Fuzzy Classification Improvement by a Pre-Perceptual Labelled Segmentation Algorithm

Ana Del Amo
Smiths Aerospace ES-GR
3290 Patterson Avenue, SE
Grand Rapids, MI 49512-1991, USA
ana.del.amo@smiths-aerospace.com

Pilar Sobrevilla and Eduard Montseny
Technical University of Catalonia,
Barcelona, Spain
eduard.montseny@upc.es
pilar.sobrevilla@upc.es

Javier Montero
Mathematics Faculty
Complutense University, Madrid, Spain
monty@mat.ucm.es

Abstract

The goal of this paper is to present how two different image processing approaches can be enhanced by merging both methodologies. We will see how the results of a perceptual labelled segmentation methodology [7] can be improved by applying a fuzzy classification algorithm [2] based on a fuzzy outranking methodology [9] as a postprocessing algorithm, and viceversa. A comparison of the individual algorithms with the combination of both algorithms will be presented in order to demonstrate the improvement. Color Bone Marrow [1] images will be used. The objective is to detect White Blood Cells. The detection of white blood cells in bone marrow microscopic images presents big difficulties because of the great variance in their characteristics and also because of staining and illumination inconsistences. On the other hand, the maturity classes of white blood cells actually represents a continuum; cells frequently overlap each other, and there is a fairly wide variation in size and shape of nucleus and cytoplasm regions within given cell classes.

---

[1]The images are from a database collected at the Ellis Fischel Cancer Center, University of Missouri

1 Introduction

Image segmentation is the process of partitioning an image into meaningful regions that correspond to objects within the scene. The partitioning process usually stops when all regions of interest have been identified. This is easy to achieve in certain controlled applications but not that easy when the outcomes are unknown. Segmentation has a unique position within a machine vision system since it can be said that it forms the bridge between low-level and high-level processing operations.

Image classification is an extremely complex problem and accuracy is a very important issue. The simplest form of digital image classification is to consider each pixel individually, assigning each to a class based upon several values measured in separate spectral bands. This type of classifier is usually referred to as a *spectral* or *point* classifier. We will apply an unsupervised fuzzy classification algorithm designed as a multicriteria decision-making problem based on fuzzy set theory.

The combination of both approaches will result in an improvement for automatic detection. The classification system we are proposing here would have an initial preprocessing part were an RGB to Smith's HSI transformation [13] will be performed before a pixel by pixel classification.

2 Selection of color patterns

The pattern's colors are to be characterized by their *hue* and *saturation* and should be distributed within the H-S map in such a way that adjacent regions should have significantly different values with respect to the characteristics on which they are uniform. Five saturation subintervals labelled Very Low -VL-, Low -L-, Medium -M-, High -Hi-, and Very High -VHi- are being used. Fifty-three colors arranged as one color belonging to the first subinterval, four within the second one, and the other forty eight colors distributed among the Medium, High and Very High subintervals are considered. Once the pattern colors C_k have been selected, the system will characterize their chromatic by means of fuzzy sets $\mu_{C_k}\{1 \le k \le 53\}$. Two membership functions $(\mu_{C_k}^S, \mu_{C_k}^H)$ will depict the color pattern typicality of both chromatic components, and would be aggregated using the geometric mean in the expression below in order to define the pattern color's membership function:

$$\mu_{C_k}(p_{ij}) = \sqrt{\mu_{C_k}^S \times \mu_{C_k}^H} \qquad (1)$$

According to Robertson [11], perceptual variables *hue*, *saturation* and *lightness* are related to the psychophysical variables *dominant wavelength*, *dye*, *purity and luminance*, respectively. Perceptual variables are dependent on the perceived environment since, for the same color stimuli, their values may vary with surrounding light changes. In order to represent human color perception, a transformation from RGB to HSI (Hue-Saturation-Intensity) has to be performed. Therefore, a color-coordinate system in compliance with human color perception is obtained. There is a plethora of these transformations in Computer Vision references ([10], [11], and [15]) with different notations (HSV, HSB, IHS, HIS, etc.). However, the general idea is basically the same in all of them, i.e., one component for the Lightness value (I, V) and two other components for the Hue and Saturation of the color. After a research on the behavior of some HSI spaces under different conditions of illumination and noise, we have chosen the Smith's HSI space [13]. In this model, components are represented in a perfect cylinder, and transformation from RGB to HSI is provided by linear equations, which makes the calculus easier and reduces the computational cost. Besides showing the maximum component reliability under illumination variations, the Smith's components present three clear advantages with respect to RGB components:

1. HSI provides high independence between their color components.

2. HSI codes human meaningful information (color dye, chromaticity degree, etc.).

3. HSI allows more efficient isolation to physical illumination phenomenon (shades, highlights, etc.).

3 Stability-Based Hue and Saturation Color Model

After studying the behavior of Hue and Saturation Smiths' components, under different illuminance conditions, their deviation predictors E_H and E_S have been defined by equations (2) and (3), respectively. In these expressions I and S are the intensity and saturation values of the considered pixel.

$$E_H = \frac{20000}{I \times S} \qquad (2)$$

$$E_S = \frac{500}{I} \qquad (3)$$

These predictors provide a degree of confidence of any H-S pair, and allow approximating color components deviations for any particular (individual) pixel, which will be particularly useful within the pixel classification process, considering components' variability. Based on deviation predictors and considering that the bigger the deviation, the smaller the stability, Hue and Saturation stability functions, noted by F_H and F_S, were introduced in [12] by (4) and (5) respectively. These functions allow assuming the variability of the chromatic components, and will be used for obtaining the similarity degrees of color image pixels and system colors.

$$F_H = \min\left\{\frac{k_H}{E_H}, 1\right\} \qquad (4)$$

$$F_S = \min\left\{\frac{k_S}{E_S}, 1\right\} \qquad (5)$$

The scalar factors k_H and k_S appearing within previous formula can be chosen to suitably fit real data degree of variance.

4 Fuzzy Classification Algorithm

In this section, we will present a schema of the algorithm in general terms. The algorithm follows three steps:

a.- RGB to Smith's HSI transformation.

b.- Color patterns membership degrees computation.

c.- Pixel by pixel image classification.

Step (a) follows the approach presented in [13] to transform the RGB values into Smith's HSI. Step (b) considers the Hue and Saturation stability functions ((2), (3)) in the final Hue and Saturation membership function computation as follows:

$$\mu_{C_k}^H(p_{ij}) = \mu_{C_k}^H(h_k - (h_k - h_{ij}) \times F_H(p_{ij})) \qquad (6)$$

$$\mu_{C_k}^S(p_{ij}) = \mu_{C_k}^S(s_k - (s_k - s_{ij}) \times F_S(p_{ij})) \qquad (7)$$

They represent the similarity degrees of Hue and Saturation values of pixel p_{ij}. h_{ij} and s_{ij} are the Hue and Saturation values of pixel p_{ij}, h_k and s_k are the central values of the corresponding Hue and Saturation membership functions, and $F_H(p_{ij})$ and $F_S(p_{ij})$ are the Hue and Saturation stability values for pixel p_{ij}. Each pixel p_{ij} in step (c) is then labelled according to the following rules:

$$\mu_{C_k}(p_{ij}) = \begin{cases} \max_{1 \le l \le 53} \{\mu_{C_l}(p_{ij}) \,|\, \mu_{C_l}(p_{ij}) > 0.15 \\ if \quad \mu_{C_l}(p_{ij}) > 0.15\} \quad \forall l \,/\, 1 \le l \le 53 \\ 0 \\ if \quad \mu_{C_l}(p_{ij}) \le 0.15 \quad \forall l \,/\, 1 \le l \le 53 \end{cases}$$
$$(8)$$

Where μ_{C_k} has been obtained as established by (1). Once the perceptual labelled segmentation has been accomplished, we apply a fuzzy point classifier. Let

$$X = \{X_{ij} \,/\, i = 1, \cdots, r \quad j = 1, \cdots, s\}$$

be the image set to be processed[2]. Let $(x_{ij}^1, x_{ij}^2, \cdots, x_{ij}^n) \in I\!\!R^n$ be the vectorial representation of a collection of features for element X_{ij}, in which x_{ij}^f is the value of feature f for element X_{ij}. We are also considering the expert knowledge, so, we are aware of the existence of three main different kinds of objects in the image. Therefore, three fuzzy classes will be considered. For each one of the classes, in which the classification will be performed, a range of valid values has to be defined. The membership function for each class C_k with respect to each f property is defined following [2]. Each object X_{ij} has an associated vector $M_k(X_{ij}) = (m_{1k}(x_{ij}^1), m_{2k}(x_{ij}^2), \cdots, m_{nk}(x_{ij}^n))$ for each class C_k, which shows the different degrees of verification each property has with respect to each class. The classification problem is at this point, solved as a multicriteria decision making problem following an outranking methodology. Each one of the pixels in the pattern image will have a representation as follows: $\mu_{ij}^u(C_k) = \mu_{ij}^{u^{ND}}(C_k) \cdot b_{ij}(C_k)$ as the membership degree to class C_s. Where $\mu_{ij}^{u^{ND}}(C_k)$ is the fuzzy set of non-dominated classes representing the

<hr>

[2] X_{ij} represent each one of the already labelled pixel $\mu_{C_k}(p_{ij})$

membership degree of pixel X_{ij} to class C_k. $b_{ij}(C_k)$ is basic pixel information. Pixel X_{ij} will verify the properties of class C_k with an intensity of

$$b_{ij}(C_k) = \min_f \left\{ m_{fs}(x_{ij}^f) \right\} \qquad (9)$$

where $m_{fs}(x_{ij}^f)$ is the membership function for each class (as defined above) C_k with respect to each f property.

## 5 Comparison of each algorithm with the merging algorithm

In order to compare the efficiency of the algorithm a bone marrow image will be processed. As has been previously mentioned, segmentation is defined as a decomposition of an image into homogeneous regions. In this context, the typical elements in a cell image are nucleus, cytoplasm, red blood cell, background, etc. The objective is to present how the segmentation of the image can be improved by a fuzzy classification algorithm that considers the color characteristics of each pixel in the image. And also how the classification algorithm can be improved by a preprocessing labelled segmentation algorithm. However, the final objective for combining both algorithms is to be able to detect White Blood Cells (WBC) more precisely.

WBC in bone marrow images show very changeable features due to the different maturity levels. The larger the cell the more immature it is. Cells become progressively smaller as they mature. The same way the nucleus of immature cells in the maturation sequence is relatively large in relation to the cytoplasm. As the cells age, the absolute and relative size of the nucleus decreases. In addition to this, cells frequently overlap each other. So, WBC can appear totally or partially hidden by other WBC or any other elements in the image. The goal here is to be able to automatically detect WBC, having a professional description of what can be identified in the image.

Figure 1 represents the original image, where, besides white cells of different maturity levels, there are nucleated red blood cells, platelets and various types of grungy. Usually, immature white cells can be characterized by a round light nucleus. The rounder and lighter the nucleus is with respect to the cytoplasm, the less mature the cell is. Mature white cells can be characterized by a darker nucleus and a lighter cytoplasm. Figure 2 represents the image processed by the segmentation algorithm [7]. Figure 3 represents the image processed by the fuzzy classification algorithm in [2]. In the original image, figure 1, the nucleus can
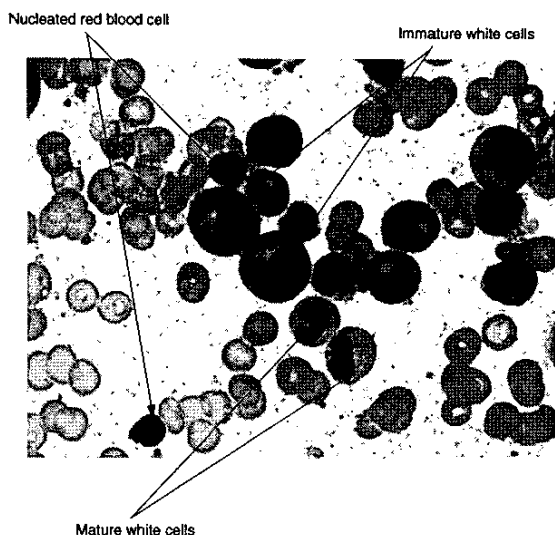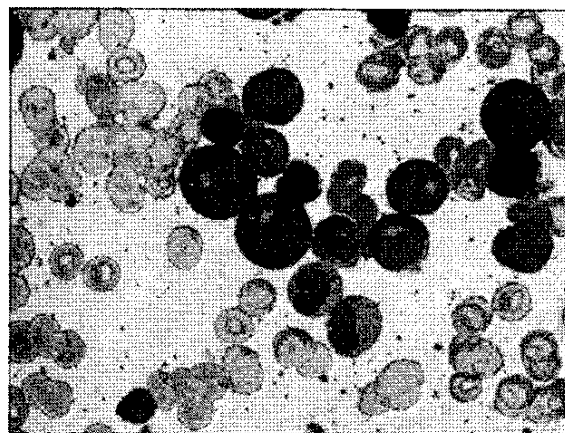
Figure 1. Original Image
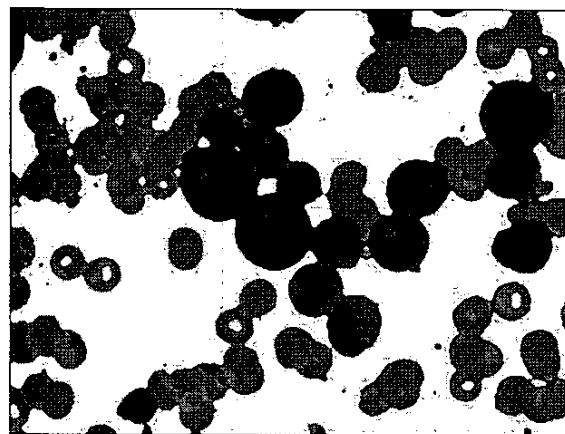
Figure 2. Segmented Image

be identified in all the cells as a blueish part. But there is not a remarkable difference between nucleated red blood cells and immature white cells. The mature white cells are identified by a lighter cytoplasm. The segmented image represents an improvement from the original image since the cytoplasm has been discriminated from the nucleus, but white mature cells and red cells are still represented by the same colors.

The nucleated red blood cells marked in the top part of the original image are classified as a whole nucleus cell, while the cell in the bottom part could be identified as just cytoplasm since is been colored as the cytoplasm part in the white cells. Figure 4 represents the image that has been preprocessed by the perceptual segmentation algorithm [7] and the fuzzy classification algorithm [2]. This image represents an improvement over figures 2 and 3. For example, mature and immature white cells have been classified in the same class and can be clearly distinguished from the immature cells represented in a lighter color. It can also be seen how the nucleolus can be distinguished in both mature and immature cells, being very noticeable the different rounded shape. As mentioned above, immature cells have a rounder and lighter nucleolus as can be clearly seen in figure 4.

Figure 3. Image Process by fuzzy clustering algorithm

6 Final comments

The goal of this research was to combine two image processing algorithms based on different approaches to obtain a more precise classification of the im-

489
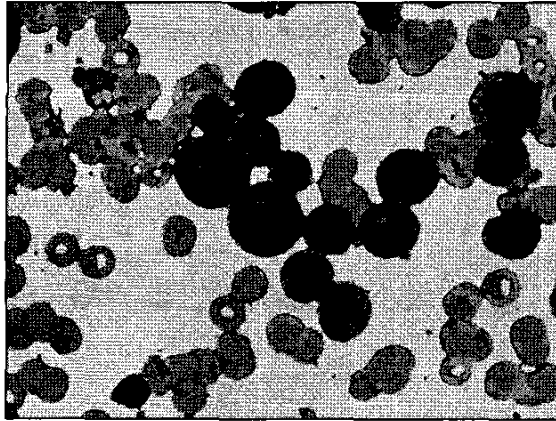
Figure 4. Segmented Image Process by fuzzy clustering algorithm

age objects. The results demonstrate that both approaches are enhanced by combining then in the proposed methodology. Also the advantage of using the fuzzy classification algorithm as a postprocessing has the advantage that we don't need to define rules to identify the colors of the different elements. Therefore, the problems that arise because of the variation of colors in the different elements of the image due to the dyes, their quality, age, etc are avoided.

References

[1] J.C. Bezdek, and S. K. Pal *Fuzzy Models for Pattern Recognition.* New York:IEEE Press 1992.

[2] A. Del Amo, J. Montero and G. Biging, "Classifying pixels by means of fuzzy relations," *Inter. J. General Systems*, Vol. 29, pp. 605–621, 1999.

[3] A. Del Amo, J. Montero, A. Fernandez, M. Lopez, J. Tordesillas and G. Biging, "Spectral fuzzy classification: an application," *IEEE Trans on Systems, Man and Cybernetics, Part C*, Vol. 32,pp. 42–48, Feb. 2002.

[4] A. Del Amo, J. Montero, and E. Molina, "Representation of consistent recursive rules", *European Journal of Operational Research*, Vol. 130, pp. 29–53, 2001.

[5] A. Del Amo, D. Gómez, J. Montero and G. Biging: "Relevance and redundancy in fuzzy classification systems". *Mathware and Soft Computing* 8:203–216, 2001.

[6] A. Del Amo, J. Montero, G. Biging and V. Cutello: "Fuzzy classification systems". *European Journal of Operational Research*, to appear.

[7] E. Montseny, P. Sobrevilla. "Application of fuzzy techniques to the design of algorithms in computer vision", *Mathware & Soft Computing*, Vol. 2-3, 1998, pp.223-230.

[8] J. R. Jensen *Introductory Digital Image Processing. A Remote Sensing Perspective.* Prentice Hall

[9] P.P. Perny and B. Roy , "The use of fuzzy outranking relations in preference modelling", *Fuzzy Sets and Systems* Vol. 49, pp. 33–53, 1992.

[10] Pratt W., *Digital image processing*, John Wiley and Sons, 1978.

[11] A.R. Robertson, "Color perception", *Physics Today*, 1992, pp. 24-29.

[12] S. Romani, E. Montseny, P. Sobrevilla. "Obtaining the Relevant Colors of an image through Stability-based Fuzzy Color Histograms", Proc. of the 12th IEEE International Conference on Fuzzy Systems, Sant Louis (MO), 2003, pp 914-919.

[13] A.R. Smith, "Color gamut transform pairs", *IEEE Trans on Computer Graphics*, Vol. 2, 1978, pp. 12-19.

[14] J. Siskos, J. Lochard and J. Lombard , "A multicriteria decision making methodology under fuzziness: application to the evaluation of radiological protection in nuclear power plants, *Fuzzy Sets and Decision Analysis* in H.J.Zimmermann, L.A. Zadeh and B.R. Gaines (eds.) North Holland, Amsterdam, 1984.

[15] D. Yagi, K. Abe, H. Nakatani, "Segmentation of color aerial photographs using HSV color models", IAPR Workshop on Machine Vision Applications. (Tokyo), 1992.