Comparison of Grayscale Conversion Methods for Malaria Classification

JunYeon¹, Jong-Dae Kim^{2,3}, Chan-Young Park^{2,3}, Yu-Seop Kim^{2,3} and Hye-Jeong Song^{2,3*}

¹Department of Computer Engineering, Hallym University, Korea ²Department of Ubiquitous Computing, Hallym University, Korea ³Bio-IT Research Center, Hallym University, Korea { mauver, kimjd,cypark, yskim01,hjsong }@hallym.ac.kr

Abstract

Malariaparasitemia is used for measuring the degree of infection by detecting the plasmodium (commonly known as the malaria parasite) in the blood of an infected patient. The most commonly used method of malarial diagnosis is counting the number of malariainfected red blood cells in a Giemsa-stained blood smear by using a microscope. This method requires expert knowledge and is prone to inter-investigator variability. Therefore, a number of studies have been conducted on automated classification techniques that can measure malarial infection rapidly and accurately. In order to detect malaria parasites by analyzing plasmodium-infected blood smear images, conversion processing is required to make the images insensitive to the luminance contrast of microscopy and staining intensity. This paper aims to identify a grayscale conversion method optimal for plasmodium-infected blood smear images by comparing the performances of various grayscale conversion methods. The grayscale conversion methods selected for the comparison are colorimetric conversion, luma coding, conversion using the green channel only, and principal component analysis (PCA)based conversion. We used 20 malaria-infected red blood cells and 20 normal red blood cells to compare the performances of these methods by obtaining thearea under the receiver operating curve (AUC) as the minimum histogram intra-class variance value for each cell image. With the AUC value of 0.9225, the PCA-based grayscale conversion method outperformed all other methods.

Keywords: Malaria, Image processing, Grayscale conversion

1. Introduction

Malaria is a serious global disease with an annual death toll of 150,000–270,000 [1]. Its pathogen is the parasite belonging to the genus Plasmodium, which accounts for 98% of its mortality rate [2]. Rapid and accurate diagnosis for a quick treatment is a crucial factor for malaria control [1].

Malaria parasitemia is used for estimating the infection by detecting the plasmodium in the blood of a malaria-infected patient. The most commonly used method of malarial diagnosis is counting the number of malaria-infected red blood cells in a Giemsa-stained blood smear by using a microscope. This method requires additional training of lab personnel to obtain diagnosis-related knowledge, and because of diverse sampling quantities and methods, the resulting diagnosis significantly depends on the performance of the microscope used and the knowledge of the individual experts. Therefore, a considerable amount of research has been conducted to investigate automated classification techniques that can measure malarial infection rapidly and accurately [5-13].One of the studies based on microscopic images was conducted in the HSV (Hue-Saturation-Value) color space, in which the hue range is identified or optimal saturation thresholds are calculated, in order to detect malaria-infected red blood cells. This method was evaluated using the Leishman-stained blood smear images [14].

In this paper, we compare the performances of four grayscale conversion methods for distinguishing malaria-infected red blood cells from normal ones in Giemsa-stained blood smear slide images.

The vital feature in distinguishing between malaria-infected and normal red blood cells is the stained cell color because the classification depends on staining intensity. Therefore, the classification efficiency can improve if the main color of each cell can be defined. However, the main color is difficult to obtain because of the large variance of microscopic images of malaria-infected cells depending on luminance and staining intensity. To address this variation-related problems, this paper compares the principal component analysis (PCA) method for obtaining the principal axis of the cell color [17] and the colorimetric (luminancepreserving) conversion, luma coding applied in another study [14], and gravscale conversion method using only the green channel [8].For the comparison, images of 20 malaria-infected red blood cells and 20 normal red blood cells are used. We compare their performances by obtaining the area under the receiver operating curve (AUC) as the minimum value of the histogram intra-class variance for each cell image. The Otsu thresholding method [16], the most popular binarization method, performs binarization by selecting a threshold value that can minimize the intra-class variance. This method is used for performance assessment because the grayscale image converted using a method to minimize intra-class variance can be more advantageous in binarization. The experimental results revealed that the PCA-based grayscale conversion method demonstrated the best AUC value (0.9225), thus proving to be the optimal method.

This paper introduces the materials and methods in section 2, presents the results in section 3, and the conclusion is made in section 4.

2. Materials and Methods

2.1. Image Acquisition

The malaria cell images used in the tests were provided by the Korea University Hospital. They were obtained directly from the Giemsa-stained thin blood smears collected in 2010 and 2011. The malaria infections detected in the microscopic images were confirmed by a malaria diagnosis expert. Malaria-infected and normal red blood cells (n = 20 each) were selected and stored. Figures 1 and 2 show the malaria-infected red blood cells and normal red blood cells respectively, as used in the experiment. The malaria-infected red blood cells can be discerned for the darker stained spots in the parts where the malaria parasites are present.



Figure1. Malaria-infected Red Blood Cells



Figure 2. Normal Red Blood Cells

2.2. Method Overview

The malaria-infected cell image analysis program used in this study was developed using MATLAB.

Figure 3 shows the overall processing flow chart for the performance verification. First, the images of the malaria-infected red blood cells and normal red blood cells are loaded along with the mask for the extraction of the cell parts alone. Then, the grayscale conversion is processed using the four methods to be compared, followed by ROC graph drawing with the minimum values of the histogram intra-class variance from the processed grayscale images and the performance assessments by obtaining their AUCs. A loaded mask is used for drawing the histograms from the grayscale images.



Figure 3. Method Overview

2.3. Conversion to Grayscale

Grayscale is a digital image that carries only the information on luminous intensity. It consists of multiple-shade grayscales ranging from low-luminosity black to high-luminosity ash color. There are different methods of converting color images into grayscale images. Luminosity is calculated by endowing the RGB channels of a color image with their respective weights.

Colorimetric grayscale conversion can be defined by Eq. (1).

 $Y = 0.2126R + 0.7152G + 0.0722B \tag{1}$

Colorimetric grayscale conversion is widely used for luminosity measurement, in which the original color image luminance is matched with the grayscale image luminance. This ensures that the color and grayscale images have the same absolute luminance. The RGB color model is converted into grayscale by using the coefficient calculated from the intensity values of the three primary colors as perceived by the human eye. In particular, human visual sense is the most sensitive to green and the least sensitive to blue. To encode gray intensity in linear RGB, the Y luminance value derived should be equivalently configured in each of the channels for red, green, and blue.

Eq. (2) represents the luma coding grayscale conversion method.

$$Y' = 0.2126R' + 0.7152G' + 0.0722B'$$
⁽²⁾

The color coordinates used here are similar to those in the color space Y'UV, which is used for video systems, such as PAL, SECAM, and NTSC, and standard color TV. A nonlinear luma component (Y') is calculated directly from the values of gamma compression via a weighted sum, which can be calculated quickly without undergoing compression or gamma expansion.

In the Y'UV and Y'IQ models used in PAL and NTSC, the REC601 luma(Y') component uses decimals to differentiate such gamma-compression values in the linear R, G, B, and Y.

Eq. (3) represents the malaria grayscale conversion method using exclusively the green channel adopted in other study.

$$Y = 0.0R + 1.0G + 0.0B$$

The parts of a malaria-infected red blood cell harboring the malaria parasites are stained violet (Figure 1) and thus, they are the most influenced by the green channel.

PCA-based grayscale conversion is a method of grayscale conversion in which the principal axis of the red blood cell color is calculated and maximum contrast is created using linear least-square fitting. This method, which was presented by Russ, uses the RGB color coordinate system. Each pixel on an RGB color image can be expressed in this color space. The best fit regression line, which minimizes the point-to-line distances in this space, can then be created via regression.

Figure 4 shows individual pixels and the best fit regression line on a malaria-infected cell image plotted in the RGB color space.

(3)



Figure 4. Color Coordinates and Best Fit Regression Line of Malaria-infected Cell Image Pixels

The best fit regression line is used as weights with which the cosines of the angles between the R, G, and B axes are obtained and converted into grayscale values.

Eq. (4) is the equation for obtaining the weights; the values of a, b, and c are applied as the weights for red, green, and blue, respectively, and converted into grayscale values.

$$\min(a,b,c) = \frac{\sum_{i=0}^{size} (ar_i + bg_i + cb_i)}{\sqrt{a^2 + b^2 + c^2}}$$
(4)

Figure 5 shows the results of comparing the images, which were generated by applying the four grayscale methods to a malaria-infected cell image, and the corresponding histograms.



Figure 5. Gray Conversion Result and Histogram

International Journal of Bio-Science and Bio-Technology Vol.7, No.1 (2015)

2.4. Performance Comparison

In order to compare the performance of the individual conversion methods, we applied the threshold values to the grayscale images converted using each of the four conversion methods, thereby using the value minimizing the intra-class variance value. The smaller this value becomes, the closer the measure approaches the maximum value, which is advantageous for binarizing the grayscale image. We evaluated the performances of the four conversion methods by obtaining the AUC values for 20 malaria-infected cell images and 20 normal cell images.

3. Results

Figure 6 shows a graphic representation of the intra-class variance values on the malaria-infected cell images.



Figure 6. Intra-class Variance in Malaria-infected Cell Image

Considering the minimum intra-class variance value of each conversion method on the graph, the method using only the green channel shows the highest value, followed by luma coding and colorimetric methods. The PCA-based method was revealed to have the lowest minimum intra-class variance value, thus proving to be the most advantageous method for binarization via thresholding.



Figure 7.The ROC Graph of the Intra-class Variance of Gray Conversion Methods

Of the ROC curves, the plus curve is the graph for the colorimetric method, long dashed curve for luma coding, and short dashed curve for the green-channel-only method. The ROC curve for the PCA-based method, represented by the solid line, demonstrates that the PCA-based method outperformed other methods in all sections except for the 90%–80% and 40%– 30% specification sections, with the best performance shown in the 65%–40% section.

Method	AUC 0.9000 0.9050		
Colorimetric			
Luma coding			
Green channel	0.8850		
PCA	0.9225		

Table 1.	The Performance	Comparison	of the 4 Gra	ayscale C	conversion M	/lethods
----------	------------------------	------------	--------------	-----------	--------------	-----------------

Table 1 lists the AUC values of the four grayscale conversion methods. With the AUC value of 0.9225, the PCA-based method outperformed the other three methods. In contrast, the green-channel-only grayscale conversion method used in other malaria studies was outperformed by all the other methods.

4. Conclusion

We performed a comparative experiment on four conversion methods in order to determine the optimal grayscale conversion method for microscopic images of malaria-infected cells.

In order to identify the best-performing method for binarizing grayscale images by using the Otsu thresholding technique, we obtained the AUCs as the minimum histogram intra-class variance values for individual cell images. The PCA-based method showed the AUC value of 0.9225, thus demonstrating the best performance.

The PCA method converts the original images into grayscale images by obtaining the principal axis of the red blood cell color, thus overcoming the variance depending on the luminance and staining intensity of the microscopic images of malaria-infected cells by efficiently determining the main color.

If a parasite classification or an automated malaria detection algorithm is developed based on this method, a high-performing automated malaria diagnosis system could be implemented.

Acknowledgments

This work (Grants No. C0137719) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Ko.

References

- M. T. Makler, C. J. Palmer and A. L. Alge "A review of practical techniques for the diagnosis of malaria", Ann Trop Med Parasitol, vol. 92, no. 4, (2008), pp. 419-433.
- [2] WHO (World Health Organization) "World malaria report 2008 world health organization", Geneva, Tech Reports, (2008).
- [3] R. E. Coleman, N. Maneechai, N. Rachaphaew, C. Kumpitak, R. Miller, V. Soyseng, K. Thimasarn, and J. Sattabongkot, "Comparison of field and expert laboratory microscopy for active surveillance for asymptomatic Plasmodium falciparum and Plasmodium vivax in Western Thailand", Am J Trop Med Hyg, vol. 67, (2002), pp. 141-144.
- [4] I. Bates, V. Bekoe, and A. Asamoa-Adu, "Improving the accuracy of malaria-related laboratory tests in Ghana", Malaria Journal, vol. 3, (2004), pp. 38.
- [5] S. Jigyasha, "Advanced Image Analysis based system for Automatic Detection of Malarial Parasite in Blood Images Using SUSAN Approach", International Journal of Engineering Science and Technology (IJEST), vol. 3, no. 6, (2011), pp. 5260.
- [6] J. Somasekar, B. E. Reddy, E. K. Reddy, and C. H. Lai, "An Image Processing Approach for Accurate Determination of Parasitemia in Peripheral Blood Smear Images", IJCA Special Issue on Novel Aspects of Digital Imaging Applications DIA, vol. 1, (2011), pp. 22-28.
- [7] Y. Purwar, S. L. Shah, G. Clarke, A. Almugairi, and A. Muehlenbachs, "Automated and unsupervised detection of malarial parasites in microscopic images", Malaria Journal, vol. 10, (2011), pp. 364.
- [8] A. Mehrjou, T. Abbasian, and M. Izadi, "Automatic Malaria Diagnosis system", Robotics and Mechatronics (ICRoM) 2013 First RSI/ISM International Conference on, (2013), pp. 205-211.
- [9] M. I. Khan, B. Acharya, B. K. Singh, and J. Soni, "Content based image retrieval approaches for detection of malarial parasite in blood images", International Journal of Biometrics and Bioinformatics, vol. 5, no. 2, (2011), pp. 97–110.
- [10] P. T. Suradkar, "Detection of Malarial Parasite in Blood Using Image Processing", International Journal of Engineering and Innovative Techhnology(IJEIT), vol. 2, no. 10, (2013), pp. 124-126.
- [11] D. A. Kurer and V. P. Gejji, "Detection of Malarial Parasites in Blood Images", International Journal of Engineering Science and Innovative Technology (IJESIT), vol. 3, no. 3, (2014), pp. 651-656.
- [12] V. V. Makapati and R. M. Rao, "Segmentation of malaria parasites in peripheral blood smear images", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, (2009), pp. 1361-1364.
- [13] F. Tek, A. Dempster and I. Kale, "Parasite detection and identification for automated thin blood film malaria diagnosis", Computer Vision and Image Understanding, vol. 114, (2010), pp. 21–32.

- [14] A. Salihah, A. N. Yusoff, Mashor, Zeehaida and Mohamed, "Colour Image Segmentation Approach for Detection of Malaria Parasites Using Various Colour Models and k-Means Clustering", WSEAS Transactions on Biology and Biomedicine, vol. 10, no. 1, (2013), pp. 41-55.
- [15] Grayscale methods :http://en.wikipedia.org/wiki/Grayscale
- [16] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Transactions on Systems, Man and Cybermetics In N/A, vol. 9, no. 1, (1979), pp. 66-66.
- [17] C. R. John and J. C. Russ, "Introduction to image processing and analysis", CRC Pr I Llc, ISBN 978-0-8493-7073-1, (2008).

Authors



Jun-Yeon, He is enrolled in a master degree in Computer Engineering major in Hallym University. He currently studies for image processing and pattern recognition. His recent interests focus on embedded system and image pattern recognition.



Jong-Dae Kim, He received the M.S. and the Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1984 and 1990, respectively. He worked for Samsung Electronics from 1988 to 2000 as an electrical engineer. He is a Professor in Department of Ubiquitous Computing, Hallym University. His recent interests focus on biomedical system and bioinformatics.



Chan-Young Park, He received the B.S. and the M.S. from Seoul National University and the Ph.D. degree from Korea Advanced Institute of Science and Technology in 1995. From 1991 to 1999, he worked at Samsung Electronics. He is currently a Professor in the Department of Ubiquitous Computing of Hallym University, Korea. His research interests are in Bio-IT convergence, Intelligent Transportation System and sensor networks.



Yu-Seop Kim, He received the Ph.D. degree in Computer Engineering from Seoul National University. He is currently a Professor in the Department of Ubiquitous Computing at Hallym University, South Korea. His research interests are in the areas of bioinformatics, computational intelligence and natural language processing.



Hye-Jeong Song, She received the Ph.D. degree in Computer Engineering from Hallym University. He is a Professor in Department of Ubiquitous Computing, Hallym University. His recent interests focus on biomedical system and bioinformatics International Journal of Bio-Science and Bio-Technology Vol.7, No.1 (2015)