Sindh Univ. Res. Jour. (Sci. Ser.) Vol. 47 (4) 699-702 (2015)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Features Used in Online Handwriting and Signature Recognition Systems: A survey

Z. JAN[++], M. SHABIR, M. A. KHAN, S.I.A SHAH* B. BALOCH**

Department of Computer Science, Islamia College University Peshawar, Pakistan

Received 21st April 2015, Revised 29th October 2015

Abstract: Handwritten characters, words, signatures, and maps recognition is an important and complex problem. It has numerous useful applications in different fields like banks, shops, teaching, data capturing and touch screen devices. Many online and offline handwriting recognition survey has been published. These surveys not clearly mention the features and type of data used. Online method relies on extraction of features from written words, characters, signatures and maps. This survey identifies features and type of data used in online handwriting recognition systems.

Keywords: Online Handwriting Recognition, Online Features, Online Types of Data

1. INTRODUCTION

One of the most important characteristic of a personal computer system is its interface, with the help of which user interact with implementation. At the early stage of computer there were no interaction of a user with the system hence users and operators were not equal. With the passage of time, computer got maturity and different user interactive devices were developed. These devices are used for the input of data into the computer system. There are three basic input methods for the computer system such as online input method, offline input method and batch input method.

1.1 Online Input Method and Hardware Devices

Online input method is a way of input data directly into the computer system at the time of interaction. There are different online hardware base devices developed for the interaction of a user with the system i.e. Microphone, Keyboards, Webcams, Mice, Joysticks, Light pen, Track ball, Graphics tablets, Gamepad, Touch pads, Plotters, Touch screen, Remote, Light gun and Steering wheel etc. These devices collect information and transform it into a series of electronic signals to the computer, are called online input method devices. Pen-based touch screen/tablet interfaces works as pointing devices that are very efficient in the field of handwriting recognition. The motion of the pen can generate some useful information e.g. X-Y coordinates, pressure and speed etc.

1.2 Offline Input Method and Hardware Devices

Offline input method is another way to input data into the computer from source documents or from two dimensional images not concerned with real time. Researcher usually use this method to process the images and developed useful application. Common devices of this method are Scanners, Optical mark reader, Optical character reader, Bar code reader, Magnetic ink character reader etc. Microscopes are also used in offline input method; the microscope devices take images of the slides during examination and store or send these images into the permanent storage area. Some of the microscopic devices are; optical, compound, stereo, X-Ray, scanning acoustic, neutron microscope, electron microscope and Olympus bx- 51 etc.

1.3 Batch Input Method and Hardware Device

Batch input method is the oldest input method. The computer systems of 1950s were based on batch input method. Those computer systems were read data from a batch of punch cards with the help of card reader.

1.4 Segmentation VS Holistic

Handwriting stroke trajectories can process in two ways (.i.e. segmentation approach and holistic approach). Segmentation is the division process of a compound object into its part. In case of words trajectory, segmentation is the separation of a word into characters. Segmentation is by self a complex task. There is a huge possibility of error due to unreliable segmentation and will affect the recognition rate. Another approach is holistic, in which the whole stroke trajectory is considered for classification and recognition. In this method the whole stroke trajectory is processed.

2. COMMON ARCHITECTURE FOR ONLINE HANDWRITING RECOGNITION

All handwriting recognition systems follow a common architecture which is shown in **(Fig. 1).**

[++]Corresponding author, Z. Jan: zahoor.jan@icp.edu.pk,
*Department of Mathematics, Islamia College University Peshawar, Khyber Pakhtunkhwa, Pakistan.
**Department of Statistics, Sindh Agriculture, University, Tando Jam

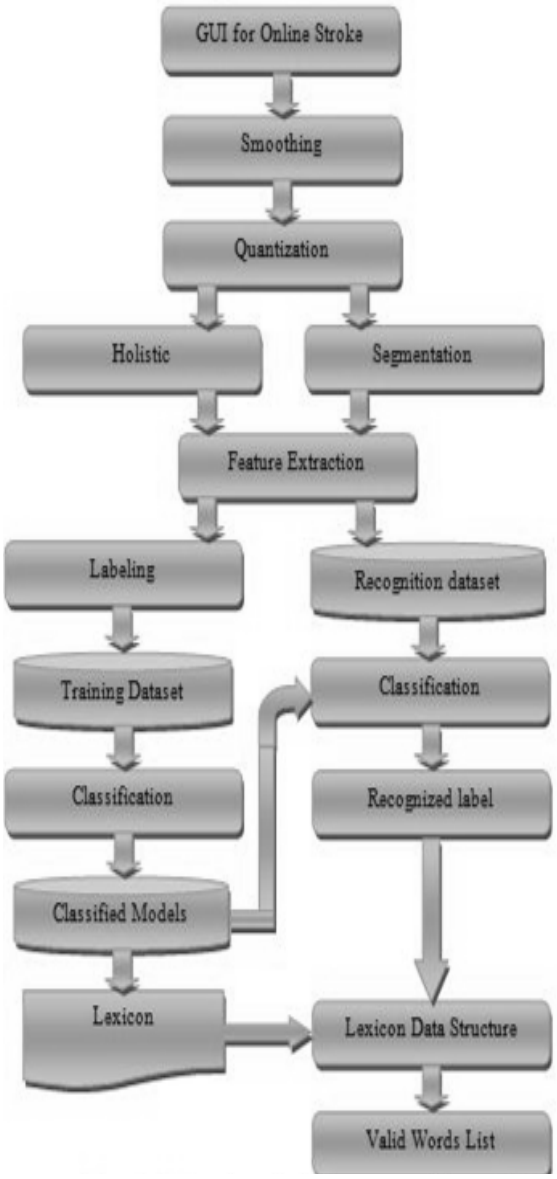


Fig 1: Online handwriting recognition architecture

The GUI reads motion and coordinates of the mouse or stylus from the user. Both stylus and mouse reporting rate is affected by the speed with which user moves the device due to which noise is generated. In order to reduce noise, the mouse/stylus readings are passed through a filter. Filter performs smoothing and minimizes zigzag noise. The output from filter then goes to resampling algorithm which tries to minimize the number of line segments representing the line. The Feature Extractor module extracts features of the stroke. Feature engineering is a very important phase and is a part of every online handwriting recognition system as shown in **(Fig 1).** A complete feature vector for a stroke is the concatenation of several computed features. Featured vectors are then labelled manually in case of supervised training. These labelled feature vectors are added to training dataset. The training dataset is then used to train classifiers; After training phase, machine-learning model is used for recognition and classification of unseen strokes, The unseen strokes are similarly obtained through GUI, perform smoothing, apply re-sample technique and features are extracted as it was in training phase and then store in recognition dataset. The classifier used in training phase must be used in recognition phase.

The labels produced after recognition is put to post-processing phase. The valid combinations of labels are then filtered out using lexicon data structure. The lexicon is derived from a valid corpus of words.

3 FEATURES AND TYPE OF DATA USED IN ONLINE HANDWRITING SYSTEM

Table 1 illustrates the online features and types of data used in online handwriting recognition. These features and types of data are extracted from Urdu, Arabic, English, Kanji, Panjabi, Persian, Telagu, Devanagari, Tamil and Nepali online handwriting characters or words related works. Some of the features and types of data are also taken from online handwriting signature verification and shapes recognition papers.

Table 1: Online handwriting features and type of data

Online Features and Types of Data Used	Online Features and Types of Data Used
Local-angle, loop-presence and super-segment. (Biadsy *et al.,* 2006).	Pen positions (x, y), pen up direction, velocity of the pen and pen down direction. (Nakai *et al.,* 2001).
Arc type, relative positions, length ratio and connection angle. (Sternby *et al.,* 2009).	Ascenders, cusps, closures, descender, dots, direction, chain code, angle, Fourier coefficients, Slop, relative position, cross correlation. Pair-wise distinction, Pencept, number of strokes, position of strokes, character sequences, velocity, force and acceleration. (Charles *et al.,* 1990).
Dot, bumps, line terminals and cusp. (Aparna *et al.,* 2004).	Length, curviness direction, direction and delayed distance. (Mcinerney *et al.,* 1999).
Point locations, curvature, second and first order derivative, the four lowest frequency Fourier coefficients, total number of Fourier coefficients and the total number of points of a generic stroke. (Toselli *et al.,* 2007).	Orientation rate, x and y coordinate, cosine angle, rate of curvature, sine angle, dx and dy, sine and cosine angles and vertical location. (Connell *et al.,* 2002).
Vertical position, aspect, writing direction, curvature, curliness, pen-up/pen-down, slope, hat, linearity, context maps, ascenders and descenders. (Jaeger *et al.,* 2001).	Differences between two consecutive points, (x,y) relative to the first point and differences of curvature between two consecutive points. (Kholmatov *et al.,* 2005).
The number of points, curve length , co-ordinates of points, size of bounding box, baseline strokes, variable-length sequences of frames, relative position and bottom strokes. (Swethalakshmi *et al.,* 2006).	Total vertical size to total horizontal size ratio, total time, number of strokes, arc length, number of horizontal and vertical coordinate sign changes, average slant, total length vertically and total length horizontally. (Krasnik *et al.,* 2004).
Base shape chose, size of stroke, angles, x and y co-ordinates, relative vertical positions, Sub stroke duration and relative horizontal position. (Artieres *et al.,* (2007).	Speed, neighbours point, directions, speed magnitude with respect to time, cosine and sine angle, pressure and the pen-tilt, strokes number, standard deviation, 4 quadrants histogram, root mean square and auto correlation. (Fuentes *et al.,* 2002).
Relative pair wise distances, x and y co-ordinate, relative pair wise angles and point density. (Izadi *et al.,* 2008).	Fourier transform, input signal components, Fourier descriptors and zero-padding. (Yanikoglu *et al.,* 2009).
Pen-up/pen-down, hat, speed, direction, x and y-coordinate, vicinity slope, curvature, vicinity curliness, vicinity aspect and vicinity linearity. (Liwicki *et al.,* 2007).	Signature length, maximum, minimum, mean, width, skew, kurtosis, maximal energy and pen-up time. (Gruber et al 2006).
Pen up/down positions, relative lengths, direction of pen movement, Euclidean distance and angle variation. (Garain *et al.,* 2002).	Pen-up, pen down, pen coordinates, direction θ and curvature. (Julita *et al.,* 2009).
Pen\pressure, velocity equivalent, x and y coordinates, angle, vicinity slope, direction of consecutive angles, average square distance and vicinity curliness. (Schenk *et al.,* 2008).	Close loop, open loop, direction, straight downward line, straight vertical up line, long horizontal left, long diagonal left, semi-circle, curves variation and cusp. (Razzak *et al.,* 2011).
Pen-up/pen-down, hat, speed, cosine and sine angles, vicinity aspect, vicinity slope, straight line, vicinity curliness and average of square distance points. (Liwicki *et al.,* 2007).	Vertical line, horizontal line, slant line, starting point & ending point, slope, writing direction, size/length and hat features. (Malik *et al.,* 2005).
Loops, cusp, endpoints, starts points and direction. (Razzak *et al.,* 2011).	Loops, cusp, endpoints, starts points, direction, size, starting points and ending point. (Razzak *et al.,* 2012).
Slope value, break point, signs of the slope values and X-Y positions. (Ahmad *et al.,* 2008).	Circle, loops, hooks and strokes. (Qiao *et al.,* 1993).

4 CONCLUSION

A large amount of advance effort is required in feature extraction to minimize the gap between human recognition and machine recognition capabilities. This paper will help to select efficient features for research works. Cleverly selection of features is a key to high accuracy. The features mentions in this review are taken from different online handwriting words, characters, maps and signatures recognition systems. By using these features high accuracy has been achieved as mentioned in their papers. We suggest some features that improve the classification and recognition accuracy i.e Hexagonal grid, Triangular grid, Spiral grid, H fractal, Star fractal, Umayyad Pattern (Muslim art) and Dado panel (Muslim art). There some other geometrical shapes and patterns that can be added to extend the number of features i.e primordial symbol, Anklets of Krishna and snake fractal.

REFERENCES:

Aparna, K., V. Subramanian, M. Kasirajan, G. V. Prakash, V. Chakravarthy and S. Madhvanath (2004). Online handwriting recognition for Tamil. Frontiers in Handwriting Recognition, IWFHR-9. Ninth International Workshop on, IEEE.

Artieres, T., S. Marukatat and P. Gallinari (2007). "Online handwritten shape recognition using segmental hidden markov models." Pattern Analysis and Machine Intelligence, IEEE Transactions on 29(2): 205-217.

Ahmad, A.T. and H. Maen (2008). "Recognition of on-line handwritten Arabic digits using structural features and transition network." Informatica 32: 275-281.

Biadsy, F., J. El-Sana and N. Y. Habash (2006). Online arabic handwriting recognition using hidden markov models, Proceedings of the 10th International Workshop on Frontiers of Handwriting and Recognition.

Connell, S. D. and A. K. Jain (2002). "Writer adaptation for online handwriting recognition." Pattern analysis and machine intelligence, ieee transactions on 24(3): 329-346.

Fuentes, M., S. Garcia-Salicetti and B. Dorizzi (2002). On line signature verification: Fusion of a Hidden Markov Model and a neural network via a support vector machine. Frontiers in Handwriting Recognition, Proceedings. Eighth International Workshop on, IEEE.

Garain, U., B. Chaudhuri and T. Pal (2002). Online handwritten Indian script recognition: a human motor function based framework. Pattern Recognition, Proceedings. 16th International Conference on, IEEE.

Gruber, C., C. Hook, J. Kempf, and B. Sick (2006). A flexible architecture for online signature verification based on a novel biometric pen. Adaptive and Learning Systems, IEEE Mountain Workshop on, IEEE.

Izadi, S. and C. Y. Suen (2008). Online Writer-Independent Character Recognition Using a Novel Relational Context Representation. Machine Learning and Applications, ICMLA'08. Seventh International Conference on, IEEE.

Julita, A., S. Fauziyah, O. Azlina, B. Mardiana, H. Hazura and A. Zahariah (2009). Online signature verification system. Signal Processing Its Applications, CSPA 5th International Colloquium on, IEEE.

Jaeger, S., S. Manke, J. Reichert and A. Waibel (2001). "Online handwriting recognition: the NPen++ recognizer." International Journal on Document Analysis and Recognition 3(3): 169-180.

Kholmatov, A. and B. Yanikoglu (2005). "Identity authentication using improved online signature verification method." Pattern recognition letters 26(15): 2400-2408.

Krasnik, S. (2004). "On-Line Handwritten Word Recognition." Retrieved at<>, CS674 Natural Language Processing: 13-18.

Liwicki, M. and H. Bunke (2007). Combining on-line and off-line systems for handwriting recognition. Document Analysis and Recognition, ICDAR Ninth International Conference on, IEEE.

Liwicki, M., A. Graves, H. Bunke and J. Schmidhuber (2007). A novel approach to on-line handwriting recognition based on bidirectional long short-term memory networks. Proc. 9th Int. Conf. on Document Analysis and Recognition.

Mcinerney, S. and R. B. Reilly (1999). Preprocessing for online handwriting recognition. Proceedings of Irish Signal and System Conference.

Malik, S. and S. A. Khan (2005). Urdu online handwriting recognition. Emerging Technologies,. Proceedings of the IEEE Symposium on, IEEE.

Nakai, M., N. Akira, H. Shimodaira and S. Sagayama (2001). Substroke approach to HMM-based on-line Kanji handwriting recognition. Document Analysis and Recognition, Proceedings. 6th International Conference on, IEEE.

Qiao, Y. and G. Leedham (1993). "Segmentation and recognition of handwritten Pitman shorthand outlines using an interactive heuristic search." Pattern Recognition 26(3): 433-441.

Razzak, M. I. and A. A. Mirza (2011). "Ghost character recognition theory and Arabic script based languages character recognition." Przeglad Elektrotechniczny (Electrical Review) 87(11): 234-238.

Razzak, M. I. (2011). Online Urdu Character Recognition In Unconstrained Environment, PhD thesis, International Islamic University, Islamabad.

Razzak, M. I. and M. T. Simsim (2012)."Feature Level Fusion for Online Arabic Script Based Languages Character Recognition." Przeglad Elektrotechniczny (Electrical Review).

Sternby, J., J. Morwing, J. Andersson and C. Friberg (2009). "On-line Arabic handwriting recognition with templates." Pattern Recognition 42(12): 3278-3286.

Swethalakshmi, H., A. Jayaraman, V. S. Chakravarthy and C. C. Sekhar (2006). Online handwritten character recognition of Devanagari and Telugu Characters using support vector machines. Tenth International Workshop on Frontiers in Handwriting Recognition.

Schenk, J., S. Schwärzler and G. Rigoll (2008). PCA in On-Line Handwriting Recognition of Whiteboard Notes: A Novel VQ Design for Use with Discrete HMMs. Int. Conf. on Frontiers in Handwriting Recognition.

Toselli, A. H., M. Pastor and E. Vidal (2007). On-line handwriting recognition system for tamil handwritten characters. Pattern Recognition and Image Analysis, Springer**:** 370-377.

Tappert, C. C., C. Y. Suen and T. Wakahara (1990). "The state of the art in online handwriting recognition." IEEE Transactions on pattern analysis and machine intelligence 12(8): 787-808.

Yanikoglu, B. and A. Kholmatov (2009). "Online signature verification using fourier descriptors." EURASIP Jou. on Advances in Signal Processing.