SIZE NORMALIZATION IN ON-LINE UNCONSTRAINED HANDWRITING RECOGNITION

Homayoon S.M. Beigi, Krishna Nathan, Gregory J. Clary, and Jayashree Subrahmonia

T.J. Watson Research Center, IBM
P.O. Box 704
Yorktown Heights, New York 10598
EMail: beigi@watson.ibm.com

Keywords: Handwriting, Recognition, Preprocessing, Deslanting, Desloping, Size Normalization

ABSTRACT

In an on-line handwriting recognition system, the motion of the tip of the stylus (pen) is sampled at equal time intervals using a digitizer tablet and the sampled points are passed to a computer which performs the handwriting recognition. In most cases, the basic recognition algorithm performs best for a nominal size of writing as well as a standard orientation (normally horizontal) and a nominal slant (normally fully upright). Here, we will discuss and provide solutions to these normalization problems in the context of on-line handwriting recognition. Most of the results presented here are also valid for Optical Character Recognition (OCR). Error rate reductions of 54.3% and 35.8% were obtained for the writer-dependent and writer-independent samples through using the following normalization scheme.

1. INTRODUCTION

In an on-line handwriting recognition system, the motion of the tip of the stylus (pen) is sampled at equal time intervals using a digitizer tablet and the sampled points are passed to a computer which performs the handwriting recognition. In most systems, the data signal undergoes some filtering process. The signal is often normalized to a standard size and its slant and slope is corrected. After normalization, the writing is usually segmented into basic units and each segment is classified and labeled. Using a search algorithm in the context of a language model, the most likely path is then returned to the user as the intended string (see Figure 1).

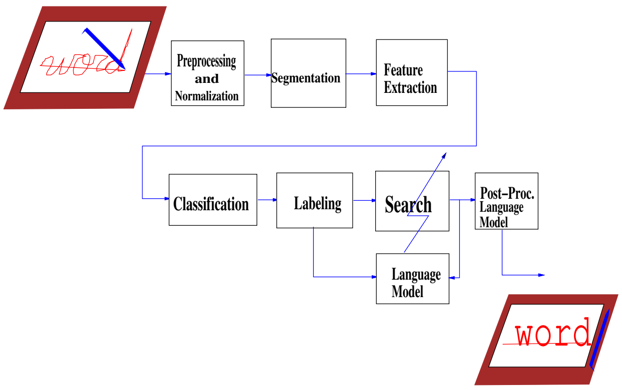

Figure 1: Generic Handwriting Recognition Process

Here, we treat size normalization of the handwriting signal in the framework of an unconstrained cursive recognition system. In such a system, the writer is allowed to connect any set of characters and is basically asked to write in his/her natural writing style. See references [1], [2] and [3] for further details on the different methods of writing and the specific recognition algorithm used in this setup.

In most cases, the basic recognition algorithm performs best for a predefined nominal size of writing. Therefore, after some basic filtering is done on the handwriting signal, it is usually desirable to scale the writing to a standard height such that the overall recognition becomes size independent. Actually, it would also be desirable to normalize the length of the writing as well (the average width of each character); this is a much harder problem and it will not be addressed here. In addition to the size of the writing, the orientation and slant of the writing should also be corrected to standard values. This is specially important when free-style cursive writing is used, say as a gesture. See [4] for the definition of a gesture. Gestures (see figure 2) are usually written with a slope (non-horizontal orientation).

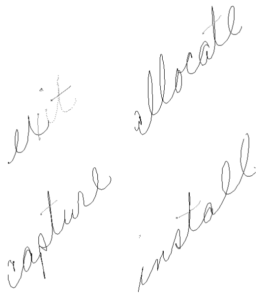

Figure 2: Example of Words in Need of Desloping

2. SIZE NORMALIZATION

To perform a size normalization, the base-line (line 1) and the mid-line (line 2) need to be estimated (see Figure 3). The area surrounded by the the base-line and the mid-line is the only part of any word which is always non-empty. This makes this area the most reliable portion of the data for usage in size normalization. Once accurate estimates of the base-line and the mid-line are obtained, a magnification factor can be computed from the ratio of the nominal mid-portion size and that of the input. The entire input data may then be scaled using the obtained magnification factor.

Furthermore, it is important to use the information provided by line 0 and line 3 of figure 3. These lines are specially useful in cases where the whole word is either totally made up of upper-case letters or lower-case letters with no ascenders or descenders – such as $\{a, c, e, o, \cdots\}$. The relative positions of line 0 versus line 1 and line 2 versus line 3, inform us of the presence of descenders or ascenders respectively. Special treatment should be employed for different cases to achieve a high accuracy.

2.1. Principal Line Estimation

Figure 3 shows the words "principal lines" written in English and Farsi (Persian). Both these scripts and those of similarly written languages (German, French, Spanish, Italian, ..., and Arabic, Ordu, ...) possess similar features related to the method of size normalization which is discussed in this paper. Therefore, all the techniques discussed here are readily applicable to other similar languages. However, these techniques are not applicable to the scripts of some other languages

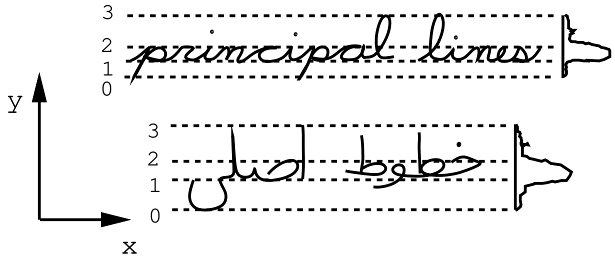

Figure 3: Principal lines of a word

such as Japanese, Chinese, Korean, Hindi, etc. See figure 4. For these languages, the entire height of the writing is basically fixed and the mean total height of words may be used to determine the normalization factor.

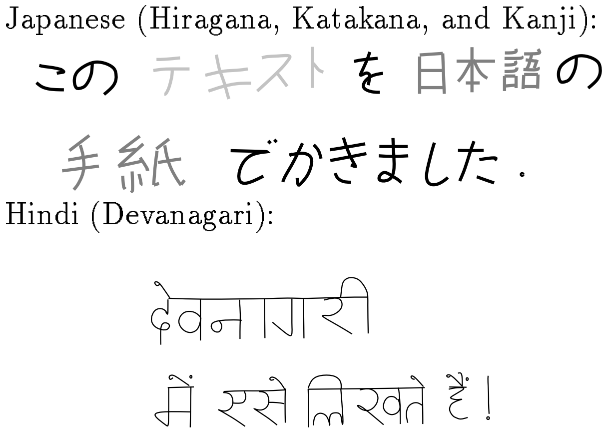

Figure 4: Samples of Japanese and Hindi Writing

Due to the nature of techniques used here, the basic normalization scheme is also applicable to OCR (Optical Character Recognition) without any changes. The basic idea behind the normalization method presented here is the use of the pixel count histogram when projected onto the $y - axis$ of the word (see Figure 3).

First, a Slope correction scheme is used to align the writing with the horizontal axis. This is necessary to stabilize the principal line estimation based on the histogram. Figure 6 shows the words of 2 after desloping was done on them. In the new horizontal orientation, more accurate estimates of the principal lines may be obtained. To compute these lines, the histogram value and slope at each point in the $y-axis$ are noted. Based on thresholds in the percentage of increase and decrease of the histogram value as well as the percentage of increase and decrease in the histogram slope, lines 1 and

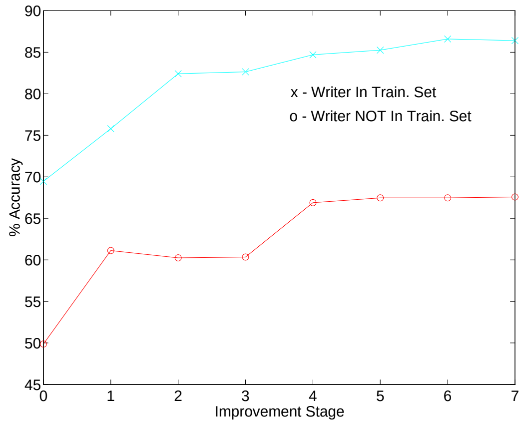

Figure 5: Accuracy Improvement Through Different Algorithm Improvement Stages

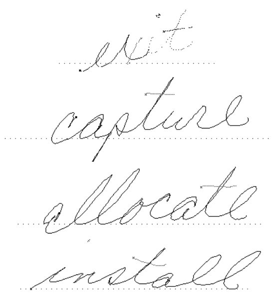

Figure 6: Example of Desloped Words

2 of figure 3 are located.

Several techniques are used to enhance the estimation of the optimal magnification factor based on the information which is obtained about the location of the four principal lines. Figure 5 shows the word accuracy of the unconstrained cursive handwriting recognition system as additional features are added to the normalization algorithm. Stage 0 shows the word recognition accuracy for the system without any normalization and Stage 1 is the simple application of the information obtained from the principal lines to magnify the writing.

The two curves shown in Figure 5 are the accuracy results for the case when the writers were members of the training set (writer-dependent) and the case when they were not part of the training set (writer-independent). In both cases, the training data was different from the test data.

In stage 2, the average height of the words with all upper-case and all lower case (in absence of ascenders or descenders) were estimated from well formed words with rich principal line information. These are words for which there are optimal number of ascenders and descenders present which makes the lines estimation very accurate. Using these running averages of all upper case and all lower case, a decision may be made about the normalization factor when no ascender or descender is present.

In stage 3 of the improvements, a new adaptive parameter was introduced in the evaluation of the magnification factor. This parameter is the average height of a word. Different writers in different occasions write in different sizes. Based on informative and rich data points, an average height is estimated with a certain forgetting factor.

Stage 4 resulted in major improvement in accuracy through the addition of a new parameter, the ratio of the magnification factors computed using the distance between lines 1 and 2 and the distance between lines 0 and 3. Comparing this ratio with an empirically found threshold, one of the two magnification factors was used.

In stage 5, an adaptive technique for evaluating the threshold of stage 4 was used to further improve the accuracy by making the line estimation more robust.

In stage 6, a maximum value was empirically generated to be used for stabilizing the adaptive evaluation of the parameter in stage 5.

Finally, in stage 7, a maximum value was empirically generated for the magnification factor. This maximum value was generated to avoid special cases which would generate very bad estimates for the magnification factor.

3. SLOPE CORRECTION

The slope correction which was mentioned in the previous section is based on the evaluation of the mean velocities in the x and y directions. The angle between these velocity vectors is used to estimate the angle by which the words should be rotated. This rotation on most occasions re-orients the writing into having a horizontal left to right flow. For special cases such as in words with all block capital letters, special provisions should be taken to avoid mis-estimation of the rotation angle. See figures 2 and 6 for some sample words before and after desloping, respectively.

4. RESULTS AND CONCLUSION

Tables of figures 7 and 8 show the details of the accuracies for the individual writers at stage 0 and stage 7 for writer-dependent and writer-independent sets respectively.

From these tables we can see that significant error rate reductions of 54.3% and 35.8% were obtained for writer-dependent and writer-independent samples through using the discussed normalization scheme. Desloping plays a big role in the general case as a preprocessor to size normalization. In addition, slope correction by itself is an important factor in the segmentation of the characters even if the size is optimal.

Please note that slope correction as a side-effect creates a slant in most cases. In general the angle of the slant created by desloping may be evaluated and the writing may be deslanted through shearing. However, since sloped writing is not naturally formed, different people write with different slants when writing with a slope. This makes the slant correction a bit more difficult since it will be writing style dependent. More sophisticated deslanting algorithms will be necessary to take care of these special cases.

5. REFERENCES

[1] Charles C. Tappert, Ching Y. Suen, and Toru Wakahara, "The State of the Art in On-Line Handwriting Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 8, Aug. 1990, pp. 787-808.

[2] T. Fujisaki, H.S.M. Beigi, C.C. Tappert, M. Ukelson, and C.G. Wolf, "On-line Recognition of Unconstrained Handprinting: a stroke-based system and its evaluation," From Pixels to Features III: Frontiers in Handwriting Recognition, S. Impedovo and J. C. Simon (eds.), Elsevier Publishers, New York, 1992, pp. 297-312.

[3] Tetsu Fujisaki, Krishna Nathan, Wongyu Cho, Homayoon Beigi, "On-line Unconstrained Handwriting Recognition by a Probabilistic Method," Pre-Proc. of IWFHRIII, Buffalo, New York, May 25-27, 1993, pp.235-241.

[4] J. Kim, "On-line Gesture Recognition by Feature Analysis," Proc. of Vision Interface '88, Edmonton, Jun. 6-10, 1988, pp. 51-55.

Writer	Accuracy (No Norm.)	Accuracy (With Norm.)	# of Words
HB	62.23%	84.24%	367
KN	83.77%	87.75%	302
TF	90.19%	86.38%	365
GC	29.44%	86.15%	231
Mean	69.45%	86.04%	316.25

Figure 7: Accuracy Results for Writers in the Training Set

Writer	Accuracy (No Norm.)	Accuracy (With Norm.)	# of Words
PS	61.64%	72.95%	292
MS	35.41%	68.38%	369
GT	53.91%	62.53%	371
Mean	49.48%	67.57%	344

Figure 8: Accuracy Results for Writers NOT in the Training Set