

ONLINE HANDWRITING RECOGNITION SYSTEM

VOON CHOI PENG

This report is submitted in partial fulfillment of requirements for the award of
Bachelor of Electronic Engineering (Computer Engineering)

Faculty of Electronic Engineering & Computer Engineering
Universiti Teknikal Malaysia Melaka

April 2009



UNIVERSITI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA

Tajuk Projek : ONLINE HANDWRITING RECOGNITION SYSTEM
Sesi Pengajian : 2008/2009

Saya VOON CHOI PENG mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan () :

SULIT*

(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

TERHAD*

(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD

Disahkan oleh:

(TANDATANGAN PENULIS)

(COP DAN TANDATANGAN PENYELIA)

Alamat Tetap:
A-5-10, JLN METRO PERDANA2,
TMN USAHAWAN KEPONG,
52100 KUALA LUMPUR,
WILAYAH PERSEKUTUAN.

WONG YAN CHIEW
Pensyarah
Fakulti Kejuruteraan Elektronik Dan Kejuruteraan Komputer
Universiti Teknikal Malaysia Melaka (UTeM)
Karung Berkunci No 1752
Pejabat Pos Durian Tunggal
76109 Durian Tunggal, Melaka

“I hereby declare that this report is the result of my work except for quotes as cited in the references.”

Signature :
Author :
Date :

“I hereby declare that I have read this report and in my opinion this report is sufficient in terms of the scope and quality for the award of Bachelor of Electronic Engineering (Computer Engineering) with Honors.”

Signature :

Supervisor's Name :

Date :

Special dedication to my loving parents, Mr. Voon Ah Soo and Mrs. Low Shaw Chan, all my siblings, my kindhearted former supervisor Mrs. Fauziah Binti Salehuddin, supervisor Mrs Wong YanChiew and thank to my dearest friends.

ACKNOWLEDGEMENT

Firstly, I take this opportunity to express my profoundest gratitude and deepest regards to all those who gave me the possibility to successfully complete this Final Year Project. I am deeply indebted to former Project Supervisor Mrs. Fauziyah Binti Salehuddin, Project Supervisor Mrs. Wong Yan Chiew and wishes to express a million thanks for hers exemplary guidance, monitoring and constant encouragement through out the development of the software project. In those moments of uncertainty and doubts when things used to turn dark without a clear understanding of the knowledge that she tried to share, her kind and patient way of explaining had indeed a soothing effect. The blessing, help and guidance given from time to time shall indeed carry me a long way in the journey of life on which I am about to embark in the near future. Last but not least, to all my well-wishers who had helped me both directly and indirectly, I virtually fall to short words to express my gratitude. Therefore, I end this acknowledgement with only two words “Thank You!” in their reminiscence.

ABSTRACT

Online Handwriting Recognition system (HWRS) is a software use to convert handwritten character into text format. This project highlights the development of online handwriting recognition system using Microsoft Visual Basic 6.0, Support Vector Machine (SVM) and, VBTablet 2.0 to recognize the input character by matching the prototypes. This may take place by writing directly on to a digitizing tablet by using stylus which is connected to the Universal Serial Bus (USB) port of a computer. Owing to the fact that each individual has its own way of presenting his/her handwriting on tablet, there is a certain level of complexity like the way of holding the stylus, the strokes use in the writing and the amount of time and pressure put on tablet which are involved in this recognition system. The general on-line recognition procedures are preprocessing, features extraction, coarse classification, detail matching and postprocessing. In this project the classifier used is the Support Vector Machine (SVM) and a new extractor is proposed consisting of time and pressure data, which provides of several more meaningful features of the handwritten character. The experiment results show that the SVM trained using features from time and pressure data makes the time and pressure technique interesting and competitive since the features has certain recognition advantages such as a convex objective function with very fast training algorithms.

ABSTRAK

Online Handwriting Recognitin System adalah satu sistem komputer untuk mengubah tulisan tangan dalam format teks. Projek ini menyetengahkan pembangunan sistem pengesahan menggunakan *Microsoft Visual Basic 6.0*, *Support Vector Machine (SVM)*, *VTablet 2.0* untuk mengenalpasti tandatangan masukan berbanding dengan pangkalan data. Tulisan tangan akan direkodkan dengan menggunakan tablet digital dan pen khas yang disambungkan dengan *Universal Serial Bus Port (USB)* pada komputer. Setiap individu mempunyai corak yang berbeza hadir pada tablet yang hendak ditulis. Oleh itu, system pengenalan ini merangkumi beberapa paras kerumitan seperti cara memegang pen khas, gerak-geri semasa menulis dan jumlah masa dan tekanan hadir pada tablet digital. Prosedur yang diguna untuk membina projek ini ialah pengolahan awal, pengekstrkan rupa bentuk, klasifikasi kasar, perbandingan perincian dan pengolahan akhir telah dibincang. Pengelasan yang digunakan dalam projek ini ialah *Support Vector Machine (SVM)* yang mengandungi masa dan data pada tekanan, dimana melengkapkan ciri/sifat tulisan tangan. Eksperimen telah menunjukkan bahawa latihan SVM yang menggunakan ciri/sifat masa dan tekanan membawa kelebihan dalam tulisan pengecaman.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	PROJECT TITLE	i
	DECLARATION STATUS OF REPORT FORM	ii
	DECLARATION	iii
	SUPERVISOR DECLARATION	iv
	DEDICATION	v
	AKNOWLEDGMENT	vi
	ABSTRACT	vii
	ABSTRAK	viii
	TABLE OF CONTENTS	ix
	LIST OF TABLES	xiii
	LIST OF FIGURES	xiv
	LIST OF ABBREVIATION	xvi
	LIST OF APPENDIX	xviii
I	INTRODUCTION	
1.1	Introduction	1
1.2	Overview	1
1.3	Project Objectives	3
1.4	Problem Statement	3
1.5	Scope Of Work	4
1.6	Thesis Outline	5

II LITERATURE REVIEW

2.1	Introduction	6
2.1.1	Handwriting recognition system	6
2.1.2	On-line and off-line handwriting recognition system	7
2.2	Brief History Of Handwriting Recognition System	9
2.3	Research And Study	10
2.3.1	Types of handwriting	11
2.3.2	Digital ink	12
2.3.3	Digitizer technology	13
2.4	Classifier Algorithm	14
2.4.1	Fuzzy neural network method	14
2.4.2	Neural Network (NN) method	16
2.4.3	Hidden Markov Model (HMM) method	20
2.4.4	Support vector machine (SVM)	21
2.5	System Development	22
2.5.1	Microsoft Visual Basic 6.0	23
2.5.2	Dynamic-link library (DLL)	26
2.5.3	Digitizing tablet	26
2.6	Conclusion	28

III PROJECT METHODOLOGY

3.1	Introduction	29
3.2	Block Diagram Of Handwriting Recognition	29
3.3	Project Methodology	30
3.3.1	Input handwriting	31
3.3.2	Data acquisition (DAQ)	33
3.3.3	Preprocessing	34
3.3.3.1	Normalization	34
3.3.3.2	Re-sampling distance	35
3.3.3.3	Re-sampling time	36
3.3.4	Feature extraction	36
3.3.5	Postprocessing	36

3.4	Flowchart Of Project	37
3.5	Functions and Algorithm of Preprocessing and Feature Extraction	37
3.5.1	Function size normalization	38
3.5.2	Function re-sampling distance	40
3.5.3	Function re-sampling time	40
3.5.4	Function based on time	41
3.5.5	Function based on stroke	41
3.6	System Architecture	43
3.6.1	Data Acquisition of input database	44
3.6.2	Online Handwriting Recognition Module	46
3.7	Conclusion	48

IV RESULT & DISCUSSION

4.1	Introduction	49
4.2	Results	49
4.2.1	The Result of stored database	49
4.2.2	Graphical user interface (GUI) of on-line handwriting recognition system	51
4.3	Experimental Results	53
4.3.1	System validation	53
4.3.2	Data analysis	55
4.3.2.1	First experiment	56
4.3.2.2	Second Experiment	58
4.4	Discussion	60
4.5	Conclusion	61

V CONCLUSION

5.1	Conclusion	62
5.2	Suggestion and Future Work	63

VI	REFERENCES	64
-----------	-------------------	----

VII	APPENDIX	66
------------	-----------------	----

LIST OF TABLE

NO	TITLE	PAGE
2.1	Personality types of handwriting	11
2.2	Description of basic debugging technique	25
4.1	Function of objects in Main GUI	53
4.2	Results Obtained from 20 writers	57
4.3	Analysis results Obtained from 25 writers	58

LIST OF FIGURE

NO	TITLE	PAGE
2.1	On-line data can be seen as a sequence of x-y-coordinates.	8
2.2	Off-line data is a simple 2D image matrix.	9
2.3	Flow Chart of Fuzzy Neural Network Techniques	14
2.4	The step of preprocessing	15
2.5	The step of features extraction	16
2.6	Flow chart of Fuzzy Neural Network Techniques	17
2.7	The step of preprocessing module	18
2.8	The MLP with one hidden layer	19
2.9	The Process of Multi layer Perceptron	19
2.10	Flow Chart of HMM methods	21
2.11	SVM Algorithm	22
2.12	Visual Basic 6.0 Form Windows	24
2.13	Wacom Bamboo Fun Digitizing Tablet and Stylus	27
3.1	Block Diagram of Online Handwriting Recognition System	30
3.2	Project Methodology Flow Diagram	31
3.3	Flow Chart of Software Creation	32
3.4	Connections from Tablet to Laptop through USB	33
3.5	Common principle of normalization	34
3.6	Results obtained on uppercase character 'D' after re-sampling	35
3.7	Flow chart of project	37
3.8	Flow chart and explanations of each pre-processing and feature extraction	38
3.9	Size normalization of character 'A'	39
3.10	Illustrations of direction $X(n)$	41

3.11	Estimation of Writing Direction	42
3.12	Estimation of curvature	43
3.13	Database created by using programming Visual Basic 6.0	45
3.14	Overview of the online handwriting recognition system	47
4.1	Data acquisition of raw data format in text files (.txt)	50
4.2	Introduction GUI of Online Handwriting Recognition System	51
4.3	Main GUI of Online Handwriting Recognition System	51
4.4	SVM Training GUI of Online Handwriting Recognition System	52
4.5	Example Character 'S' is verified	54
4.6	Example curvature Character 'b' is not been verified	55
4.7	Sample of Cinderella type of handwriting characters	56
4.8	Recognition accuracy of different number of character	59

LIST OF ABBREVIATION

ADC	-	Analog to Digital Converter
API	-	Application Programming Interface
DAC	-	Digital to Analog Converter
DAQ	-	Data Acquisition
DLL	-	Dynamic Link-Library
FHR	-	Freestyle Handwriting
GAC	-	Global Assembly Cache
GUI	-	Graphical User Interface
HMM	-	Hidden Markov Model
HWRS	-	Handwriting Recognition System
ICR	-	Isolated Handprint Character Recognizer
IDE	-	Integrated Development Environment
ISA	-	Industry Standard Architecture
MLP	-	Multilayer Perceptrons
NN	-	Neural Network
OCR	-	Optical Character Recognition
OS	-	Operating System
PCI	-	Peripheral Component Interconnect
RAD	-	Rapid Application Development
RAM	-	Random-access Memory
SDK	-	Software Development Kit
SVM	-	Support Vector Machine
TDNN	-	Time Delay Neural Network
TTL-IO	-	Transistor-transistor Logic Input Output
UI	-	User Interface
USB	-	Universal Serial Bus

VB - Visual Basic

LIST OF APPENDIX

NO	TITLE	PAGE
A1	Coding of Introduction GUI	66
A2	Coding of Main GUI	67
A3	Coding of SVM Training GUI	78
A4	Coding of SVM Declaration	83
A5	Coding of Processing and Feature Extraction	86
B	Online Handwriting Recognition System Instruction Manual	100

CHAPTER I

INTRODUCTION

1.1 Introduction

This chapter covers introduction of project. To outline the chapter, it is first discussed the introduction of the project. Then it is followed by the problem that is trying to be improved. Next, the objectives are presented and the remaining sections described the work scope and the structure of this report.

1.2 Overview

Online Handwriting Recognition System is used to convert human handwritten character into text format. This may go on with writing directly on a digitizing tablet by using stylus. The goal of the project is to develop adaptive methods for on-line recognition of handwritten characters. In this case, adaptation means that the system is able to learn new writing styles during its normal use. Due to the learning, the user can use natural style of writing instead of some constrained style. Typical applications for on-line handwriting recognition are small hand-held devices, for example PDAs and hand phone, in which it replaces a keyboard. On-line handwriting recognition system can also be used in many real-world applications. This system when implemented can actually bring more benefits and efficiencies towards daily life of being more convenient. There are some applications that can

probably implement this system such as recognize handwritten mathematical expressions and also improving classroom lecture with technology as a digital educational at low cost.

Handwriting recognition is the ability of a computer to receive and interpret intelligible handwritten input. The image of the written text may be sensed "off line" from a piece of paper by optical scanning (optical character recognition). Alternatively, the movements of the pen tip may be sensed "on line", for example by a pen-based computer screen surface. Handwriting recognition principally entails optical character recognition. However, a complete handwriting recognition system also handles formatting, performs correct segmentation into characters and finds the most plausible words. On-line handwriting recognition involves the automatic conversion of text as it is written on a special digitizer, where a sensor picks up the pen-tip movements $X(t), Y(t)$ as well as pen-up/pen-down switching. That kind of data is known as digital ink and can be regarded as a dynamic representation of handwriting. The obtained signal is converted into letter codes which are usable within computer and text-processing applications.

A new approach for the recognition of handwritings based on the Support Vector Machine (SVM) will be discussed. The Support Vector Machine is a new type of learning machine for pattern recognition and regression problems, which constructs its solution (decision function f) in terms of a subset of the training data and the Support Vector. In recent years Support Vector Machines (SVM) becomes very popular. As they provide very good results for various pattern recognition problems, they also seem to be a good choice for online handwriting recognition. Compared to most methods used for handwriting recognition such as Hidden Markov Models (HMM) or Dynamic Time Warping (DTW), SVM, which is based on the principle of structural risk minimization, has various advantages such as a convex objective function with very fast training algorithms.

On the other hand, SVM typically are applied to data sets containing feature vectors of fixed length and not to problems dealing with time series of variable length such as online handwriting recognition. SVM is mainly used in classification and regression problems. In classification it involves estimation of the decision

function, f using a set of training data with the labels that will correctly classify unseen test examples. Meanwhile, for regression, it is the estimation of real-valued functions, which is carried out in analogous manner to the case of pattern recognition.

1.3 Project Objectives

To achieve the goal of this project, some objectives are defined as a guided:

- i) To study a method of programming development for support vector machine based design by using Visual Basic 6.0.
- ii) To design and develop an online handwriting recognition system by writing through digitizing tablet by obtaining database using Visual Basic 6.0.
- iii) To study and gain knowledge for classifier method in Support Vector Machine (SVM).

1.4 Problem Statement

In recent years, several attempts were made to produce ink pens that include digital elements, such that a person could write on a paper, and have the resulting text stored digitally. A Tablet PC is a special notebook computer that is outfitted with a digitizer tablet and a stylus allowing a user to handwrite text on the unit's screen. The operating system recognizes the handwriting and converts it into typewritten text. The success of these products is yet to be determined.

On-line handwriting recognition systems currently available in the market have functions limited to converting handwritings into typewritten text. It has not achieved widespread use in either desktop computers or laptops. It is still generally accepted that keyboard input is both faster and more reliable. However, this system will be improved by using SVM classifier method which consist features of time and pressure obtained from database.

The technique is far sophisticated than a simple analysis of a character. As a user writes on a pressure-sensitive tablet, the software records the character shape, writing speed, stroke holder, off tablet motion, pen pressure and timing. These characteristics uniquely identify a person and cannot be mimicked or stolen. The system accuracy is improved but it not yet achieved 100% accurate.

1.5 Scope Of Work

This project basically is to develop an online handwriting recognition system by using Visual Basic 6.0 and supported files which are ActiveX control (cwui.ocx), Tablet Devices Wintab API Wrapper (VBTablet.dll) and SVMCORE.dll. The system developed will grab the written texts by the user who writes on adapted digitizing tablet via capturing the input signal. Then the input signal will be matched the database in recognizer machine and transforms the handwritten character into text format. In addition to transforming handwriting into text format, handwriting recognition opens a vast scope of possibilities: from searching for handwritten notes to triggering actions by writing a symbol. In text recognition, the mainstream solutions for data input can be identified by online data input. The input signal is captured by stylus on digitizing tablet. The information on strokes and trajectories is mathematically represented in a digital ink signal composed of a sequence of 2D points by timing order. The input signal is always plotted according to a matrix with $X(t)$ and $Y(t)$ axes and a point of origin.

Besides, the performance of the recognition system is improved by using SVM classifier algorithm to recognize character. Implementing SVM in this project requires a specific program for it to work properly on recognizing character. SVM is actually a supervised training algorithm in which has to be trained first before it can be used for recognize characters. In order to see the performance of SVM in the recognition module, an experiment to test the SVM with real data captured through digitizing tablet will be carried out. Experimental result has to be as low as possible and within an acceptable range.

1.6 Thesis Outline

This thesis consists of five chapters. Chapter 1 discusses on the project introduction. It started by presenting the importance of this project and motivation of the projects. The objective of project, problem statement, scopes of work and methodology will also be presented. Chapter 2 introduces the literature review which discusses the literature review and all information that are related to handwriting recognition. Every facts and information which found through journals or other references will be compared and the better methods have been chosen for the project. Beside that, this chapter also includes the concept and fundamental of the handwriting recognition system. Chapter 3 is regarding the project methodology that consists of two parts which are preprocessing and postprocessing. The system development as stated in the project methodology which includes the design of data acquisition module, a pre processing module, a normalization and re sampling module, a feature extraction module, a classifier module and a decision module. Chapter 4 covers all data and analysis results from the recognition database. Experiments included in this chapter put in forms of table, figure and discussions. Finally, chapter 5 summarizes the whole project. The conclusion, suggestions and recommendations for improvements which can be implemented in future are discussed within this chapter.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

This chapter covers on the literature review on the topic by past researches that have been done. Some of the findings, theories behind the research and also justification are highlighted to the research. Few recognition classifiers are discussed on this chapter.

2.1.1 Handwriting recognition system

Handwriting has continued to persist as a means of communication and recording information in day-to-day live even with the introduction of new technologies. Given its ubiquity in human transaction, machine recognition of handwriting has practical significance as in reading handwritten notes in PDA, in postal address on envelopes or in handwritten field in forms. The handwritten language needs to be transduced into electronic data and thus perform with recognition algorithms. The overviews of methods for online handwriting recognition system are discussed.

The modern inventions in computer and communication technologies such as word processors, fax machines and e-mail are having their impact on handwriting