Journal of Engineering Sciences

Vol 13, Issue 03, MARCH/2022
ISSN NO:0377-9254

Handwritten Text Recognition Through Writer Adaptation

Mr. Pradeep Kumar Mishra,Asst. Professor, Mtech(phd),Department of CSE, pradeepkumar.mishra@sharda.ac.in

Likith Pulluri,Department of CSE,2018010435.likith@ug.sharda.ac.in

Yashwanth Konakanchi,Department of CSE,2018004614.konakanchi@ug.sharda.ac.in

Divyaprabha Gangadharan,Department of CSE,2018005685.divyaprabha@ug.sharda.ac.in

Asima Nayak,Department of CSE,2018016430.asima@ug.sharda.ac.in

ABSTRACT: Handwritten Text Recognition (HTR) remains a challenging problem to date, largely due to the varying writing styles that exist amongst us. Prior works however generally operate with the assumption that there is a limited number of styles, most of which have already been captured by existing datasets. Handwritten text recognition is a problem rarely studied out of specific applications for which lexical knowledge can constrain the vocabulary to a limited one. In the case of handwritten text recognition, additional information can be exploited to characterize the specificity of the writing. This knowledge can help the recognition system to find coherent solutions from both the lexical and the morphological points of view. We present the principles of a handwritten text recognition system based on the on-line learning of the writer shapes. The proposed scheme is shown to improve the recognition rates on a sample of fifteen writings, unknown to the system.

Keywords- *Handwritten Text Recognition (HTR), on-line learning.*

1. INTRODUCTION

Handwritten Text Recognition (HTR) has been a longstanding research problem in computer vision. As a fundamental means of communication, handwritten text can appear in a variety of forms such as memos, whiteboards, handwritten notes, stylus-input, postal automation, reading aid for visually handicapped, etc . In general, the target of automatic HTR is to transcribe hand-written text to its digital content, so that the textual content can be made freely accessible. Handwriting recognition is inherently difficult due to its free flowing nature and complex shapes assumed by characters and their combinations. Torn pages, and warped or touching lines also make HTR more challenging. Most importantly however, handwritten texts are diverse across individual handwritten styles where each style can be very unique, while some might prefer an idiosyncratic style of writing certain characters like ‘G’ and ‘Z’, others may choose a cursive style with uneven line-spacing. Modern deep learning based HTR models mostly tackle these challenges by resourcing to a large amount of training data. The hope is that most style variations would have already been captured

because of the data volume. Albeit with limited success, it had become apparent that these models tend to over-fit to styles already captured, but generalizing poorly to those unseen.

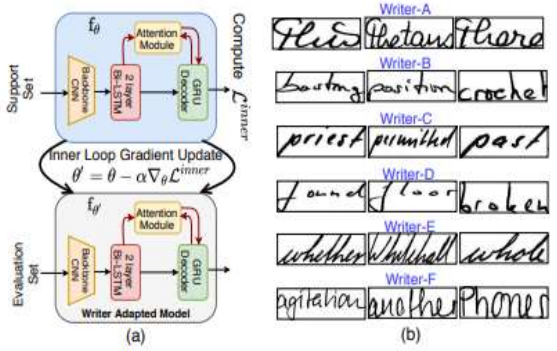

Fig.1: (a) During inference, exploits additional handwritten images of a specific writer through support set, and gives rise to a writer adapted model via single gradient step update. (b) Varying styles across different writers

Off-line handwriting recognition has given rise to numerous studies until now. Despite constant efforts, one can notice how difficult the task is when considered out of restricted applications such as bank cheques, postal addresses or form processing. For these reasons, few studies have dealt with handwritten text recognition until now. Recently, some studies have been published which introduce a syntactical analysis as a post-processing stage using language models [5,8]. From a methodological point of view, Lorette et al. [6] set the difficulty of handwritten text recognition as a scene analysis problem.

2. LITERATURE REVIEW

2.1 A Survey of Methods and Strategies in Character segmentation [1] :

R. Casey, E. Lecolinet et al., provided that Character segmentation has long been a critical area of the OCR process. The higher recognition rates for isolated characters vs. those obtained for words and connected character strings well illustrate this fact. A good part of recent progress in reading unconstrained printed and written text may be ascribed to more insightful handling of segmentation. This paper provides a review of these advances. The aim is to provide an appreciation for the range of techniques that have been developed, rather than to simply list sources. Segmentation methods are listed under four main headings. What may be termed the "classical" approach consists of methods that partition the input image into subimages, which are then classified. The operation of attempting to decompose the image into classifiable units is called "dissection." The second class of methods avoids dissection, and segments the image either explicitly, by classification of prespecified windows, or implicitly by classification of subsets of spatial features collected from the image as a whole. The third strategy is a hybrid of the first two, employing dissection together with recombination rules to define potential segments, but using classification to select from the range of admissible segmentation possibilities offered by these subimages. Finally, holistic approaches that avoid segmentation by recognizing entire character strings as units are described.

2.2 Maximum Likelihood from incomplete data via the EM algorithm [2]:

A. P. Dempster, N. M. Laird, D. B. Rubin et al., said that A broadly applicable algorithm for computing maximum likelihood estimates from incomplete data is presented at various levels of generality. Theory showing the monotone behaviour of the likelihood

and convergence of the algorithm is derived. Many examples are sketched, including missing value situations, applications to grouped, censored or truncated data, finite mixture models, variance component estimation, hyperparameter estimation, iteratively reweighted least squares and factor analysis.

2.3 An HMM-Based Approach for Off-Line Unconstrained Handwritten Word Modeling and Recognition [3] :

A. El-Yacoubi, M. Gilloux, R. Sabourin, C.Y. Suen et al., Describes a hidden Markov model-based approach designed to recognize off-line unconstrained handwritten words for large vocabularies. After preprocessing, a word image is segmented into letters or pseudoletters and represented by two feature sequences of equal length, each consisting of an alternating sequence of shape-symbols and segmentation-symbols, which are both explicitly modeled. The word model is made up of the concatenation of appropriate letter models consisting of elementary HMMs and an HMM-based interpolation technique is used to optimally combine the two feature sets. Two rejection mechanisms are considered depending on whether or not the word image is guaranteed to belong to the lexicon. Experiments carried out on real-life data show that the proposed approach can be successfully used for handwritten word recognition.

2.4 An Architecture for Handwritten Text Recognition System [5]:

G. Kim, V. Govindaraju, S. N. Srihari et al., said that This paper presents an end-to-end system for reading handwritten page images. Five functional modules included in the system are introduced in this paper: (i) pre-processing, which concerns introducing an image representation for easy manipulation of large page images and image handling procedures using the image representation; (ii) line separation, concerning text line detection and extracting images of lines of text from a page image; (iii) word segmentation, which concerns locating word gaps and isolating words from a line of text image obtained efficiently and in an intelligent manner; (iv) word recognition, concerning handwritten word recognition algorithms; and (v) linguistic post-processing, which concerns the use of linguistic constraints to intelligently parse and recognize text. Key ideas employed in each functional module, which have been developed for dealing with the diversity of handwriting in its various aspects with a goal of system reliability and robustness, are described in this paper. Preliminary experiments show promising results in terms of speed and accuracy.

2.5 A full English Sentence Database for Off-Line Handwriting Recognition [8]:

U. V. Marti, H. Bunke, et al discussed In this paper we present a new database for off-line handwriting recognition, together with a few preprocessing and text segmentation procedures. The database is based on the Lancaster-Oslo/Bergen(LOB) corpus. This corpus is a collection of texts that were used to generate forms, which subsequently were filled out by persons with their handwriting. Up to now (December 1998) the database includes 556 forms produced by approximately 250 different writers. The database consists of full English sentences. It can serve as a basis for a variety of handwriting recognition tasks. The main focus, however, is on recognition techniques that use linguistic knowledge

beyond the lexicon level. This knowledge can be automatically derived from the corpus or it can be supplied from external sources.

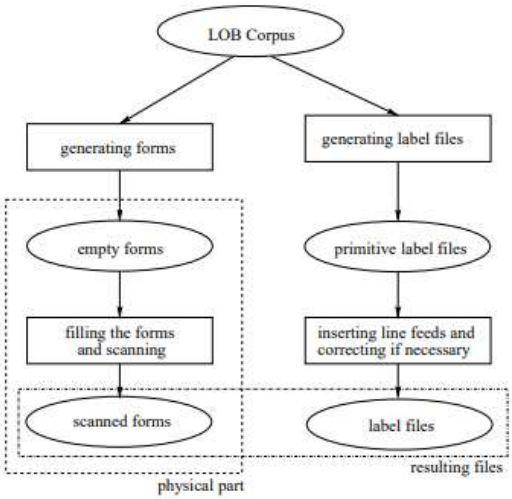

Fig.2: Database acquisition image

2.6 Defining writer's invariants to adapt the recognition task [9]:

A. Nosary, L. Heutte, T. Paquet and Y. Lecourtier et al., said that Investigates the automatic reading of unconstrained omni-writer handwritten texts. This paper shows how to endow the reading system with adaptation faculties for each writer's handwriting. The adaptation principles are of major importance for making robust decisions when neither simple lexical nor syntactic rules can be used, e.g. for a free lexicon or for full text recognition. The first part of this paper defines the concept of writer's invariants. In the second part, we explain how the recognition system can be adapted to a particular handwriting by exploiting the graphical context defined by the writer's invariants. This adaptation is guaranteed, thanks to the writer's invariants, by activating interaction links over the whole text between the recognition procedures for word entities and those for letter entities.

3. IMPLEMENTATION

In the framework of handwritten text recognition, we have so far favoured a batch and unsupervised approach. To adapt our recognition system to the handwriting, it is necessary to have at our disposal the models of the writer allographs. The adaptation scheme thus roughly relies on the two steps of the EM algorithm:

1. The first one is an estimation step of the missing data and allows to label the observations as characters or rejects;
2. The second one allows to re-estimate the character models.

The omniwriter word recognition system we have presented allow to initialize the system not too far from the desired solution in order to converge towards a better recognition solution. This is the condition for bootstrapping the adaptation system. Figure 3 summarizes the whole organization of the system. The initial omni-writer system uses intrinsic morphological knowledge (IMK), i.e. the feature vector extracted on each grapheme, to propose a first interpretation (ISK). The unsupervised labelling of the graphemes by means of recognition results produces a contextual interpretation (CSK) that allows to learn the writer references by means of the invariants (CMK). The last important point that must be specified concerns the learning of the writer references.

The adaptation scheme of the recognition system we propose is thus justified by the existence of such a contextual information on the segmented patterns that

can act as a morphological constraint, provided that we are able to break with interpretation independence. This concerns the adaptation of models, but it is also possible to adapt the input representations (observations). In the framework of our study, the writer invariants allow to consider an adaptation of the system to the input representations.

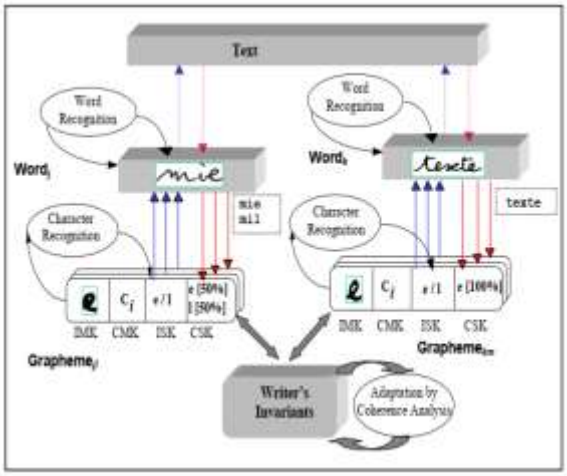

Fig.3: System architecture.

We have chosen to work in the writer space defined by the sets of invariant patterns that can be detected on each handwritten page. Indeed, each invariant pattern defines a neighborhood adapted to each grapheme and allows to infer new interpretations. This point needs to be highlighted in the framework of our approach by explicit segmentation since it enables to distinguish the rare representations from the frequent representations for which the adaptation can be performed in a robust manner. As for character modelling, we have chosen a non parametric approach because the re-estimation of parametric models might not fit properly in some cases, especially when the number of samples available is limited in a handwriting page. Therefore, in this study, we propose a writer adaptation of the input representations (observations) as well as of the character models.

5. METHODS

A. Handwritten Word Recognition:

One can distinguish segmentation-based analytical approaches and segmentation-free approaches. The latter has the advantage of integrating segmentation during the learning phase. But many problems still remain for off-line segmentation free approaches. Among them, the choice of a convenient 1D representation of a 2D input bitmap is still an open issue. Therefore, few studies have been proposed yet. On the contrary, in the case of on-line handwriting recognition, segmentation free approaches are widely used. Indeed, the temporal 1D representation of the on-line signal is well-suited to segmentation-free approaches, as well as for speech recognition.

B. Adaptation:

The word recognition system presented above has been designed without taking into account the distinctive handwriting features of the writer. We have introduced in an earlier study the concept of writer's invariants, which can be defined as the redundant elementary patterns extracted from writer's handwriting.

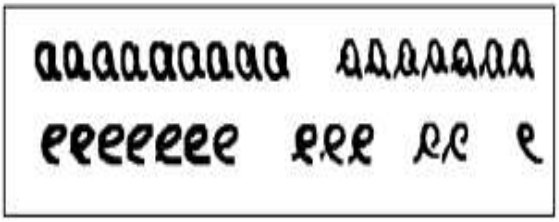

Fig.4: Samples of invariant clusters extracted from a handwritten page

Figure 4 gives an example of the clusters that we obtained using an adapted clustering method. As one can see, the proposed method succeeds in determining regular patterns that occur in the text.

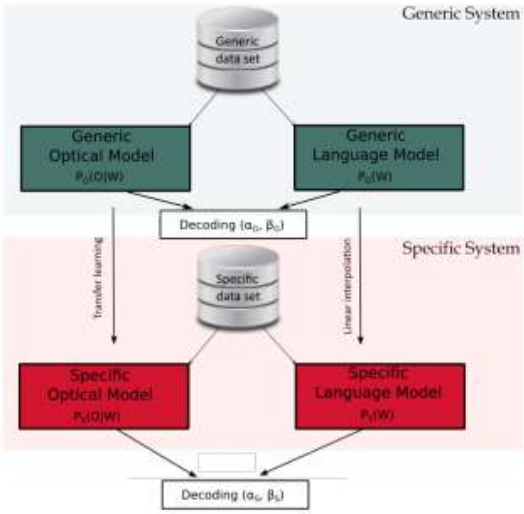

Fig.5: writer adaptation

C. Word Recognition:

Segmentation free approaches rely mainly on Hidden Markov Models or recurrent neural networks to modelize cursive characters, whereas most segmentation based approaches use either Hidden Markov Models or dynamic programming to find the best segmentation hypothesis. In this case, characters are described in a static fixed size representation space using the most invariant features that can be constructed among the various writing styles. Our previous work has motivated the choice of this approach mainly for rapid prototyping purposes in order to assess our adaptation strategy to handwriting.

5. EXPERIMENTAL RESULTS

It is rather difficult to evaluate practically the adaptation scheme we propose since its implementation requires a complete page of handwriting. Now, it is well known that a precise localization of the handwriting information in a handwritten page is still a very sensitive step because of the potential variability in the handwriting line and word layout (inclination, alignment, spacing, superposition, ...). The line segmentation process has been controlled manually so as to test the adaptation scheme only on the correct word images. The adaptation principle has been evaluated on the 15 texts of the test database we have presented above.

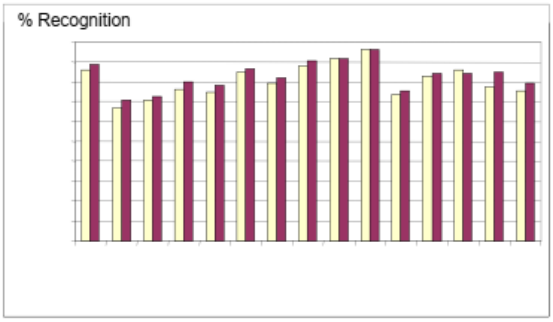

Fig.6: Adaptation results per writer.

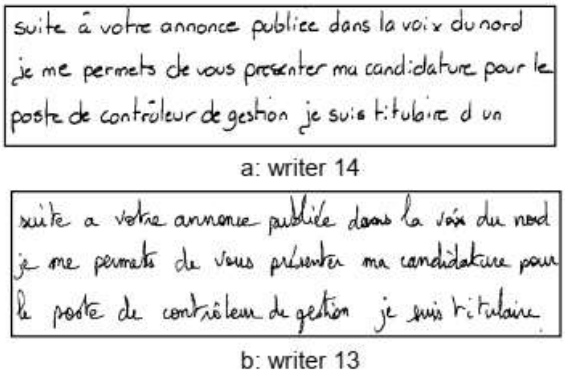

Fig.7: Samples of handwritings: (a) writer 14; (b) writer 13

Figure 7 presents two samples extracted from two handwritings. The first sample (Figure 7.a) is extracted from writer 14, which can be qualified as script type, regular with many natural separations

between letters. The second sample (Figure 7.b) is extracted from a handwriting on which the adaptation did not permit performance improvements (writers 13). In fact, this handwriting is quite variable: there is not enough regularity to exploit. Therefore, letter allographs constitute groups of singletons as a result of the invariants determination.

The precise analysis of these results, which cannot be developed here, shows a direct correlation of the adaptation performance with the quality of the writing eg. better improvement for stable handwritings (cf. Figure 8). For this purpose, the quality of handwritings has been quantified thanks to the variability measure of a handwritten text T defined in [9], say E(T). This entropy based measure takes small values on stable handwritings while unstable handwritings have their variability measure close to 1.

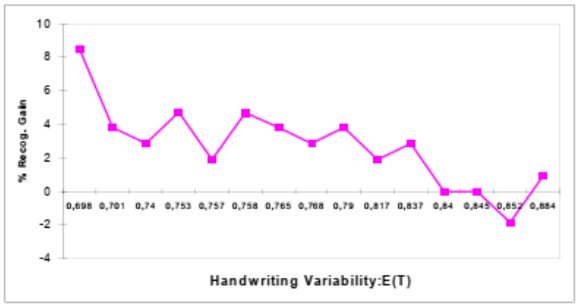

Fig.8: Correlation of adaptation performance with the variability measure of handwriting E(T).

6. CONCLUSION

In this communication, we have presented a new principle of handwritten text recognition based on writer adaptation. This adaptation is realized by iterating word recognition steps that allow to label the writer representations (allographes) on the whole text, and reestimation of character models. The originality of this approach relies in the ability of the system to learn in an unsupervised manner the writer specificities. The most particular point in this unsupervised learning principle lies in the control of lexical decisions, which directly influence the knowledge inferred at the level of writer allographs. This last point is beyond the scope of our future research.

REFERENCES

[1] R. Casey, E. Lecolinet, A Survey of Methods and Strategies in Character segmentation, IEEE-PAMI, Vol. 18, pp. 690-706, 1996.

[2] A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum Likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society, 39: pp. 1-39, 1977.

[3] A. El-Yacoubi, M. Gilloux, R. Sabourin, C.Y. Suen, An HMM-Based Approach for Off-Line Unconstrained Handwritten Word Modeling and Recognition, IEEE-PAMI, Vol. 21, No. 8, pp. 752-760, August 1999.

[4] L. Heutte, T. Paquet, J.V. Moreau, Y. Lecourtier, C. Olivier. A structural/statistical feature based vector for handwritten character recognition. Pattern Recognition Letters, Vol. 19, No. 7, pp. 629-641, 1998.

[5] G. Kim, V. Govindaraju, S. N. Srihari, An Architecture for Handwritten Text Recognition System, IJDAR, Vol. 2, pp. 37-44, 1999.

[6] G. Lorette, Y. Lecourtier, Reconnaissance de textes manuscrits hors-ligne : un problème d'analyse de scène ?, Bigre 80 , CNED'92, pp.109-135, Nancy, juillet 1992. -4 -2 8 6 4 2 0 10 0,698 0,701 0,74 0,753

0,757 0,758 0,765 0,768 0,79 0,817 0,837 0,84 0,845 0,852 0,884 Handw riting Variability:E(T) % Recog. Gain

[7] Y. Lu, M. Shridhar, Character Segmentation in Handwritten Words-an overview, Pattern Recognition, Vol. 29, No. 1, pp. 77-96, 1996.

[8] U. V. Marti, H. Bunke, A full English Sentence Database for Off-Line Handwriting Recognition, ICDAR'99 , pp. 705-708, Bengalore, 1999.

[9] A. Nosary, L. Heutte, T. Paquet and Y. Lecourtier, Defining writer's invariants to adapt the recognition task, Proc. ICDAR'99, India, pp. 765-768, 1999.

[10] G. Kim, V. Govindaraju, Handwritten Phrase Recognition as Applied to Street Name Recognition, Pattern Recognition, Vol. 31, No. 1, pp. 41-51, 1998.

[11] C. Viard-Gaudin, P. M. Lallican, S. Knerr, P. Binter, The IREST On/Off (IRONOFF) dual handwriting database, ICDAR'99, pp. 455- 458, Bangalore, 1999.

[12] J. luc Gauvain and C. hui Lee, "Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chains," IEEE Transactions on Speech and Audio Processing, vol. 2, pp. 291–298, 1994.

[13] K. Ait-Mohand, T. Paquet, and N. Ragot, "Combining structure and parameter adaptation of hmms for printed text recognition," IEEE PAMI, vol. 36, no. 9, pp. 1716–1732, 2014.

[14] I. Marosi, "Industrial ocr approaches: architecture, algorithms, and adaptation techniques," in Proc. SPIE. 6500, Document Recognition and Retrieval XIV. SPIE, 2007, pp. –.

[15] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," IEEE PAMI, vol. 39, no. 11, pp. 2298–2304, 2017.