Multi-agent Systems for Arabic Handwriting Recognition

Youssef Boulid*, Abdelghani Souhar and Mohamed Elyoussfi Elkettani

University Ibn Tofail, Faculty of science, 14000, Kenitra (Morocco)

uniR
LA UNIVERSIDAD
EN INTERNET

Abstract

This paper aims to give a presentation of the PhD defended by Boulid Youssef on December 26th, 2016 at University Ibn Tofail, entitled "Arabic handwritten recognition in an offline mode". The adopted approach is realized under the multi agent paradigm. The dissertation was held in Faculty of Science Kénitra in a publicly open presentation. After the presentation, Boulid was awarded with the highest grade (Très honorable avec félicitations de jury).

Keywords

I. Introduction

On December 26th, 2016, Boulid Youssef defended his PhD thesis related with Arabic handwritten recognition [1]. The thesis was supervised by Prof. Mohamed Elyoussfi and co- supervised by Prof. Abdelghani Souhar. The assessing committee of the PhD dissertation was composed of Prof TOUAHNI Raja, Prof SADIQ Abdelalim, Prof AIT KERROUM Mounir, Prof BENATTOU Mohammed from Faculty of Science, Kénitra and Prof TABII Youness from National School of Applied Sciences, Tétuan. The thesis has been read and approved by this committee. All of them were present at the presentation. The main publications associated with the PhD thesis are [2-5].

II. Thesis Summary

Handwritten recognition is a very broad subject of research and depending on the quality of the document to recognize, there is a multitude of problems that can be encountered. The pattern recognition process is often adopted in the design of handwriting recognition systems, which consists mainly of four stages: the pre-processing stage, which concerns the preparation of the document in terms of normalization and suppression of noise, the segmentation stage, which concerns the detection of lines, words and then the segmentation of those words into characters. The third stage concerns the feature extraction, in which the designer must choose or design the adequate characteristics to be extracted from the character that allow to minimize the intra-class variance while maximizing the interclass variance.

The fourth stage involves learning and testing, in which a learning algorithm is used to recognize new letters or new words based on those already learned. To these stages is added a post-processing stage which concerns the verification of the recognized words using a lexical and semantic analysis.

Several researchers propose techniques that respond under certain conditions to a precise problem of a given stage of the process. While the major problem lies in the collaboration between these different techniques since the followed process is often executed in a sequential manner, which is a handicap since the errors in the first stages will propagate in the following stages and thus influence the result of the recognition.

As for a human reader, he has several strategies when facing a document, such as: complete reading (word-by-word reading), the inspection (searching in specific regions of the document), and the overview of the document which give him the capability to read documents he has not seen before.

From this point of view we are interested in analyzing the problems of the recognition of handwritten document by taking inspiration from the mechanisms of what we think the human reader uses during the reading process. This problem is modeled under the multi-agent systems paradigm while taking into consideration the specific characteristics of the Arabic language.

In this context, the contribution of the thesis concerns the recognition of handwritten Arabic documents and precisely the pre-processing, the line segmentation and the character recognition stages [1].

Generally, there are two ways for document noise removal; either detection and suppression of noise, this is possible when the patterns of the noise have independent characteristics that could differentiate them from the textual content, or extraction of textual content while ignoring the noise, in this case the contextual information and the prior

---

* Corresponding author.

E-mail address: y.boulid@gmail.com

knowledge about the text are required.

For the problem of text line segmentation, generally there are three approaches: the first one focuses on the regions separating the text lines, the second one searches the connected component that constitutes the lines, while the third one searches for the baseline of each word and regroup those that participate in the same line.

Feature extraction methods could be classified in two categories: the structural features, that extract geometrical and topological properties such as the number and position of diacritical points, number of connected components, presence of loops, orientation of curves, location of intersections; and the statistical features such as histograms of projection profile of transition, moments, histogram of gray level distribution, Fourier descriptor, freeman chain code…etc.

Based on the mechanisms of scanning, which humans can use when reading a document, we have divided the problem of noise removal into two collaborative agents. The first one is responsible for the estimation of global parameters of the document and the creation and affectation of noise removal agents into different regions of the document. Based on the nature of the Arabic script, we have found that the notion of intersection (pixel position between characters in a cursive word) could be used in a manner that allows us to distinct between textual and non-textual content. The percentage of intersections according to the total area of the component is higher than 50% in the case of noise. The treatments that the agents execute are: the suppression of salt and pepper noise based on the average stroke width of the text, the classification of content into textual and non-textual based on the percentage of intersections and finally, the distinction of noise similar to text using the contextual information [2].

Based on the mechanisms of reading word after word, we have modeled the problem of line segmentation as an agent based on utility that integrates the Markov Decision Process. The proposed approach detects the connected component in the same line by using knowledge about characteristics and disposition of the components in the document [3].

Problems may occur when adjacent line are touching due to narrow gaps between them, where we may find words that belongs to different lines, but are linked together. To overcome this and inspired by the mechanisms of perception involved in the process of reading, we have modeled the problem as three collaborative agents. The first one is responsible for the estimation of global parameters of the document and for the line extraction. The second one is responsible for the detection of the first component in the line and the detection of components that belong to the same line. As to the third one is responsible for splitting and segmentation of touching characters and words [4].

Based on the fact that Arabic is written from right to left, we have found that extracting features from the right portion rather than from the whole character's image allows enhancing the recognition rate. The Arabic script uses the information about the baseline to differentiate between some character having similar shapes. Once the word is correctly segmented, we recognize each one of its characters, but when the character is isolated we no longer have the information about the baseline. Also, the existing datasets for Arabic handwritten letters do not integrate this information. Extracting textual and structural features after a proper decomposition of the character allows us to increase the recognition rate. This solution allows compensating the lack of the baseline information [5].

To overcome the problem that resides in the traditional approaches, which is the use of the phases in the recognition process in a sequential manner, we have proposed an agent-based modeling offering the possibility to implement different strategies of human reading. According to the homogeneity of the document, this latter is divided into regions, where each region contains a set of agents from different level of the recognition process that could collaborate locally and also between different regions. Each agent has a memory allowing him to track different actions that he performs and the possibility to undo an action in order to correct it if necessary.

The obtained results are encouraging, although we are still in the preliminary stages of the design of a handwritten Arabic recognition system achieving at least the human performance.

Finally, the need of a platform, allowing collaboration between the different stages of the recognition process, is necessary. We believe that such a platform should be based on multi-agent systems offering the possibility of implementing and integrating the different recognition stages in parallel.

References

[1] Boulid, Y. Arabic handwritten recognition in an offline mode. PhD thesis, Faculty of Science Kénitra, University Ibn Tofail, December 2016.

[2] Boulid, Y., & Elkettani, M. Y. (2014). Approach for Arabic Handwritten Image Processing: Case of Text Detection in Degraded Documents. International Journal of Computer Applications, Vol. 101, N° 14, Pages 35-42.

[3] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2016). Detection of Text Lines of Handwritten Arabic Manuscripts using Markov Decision Processes. International Journal of Interactive Multimedia and Artificial Intelligence, (Special Issue on Artificial Intelligence Underpinning), Vol. 4, Nº1 Pages 31-36.

[4] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2016). Segmentation approach of Arabic manuscripts text lines based on multi agent systems. International Journal of Computer Information Systems and Industrial Management Applications, Vol. 8, Pages 173 - 183.

[5] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2017). Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language. International Journal of Interactive Multimedia and Artificial Intelligence, (Regular Issue), Vol. 4, Nº4 Pages 45-53.

Youssef Boulid

Youssef Boulid received his M.S. degree in Decision Support Systems and Project Management in 2012 and PhD degree in Computer Science in 2016 from University Ibn Tofail, Faculty of science, Kenitra- Morocco. Now he works as a researcher at Maxware Technology Kénitra-Morocco. His research interests include image processing, handwritten document analysis, Arabic handwritten recognition and artificial intelligence.

Abdelghani Souhar

Abdelghani Souhar received M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from University Mohammed 5 in Rabat - Morocco. Now he is a Professor at university Ibn Tofail in Kenitra - Morocco. His research interests include CAE, CAD and Artificial Intelligence.

Mohamed Elyoussfi Elkettani

Mohamed Elyoussfi Elkettani received M.S. degree in applied mathematics in 1980 and PhD degree in Statistics in 1984 from Orsay Faculty of Science, University of Paris XI. Now he is a Professor at university Ibn Tofail in Kenitra - Morocco. His research interests include Multivariate statistics and Image recognition algorithms.