Middle-East Journal of Scientific Research 12 (12): 1861-1865, 2012
ISSN 1990-9233
© IDOSI Publications, 2012
DOI: 10.5829/idosi.mejsr.2012.12.12.1232

An Overview of Handwriting Recognition Systems

S. Rajeswari

Department of ECE,
Bharath University, India

Abstract: Handwriting is a uniquely human act, performed daily by millions of people. "Graphology" would be the first idea to hit your mind if you are about to present the concept of "decoding", (or rather the more non-scienti•c determination of some psychological character traces of the writer). But the more mundane and more frequently overlooked, "decoding" of handwriting is the process of figuring out the letters and symbols written for recognition. Handwriting recognition cannot be under estimated, it has its own huddles in the process of recognition. A common complaint of people is that they couldn't read their handwriting. A very useful one for handwriting recognition and one that is used in several recognizers, is a neural network.

Key words: Structural Techniques• Aspect Ratio • Stylus • Graffiti,• Personalization wizard

INTRODUCTION

Handwriting Recognition: Paperdocuments, photographs, touch-screens and other devices are the sources of feed for computer to interpret characters for recognition. The image of the written text may be sensed "off line" from a piece of paper by optical scanning or intelligent word recognition [1]. Alternatively, the movements of the pen tip may be sensed "on line", for example by a screen surface with pen as an input device.

Handwriting recognition principally entails optical character recognition. However, formatting and correct segmentation are also falls under the process of hand writing recognition.

There are two forms of recognition

- Off line recognition ii) on line recognition

In Off-line handwriting recognition, the text in an image are converted into letter codes,which are used by text processing application developed for the purpose of hand writing recognition. Off line handwriting recognition is not easy to recognize and, as of today, OCR engines are primarily focused on machine printed text and ICR for hand "printed" (written in upper case letters) text. There is no OCR engine that supports handwriting recognition as of now.

Problem Domain Reduction Techniques: We have to narrow the problem domain to increase the accuracy of handwriting recognition systems.

Primary Techniques:

- Have to specify specific character ranges
- Utilization of specialized forms

Character Extraction: Off-line character recognition often involves scanning a form or document written sometime back. This requires to extract each and every characters in scanned image. Tools exist that are capable of performing this step [2, 14]. The most common hurdle we face is when connected characters are considered as a single sub-image containing both characters. This causes a major problem in the recognition phase. But still we have handful of algorithms to face this problem at recognition stage

Character Recognition: Recognition engine helps to match each and every extracted character to a matching computer character for further processing.

Methodologies of Cr System: We focus on the methodologies of Character Recognition systems, emphasizing the off-line handwriting recognition issues that we face. For most of the system hierarchical

Corresponding Author: S. Rajeswari, Department Of ECE, Bharath University, India.

FIFTH THIRD EX1024
Page 1 of 5

approach, would befrom pixel to text, as follows: Pixels Features Characters Subwords Meaningful Text.This bottom up approach varies a great deal, depending upon the type of the CR system we use and the methodology we follow [3,9]. The expert review in the field of character recognitionindicates that the above mentioned hierarchical tasks are grouped in the differentphases of the CR for Pre-processing, segmentation, representation, training and recognition, post processing. The process of CR will completes after following theses stages with at most perfection

Pre-processing: Based on the source of input, raw data should undergo number of preliminary phases to make it suitable for CR system [4]. The process varies from one source of input to other source of input. Pre processing facilitated CR system to operate accurately by providing data which more suitable to it. The main objectives of pre processing are:

- Noise reduction,
- Normalization of the data,

Segmentation: Segmentation is an important stage, as segregation of words, lines or characters has a huge impact on how CR system handle the input [6,13]. The four types of segmentation are

External Segmentation
Internal Segmentation
Explicit Segmentation
Implicit Segmentation

Representation: Image representation plays vital roles in a recognition system. To put it simple, gray-level or binary images are feeded to a recognizer [4,12. However, in most of the available recognition systems, inorder to get rid ofcomplexity and totune-in the accuracy of the algorithms, a more compact and characteristic representation is required. Hundreds of documents provide feature extraction methods for character recognition. Image representation methods are categorized in three major groups as:

-Global Transformation and Series Expansion
Statistical Representation, Geometrical and Topological Representation

Training and Recognition Techniques: CR systems usesdifferent methodologies pattern recognition,it assigns a sample which is unknown into a predefined class. Below approaches to pattern recognitions can explain about CR, as suggested as

- Template Matching,
- Statistical Techniques,
- Structural Techniques,
- Neural Networks.

Post Processing: It is well known that humans read by context up to 60% for careless handwriting. The review of the recent CR research indicates minor improvements, when only shape recognition of the character is considered. This is done in the post processing stage with a feedback to the early stages of CR. The simplest way of incorporating the context information is the utilization of a dictionary for correcting the minor mistakes of the CR systems.

Feature Extraction: Feature extraction works in a similar fashion to neural network recognizers however, Its mandatory for programmers to manually determine the properties they feel as important

Some Example Properties Might Be:

- Aspect Ratio
- Percent of pixels above horizontal half point
- Percent of pixels to right of vertical half point
- Number of strokes
- Average distance from image center
- Is reflected y axis
- Is reflected x axis

Recognizer gain more control over properties considered for recognition by using these approaches. Still any system using this approach requires more development time than a neural network because of the reason that properties are not learned automatically.

On-line Recognition: On-line handwriting recognition will automatically recognize input characters on the go, input source for this varies based on the technology used (example: PDA) Input character should be recognised immediately once it is given as input,sensor will be picking up a pen tip movement and pen switchings. This is called digital ink and it's a source of input for character recognition. The signal captured from input device will be converted into a matching computer input and computer will proceed with further processing.

FIFTH THIRD EX1024
Page 2 of 5

The elements of an on-line handwriting recognition interface typically include:

- a pen or stylus, user should writ with this.
- a touch sensitive surface, which may be integrated with, or adjacent to, an output display.
- a software application which interprets the movements of the stylus across the writing surface, translating the resulting strokes into digital text.

General Process: The process of online handwriting recognition can be broken down into a few general steps:

- preprocessing,
- feature extraction and
- classification.

Preprocessing discards irrelevant information that can affect CR system performance..[3] This can put speed and accuracy at risk. Pre processing usually has phases like binarization, normalization, sampling, smoothing and denoising.[4] The second step followed will be feature extraction. From 2 or more dimensional vector field fetched from pre-processing algorithm, higher dimensional data is fetched. This step is for projecting important information for recognition model. This data may include data like pen pressure, velocity or the changes of writing direction. The last important and crucial step is classification. In this step various models are used to map the extracted features to different classes and thus identifying the characters or words the features represent.

Hardware: Commercial products with handwriting recognition as aalternative for keyboard input came into market in the early 1980s. Examples include handwriting terminals such as the Pencept and the Inforite point-of-sale termin With the advantage of the huge consumer market for personal computers, several commercial products were introduced to replace the keyboard and mouse on a personal computer with a single pointing/handwriting system, such as those from PenCepCIC and others. The first commercially available tablet-type portable computer was the GRiDPad from GRiD Systems, released in September 1989. Its operating system was based on MS-DOS.

Advancements in electronics allowed the computing power necessary for handwriting recognition to fit into a smaller form factor than tablet computers and handwriting recognition is often used as an input method for hand-held PDAs. The first PDA to provide written input was the Apple Newton, which exposed the public to the advantage of a streamlined user interface. However, the device was not a commercial success, owing to the unreliability of the software, which tried to learn a user's writing patterns. By the time of the release of the Newton OS 2.0, wherein the handwriting recognition was greatly improved, including unique features still not found in current recognition systems such as modeless error correction, the largely negative first impression had been made. Palm later launched a successful series of PDAs based on the Graffiti recognition system. Graffiti improved usability by defining a set of "unistrokes", or one-stroke forms, for each character. This narrowed the possibility for erroneous input, although memorization of the stroke patterns did increase the learning curve for the user. The Graffiti handwriting recognition was found to infringe on a patent held by Xerox and Palm replaced Graffiti with a licensed version of the CIC handwriting recognition which, while also supporting unistroke forms, pre-dated the Xerox patent. The court finding of infringement was reversed on appeal and then reversed again on a later appeal. The parties involved subsequently negotiated a settlement concerning this and other patents Graffiti (Palm OS).

A Tablet PC is a special notebook computer that is outfitted with a digitizer tablet and a stylus and allows a user to handwrite text on the unit's screen. The operating system recognizes the handwriting and converts it into typewritten text. Windows Vista and Windows 7 include personalization features that learn a user's writing patterns and/or vocabulary for English, Japanese, Chinese Traditional, Chinese Simplified and Korean. The features include a "personalization wizard" that prompts for samples of a user's handwriting and uses them to retrain the system for higher accuracy recognition. This system is distinct from the less advanced handwriting recognition system employed in its Windows Mobile OS for PDAs.

Although handwriting recognition is an input form that the public has become accustomed to, 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. As of 2006, many PDAs offer handwriting input, sometimes even accepting natural cursive handwriting, but accuracy is still a problem and some people still find even a simple on-screen keyboard more efficient.

FIFTH THIRD EX1024
Page 3 of 5

Software: Initial software modules could understand print handwriting where the characters were separated. Commercial examples came from companies such as Communications Intelligence Corporation and IBM. In the early 90s, two companies, ParaGraph International and Lexicus came up with systems that could understand cursive handwriting recognition. ParaGraph was based in Russia and founded by computer scientist StepanPachikov while Lexicus was founded by Ronjon Nag and Chris Kortge who were students at Stanford University. The ParaGraphCalliGrapher system was deployed in the Apple Newton systems and Lexicus Longhand system was made available commercially for the PenPoint and Windows operating system. Lexicus was acquired by Motorola in 1993 and went on to develop Chinese handwriting recognition and predictive text systems for Motorola. ParaGraph was acquired in 1997 by SGI and its handwriting recognition team formed a P andI division, later acquired from SGI by Vadem. Microsoft has acquired CalliGrapher handwriting recognition and other digital ink technologies developed by P andI from Vadem in 1999. Wolfram Mathematica (8.0 or later) also provides a hand writing or text recognizing function can be called by writing command TextRecognize user can then drag the picture to be analysed on the place of "^".

DISCUSSIONS

In this study, we have overviewed the basic approaches in the CR domainattemptingto bring out the present status of CR research. Although each of themethods mentioned have their own advantages and limitations, the success level is achieved by each method. However, taking in to account various databases, constraints and sample spaces it is difficult to comment about the success of recognition methods, especially in termsof recognition rates. An improved approach in almost all stages of CR research is needed in case of texts which are handwritten under poor conditions or for free style handwriting. Number digit or limited vocabulary form (bank checks, envelopes and forms designed for specific applications) recognition is a popular application area.

REFERENCES

1. Nafiz Arica, FatosT.YarmanVural, "Overview of Character Recognition Focused on offline handwriting,"IEEE Transactions on systems, man and Cybernetics-part c: applications and Reviews, 31(2): 2001.
2. Plamondon, R., 1993. Special Issue on Hand-Written Processing and Recognition, Pattern Recognition, 26(3): 1993.
3. Said, H.E.S., T.N. Tan and K.D. Baker, "Personal Identific.
4. ation Based on Handwriting", Pattern Recognition, vol.33, no.1, pp.149- 160, 2000.
5. Yetrie Ludang, "A Brief Review on the Role of Community to Lessen the Rate of Global Warming", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 69-75. 2010. Ihsan Nuri Demirel, "Knowledge Level of National Education Managers upon In-Service Education Concepts Such as "Holiday Activity Logic, Expectation Difference, Concept Visibility, Attitude of Workers and Upper Stage Managers"", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 08-14, 2010.
6. Ihsan Nuri Demirel, "Conceptual Description, Hastiness in Obtaining Result, Voluntary Participation, Assumption of Inefficiency", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 15-21, 2010.
7. Saman Shahbaz and Muhammad Qaiser Shahbaz, "On Bivariate Concomitants of Order Statistics for Pseudo Exponential Distribution", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 22-24, 2010.
8. Hadji Mohammad Takalloozadeh, "Effects of Host Plants and Various Temperatures on Population Growth Parameters of Aphis gossypii Glover (Hom.: Aphididae)", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 25-30. 2010.
9. Sobhy, E. Elsilk and A. Ahmed El-barbary, 2010. "Antimicrobial Activity Studies on Some 1,2,4-triazine Derivative *in vitro*", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 31-36.
10. Siva Kumar, J., T. Siva Kumar, P. Santhanam and M. Masilamaniselvam, 2010. "Effect of Mutation on Antimicrobial Activity of Actinomycetes from Western Ghat's of India", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 37-44.

A. Anis Najiha, A.Y. Tajul, M.H. Norziah and W.A. Wan Nadiah, 2010. "A Preliminary Study on Halal Limits for Ethanol Content in Food Products", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 45-50.

11. Marjan Panahi, Rahim Naseri and Reza Soleimani, "Efficiency of Some Sweet Corn Hybrids at Two Sowing Dates in Central Iran", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 51-55, 2010.

FIFTH THIRD EX1024
Page 4 of 5

12. Zafar, S. Khan, N. Vishal Shinde and P. Nilesh, 2010. Bhosle and Sahera Nasreen, "Chemical Composition and Antimicrobial Activity of Angiospermic Plants", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 56-61.
13. Rahman Mohseni-Astani, Poorya Haghparast and Sahab Bidgoli-Kashani, "Assessing and Predicting the Soil Layers Thickness and Type Using Artificial Neural Networks - Case Study in Sari City of Iran", Middle-East Journal of Scientific Research, ISSN:1990-9233, 6(1): 62-68, 2010.
14. Yetrie Ludang, 2010. A Brief Review on the Role of Community to Lessen the Rate of Global Warming, Middle-East Journal of Scientific Research 6 (1): 69-75.

FIFTH THIRD EX1024
Page 5 of 5