

# Classification of Blood Types by Microscope Color Images

S. M. Nazia Fathima

**Abstract**—Blood typing is a method to tell what specific type of blood a person has. It is a mandatory that everyone should know their blood type. It is extremely useful in blood transfusions, donation, accidents and other emergencies. The blood type testing is typically made in laboratories by technicians. Such a procedure presents undesirable drawbacks: slowness and it presents non standardized accuracy since it depends on the operator's capabilities and tiredness. This paper presents a methodology to achieve a semi automated system for classification of blood types by microscope color images. This paper concerns with the ABO and Rh blood typing systems. The classification of blood types in microscopy images allows identifying the blood groups and Rh factor accurately. The proposed system first performs image pre-processing by histogram equalization and color correction and then a color space conversion from RGB to HSI is done. Then it extracts the color and texture features of the images using cumulative histogram and Haralick method respectively. Finally it classifies the blood type by support vector machine (SVM).

**Index Terms**—Blood group, color correction, color space conversion, cumulative histogram, Haralick, SVM.

## I. INTRODUCTION

Blood typing is a method to tell what specific type of blood a person has. The differences in human blood are due to the presence or absence of certain protein molecules called antigens and antibodies. The antigens are located on the surface of the red blood cells and the antibodies are in the blood plasma. Individuals have different types and combinations of these molecules [1].

According to ABO and Rh blood grouping systems, a person can belong to either of following 8 blood groups:

A Rh+, A Rh-, B Rh+, B Rh-, AB Rh+, AB Rh-, O Rh+ and O Rh-.

The blood grouping is done in laboratories by slide test which is a manual method. Most of the techniques applied are still based on the principle of interaction between antigen and antibody and subsequent agglutination of RBCs (positive result). The absence of agglutination indicates the lack of interaction (negative result) [2].

This manual blood grouping presents serious drawbacks such as incorrect blood grouping and wrong typing in the report.

Hence, it is necessary to develop an automated system for blood group identification based on the microscope color images of blood.

In this paper, SVM method was used on microscope images of blood (after adding antigens) in 8 categories.

## II. IMAGE CLASSIFICATION

Acquisition of blood images is usually done using a digital camera mounted in the optical path of the microscope. Very often, very high resolution cameras are employed to gain as much direct information as possible. We have considered 10 color images for each category of blood.

### A. Image Pre-processing

The role of pre-processing is to segment the interesting pattern from the background. The pre-processing also defines a compact representation of the pattern. Generally, noise filtering, smoothing and normalization are done in this step and thereby improve the quality of the images. Images are pre-processed prior to image segmentation and feature extraction. Techniques used here are Histogram Equalization and Color Correction.

Histogram Equalization is a method in image processing of contrast adjustment using the image's histogram. Histogram provides a global description of the appearance of the image [3]. The original image (Fig. 1) and histogram equalization results (Fig. 2) are shown.



Fig. 1. Original image.

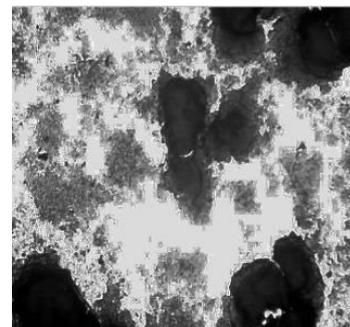


Fig. 2. Image after histogram equalization.

In microscope images, colors are rarely balanced correctly. Color correction is performed by equalizing each color

Manuscript received February 4, 2013; revised July 18, 2013.

S. M. Nazia Fathima is with the Department of Computer Science and Engineering, Sethu Institute of Technology, Virudhunagar, Tamilnadu, India (e-mail: smaizan@gmail.com).

means. This algorithm is a linear translation of the histogram. We add to each pixel the difference between desired mean value and the mean of the channel. We do that for each RGB channel. Values can't overrun the interval of color. Our method is a compromise that produces at the output a more pleasant image without any knowledge on the environment.

### B. Region of Interest Selection

A rectangular region in the image is selected which is the region of interest. The selected region is shown in the Fig. 3.

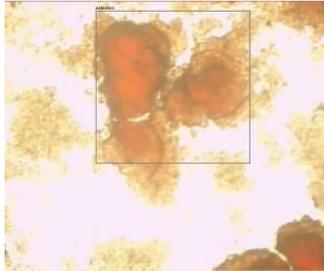


Fig. 3. Selecting the region of interest.

### C. Color Space Conversion

The microscope images are in RGB color space. Hence there is a need for conversion from RGB to HSI color space model. The HSI color space is very important and attractive color model for image processing applications because it represents colors similarly how the human eye senses colors [4]. The HSI color model represents every color with three components: hue (H), saturation (S), intensity (I). The Hue component describes the color itself in the form of an angle between [0, 360] degrees. 0 degree mean red, 120 means green 240 means blue. 60 degrees is yellow, 300 degrees is magenta. The Saturation component signals how much the color is polluted with white color. The range of the S component is [0, 1]. The Intensity range is between [0, 1] and 0 means black, 1 means white.

Conversion from RGB to HSI,

$$I = (R + G+B)/3 \quad (1)$$

$$S = 1 - (3/(R+G+B)) \times a \quad (2)$$

where  $a$  is the minimum of R, G and B

$$H = \cos^{-1} \left\{ \frac{0.5 * [(R-G) + (R-B)]}{\sqrt{[(R-G)^2 + (R-B)(G-B)]}} \right\} \quad (3)$$

If  $S = 0$ , H is meaningless.

If  $(B/I) > (G/I)$  then  $H = 360 - H$  since H is an angle in degrees we then normalize to 0, 1 with  $H=H/360$ .

Fig. 4 shows the image sample after color space conversion.

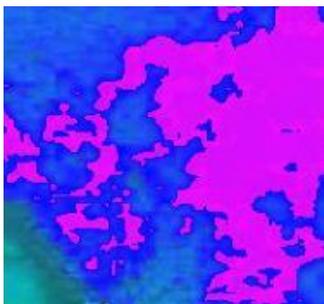


Fig. 4. Image sample after color space conversion.

### D. Spot Marking

The spot marking is done which involves marking of the spots of normal surface or clumps in the image. The spot marking is done manually. Fig. 5 shows the image with spots marked (green spots).

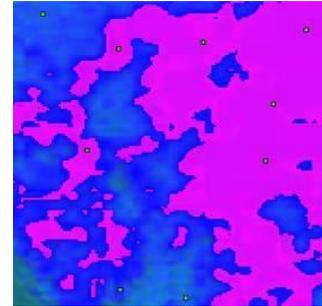


Fig. 5. Spot marking.

### E. Color Feature Extraction

After pre processing and determination of ROI to the images, you can extract characteristics needed from the image. These extracted features are organized in a database, as the input of classification system. The features extracted here are color and texture.

It is essential to quantify HSI space component to reduce computation and improve efficiency. At the same time, because the human eye to distinguish colors is limited, do not need to calculate all segments. Unequal interval quantization according the human color perception has been applied on H, S, I components as follows:

Based on the color model of substantial analysis, we divide color into eight parts. Saturation and intensity is divided into three parts separately in accordance with the human eyes to distinguish.

In accordance with the different colors and subjective color perception quantification, quantified hue (H), saturation(S) and Intensity (I) are showed as equation (6, 7, 8).In accordance with the quantization level above, the H,S,I three dimensional feature vector for different values of with different weight to form one dimensional feature vector named

$$G:G=Q_s Q_v H+Q_v S+I. \quad (4)$$

where  $Q_s$  is quantified series of S,  $Q_v$  is quantified series of I. Here we set  $Q_s = Q_v = 3$ , then

$$G=9H+3S+I \quad (5)$$

$$H = \begin{cases} 0 & \text{if } h \in [316, 20], \\ 1 & \text{if } h \in [21, 40], \\ 2 & \text{if } h \in [41, 75], \\ 3 & \text{if } h \in [76, 155], \\ 4 & \text{if } h \in [156, 190], \\ 5 & \text{if } h \in [191, 270], \\ 6 & \text{if } h \in [271, 295], \\ 7 & \text{if } h \in [296, 315] \end{cases} \quad (6)$$

$$S = \begin{cases} 0 & \text{if } s \in [0, 0.2], \\ 1 & \text{if } s \in [0.2, 0.7], \\ 2 & \text{if } s \in [0.7, 1] \end{cases} \quad (7)$$

$$I = \begin{cases} 0 & \text{if } v \in [0, 0.2], \\ 1 & \text{if } v \in [0.2, 0.7], \\ 2 & \text{if } v \in [0.7, 1] \end{cases} \quad (8)$$

In this way, three-component vector of HSI form one-dimensional vector, which quantize the whole color space for the 72 kinds of main colors. So we can handle 72 bins of one-dimensional histogram. In this manner, we represent the one-dimensional vector  $G$  by constructing a cumulative histogram of the color characteristics of image after using non-interval HSI quantization for  $G$  [5].

#### F. Texture Feature Extraction

The texture features extracted are: angular second-order distance, contrast, entropy, anti-difference distance, relevant, mean of sum, mean of difference, entropy of sum, entropy of difference, variance, variance of sum, variance of difference. Entropy, entropy of sum and entropy of difference are used to measure the changes of the coefficients of co-occurrence matrix and complex levels of texture changes in the image. Variance, variance of sum and variance of difference reflect the mutative characteristics of gray-scale images [6].

#### G. SVM Classifier

The support vector machine (SVM) is a popular classification technique. A classification task usually involves separating data into training and testing sets. Each instance in the training set contains one target value and several attributes. The goal of SVM is to produce a model (based on the training data) which predicts the target values of the test data given only the test data attributes. In the parlance of SVM literature, a predictor variable is called an attribute, and a transformed attribute that is used to define the hyper plane is called a feature. The task of choosing the most suitable representation is known as feature selection. A set of features that describes one case is called a vector. So the goal of SVM modelling is to find the optimal hyper plane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyper plane are the support vectors [7].

For multiclass SVM, it is possible to combine multiple binary classifiers. The one-against-one approach is utilized in this paper for multiclass SVM because it provides fast training. In the one-against-one method,  $K(K-1)/2$  binary classifiers are trained and  $K(K-1)/2$  binary tests are required to make a final decision, where  $K$  is the total number of classes. Each outcome gives one vote to the winning class, and the class with the most votes is selected as the final result.

Key of Support Vector Machine is kernel function. A kernel function is a function that corresponds to an inner product in some expanded feature space. Kernel functions used in SVM must satisfy Mercer's condition which requires the kernel to be a continuous symmetric kernel of a positive integral operator. Low-dimensional space vector set is usually difficult to be divided, and the solution is to map them to a high-dimensional space. However, the difficulties of this approach are the increasing of computational complexity,

while the kernel function is just the solution to this problem. In other words, as long as the choices of appropriate kernel function, we can get the classification function of high-dimensional space. Support vector machine is constituted to the training sample set and the kernel function. Thus using different kernel function  $K(x_i, x_j)$  can construct different types of learning machines with non-linear decision-making side on the input space, and lead to different support vector algorithms. Now commonly used kernel functions are mainly polynomial kernel function, radial basis function, and Sigmoid function, etc.

The choice of kernel function, RBF function is often the first to be considered. RBF function via non-linear transformation maps the samples to a higher-dimensional feature space; therefore, it provides the possibility to solve the nonlinear problem of the samples. RBF kernel function is only one parameter, the polynomial kernel function and Sigmoid parameters have more than the RBF kernel function [8].

### III. BLOOD SAMPLES

The following are the blood samples of various blood groups taken under the microscope after adding the antigens.

#### A. O Positive Group

Fig. 6 (a, b, c) shows the O Positive blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

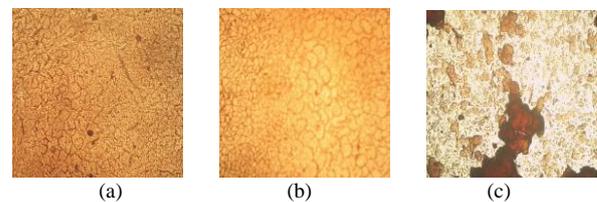


Fig. 6. O Positive blood group sample.

#### B. O Negative Group

Fig. 7 (a, b, c) shows the O Negative blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

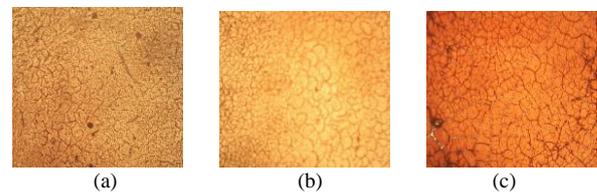


Fig. 7. O Negative blood group sample.

#### C. B Positive Group

Fig. 8 (a, b, c) shows the B Positive blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

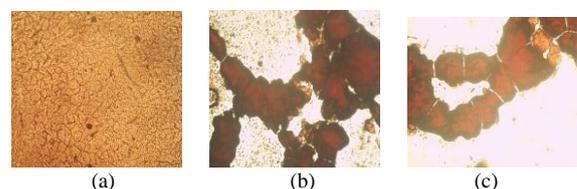


Fig. 8. B Positive blood group sample.

#### D. B Negative Group

Fig. 9 (a, b, c) shows the B Negative blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

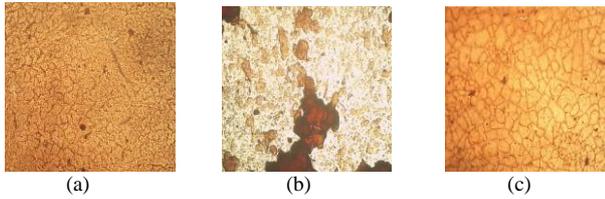


Fig. 9. B Negative blood group sample.

#### E. A Positive Group

Fig. 10 (a, b, c) shows the A Positive blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

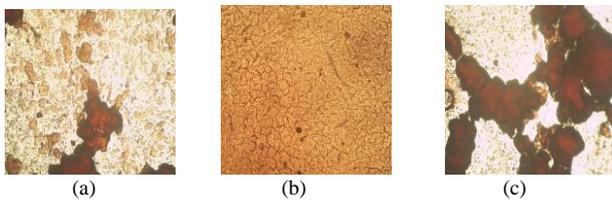


Fig. 10. A Positive blood group sample.

#### F. A Negative Group

Fig. 11 (a, b, c) shows the A Negative blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

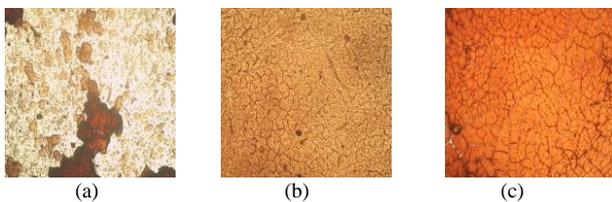


Fig. 11. A Negative blood group sample.

#### G. AB Positive Group

Fig.12 (a, b, c) shows the AB Positive blood group sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

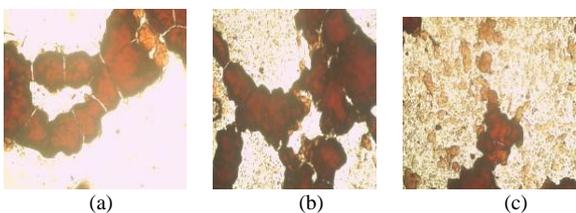


Fig.12. AB Positive blood group sample.

#### H. AB Negative Group

Fig. 13 (a, b, c) shows the AB Negative blood group

sample on adding Antigen-A, Antigen-B and Antigen-D respectively.

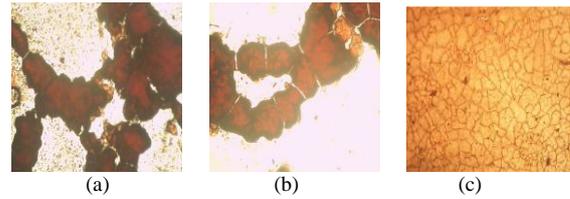


Fig. 13. AB Negative blood group sample.

#### IV. CONCLUSION

This paper discusses the SVM classifier for automatic classification of eight blood types. Because SVM is based on structural risk minimization principle, it has high recognition rate and robustness [9].

Such automated classification of blood types overcomes the drawbacks of the manual blood typing.

#### ACKNOWLEDGMENT

The author would like to thank Jaya Kumar, Research Associate, Sethu Institute of Technology, Virudhunagar, for initiation of this project and then to Meenakashi Mission Hospital, Madurai, Tamilnadu, India for giving necessary permission to use their laboratory and to all other who helped to undertake this project.

#### REFERENCES

- [1] G. Daniels, *Human Blood Groups*, 2nd ed. Blackwell Science, 2002.
- [2] B. H. Estridge, A. P. Reynolds, and N. J. Walters, *Basic Medical Laboratory Techniques*, 4th ed. Thomson Learning, 2000.
- [3] The Igeland website. [Online]. Available: [http://www.igeland.net/bildebehandling/local\\_histogram\\_equalization.html](http://www.igeland.net/bildebehandling/local_histogram_equalization.html)
- [4] M. H. Asmare, V. S. Asirvadam, and L. Iznita, "Color space selection for color image enhancement applications," *ICSAP*, pp. 208-212, 2009.
- [5] B. Jyothi, Y. M. Latha, and V. S. K. Reddy, "Medical image retrieval using multiple features," *Research India Publications*, vol. 3, pp. 387-396, 2010.
- [6] A. Porebski, N. Vandenbroucke, and L. Macaire, "Haralick feature extraction from LBP images for color texture classification," *IPTA*, pp. 1-8, 2008.
- [7] The support-vector-machine website. (2005). [Online]. Available: <http://www.support-vector-machines.org/>
- [8] H. Yan, X. G. Zhang, and Y. D. Li, "Kernel-based maximal-margin Clustering algorithm," *Journal of Tsinghai University (Science and Technology)*, vol. 42, no. 1, pp. 132-134, 2002.
- [9] Z. Q. Song, H. Cui, and Y. A. Hu, "Research and development of support vector machine theory," *Journal of Naval Aeronautical Engineering Institute*, vol. 23, no. 2, pp.143-148, 2008.



**S. M. Nazia Fathima** was born in Dubai, UAE in 1987. She done masters degree in Computer Science and Engineering. Her research interest is image processing.