Semi-automatic determination of allograph duration and position in on-line handwritten words based on the expected number of strokes

Poster presentation at the IWFHR-5 in Colchester 1996

Janek Mackowiak, Lambert Schomaker, Louis Vuurpijl
Nijmegen Institute for Cognition and Information (NICI)
P.O. Box 9104, NL 6500 HE, Nijmegen, The Netherlands
Tel: +31-80-615981, Fax: +31-80-615938 E-mail: mackowiak@nici.kun.nl

1 Introduction

A semi-automatic character labeling ('truthing') procedure is presented, which uses an initial un-supervised algorithm to estimate the starting stroke and number of velocity-based strokes **(VBS)** per allograph, followed by minimal user interaction to improve the estimates. In this study, a stroke is defined as the trajectory of the pen tip between two consecutive minima in the absolute pen-tip velocity [3,4,5,6], see Figure 1). Such a procedure is useful since the complete manual labeling of handwriting takes a lot of time, in the order of one hour per 100 written words. Basically, the question in this study is: How far can one get in estimating character locations, starting with only the word label and the XY(t) coordinates, while using velocity-based segmentation into strokes.

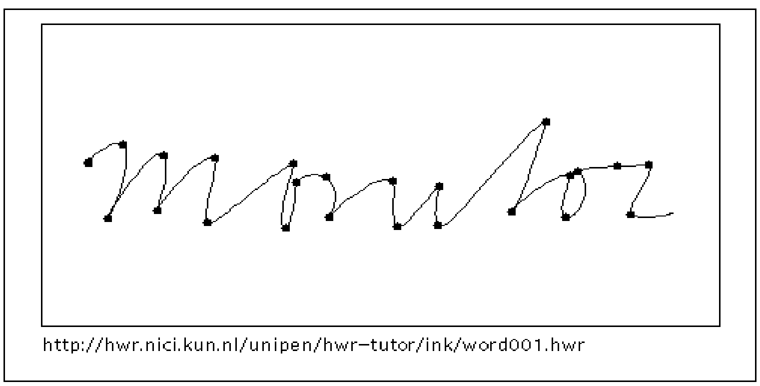

Figure 1. Division of a handwritten word into strokes

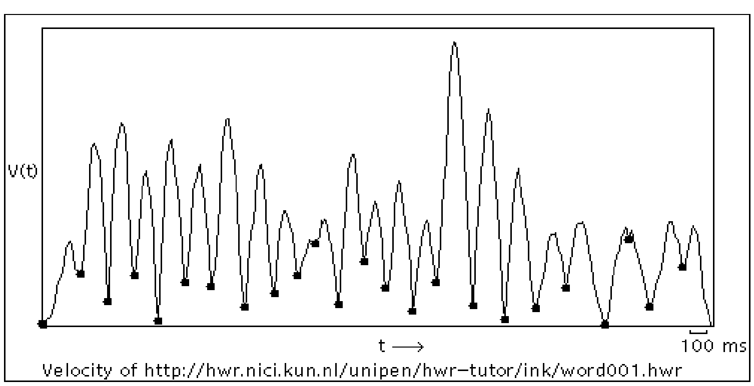

Figure 2. Minima in the velocity of the same word

A basic problem in on-line handwriting recognition is the variation in handwriting styles (see Figure 3 and 4). Some [2] believe that the ultimate goal is the writer-independent handwriting recognition system, we think that a system which is able to learn the peculiarities of the handwriting style of a given writer will have a much higher performance asymptote.

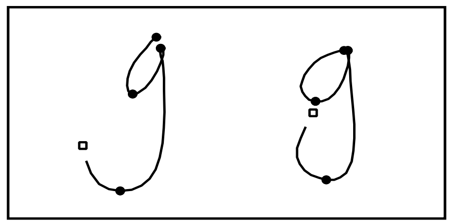

Figure 3. Two instances of a four-stroked <g> allograph

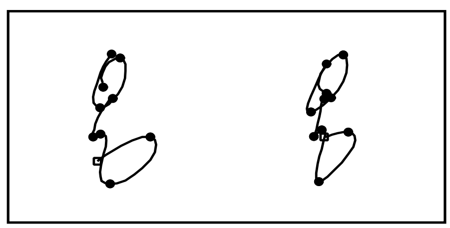

Figure 4. Two instances of a nine-stroked <g> allograph

The drawbacks of manual labeling are that, apart from the amount of user time required, the user must apply systematic and consistent criteria in segmenting and labeling characters (Figure 5)

Figure 5. An example of the difficulties that may arise in the manual labeling of a word: Guess the word, and the location of its characters

2 The problem

How can we predict or estimate the location and duration of allographs in a handwritten word, using **minimal advance knowledge** of allograph shape.

In an earlier study [1], a solution for the problem was coined, on the basis of simultaneously solving a set of linear equations. This latter idea is pursued in the current study.

---

3 Method

$$\tilde{N}_w - (C_w - 1) = \beta_0 + \sum_{i=1}^{26} \beta_i n_{iw} + \varepsilon \qquad (1)$$

- $\tilde{N}_w$ = the estimated number of strokes in a word
- n_{iw} = the number of instances of each letter of the alphabet in this word (may be zero)
- β_i = the estimated number of strokes of a letter i
- $C_w - 1$: we (heuristically) assume that the number of connecting strokes between letters in a handwritten word equals the number of characters minus one. A pen-lift counts as a single VBS. Of course this is simplistic: small ligatures at the beginning and end of isolated characters are omitted (but we'll see how far we get)
- β_0 = an indication of the number of connecting strokes this writer produces *in excess* of the expected strokes between characters $C_w - 1$

The model is solved by using the known number of strokes N_w as output, and the number of letters n_{iw} used from the alphabet as the input, in order to find the coefficients β of the model by linear regression (regress in Unix `|Stat`). The vector $\vec{\beta}$ can be dubbed as *character-duration vector*.

For this study we used earlier recorded handwriting data of 92 writers (male and female) from the US, who all wrote 500 lowercase words. This amounts to 46000 handwritten words.

actual Nstr	word	letters
33	agreement	1.a, 3.e, 1.g, 1.r, 1.m, 1.n, 1.t
7	bell	1.b, 1.e, 2.l
11	blade	1.a, 1.b, 1.d, 1.e, 1.l
22	brain	1.a, 1.b, 1.i, 1.n, 1.r
.	.	.
.	.	.
24	young	1.g, 1.n, 1.o, 1.u, 1.y

Table 1: Example of the input data format for the stroke estimation procedure

4 Results

Results after application of the model on all data, separately for each writer, yielding writer-specific character-duration vectors.

error	Mean	sd	Min	Max
ϵ	-0.024	0.52	-1.73	1.47
$\|\epsilon\|$	1.31	0.25	0.48	2.09
σ_ϵ	1.76		0.85	3.06

Table 2: Errors of the character-duration estimate model, expressed in average number of strokes per word (N = 92 writers, 500 words per model estimation)

If we take the word *thunder* in Figure 6 as an example, we can clearly see the flaws of our estimation procedure. Because of forward propagation of errors in the estimation of the allograph duration (in strokes) the location of for example the allograph <e> is not correctly estimated.

5 Conclusion

- The procedure yields reasonable results with minimal knowledge about allograph shapes.
- Not all problems are solved with this procedure.
- Knowledge about ligatures is not taken into account in the equation.

If we **do** use knowledge about ligatures, it is expected that the semi-automatic determination of the location and duration of allographs in handwritten words yields better results. A second experiment implemented the ligature knowledge by using a matrix M_{uv} of ligatures containing information about the expected number of strokes between letters u and v in the alphabet. This gives the following equation:

$$\tilde{N}_w = \beta_0 + \sum_{i=1}^{26} \beta_i n_{iw} + \sum_{j=1}^{m-1} M_{\lambda_w(j),\lambda_w(j+1)} + \varepsilon \quad (2)$$

where $\lambda_w(j)$ is letter j from the word and m is the actual number of letters in the word.

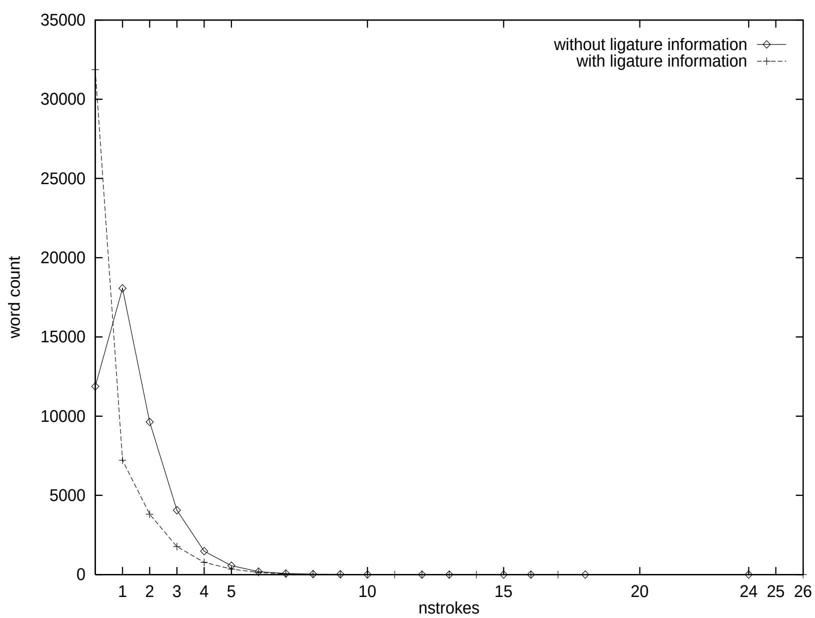

Figure 6. This graph shows the histogram of $|\epsilon|$ in number of strokes and the number of times this error occurs in the 46000 words for both models (eq 1 and eq 2)

Figure 6 shows $|\epsilon|$ for the basic model, and for the improved model. Indeed, the improved model shows a better performance, at the cost however, of substantial advance knowledge (M_{jk}). we can see that the mean number of errors in the estimation procedure without ligature information is 1,31 stroke absolute difference between actual and predicted number of strokes per word. For the procedure in which we use ligature information the mean is 0,73. The area under both curves should be the same.

Figure 7. A picture of the result of the estimation procedure (eq 1) on words containing allographs <d>

6 References

[1] Teulings, H.-L., & Schomaker, L.R.B. (1992). Un-supervised learning of prototype allographs in cursive script recognition using invariant handwriting features. In J.-C. Simon & S. Impedovo (Ed.), *From Pixels to Features III* (pp. 61-73), Amsterdam: North-Holland.

[2] Manke, S., Finke, M., & Waibel, A. (1995). *NPen*$^{++}$: A Writer Independent, Large Vocabulary On-Line Cursive Handwriting Recognition System. In Kavanaugh, M., & Storms, P. (Eds.), *Proceedings of the Third International Conference on Document Analysis and Recognition* (pp. 403-408), Los Alamitos: IEEE Computer Society Press.

[3] Teulings, H.-L., Schomaker, L.R.B., Morasso, P., & Thomassen, A.J.W.M. (1987). Handwriting-analysis system. In Plamondon, R., Suen, C.Y., Deschênes, J.-G., & Poulin, G. (Eds.), *Proceedings of the Third International Symposium on Handwriting and Computer Applications* (pp. 181-183), Montreal: Ecole Polytechnique.

[4] Morasso, P. (1986). Human movement Understanding. In Morasso, P., & Tagliasco, V. (Eds.), *Trajectory formation*, Amsterdam: North-Holland.

[5] Plamondon, R., & Maarse, F.J. (1989). An evaluation of motor models of handwriting, *IEEE Trans. Syst. Man Cybern*, **19**, 1060-1072.

[6] Schomaker, L.R.B. (1991). Simulation and recognition of handwriting movements: A vertical approach to modeling human motor behavior, *Doctoral Dissertation*, University of Nijmegen, NICI, The Netherlands.