ and } M_{20}.$$

The barycenter coordinates are :

$$B_{x_i} = M_{10} / M_{00} \quad \text{and} \quad B_{y_i} = M_{01} / M_{00}$$

The sharpness is :

$$S_i = -\ln \left(\frac{\alpha_i}{I_{\max}} \right) / \text{Area}(A_{\alpha_i})$$

I_{\max} : barycenter intensity ; α_i : level-cut level ; $\text{Area}(A_{\alpha_i})$: area of level-cut A_{α_i}

6.2 Weight function

Images are not real membership functions, even if they approximate the belonging function to class « Object », thus we are not strictly in fuzzy logic formalism. It is important to normalize the function for each object.

The weight function privileges to much high level-cuts. In our images these regions might be very small (3 or 4 pixels), consequently computed parameters may be wrong. For instance, an elliptic object may have an high level-cut composed of 4 pixels (and then supposed to be circular). This will affects strongly the global ellipticity, and sharpness may not be defined. Moreover, we cannot use the same weight for all level-cuts of a same object. As low level-cuts (very noisy) are more numerous than high level-cuts the final result will be affected.

The chosen weight function minimizes low level-cuts and moderately privileges high level-cuts, and of course normalizes the function for each object.

Expression is : $m(A_{\alpha_i}) = \alpha_i / \sum_{j=1}^n \alpha_j$ It respects the constraint : $\sum_{i=1}^n m(A_{\alpha_i}) = 1$

α_i is current level for the region, A_{α_i} are each level-cuts of the object.

6.3 Combined criteria

Then we define combined criteria for each object. The values of order 2 moments depend on level-cut area. Level-cuts with small area have low moments, and our high level level-cuts have low area. To minimize a bit more the influence of high level-cuts in combined ellipticity computation, the ellipticity is not computed for each level-cut and then combined. But we first combine order 2 moments, to get combined moments, and then we compute combined ellipticity using them. Thus medium regions with good resolution and good signal/noise ratio are privileged.

Combined moments are : $CM_{jk} = \sum_{i=1}^n m(A_{\alpha_i}) \cdot M_{jk_i}$

Combined barycenter is : $CB_X = \sum_{i=1}^n m(A_{\alpha_i}) \cdot B_{X_i}$; $CB_Y = \sum_{i=1}^n m(A_{\alpha_i}) \cdot B_{Y_i}$

Combined orientation is : $CO = \frac{1}{2} \tan^{-1} \left(\frac{2 \cdot FM_{11}}{FM_{20} - FM_{02}} \right)$ For circular objects CO is forced to 90° instead of 45°

Combined sharpness is : $CS = \sum_{i=1}^n m(A_{\alpha_i}) \cdot S_i$

Combined ellipticity is [3] : $CE = \sqrt{\frac{(CM_{20} + CM_{02}) + \sqrt{(CM_{20} - CM_{02})^2 + 4 \cdot CM_{11}^2}}{(CM_{20} + CM_{02}) - \sqrt{(CM_{20} - CM_{02})^2 + 4 \cdot CM_{11}^2}}}$

7. Judgments associated to criteria

Ellipticity and sharpness have different ranges. A medium ellipticity has not the same value as medium sharpness. To judge this two values in a similar way we associate a fuzzy set to each criteria $\mu_E(CE)$ and $\mu_S(CS)$, coding them in $[0, 1]$ domain. Thus the criteria are homogeneous.

7.1 Decision rules

The "Star" membership degree, SCF (Star Certainty Factor), depends on $\mu_E(CE)$ and $\mu_S(CS)$. The decision uses rules simulating human-like behavior. This rules are activated in the following order :

1. If sharpness is very high (i.e. $\mu_S(CS) = 1$), then SCF is $\mu_E(CE)$.
If sharpness is high, then it is either a star or an object with sharp nucleus, we judge only with ellipticity.
2. If object is very elliptic (i.e. $\mu_E(CE) = 0$), then it is not a star SCF = 0.
If ellipticity permits to assert it is not a star, then ignore sharpness. It squeezes doubt with sharp galaxies.

3. If ellipticity is very low (i.e. $\mu_E(CE) = 1$), then SCF is $\mu_S(CS)$.
If object is circular, then it is either a star or a stretched circular object, only sharpness is useful.
4. If none of former rule is applied, then SCF is the medium value of $\mu_E(CE)$ and $\mu_S(CS)$.

And a contradiction between the two values gives 0.5, this is a fully doubt.

The certainty factor for class « Other object », OCF, is : $OCF = 1 - SCF$.

We obtain SCF and OCF measures for each object. This is extended to each object pixel (x, y) to compute ISCF and IOCF images as follows :

$$ISCF(x, y) = \sum_{\alpha_i / (x, y) \in A_{\alpha_i}} m(A_{\alpha_i}) \cdot SCF ; \quad ISCF(x, y) = \sum_{\alpha_i / (x, y) \in A_{\alpha_i}} m(A_{\alpha_i}) \cdot SCF$$

8. Results

We first test the process on synthesis images without noise with known catalog. Fully good answers represent 70% (CF=1 and is right). Fully bad answers represent 3% (CF=1 and is wrong). If till fully doubt answers (CF=0.5) are considered good we obtain 85% of success. Of course this result depends on the content of the image.

Then we test on noisy synthesis images. With a Poissonian additional noise (68 ± 2). Fully good answers represent 60%. Fully bad answers 8%. Answers till fully doubt 85% too. Fully bad answers concern mainly very weak objects. The quality of classification depends only on dispersion of noise and not on mean value. Typically a good image has a dispersion between ± 1 and ± 3 with short exposure time, the noise increase with exposure time (i.e. sky background) following a Poisson's law.

The test with real noisy images gives a similar result, when sky background is uniform. One of the images, without catalog, has been processed without noise reduction, the noise dispersion is ± 5 . All very weak objects are classified in « Other » class. The problem is due to low quantization rate. A good classification should give 0.5 (fully doubt) on this objects.

Modifying ellipticity and sharpness functions ($\mu_E(CE)$ and $\mu_S(CS)$), we record a problem. If classification of weak objects is good, classification of bright objects is less, and inversely. This is due to low quantization rate too. A weak star has a lower sharpness than a bright star. For some small objects, it is impossible to distinguish stars and other objects if they are circular or with a low ellipticity. We must force CF to 0.5 for this weak objects, only spectral analysis may improve the result.

Two other problems are the "seeing" value (blurring due to atmospheric perturbations during exposure), the star sharpness is different depending on images, and the telescope motion quality which modifies ellipticity of stars. When the seeing is bad it is more difficult to discriminate weak objects. When the motion of the telescope is bad stars are elliptic with the same value in the same direction.

The deep sky test image is provided with a catalog of another classic segmenter : the astronomical image processor IRAF. It does not specify classes but only shape parameters, area, compactness, order 2 moments, and location. The whole of these

parameters is not absolutely right. The comparison between the catalogs give about 75% of success. This is a good result regarding errors of each segmenter and the disparity of criteria. Only the main bright objects are identified, and the classification is right for them. It is impossible to say if classification for weak objects is right or not. We do not have deep sky image with a known catalog (truth) to perform a good comparison.

8.1 Improvements

The quality of classification depends on noise, but modification of level-cuts (number and levels) may improve the results. Increase the doubt for very weak objects depending on noise dispersion.

When a galaxy with high sharpness nucleus is close to a star, it is recognized as a star. Image-adaptive level-cuts is possible to improve the classification.

It is important to take notice of seeing and star ellipticity values. Evaluation of medium sharpness and medium ellipticity of stars permits to adjust sharpness and ellipticity thresholds.

The sky background rejection used in the process is not sufficient, it is just a test operator. We need an operator able to remove non-uniform sky with interpolation method.

9. Conclusion

According to astronomers this is a good result regarding to classic processes and visual analysis. This process is only the first step in astronomical objects classification. The obtained membership degrees allow revision of decisions, and flexible human-like reasoning. Membership degree images may be used by neural networks in combination with spectrographic data.

This method is a new approach allowing the use of fuzzy concepts all along the classification process.

10. References

1. DUBOIS and H. PRADE : « Théorie des possibilités ». Ed. MASSON 1985.
2. DUBOIS and H. PRADE : « Possibility Theory, an approach to the computerized processing of uncertainty », Plenum Press, New-York, 1988.
3. REED TEAGUE : « Image analysis via the general theory of moments ». J. Optical Society of America, Vol.70, N°8, August 1980, p.920-921.
4. SERRA : « Image Analysis and Mathematical Morphology », Academic press, 1982.
5. SHAFER : « A Mathematical Theory of Evidence », Princeton University Press 1976.
6. THONNAT and A. BIJAOUTI : « Knowledge Based Classification of Galaxies » from *Knowledge Based System in Astronomy*, A. Heck & F. Murtagh, Springer Verlag, p.121-159, Berlin, 1989.
7. A. ZADEH : *Fuzzy Sets. Information and Control* 8:338-353, 1965. Reprinted in « Fuzzy Models for Pattern Recognition », J.C. Bezdek and S.K. Pal eds., IEEE Press, 1992 p.35-45.