Handwritten Text Normalization by using Local Extrema Classification

J. Gorbe-Moya, S. España-Boquera, F. Zamora-Martínez and M. J. Castro-Bleda

Departamento de Sistemas Informáticos y Computación
Universidad Politécnica de Valencia, Valencia, Spain

Abstract. This paper proposes a method to normalize handwritten lines of text based on classifying a set of local extrema with supervised learning methods. The points classified as lower baseline are used to accurately estimate the slope and the horizontal alignment. A second step computes the reference lines of the slope and slant corrected text in order to normalize the size. Experimental comparison with other well known technique has been performed showing an improvement in the recognition accuracy using HMMs.

1 Introduction

Handwritten text recognition is one of the most active areas of research in computer science and it is comparatively difficult because of the high variability of writing styles. Automatic handwriting recognition systems must include several preprocessing steps for the purpose of reducing variations in the handwritten texts as much as possible.

For off-line handwriting recognition, this preprocessing typically relies on slope and slant correction and normalization of the size of the characters. With the slope correction, the handwritten word is horizontally rotated such that the lower baseline is aligned to the horizontal axis of the image. Slant is the clockwise angle between the vertical direction and the direction of the vertical text strokes. Slant correction transforms the word into an upright position. Ideally, the removal of slope and slant results in a word image independent with respect to such factors. Finally, size normalization tries to make the system invariant to the characters size and to reduce the empty background areas caused by the ascenders and descenders of some letters.

Most of handwriting recognition systems comprise the detection of the different areas of the cursive script: the main body area (area between the upper baseline and the lower baseline), the ascenders, and the descenders (see the image from Figure 1 for an example). These areas can be detected by means of horizontal histogram projection [1–3] or also by obtaining the upper and lower contours of the image [4] after applying the "Run-Length Smoothing Algorithm" [5]. None of these methods track baselines and local extrema accurately, in the sense that they do not classify those points as belonging or not belonging to these baselines. Our approach to image normalization consists in automatically detecting and classifying those local extrema by using neural networks. Some previous work on similar ideas was presented in [6, 7].

Gorbe-Moya J., España-Boquera S., Zamora-Martínez F. and Castro-Bleda M. (2008).
Handwritten Text Normalization by using Local Extrema Classification.
In *Proceedings of the 8th International Workshop on Pattern Recognition in Information Systems*, pages 164-172

2.5 Size Normalization

Size normalization tries to minimize the variations in size and position of the three zones (main body area, ascenders, descenders) which constitute the text line. Furthermore, the normalized size of ascenders and descenders is reduced with respect to the body since they are not as informative (the presence or absence of ascenders and descenders is preserved, also the width, but the actual height is not as important).

After slope and slant correction, the local extrema are computed again using the same method described above, and classified into five classes by using the second MLP. The points belonging to the same class are used to obtain the four reference lines by linear interpolation. These lines comprise the three zones to be normalized. The normalization process is performed for each column of the image by linearly scaling the three zones to a fixed height. Ascenders and descenders are reduced to 20% and 10% of the final image height respectively (see Figure 6).

It should be noted that our normalization technique does not maintain the aspect ratio as other methods [4], but this avoids the problem of size caused by a bad classification of the three areas (see Figure 7).

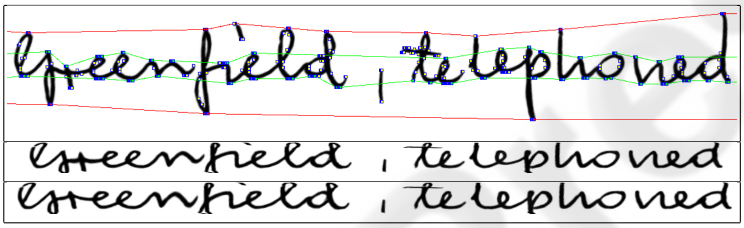

Fig. 6. Image normalization example (from up-to-down): image with slope and slant corrected and local extrema labelled by the MLP; image normalized by using the points labelled by the MLP; image normalized by using the points labelled by a human in order to observe the effects of MLP classification error in the result image.

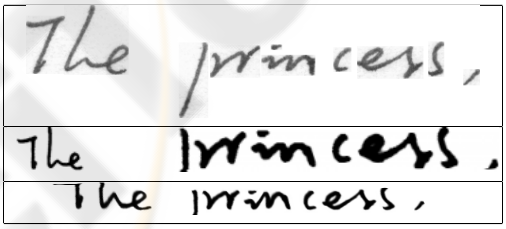

Fig. 7. Comparison of two different normalisation techniques (from up-to-down): The top figure is the original image extracted from IAM database. The middle figure has been normalized using the "second maximum" technique described in [4]. The bottom figure has been normalized with our proposed method. As can be observed, our method does not preserve the aspect ratio but does not distort the entire segment width in case of mistake.

3 Experiments

3.1 IAM Corpus

In order to test the proposed size normalization technique, a handwriting recognition experiment with the version 3.0 of the IAM-database has been conducted. The IAM-database [15, 16] is publicly accessible and freely available upon request for non-commercial research purposes[1]. The corpus is based on the Lancaster/Oslo/Bergen Corpus (LOB) [17]. The version 3.0 of this database consists of 5 685 sentences comprising about 115 000 word instances produced by 657 writers, without restrictions on the writing style or the writing instrument used.

The subset of the IAM-database used in this work consists of 2124 training sentences and 200 test sentences, with a closed vocabulary composed of 8 500 words.

3.2 Image Cleaning

A neural network filter has been trained to estimate the value of a cleaned pixel given a square of 11×11 pixels that was centered at the pixel to be cleaned (see Figure 2). The MLP had two hidden layers of 32 and 16 neurons respectively with the logistic activation function and a unique output with the linear activation function. The error function was the mean square error and the net was trained with the on-line version of backpropagation with momentum term algorithm.

The patterns used to train the net were obtained by mixing IAM-db original noisy images cleaned by hand and artificially noised images. An example of two fragments of image used to train the network are shown in Figure 8 (up and middle). Figure 8 (down) shows an example of an image cleaned with the neural filter.

3.3 Local Extrema Classification

A total of 773 lines of the IAM-db corpus have been manually labelled using a bootstrapping technique: first, a horizontal projection algorithm has been used to classify the points of a subset of images which have been manually corrected using a graphical tool designed to this purpose (see Figure 9); these images have been used to train a MLP to classify the rest of the lines, which have also been manually corrected.

The 773 labelled lines have been used to train two MLPs which classify points transformed into 50×30 patterns as described in Section 2. A total of 723 lines have been used as training data and the remaining 50 as validation data. Table 1 shows some statistics about these sets. Both MLPs use the logistic activation function in the three hidden layers and the softmax in the output layer. The sizes of the hidden layers are 70, 20, 10 for the first network (which has two outputs) and 70, 70, 20 for the second (which has five outputs). They have also been trained using the on-line version of backpropagation with momentum term algorithm.

[1] http://www.iam.unibe.ch/ zimmerma/iamdb/iamdb.html

Table 1. Statistical information about the number of local extrema of each class.

	Training	Validation
lines	723	50
words	5 249	353
points	430 929	29 965
ascenders	6.08 %	6.09 %
upper baseline	22.13 %	21.87 %
lower baseline	36.01 %	35.74 %
descenders	2.22 %	2.61 %
rest	33.56 %	33.68 %

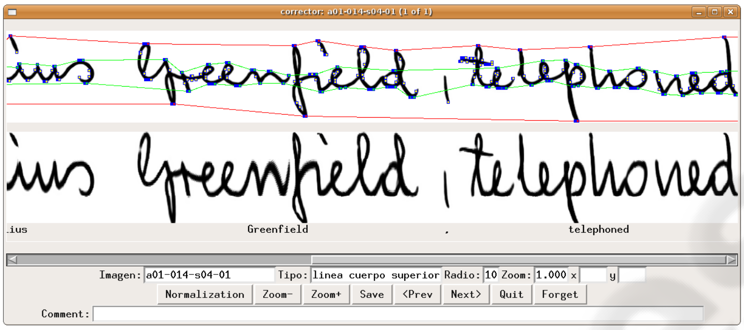

Fig. 9. Graphical tool used to manually supervise the local extrema classification.

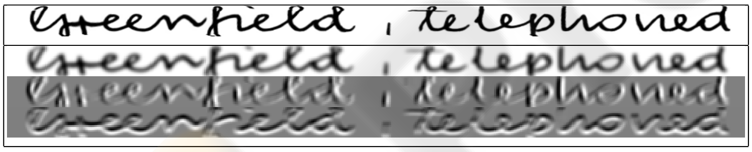

Fig. 10. An example of the graphical representation of the features extracted for the experiments (from up-to-down): preprocessed image, normalized gray level, horizontal gray level derivative, vertical gray level derivative.

in [3, 8] and by the “second maximum” normalization technique described in [4]. This experiment obtained a word error rate (WER) of 22.86%.

The same experimentation has been performed with the preprocessing methods proposed in this work, obtaining a WER of 18.25%, which is significantly better.

4 Conclusions

We have presented a new technique to remove the slope and to normalize handwritten text line images by labeling local extrema.

The proposed method outperforms the baseline experiment, obtaining a roughly 20 percent relative improvement of the WER, showing in this way the practical importance of the preprocessing stage for handwritten text recognition.

Acknowledgements

The authors wish to give special thanks to Moisés Pastor Gadea for his help in providing all the software and the data needed to reproduce his baseline experiment.

This work has been partially supported by the Spanish Ministerio de Educación y Ciencia (TIN2006-12767) and by the BPFI 06/250 scholarship from the Conselleria d'Empresa, Universitat i Ciencia, Generalitat Valenciana.

References

1. Burr, D.J.: A normalizing transform for cursive script recognition. In: Proc. 6th Int. Conf. Pattern Recognition, Munich (1982) 1027–1030
2. Bozinovic, R.M., Srihari, S.N.: Off-line cursive script word recognition. IEEE Trans. on PAMI 11(1) (1989) 68–83
3. Vinciarelli, A., Luettin, J.: A new normalization technique for cursive handwritten words. Pattern Recognition Letters 22(9) (2001) 1043–1050
4. Romero, V., Pastor, M., Toselli, A.H., Vidal, E.: Criteria for handwritten off-line text size normalization. In: Procc. of The Sixth IASTED international Conference on Visualization, Imaging, and Image Processing (VIIP 06), Palma de Mallorca, Spain (2006)
5. K. Y. Wong, R.G.C., Wahl., F.M.: Document Analysis system. IBM Journal of Research and Developement 26(6) (1982) 647–655
6. Hennig, A., Sherkat, N.: Exploiting zoning based on approximating splines in cursive script recognition. Pattern Recognition 35(2) (2002) 445–454
7. Simard, P., Steinkraus, D., Agrawala, M.: Ink normalization and beautification. Document Analysis and Recognition, 2005. Proc. Eighth Int. Conference on (2005) 1182–1187 Vol. 2
8. Pastor, M., Toselli, A., Vidal, E.: Projection profile based algorithm for slant removal. In: Proceedings of the 2004 International Conference on Image Analysis and Recognition (ICIAR04), Porto, (Portugal) (2004)
9. Stubberud, P., Kanai, J., Kalluri, V.: Adaptive Image Restoration of Text Images that Contain Touching or Broken Characters. In: Proc. ICDAR. Volume 2. (1995) 778–781
10. Egmont-Petersen, M., de Ridder, D., Handels, H.: Image processing with neural networks – a review. Pattern Recognition 35(10) (2002) 2279–2301
11. Suzuki, K., Horiba, I., Sugie, N.: Neural Edge Enhancer for Supervised Edge Enhancement from Noisy Images. IEEE Trans. on PAMI 25(12) (2003) 1582–1596
12. Hidalgo, J.L., España, S., Castro, M.J., Pérez, J.A.: Enhancement and cleaning of handwritten data by using neural networks. In: Pattern Recognition and Image Analysis. Volume 3522 of LNCS. Springer-Verlag (2005) 376–383 Proc. IbPRIA 2005.
13. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press (1995)
14. Gonzalez, R., Woods, R.: Digital Image Processing. Addison-Wesley Publishing Co. (1993)
15. Marti, U., Bunke, H.: A full English sentence database for off-line handwriting recognition. In: 5th Proc. ICDAR. , Bangalore (1999)
16. Marti, U.V., Bunke, H.: The IAM-database: an English sentence database for offline handwriting recognition. Int. Journal on Document Analysis and Recognition 5 (2002) 39–46
17. Johansson, S., Atwell, E., Garside, R., Leech, G.: The Tagged LOB Corpus: User's Manual. Norwegian Computing Centre for the Humanities, Bergen, Norway (1986)
18. Toselli, A.H., et al.: Integrated Handwriting Recognition and Interpretation using Finite-State Models. Int. Journal of Pattern Recognition and Artificial Intelligence 18 (4) (2004) 519–539
19. Bunke, H.: Recognition of Cursive Roman Handwriting – Past, Present and Future. In: Proc. ICDAR 2003, Edinburgh, Scotland (2003)