Pattern Recognition 58 (2016) 172–189

Contents lists available at ScienceDirect

Pattern Recognition

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

A multi-objective approach towards cost effective isolated handwritten Bangla character and digit recognition

Ritesh Sarkhel, Nibaran Das*, Amit K. Saha, Mita Nasipuri

Computer Science and Engineering Department, Jadavpur University, Kolkata 700032, India

ARTICLE INFO

Article history:
Received 22 September 2015
Received in revised form
22 March 2016
Accepted 13 April 2016
Available online 22 April 2016

Keywords:
Feature set
Region sampling
Handwritten character recognition
Multi-objective evolutionary algorithm
Harmony search
NSGA-II
AFS theory

ABSTRACT

Identifying the most informative local regions of a handwritten character image is necessary for a robust handwritten character recognition system. But identifying them from a character image is a difficult task. If this task were to be performed incurring minimum possible cost, it becomes more challenging due to having two independent, apparently contradicting objectives which need to be optimized simultaneously, i.e. maximizing the recognition accuracy and minimizing the associated recognition cost. To address the problem a multi-objective approach is required. In the present task, two popular multi-objective optimization Algorithm (1) a Non-Dominated Sorting Harmony-Search Algorithm (NSHA) and (2) a Non-Dominated Sorting Genetic Algorithm-II (NSGA-II, Deb et al., 2002 [18]) are employed for region sampling separately. The method objectively selects the most informative set of local regions using the framework of Axiomatic Fuzzy Set (AFS) theory, from the sets of pareto-optimal solutions provided by the multi-objective region sampling algorithms. The system has been evaluated on two isolated handwritten Bangla datasets, (1) a dataset of randomly mixed handwritten Bangla Basic and Compound characters and (2) a dataset of handwritten Bangla numerals separately, with SVM based classifier, using a feature set containing convex-hull based features and CG based quad-tree partitioned longest-run based local features extracted from the selected local regions. The results have shown a significant increase in recognition accuracy and decrease in recognition cost for all the datasets. Thus the present system introduces a cost effective approach towards isolated handwritten character recognition systems.

1. Introduction

Optical Character Recognition (OCR) is an active area of research. While there are many systems commercially available for recognizing printed text [1–4], their success is yet to be extended to handwritten characters. Several reasons can be cited to explain this apparent anomaly. Shape and size of handwritten characters vary from one individual to another. It may even vary for a single individual from time to time, depending on various factors. These challenges make the task of recognizing handwritten characters very difficult. Researchers all around the world have proposed several methods [5] for handwritten character recognition, but most of them are focused on Roman scripts [6], concentrating on English and other European languages. Among Asian languages, Chinese [7], Japanese, Korean languages are dominant in the literature. Indian scripts like Malayalam, Tamil, Telugu, and Hindi have started to get attention of the researchers during past decade [8,9], but development of OCR for complete Bangla script [10] has not received much attention from researchers until recently. Bangla is the second most popular script in India and the fifth most popular script in the world [11]. Bangla alphabet contains some of the most intricate and complex characters, which differ from one another only by a single period, a modifier *ref* or an upper horizontal line or *Matra,* as shown in an example in Fig. 1. Bangla alphabet contains about 50 Basic characters (11 vowels and 39 consonants) and more than 334 Compound characters [12]. Samples of a few of Bangla Basic and Compound characters is shown in Fig. 2.

One of the most common approaches taken up by OCR researchers is *zoning*, i.e. dividing the character image into several *zones* or *local regions* [13] and generating the invariant *local feature set* by extracting features from every *local region*. There are several different *zoning methods* [13] mentioned in the literature, but most of them can be classified into two major categories: *static* [4,10] and *dynamic zoning methods* [13]. *Static zoning methods* divide a handwritten character image into a fixed set of overlapping or non-overlapping windows, where the number of windows is fixed. Basu et al. used static zoning method in [14] and sub-divided the handwritten numerals' image into 9 fixed-sized, overlapping *local*

* Corresponding author. Tel./fax: +91 3324146766.
E-mail address: nibaran@gmail.com (N. Das).

http://dx.doi.org/10.1016/j.patcog.2016.04.010
0031-3203/

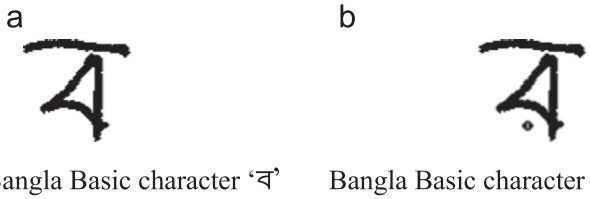

Fig. 1. Similarity in shape and size between different Bangla characters. (a). Bangla Basic character 'ব' (b) Bangla Basic character 'র'.

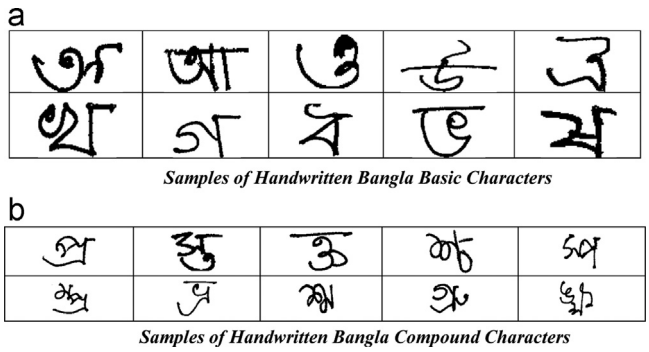

Fig. 2. Samples of handwritten Bangla characters.

regions and extracted longest-run based features form each sub-region. On other hand, *dynamic zoning methods* sub-divide a handwritten character image into local regions by dynamically creating windows based on some statistical or topological feature of that specific character. Cao et al. [8] proposed a similar technique to generate a hierarchical feature-space based on a quin-tree partition of the character image, where zones were dynamically created based on the centroid of the contour segment of the character residing in the parent zone. Das et al. [10,15] have used a GA based selection mechanism to find out the most optimal set of local regions for recognition of handwritten Bangla numerals.

In those papers, the researchers have emphasized on achieving better recognition accuracies, but associated recognition costs incurred in the process were not taken into consideration. For example, Das et al. presented a two pass approach towards handwritten character recognition in [10], which produced a significant increase in the recognition accuracy but at the cost of a recognition cost which is almost 8.5 times than the average per character recognition cost incurred by traditional single pass approach towards handwritten character recognition. This may prove to be undesirable to users who want to use such a system for real-life applications. An extensive study of recognition accuracy versus associated recognition cost is undertaken in our experimental setup to investigate the scope of a practical optical character recognition system, in terms of both recognition accuracy and associated recognition cost.

In the present work, a multi-objective approach towards optical character recognition (OCR) is proposed, which attempts to find a trade-off between the recognition accuracy achieved by the system and its associated recognition costs. In real life applications of an OCR system, insignificant increase in recognition accuracy at the expense of high recognition cost may not be acceptable to the users of the system. In such cases, a multi-objective approach can provide the user with a set of good solutions. In the present work, framework of a novel, multi-objective isolated handwritten character recognition system is proposed. There are several variants of multi-objective Evolutionary Algorithms [16] present in the literature. A *Non-dominated Sorting Harmony-search Algorithm (NSHA* [17]*)* based region sampling method and a *Non-dominated Sorting Genetic Algorithm – II (NSGA-II* [18]*)* based region sampling method is introduced in our present work. These two multi-objective region sampling algorithms mark one of the contributions of the present work. Both of the region sampling algorithms are employed over the decision space separately. These algorithms have two objective functions: – (1) maximizing handwritten character recognition accuracy and (2) minimizing associated recognition costs. In our experimental setup, recognition accuracy is measured using an SVM based classifier and recognition costs are measured by: (i) average time taken by the recognition system to recognize each handwritten character in the test-set and (ii) the number of local regions used to represent each handwritten character in the test-set. Two sets of pareto-optimal solutions provided by these two algorithms are then combined using *Axiomatic Fuzzy Set (AFS) Theory* [19]. The multi-objective region sampling algorithms and the AFS theory based approach to objectively combine the pareto-optimal solutions provided by the multi-objective algorithms mark one of the contributions of the present work.

The proposed method tries to find an objective solution over the decision space, while providing an optimal trade-off between recognition accuracy and corresponding recognition costs, making it suitable to use in practical applications. The present work has been evaluated on datasets of isolated handwritten Bangla characters and handwritten Bangla numerals separately. Results from these experiments have been compared with some of the other popular handwritten character recognition methods present in the literature, to prove its superiority.

The rest of the paper is organized as follows: in Section 2, a brief overview on multi-objective evolutionary algorithms based region sampling techniques is presented, basics of Axiomatic Fuzzy Set (AFS) theory is introduced in Section 3; Section 4 describes the featureset and our present work is discussed in details in Section 5, experimental results are presented in Section 6. Finally, a brief conclusion is drawn based on the results gathered from the experiments.

2. Motivation behind using multi-objective evolutionary algorithms for region sampling

Region sampling based OCR systems try to identify the most discriminative set of local regions from handwritten character