Classification of the galaxy Milky Way using variable precision rough sets technique

Ban Sabah Ismael

Department of Astronomy and Space, College of Sciences, University of Baghdad

E-mail: ban_alsabah77@yahoo.com

Abstract

Astronomy image is regarded main source of information to discover outer space, therefore to know the basic contain for galaxy (Milky way), it was classified using Variable Precision Rough Sets technique to determine the different region within galaxy according different color in the image. From classified image we can determined the percentage for each class and then what is the percentage mean. In this technique a good classified image result and faster time required to done the classification process.

Key words
Image classification, classification techniques, galaxy, Milky way, Astronomy Image.

Article info.
Received: Apr. 2014
Accepted: May. 2014
Published: Sep. 2014

تصنيف مجرة درب اللبانه بأستخدام تقنية مجموعه تقريبيه متغيرة الدقه

بان صباح اسماعيل

قسم الفلك والفضاء, كلية العلوم, جامعة بغداد

الخلاصة

تعتبر الصور الفلكية هي المصدر الرئيسي حتى نكتشف الفضاء الخارجي, لذلك حتى نتعرف على ما تحتويه مجرة (درب اللبانة), قمنا بتصنيف هذه المجرة باستخدام تقنية مجموعه تقريبية متغيرة الدقة لتحديد المناطق داخل هذه المجرة وفقا للألوان الموجودة في الصورة. حصلنا على نسب مئوية لكل منطقه من الصورة المصنفة. في هذه التقنية حصلنا على نتائج تصنيف جيده وسرعه في انجاز عملية التصنيف.

Introduction

Our Galactic Home (Milky way) A glance up at the night sky reveals a broad swath of light. Described by the ancients as a river, as milk, and as a path, among other things, the band has been visible in the heavens since Earth first formed. In reality, this intriguing line of light is the center of our galaxy, as seen from one of its outer arms[1].

The Milky Way is a barred spiral galaxy, about 100,000 light-years across. If you could look down on it from the top, you would see a central bulge surrounded by four large spiral arms that wrap around it. Spiral galaxies make up about two-third of the galaxies in the universe [2].

The Milky Way does not sit still, but is constantly rotating. As such, the arms are moving through space. The sun and the solar system travel with them. The solar system travels at an average speed of 515,000 miles per hour. Even at this rapid speed, the solar system would take about 230 million years to travel all the way around the Milky Way[1].

Curled around the center of the galaxy, the spiral arms contain a high amount of dust and gas. New stars are constantly formed within the arms. These arms are contained in what is called the disk of the galaxy. It is only about 1,000 light-years thick. At the center of the galaxy is the galactic bulge. The heart of the Milky Way is crammed full of gas, dust, and stars. The bulge is the reason that you can only see a small percentage of the total stars in the galaxy. Dust and gas within it are so thick that you can't even peer into the bulge of the Milky Way, much less see out the other side[3].

Variable precision rough sets

Rough sets theory is a new mathematical tool in data mining area to deal with vagueness and uncertainty data, which can analyze and deal with various imprecise and incomplete information[4,5]. However, traditional rough sets are very sensitive to even small misclassification errors which restrict its application greatly. Hence, it is necessary to increase the system redundancies. Here, we mainly introduce the Variable Precision Rough Sets (VPRS) model. And VPRS is also taken [4,6].

In conventional rough sets, universe *U* is known and conclusion is only suitable for objects belonging to *U*. It is very difficult to satisfy constrains in practice. To solve the problem, a method must be found to generalize conclusions obtained from sample data to a more wide area. VPRS is proposed by Ziarko to solve the problem[5].

Let X, Y be non-empty sets in finite field. If there is exist x ∈Y for all x ∈ X, we call that $X \subseteq Y$.

It is obviously that no misclassification errors are allowed for in the condition. A new idea is presented in VPRS which give a new measurement method on inclusion relation as follows.

$$c(X,Y) = \begin{cases} 1 - \frac{card(X \cap Y)}{card}(X) & if\ card(X) > 0 \\ 0 & if\ card(X) = 0 \end{cases} \quad (1)$$

where card(*) denote cardinal number of sets. C(X,Y) denote degree of misclassification set X into Y. That is to say, there are c(X,Y) *100% elements misclassified. Obviously, X ≤Y when c(X,Y)=0. Therefore we can give an admissible misclassification error $\beta(0 \leq \beta \leq .5)$ .According to the definition, there is:

$$Y^{\beta} \supseteq X \qquad if\ and\ only\ if\ c(X,Y) \leq \beta \quad (2)$$

Suppose that U is universe, R is indiscernibility relation on U.R^* $=\{E1,E2,\dots\dots\dots E_n\}$ are partitions of equivalent classes on U.

B- Lower approximation (β-position region of set X),

$$R_{\beta}\,X = U\,\{\,E\ \in\ R^{*}: c(E,X)\ \leq \beta\} \quad (3)$$

B- upper approximation (β-negative region of set X)

$$R_{\beta}X = U\,\{\,E\ \in\ R^{*}: c\ (E,X)\ < 1 - \beta\} \quad (4)$$

β- Boundary region,

$$BNR_{\beta}\,X = U\,\{\,E\ \in\ R^{*}:\ \beta < c(E,X) < 1 - \beta\} \quad (5)$$

B-negative region

$$NEGR_{\beta}X = U\{\,E\ \in R^{*}: c\ (E,X)\ \geq 1 - \beta\} \quad (6)$$

Ziarko give a very important definition in VPRS namely quality of classification.

$$\gamma(P,Q,\beta) = card\ (POS(P,Q,\beta))/\ card\ (U) \quad (7)$$

In which POS(P, Q, β) is a β-position region on part ion Q*

Attribute reduction and optimal set of attribute are the most important conception in rough sets model. VPRS provide us two important criteria [6],

1. $\gamma(P, Q, \beta) = \gamma(RED(P, Q, \beta), Q, \beta)$
2. No attribute can be eliminated from $RED(P, Q, \beta)$ without affecting the requirement1.

There have been many algorithms for attribute reduction. Optimal reduction can be derived from combined minimum cost criterion naturally if it is possible to assign a cost function to attributes. In the absence of attribute cost function, two basic approaches were presented by Ziarko in which optimal reduction can be determined according to the number of attributes and rules [5].

Experiment analysis

In this search we was take Astronomy image to classified using Variable Precision Rough Sets Technique, in this technique we will classified the Milky Way image in seven classes according to different region and contain such that; the old stars, new star, dust, gas, heart and arms of galaxy. The classified technique was done used Visual Basic language.

Results and discussion

A Milky Way galaxy Multiband image in this study consisting (256x256) shown in Fig.1. Fig.2 shows the histogram of Multiband image for each band, the histogram shows the Data value and its frequency distribution in original image. Fig.3 shows the classified image of seven classes, in its color is assigned arbitrary, Fig.4 shown the histogram of a segment of image data may exhibit peaks at the location of classes or clusters. Table 1 shows the statistical properties for each band of Original image. Table 2 represents the calculated number of each class and evaluated the percentage of each cluster. The total percentages for all classes not equal to 100%, perfected because there are several points are not included in any class of classified image. Therefore; the rate of this study represents the classified points which are included in the classes of classified image.

Fig.1: Original Milky way galaxy.

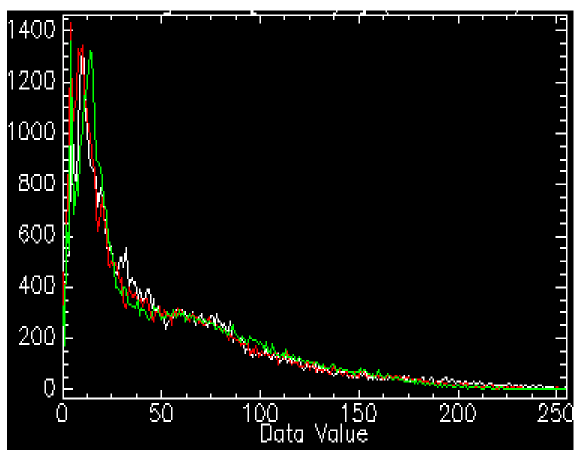

Fig.2: The histogram of the original Milky Way galaxy (three bands).

Conclusions

Obviously from classified image the following important notes.

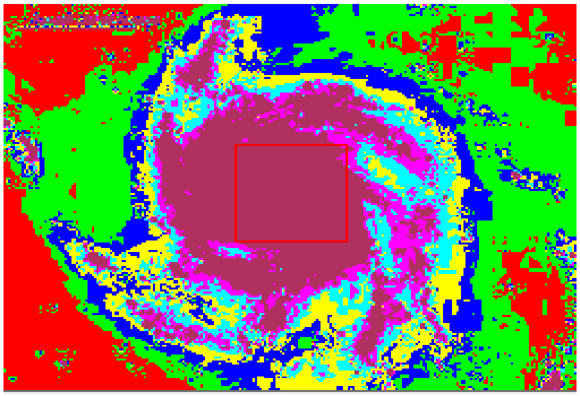

Fig.3: The classified Milky Way galaxy.

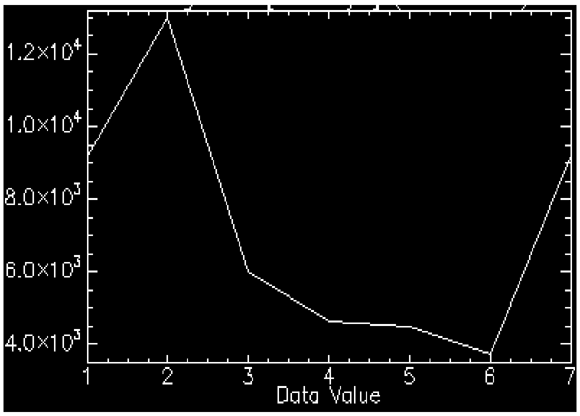

Fig.4: The histogram of classified Milky Way.

Table 1: The statistical properties of each band

Basic stats	Min	Max	Mean	Stdev
Band 1	0	255	51.963	49.238
Band 2	0	255	48.523	46.563
Band 3	0	255	50.248	45.673

Table 2: The classes, (NPTS) and percent for each one of them.

DN	NPTS	Total	Percent
1	9176	9176	18.2454
2	12973	22149	25.795
3	12973	28180	11.994
4	4641	32821	9.228
5	4476	37297	8.9
6	3714	41011	7.3849
7	9265	9265	18.311

1- The percentage of the heat of galaxy is a nearly (18.2) from the total size of the Milky Way galaxy.

2- The spiral arms of the galaxy is about (25.8%) of the galaxy. Clear the major class location in spiral arms (25.795%) that contain the a new stars are constantly formed within the arms therefore; according this result we can consider the galaxy(milky way) is one from young galaxies, more detail see[7],

3- The spiral arms contain a high amount of dust and gas. New stars are constantly formed within the arms the percentage of it is (11.99).

4- The bulge and the arms are the most obvious components of the Milky Way, but they are not the only pieces. The galaxy is surrounded by a spherical halo of hot gas, old stars and globular clusters. Although the halo stretches for hundreds of thousands of light-years, it only contains about two percent as many stars as are found within the disk.

References

[1] Eric J. Chaisson , Steve McMillan , Eric Chaisson "Astronomy: A Beginner's Guide to the Universe" (3rd Edition) Paperback – Unabridged, July 25, 2000.

[2] D. John. Fix "Astronomy Journey to the cosmic frontier" Forth Edition, McGraw-Hill, 2006.

[3] G. de Vaucouleurs, in IAU Symposium No. 58, The Formation and Dynamics of Galaxies, ed. J. R. Shakeshaft (Boston: Reidel), (1974) p.1.

[4] P. D. Heermann, N. Khazenie, IEEE Trans on Geoscience and Remote Sensing, 30 (2010)1.

[5] W. Ziarko, Journal of computer and System Sciences, 1993 (1993) 39-59.

[6] A. S. Chintan, M.K. Arora, and K. V. Pramod, International Journal of Remote Sensing, 25 (2004) 481–487.

[7] Bushra Ali "Simulation methods on Rotation curve of spiral Galaxies" Ph.D Thesis, University of Baghdad, College of Science, (2011).