New Algorithm of Straight or Curved Baseline Detection for Short Arabic Handwritten writing

Houcine BOUBAKER , Monji KHERALLAH , Adel M. ALIMI
Research Group on Intelligent Machines (REGIM), National School of Engineers ENIS,
University of Sfax, BP 1173, Sfax 3038, Tunisia
houcine-boubaker@ieee.org , monji.kherallah@enis.rnu.tn , adel.alimi@ieee.org

Abstract

In this paper we present a new method of baseline detection of online or offline short handwriting. This work is part of a large project for the edification of a dual online / offline Arabic handwriting recognition system. Compared to the existing approaches in the literature, this new method brings three specific novelties: First, the consideration of the agreement between the alignment of the points and their trajectory tangent directions for the detection of aligned points regroupings. Then, the consideration of a topologic characteristics specific to the used writing language, to value the pertinence of the pretender points regroupings to be recognized as baseline. Finally, we showed the aptitude of the algorithm to detect curved baseline.

1. Introduction

The baseline detection is an essential stage in a parametric or structural features extraction process [1] [4] [5], in particular for the segmentation of a semi cursive writing as the Arabic. Indeed, the detection of the baseline associated to the application of topologic and logical rules makes it possible to segment the handwriting in graphemes or in visual codes delimiting downward shafts, ascending shafts, closures, open curves, or legs [7]. A parametric or structural modeling of the detected segments consents then to recognize the treated handwriting text [2] [7].

The objective of the presented work is the baseline detection of online or offline skeleton handwriting. We exposed it in 4 paragraphs. First, we introduce in the next section the state of art and the specificity of the application. Then we present in the third section, the principle of the elaborated algorithm for the baseline detection. In continuation we joined an assessment procedure to optimize proposition select decision. In the fifth section, we study the faculty of the algorithm to detect curved baselines. Finally we conclude by presenting statistic results and perspectives.

2. The state of the art

By definition, the baseline is the virtual line on which cursive or semi cursive writing characters are aligned and/or joined. Indeed, it represents in writing as well as in reading, a reference for the vertical positioning of every character and ligatures graphics as well as for the distinction of the graphemes that they compose [1] [10].

Several baseline detection methods are proposed in the literature but that they apply often on large enough text block. The most known are:

- The histogram method: The projection of the writing tracing points according to a predefined direction makes it possible to detect the levels of the baselines that coincide with the local maximums of these histograms [4] [5]. This method has the defect to be very sensitive to the skew [3].
- The Hough transform method: The transposition in the polar coordinates space of the tracing points makes it possible to detect agglomerations of point images intersection defining the angle of writing lines skew [8] [9]. This method is expensive enough in calculation time and it is applicable on text blocks (of several lines).
- The entropy method: The projection of the words contour ordinates according to several rotation plans. The calculation of the entropy of each of these histograms permits to find the orientation according to which the word is the most compact [6]. This method has also the defect to be expensive in term of calculation time.

Other published baseline detection methods applicable on short handwritten texts [3] do not tolerate on the other hand even weak non alignment of its words.

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.265

IEEE computer society

Since the context of our application: trajectories of a few number of arabic words online handwritten (short messages) or offline handwriting skeleton (postal addresses) which are not strictly aligned, is maladjusted to the existing baseline detection approaches, we have been proposed to develop a new method better adapted to this context in order to assure the success conditions to the ulterior stages: to know the segmentation, the parameters extraction and the text handwriting recognition.

3. A new method of semi cursive handwriting baseline detection

Conversely to the existing alignment detection methods (histogram, Hough transform, entropy methods) based solely on geometric features, we propose in our approach to consider two types of features:

- Geometric features: alignment of point neighbourhoods (concord between interpolation directions and trajectory tangents directions).
- Topologic features: rules and logical conditions associate to the arrangement of the characters on the baseline.

The developed detection process consists of two stages: the first one is a basic stage permitting the detection of the points regroupings of aligned neighbourhood. The second stage measures by qualitative parameters, the verification level of some topologic conditions by the most numerous points regroupings found in the first stage.

3.1. Basic Stage: Detection of the alignments of neighborhoods

The developed method considers that the baseline can be slanted of an angle α_{LB} that we suppose limited in order to alleviate the calculation load.

$$-\alpha_{\lim} \le \alpha_{LB} \le \alpha_{\lim} \qquad (1)$$

The principle of the method consists in looking among the $\{M\}_{Str}$ trajectory points set, of which the $\alpha_{tg\,M}$ tangent slant angle justifies (1), for the points regroupings verifying the alignment of their neighborhoods:

$$\forall M_{n,i} \text{ and } M_{n,j} \in \{M\}_n \text{ we have } \Delta\alpha_{i,j} + \Delta\alpha_{j,i} < \Delta\alpha_{\lim} \qquad (cI)$$

With $\Delta\alpha_{\lim}$ is the tolerance limit of the absolute deviation angles between the trajectory tangents. And :

$$\Delta\alpha_{i,j} = \left|\alpha_{tg\,M_{n,i}} - \text{the slant angle of the straight line}\,(M_{n,i}, M_{n,j})\right|$$

$$\Delta\alpha_{j,i} = \left|\alpha_{tg\,M_{n,j}} - \text{the slant angle of the straight line}\,(M_{n,i}, M_{n,j})\right|$$

(See figure 1)

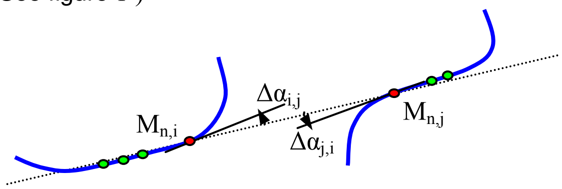

Fig. 1 – verification of the trajectory neighborhoods alignment

A point candidate M_k verifying the conditions of affectation (cI) to several regroupings $\{M\}_{1,\dots,q}$, is assigned to the regrouping of rank *m*: $\{M\}_m$ where agrees best its trajectory tangent direction with those of the other members as well as with the directions of interpolation (M_k , $M_{m,i}$) in accordance with the following criterion (cII) :

$$\Delta\theta_{M_k}(m) = \underset{n=1,\dots,q}{\mathrm{Min}}\left\{\Delta\theta_{M_k}(n)\right\} \qquad (cII)$$

$$\textit{with}: \quad \Delta\theta_{M_k}(n) = \frac{1}{N_n}\cdot \sum_{M_{n,i}\in\{M\}_n}\left\{\Delta\alpha_{k,i} + \Delta\alpha_{i,k}\right\}$$

Where N_n is the initial size of the $\{M\}_n$ regrouping and $m \in \{1,\dots,q\}$.

A new points regrouping is initialized when the point candidate M_k is not included by any already constituted regrouping, following the affectation condition (cI). So in this case, the point M_k will form the first element of the new points regrouping

In order to reduce the calculation load, each regrouping is represented respect to a new candidate point, by its point barycentre M_{bc} provided with the average direction of the tangents to the trajectory in its elements. The calculation of the barycentre coordinates (X_{bc}, Y_{bc}) and the average direction of the tangents $T_{bc} : (A\cdot x) + B$ associated to a given regrouping is up to date every time that a new $M_k\left(X_{M_k}, Y_{M_k}\right)$ point integrates it.

$$\begin{bmatrix} X_{bc} \\ Y_{bc} \end{bmatrix}_{current} = \left[N \cdot \begin{bmatrix} X_{bc} \\ Y_{bc} \end{bmatrix}_{precedent} + \begin{bmatrix} X_{M_k} \\ Y_{M_k} \end{bmatrix} \right] \cdot \frac{1}{N+1}$$

$$\begin{bmatrix} A \\ B \end{bmatrix}_{current} = \left[N \cdot \begin{bmatrix} A \\ B \end{bmatrix}_{precedente} + \begin{bmatrix} A_{M_k} \\ B_{M_k} \end{bmatrix} \right] \cdot \frac{1}{N+1} \qquad (2)$$

With N is the initial size of the treated regrouping before the integration of the point M_k.

To the term of this first phase, the starting points set $\{M\}_{Str}$ is decomposed in several regroupings (see figure 2.a) whose number also depends a lot on the trajectory layout as well as on the rigidity of the points inclusion criterion (cI) measured by the value of supportable deviations limit $\Delta\alpha_{\lim}$. Practically an over decomposition of the starting set $\{M\}_{Str}$ do not support the baseline detection, but it permits to obtain regroupings of very arranged points in terms of the regroupings formation criteria (cI) and (cII). A second phase would be then necessary to form regroupings of bigger sizes. This phase is based on the migration of the points from a small size points regrouping to the nearest bigger size one related to the (cI) and (cII) criteria supported by a weighting coefficient proportional to the logarithm of the size of the new reception regrouping. So a M_k point belonging initially to a given regrouping $\{M\}_p$, migrates to integrate another regrouping $\{M\}_m$ if :

$$\Delta\theta'_{M_k}(m) = \underset{n=1,\dots,q}{\mathrm{Min}}\left\{\Delta\theta'_{M_k}(n)\right\} \qquad , \textit{with}: \qquad (3)$$

$$\Delta\theta'_{M_k}(n) = \left[\frac{1}{N_n}\cdot\sum_{M_{n,i}\in\{M\}_n}\left\{\Delta\alpha_{k,i} + \Delta\alpha_{i,k}\right\}\right]\cdot\left[\frac{\mathrm{Log}(1+N_n)}{N_n}\right]^d$$

Where *d* is the factor of the weighting coefficient power, *q* the total numbers of regroupings, and the regrouping index $m \in \{1,\dots,q\}$.

The baseline detection at this stage of the treatment, consist in looking for the most numerous regrouping among the points regroupings of that are constituted (see figure 2.a). The first tests on the IFN/ENIT database of Tunisian town names give a preliminary baseline detection rate of 94.3%.

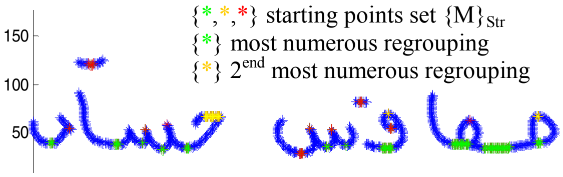

Fig. 2 – a/ constitution of the points regroupings and preliminary baseline detection by identification of the most numerous regrouping (green)

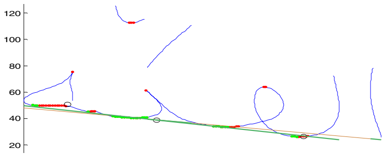

Fig. 2 – b/ detection of slanting base line

3.2. Verification of the topologic conditions

The examination of baseline detection errors shows that they are classified in two cases:

- Confusion of the baseline with the lower limit line for the cases of words composed essentially or exclusively of isolated character or of legs as ‘ر , و, ز , ن’ (example see figure 3.a).
- Confusion of the baseline with the median zone line, or the superior limit line, due to the writing style or to the presence of particular calligraphic effects (example see figure 3.b).

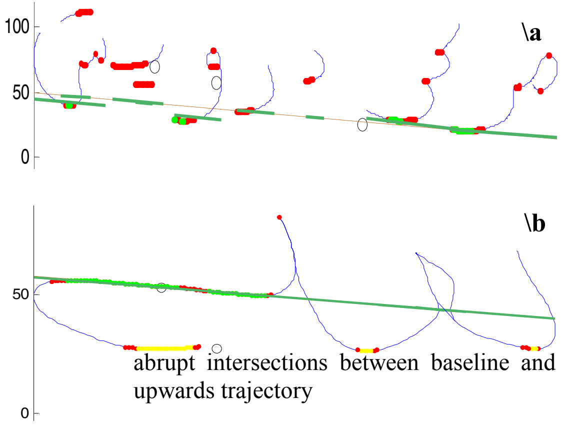

Fig. 3 – Examples of baseline detection errors: a/ Confusion of the baseline with the lower limit line. b/ Confusion of the baseline with the superior line of the median zone

The detection of these errors and their treatment impose the junction of a second stage consisting in estimating whether the line carrying the most numerous points regrouping verifies some logical and topologic conditions that characterize a baseline:

- The baseline doesn't cut vertical writing trajectory solely in the cases of the tracings of ‘legs’ (vertical or oblique) or ‘pockets’ and generally in the sense of the top downwards (examples: 'ل , ر, و, ز, ن, م').
- The absolute curvature angle of the grapheme trajectories drawn below the baseline for the arabic writing is generally limited to $\theta_{L\_under} = \frac{3\pi}{2}$.
- The continuous curvature angle of the graphemes segmented in over some baseline (according to its hypothetic position) doesn't exceeds $\theta_{L\_over} = 3\pi$.

The values of the last two thresholds θ_{L_under} and θ_{L_over} are estimated by experimental measures made on several examples extracted from the IFN / ENIT database.

- The barycentre of the segments which support an arabic characters or pseudo words on the baseline is centred or leaned on the right compared to the horizontal limits of their container box.

Several solutions can be considered to discern and to treat the baseline detection errors. In our case we opted for a function of assessment considering the first three most extended regroupings in term of number of points (of which the probability that one of them carries the baseline is ≈ 100%) in order to optimize the detection result. This cost function excels the size of the points regrouping (*npt*) and penalize:

- The number of irregular (abrupt) intersections (*nii*) between the line carrying the points regrouping candidate and the writing trajectory evolving upwards in a vertical direction: (see figure 3.b).
- The overtaking of the curvature angles limits θ_{L_under} and θ_{L_over} respectively by the strokes below the baseline (nda_{under}) and the graphemes in over of the baseline (nda_{over}).
- The bending on the left (*bbl*) of the barycentre of the segments which support the pseudo words on the baseline respect to the horizontal limits of their container box of width *Wid* ($rep_{lag} = \dfrac{bbl}{Wid}$).

The function of assessment *S* that takes into account these different parameters is expressed by the following formula (4) :

$$S = (\alpha_1 \cdot npt) - (\alpha_2 \cdot nii) - (\alpha_3 \cdot nda_{under}) - (\alpha_4 \cdot nda_{over}) - (\alpha_5 \cdot rep_{lag}) \qquad (4)$$

In order to estimate correctly the weighting coefficients α_1, α_2, α_3, α_3 and α_5, we assimilated the *S* function to the output of an ADALINE network simple layer trained according to the 'least mean square' rule. During the training phase, the system extracts for each sample the five parameters for every one of the first three most numerous point regrouping that will be put to the input of the network. On the other hand, its output value is manually assigned for every regrouping as follows:

S = 1: if the regrouping is judged like baseline
S = -1: if not

In the classification phase, system associates the baseline to the most numerous point regrouping of which the output of the *S* assessment function is positive. The figures 4.a and 4.b show some examples of correct results of baseline detection thanks to the consideration of the topologic conditions show compared with the results of the basic stage (figures 3.a and 3.b)

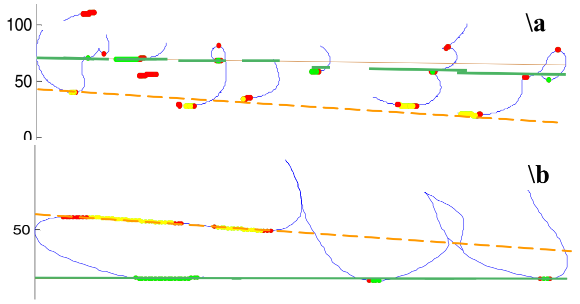

Fig. 4 a,b– Examples of detected baseline correction (green) obtained thanks to the consideration of topological conditions

Other examples:

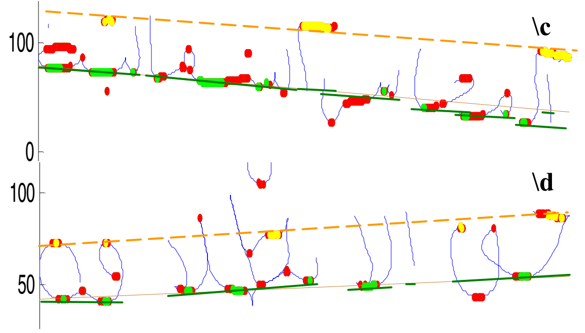

Fig. 4 c.d– Other results of baseline correction

4. Detection of curved baselines

To follow a strictly straight virtual baseline in handwriting is a coercive condition for the writer and not practical for the use of an online or offline handwriting recognition system. Of this fact we designed our algorithm in order to be able to detect the curved baselines. So the adopted approach that consists to detect points regroupings and no of the straight directions can be arranged in order to be able to seize the points segments of the virtual laying of the writing on a curved baseline (see figure 5). Indeed it is sufficient to define for every point candidate a neighbourhood of a given ray on which we apply the points regroupings integration and initialization criteria (cI) and (cII) (see figure 6). i.e. that a point candidate M_k is affected to the regrouping of rank *m*: where agrees best the direction of its tangent to the trajectory with those of the other members of this regrouping included in a neighbourhood of R_V ray surrounding of M_k as well as with the directions of interpolation $(M_k, M_{m,i})$ consistently to the following criterion (cIII):

$$\Delta\theta(m, M_k) = \underset{n=1,\dots,q}{\mathrm{Min}} \{\Delta\theta(n, M_k)\} \qquad \text{(cIII)}$$

$$\textit{with},\ \Delta\theta(n, M_k) = \frac{1}{N_n} \cdot \sum_{M_{n,.i} \in \{M\}_n \text{ such as } [M_k\, M_{n,i}] < R_V} \{\Delta\alpha_{k,i} + \Delta\alpha_{i,k}\}$$

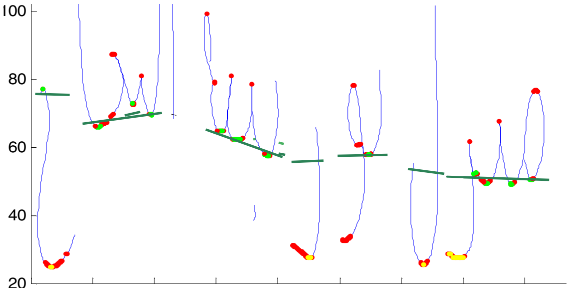

Fig. 5 – Curved baseline detection

The size of the R_V ray defines the width of the neighborhood in which the included baseline stroke can be approximated by a straight segment. The lower limit value of R_V is given by the average of the pseudo words widths.

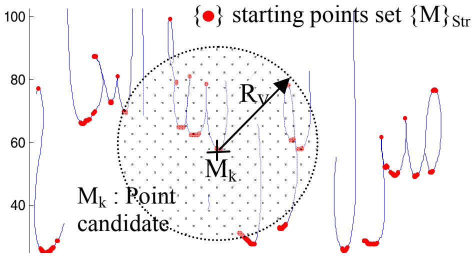

Fig. 6 – Neighborhood for criteria Verification

5. Results and discussion

The elaborate system is tested on the IFN / ENIT database of Tunisian town names for the offline handwriting samples and on the ADAB database for the online handwriting samples. Statistics of these tests done on a set of 1000 samples show that the baseline correct detection rate R_{cd} is of 97.9%. This final result improved in relation to that gotten by the algorithm basic stage of 94.3% returns to the integration of an assessment function of the detected points regroupings based on topologic parameters. The survey of the tests statistics shows also that the effect on the baseline detection rate of the pseudo words number or the length in pixels of the treated text is less influencing to the output of the assessment function (the 2nd stage of the algorithm) that to the output of the basic stage (the 1st stage) as to been noted in [3].

6. Conclusion

In this article we presented a new method of baseline detection for online handwriting or skeleton of offline handwritten writing. This method brings three specificities: first we consider the agreement between the alignment of the points and the direction of their tangents to the trajectory for the detection of aligned points regroupings. Then we use a topologic conditions specific to the language of writing (arabic in our case) to value the relevance of the points regroupings extract on the first stage to be recognized like baseline. Finally the aptitude to detect curved baselines thanks to the local application of the regroupings integration criteria. The tests statistics show first the robustness of the algorithm basic stage and the significance of the assessment function based on topologic parameters, to improves the detection rate to 97.9%. This advantageous result invites to adopt the algorithm in a more complete system of segmentation, modelling and recognition of online / offline handwritten short texts.

Acknowledgement

The authors acknowledge the financial support of this work by grants from the General Direction of Scientific Research and Technological Renovation (DGRST), Tunisia, under the ARUB program 01/UR/11/02.

7. References

[1] A. M. Alimi, Evolutionary computation for the recognition of on-line cursive handwriting, *IETE Journal of Research*, Volume 48, Issue 5 SPEC., September 2002, pp 385-396

[2] A. M. Alimi, Evolutionary neuro-fuzzy approach to recognize on-line Arabic handwriting, *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 1997, Volume 1, pp 382-386

[3] M. Pechwitz, V. Märgner. Baseline Estimation For Arabic Handwritten Words. *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)*, 2002, pp 479 – 484

[4] B. Al-Badr and S. A.Mohmond. Survey and bibliography of Arabic optical text recognition. *Signal Processing* 1995, pp 49–77.

[5] A. Amin. Off-line Arabic character recognition: The state of the art. *Pattern Recognition*, 1998, pp 517–530.

[6] M. Côté, M. Chériet, C. Suen and E. Lecolinet (1996), Détection des Lignes de Base de Mots Cursifs à l'aide de l'Entropie, *Colloque sur l'Intelligence Artificielle dans les Technologies de l'Information*, Montréal Canada, may 1996.

[7] M. Kherallah, F. Bouri, A. ALIMI. Toward an Online Handwriting Recognition System Based on Visual Coding and Genetic Algorithm. *Proceedings of the I. C. on adaptive and Natural computing Algorithm ICANNGA*, Portugal, 2005

[8] Z. Razak, K. Zulkiflee, M. Yamani, I. Idris. Off-line Handwriting Text Line Segmentation: A Review. *IJCSNS International Journal of Computer Science and Network Security*, VOL.8 No.7, July 2008

[9] Likforman-Sulem L., A. Hanimyan, C. Faure (1995) A Hough based algorithm for extracting text lines in handwritten documents, *In Proc. of the Third Int. Conference on document analysis and recognition (ICDAR)*, Montreal, Canada, August 1995, pp. 774-777

[10] M. Arivazhagan, H. Srinivasan, and S. N. Srihari, "A Statistical Approach to Handwritten Line Segmentation", in . Document Recognition and Retrieval XIV, Proceedings of SPIE, San Jose, CA, February 2007, pp. 6500T-1-11