LEARNING TOOLS FOR BLOOD CELL SEGMENTATION AND EXTRACTION TECHNIQUES

POON CHEE LIM

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Electrical - Computer and Microelectronic System)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JANUARY 2013

LEARNING TOOLS FOR BLOOD CELL SEGMENTATION AND EXTRACTION TECHNIQUES

POON CHEE LIM

UNIVERSITI TEKNOLOGI MALAYSIA

To my parents and God for always being there.

ACKNOWLEDGEMENT

During the writing of this thesis I was assisted by many individuals with whom I would like to share my sincerest gratitude.

First and foremost I would like to express my appreciation to Dr. Nasrul Humaimi Mahmood, my project supervisor for all his time, effort and guidance in ensuring that I am able to complete this thesis. I would also like to thank him for all his advice and also for the motivation he has given me throughout this time.

Next I would like to give my heartfelt thanks to my parents who have supported me the whole time.

I would also like to express my thanks to all my friends who have been supportive during this time and even providing some light hearted entertainment throughout the writing of this thesis. They also helped give me ideas and encouragement.

Finally, I would also like to thank Faculty of Electrical Engineering, Universiti Teknologi Malaysia and School of Graduate Studies (SPS UTM) for their support in this project.

ABSTRACT

Blood cell segmentation and identification is vital in the study of blood as a health indicator. A complete blood count is used to determine the state of a person's health based on the contents of the blood in particular the white blood cells and the red blood cells. The main problem arises when massive amounts of blood samples are required to be processed by the haematologist or Medical Laboratory Technicians. The time and skill required for the task limits the speed and accuracy with which the blood sample can be processed. This project aims to provide user-friendly software based on MATLAB allowing for quick user interaction with a simple tool for the segmentation and identification of red and white blood cells from a provided image. The project presents the solution in a modular framework allowing for future development within a structured environment. In order to perform the segmentation, this project uses k-means clustering and colour based segmentation using International Commission on Illumination L*a*b* (CIELAB) colour space coupled with polygon information of the region of interest. The project integrates these methods into a flow within a Graphical User Interface (GUI) with customizable variables to handle changing input images. The result of the project is a working GUI with the capability to accept user interaction. The completed project is able to obtain quick and accurate blood cell segmentation of both red and white blood cells. The accuracy of this project ranges from 64% to 87% depending on the type of processing used and the type of cells being extracted.

ABSTRAK

Segmentasi dan pengenalan sel darah adalah penting dalam kajian darah sebagai petunjuk kesihatan. Pengiraan darah lengkap digunakan untuk menentukan tahap kesihatan seseorang berdasarkan kandungan sel darah putih merah. Masalah timbul apabila kuantiti sampel darah yang perlu diproses oleh haematologist atau Juruteknik Makmal Perubatan adalah besar. Masa dan kemahiran yang diperlukan menghadkan kelajuan dan ketepatan pemprosesan sampel darah. Projek ini menyediakan perisian yang membolehkan interaksi pengguna yang cepat dan mudah digunakan untuk segmentasi dan pengenalpastian sel darah merah dan putih dari imej yang disediakan. Projek ini membentangkan penyelesaian dalam bentuk rangka modular yang membenarkan pembangunan masa depan dalam persekitaran teratur. Dalam usaha untuk melaksanakan segmentasi, projek ini menggunakan pengelompokan *k-means* dan segmentasi berasaskan warna dalam ruang warna "*International Commission on Illumination L*a*b**" bergandingan dengan penggunaan maklumat poligon. Projek ini mengintegrasikan kaedah ini dalam antara muka grafik pengguna (*GUI*) beserta dengan pembolehubah yang boleh diubah untuk memproses input yang berlainan. Hasil projek adalah *GUI* yang mempunyai keupayaan untuk berinteraksi dengan pengguna. Secara keseluruhan, projek ini berupaya mengendalikan segmentasi sel darah dengan cepat dan tepat untuk sel darah merah dan putih. Ketepatan projek ini adalah diantara 64% sehingga 87% bergantung kepada proses yang digunakan dan jenis sel yang hendak diekstrak.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

WBC	-	White Blood Cells
RBC	-	Red Blood Cells
MLT	-	Medical Laboratory Technicians
GUI	-	Graphical User Interface
ROI	-	Regions of Interest
GVF	-	Gradient Vector Flow
CIELAB	-	International Commission on Illumination L*a*b* colour space

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

Blood is a vital component in human bodies. It functions to transport oxygen from the lungs to the rest of the tissues in the body and carbon dioxide from those tissues back to the lungs to be expelled. Blood also carries the waste products and toxins from all over the body to the liver to be excreted. Body temperature regulation is also one the main functions of the constant flow of blood throughout the body. Finally, there are components in the blood that act to close up wounds and attack foreign bodies that can cause disease.

Due to the important functions of blood, the monitoring of blood components is of use to physicians in the determination of a person's state of health as well as the identification of any disease a person might have. The procedure of determining the contents of a person's blood is called a complete differential blood count or CBC.

Blood samples normally contain white blood cells (WBC), red blood cells (RBC), plasma and platelets. The composition of RBCs in the blood is an indication of the function of the blood to transport agents such as oxygen, nutrients and waste. WBCs on the other hand are an indication of the body's ability to combat disease. It

is also an indication of various types of diseases already infecting a body. Due to this, more emphasis has been given to the counting of WBC. However, the RBC count also has value and should not be neglected. The final components of blood are the platelets and plasma. Platelets are not clinically important so they are usually ignored, while the plasma is basically the background of a blood smear image.

Blood smear analysis is a common procedure in medical facilities. It is usually performed by a haematologist or trained Medical Laboratory Technicians (MLT) to study the contents of the blood in particular the red and white blood cells. The current method of manual peripheral blood smear analysis is tedious and prone to error. Some of the problems faced when performing manual counting is time consumption. Any interruption in the process will necessitate a recount which is both time and energy consuming for the MLT. A second problem is the training required to get an MLT up to speed on the differentiation of the RBC and WBC. The amount of effort involved is also an issue. With hundreds of samples, each with hundreds of cells, the amount of counting in total is extremely demanding on the MLT. Finally, there is the problem of data retrieval. Manual blood counts are not computerized and thus retrieving their data or even data entry into a computer system is an extremely time consuming task.

Currently, there are already programs in existence that perform automated blood cell counts that focus on either WBCs or RBCs. However, these programs are not cheap and their processing methods are not open sourced.

1.2 Objectives

The objective of this project is to provide an open sourced MATLAB solution to the counting of both RBCs and WBCs through an interactive Graphical User Interface (GUI). It aims to be robust as well as automated. Blood features will be clearly marked to assist medical students and MLTs as well as any interested parties in the detection of RBCs and WBCs.

In addition to creating a tool to aid blood cell segmentation, this project also aims to create a framework for a modular GUI which can be used as the basis for future work and improvement in the field of blood analysis.

1.3 Scopes of Work

This project is limited to 108 images from the "Acute Lymphoblastic Leukaemia Image Database for Image Processing" database set up and maintained by Fabio Scotti, Department of Information Technology from the Università degli Studi di Milano and are obtained from the image repository provided by the M. Tettamanti Research Center for Childhood Leukemia's and Hematological Diseases, Monza, Italy. These images use Wright's Stain which colours the WBC purplish to ease identification. The focus of this project is the segmentation and counting of platelets, RBCs and WBCs through image processing methods. The focus of this work is more towards image processing because neural network based recognition systems will still rely on basic feature extraction that is done in this project.

1.4 Thesis Overview

This thesis will be made up of five chapters detailing the project from beginning to end. Chapter 1 describes the background, objective and scope of the project as well as an overview of the thesis. Chapter 2 covers the literature review of works related to the project as well as an overview of blood cells and their segmentation methods. Chapter 3 is an in depth description of the methodologies used in this project along with a description of the completed system that is the result of this project. Chapter 4 discusses and analyses the results of the project. Finally chapter 5 will contain the conclusion, challenges and solutions in the current work and future work proposed.

REFERENCES

[1] Ananda, M.Y., Sharif, J.M., Ngadi, M.A., Muhammad, D., Jamil, M.M.A. (2008). A Review Of Human Blood Cells Segmentation, Counting Cell, And Three-Dimensional Visualization For Helping Diseases Analyses. *Proceedings of the 3rd International Conference on Postgraduate Education (ICPE-3)*, 16-17 December. Penang, Malaysia.

[2] Gonzalez & Woods, (2002). *Digital Image Processing*. (2nd ed.). Upper Saddle River, New Jersey: Prentice-Hall.

[3] Hiremath, P.S., Bannigidad, P., Geeta, S. (2010). Automated Identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images. *IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR*, 59.

[4] Pelgur, A. (2011). *Determination of Types and Number of Leukocytes in Blood Samples*. Baskent University.

[5] Poomcokrak, J., Neatpisarnvanit, C. (2008). Red Blood Cells Extraction And Counting. *The 3rd International Symposium on Biomedical Engineering*. 2008. Thailand.

[6] Sadeghian, F., Seman, Z., Ramli, A.R., Kahar, B.H.A., Saripan, M. (2009). A Framework for White Blood Cell Segmentation in Microscopic Blood Images Using Digital Image Processing. *Biological Procedures Online*. 11(1), 196-206.

[7] Sanpanich, A., Iampa, W., Pintavirooj, C., Tosranon, P. (2008). White Blood Cell Segmentation by Distance Mapping Active Contour. 2008 *International Symposium on Communications and Information Technologies*. 21-23 October. Don Chan Palace, Vientiane, Laos, 251-255.

[8] Sinha, N., Ramakrishnan, A.G. (2002). Blood Cell Segmentation Using EM Algorithm. *Proceedings of the Third Indian Conference on Computer Vision, Graphics & Image Processing*. 16-18 December. Ahmadabad, India.

[9] Vemula, V. (2007). *Cell Modelling and Cellular Dynamics.* Master's Thesis, Indian Institute of Science, Bangalore.

[10] Vromen, J., McCane, B. (2006). Red Blood Cell Segmentation Using Guided Contour Tracing. *SIRC 2006 - The 18th Annual Colloquium of the Spatial Information Research Centre*. 6-7 November. University of Otago, Dunedin, New Zealand.

[11] D'Souza, A.C. (2007). *Automated Counting Of Cell Bodies Using Nissl Stained Cross-Sectional Images*. Master's Thesis, Texas A&M University.