Exploiting Zoning based on Approximating Splines in Cursive Script Recognition

A. Hennig, N. Sherkat

The Nottingham Trent University, Department of Computing

Burton Street, Nottingham NG1 4BU

Tel: +44 (0) 115 9486032

Fax. +44 (0) 115 9486518

Email: ns@doc.ntu.ac.uk

http://www.doc.ntu.ac.uk

Abstract

Because of its complexity, handwriting recognition has to exploit many sources of information to be successful, e.g. the handwriting zones. Variability of zone-lines, however, requires a more flexible representation than traditional horizontal or linear methods. The proposed method therefore employs approximating cubic splines. Using entire lines of text rather than individual words is shown to improve the zoning accuracy, especially for short words. The new method represents an improvement over existing methods in terms of range of applicability, zone-line precision and zoning-classification accuracy. Application to several problems of handwriting recognition is demonstrated and evaluated.

Keywords: Zone-line extraction, handwriting recognition, approximating spline functions, zone prominence, point classification, fuzzy logic

1 Introduction

Recognising a page of cursive handwriting involves many difficult tasks. After lines of text have been identified, letters and words have to be recognised and incorrect alternatives eliminated. As shape, size and orientation of letters and words can vary widely amongst different writers and even within a single scribe, ambiguities arise on many levels. Allowing for style variation, some groups of letters (e.g. 'e' and 'l', 'C' and 'c', 'P' and 'p') can thus appear to be written with almost identical shapes. They differ, however, in size and position relative to the word they form. The information about size, position and orientation of a word or a line of text is referred to as zoning information. Exploiting zoning context provided by neighbouring letters is therefore assumed to contribute towards recognition of single letters, by verifying hypotheses produced by a letter recognizer. Simple features describing hand-written words have often been used to restrict the search in a dictionary of possible words. Such features often require zoning information to identify them reliably (e.g. the presence, number and position of ascenders, descenders or diacritical marks [1]).

Furthermore, many heuristics used in handwriting recognition use empirically determined thresholds that are based on zoning information, e.g. to guess the width of an inter-word gap [2] or the maximum number of letters in a word [3]. Zoning information can also be used to normalise the image of the handwriting before recognition. If zoning information has been extracted incorrectly, however, normalising script introduces irrecoverable errors. Some of these errors could be avoided if the script is left unchanged and zoning is only used for more flexible identification and verification steps - even if only approximate and uncertain.

1.1 Zoning Problems

Ideally, the *zones* of handwriting (here called *middle, upper* and *lower zone*, containing letters like 'a' and 'm', protruding parts of ascenders as in 'd', and descenders as in 'g' respectively) are horizontally aligned. The lines of separation between zones will be called *zone-lines* (the *top-, half-, base-* and *bottom-line*). Real-life script however, often follows an overall upward or downward slope, which is often not constant across the width of the page. This results in tilted or even 'curved' baselines (Fig. 1a and d). The height of zones is rarely constant throughout a single word (e.g. Fig. 1e, 'one'), neither are relative sizes of the outer zones in relation to the middle zone. Many writers use characters of a style that can exhibit 'unexpected' zoning. This can occur at any place in the word but has more frequently been found for the first letter of a word, e.g. the enlarged introductory ligatures in 'c' and 'p' of Fig. 1b. Another writer-induced problem is ambiguous ascenders or descenders, i.e. two or more subsequent ascenders (or descenders), that vary substantially in height (Fig. 1d). This variation can approach or even exceed the height difference between ascenders and elements of the middle zone and thus complicate or disable their discrimination. If the highest middle zone element exceeds the lowest ascender, the two zones can no longer be separated by horizontal lines (Fig. 1a).

Further to the writer-induced problems, zoning estimation can also be negatively affected by the lack of information. Some short words make it difficult to determine their overall orientation as their specific pattern of ascenders and descenders causes them to appear 'tilted' (e.g. 'dog' in Fig. 1c). This is mainly a problem of zoning scope, as surrounding words could provide the missing information and rectify the error. If, however, the 'tilted' word appears alone on a line, e.g. at the end of a paragraph, the problem remains. Information about the outer zone-lines is often missing from the script (e.g. the descender in 'attains' in Fig. 1d if
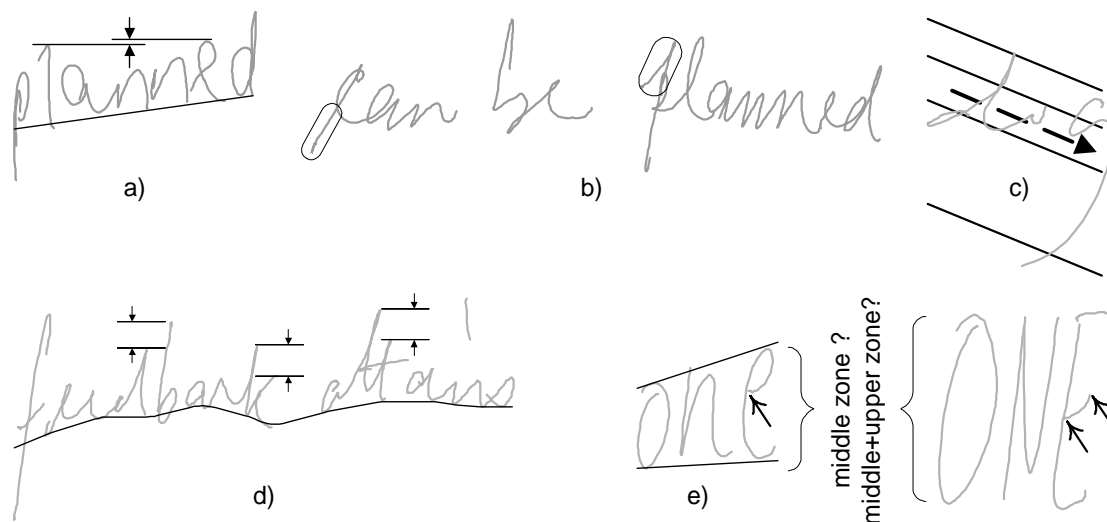
Fig. 1: Zoning problems: a) tilted baseline and inseparable top- and half points, b) unexpected zoning, c) short 'tilted' words d) ambiguous ascender and curved baseline, e) inconsistent zone height and lack of half-line information.

written separately). The position of the affected zone-line (the bottom-line in Fig. 1d) is therefore undefined and might have to be obtained heuristically. In some words, none or only little information indicating the position of the half-line can be found. This is particularly true for short middle-zone words and words that are written in capital letters only (Fig. 1e).

1.2 Zoning estimation in the literature

The various methods for zone-line estimation described in the literature can be characterised by a) the amount of text the zone-lines are based on (e.g. a zoning scope of one or more words); b) the source of information used to determine the lines; c) the representation of the zone-lines and d) the various restrictions required by the method, e.g. maximal rotation of text, availability of word boundaries, etc.

Most methods determine zone-lines on a per-word basis as wordwise zoning allows a certain variability of the height of zones between subsequent words and because most recognition systems accept wordwise input. The correlation in height and position between subsequent
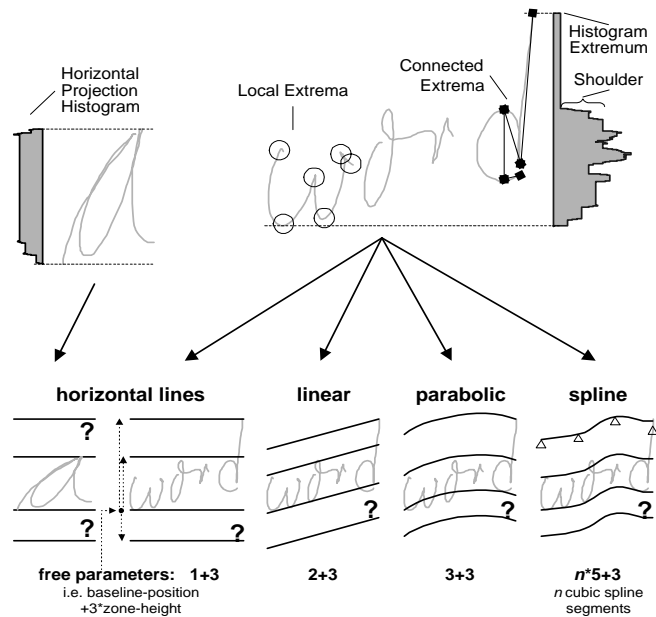
4

Fig. 2: Sources of zoning information and representation of zone-lines.

words is exploited in the method presented here as an entire line of text is used for the extraction of zone-lines. Original ideas of this approach were initially presented in [2] and more elaborately in [3].

Horizontal projection histograms (Fig. 2) are frequently used to locate the zone-lines [4-7]. The outer extrema of the histogram and the 'shoulders' found near the maximum are used (Fig. 2). If, however, no ascenders or descenders are present in the script, the histogram's extrema indicate the inner lines rather than the outer ones. The location of the outer lines is then undefined and has to be derived otherwise, e.g. by using an empirically determined 'usual' height of the outer zones in relation to the middle zone. Therefore the histogram has to be pre-classified into one of the four *word classes* (i.e. words containing ascenders, descenders, both or neither) in order to choose the appropriate heuristics to determine the zone-lines. This classification, however, has been found error-prone [6]. Moreover, the histogram can become rather flat - especially for short words - which might complicate reliable classification or identification of the shoulder.

5

As zones are often defined to envelop selected parts of the script, the vertical local minima and maxima are sometimes deemed sufficient to define the zone-lines. Methods using local extrema [6-8], however, have to base their decision on few indicators, which can make them more susceptible to noise. Furthermore, each extremum has to be pre-classified into one of the four zone-lines, e.g. based on the projection histogram. Again, this classification is susceptible to errors.

The method presented here uses the connectivity of local extrema in order to identify parts of the ink relevant to zoning. 'Connectivity' describes which pairs of extrema are connected by ink. Isolated diacritical marks, for example, can thus be identified and discarded, as they are small segments that are not connected to the word's body.

Most methods use horizontal lines to represent the zone-lines of words, which requires determination of four parameters (Fig. 2). If intervals are used to accommodate variation of extrema on the inner zone-lines [7] the number of parameters rises to six. Diagonal parallel lines [9,10] and equidistant parabolas [11] have also been used. While more flexible representations allow higher variability of ink , they also require determination of an increasing number of parameters and hence need to exploit more information. Some methods [6,7] aim to provide a set of clear thresholds above or below which ink is deemed to belong to certain zones, assuming that such horizontal separation is possible. Approximating methods [9,11] attempt to minimise the overall distance of selected points from their respective zone-lines. Neural networks have also been used to estimate position and height of the middle zone [12], but only few details of the network's architecture and training are provided by Seiler.

As zone-line extraction is usually mentioned under pre-processing, few papers provide a detailed evaluation. Performance is often assessed by visual inspection and quoted in terms

of improvements in the final letter or word recognition rate. This, however, depends partly on the recognition algorithm used and its capability to deal with zoning errors. Guerfali [6] proposes measures for the variability of the middle zone and for the error in detecting inner zone-lines. These measurements, however, are based on the assumption that horizontal zone-separation is possible and focus on the efficacy of thresholds identifying ascenders and descenders rather than on the description of the zones itself.

1.3 Design Choices

Further to the two above classes of problems (writer-induced and lack of information), zoning estimation is also impaired by mutual dependencies between individual estimation steps. As an example, zone-lines could be defined in terms of previously recognised letters in the word, which is probably how human beings define the zoning of more complex script. Automatic letter recognition, however, usually depends on some kind of zoning information to distinguish between at least some pairs of letter hypotheses. Another cyclical problem often encountered is the need to classify extrema to belong to one of the four zone-lines in order to be able to define the zone-lines, which could then in turn be used to classify the extrema. Similarly, the histogram projection method requires the classification of the word into one of the four word-classes in order to interpret the histogram, which would then yield the zone-lines to allow histogram classification.

In order to address these problems, the following choices of design have been made. As zone-lines are often curved and smooth, cubic spline functions have been chosen as a representation. In order to address the problems of 'unexpected' writing approximating spline functions have been employed as they can disregard irregular parts of the script. Approximating spline functions also address the problem of ambiguous ascenders and descenders. As the height of zones can vary, an independent spline is used for each zone-line

7

(i.e. zone-lines are not equidistant as in Fig. 2). The problem of insufficient information is mitigated by using an entire line of text instead of single words. The problems of interdependency are circumvented by dividing the estimation process into small steps, each step solely depending on already available information. Preliminary estimates of zone-lines are obtained and then successively refined based on the contextual information available at the time. The first estimate of the baseline (called a '*guideline*'), for example, roughly indicates the lower half of the middle zone. Based on this guide, descenders can be identified and ignored to obtain an improved estimate (a second guideline), which will ultimately lead to a well-corroborated estimation of the final baseline.

2 Methodology

The script on a page is segmented at local horizontal extrema into small parts, the '*ink-boxes*' representing the connectivity information. Each ink-box is attributed by its bounding box, the amount of ink covered by the box as well as a representing point and a tolerance measure. The extraction of these ink-boxes from dynamic handwriting data is described in more detail elsewhere [13], as is a simplified extraction method from static handwriting images 3. The representing point is chosen with respect to the desired zone-line, e.g. the bottom of the box during estimation of the baseline. Contextual information provided by overall statistics and previously obtained guidelines is used to adjust both representative point and tolerance. In order to eliminate undue oscillation of the approximating cubic spline function [13] (denoted by $g(x)$; *spline* for short), the spline's control points are obtained by resampling the representative points and their associated tolerances.
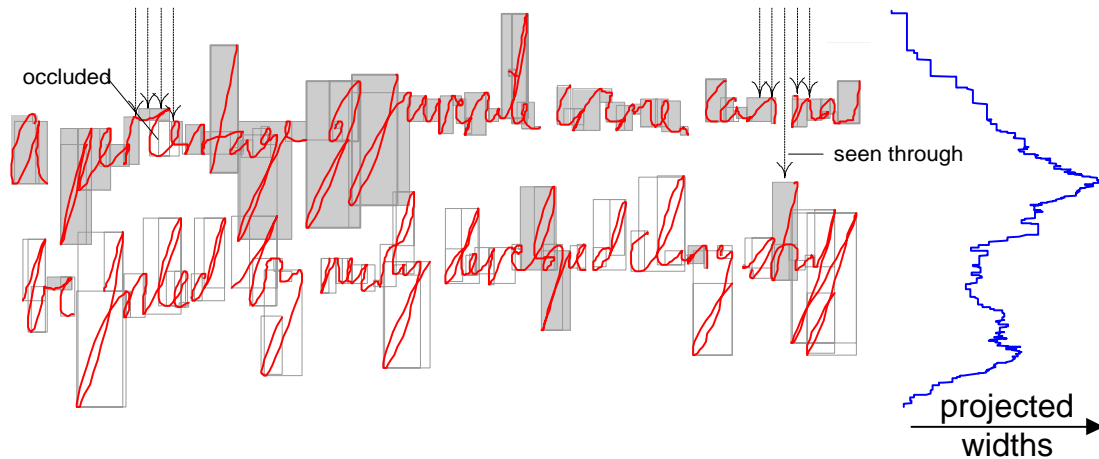
Fig. 3: 'Vertical Descend' marks ink-boxes of the top line, but includes spurious boxes 'seen' through gaps. The width-projection histogram shows no clear local minimum, let alone a gap.

The resampling ratio and the tolerances determine the overall smoothness of the resulting spline. The reach of the triangular weighting function used for weighted averaging during resampling models the distance of influence of individual ink-boxes. Boxes representing the relatively rare descenders, for example, define the height of the lower zone over a wider distance than the many boxes representing middle zone segments. The estimate of the bottom line should consequently employ a larger distance of influence. As the bottom line is also expected to be smoother than the half-line, the resampling ratio should be smaller. Resampling also compensates for an overall upwards or downwards trend of the script by projecting the ordinate of the representative points along the assumed axis of the script. Initially, the linear regression of the ink-boxes is employed as an estimate of the axis. The centre-line (i.e. the zone-line representing the centre of the middle zone) is then estimated and used as a more accurate description of the axis (see section 2.2 for details).

9

2.1 Line-Segmentation

From the list of all ink-boxes, the line-segmentation step attempts to coarsely identify the majority of boxes forming the topmost line of text. Based on these topmost boxes, the zone-lines are estimated. As the line of text is then encompassed by the top and bottom zone-line, ink-boxes belonging to the topmost text-line can be more accurately identified and removed from the set of ink-boxes. The method continues to identify, approximate and remove the next-highest lines of text until no more boxes remain.

The simplest way to segment pages into lines of text is horizontal projection of the ink into a (vertical) histogram (Fig. 3), in which gaps indicate line-breaks. This method, however, depends on horizontal text. Even if skewed text is identified and corrected beforehand, the lines must be clearly spaced, as the histogram might otherwise not contain the required gaps. Additional spurious gaps might appear if diacritical marks are written far above the rest of the text-line. If the text is enclosed in boxes, the histogram might not contain any gaps at all, rendering segmentation impossible.

The 'Vertical Descent' segmentation has therefore been developed, which assumes that lines are arranged relatively smoothly and near-horizontally, and that a text-line 'hides' most of subsequent lines if one descends vertically. Only the few segments 'seen' from above are considered for the topmost text-line (Fig. 3). This set can contain spurious boxes ('seen' through the gaps in the topmost text-line or 'around' the line) and usually does not include all boxes that actually belong to the line of text (the 'occluded' boxes). The approximating spline function based on the initially selected boxes, however, is precise enough to serve as a preliminary indication of the text's location. Spurious boxes far below the topmost line can thus be identified and ignored and the guideline updated accordingly. The set is then extended by boxes near to the new guideline, thus including originally occluded boxes. The

guideline is then made to follow the boxes more closely by increasing the resampling rate, which allows identification of discontinuities. Discontinuities usually occur if the set describes two different lines of text with large gaps in the higher line. The guideline might then alternate between the higher line and - inside the gaps - the lower line. Detected discontinuities are thus used to restrict the set to ink-boxes of the higher line and the guideline is updated accordingly. The set is again amended by boxes surrounding the guideline in order to contain the majority of boxes that form the line of text.

2.2 Zone-line Estimation

Individual zone-lines are estimated based on the set of boxes. First the centre-line is established and used as the text axis during resampling of the remaining splines. The centre line is expected to be best described by the centroid of boxes forming the middle zone. Centroids of Ascenders and descenders would unduly influence the spline and should be discarded. They could be identified by contextual information such as the average height of all boxes in the line of text or their position in relation to the text. The average height is easily obtainable from the set of boxes and is used to modify the tolerances assigned to each representative point. Increasing the tolerance allows modelling the certainty to which a box is not an ascender (or descender) and thus the degree to which it is "ignored"[1]. Avoiding a crisp decision also limits the effect of incorrectly identified ascender/descender-boxes, which might contain useful information about the centre-line.

Contextual size-information, however, is unable to distinguish between middle-zone boxes and diacritical marks or 'open-ended loops' attached to an ascender/descender (Fig. 4A).

---

[1] Subsequently, "ignored" or "eliminated" ink-boxes are only reduced in influence by increasing the tolerance.
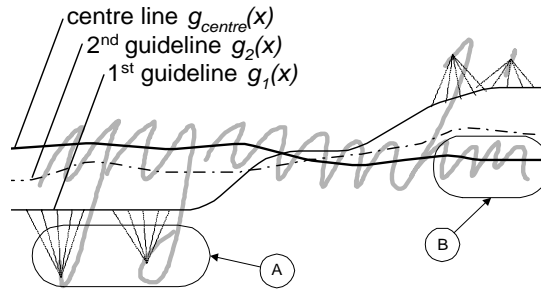
Fig. 4: Estimation of the Centre-line via

guidelines

This often leads to a distorted first guideline g_1, which can nevertheless provide the contextual position-information needed to identify and eliminate the affected boxes for calculation of g_2.

In some cases, the elimination of outliers described above is affected by errors propagated from a heavily distorted first guideline (Fig. 4B). The second guideline is therefore used to restrict the estimation of the final centre-line to middle-zone boxes only.

Various strategies could be adopted for the extraction of the remaining zone-lines. The half-line for example, could be extracted first. Ink-boxes that extend above the half-line could then be identified as ascenders whose upper borders define the top-line in turn. An error in the half-line, however, might affect the top-line as well. In order to minimise such dependencies and error-propagation all zone-lines and their intermediate guidelines are calculated independent of other zone-lines, referring only to the centre-line during resampling. Their estimation follows similar reasoning and utilises the same contextual information as demonstrated for the centre line. Two to three intermediate guidelines are used per zone-line (see [3]) for details).

The obtained zone-lines exhibit the following properties. As the splines are approximating the representative points, the zone-line does usually not reach all extrema found in the script

if they are of inconsistent height (e.g. the tops of 'f' and 'l' in). This compromise can not only deal with contradicting ascenders/descenders but also restricts the impact of potentially misclassified boxes. If no boxes can be found within the distance of influence, the outer zone-lines approach the respective inner lines (exaggeratedly so in Fig. 6). While the outer lines could be kept at a minimum distance from the middle zone, this would require an empirically obtained average height of the outer zones. The heuristics, however, might not correspond to the script at hand. Using empirical heights would also fail to indicate the lack of precise information, which could be used to adjust the method classifying points.

2.3 Zoning Classification

The most frequent use of zoning-information is the classification of a given point as one belonging to the top-, half-, base- or bottom-line. Based on point classification, the zoning-class of strokes, letters and words can be established or diacritical marks can be identified by their position.

For this, the vertical position of a point under investigation is normalised with respect to the middle-zone at its horizontal position. The point is drawn into the prominence plot, showing the point's vertical position against the prominence of the relevant outer zone (Fig. 5c). The prominence is defined as the height of the outer zone relative to the middle zone. If the zoning is prominent, the classification uses simple triangular fuzzy sets based on all four zone-lines to obtain a fuzzy classification of the point (Fig. 5a). In non-prominent areas, the fuzzy sets are based on the middle zone alone (Fig. 5b). Fuzzy-logic has been used to cope with points of ambiguous zoning, which can be identified and marked as 'uncertain' in zoning class.
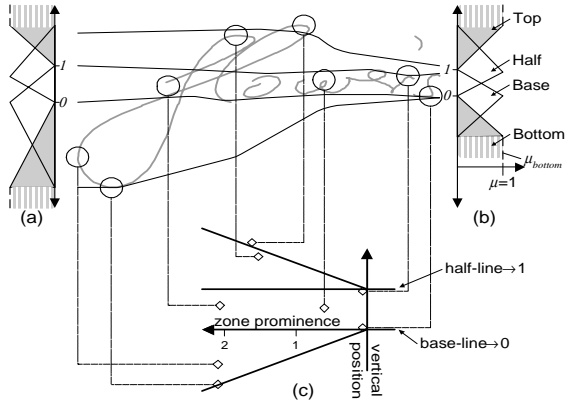
Fig. 5: Zone prominence and fuzzy
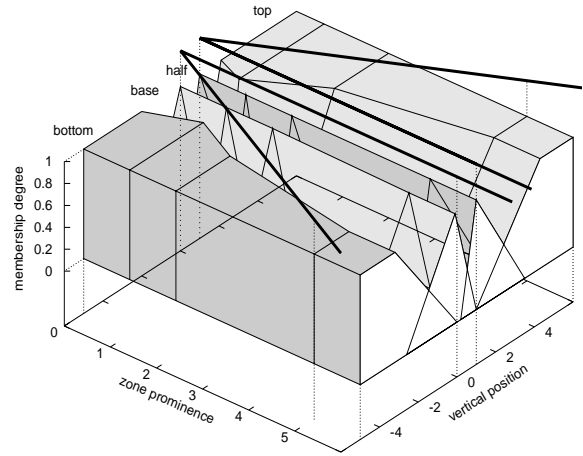
point classification.

Fig. 6: 2D fuzzy mapping for prominence-

based point classification.

In order to avoid a crisp decision of prominence or non-prominence, a 2-dimensional fuzzy relation has been defined to exploit both the observed zoning prominence and the vertical position of the point (Fig 6). The mapping at the far ends of F_{2D} (i.e. at prominence=0 and ≥ 6) corresponds to the two fuzzy set systems depicted in Fig. 5b and 5a, respectively.

3 Experimental Evaluation

The proposed method has been evaluated extensively on 300 pages of handwriting, i.e. 12 sentences written by 25 writers, one sentence per page, resulting in 3750 individual words. This set has been collected twice, for training and testing purposes respectively. The sets were collected as dynamic data using a NCR3125 pen-computer and electronically converted into images for evaluations on static data. Some tests require reference zone-lines, which have been provided manually for a smaller subset (2 sentences by 25 writers, i.e. 50 pages). To evaluate the accuracy of human zone-line estimation, one subject was asked to mark the four main zone-lines on images containing blackened ink-boxes only, i.e. without the aid of the textual content of the page (1 sentence, 25 writers, i.e. 25 pages).

14

3.1 Line-segmentation

The accuracy of the vertical descent line segmentation has been manually established from the 300 test set pages, several containing overlapping lines as in Fig. 3. While simple horizontal projection failed in 18% of all pages (values refer to dynamic data unless mentioned otherwise, 17% for static data), vertical descent has been found to be correct in all cases (0.3% error for static data). Even for a modest rotation of 4°, the error rate of horizontal projection rose to 37% (35%), while vertical descent failed only in 1% for pages rotated by 10°.

3.2 Zone-line accuracy

The accuracy of the estimated zone-lines is evaluated by the *mislocation error M*, the average vertical distance of the points constituting the reference zone-lines from the respective estimated zone-line. Normalisation by the average middle-zone height (by reference zone-lines) renders M size-independent. The error E proposed by Guerfali [6]) evaluates the maximum distance between reference and estimate normalised by the maximum height of middle zone and the appropriate outer zone, focuses on the accuracy of detecting the middle zone. Using the maximum instead of the average yields the worst-case error rather than the average error. Worst-case error is important for threshold-based point classification, but assumes that zone-lines can be separated horizontally, which has been found impossible in 6% of the words. If the obtained zone-lines are used for calibration of subsequent methods (e.g. the height of middle zone for filtering strength [7]), the average error is more appropriate, especially as zones need not be horizontally separable.

The mislocation error M of 15.6% for linewise zoning (19.2% for static data) has been observed to outperform the human estimate on blackened boxes (19.4%). Applying the new

method to individual words instead of lines of text, the performance degraded to 17.0% (22.2% static), which demonstrates the benefit of extended zoning scope. The accuracy of optimal horizontal zone-lines (i.e. the wordwise average of references) yielded an *M* of 22.6%, which suggests that horizontal lines are often inappropriate as representation of zone-lines. The mislocation of a real system using horizontal zones [7] has been found worse (26.4%) than the theoretical optimum, as the classification of points used for zoning extraction is sometimes incorrect. The optimal skewed zone-lines error (i.e. the wordwise least square linear regression of references) has been found to be slightly better than the spline-based equivalent (16.4% instead of above 17.0%). However, this is expected to degrade as a result of imperfect point classification in a real system. While this finding suggests that linear zoning might be similarly accurate to spline-based zoning, it does not question the superiority of increased zoning scope (15.6%), which would be difficult to exploit with single lines representing an entire line of text.

3.3 Zoning classification of points

The accuracy of classifying points is measured by the *misclassification error C*, which is the percentage of reference points that have been classified as the wrong zoning class. The misclassification of the new method (2.3%, 4.8% static data) has been found to outperform both the human judgement (4.6%) and optimal zone-lines (2.6% horizontal, 2.4% skewed). Applying splines to individual words performs slightly worse (3.9%, 6.5% static data) as less contextual information was exploited. An alternative method using (wordwise) horizontal lines that - unlike the optimal zones - did not have human-supplied reference points performed far worse (10.2%) than the theoretical optimum. This reiterates the superiority of increased zoning scope and the use of approximation splines. Most misclassifications (59% of the 2.3% above) were observed between half- and base line points, which is a result of

frequently poorly aligned script in the middle-zone. Most of the remaining outer-zone misclassification errors, which is the more important error for recognition , occurred when the affected zone was of low prominence (30% upper zone, 11% lower zone).

As outer zone-lines are not artificially kept at a minimum distance from the middle zone and as a consequence of using fuzzy logic, the certainty of point classification can be evaluated. The *confusion* value (the membership degree of the winning zoning class divided by the degree of the second best class) has therefore been defined. Points with confusion exceeding 65% were marked 'uncertain' the remaining ones 'certain'. It has been found that 61% of wrongly classified points have been marked 'uncertain', which indicates that the system is often aware of its deficiencies for certain points and cases of zoning. The distribution of misclassification errors and some typical examples are shown in Fig. 7.

The misclassification and mislocation errors have been obtained for both static and dynamic data as well as for various angles of rotation of the entire page. Zoning usually performed better for dynamic data, which is the result of the more accurate extraction of the connectivity information in ink-boxes (Fig. 8). A further common observation is the superiority of increased zoning scope (ref. inset in Fig. 8). It has been found, that zoning is reasonably stable against a limited rotation of the page (i.e. inside [-20°,+30°]). If page rotation is detected and corrected before zoning estimation, the above range signifies the permissible rotation variation of individual text-lines. In our data set, the overall page rotation has always been found to be less than 2.5°, the rotation inhomogeneity less than 0.7°, i.e. both well inside the interval of stability.
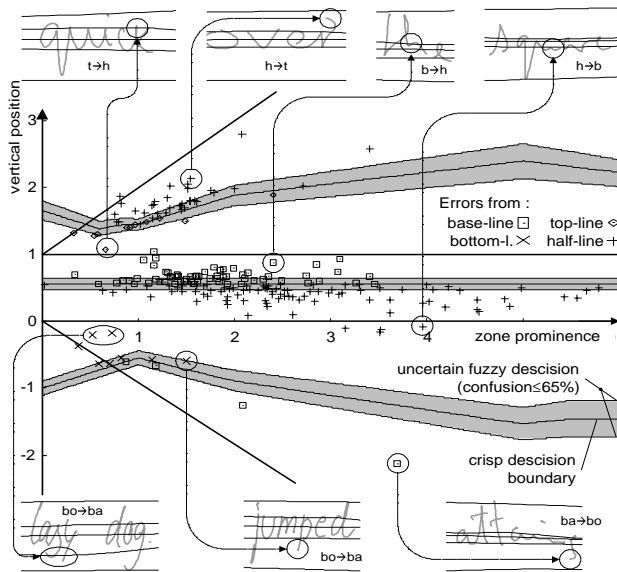
Fig. 7: Misclassified points and zoning examples

Fig. 8: Zoning accuracy versus rotation

3.4 Other applications of zoning

The zone-line information obtained by the proposed method has been applied to several problems related to the recognition of handwriting. Based on the fuzzy classification of the upper and lower limit of ink-boxes, boxes representing ascenders, descenders, combined ascenders/descenders and diacritical marks have been identified. Fuzzy classification allows extending this set by their less certain variations. A 'potential descenders', for example, is a ink-box that does not reach far enough towards the bottom line (e.g. the 'p' in 'jumped' Fig. 7).

While establishing the rotation of a page, upside-down pages have to be discriminated from the pages that are the right-way-up. A method of identifying inversion by means of diacritical marks has been proposed earlier [15], but did not perform well for hand-written pages. It was based on i-dots whose great variability complicated reliable extraction by shape information alone. Diacritical marks identified by their position relative to the estimated zone-lines have therefore been used. In English text, most marks are expected to be above the half line. If for
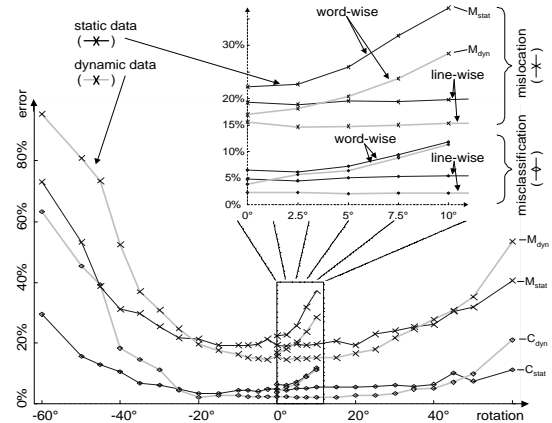
18

a certain page, more marks are found below the 'base-line' than above the 'half-line', the page has most likely been presented upside-down and should therefore be inverted. Furthermore, the number of ascenders is expected to be much higher than the number of descenders (in English text, ascenders are about 3.2 times more frequent than descenders). Ascenders and descenders identified by zoning information have thus been used to provide a second indicator for upside-down pages. On the set of 300 pages, the diacritical mark indicator has been observed to fail in 2.2% of pages (16.4% static data), as/descenders ratio failed in 2.7% (8.0%). A cascaded combination of both indicators uses the more reliable one (diacritical marks) if a clear verdict has been rendered, the as/descenders indicator otherwise. The upside-down error was thus reduced to 0.5% (4.8%), which is a significant improvement from the 9.6% correctness (static data) obtained from the earlier method [15] on a much simpler test set containing largely printed text, i.e. easily identifiable diacritical marks.

The word-class can also be easily derived from the prominent ink-boxes found above. Based on word-class, the size of a lexicon of word alternatives can be reduced dramatically by eliminating alternatives with incompatible zoning, i.e. perform holistic dictionary reduction. In comparison with a method based on horizontal zoning [7], the percentage of incorrectly rejected word alternatives has improved from 16.1% to 5.6% at similar effectiveness (a 4k dictionary was reduced to 41.2% and 39.8% respectively). In order to reduce the false-reject rate, the second likely word-class is accepted by Powalka [7] irrespective of congruency of word image and secondary word-class. The fuzzy classification available in the new method permits us to use relaxed matching only in cases where image and secondary classes correspond. While the false-reject rate was improved to 6.3%, the effectiveness decreased dramatically (64.7%) [7]. The fuzzy classification method, however, improves the false-reject rate more dramatically (to 1.2%) with only a small decrease in effectiveness (42.4%). As above, due to the simplified extraction method of ink-boxes, the word classification is less

effective for static ink (11.6% crisp, 5.6% fuzzy word classification at 38.2% and 41.9% effectiveness respectively).

If instead of the binary feature 'presence of prominent ink-boxes of a certain type' the number of prominent ink-boxes is used, the effectiveness can be improved to 19.9% (at 1.95% false-reject rate). To achieve this, ranges of expected prominent box counts are defined for every character, allowing for style variation. The accumulated ranges are compared against the number of extracted prominent boxes. Frequently observed effects affecting the box count are modelled and the accumulated range is updated accordingly. This includes the observation of an additional combined as/descenders, which are usually restricted to letter 'f's and in the ligature of combinations like 'gh'. Similarly, initial or final letters of a word have often been observed to be oversized (e.g. Fig. 7, the oversized initial 's' in 'square' and oversized final letters 'r' in 'over' and 's' in 'attains').

4 Conclusion

The proposed method has been found to provide zoning information for a much larger set of problems than previously published. Zone-lines are no longer restricted to straight lines (horizontal or linear) or parallel curves. During zoning estimation, extrema are not classified crisply and therefore not affected by isolated errors, which occur frequently as a result of ambiguous ascenders/descenders and writing of unexpected or unclear zoning. Errors occurring during estimation are further reduced by the iterative estimation of the zone-line by means of guidelines. The most important improvement over traditional methods, however, is the increase of the zoning scope to entire lines of text, which aids disambiguation especially of short and 'tilted' words.

The obtained zone-lines have been found accurate in terms of their distance from the real zone-lines (i.e. the mislocation) used to describe the scripts vertical size. The classification of points based on the zone-lines has been observed to be of superior accuracy compared to other methods and even the human performance (i.e. lower misclassification error of 2.5% rather than 10.2% [7] and 4.6%, human). Classification of points utilises a 2-dimensional fuzzy mapping, which avoids a further crisp decision as to the presence of outer zones by defining a zone-prominence measure. Point classification uses fuzzy logic also, which defers a crisp decision further. The extension of the zoning-scope to entire lines of text rather than words has been found to improve performance significantly compared to word-wise zoning. Point classification has successfully been applied to various handwriting-related problems such as upside-down detection of pages and holistic word recognition.

The proposed method is therefore considered a significant improvement over other zoning methods, in terms of increased number of zoning-problems addressed, improved performance and wide-ranging applicability.

5 References

1.   Hennig, A., N. Sherkat, and R. J. Whitrow, "Obtaining Words from a Sparse Graph of Letter Candidates: The Wildcard Combination Method", *8th Conference of the International Graphonomics Society IGS'97*, Genova, Italy, 24 August 1997, pp. 89-90.

2.   Hennig, A., N. Sherkat, and R. J. Whitrow, "Zone Estimation for Multiple Lines of Handwriting Using Approximating Spline Functions", in *Proceedings of the Fifth International Workshop on Frontiers in Handwriting Recognition IWFHR-5*, Essex, England, September 1996, pp. 325-328.

3.   Hennig, A., *Recognising a Page of Unconstrained Cursive Handwriting*, PhD Thesis, The Nottingham Trent University, 1 June 1998.

4.   Burr, D. J., "A Normalizing Transform for Cursive Script Recognition", in *Proc. 6th Int. Conf. On Pattern Recognition*, Munich, vol. 2, October 1982, pp. 1027-1030.

5.   Leroux, M. and J. C. Salome, "Recognition of Cursive Writing. A Method of Segmentation of Words into Letters", in *ICARV*, Singapore, September 1990, pp. 1211-1215.

6.   Guerfali, W. and R. Plamondon, "Normalizing and Restoring On-Line Handwriting", in *Pattern Recognition*, vol. 26, 1993, pp. 419-431.

7.   Powalka, R. K., N. Sherkat, and R. J. Whitrow, "Feature Extraction: On the Importance of Zoning Information in Cursive Script Recognition", in *Progress in Image Analysis and Processing*, ed. S. Impedovo Ed., World Scientific, 1994, pp. 342-349.

8.   Brown, M. K. and S. Ganapathy, "Preprocessing Techniques for Cursive Script Word Recognition", in *Pattern Recognition*, vol. 16, 1983, pp. 447-458.

9.   Caesar, T., J. M. Gloger, and E. Mandler, "Preprocessing and Feature Extraction for a Handwriting Recognition System", in *Proceedings of the Second Int. Conf. On Document Analysis and Recognition ICDAR'93*, Tsukuba Science City, Japan, October 1993, pp. 408-411.

10.		Caesar, T., J. M. Gloger, and E. Mandler, "Estimating the Baseline for Written Material", in *Proceedings of the Third International Conference on Document Analysis and Recognition ICDAR'95*, IEEE, August 1995, pp. 382-385.

11.		Bengio, Y., Y. Le Cun, and D. Henderson, "Globally Trained Handwritten Word Recognizer Using Spatial Representation, Convolutional Neural Networks and Hidden Markov Models", *Advances in Neural Information Processing Systems*, no. 6, 1994, pp. 937-944.

12.		Seiler, R., M. Schenkel, and F. Eggiman, "Cursive Handwriting Recognition: Off-Line Versus On-Line Recognition", in *Proceedings of the Fifth International Workshop on Frontiers in Handwriting Recognition*, Essex, England, September 1996 *IWFHR'96*, pp. 23-28.

13.		Hennig, A., N. Sherkat, and R. J. Whitrow, "Recognising Letters in on-line Handwriting with Hierarchical Fuzzy Inference", *Fourth International Conference on Document Analysis and Recognition ICDAR'97*, Ulm, Germany, 18 August 1997, pp. 936-940.

14.		Reinsch, C., "Smoothing by Spline Functions", *Numerische Mathematik*, vol. 10, 1967, p. 177.

15.		Hennig, A., G. Raza, N. Sherkat, and R. J. Whitrow, "Detecting a Document's Skew: A Simple Stochastic Approach", *Vision Interface'97, the Eleventh Canadian Conference on Computer Vision, Signal and Image Processing, and Pattern Recognition*, Kelowna, Canada, May 1997, pp. 97-102.