HMM-based Handwritten Word Recognition System by using Singularities

Sebastiano Impedovo, Anna Ferrante and Raffaele Modugno
Dipartimento di Informatica – Università di Bari – Via Orabona, 4 – 70126 Bari – Italy
Centro "Rete Puglia" – Università di Bari – Via G. Petroni, 15/F.1 – 70126 Bari - Italy
impedovo@di.uniba.it

Abstract

This paper presents a new approach for Handwritten Word Recognition based on Hidden Markov Model theory and the sliding window technique. The new approach uses specific singularity markers to support the recognition phase: the Static Marker and the Dynamic Marker. Moreover, different strategies for sliding window step are considered: Regular Step and Progressive Step. Experimental results showing the improvements obtained for basic word lexicon recognition are reported in the paper.

1. Introduction

Handwritten recognition is a complex process, in fact during the last thirty-five years a lot of approaches have been proposed and several algorithms have been developed for handwritten digit, character and word recognition in a large variety of application fields [8, 9, 14, 15, 17, 23].

Without any doubt the total information about a grapheme, representing a handwritten trace, includes both the shape and the dynamics of its tracing process but in written words however, the history of this tracing is completely disregarded. On the other hand, for several millennia, the nature of the writing process has ensured knowledge transmission from generation to generation among humans. For this reason, for a long time researchers were not very interested in using algorithms for time dependent processing as theory of Markov chains. However, several attempts to use Markov chains theory for character recognition were made in the '70 years [10, 11] but based on the poor experimental results achieved no long Markov theory use attracted the scientific community.

In fact, after a better understanding of the computational theory and specifically after the successful formulation made by Cooley and Tukey [4] of the FFT algorithms, the Neural Network computation [7] and Rabiner's idea to use dynamic programming to handle Markov chains of the first order for speech recognition [19], the theory of Hidden Markov Model (HMM) [18] has been used at beginning of the '80 years also for off-line Handwritten Word Recognition (HWR) [2, 3, 16]. For this purpose, one of the most interesting technique in using the HMM for HWR, is the artificially recovery of the time dependence of a trace by using a sliding window [25].

Some advancement of the approach with the aim of investigating human mechanisms for handwritten recognition was developed. Furthermore, some experiments to develop very robust algorithms for word recognition have been also proposed in the last years by some researchers [21, 22, 24].

This paper presents a new version of a HWR system, based on the HMM theory and the use of the sliding window technique together new singularity markers and different sliding window step strategies.

The algorithms and prototypes have been developed by using an Integrated Development Environment (IDE) software based on a visual programming language.

In order to describe the progress and results obtained, this paper reports: in Section 2, shortly the HMM theory and the singularity markers; in Section 3, the recognition approachs based on the sliding window technique and the HMM; in Section 4, an overview of the system prototype with experimental results. Finally, Section 5 reports the conclusions and some suggestions for future improvements of the system.

2. HMM and Singularity Markers

This section presents the HMM theory according to the singularity-based approach for HWR and the new strategies based on a marking procedure to implement the recognition algorithms. An HMM is a double stochastic process. The first process is not observable or hidden, the second one produces a sequence of observable symbols according to the probabilistic law that associates the given observed symbol to the hidden state. An HMM is defined as a five-tuple

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.73

$\lambda =< Q,O,A,B,\pi >$, where “Q” is the finite set of hidden states, “O” is the set of observable symbols, “A” is the state transaction matrix, “B” is the matrix of the emission probabilities of the symbols in each state and “π” is the set of the initial state probabilities [18].

In handwritten word recognition the set of observations is obtained by feature extraction from the bitmap image. HMMs are successively trained according the following steps: prototypes creation, inizialization and re-estimation. In the first step, the number of states, the number of Gaussian’s components in the mixture for state and the topology are defined and adopted [1, 6, 25]. The HMM parameters for the inizialization step are achieved by the Viterbi algorithm, finally the re-estimation step utilizes the Baum-Welch algorithm [18].

In this paper, a new approach based on the singularity definition is proposed in order to improve recognition performance. The approach moves from the observation that an handwritten cursive word generally consists of a periodical or quasi-periodical part, called regular part, and of a non-periodical part, called singular part or singularity [21, 22].

In this work, the singularity is used to denote the ascending and descending word strokes that lie respectively in the upper and lower region of the word image. The horizontal histogram of the bitmap is used to locate the upper and the lower baseline by taking into account the inflection points. The image is divided into three areas, of course it is evident that the greatest part of the word under process is located in the central region (regular part), the ascenders are over the upper baseline and the descenders are under the lower baseline as shown in Figure 1.

The singularities are successively detected by processing the vertical histogram of the bitmap in the upper and in the lower regions. Scanning the histogram, for each block of consecutive foreground pixels, two singularity tags are generated: the first one (starting tag) corresponds to the supposed starting point of the singularity, the second (end tag) corresponds to the supposed end point. In order to discard not significant traits, the measure of the size of the distance between the starting and end tag is computed. This can be easily done by measuring the diagonal of the rectangular region including the singularity, or by considering the weight of the vertical histogram related of the singularity.

Successively, the proposed approach highlights the detected singularities using two specific markers of singularity:

- *Static Marker*;
- *Dynamic Marker*.

The approach helps the HMM in processing the entire handwritten word.

- The *Static Marker* acts on an extention of the bitmap image control box obtained by adding two rows on top image and including in these added rows a fixed marker above the singularities. Practically the marker consists of two lines of 16 pixels and it is positioned from the starting point of the detected singularity. The procedure is sketched in Figure 1.

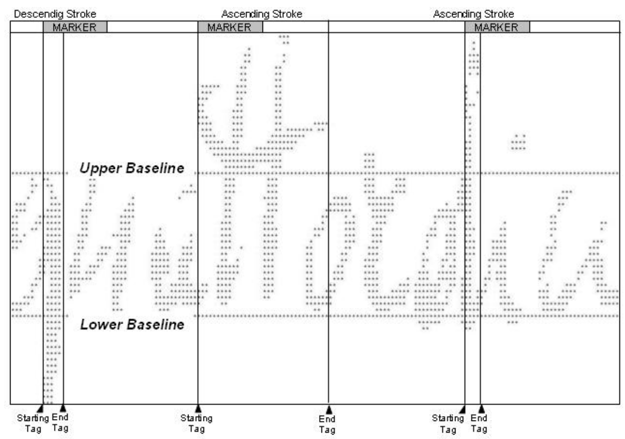

Figure 1. *Static Marker* of singularity.

- In the dynamic approach the total area of the bitmap image between starting and end tags is considered and tagged by weighting the pixels by a twice factor when compared with the regular part. The Figure 2 rappresents the singularities twiced weight by the darker traces.

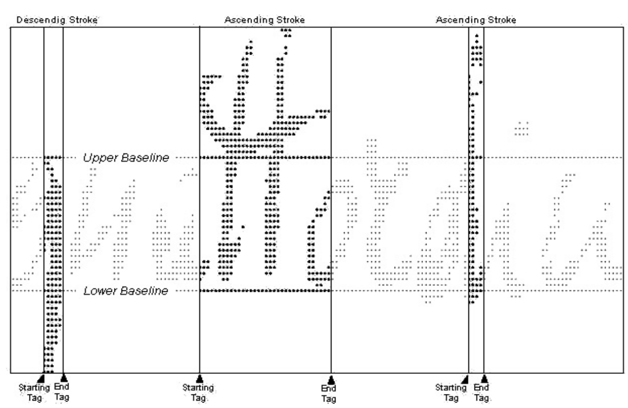

Figure 2. *Dynamic Marker* of singularity.

3. HWR: system overview

The HWR system has been developed highlighting three different approaches in terms of sliding window technique and HMM construction phase.

- The Baseline Sliding Window (BSW) approach uses the sliding window technique in order to extract the features from the bitmap image. The window is divided into sub-regions (5 rows and 4 columns) defined on the basis of the upper and the lower baseline. The baselines are calculated by horizontal histogram of the image as already described in section 2. The upper row contains the upper image part (ascenders) and the lower row contains the lower image part (descenders). The middle part is the core region of the image: it is processed by three rows having the same height. The sliding window has a fixed base's length and it moves from left to right across the handwritten word. The ratio between the number of foreground pixels, included in each sub-region, and the total number of pixels in the related window is considered as a vector element. At the end of the process on the whole word, the vectors obtained are considered as observations for the HMM construction.

- The Average Baseline Sliding Window (ABSW) is introduced in order to reduce the influence of the error in baseline detection on the performance of the system. Three windows having different position of the core region are considered: the first one is the BSW, the second one has a reduced upper area in the advantage of the symbols representation in the lower one, finally the third one is the complementary of the previous. For each frame three feature vectors ($f_{j,middle}$, $f_{j,up}$, $f_{j,down}$) are extracted using the different windows, then a weighted mean is computed for corresponding elements in vectors. The most decisive weight "p" (empirically estimated) is assigned to the first vector $f_{j,middle}$.

- The Multi-Branch (MB) approach [26], is also considered. Here the baseline shift concept is extended and 3 HMMs are used to model each word. The first one is trained by using the sliding window centred on the core region for feature extraction (BSW). The second and the third HMMs are obtained by using the sliding window with the upper and the lower shift already described for ABSW. The 3 HMMs are combined in a unique Multi-Branch model having a parallel left to right topology.

Moreover, the paper presents two strategies for sliding window step: the *Constant Step* and the *Progressive Step*.

- In the *Constant Step* approach the sliding window is moved column by column (4 pixel) from left to right, and for each position, a feature vector is extracted. This is the classic method for sliding window step.

- In the *Progressive Step* approach the sliding window is moved from left to right of one pixel at the first step, of two pixels at the second step, of three pixels at the third step, while following windows are moved column by column (4 pixel). This method extracts more observations in the initial part of image where a great level of ambiguity in the discriminating content generally happens as for instance the words "settanta", "ottanta", "novanta".

4. Application and Experimental Results

This section reports some experimental results of the application of the proposed methods for the recognition of basic words in the legal amounts extracted from italian bank checks. The automatic bank-check processing is a complex system.

In this paper a prototype consisting of several modules is considered. The main modules integrated into the system are: the image acquisition and pre-processing module, the layout processing module, the courtesy amount recognition module, the legal amount recognition module, the amount validation module and the signature verification module. Each processing module is indipendent and it includes and combines several sub-modules in a hierarchical structure [12].

In particular, the legal amount recognition module, which already contains the algorithms based on HMM approach [13], has been enriched with new software objects that implement the strategies proposed in this paper. Infact, the singularity markers and the sliding window step strategies have been utilized for the recognition of the 48 basic words used in the italian lexicon to represent legal amounts on bank checks. For this reduced lexicon here considered, the basic words are grouped according to the sequence of singularities they contain [12] as shown in Table 1. In particular the singularities are generated by the presence of the following characters: "d", "l", "t", "tt", "ll" that generates ascenders and "q" that generates descenders.

Sing.	***Basic Words***
none	uno, sei, nove
d	due, dieci, undici, sedici, diciannove
t	tre,venti,ventuno,sessanta,sessantuno,novanta, novantuno, cento
q tt	quattro
q	cinque
tt	sette, otto
d d	dodici
t d	tredici
q tt d	quattordici
q d	quindici
d tt	diciassette, diciotto
t tt	ventotto, sessantotto, novantotto
t t	trenta, trentuno
t t tt	trentotto
q t	quaranta, quarantuno, cinquanta, cinquantuno
q t tt	quarantotto, cinquantotto
tt t	settanta, settantuno, ottanta, ottantuno
tt t tt	settantotto, ottantotto
ll	mille
l	mila , milione, unmilione

Table 1. Basic word for italian legal amount production

In this specific domain, $\lambda_k = < A_k, B_k, \pi_k >$ indicates the HMM for the k-th basic word class ($k \in \{1,...,48\}$) and the left-to-right topology is used. Moreover, 30 states for the single HMM and 3 Gaussians components in the mixture for each state have been considered.

The prototyping phase of the recognition system has been realized by Visiquest, an IDE software based on a visual programming language which supports the entire software development cycle [20]. The software objects are dynamically connected to compose a high-modular and easy-to-reuse data-flow diagram. The algorithms grouped into toolboxes are divided into subcategories according to their specific functions. This organization allows a very fast and immediate access to each software component during the development and maintenance phases.

In the experimental phase the two approaches for sliding window step have been considered: the *Constant Step* and the *Progressive Step*. In each case, the experiments have been developed on the three different recognition models (BSW, ABSW, MB) appling no singularity marking procedure and both *Static* and *Dynamic Markers*.

For the experiments a training and a test set database have been used, each one consisting of a collection of 9600 bitmap images (200 black and white images for each basic word) [5].

Table 2 reports the results, on the test set database, obtained using the *Constant Step* strategy for sliding window. The best results, in terms of recognition rate, are obtained with the MB recognition strategy. In particular, both *Static* and *Dynamic Markers* increase the basic word recognition rate when compared to the results obtained with no marker procedure. The *Static Marker* and the *Dynamic Marker* approaches lead to a recognition rate of 82.54% and 83.40%, respectively.

By considering of the marking approaches, the ABSW works better than the BSW, only when no marker procedure is applied. On the contrary, MB is superior to BSW and ABSW whatever approach for singularity marking is considered.

	No Mark.	Static Mark.	Dynamic Mark.
BSW	81.34%	82.06%	82.73%
ABSW	81.69%	81.82%	82.73%
MB	81.83%	82.54%	83.40%

Table 2. Recognition rates: *Constant Step*.

Further improvements in system performance have been obtained by appling the *Progressive Step* strategy to the MB approach. In particular, considering the best results obtained with the *Constant Step* using the MB recognition approach on the different singular marking approaches, the Table 3 reports the recognition rates obtained with the *Progressive Step*. In particular, as Table 3 shows, when MB approach is considered with the *Progressive Step* strategy, the hightest rate is obtained considering the *Dynamic Marker* procedure with a recognition rate of 83.86%.

	No Mark.	Static Mark.	Dynamic Mark.
MB	82.89%	83.12%	83.86%

Table 3. Recognition rates: *Progressive Step*.

5. Conclusions

This paper presents a new investigation in the handwritten word recognition. Together with the pre-existing ones, a new approach for basic word recognition, based on HMM theory and the sliding window technique is developed.

The new approach is based on singularity markers, which aims to evaluate the ascenders and descenders in an handwritten word and to increace the HMM performance.

The experimental results obtained on different recognition approaches are encouraging, in particular the MB approach using the *Dynamic Marker* and the *Progressive Step* for the sliding window, reports the best performance.

6. References

[1] J. A. Bilmes, “A Gentle Tutorial of EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Model”, Technical Report ICSI-TR-97-021, University of Berkeley, 1998.

[2] J.J. Bozinovic and S.N. Srihari, “Off-line cursive word recognition”, IEEE Trans. Pattern Anal., Machine Intell. 11 (1989), 68-83.

[3] R. L. Cave and L. P. Neuwirth, “Hidden markov models for English” in Proc. of the Sym. on the Application of Hidden Markov Models to Text and Speech, ed. J. D. Ferguson, Princeton , 1980, 16-56.

[4] J.W. Cooley, J.W. Tukey, "An algorithm for the machine calculation of complex Fourier series," Math. Comput., Vol. 19, 1965, pp. 297–301.

[5] G. Dimauro, S. Impedovo, R. Modugno, G. Pirlo “A New Database for Research an Banck-check Processing”, Proc. 8th IWFHR, Niagara-on-the-Lake, Ontario, CA, Aug. 6-8 2002, pp. 524-528.

[6] S. Günter, and H. Bunke, “Optimizing the Number of States, Training Iterations and Gaussian in an HMM-Based Handwritten Word Recognizes”, Proc. of 7th ICDAR, Edinburgh, Scotland, Aug. 3-6 2003, pp. 472-476.

[7] D.H. Hubel and T.N. Wiesel, "Brain Mechanisms of Vision", Science, Vol. 29, pp. 130-144, 1982.

[8] S.Impedovo, J. C. Simon, “From Pixels to Feature III” North Holland 1992 pp. 1-515.
[9] S.Impedovo, “Fundamentals in Handwriting Recognition” NATO ASI Series F Vol. 124, pp. 1-494
[10] S. Impedovo, N. Abbattista, B. Marangelli, "Markov Chains as Model for Analysis and Synthesis of Handwritten Numerals", Int. Conf. on Digital Signal Processing, IEEE ASSP Society press, Florence, Italy, Aug. 30- Sept. 2, 1978.
[11] S. Impedovo, N. Abbattista, B. Marangelli, "A Markov Model for Pattern Organization Measurements", Proc. of the XIV Yugoslav International Symposium on Information Processing - INFORMATICA 79, Bled, Yugoslavia, Oct.1-6 1979, pp 1-3.
[12] S. Impedovo, G. Dimauro, G. Pirlo, A. Salzo, "Automatic Bank-Check Processing: A New Engineered Approach", Automatic Bank-Check Processing, S. Impedovo, P. S. P. Wang, H. Bunke editors, Series in Machine Perception and Artificial Intelligence, vol. 28, World Scientific publishing, 1997, pp. 5-42
[13] S. Impedovo, M.G. Lucchese “Improving a Bank-Check Processing System with New HMM-based Algorithms”, Proc. of 5th International Conferece on Applied Informatics and Comunications, Malta, Sep. 15-17, 2005, WSEAS publishing, pp. 207-209.
[14] S. Impedovo, P.S.P. Wang, and H. Bunke, Automatic Bankcheck Processing, World Scientific Publ., Singapore, 1997.
[15] G. Kim, and V. Govindaraju, "Bankcheck recognition using cross validation between legal and courtesy amounts", Automatic Bankcheck Processing, S. Impedovo, P.S.P. Wang and H. Bunke Eds. World Scientific Publ., 1997, pp.195-212.
[16] A. Kundu, Y. He, and P. Bahl, “Recogntion of handwritten word: First and second order hidden Markov model based approach”, Pattern Recogntion J., 22 (1989) 23-297.
[17] R. Plamondon and S. N. Srihari, “Online and offline handwriting recognition: A comprehensive survey” IEEE Trans. Pettern Anal. Machine Intell., vol. 22, no. 1, pp. 63-84, 2000.
[18] L. R. Rabiner, and B. Juang, "An Introduction to Hidden Markov Model" IEEE Signal Processing Magazine, 1986, Vol. 3, pp. 4-16.
[19] L. Rabiner, B, H. Huang, “Fundaments of Speech Recognition”, Prentice Hall, 1993.
[20] J. R. Rasure, and K. Konstantinides, "The Khoros software development environment for image and signal processing", IEEE Transactions on Image Processing, Vol. 3, 1994, pp.243-252.
[21] J.C. Simon, “Off-Line Cursive Word Recognition”, Proc. of the IEEE, Vol. 80, NO. 7, July 1992, pp. 1150-1161
[22] J.C. Simon, and O. Baret, "Cursive words recognition", From Pixels to Features III-Frontiers in Handwriting Recognition, S. Impedovo and J.C. Simon Eds, Elsevier Publ., 1992, pp. 241-260.
[23] C. Y. Suen, “Frontiers in Handwriting Recognition”, Centre for Pattern Recognition & Machine Intelligence, Concordia University, Montreal, Canada, 2-3 April, 1990.
[24] A. Vinciarelli, “A survey on off-line Cursive Word recognition”, Pattern Recognition 35, 2002.
[25] A. Vinciarelli, and J. Luettin, "Off-Line Cursive Script Recognition Based on Continuous Density HMM", Proc. of 7th IWFHR, Sep. 11-13 2000, pp. 493-498.
[26] W. Wang, A. Brakensiek, A. Kosmala, G. Rigoll, “Multi-branch and two-pass HMM modeling approaches for off-line cursive handwriting recognition”, Proc. Sixth International Conference on Document Analysis and Recognition, 2001, pp. 231 – 235.