Pattern Recognition Letters 35 (2014) 113–119

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Off-line hand written input based identity determination using multi kernel feature combination

Ehtesham Hassan [a,*], Santanu Chaudhury [b], Nivedita Yadav [a,b,c,d], Prem Kalra [c], M. Gopal [d]

[a] Department of Electrical Engineering, IIT Delhi, India
[b] School of Information Technology, IIT Delhi, India
[c] Department of Computer Science and Engineering, IIT Delhi, India
[d] School of Engineering, Shiv Nadar University, Dadri, UP, India

ARTICLE INFO

Article history:
Available online 16 May 2013

Keywords:
Multiple kernel learning
Writer identification
Writer verification

ABSTRACT

The paper presents a novel framework for the application of multiple features for handwritten data based identity recognition. Different types of features characterise different facets of the handwriting. We have designed a scheme for multiple feature based identity establishment using multi-kernel learning. A new formulation for multi-kernel learning using genetic algorithm has been presented. The efficacy of the framework using individual and combination of features is demonstrated for Devanagari script input.

1. Introduction

The paper presents novel identity determination framework with off-line handwritten data by the application of multiple features. The framework presents novel multiple kernel learning (MKL) formulation for the identity determination problem. A set of novel features are introduced, which are optimally combined by the proposed MKL formulation in kernel space for improving the identification results. The paper considers both the aspect of identity determination, i.e., identification and verification. The proposed MKL formulation is applied for both the problems where a novel Genetic algorithm based solution is presented for solving the MKL. The application of the proposed framework is demonstrated on handwritten Devanagari script input which has not been explored much in the literature (Shaw et al., 2008).

The handwritten input based identification is a challenging task due to the possible variations in directions and shapes of constituent strokes of writing samples. In addition, the problem is compounded by within class variations due to different writing conditions. The classifier combination have been the preferred approach to handle the complexity of the problem (Bajaj et al., 1997; Garain et al., 2009; Bartolini et al., 2010). For Latin and other Asian script writer recognition, many feature representations and decision models have been proposed (Siddiqi et al., 2009; Srihari et al., 2001; Zois et al., 2000). In (Bulacu et al. (2007)), Bulacu and Schomaker have presented extensive study on combination of texture and allographic features for Latin script writer verification and identification. Srinivasan et al. (2007) demonstrated application of orientation, structural, statistical and morphological features for writer identification. Schlapbach and Bunke (xxxx) presented GMM and HMM based modelling for identification. The recent work by Behzad and Mohsen Helli et al. (2010) proposed graph based feature representation computed over the Gabor filter responses for Persian writer identification. Siddiqui and Vincent Siddiqi et al., 2010 presented the comparative analysis of codebook based and contour base features for writer recognition. The codebook based features exploits the primitive level redundancy in the written text and identifies the writer invariant equivalence groups of primitives. The contour feature represents the boundary characteristics of gray and binary patterns by exploiting the gradient information in local and global fashion. In this paper, feature combination approach is used to improve the writer recognition, where learning based framework is followed for generation of optimal combination. Feature combination through learning have proved to improve the performance of many recognition problems Yin et al., 2005; Fang et al., 2002; Fang et al, 2009. We demonstrate combination of features for identity determination through learning. The existing work in the related domain have addressed the problem at page, paragraph and line level. These methods have inherent limitation as the document may have varying line, paragraph and word counts. The present work proposes feature combination framework for handwritten document recognition which could be directly applied for identity determination at word level. The application of the framework is demonstrated on Devanagari script collection, where word formulation is characterized by combination of set of characters exhibiting complex shapes

* Corresponding author. Tel.: +91 9911871941.
E-mail addresses: hassan.ehtesham@gmail.com (E. Hassan), santanuc@ee.iitd.ernet.in (S. Chaudhury), nivedita@cse.iitd.ernet.in (N. Yadav), pkalra@cse.iitd.ernet.in (P. Kalra), mgopal@snu.edu.in (M. Gopal).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.04.032

and large set of modifiers. The brief of the proposed concept and experimental evaluation on small data collection is presented in Hassan et al. (2010). In this work, we extend the data collection and increase the size of the feature set for rigorous validation and analysis of proposed feature combination framework. The paper introduces a set of novel features for handwritten word/syntax object representation. The features primarily exploit the shape information and statistical attributes of the word object. The brief description of features are as follows:

- Shape descriptor: The feature represents the structural organization of the word object by 2-D histogram.
- Envelope curve coding: The feature exploits the outer boundary information of the word object. The outer boundary information is represented by upper and lower envelope curves of the word. The second feature defines a coding scheme based on the envelope curves. The coding scheme represents the pattern of envelopes in different regions of image box.
- Envelope histogram of gradients: The feature represents the outer boundary gradient information which is considered to invariant to small disturbances.
- Computational feature set: represents the statistical and morphological attributes of the bounding box containing the word object.
- Fourier Mellin transform: The feature presents integral transform based rotation and scale invariant description of gray scale objects.

The structure of the paper is as follows. The details of the different feature representations used in this paper are presented in the Section 2. The feature combination framework: the problem formulation and solution framework is presented in Section 3. The experimental evaluation and analysis of proposed concepts are presented in the Section 4. The final section concludes and presents perspective of the work.

2. Feature description

The details of the features used in the paper are presented as follows.

2.1. Shape descriptor (SD)

The shape descriptor proposed in (Hassan et al. (2009)) represents the structural organization of an object in 2D histogram and is fundamentally based on shape context (Belongie et al., 2002). For the set of n descriptor points on object shape, there are $n(n-1)$ point-pair arrangements. The histogram represents distance and orientation based distribution of these point-pair arrangements. We name the histogram as point distribution histogram (*pdh*) and consider its Fourier coefficient as shape descriptor. The *pdh* is computed as sum of shape contexts for the point set. The empirical evaluation have shown that robustness of descriptor to broken links and noisy ink dots increases by considering partition based approach for descriptor computation. In this case, the word image is divided in fix partitions of equal width (Fig. 1). The *pdh* for each partition is computed w.r.t. points lying in that partition. The partition based approach for descriptor computation nullifies the affect of distortion or noise in one partition to other partitions. The descriptor computational steps are summarized as follows:

- Descriptor points extraction followed by *pdh* computation for all the partitions.
- Descriptor is defined as $\|Fourier(h_1 : \ldots : h_{parts})\|, h_i$ represents the *pdh* for i^{th} partition

The global shape features are predominantly represented by low frequency Fourier coefficients and local shape features like internal contours, sharp curves and broken links contribute in high frequency Fourier coefficients. We follow grid based approach to extract semantic information inherent in the inner contours of word shape to define their unique representation. For descriptor point extraction, we overlay a logical grid over the word shape. The transition points $\{1 \to 0, 0 \to 1\}$ over grid are selected as descriptor points (The red points in Fig. 3). The process of point extraction gives due importance to inner contours for defining representation. Using the extracted set of points Shape descriptor (SD) is computed as discussed earlier.

2.2. Envelope curve coding (EC)

The outer boundary of handwritten word sample remains invariant to different writing conditions for a writer. We extract the boundary information by detecting the lower and upper envelope curves of the word across the principal axis of the word (Fig. 2). We define a coding scheme for defining the representation based on the envelope pattern. The feature therefore represents the local information about the word envelope. For our feature computation we have divided the each envelope image (lower and upper envelopes) using 4×6 grid. For each division in the image created by the grid, the following coding scheme defined the pattern of envelope curve.

- 0:If envelope does not pass through the division.
- 1:If envelope passes through the division, but no prominent peak or valley appears or both of them appear in the envelope pattern in the division.
- 2:If the envelope exhibit a prominent peak in the division.
- 3:if the envelope exhibit a prominent valley in the division.

The peaks and valleys are detected by processing local information of the envelope curve (Fig. 4). The peak or valley in the envelope pattern is considered prominent, if the rise or fall computed in y-direction is greater than one fourth of the height of the division. Therefore, for each envelope we get a 4×6 vector as representation. The word representation is defined by concatenating the upper and the lower envelope curve codes.

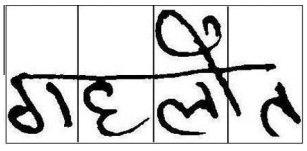

Fig. 1. Partitions in the image.

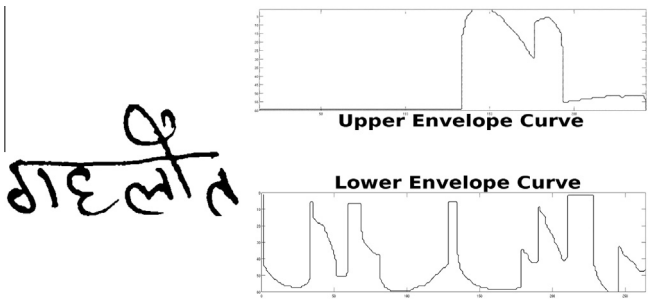

Fig. 2. Envelope curve of the word.