A Ballistic Stroke Representation of Online Handwriting for Recognition

Prabhu Teja S and Anoop M Namboodiri
Center for Visual Information Technology
International Institute of Information Technology, Hyderabad - 500032, India
prabhuteja.s@research.iiit.ac.in, anoop@iiit.ac.in

***Abstract*—Robust segmentation of ballistic strokes from online handwritten traces is critical in parameter estimation of stroke based models for applications such as recognition, synthesis, and writer identification. In this paper we propose a new method for segmenting ballistic strokes from online handwriting. Traditional methods of ballistic stroke segmentation rely on detection of local minima of pen speed. Unfortunately, this approach is highly sensitive to noise, in sensing and in both spatial and temporal dimensions. We decompose the problem into two steps, where the spatial noise is filtered out in the first step. The ballistic stroke boundaries are then detected at the local curvature maxima, which we show to be invariant to temporal sampling noise. We also propose a bag-of-strokes representation based on ballistic stroke segmentation for online character recognition that improves the state-of-the-art recognition accuracies on multiple datasets.**

I. Introduction

As the primary computing and communicating device of most people is becoming smaller from the desktops or laptops to tablets and smart phones, the interface for such devices for control and data entry should be made robust, efficient, and intuitive. Touch based interfaces to control such devices have made tremendous advances with the advent of capacitive touch screens and gesture recognition algorithms. However doing any amount of data entry using touch based interfaces is still a challenging task. On-screen keyboards development is in a nascent stage to be used extensively, effectively, especially for scripts with large number of symbols such as Indic and Han scripts. An effective alternative would be to use handwriting recognition, which allows a larger character set to be input to devices with limited screen size. Unfortunately the sensitivity and robustness of most touch and even pen based interfaces makes handwriting recognition difficult when compared to dedicated pen interfaces for Tablet PCs or stand alone digitizers.

The error in sensing the pen or finger tip, both in spatial and temporal dimensions has adverse affects on any recognition algorithm. Consider the problem of detecting the ballistic strokes from pen or finger movement in a mobile device. The most common approach is to detect minima in pen velocity as boundaries of ballistic strokes. However, as Figure 1 demonstrates, the noise introduced in such devices would make the process extremely difficult. The error in digitizing devices may be categorized into two: spatial and

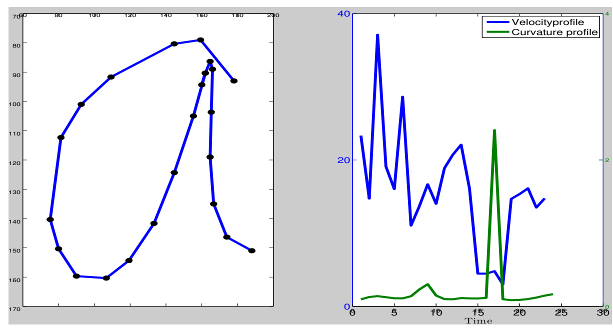

Figure 1: The letter *a* with four ballistic strokes. The velocity profile has seven minima indicating eight strokes, while the proposed curvature profile computation gives three maxima indicating four ballistic strokes.

temporal. Spatial error is introduced due of limits of spatial resolution of the digitizer and noise within the sensing hardware. The temporal sampling error is caused by delays in acquiring a sample due when the recording hardware is busy, and is especially detrimental to velocity based stroke detection as seen from Figure 1 or in the computation of dynamic features for recognition.

In this paper, we develop a robust method for segmentation of ballistic strokes. The detected strokes can be used for recognition of handwritten characters as well as other applications. We also propose a fixed length representation of a trace from the constituent ballistic strokes, which we refer to as Bag-of-strokes as explained later. Through the rest of the paper, we refer to as trace, the handwriting that has linguistic sense and correctness. The terms ballistic stroke and stroke are used interchangeably.

Various representations of online handwriting have been investigated. The most famous and effective representations have been found to be re-sampling of points, with equi-distant resamples performing better compared to equi-time resampling, random re-sampling[1]. Freeman codes have also been investigated to be a effective representation of online handwriting. Willems *et al.*[2] have surveyed various features and have categorized them into spatial, dynamic and force based features depending on their domain of inquiry. Stroke segmentation is vital for various tasks such as handwriting modeling, writer identification. Human handwriting modeling based on dynamics of hand movement was popularized by Plamondon [3], where he proposed

1520-5363/13 $26.00 © 2013 IEEE
DOI 10.1109/ICDAR.2013.175

that the handwriting output is the result of an agonist and antagonist systems. Several oscillatory models were also proposed (Hollerbach [4], Gangadhar [5]) for the purpose of generation of cursive handwritten characters. In the context of kinematic models of handwriting, Sigma-lognormal model (O'Reilly [6]) describes speed of handwriting to be of amplitude scaled log-normal with each stroke spatially being a part of a circle and Beta-elliptic model (Bezine [7]), which describes speed as beta-function and the spatial trajectory as a part of an ellipse, are important. Both the models describe the basic unit of handwriting called stroke defined as one movement of hand, which has an asymmetric Gaussian speed curve with a chance of inversion in direction at the starting or ending of the stroke. The handwriting generative models have been able to describe complex movements such as signatures well. The Beta-Elliptic model has also been applied to handwriting recognition using genetic algorithms and Fuzzy-SOM (Kherallah [8],[9]).

All of the above work assume that the data obtained from the digitizer is clean and can be segmented easily using the velocity profile of the pen/finger trace. With increasing use of mobile devices for recognition, there is a trade-off in the quality of the data being dealt with. Input data to PDAs and tablets are affected by sensor limitations, environmental conditions such as humidity,temperature and EMI from various components. Hence the problem of robust segmentation of ballistic strokes is complementary to the existing works on modeling of hand movement, where segmented strokes are used as basis to estimate the model. We now look into the details of the segmentation process in Section II, before developing a representation for traces suitable for recognition (see Section III).

II. Curvature based stroke segmentation

Traditional method for ballistic stroke segmentation divides a trace at instants of time when speed profile reaches a local minimum [10]. While this method is ideal for non-noisy data, it falters in presence of noise in data capture. There are two major types of distortions induced during data capture. The spatial error described in section I can be mitigated to a large extent by smoothing and parametric curve fitting such as splines due to the smooth nature of hand movements [11]. Due to temporal noise distortions calculation of velocity becomes erroneous. Most devices specify a temporal frequency of data capture. However, the actual instances of recording of pen tip usually varies slightly from it and the interval between such recordings does not remain constant.

In online handwriting, the trace of a pen or finger is recorded as the sequence: $(x_t, y_t), 0 \leq t \leq N$. The instantaneous acceleration of the pen tip is proportional to the vectorial sum of the force applied by the hand and the frictional force. Since the force applied on the pen by the fingers is finite and bounded, the third derivatives functions x_t & y_t should exist. Based on these, we consider cubic splines to represent and model handwriting. Specifically, we choose a basis spline representation, where the order of the basis function is chosen to be 3. Splines impose the smoothness constraint, which minimizes the integral of the squared second derivative along the curve [12], i.e.,

$$\int S''(t)^2 dt \leq \int G''(t)^2 dt \tag{1}$$

where $S(t)$ is the spline function, $G(t)$ is the set of all functions that have continuous second derivatives.

The *two-thirds power law* [13] states an inverse cubic relationship between tangential hand-speed and curvature of trajectory of motion. This inverse relationship allows us to use the maxima in curvature of a trace to segment the ballistic strokes instead of minima of velocity.

Curvature κ at point on a curve, is the measure of rate of change of the tangent at that point to neighboring points. In arc length parameterization, curvature may be defined as [14] the rate of change of tangential angle with arc length.

$$\kappa = \frac{d\phi}{ds} \tag{2}$$

After spline fitting to handwriting signals, it can be assumed that there is no spatial sensing noise that deviates from the handwriting curve. Spline evaluation at integral units of time can be understood as dropping the noisy point to the point closest on the spline curve. Ideally, that might not be the correct point given that time instant. In other words, now there is a deviation from the ideal point along the spline curve. This cannot be traditionally termed as error in spatial sampling but as an error in time sampling i.e., sampling not being done at uniform intervals of time.

Lemma 2.1: If the cubic spline fit of the handwriting trace retains only spatial errors along the curve in addition to temporal errors, the computed curvature values and hence the ballistic stroke segmentation based on it is independent of spatial and temporal noise.

Proof: As noted before, any error in the spatial location of the pen tip is indistinguishable from an error in the time instant of sampling of the curve. Hence we can assume that after the spline fit, we have a curve with correct spatial trajectory, which is probably erroneous in the temporal points of sampling. If we use Equation 2, which is independent of time as the definition of curvature, the resulting value of curvature would be accurate. ■

III. Circle based representation of strokes

A ballistic stroke, spatially, can be described as a pivotal movement of the hand along the arc of a circle [6]. Each basic unit i.e., a stroke is hypothesized to be an arc of a circle. In order to spatially characterize a stroke, the center, radius of the circle, the starting and ending points of the arc are required. To each of the strokes calculated from the

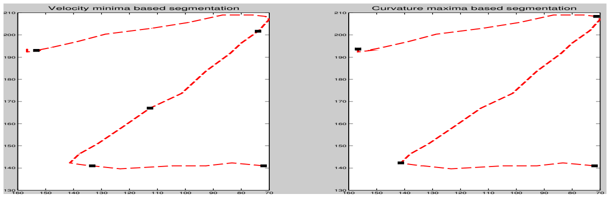

Figure 2: The character z segmented by velocity minima and curvature maxima. Note that the spurious stroke boundary is removed and the location of boundaries are corrected by the curvature based segmentation.

curvature profile, a circle is fit by minimizing the sum of squared radial deviations.

$$\min_{x_0,y_0,r} \sum_{i}^{m} x_i^2 + y_i^2 - 2x_0x_i - 2y_0y_i + x_0^2 + y_0^2 + r^2$$

Let $-2x_0 = a_1, -2y_0 = a_2, x_0^2 + y_0^2 + r^2 = a_3$. For $i = 1 \dots m$, the above equation can be represented in matrix form as,

$$\begin{pmatrix} X & Y & \mathbf{1} \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}^T = -\left(X \circ X + Y \circ Y\right) \quad (3)$$

where $A \circ B$ denotes the Hadamard product of two matrices A & B, X is a column vector of x co-ordinates and Y is a column vector of y co-ordinates of the points to which a circle has to be fit, $\mathbf{1}$ is a column vector of 1 of length m. The least-squares fit to the system of 3 can be solved by calculating the Moore-Penrose pseudoinverse. The center and the radius of the circle can be computed from 3 by

$$x_0 = -\frac{a_1}{2}, y_0 = -\frac{a_2}{2}, r = \sqrt{\frac{a_1^2 + a_2^2}{4} - a_3}.$$

An equivalent of describing the circular arc with the starting and ending points is taking the angles of the lines joining the center of the circle and two ends of the stroke with x-axis. Thus the representation $(x_0, y_0, r, \theta_s, \theta_e)$ describes completely a stroke in space, where θ_s, θ_e denote the angles at the starting and ending points of the stroke.

Each trace is normalized in space to fit into a unit square. The sensitivity of the circle parameters (r, x_0, y_0) to curvature of the arc being fit is high. *i.e.*, the values of (r, x_0, y_0) become exponentially large as the the arc approaches a straight line. In order to mitigate this problem, a monotonic function (say hyperbolic tangent) is used to map the values from $\mathbb{R}$ to $[0, 1]$. Also, it is intuitive that the variation of the mean of a stroke is less compared to the variations in the center of the fitted circle. So we substitute the mean coordinates of the ballistic stroke in the place of center (x_0, y_0), which we will continue to denote the mean coordinate rather than the center of the fitted circle. The hyperbolic function $\tanh$ is applied only to the radius dimension, as it is the only unbounded ordinate in the representation. Thus the problem that arise during fitting a circle to a straight line is mitigated by using the center of the line and the radius will only tend to 1.

To apply the above model for character recognition, for all the characters in the training set, the strokes are found out and their circle fits are calculated. The circular features *i.e.*, $(x_0, y_0, r, \theta_s, \theta_e)$ of all the strokes are clustered using k-means into k clusters and the centroids of the clusters are stored. Each character is represented as a Bag-of-Words (BoW) model, with the words defined as these centroids.

The Bag-of-Words(BoW) model is a representation that has traditionally been used in text document retrieval. It uses an unordered set of words, which are pre-determined to best represent the content of the document and the documents are represented as a histogram whose bins are given these set of *words*, ignoring the semantics, grammar and word order. This representation has been extended to images as Bag-of-Visual-Words (BoVW), where the equivalent to words are the interest points detected in the images. They have been applied problems like texture recognition, object categorization, natural scene categorization[15].

The feature representation for each character is a k-bit feature vector corresponding to k centroids, which is initialized to zero. For each character's strokes, the circular representation is found. The distance to each of the centroids is calculated and the bit corresponding to the centroid that is closest to the given stroke is set. Thus the entire character is encoded as a string of k-bits. Our representation differs from the traditional Bag-of-Words in such a way that indicator functionality is described rather than a histogram, and we refer to the representation as Bag-of-Strokes. An example of this process for the Malayalam vowel symbol *au* is illustrated in 3.

IV. Experimental Results and Analysis

We report results on three different representations of the traces. The popular traditional representations are equitime samples and equidistant samples [1]. It has been shown, experimentally, that equidistant resampling has a better discriminative feature space than equitime resampling [16], [1]. One of the problems of using equidistant resampling for recognition of Indic languages is that it, on certain occasions, minor discriminative details such as Cusps *vs* small loops are lost. Most of these smaller details happen at areas of high curvature. Thus we also experiment with an alternative method for sampling that is based on curvature.

The curvature is calculated from the spline representation as

$$\kappa(t) = \left| \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{\sqrt{(\dot{x}^2 + \dot{y}^2)^3}} \right|, \quad (4)$$

where $x(t)$ & $y(t)$ are the x and y trajectories with time. The cumulative curvature is calculated

$$h(t) = \int_0^t \kappa(t)\, dt$$

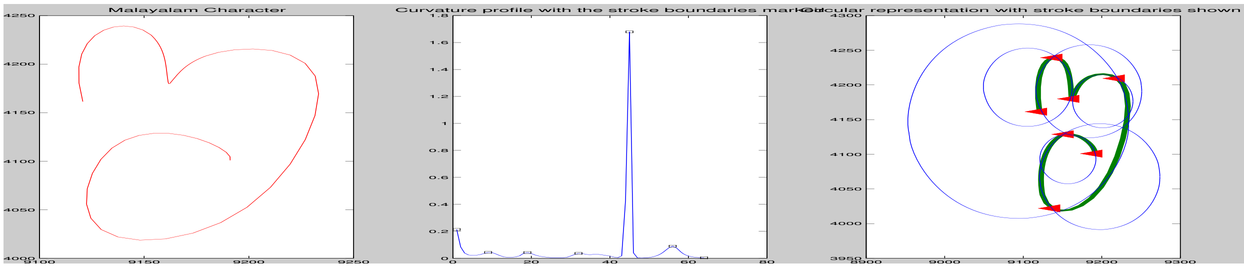

Figure 3: Example of segmentation and representation

The inverse of $h(t)$ is computed and is evaluated at linearly spaced points in the range $[0, max\{h(t)\}]$. This gives the time-intervals $\{t_i\}$ at which the handwriting has to be resampled. The original handwriting is then interpolated at $\{t_i\}$ to get a resampled version in accordance with the curvature of handwriting. We found that, curvature weighted sampling when clubbed with equidistant resampling performs better than either of them individually (See Table I).

A. *Malayalam Dataset*

Malayalam is a Dravidian language that is predominantly spoken in the state of Kerala. We have collected a significant amount of natural handwritten data for Malayalam. Modern Malayalam script has 13 vowels, 36 consonants, and 5 half-consonants. In addition to these, there are several symbols for multiple consonant combinations. We have included the ones currently used in our dataset, ignoring the ones that have fallen out of usance. Unique traces in the data were identified and the complete dataset was labelled accordingly. Our Malayalam dataset contains 106 different traces or classes to be identified, many of them are very similar in shape to others. The actual data was collected as a set of words that were chosen to cover all the trace classes and the set of words were written by over 100 writers. After removal of characters that were incorrectly written, we had 8966 traces in our final dataset. The data was collected using Genius G-Note 7000 digital ink pad.

The optimal number of clusters for Bag-of-Strokes (BoS) was experimentally determined to be $k = 150$. Using the BoS representation, we trained a Support-Vector machine classifier (LIBSVM [17] implementation) with RBF kernel ($\sigma = 0.2$). All experimental results are reported as averages in 3-fold cross validation trials. We are able to achieve an accuracy of 94.55% with BoS representation as compared to 85.75% using the representation proposed by Arora *et al.* [16]. Concatenating the equidistant resampling and the BoS representation results in a further improvement in accuracy (97.7%) with a linear kernel in SVM, showing that this method adds information complementary to the one given by the resampling.

B. *UJI Penchars*

To test the proposed stroke segmentation and classification, a lower case character subset of publicly available UJIPenchars2 [18] is used. The classification task is of 26 classes. Each class on an average has 120 samples and total number of samples used is about 3116, with some of them being removed as they were not recognizable as corresponding to the class specified.

Using the Bag-of-Words representation, we trained a Support Vector Machine (SVM) classifier. We used the freely available software LIBSVM [17] for classifier design. We performed 4-fold cross validation test on the data. We achieved a peak accuracy of 95.8% using 140 words and RBF kernel with $\sigma = 0.2$, which is higher than [19] method of using approximate DTW with multi-layer perceptron who report an accuracy of 91.8%.

C. *Cross-language recognition*

The ballistic stroke is considered to be the most basic unit of handwriting. Any complex movement can be broken down into a set of ballistic strokes. These basic units can be comprehended to be language independent. In-order to test this hypothesis, we used the clusters that were generated in the optimum case of Malayalam data and used those to define the words for the UJIPenchars dataset. The resultant accuracy of such a representation is 95.0%. The apparent 1% decrease in the percentage accuracy is because of the shape characteristics of Malayalam which has lot more complex structures than English does and hence the distribution of centroids is affected by that.

D. *Data from Capacitive sensing device*

Capacitive sensors are inherently more noisy and less accurate than resistive sensing devices technology [20]. The various factors that affect the accuracy of sensing are input current noise of amplifier circuits, shot noise effect, excess noise of semiconductor junctions, thermal in resistors apart from the electromagnetic interference. In-order to verify the robustness of our algorithm, recognition experiments were ran on small handwriting dataset collected from Google Nexus 7 tablet and a Samsung Galaxy SII mobile phone. 5 participants took part in the data collection for 26 lower case English alphabets, with each of the participants writing each character at-least 10 times. The total number of characters in the database is 1380, giving an average of 53 samples per class.

	Equidistant Sampling(ED)	Curvature weighted sampling(CS)	ED+CS	Bag of Strokes representation	ED+CS+BoS
Malayalam	84.40	81.75	85.76	94.55	97.75
UJIPenchars	82.51	76.05	86.70	95.8	96.5
Touch Screen data	95	94.5	95.58	93.9	96.2

Table I: A table showing percentage accuracies of recognition of various datasets with several feature extraction techniques

As seen from Table I, the proposed method performs comparably to equidistant resampling, and provides an improvement over velocity based stroke segmentation, which gives an accuracy of 91.9% on the same dataset (compared to 93.9%) . The result shows the robustness of the proposed segmentation approach to noise in the data. Moreover, the information in the representation complements resampling based methods and the combined accuracy is even higher.

V. Conclusion

In this paper, we presented an alternative method for ballistic stroke segmentation that uses curvature at any given point by using a parametric representation of handwriting. We observed that this method of stroke segmentation is more robust to sensing and digitizing noise. We proposed a circle based representation using the bag of words technique, where the entire character was represented as a bit vector. The resulting representation is both efficient and effective for character recognition on mobile devices with limited processing and memory capacity. We used the UJIPenchars dataset and a large-class Malayalam dataset to evaluate our algorithm and to show its satisfactory performance. We are currently working on applying the segmentation technique and the representation developed to related problems such as writer identification and script recognition.

Acknowledgment: We thank the Technology Development in Indian Languages (TDIL) initiative of the Department of Electronics and Information Technology (DeitY), India for the financial support of this work.

References

[1] R. H. Kassel, "A comparison of approaches to on-line handwritten character recognition," Ph.D. dissertation, Massachusetts Institute of Technology, 1995.

[2] D. Willems and R. Niels, "Definitions for features used in online pen gesture recognition," *Nijmegen Institute for Cognition and Information Radboud University Nijmegen, Nijmegen, The Netherlands*, 2008.

[3] R. Plamondon, "A kinematic theory of rapid human movements," *Biological Cybernetics*, vol. 72, no. 4, Mar. 1995.

[4] J. Hollerbach, "An oscillation theory of handwriting," *Biological Cybernetics*, vol. 39, no. 2, Mar. 1981.

[5] G. Gangadhar, D. Joseph, and V. Chakravarthy, "An oscillatory neuromotor model of handwriting generation," *IJDAR*, vol. 10, no. 2, 2007.

[6] C. OReilly and R. Plamondon, "Development of a sigmalognormal representation for on-line signatures," *Pattern Recognition*, vol. 42, no. 12, pp. 3324–3337, Dec. 2009.

[7] H. Bezine, A. Alimi, and N. Sherkat, "Generation and analysis of handwriting script with the beta-elliptic model," in *Proceedings of the Ninth IWFHR*, Tokyo, Japan, Oct. 2004.

[8] M. Kherallah, L. Haddad, A. M. Alimi, and A. Mitiche, "On-line handwritten digit recognition based on trajectory and velocity modeling," *Pattern Recognition Letters*, vol. 29, no. 5, Apr. 2008.

[9] M. Kherallah, F. Bouri, and A. M. Alimi, "On-line arabic handwriting recognition system based on visual encoding and genetic algorithm," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 1, Feb. 2009.

[10] H. Bunke and P. Wang, *Handbook Of Character Recognition And Document Image Analysis*. Hackensack, NJ: World Scientific, 1997.

[11] A. M. Namboodiri and A. K. Jain, "Robust segmentation of unconstrained online handwritten documents." ICVGIP, 2004.

[12] C. De Boor, "A practical guide to splines," 2001.

[13] F. Lacquaniti, C. Terzuolo, and P. Viviani, "The law relating the kinematic and figural aspects of drawing movements," *Acta Psychologica*, vol. 54, no. 1-3, Oct. 1983.

[14] E. Kreyszig, *Differential Geometry*. Dover Books, 1991, ch. Principal Normal, Curvature, Osculating Circle.

[15] R. Szeliski, *Computer vision: algorithms and applications*. Springer, 2010.

[16] A. Arora and A. Namboodiri, "A hybrid model for recognition of online handwriting in indian scripts," in *ICFHR*. IEEE, 2010.

[17] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2011.

[18] D. Llorens *et al.*, "The UJIpenchars database: A pen-based database of isolated handwritten characters," in *Proceedings of the Sixth ILRE*, Marrakech, Morocco, May 2008.

[19] M. Castro-Bleda *et al.*, "Improving a DTW-based recognition engine for on-line handwritten characters by using MLPs," in *Proceedings of the ICDAR*, Barcelona, Spain, Jul. 2009.

[20] L. K. Baxter, *Capacitive Sensors: Design and Applications*. Piscataway, NJ: IEEE Press, 1996.