# FEATURE ENHANCEMENT FOR EXTRACTING ON-LINE ISOLATED HANDWRITTEN CHARACTERS

MUHAMMAD FAISAL ZAFAR

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

AUGUST 2006

Dedicated to my beloved father (late, 7 Dec. 2005)

# ACKNOWLEDGEMENT

All the praise for The **Almighty Allah SWT** for His Blessings and uncountable awards towards me in spite of all of my weaknesses and serious faults. Hundreds of millions **Darood** upon our **Holy Prophet** (Sallallaho alaihi wasallam) being a trivial Ummati.

I wish to thank and express my deepest gratitude and appreciation to the following individuals and organization that enable and motivate me in completing my study and research:

**Associate Professor Dr. Dzulkifli Mohamad** who greatly helped me in every way I needed to go through this study. His personal kindness, skill, patience and guidance has been given and coloring my daily life. As my supervisor, he also becomes my parent.

**Associate Professor Daut Bin Daman**, for his encouragement, guidance, moral support and friendship. Associate Professor **Dr. Siti Mariyam Bt. Hj. Shamsuddin** who always was kind and helpful being Head of Department. **En. Razib Othman** who is my lovely friend and well wisher. All respectable teachers of GMM for their valuable support all the time.

**Professor Dr. Rahmalan and Professor Dr. Safaai Deris** from **School of Graduate Study,** for their full support and cooperation in my thesis submission.

I'm also thankful to my friends Abdul Majid, Anjum Iqbal, Adnan yonus, Imran Ghani, Adil Khattak, Kashif Saleem, Mubarak Ali, Mahmoud Ali, Abdurrahman, Adil Ali, Khalid Sb and Jehan. They were always available for my help whenever I needed. It will be dishonest if I don't thank to all staff of our faculty who gave their best cooperation during my research work.

My fellow postgraduate students Chu Kai Chuan,  Lai Chui Yen, Moon Tan, Ijat, Ferhan, Fadni, Ijam, Masroor and so many others should also be recognized for their valuable suggestions and support.

My parents, **Muhammad Hanif** (late) and **Razzia Beghum**, who gave me a real love, pray, support, and all they have. Special pray for my beloved father who passed away at my final stage of thesis submission. Together with my brothers and sisters, they have been supporting me morally in ups and downs. And to my wife **Laila Khalid** who gives a meaningful love and care to me. The sweet company of my kids, **Jowairia**, **Saleem** and **Laiba** also kept me fresh and alive.

I have no words to pay thanks to a great personality, **Mian Muhammad Bashir Hayat**, who actually the driving force behind all of my efforts and successes and his affection is the real assets for me in this world as well as the next.

## ABSTRACT

The study of online handwriting recognition has gained an immense interest among the researchers especially with the increase in use of the *personal digital assistant* (PDA). The large number of writing styles and the variability between them make the handwriting recognition a challenging area to date. The present tools for modelling are not sufficient to cater for the various styles of human handwriting. Furthermore, the techniques used to get appropriate features, architecture and network parameters for online handwriting recognition are still ineffective. The success of any recognition system depends critically upon how far a set of appropriate numerical attributes or features can be extracted from the object of interest for the purpose of recognition. Thus the aim of this research work is to propose novel feature extraction methods to facilitate a system or device to achieve satisfactory online handwriting recognition. Two new simple and robust methods based on annotated image and sub-character primitive feature extractions have been proposed. The selection of features is based mainly on their effectiveness. Using the proposed techniques and a neural network based classifier, several experiments were carried out using the UNIPEN benchmark database. The techniques are independent of character size and can extract features from raw data without resizing. The maximum recognition rates achieved are 94% and 92% for annotated image and sub-character primitive methods respectively.

# ABSTRAK

Kajian pengecaman tulisan tangan semakin mendapat perhatian para penyelidik, khususnya apabila penggunaannya telah diaplikasikan di dalam peralatan keperluan era baru seperti *personal digital assistant* (PDA). Kepelbagaian gaya tulisan dan kewujudan beberapa pembolehubah yang boleh mempengaruhi gaya tulisan menjadikan pengecaman tulisan tangan satu bidang kajian yang agak mencabar pada hari ini. Peralatan pemodelan yang sedia ada pada hari ini masih tidak mampu menangani kepelbagaian gaya tulisan tangan manusia. Tambahan pula teknik untuk mendapatkan parameter kesesuaian ciri, senibina dan rangkaian untuk mengecam tulisan tangan secara atas talian masih juga kurang berkesan. Keberkesanan suatu sistem pengecaman adalah bergantung sepenuhnya kepada sejauhmana set ciri atau sifat numerik yang sesuai dapat diekstrak daripada objek yang hendak dicam. Oleh itu, tumpuan utama kajian ini adalah untuk mencadangkan kaedah baru pengekstrakan ciri bagi membantu sistem atau alat untuk mendapatkan satu pengecaman tulisan tangan secara atas talian yang lebih berkesan. Dua kaedah baru yang mudah dan tegar berasaskan pengekstrakan ciri imej teranotasi dan primitif sub-huruf telah dibangunkan. Pemilihan ciri dilakukan hanya berdasarkan kepada keberkesanan. Dengan menggunakan kaedah yang telah dibangunkan ini bersama pengelas rangkaian neural, beberapa pengujian telah dilakukan dengan menggunakan data daripada pangkalan data piawai UNIPEN. Teknik ini didapati tidak terhad kepada saiz huruf dan mampu mengesktrak ciri daripada data mentah tanpa perlu pensaizan semula. Kadar pengecaman tertinggi yang telah dicapai adalah 94% untuk imej teranotasi dan 92% untuk primitif sub-huruf.

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

Online handwriting recognition is one of the very complex and challenging problems (Plamondon and Srihari, 2000; Xiaolin and Yeuug, 1997; Plamondon and Privitera, 1999) because of variability of size, writing style of hand-printed characters (Verma *et al*., 2004), and duplicate pixels caused by a hesitation in writing or interpolation of non-adjacent consecutive pixels caused by fast writing. As mentioned in the literature (Parizeau *et al*., 2001; Jaeger *et al*., 2001), the feature extraction plays an important role in the overall process of handwriting recognition. Many feature extraction techniques (Parizeau *et al*., 2001; Jaeger *et al*., 2001; Chakraborty B. and Chakraborty G., 2002; Ping and Lihui, 2002; Gomez *et al*., 1998; Trier *et al*., 1996; Hammandlu *et al*., 2003) have been proposed to improve overall recognition rates; however most of them are dependent on the size and slope of handwriting characters. They require very accurate resizing, slant correction procedure or technique, otherwise they achieve very poor recognition rates. Also most of existing techniques use only one characteristic of a handwritten character. This research focuses on an annotated image feature extraction technique that does not use resizing of a character but it uses overall characteristics of a character and combines them to create a global feature vector. Another approach based on sub-character primitive features of a character has also been experimented. This approach was previously adopted for trademark matching (Zafar, 2003) and now has been extended and modified for isolated character recognition.

To estimate the relative performance of different proposed feature representations, experiments have been conducted using the same UNIPEN data sets using a neural network classifier, namely a multilayer perceptron with backpropagation (Verma *et al*., 2004), trained with a fixed set of parameters. In this way, even if backpropagation may not be the best and the fastest learning algorithm for all situations and problems, it is assumed in this study that its limitations will not change the relative ordering of the different representations, nor it will affect greatly their performance gains on a given data set. Besides, it has been experienced that, when used correctly, it can in fact perform very well on large noisy data sets like UNIPEN that contains broad within class deviations (Parizeau *et al*., 2001).

## 1.1 Handwriting Recognition

Two classes of handwriting recognition systems are usually distinguished: online systems (Tappert et al., 1990); (Anoop et al.,2004);( Cheng-Lin et al., 2004) for which handwriting data are captured during the writing process, which makes available the information on the ordering of the strokes, and offline systems (Steinherz et al., 1999) for which recognition takes place on a static image captured once the writing process is over.

### 1.1.1 On-Line Handwriting and Personal Digital Assistants (PDAs)

Handwriting recognition can be approached from both perspectives, and the current focus of the market is on-line handwriting recognition. With the increase in popularity of portable computing devices such as PDAs and handheld computers (Evan, 2005), (Pen Computing Magazine, 2005), non-keyboard based methods for data entry are receiving more attention in the research communities and commercial sector. Large number of symbols in some natural languages (e.g., Kanji contains 4,000 commonly used characters) make keyboard entry even a more difficult task (Scott, 2000). The most promising options are pen-based and voice-based inputs. Digitizing devices like SmartBoards (Smart Technologies Inc. Homepage, 2005) and computing platforms such as the IBM Thinkpad TransNote (IBM ThinkPad

TransNote, 2005) and Tablet PCs (Windows XP Tablet PC Edition Homepage, 2005) have a pen-based user interface.

In current PDAs, people use input methods which differ from the natural writing habit, e.g., the widespread Graffiti. In Graffiti, each character or symbol is obtained with a specific shape written in one or two strokes as shown in Figure 1.1. These specific symbols are not very user-friendly because user needs to learn the specific shapes. Other systems use a more natural input; however, they still rely on restricted writing styles. Thus, in the majority of these devices, the handwriting input method is still not satisfoatory (Bahlmann and Burkhardt, 2004; Bouteruche *et al*., 2005). Even more difficult for online recognition, a writing which looks similar in a graphical (i.e., offline) representation, can have a different sequential (i.e., online) representation (Bahlmann and Burkhardt, 2004).

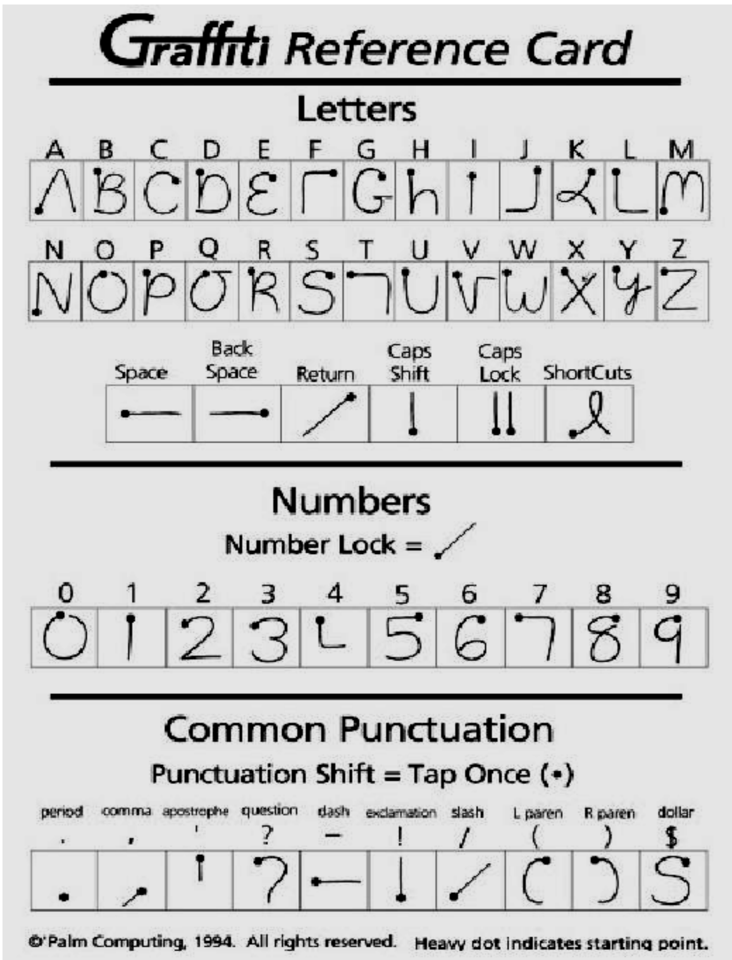


**Figure 1.1:** The Graffiti Character Set. Reproduced from (Scott, 2000)

### 1.1.2 Off-Line Handwriting Recognition

Off-line handwriting recognition focuses on documents that have been written on paper at some previous point in time. Information is presented to the system in the form of a scanned image of the paper document. The literature contains many studies on the recognition of isolated units of writing such as characters, words or strings of digits, which are important subtasks of many applications such as reading texts from pages (Marti and Bunke, 2001), postal addresses (Yacoubi *et al*., 2002), and processing of dates (Morita *et al.,* 2002), courtesy (Oliveira *et al*., 2002) and legal (Gorski *et al*., 2001) amounts on cheques.

Off-line data is two-dimensional in structure because of its image representation and has a typical size of a few hundred kilobytes per word. Since an image has no granted provision to distinguish its foreground and background, the first step of an off-line recognition, called "thresholding" (Liu and Srihari, 1997; Otsu,1978; Sahoo, 1988), is to separate the foreground pixels from the background in the input. Unlike on-line handwriting, a written image also has a line thickness whose width depends on the writing instrument used and the scanning process. Hence the next processing step is to apply a class of techniques called "thinning" or "skeletonization" (Plamondon *et al*., 1993) which tries to shed out redundant foreground pixels from the input. These early preprocessing steps are necessary for off-line recognition but are in general expensive computationally and imperfect, and may introduce undesirable artifacts in the result, for example, "spurs" in the thinning process (Lam *et al*., 1992).

### 1.1.3 Comparisons of On-Line and Off-Line Recognition

An aspect of on-line handwriting recognition, that sets it apart from off-line handwriting recognition, is the temporal input sequence information provided directly by the user. The digitizer naturally captures the temporal ordering information when it samples the points on the contour that the user is forming. Hence on-line data has one-dimensional structure and has a typical size of a few hundred bytes per word. This dynamic information provides clean foreground separation and perfect thinning, and the on-line recognition can bypass the preprocessings that are required by the off-line recognition process. Also the difference in input representation leads to a large difference in the size of the input data. As mentioned above, on-line data, in general, is at least an order of magnitude more compact compared to off-line data because of the different dimensionalities in representation. The difference in the data size also results in substantial difference in the processing time (Jong, 2001).

Meanwhile, an advantage of off-line recognition's image representation is that it is insensitive to variations in the ordering of the strokes contained in handwriting. See Figure 1.2. That is, the same handwriting may have been formed in many different orders of strokes, but the completed written image looks the same and has the same representation. This is not the case for on-line data since different orderings of the strokes will result in different representations even though the completed image is the same. Fortunately, each character class has certain regularity in stroke orderings so that the number of different stroke orders is not large in most cases. In overall comparison, the advantages of on-line handwriting recognition outweigh its disadvantages and on-line recognizers achieve consistently higher accuracy and run faster than the off-line recognizers do. Because of the benefits of on-line recognition, some efforts have studied the interchangeability of the representations (Doermann and Rosenfeld,1995; Nishida, 1995; Fevzi and Ethem,1997; Plamondon and

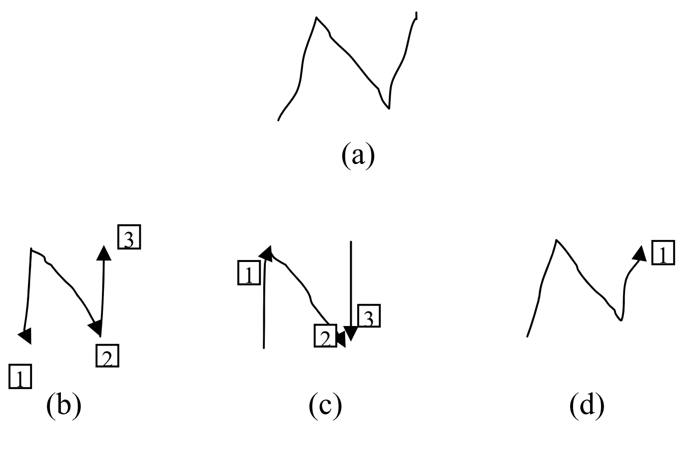


**Figure 1.2:** (a) image of "N" and the three different orders that it could have been written indicated in boxed numbers in (b), (c) and (d).

Privitera,1999). The rationale is that if we have a means to convert off-line data to an on-line version and apply the on-line processing techniques, then we would achieve a level of performance comparable to on-line recognition, on the off-line data. However, the interchangeability has proven to be asymmetric: the conversion from on-line data to off-line version is not hard but the other direction has turned out to be difficult and has led to only limited success (Jong, 2001).

## 1.2 Research Problem Statement

In recent years, several handheld devices like PDAs are in operation, which use the feature of online handwriting recognition. In online handwriting recognition, existing challenges are to cope with problems of various writing fashions, variable size of the same character, different stroke orders of the same letter, and efficient data presentation to the classifier. The similarities between distinct character shapes and the ambiguous writing further complicate the dilemma. A solitary solution of all these problems lies in the intelligent and appropriate extraction of features from the character at the time of writing. The present tools for feature extraction are not yet sufficient to handle online variations in handwriting and there is a lack of techniques, which can find appropriate features. Also the existing techniques are quite complicated and computationally very expensive which are not suitable for small devices. Thus, there is demand to propose feature extraction schemes which are computationally less expensive and are able to handle the above mentioned problems efficiently.

## 1.3 Objectives

The objective of this work is to produce a new feature extraction method for online isolated handwritten characters. The proposed technique is based on a new annotated feature method and the extension of sub-character primitive features.

## 1.4 Research Scope

Scope of this research consists of the following problems:

1. Only online recognition is performed.
2. Only upper case alphabets are considered for isolated character recognition
3. Only section 1b of UNIPEN Train-R01/V07 data set is used for training and testing.
4. Standard backpropagation neural network has been used for classification

## 1.5 Thesis Contributions

The main contributions of this thesis are:

- On-line handwritten scripts are usually dealt with pen tip traces from pen-down to pen-up positions. However, the data obtained needs a lot of preprocessing including filtering, smoothing, slant removing and size normalization before recognition process. Instead of doing such lengthy preprocessing, simple and robust annotated image features have been proposed and experimented which showed very encouraging results. The entire process requires no preprocessing and size normalization.
- Sub-character primitive feature set, previously designed for trademark matching, has been extended and modified for online isolated handwritten characters. For handwriting recognition, this approach has succeeded in having robust pattern recognition features, while maintaining features' domain space to a small, optimum quantity.
- A novel smoothing algorithm has been proposed and implemented for the direction encoded data.

## 1.6 Thesis Organization

Chapter 2 presents the literature review in detail. Chapter 3 gives the description of the methodology of the proposed system for isolated handwritten character recognition using annotated image features. In Chapter 4 methodology for sub-character primitive feature extraction has been discussed in detail. In chapter 5 the thesis has been concluded. In the end, references have been given.

# REFERENCES


Alessandro Vinciarelli (2002). A survey on off-line Cursive Word Recognition. *Pattern Recognition*, 35 (2002) 1433–1446.

Anoop M. Namboodiri, Anil K. Jain. (2004). Online Handwritten Script Recognition. *IEEE Trans. PAMI. 26*(1): 124-130

Anquetil, E. and Bouchereau, H. (2002). Integration of an On-line Handwriting Recognition System in a Smart Phone Device. *Proceedings of the sixteenth IAPR International Conference on Pattern Recognition*, Québec, 192-195.

Ayat N. E., Cheriet M., and Suen C. Y. (2002). Optimization of the SVM kernels using an empirical error minimization scheme. *In Proc. of the International Workshop on Pattern Recognition with Support Vector Machine*, Niagara Falls-Canada, August, pp. 354-369.

Bahlmann C. and Burkhardt H.(2004). The writer independent online handwriting recognition system flog on hand and cluster generative statistical dynamic time warping. *IEEE Trans. PAMI*, 26(3): 299-310, 2004.

Bishop M. (1995). *Neural Networks for Pattern Recognition*. Oxford Univ. Press, Oxford-U.K.

Bortolozzi, F., Britto Jr,. A., Oliveira, L. S. and Morita, M., (2005). *Recent Advances in Handwriting Recognition*. In Umapada Pal et al editors, Document Analysis, 1-31.

Bote_Lorenzo, M., Dimitriadis, Y., and Gomez-Sanchez, E., (2003). "*Automatic Extraction Of Human-recognizable Shape And Execution Prototypes Of Handwritten Characters*", Pattern Recognition, 36(7), 1605-17.

Bouteruche F., Deconde G., Eric Anquetil, and Eric Jamet, (2005). Design and evaluation of handwriting input interfaces for small-size mobile devices. *Proceedings of the 1st Workshop on Improving and Assessing Pen-Based Input Techniques,* 49-56, Edinburgh, Scotland.

Britto Jr., Sabourin R., F. Bortolozzi, and C. Y. Suen (2001). A two-stage hmm based system for recognizing handwritten numeral strings. In *Proc. 6th International Conference on Document Analysis and Recognition*, Seattle-USA, 396-400.

Burr D. J.,(1983). Designing a handwriting reader. *IEEE Trans. PAMI*, 5(5): 554-559.

Byun H., and Lee S. W. (2002). Applications of support vector machines for pattern recognition. In *Proc. of the International Workshop on Pattern Recognition with Support Vector Machine*, Niagara Falls-Canada, August, pp. 213-236.

Cai J. and Z. Q. Liu (1999). Integration of structural and statistical information for unconstrained handwritten numeral recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(3):263-270.

Chakraborty, B., and Chakraborty, G., (2002). A New Feature Extraction Technique for On-line Recognition Of Handwritten Alphanumeric Characters. *Information Sciences*, 148(1), 55-70.

Chan K.F. and Yeung D.Y., (1998) " Elastic structural matching for on-line handwritten alphanumeric character recognition," *Tech. Report, Department of Computer Science*, Hong Kong University of Science and Technology.

Cheng-Lin Liu, Stefan Jaeger, and Masaki Nakagawa (2004). Online Recognition of Chinese Characters:The State-of-the-Art. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(2), 198-203.

Dimauro G., Impedovo S., Pirlo G., Salzo A. (1997). Automatic bankcheck processing: A new engineered system. In S. Impedovo et al, editor, *International Journal of Pattern Recognition and Artificial Intelligence*, World Scientific, 467-503.

Doermann D. S. and Rosenfeld A. (1995). Recovery of Temporal Information form Static Images of Handwriting. *Int' l J. Computer Vision*, vol. 15: 143-164.

Duda Richard O., and Peter E. Hart (1993). *Pattern Classification and Scene Analysis*. New York, Wiley.

Duda R. O., Hart P. E., and Stork D. G. (2001). *Pattern Classification*. John Wiley and Sons, second edition edition.

Dzulkifli Mohamad and Zafar, M. F., (2004). Recognition Of Complex Patterns Using A Novel Proposed Feature Vector by Backpropagation Neural Network", *Pattern Recognition and Image Analysis***,** 3(14), Scientific Council "Cybernetics", the Russian Academy of Sciences, Moscow, Russia.

Dzulkifli Mohamad and Zafar, M. F., (2006). Pattern Matching Based on Robust Features Using Counter-propagation Neural Network", *Journal of Technology and Commerce* (4)/ITB.

Evan Koblentz **(**May 2005). The Evolution Of The PDA: 1975-1995. *Computer Collector Newsletter* , version 0.993.

Favata J T. (2001). Offline general handwritten word recognition using an approximate beam matching algorithm. *IEEE Trans PAMI,* 23(9):1009–1021.

Fevzi Alimoglu and Ethem Alpaydin. (1997). Combining Multiple Representations and Classifiers for Pen-based Handwritten Digit Recognition. *Proc. of the 4th International Conference Document Analysis and Recognition* (ICDAR '97).

Fleetwood, M.D., Byrne, M.D., Centgraf, P., Dudziak, K.Q., Lin, B. and Mogilev, D. (2002). An evaluation of text-entry in palm OS – Graffiti and the virtual keyboard. *Proceedings of Human Factors and Ergonomics Society 46th Annual Meeting*; Baltimore, 23-27 September; 617-621.

Freeman H.. (1974). Computer processing of line-drawing images. *Computing Surveys*, 6(1):57-97.

Gader P. D., Forester B., Ganzberger M., A. Billies, B. Mitchell, M. Whalen, T. Youcum (1991). Recognition of handwritten digits using template and model matching. *Pattern Recognition*, 5(24):421-431.

Gader P. D., . Keller J. M, and Cai J. (1995). A fuzzy logic system for detection and recognition of street number fields on handwritten postal addresses. *IEEE Transactions on FuzzySystems*, 3(1):83-95.

Gomez Sanchez, E., Gago Gonzalez, J.A., Dimitriadis, Y.A., Cano Izquierdo, J.M., and Lopez Coronado, J., (1998). Experimental Study of A Novel Neuro-fuzzy System For On-line Handwritten UNIPEN Digit Recognition. *Pattern Recognition Letters*, 19(3), 357-364.

Gorski N., Anisimov V., Augustin E., Baret O., and Maximov S. (2001). Industrial bank cheques processing: the A2iA checkreader. *IJDAR*, 3:196–206.

Govindan, V. K.. and Shivaprasad, A. P. (1990), Character Recognition-a Review, *Pattern Recognition,* 23 (1990) 671-683.

Guillevic D., and Suen C. Y. (1995). Cursive script recognition applied to the processing of bank cheques. In *Proc. of 3th International Conference on Document Analysis and Recognition*, Montreal-Canada, August, pp. 11-14.

Hammandlu, M., Murali Mohan, K. R., Chakraborty, S., Goyal, S. and Choudhury, D. R., (2003). Unconstrained Handwritten Character Recognition Based On Fuzzy Logic. *Pattern Recognition*, 36(3), 603-623.

Haykin, Simon, (1994). *Neural Networks, A Comprehensive Foundation*. Newyork: Macmillan.

Hiroto Mitoma, Seiichi Uchida, and Hiroaki Sakoe. (2005). Online character recognition based on elastic matching and quadratic discrimination. *Proceedings of 8th International Conference on Document Analysis and Recognition* (ICDAR 2005, Seoul, Korea).1(2):36—40.

Hu J., Brown M. K. and Turin W. (1996). HMM Based On-Line Handwriting ecognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*,18(10): 039-1044.

Hu J., Lim,S., and Brown, M., (1998). HMM based writer independent on-line handwritten character and word recognition. *International Workshop on Frontiers in Handwriting Recognition*, 143–155.

Hu J., Lim,S., and Brown, M., (2000). Writer independent on-line handwriting recognition using an HMM approach, Pattern Recognition 33 (1), 133–147.

IBM ThinkPad TransNote, (2005). http://www-132.ibm.com/content/search/ transnote.html

Ikeda K., Yamamura T., Mitamura Y., Fujiwara S., Y. Tominaga,and T. Kiyono.(1981). On-line recognition of handwritten characters utilizing positional and stroke vector sequence. *Pattern Recognition*, 13(3): 191-206, 1981.

Ishigaki K. and Morishita T., (1988). A Top-Down Online Handwritten Character Recognition Method via the Denotation of Variation. *Proc. 1988 Int'l Conf. on Computer Processing on Chinese and Oriental Languages,* 141-145.

Jaeger, S., Manke, S., Reichert, J., and Waibel, A., (2001). Online Handwriting Recognition: The NPen++ Recognizer. *International Journal on Document Analysis & Recognition*, 3(3), 169-180.

Jain K, Duin R. P. W., Mao J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4-37.

Jong Oh. (2001). An On-Line Handwriting Recognizer with Fisher Matching, Hypotheses Propagation Network and Context Constraint Models. *PhD thesis, Department of Computer Science New York University* USA.

Jungpil Shin (2002). On-Line Handwriting Character Recognition Using Stroke Information" Hendtlass and M. Ali (Eds.): IEA/ AIE 2002, LNAI 2358, 703–714, @ Springer-Verlag Berlin Heidelberg 2002.

Kim H. Y. and Kim J. h. (1998). Handwritten korean character recognition based on hierarchical randon graph modeling. In *Proc. 6th International Workshop on Frontiers of Handwriting Recognition*, Taegon-Korea, August, pp. 557-586.

Kittler J., Hatef M., R. Duin, and J. Matas (1998). On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226-239.

Koerich L. (2002). *Large Vocabulary Off-Line Handwritten Word Recognition*. PhD thesis, École de Technologie Supérieure, Montreal-Canada.

Koerich A. L., Sabourin R., Suen C. Y. (2003). Large vocabulary off-line handwriting recognition:A survey. *Pattern Anal Applic,* 6: 97–121.

Kohonen T. (1987). Adaptive, Associative, And Self-Organizing Functions In Neural Computing. APPLIED OPTICS 26(23):4910-4918.

Költringer, T. and Grechenig, T. (2004). Comparing the Immediate Usability of Graffiti 2 and virtual keyboard. *Proceedings of CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, Vienna, 24-29 April, 1175-1178.

Kundu, He Y., and Chen M. (2002). Alternatives to variable duration hmm in handwriting recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1275-1280.

Kurtzberg J. M., (1987). Feature analysis for symbol recognition by elastic matching. *Technical Report*, IBM Research Center.

Kurtzberg J. M. and Tappert C. C. (1982). Segmentation procedure for handwritten symbols and words. *IBM Tech. Disclosure Bull*, vol. 25: 3848-3852, Dec. 1982.

Lam L., Lee S. W. and Suen C. Y. (1992).Thinning methodologies: A Comprehensive Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14: 869-885.

Lazzerini, and Marcelloni F. (2000). A linguistic fuzzy recognizer of off-line handwrriten characters. *Pattern Recognition Letters*, 21(4):319-327.

LeCun Y., Bottou L., Orr G. B., Muller K. R. (1998a). Eficient backprop. In Orr G. and K. Miller, editors, *Neural Networks: Tricks of the Trade*. Springer.

LeCun Y., Bottou L., Bengio Y., and Haffner P. (1998b). Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE*, 86(11): 2278-2324.

Leedham Graham, Guerfali Wacef, and Plamondon, R. (2002). Special Issue: Handwriting Processing and Applications. *Pattern Recognition, (*35) 5: 981-982.

Lethelier, Leroux M., and Gilloux M. (1995). An automatic reading system for handwritten numeral amounts on french checks. *In Proc. 3th International Conference on Document Analysis and Recognition*, Montreal-Canada, August, pp. pages 92-97.

Ling, M. Lizaraga, N. Gomes, and A. Koerich (1997). A prototype for brazilian bankcheck recognition. In S.Impedovo et al, editor, *International Journal of Pattern Recognition and Artificial Intelligence*,World Scientific, pp. 549-569.

Liu, X.Y., and Blumenstein, M. (2004). Experimental analysis of the modified direction feature for cursive character recognition. *International Workshop on Frontiers in Handwriting Recognition*, 353–358.

Liu Y. and Srihari S. N. (1997).Document Image Binarization Based on Texture Features," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(5):1-5.

MacKenzie, I.S. and Soukoreff, R.W. (2002). Text entry for mobile computing: Models and methods, theory and practice. *Human-Computer Interaction*, 17, 147-198.

MacKenzie, I.S. and Zhang, S.X. (1997). The immediate usability of Graffiti. *Proceedings of the conference on Graphics interface '97*, Toronto, 129-137.

Mantas, J. (1986), An Overview of Character Recognition Methodologies, *Pattern Recognition,* 19 (1986) 425-430.

Marti U. and Bunke H. (2001). Text line segmentation and word recognition in a system for general writer independent handwriting recognition. *In Proc. 6*th *ICDAR*: 159–163.

Matan O., Burges J. C., LeCun Y., Denker J. S. (1992). Multi-digit recognition using a space displacement neural network. In J. E. Moody, S. J. Hanson, and 165 R. L. Lippmann, editors, *Advances in Neural Information Processing Systems*, volume 4, Morgan Kaufmann, pp. 488-495.

Mico L., and Oncina J. (1999). Comparison of fast nearest neighbour classifier for handwritten character recogniton. *Pattern Recognition Letters*, 19(3-4):351-356.

Morita M, Sabourin R., F. Bortolozzi, and C. Suen. (2002). Segmentation and recognition of handwritten dates. *In Proc. 8th IWFHR*: 105–110.

Nag R., K. Wong H., and Fallside F., (1986). Script recognition using hidden Markov models. *Proc. ICASSP*, 3(4): 2071–2074.

Nakai M., Akira N., Shimodaira H., and Sagayama S. (2001). Substroke approach to HMM-based on-line Kanji handwriting recognition," *Proc. ICDAR*, 491–495, 2001.

Nishida H. (1995). An Approach to Integration of Off-Line and On-Line Recognition of Handwriting. *Pattern Recognition Letters*, 16(11): 1213-219.

Nouboud, F., and Plamondon, . (1990), On-Line Recognition of Handprinted Chara.cters: Survey and Beta Tests, *Pattern Recognition,* 23 (1990) 1031-1044.

Odaka K., Wakahara T. and Masuda I., (1986). Stroke-Order-Independent On-Line Character Recognition Algorithm and Its Application. *Rev. Electrical Comm. Laboratories*, 34(1): 79-85.

Oliveira L. S., and Sabourin R. (2004). Support Vector Machines for Handwritten Numerical String Recognition, 9th *International Workshop on Frontiers in Handwriting Recognition*, October 26-29, Kokubunji, Tokyo, Japan, pp 39-44.

Oliveira L. S., Sabourin R., Bortolozzi F., and C. Y. Suen. (2002). Automatic recognition of handwritten numerical strings: A recognition and verification strategy. *IEEE PAMI*,
24(11):1438–1454.

Okamoto M. and Yamamoto K., (1999). On-line handwritten character recognition method using directional features and clockwise/counterclockwise direction-change features. In *International Conference on Document Analysis and Recognition*, 491–494.

Otsu N. (1978).A Threshold Selection Method from Gray-Scale Histogram. *IEEE trans. Systems, Men and Cybernetics*, vol. 8, 62-66.

Parizeau, M., Lemieux, A., and Gagne, C., (2001). Character Recognition Experiments Using Unipen Data. Proceedings, *Sixth International conference on Document Analysis and Recognition*, 481-485.

Pen Computing Magazine: PenWindows, (2005). Availabale at: http://www.pencomputing.com/PenWindows/index.html, Peter N. Belhumeur, J. P. Hespanha, and David J. Kriegman, (1997). Eigenfaces vs. Fisher Faces: Recognition Using Class Specific Linear Projection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7).

Ping, Z., and Lihui, C., (2002). A Novel Feature Extraction Method And Hybrid Tree Classification For Handwritten Numeral Recognition. *Pattern Recognition Letters*, 23(1), 45-56.

Plamondon R., (1993). Special issue on handwriting processing and recognition. *Pattern Recognition* ed.

Plamondon R. and Leedham G. (1990). Computer Processing of Handwriting. eds.: *World Scientific*, Singapore.

Plamondon R. and Privitera C. M. (1999). The Segmentation of Cursive Handwriting: An Approach Based on Off-Line Recovery of the Motor-Temporal Information. *IEEE Trans. Image Processing*, 8(1): 80-91.

Plamondon Rejean, and Sargur N. Srihari, (2000). On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey. 1EEE Transactions on Pattern Analysis and Machine Intelligence. 22(1): 63-84.

Plamondon R., Suen C. Y., Bourdeau M and Barriere C. (1993). Methodologies for Evaluating Thinning Algorithms for Character Recognition. *Int'l J. Pattern Recognition and Artificial intelligence*, special issue on thinning algorithms, 7(5), 1247-1270.

Powalka R Kazimierz. (1995). An algorithm toolbox for on-line cursive script recognition. *PhD thesis The Nottingham Trent University*.

Rabiner L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2): 257-286.

Ratzlaff E. H., (2003). Methods, Report and Survey for the Comparison of Diverse Isolated Character Recognition Results on the UNIPEN Database. *Proceedings of the Seventh International Conference on Document Analysis and Recognition* (ICDAR 2003)

Rigoll G.and Kosmala A., (1998). A systematic comparison between on-line and o_-line methods for signature verification with hidden markov models. In *International Conference on Pattern Recognition*, 1755–1757.

Ripley B. D., (1996). *Pattern Recognition and Neural Networks*. Cambridge, Cambridge University Press.

Sahoo P. K., Soltani S., Wong A. K. C. and Chen Y. C. (1988). A Survey of

Thresholding Techniques," *Computer Vision, Graphics and Image Processing*, vol. 41, 233-260,

Sears, A. and Arora, R. (2002). Data entry for mobile devices: an empirical comparison of novice performance with Jot and Graffiti. *Interacting with Computers*, 413-433.

Schomaker, L.R.B. (1993). Using Stroke- or Character-based Self-organizing Maps in the Recognition of On-line, Connected Cursive Script. *Pattern Recognition* , 26(3): 443-450.

Schurmann J. (1996). *Pattern Classification* - A unified view of statistical and neural approaches. Wiley interscience.

Scott D. Connell. (2000).Online Handwriting Recognition Using Multiple Pattern Class Models. *PhD Thesis, Dept. of Computer Science and Engineering, Michigan State University* USA.

Shridhar M. and Badreldin A. (1986). Recognition of isolated and simply connected handwritten numerals. *Pattern Recognition*, 19(1):1-12.

Simner M., Hulstijn W., and Girouard P. (eds.1994). Forensic: Developmental and Neuropsychological Aspects of Handwriting. *Special issue, J. Forensic Document Examination*.

Simner M. L., Leedham C. G. and Thomassen A. J. W. M. (eds.1996). Handwriting and Drawing Research: Basic and Applied Issues. *Amsterdam: IOS Press*.

Smart Technologies Inc. Homepage, (2005). http://www.smarttech.com/

Steinherz T., Rivlin E., and Intrator N. (1999). Offline Cursive Script Word Recognition—A Survey. *Int'l J. Document Analysis and Recognition*, vol. 2, 90-110.

Suen, C. *Y.,* Berthod, M., and Mori, S. (1980), Automatic Recognition of Ha.nd printed Character-the State of the Art, *Proceedings of IEEE.* 68 (1980) 469-487.

Tappert C. C.,(1982). Cursive script recognition by elastic matching. *IBM J. Res. Develop*. 26:765-771, Nov. 1982.

Tappert C.C. (1984). Adaptive on-line handwriting recognition. *IEEE Seventh Int. Conf. Proc. on Pattern Recognition*, 1004-1007.

Tappert C.C., Suen C.Y., Wakahara T.(1990). The state of the art in on-line handwriting recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(8), 787-808.

Trier, Ø. D., Jain, A.K., and Taxt, T., (1996). Feature Extraction Methods For Character Recognition – A Survey. *Pattern Recognition*, 29(4), 641-662.

Van Galen G. P. and Morasso P. (1998). Neuromotor Control in Handwriting and Drawing. *Acta Psychologica*, Amsterdam: North Holland 100(1-2): 236.

Van Galen G. P. and Stelmach G. E. (1993). Handwriting: Issues of Psychomotor Control and Cognitive Models. *Acta Psychologica*, special volume, Amsterdam: North Holland.

Vapnik V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag, New York-USA.

Verma, B., Lu, J., Ghosh, M., and Ghosh, R. (2004). A Feature Extraction Technique for Online Handwriting Recognition. IEEE International Joint Conference on Neural Networks, IJCNN'04, Hungary, 1337-43.

Vuurpijl L, and Schomaker L. (2000). Two-stage character classification: A combined approach of clustering and support vector classifiers. *Proc 7th International Workshop on Frontiers in Handwriting Recognition*, Amsterdam, Netherlands, 423–432.

Wang F, Vuurpij L, and Schomaker L. (2000). Support vector machines for the classification of western handwritten capitals. *Proc 7th International Workshop on Frontiers in Handwriting Recognition*, Amsterdam, Netherlands; 167–176.

Wakahara T., Suzuki A., Nakajima N., Miyahara S., and Odaka K., (1996). Stroke-Number and Stroke-Order Free On-Line Kanji Character Recognition as One-to-One Stroke Correspondence Problem. *IEICE Trans. Inf. & Syst.*, E79-D(5): 529–534.

Wan J., Wing A. M. and Sovik N. (1991). Development of Graphic Skills: Research, Perspectives and Educational Implications. *London: Academic Press*.

Watt M.Stephen and Xiaofang Xie (2005). Prototype Pruning by Feature Extraction for Handwritten Mathematical Symbol Recognition. In Ilias S. Kotsireas, editor, *Maple Conference 2005*, 423-437, Waterloo, Ontario, Maplesoft.

Windows XP Tablet PC Edition Homepage, (2005). http://www.microsoft.com/windowsxp/tabletpc/default.asp

Winkler H-J., (1996). Hmm-based handwritten symbol recognition using on-line and off-line features. In *International Conference on Acoustics Speech and Signal Processing*, 3438–3441.

Xiaolin, L., and Yeung, D.-Y. (1997). On-line Handwritten Alphanumeric Character Recognition Using Dominant Points In Strokes. *Pattern Recognition*, 30(1), 31-44.

Xie S. L., and Suk M. (1988). On machine recognition of hand-printed chinese character by feature relaxation. *Pattern Recognition*, 21(1):1-7.

Xu Q. (2002). *Automatic Segmentation and Recognition System for Handwritten Dates on Cheques*. PhD thesis, Concordia University, Montreal-Canada, December.

Yacoubi A. E., Gilloux M., and J. Bertille. (2002). A statistical approach for phrase location an recognition within a text line: An application to street name recognition. *IEEE PAMI*, 24(2):172–188.

Yacoubi El, Gilloux M., Sabourin R., and Suen C. Y. (1999). An hmm-based approach for offline unconstrained handwritten word modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):752-760.

Yoshida K. and Sakoe H. (1982). Online handwritten character recognition for a personal computer system. *IEEE Trans. Consumer Electronics*, CE-28(3): 202-209, 1982.

Zafar M. F. (2003). Trademark Matching Algorithm Based On Simplified Feature Extraction. Master Thesis, FSKSM, University Technologi Malaysia.

Zafar M. F., and Dzulkifli Mohamad (2005). Comparison of Two Different Proposed Feature Vectors for Classification of Complex Image. *Jurnal Teknologi*, Universiti Teknologi Malaysia, 42(D):65-82.

Zhang G. P. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man, and Cybernetics* - Part C: Applications and Reviews, 30(4):451-462.

Zhang, Fu M., Yan H., and Fabri M. A. (1999). Handwritten digit recognition by adaptativesubspace self organizing map. *IEEE Trans. on Neural Networks*, 10:939-945.

Zhou J (1999). Recognition and Verification of Unconstrained Handwritten Numeral. *PhD thesis*, Concordia University, Montreal-Canada, November.