CISC 871: Fuzzy Logic and Fuzzy Algorithms

Term Paper

Exploring Space with Fuzzy Logic

Author:
Brenna Douglas (6302386)

Professor:
Dr. Robin Dawes

April 28, 2012

Abstract

The applications of fuzzy logic and fuzzy algorithms to the classification of celestial objects will be explored through various prominent papers in the field. Specifically, these papers will address issues concerning star and galaxy separation, classification of stars into their spectral types, classification of galaxies into the main galaxy types, and comparing the results to their neural network counterparts. Additionally, the task of improving astronomical image processing through the use of fuzzy logic based algorithms will be explored. A brief discussion will follow regarding the application and practicality of fuzzy logic to astronomical data mining and astronomical image processing.

1 Introduction

In the past few years alone, the amount of data that has been collected for astronomical images has been growing exponentially. Ideally, these images will be used to answer fundamental questions about the universe, such as its composition, the existence of terrestrial planets with the possibility of life, or finding asteroids that pose a threat to the Earth. In addition to an increase in number of robotic telescopes, many new sky surveys, both ground-based and space-based, have been launched or are in the planning phases.

The Sloan Digital Sky Survey, which began collecting data in 2000, has since collected spectral data for more than a million objects and mapped more than 35% of the sky [3]. Pan-STARRS is another ambitious venture that went online in 2008 with the goal of covering the entire sky every 10 days while collecting 13 terabytes of data per night [7]. Additionally, space based missions, such as Kepler and JWST ([1], [2]), are also continuously adding to the astronomical amount of data already gathered.

With this vast amount of data pouring in daily, the task of identifying specific celestial objects is more than just an overwhelming task for humans; it is impractical. Most of our knowledge of galaxy classification is still based on the work of several dedicated observers. For example, one study consisted of three people inde-

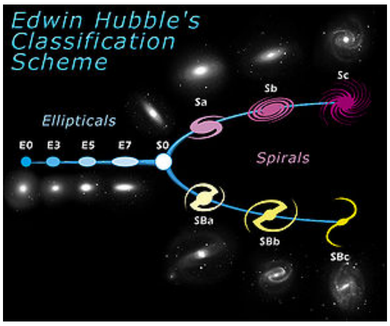

Figure 1: The Hubble sequence for galaxy classification. Invented by Edwin Hubble in 1926, the scheme divides galaxies into three classes based on appearances.

pendently classifying 2253 objects, with the final classification taken as the mean of the three observations [4]. Galaxy Zoo is yet another attempt to classify celestial objects, but makes use of the power of the internet. Upon its launch in 2007, Galaxy Zoo consisted of a data set of one million from the Sloan Digital Sky Survey. Galaxy Zoo users were then given the simple task of classifying images of galaxies into two types, and were given the necessary information to do this. The results were great; each object had multiple classification sources and it was shown that the results were as good as those completed by professional astronomers [10].

These both provide excellent databases of classified objects, however, the size of currently classified object pales in comparison to the amount of data continuously being recorded. Therefore, it is apparent that an automated classification system that performs as well as a professional astronomer is needed. As this is essentially a data mining problem, many people have already applied various data mining techniques to this task. Here we will explore the application of fuzzy logic and algorithms to this problem.

2 Separating Stars and Galaxies

The night sky contains an abundance of different kinds of objects. There are stars, planets, galaxies, clusters, black holes, and nebulae, to name a few. Even within these categories, each object has its own classification system. For example, galaxies can be classified into spiral, barred-spiral, irregular, or elliptical galaxies. Figure 1 is a common, transparent version of a classification scheme for galaxies, but not the

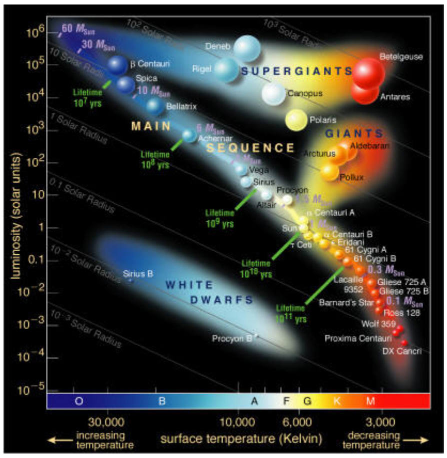

Figure 2: The Hertzsprung-Russell diagram for star classification demonstrates the relationship between spectral types, temperature and absolute magnitudes of stars.

only one. The Hertzsprung-Russell Diagram, as seen in figure 2, demonstrates one of the more common diagrams for star classification. There are other classifications schemes, for both stars and galaxies, as well as the other types of celestial objects, and for most cases, classification means taking these objects and categorizing them into discrete sets. Considering the vast amount of data as previous discussed, classifying a random object into its correct class becomes a daunting task.

Therefore, many people have taken a more simplistic approach for classification. In 2000, Mähönen and Frantti attempted to tackle this challenge by simplifying it to a problem of distinguishing stars from galaxies only. Previous works had already made attempts to use various types of neural networks for star/galaxy separation, but Mähönen and Frantti take a different approach and demonstrate an automatic classification system which uses fuzzy set reasoning [12].

[12] used an extension of the *k-means clustering* algorithm, the *fuzzy c-means* clustering algorithm. Similar to the k-means clustering algorithm, the fuzzy c-means algorithm attempts to create c cluster centers in the space, and then measures the distance of each point to the center of each cluster. Unlike the k-means clustering algorithm, the fuzzy c-means algorithm computes the degree to which each point belongs to a cluster. This allows points farther away from the center to be in the cluster to a lesser degree than those points close to the center.

Measurements for each image, such as elliplicity, average transmission over image, and the image gradients, were computed. In total, fourteen measurements were

COMPARISON BETWEEN TRUE-CLASSIFICATION PROBABILITIES			
Size	Fuzzy	BP	SOM
200	77	84	86
400	84	92	93
600	96	96	97
800	94	98	99
1000.........	95	98	99
1200.........	99	99	99

Figure 3: Taken from [12], the results of the fuzzy classifier (Fuzzy), a back-propagation network (BP), and self-organizing maps (SOM). Values are in percentages. Size represents the diamater of the object.

made and used as the parameters to define each image. Using a data set consisting of 5528 stars and 3717 galaxies, [12] compared the results of their Fuzzy classifier to two neural network approaches (using back propagation and self-organizing maps) over various subsets of the data. The results can be seen in figure 3.

Although the results are not quite as good as those from neural networks, it is pointed out that the fuzzy classifier has its own advantages. The output of the classifier can be used directly as an estimate of how reliable the classification is. For examples, classification memberships lie within the range [0,1], and those classifications which are in the middle of the range (around 0.5) are considered very unreliable. These objects can then be sent for more sensitive processing, whether it be a different classification system, or a human observer.

Additionally, the fuzzy classifier can be used for preclassification for a different classification system. The paper suggests using it as a preclassifier for their neural networks, but discusses one of the key issues is to determine how to weight the output from the fuzzy preclassification.

3 Using a Different Separation Approach

The results from [12] demonstrate the use of fuzzy logic for star/galaxy separation, but is not the only paper to do so. Longo et al. also aimed to create a classification system based on unsupervised learning to perform star/galaxy separation [11]. However, their approach was different than that of the previous paper. Their idea is to first determine two prototypes for the stars and galaxies classes to be used as references for each catalog object. Through fuzzy logic, the degree of similarity with respect two each prototype can then be computed.

To first compute the prototypes, Self-Organizing Maps ([8]) were used. A Self-Organizing Map is essentially a type of neural network, and in this usage, the first level is composed with as many nodes as there are object features. The second level is composed of two nodes, one for stars, and one for galaxies. Therefore, the final weights represent the prototypes for each class. Object features were measured from various photometic and astrometric properties of each image, and the seven most significant features were used.

Once the prototypes are created, the similarities of objects in the catalog to the prototypes is computed. To compute the similarity, the paper uses Lukasiewicz algebra. In Lukasiewicz algebra, the binary operations $\rightarrow$ and $\otimes$ (the residuum and t-norm, respectively), are defined by

$$x \rightarrow y = min\{1, \sqrt{1 - x + y}\} \tag{1}$$

$$x \otimes y = \sqrt{max\{0, x + y - 1\}} \tag{2}$$

From this, we can use the bi-residuum to interpret fuzzy logic equivalence. The bi-residuum can be defined as

$$x \leftrightarrow y = (x \rightarrow y) \wedge (x \leftarrow y) \tag{3}$$

Using Lukasiewicz algebra, this becomes

$$x \leftrightarrow y = 1 - max(x, y) + min(x, y) \tag{4}$$

Using $\mu_X(x)$ to represent the membership of element x in set X, then fuzzy similarity S can be described by

$$S(x, y) = \mu_X(x) \leftrightarrow \mu_X(y) \tag{5}$$

Then, to compute the total fuzzy similarity over all n features, we use the equation

$$S(x, y) = \frac{1}{n} \sum_{i=1}^{n} S_i(x, y) \tag{6}$$

Using 6, we now have a means to compute the similarity between the prototypes. Once computed, we have a measure of similarity between the objects in the catalog and the two prototypes. The final step for classification is to defuzzify. That is, an object from the catalog is classified as a member of the star class or the galaxy class based on the maximum similarity with respect to each prototype.

For the actual experiment, 10,000 objects were randomly chosen from a catalog of 231,000 labeled star and galaxy objects. They obtained a 77.38% classification, with 2867 galaxies and 4871 stars correctly classified, in incorrectly classified 590 star objects and 1672 galaxy objects. These results at least perform better than chance, however, for practical use it is not reasonable. Additionally, the method from the previous paper seemed to perform much better. Nevertheless, these were preliminary results and the potential for improvement exists.

4 Galaxy Morphology

While the previous papers aimed to simply separate star and galaxy objects into their respective classes, others attempted to broaden the abilities of their algorithms. In a more recent paper, Gauci et al. aspired to design an intelligent algorithm with the same accuracy as humans that could distinguish between spiral galaxies, elliptical galaxies, and stars or unknown objects. Here, the input object requirements are relaxed, as it is not necessary for the object to be a star or galaxy object. While classifying galaxies into elliptical or spiral galaxies is a bit simplified (see Figure 1 as an example of more detailed classification), it clearly is much more complex than simply determining a galaxy.

As mentioned before, most of our current knowledge of galaxy classification is based on the work of several dedicated observers who visual inspect and catalog thousands of galaxies. The Galaxy Zoo project has provided quite an extensive collection of about a million labeled objects. This is a great accomplishment and

GZ	FIS (Subtractive Clustering) E	S	U
E	**96.68 %**	3.25 %	0.07 %
S	7.85 %	**92.09 %**	0.06 %
U	18.58 %	26.11 %	**55.31 %**

Classifier	Accuracy
FIS (Subtractive Clustering)	94.568 %

Figure 4: Taken from [5], the results of the fuzzy inference system.

extremely useful to the field of astronomy and astrophysics, as the results can be used for data sets for experiments, such as new machine learning algorithms for astronomical data.

A million may seem like a large number, but the distribution of objects is not even. We have many bright, old galaxies, but we have a limited number of dim galaxies, dark-matter galaxies, or dwarf galaxies, simply because they are harder to see. Likewise, young galaxies and proto-galaxies that formed soon after the big bang are not well documented compared to common spiral or elliptical galaxies, either. Additionally, even for humans it can be difficult to distinguish what classification to use for a galaxy. For example, when two galaxies collide, depending on the extent of the current collision, it might be difficult to distinguish if they should be classified as two separate galaxies, or one galaxy.

While we may not have as many examples of these, they are certainly interesting and obviously offer a lot of useful information about the universe. Therefore, although a million objects is a lot, there are a lot of categories of galaxy classes that are completely underrepresented. It then becomes apparent that intelligent algorithms for automated classification is a huge challenge and being able to classify galaxies into two subcategories is a great step in the right direction.

[5] compared results from a fuzzy inference system and results from various decision tree algorithms. Specifically, they used Random Forests, the Classification and Regression Tree (CART) scheme, and the C4.5 decision tree learner. They took their training and testing samples from Galaxy Zoo catalog, and computed various photometric and spectral attributes for each object, using a total of thirteen attributes. Such measurements include (but not limited to) the DeVancouleurs fit

		CART		
		E	S	U
GZ	E	**97.28 %**	2.69 %	0.03 %
	S	5.05 %	**94.74 %**	0.20 %
	U	5.15 %	11.32 %	**83.53 %**

		Random Forest (10 trees)		
		E	S	U
GZ	E	**98.23 %**	1.75 %	0.02 %
	S	4.73 %	**95.17 %**	0.11 %
	U	3.97 %	11.77 %	**84.27 %**

		C4.5 (0.25 confidence)		
		E	S	U
GZ	E	**97.18 %**	2.79 %	0.03 %
	S	5.02 %	**94.80 %**	0.18 %
	U	4.41 %	10.15 %	**85.44 %**

		Random Forest (50 trees)		
		E	S	U
GZ	E	**98.21 %**	1.77 %	0.02 %
	S	3.78 %	**96.10 %**	0.12 %
	U	3.24 %	10.15 %	**86.62 %**

		C4.5 (0.1 confidence)		
		E	S	U
GZ	E	**97.31 %**	2.66 %	0.03 %
	S	5.01 %	**94.82 %**	0.17 %
	U	4.56 %	10.00 %	**85.44 %**

Classifier	Accuracy
CART	96.227 %
C4.5 (0.25 confidence)	96.203 %
C4.5 (0.1 confidence)	96.288 %
Random Forest (10 trees)	96.979 %
Random Forest (50 trees)	97.331 %

Figure 5: Taken from [5], the results of the various decision tree algorithms used.

axis ratio, concentration, star log likelihood, and adaptive fourth moment.

Of interest to this paper is their use of fuzzy logic. They made use of a fuzzy inference system, where *if-then* rules that deal with fuzzy consequents and fuzzy antecedents are defined. First, the degree of truth for each antecedent is computed, which then can be used to compute the truth of each consequent. The resulting consequents are then weighted and combined by standard logical operators. The standard *min* and *max* functions are used as the t-norms and s-norms, and the final result is defuzzified to obtain one value. Disappointingly, the paper did not include the inference rules used. Although it would probably require a lot of background knowledge of properties of each class (elliptical galaxies, spiral galaxies, stars), examples would have been appreciated.

The results for the fuzzy inference system can be seen in figure 4. It was able to correctly classify 96% of the elliptical galaxies and 92% of spiral galaxies, which is good news. However, it was only able to classify 55% of the unknowns correctly. Looking at the confusion matrix, it is apparent that for both galaxy types, they were rarely classified as unknowns incorrectly. The unknowns did slightly better than chance (chance would be 33%), and the overall classification had an accuracy of 94%. Depending on the goals for the output, this could be a reasonable result.

For comparison, the corresponding decision tree algorithm results can be seen in figure 5. We can see that the decision trees have an overall accuracy of at least 96%. Additionally, unknowns are all classified correctly at least 83% of the time. From these results we could suspect that decision trees provide a better algorithm than fuzzy inference trees.

However, it is important to note that these calculations were based off of parameters that were calculated in the i-band. That is, they were measured in the infrared range. The paper also performed calculations in the r-band (the red range) input parameters, as well as spectra input parameters, and classification using the various decision tree methods were performed with these input parameters. It is interesting that they did not use these measurements towards the fuzzy inference system. Perhaps it was due to only a slight improvement in overall classification (about 0.5%) for decision trees, although no explanation was given.

It seems as though there is a lot of room for improvement and growth in this particular application. Indeed, the paper does discuss that further study could be apply to the fuzzy inference system, to both improve the rules used and the final results. It was not explained why the spectral parameters and the r-band parameters were not used, but this is another area that could be explored. Additionally, like many other papers, they suggest that this technique could be used as a preliminary classification technique. Therefore, it would suggest that hope is not yet lost for the application of fuzzy logic and astronomical object classification.

5 Fuzzy Reasoning and Stars

The previous papers are essentially various ways to separate stars and galaxies, but do not get much more complex. Rodriguez et al. look at the problem of classification differently. Their goal is to create an intelligent system for the analysis and classification of low-resolution optical spectra of super giant, giant and dwarf stars, with luminosity levels I, III, and V, respectively [14].

Refer to 2 to see these classifications. Essentially their goal is to determine main sequence stars from giants and from super giants, but ignore the white dwarfs. A luminosity level of II is for stars classified as "bright giants" and a luminosity level of IV is for stars classified as "sub-giants" (these are not in figure 2). Both of these classes straddle the boundary between super giants and giants, and giants and main sequence stars. It makes sense to avoid attempting to include them in the classification in the beginning stages of start classification, as their classification can be tricky, even for an expert.

Using stellar spectroscopy, the physical conditions (temperature, pressure, etc.) and chemical components of stars can be measured. A stellar spectrum for a star can be collected by using a telescope with the appropriate spectrographs and detectors mounted. In addition to the luminosity level classes, stars can be put into classes based on their temperature using the Morgan-Keenan system [9]. These classes are O, B, A, F, G, K, and M, with O class stars being the hottest, and M class stars being the coolest. (This can also be seen in figure 2). Stars that belong to specific temperature and luminosity classes will produce different spectrographs. Using luminosity, temperature, and spectral type, a star can be classified into the classes super giant, giant and dwarf (or main sequence) stars.

Again, as is the case with most astronomical data mining, classifications based on the spectrographs were mostly carried out by hand by dedicated experts, and it is very time-consuming and requires a lot of human resources. Therefore, the authors created a process to simulate the behaviour of human experts using fuzzy and knowledge-based reasoning based on a set of uncertain classification criteria derived from previous experience. From the spectral features, ten molecular bands and nine emission/absorption lines and their relationships were found to be the main reasoning criteria used by human experts for manual processing of stellar classification.

Typically, human experts visually observe the spectral features and obtain a preliminary classification, which includes the spectral type, the luminosity and the global group (early, intermediate, and late). From there, they compare each spectrum to the reference catalog to obtain the spectral sub-type. Sometimes, it is difficult to obtain the spectral sub-type interval, in which case they are classified with two numbers to indicate the sub-type interval. Because the human reasoning itself includes uncertainty and imprecision, it is easy to see that fuzzy logic could

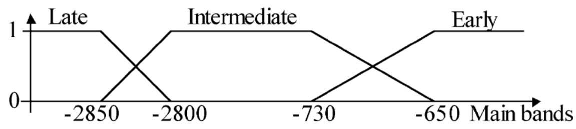

Figure 6: Taken from [14], an example fuzzy variable with its corresponding fuzzy set for 'Global clssification in luminosity I'.

Technique	Global temperature (%)	Spectral type (%)	Luminosity (%)
Human expert A	99	92	81
Human expert B	95	85	70
Expert systems	96.5	88	65
Expert Systems with fuzzy logic	98.6	90.3	78.2
Backpropagation networks	97	95.4	81

Figure 7: Taken from [14]. The performance of the automatic classification techniques and two human experts.

be of use.

Rodriguez et al. used fuzzy *if-then* production rules to model the reasoning followed by experts in the field. The antecedent conditions of these rules refer to the values of the measured parameters (which are stored in a current facts base). The consequents allude to the three levels of spectral classification. Fuzzy sets and membership functions were determined by the values of the spectral features in the guiding catalog spectra. That is, they defined as many fuzzy variables as there were classification levels for each luminosity class. An example can be seen in figure 6.

Additionally, the outputs from each consequent were weighted based on their importance to the classification learned through experience and knowledge of human experts. The results from each rule were combined using the Max-product method, and the defuzzification of the data into a crisp output was performed using the fuzzy-centroid method, which essentially favours the rule with the greatest area. The system then can take as inputs the values of the spectral features and have a leveled classification (I, III, or V) with an associated truth-value and the explanation of the system's reasoning as output. If the truth value for a classification is significantly small, and alternative classification could be included.

Similar to the other works, this paper also demonstrated the application of neural networks. In this case, they used backpropagation networks. In addition to the expert system with fuzzy logic, as described above, they included results

Type/luminosity	Spectra number	Success rate (%)
B	45	92
A	22	88
F	36	91
G	25	80
K	34	89
M	38	93
I	66	98
III	56	89
V	78	82

Figure 8: Taken from [14]. The system performance for spectral types and luminosity.

from an expert system that did not incorporate fuzzy logic. The results can be seen in figure 7. Additionally the expert system with fuzzy logic performance for temperature and luminosity is shown in figure 8. The expert systems were able to classify stars with an error rate below 20%, and errors that were made were usually explainable. For example, as seen in figure 8, spectral types G, A, and luminosity type V have high error rates, but the catalog also contained few stars of these types. More data for these stars would lead to more refined features used and better fuzzy rules.

Figure 7 shows the overall results. Neural networks are shown to provide better results for the spectral types and luminosities, whereas expert systems with fuzzy logic are more suitable for classifying the star's global temperature. Overall, each system reached a global success rate of around 90%. It is interesting to note that, in some cases, the automatic classification techniques performed slightly better than the two human experts. The results are promising, and future works discussed by the paper include incorporating the expert system with fuzzy logic and the back-propagation network to improve results.

6 Improving Astronomical Image Processing

The previously discussed papers all aimed to automate classification of astronomical images in some way. However, obtaining automatic identification of astronomical objects is also a challenge that has been exacerbated by the rise of digital imaging and the vast amount of astronomical data. Questions about the universe, such as the abundance of Earth-like planets, are being explored with robotic telescopes. The success largely depends on the accuracy of automated real-time processing of images, never seen by humans, to distinguish between known astronomical objects and new astronomical objects. This is a difficult task itself, as many objects are extremely faint, objects are moving (such as asteroids or comets), and objects can change (such as a massive stars turning supernova, and then forming a nebula). Objects in space change all the time, but this is not the only problem of astronomical imaging. The equipment used can make this problem more challenging as well. Slight shifts in the orientation of the camera, imperfections in the CCD, and inaccuracies of the optics are just a few problems that can arise.

Typically, the process for pipeline processing of astronomical images depends on algorithmic decisions. This requires detecting and isolating single objects in the image. The first step is usually to locate objects already known in a catalog. Computing the topocentric coordinates at a specific time of a known star and then transforming those coordinates to image coordinates would be a simple way to find the expected locations of any star in the frame. However, as previous mentioned, other factors make this a difficult challenge.

Shamir and Nemiroff made it their objective to solve this problem in their 2005 paper, [15] (as well as their 2006 paper, [16]). They present an algorithm which uses fuzzy logic to transform celestial coordinates into (x, y) image coordinates, despite various noise challenges (specifically for wide-angle non-linear optical distortions, slight optical imperfections, and small unrecorded shifts in orientation).

The first step of the algorithm is to manually identify reference stars. It is important to note that, while this process does involve humans, the task of locating a few familiar reference stars in a frame is trivial compared to identifying all stars in the frame. The next step is to build the fuzzy logic models (the rules). The models are based on the following two transformation functions:

$$f_1 : azimuth \longmapsto angle \tag{7}$$

$$f_2 : altitude, azimuth \longmapsto distance \tag{8}$$

The first rule has one antecedent variable, *azimuth*, and one consequent variable, *angle*. The second rule takes two antecedents, *azimuth* and *altitude*, and has

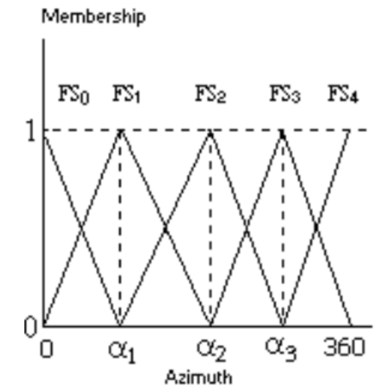

Figure 9: Taken from [15]. The example membership functions of four fuzzy sets (FS_0 to FS_1) created for the four example reference stars for equation 7. Assumes that $\alpha_1 < \alpha_2 < \alpha_3$.

one consequent variable, *distance*. Azimuth is a measurement in angles from a reference point (usually North) and the perpendicular projection of the object onto the horizon. Altitude is the measurement of the height of the object above the horizon, also measured in degrees. These are easily computed for known stars at a given time. The *angle* and *distance* are both measurements of the camera, in degrees and pixels (respectively).

For each reference star, a fuzzy set and a fuzzy rule is added to the model. Figure 9 is an example of the fuzzy logic model for equation 7 using four reference stars with (azimuth, altitude, angle, and distance) of $(0, \epsilon_0, \theta_0, R_0)$, $(\alpha_1, \epsilon_1, \theta_1, R_1)$, $(\alpha_2, \epsilon_2, \theta_2, R_2)$, and $(\alpha_3, \epsilon_3, \theta_3, R_3)$. In general, membership functions are built as triangles that peak at the reference value, and intersects the x-axis at the reference value of its neighboring points. Because of this, almost all azimuth values belong to two fuzzy sets (the exception are those at the maximum of a particular fuzzy set). Additionally, each reference star i adds the rule

$$FS_i \longmapsto \theta_i \tag{9}$$

To compute the angle for any other stellar object, it is simply a matter of computing the membership value for each fuzzy set. The results of the rules are then combined using a weighted average defuzzification method. The results are similar to those of a linear interpolation [15]. The technique for building the model of equation 8 is similar.

This fuzzy logic based transformation algorithm has been tested and is in practical use with the Night Sky Live project [13]. The Night Sky Live project uses fish-eye lens cameras to take images of the entire night sky, which are then analyzed. The data is made freely available for scientific or public use. It continuously tracks the objects in the night sky, and any non-cataloged bright objects are immediately detected.

Before this algorithm was used, a previous Night Sky Live identification algorithm employed a straight forward analytic transformation and was only accurate for stars with a magnitude of about 3.5 in best cases. The current algorithm has dramatically improved the identification and has practically 100 percent accuracy for identification of stars down to a magnitude of 5.6. To put this in perspective, magnitudes measure the brightness of stellar objects. The more positive the magnitude, the more faint the object is. The sun is currently the brightest object in the sky, with a magnitude of about -27. The full moon has a magnitude around -13. The brightest star in the night sky, Sirius, has a magnitude of -1.4. Objects with magnitudes higher than 6 are no longer visible to the naked eye.

These are obviously exciting results. For future works, the authors aim to improve the accuracy for stellar objects with even fainter magnitudes. The Night Sky Live project's equipment can currently only produce images of objects brighter than a magnitude of 6.8, so to improve the algorithm to include these faint objects would be a first step. Also, it would be fascinating to see how this technique works with deep and ultra-deep field telescopes, such as the Hubble telescope, which have been able to take images of objects created soon after the big bang [6].

7 Discussion and Conclusion

The need for an automated data mining system for astronomical data is clearly apparent. However, the task does not seem to be as straight forward as simply creating an algorithm that will classify any celestial object. Many papers have attempted to tackle smaller portions of the problem. As discussed in this paper, various works have attempted to use fuzzy logics, including fuzzy similarity, fuzzy inference, and fuzzy clustering algorithms, to tackle the problem of star and galaxy separation ([5], [12], [11]). Other works attempt to choose even further sub-problems, such as classifying specific types of object. [14] attempted to do this with stars. The results demonstrate that it is possible to create automatic classification techniques for stars that perform as well as human experts

Additionally, improving classification algorithms is not the only way people have tried to improve the data mining of the astronomical data constantly pouring in. Works, such as [15] and [16], aimed to improve the images obtained by telescopes. In this way, classification techniques would have less interference from noise due to shifts in orientation, wide-angle non-linear optical distortions, and alight optical perfections.

There are no known techniques that have the ability to be as good as human observers when it comes to general classification, but there are many techniques that do very well at solving the sub-problems of classification. Additionally, fuzzy logic has been shown to produce reasonable results when compared to non-fuzzy logic algorithms counterparts and have also been put to practical use. While using fuzzy logic alone might not be the optimal solution, integrating fuzzy logic as a preclassifier or using it for improved astronomical image processing could lead to an overall optimal result. Current applications of fuzzy logic are promising, and also open the door to many new future works and applications. There remains much to be studied about the applications and practicality of fuzzy logic and its employment for astronomical data mining.

References

[1] Borucki, W.J. et al. ”Kepler Mission: a mission to find Earth-size planets in the habitable zone.” *Proceedings of the Conference on Towards Other Earths.* (2003): 69–81.

[2] Clampin, M. et al. ”Status of the JWST observatory design.” *American Astronomical Society Meeting 203, 121.01.* 35 (2003): 1408.

[3] Donald G. York et al. ”The sloan digital sky survey: technical summary.” *Astronomical Journal.* 120.3 (2000): 1579-87.

[4] Fukugita, M. et al. ”A catalog of morphologically classified galaxies from the sloan digital sky survey: North equato-rial region.” *The Astronomical Journal*, 134 (2007): 579-593.

[5] Gauci, Adam, Kristian Zarb Adami, John Abela. ”Machine Learning for Galaxy Morphology Classification” *Computer.* 8 (2010): 1-8.

[6] ”Hubble's Deepest View Ever of the Universe Unveils Earliest Galaxies”. March 9, 2004. NASA. Retrieved 2008-12-27. <http://hubblesite.org/newscenter/archive/releases/2004/07/>

[7] Kaiser, Nicholas. ”Pan-STARRS: a wide-field optical survey telescope array.” *Proceedings of the SPIE.* 5489 (2004): 11-22.

[8] Kohonen, T. *Self-Organizing Maps.* Berlin: Springer, 1995.

[9] Kurtz, M. J. *The MK process and stellar classification of stars.* Toronto: David Dunlap Observatory, 1984.

[10] Lintott, Chris J. et al. ”Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey”. *Monthly Notices of the Royal Astronomical Society.* 389.3 (2008): 1179-1189.

[11] Longo, G et al. “Advanced data mining tools for exploring large astronomical databases.” *Proceedings of SPIE.* 4477 (2001): 61-75.

[12] Mähönen, Petri, Tapio Frantti. ”Fuzzy classifier for star-galaxy separation.” *The Astrophysical Journal.* 541.1 (2000): 261-263.

[13] Nemiroff, R.J and L. Shamir. ”Toward a Continuous Record of the Sky.” *Astronomical Society of the Pacific.* 3.761 (1999): 886-897.

[14] Rodrıguez, Alejandra, Bernardino Arcaya, Carlos Dafontea, Minia Manteigab, and Iciar Carricajob. ”Automated knowledge-based analysis and classification of stellar spectra using fuzzy reasoning.” *Expert Systems with Applications.* 27.2 (2004): 237–244.

[15] Shamir, Lior, and Robert J Nemiroff. “A fuzzy logic based algorithm for finding astronomical objects in wide-angle frames.” *Publications of the Astronomical Society of Australia.* 22 (2005): 111-117.

[16] Shamir, Lior, and Robert J Nemiroff. ”Astronomical pipeline processing using fuzzy logic.” *Applied Soft Computing.* 8.1 (2006): 79-87.