Knowledge-Based Systems 24 (2011) 680–688

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Performance of hidden Markov model and dynamic Bayesian network classifiers on handwritten Arabic word recognition

Jawad H. AlKhateeb [a,*], Olivier Pauplin [b], Jinchang Ren [c], Jianmin Jiang [b]

[a] Faculty of Science and Information Technology, Al-Zaytoona University of Jordan, Amman, Jordan
[b] School of Informatics, University of Bradford, Bradford BD7 1DP, United Kingdom
[c] Centre for excellence in Signal and Image Processing, Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, United Kingdom

ARTICLE INFO

Article history:
Received 6 May 2010
Received in revised form 26 January 2011
Accepted 9 February 2011
Available online 15 February 2011

Keywords:
Off-line handwritten recognition
Hidden Markov model (HMM)
Dynamic Bayesian network (DBN)
Performance evaluation
IFN/ENIT database

ABSTRACT

This paper presents a comparative study of two machine learning techniques for recognizing handwritten Arabic words, where hidden Markov models (HMMs) and dynamic Bayesian networks (DBNs) were evaluated. The work proposed is divided into three stages, namely preprocessing, feature extraction and classification. Preprocessing includes baseline estimation and normalization as well as segmentation. In the second stage, features are extracted from each of the normalized words, where a set of new features for handwritten Arabic words is proposed, based on a sliding window approach moving across the mirrored word image. The third stage is for classification and recognition, where machine learning is applied using HMMs and DBNs. In order to validate the techniques, extensive experiments were conducted using the IFN/ENIT database which contains 32,492 Arabic words. Experimental results and quantitative evaluations showed that HMM outperforms DBN in terms of higher recognition rate and lower complexity.

1. Introduction

Handwriting recognition (HWR) is a mechanism for transforming the written text into a symbolic representation, which plays an essential role in many human–computer interaction applications including cheque verification, automatic mail sorting, office automation as well as natural human–computer interaction [1]. HWR for Latin languages has been conducted and significant achievements have been made. However, there has been less work in Arabic handwriting recognition. This is due to the complexity of the Arabic language and lack of public Arabic handwriting databases. In general, HWR can be categorized into two distinct types: online and off-line based systems. Recognition in online systems uses the dynamics of writing by following the pen movement. Recognition in off-line based systems is based solely on an image of the written text. Online recognition is easier because it can make use of the additional information not available to the off-line systems such as the strength and sequential order of the writing [2]. However, online recognition is not possible in many applications so in this paper, we focus on the off-line recognition of handwritten Arabic text.

The recognition of handwritten Arabic scripts can be divided into segmentation based or segmentation free approaches. The former segments words into characters or letters for recognition and can be regarded as an analytical approach. The latter, which can be regarded as a global approach, takes the whole word image for recognition and therefore needs no segmentation. Although the global approach makes the recognition process simpler, it requires a larger input vocabulary than analytical approach [3].

This paper focuses on the Arabic handwritten word recognition phase and introduces new methods for extracting features. Several experiments have been conducted using the IFN/ENIT benchmark database [4] and our algorithm showed the best recognition rate among the existing work reported using the same database. The remainder of this paper is structured as follows: Section 2 presents the literature review while Sections 3 and 4 describe the proposed method in terms of pre-processing and feature extraction; Section 5 describes the HMM classification process in details; experimental results are presented in Section 6. The paper ends with conclusions and suggestions for further work.

2. Literature review

Khorsheed and Clocksin [5] presented a technique for the word can be recognized as single unit which depends on a predefined lexicon. Using the skeleton of the word based on the Stentiford's algorithm [6], all segments were extracted for recognition into feature vector. The extracted the structural features from Arabic cursive text in three consecutive steps: segment extraction, loop extraction and segment transformation. A 8-dimensional feature

* Corresponding author.
E-mail addresses: jawadalkhatib@yahoo.com, jawad.kh@alzaytoonah.edu.jo (J.H. AlKhateeb), o.pauplin@bradford.ac.uk (O. Pauplin), Jinchang.ren@strath.ac.uk (J. Ren), j.jiang1@bradford.ac.uk (J. Jiang).

0950-7051/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2011.02.008

vector was created for each segment. Using vector quantization (VQ) [7], each vector was mapped to the nearest symbol in the codebook resulting in a sequence of observation which is fed into HMM. The Viterbi algorithm [8] is used to form a codebook of 76 symbols by apportioning the training samples into several classes. The technique was tested with a lexicon of 294 words acquired from a different text sources using the HMM. Recognition rates of up to 97% were achieved.

Khorsheed [9] presented another holistic recognition system for recognizing Arabic handwritten words. Pre-processing tasks performed included using the Zhang-Suen thinning algorithm [10] to generate the skeleton graph. Structural features for the handwritten script were extracted after skeletonization by decomposing the word skeleton into a sequence of links with an order similar to the word writing order. Using the line approximation [6], each line was broken into small line segments, which were transferred into a sequence of discrete symbols using VQ [7]. Then an HMM recognizer was applied with image skeletonization to the recognition of an old Arabic manuscript which can be found in [11]. One HMM was performed using 296 states on the 32 character models. Each model is left to right HMM with no restriction jump margin. The system was tested on 12,960 recognition tests associated with 405 character samples of a single font extracted from the single manuscript. The recognition rates achieved was 72% without spelling check and 87% with spelling check.

Pechwitz and Maergner [12] presented an off-line recognition system for the Arabic isolated handwritten words. They validated their system using the IFN/ENIT benchmark database [4] which consists of four sets (*a*, *b*, *c*, and *d*). They used a sliding window based on the image representation of the word image using pixel values as main features. The sliding window is shifted across the word image from right to left and generates the feature vector. The word image is a gray normalized image. The Karhunen Loeve Transformation (KLT) is performed in order to reduce the feature vector dimension. They used Semi Continuous HMMs (SCHMM) classifier for recognition. They used sets *a*, *b*, and *c* for training and set *d* for testing their system. The recognition rate achieved was 89%.

El Abed and Margner [13] presented an Arabic isolated handwritten word recognition based on HMM. They used the sliding window approach for extracting the pixel features. They used the skeleton direction based feature extraction technique where each word image was splitting into uniform vertical frames and each word image was split into five horizontal zones with equal height. The lengths of all lines in each zone frame were calculated in four directions to form a 20 dimensional feature. They used HMM for recognition and they validated their system using the IFN/ENIT benchmark database [4]. They used sets *a*, *b*, and *c* for training and set *d* for testing their system. The recognition rate achieved was 89.1% for top 1 and 96.4% for top 10.

In [14], a similar sliding window is used in such a window, the word image is divided into vertical overlapping frames with a constant width and variable heights. They validated their system using the IFN/ENIT benchmark database [4]. The sliding window is shifted from right to left and a feature vector is calculated for each frame. For each frame, 24 features are extracted using foreground pixel densities and concavity features. In addition there are 15 baseline independent features. They used HMMs classifier for recognition based on character modeling. Each character has a left right topology. Their HMM model had four states for each character model resulting 159 character models in total. They used sets *a*, *b*, and *c* for training and set *d* for testing their system. The recognition rate achieved was 75.41%.

Al-Hajj et al. [15] presented a two stage system for recognizing Arabic handwritten words. The first stage system was based on three HMM based classifiers which used pixel features in [14]. Each HMM classifier produces the best ten candidates (top 10) based on the likelihood. The second stage combined the three HMM classifiers by fusing the candidates provided by the HMM classifiers. They used three schemes for combining the classifiers: the sum rule, the majority vote rule, and the neural network based combining classifier. Different combinations were experimented using the IFN/ENIT benchmark database [4] and the recognition rate achieved was 90.96%.

Al-Hajj Mohamad et al. [16] presented an off-line recognition system for handwritten Arabic words of city names using the IFN/ENIT benchmark database [4] based on HMM. They used the sliding approach for extracting the features [14]. Their system relies on combining three homogeneous HMM classifiers in order to increase the system performance. They used the same three schemes for combining the classifiers in [15]. The recognition rate achieved was higher than 90% depends on the candidates. It is 90.26% for top 1, 94.71 for top 2, and 95.68% for top 3. It is important to mention that this system used sets *a*, *b*, and *c* for training and set *d* for testing.

Benouareth et al. [17,18] presented an off-line unconstrained handwritten Arabic word recognition based on semi-continuous hidden Markov models (SCHMMs) with explicit state duration. Statistical and structural features were utilized on the basis of the adopted segmentation in which implicit word segmentation is used to divide images into vertical frames of constant and variable width for feature extraction. Based on maxima and minima analysis of the vertical projection histogram, morphological complexity of the Arabic handwritten characters is further considered. They used SCHMM for recognition and they validated their system using the IFN/ENIT benchmark database [4]. They used sets *a*, *b*, and *c* for training and set *d* for testing their system. The recognition rate achieved with uniform segmentation was 81.02.1% for top 1 and 91.74% for top 10. The recognition rate achieved with non uniform segmentation was 83.791% for top 1 and 92.12% for top 10.

Likforman-Sulem and Sigelle [19,20] presented a new approach for off-line printed character recognition based on DBN. Their model consists of coupling two HMMs in various DBN architectures. The image rows and image columns of the coupled HMMs were used as the main observations. Their system has been evaluated using various DBN architectures and achieved a recognition rate of 98.3% with the vertical HMM, and 93.7% with the horizontal HMM. However, when testing degraded letters the recognition rate went down such as 93.8% with the vertical HMM and 88.1% with the horizontal HMM.

In this paper, we proposed an off-line recognition system for the handwritten Arabic cursive using HMM. We split the description of the system into three stages: preprocessing, feature extraction, and classification.

3. Preprocessing

The main aim of the preprocessing is to enhance the inputted signal and to represent it in a way which can be measured consistently for robust recognition. Here the preprocessing stage involves scanning the paper document, removing noise, binarizing the images, segmenting lines and words and estimating baselines. These steps are strongly dependent on the quality of the paper document. As samples of words provided in the IFN/ENIT database have been manually separated and binarized during the development stage [4], the only processes still needed are estimation of the baseline and normalization. Although not needed here, we have investigated how to generally segment words and this and our technique to estimate the baseline can be found in [21].

In an ideal handwriting model, the word has to be written in a horizontal way with both ascenders and descenders aligned along the vertical direction. These conditions in real data are rarely