Applied Soft Computing 62 (2018) 315–327

Contents lists available at ScienceDirect

Applied Soft Computing

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

Evaluation of a novel fuzzy sequential pattern recognition tool (fuzzy elastic matching machine) and its applications in speech and handwriting recognition

Sina Shahmoradi *, Saeed Bagheri Shouraki

Artificial Creatures Lab (ACL), Department of Electrical Engineering, Sharif University of Technology, Azadi Ave, Tehran, Iran

ARTICLE INFO

Article history:
Received 23 November 2016
Received in revised form
26 September 2017
Accepted 21 October 2017
Available online 31 October 2017

Keywords:
Fuzzy sequential pattern recognition
Fuzzy elastic pattern
Speech recognition
Handwriting recognition
Hidden markov model

ABSTRACT

Sequential pattern recognition has long been an important topic of soft computing research with a wide area of applications including speech and handwriting recognition. In this paper, the performance of a novel fuzzy sequential pattern recognition tool named “Fuzzy Elastic Matching Machine” has been investigated. This tool overcomes the shortcomings of the HMM including its inflexible mathematical structure and inconsistent mathematical assumptions with imprecise input data. To do so, “Fuzzy Elastic Pattern” was introduced as the basic element of FEMM. It models the elasticity property of input data using fuzzy vectors. A sequential pattern such as a word in speech or a piece of writing is treated as a sequence of parts in which each part has an elastic nature (i.e. can skew or stretch depending on the speaker/writer’s style). To present FEMM as a sequential pattern recognition tool, three basic problems, including evaluation, assignment, and training problems, were defined and their solutions were presented for FEMMs. Finally, we implemented FEMM for speech and handwriting recognition on some large databases including TIMIT database and Dr. Kabir’s Persian handwriting database. In speech recognition, FEMM achieved 71% and 75.5% recognition rates in phone and word recognition, respectively. Also, 75.9% recognition accuracy was obtained in Persian handwriting recognition. The results indicated 18.2% higher recognition speed and 9–16% more immunity to noise in speech recognition in addition to 5% higher recognition rate in handwriting recognition compared to the HMM.

1. Introduction

Sequential data often emerge in the measurement of time series, e.g. the acoustic features at the time frames of speech signals [1] and kinematic and biological signals in human movements [2]. Hence, sequential pattern recognition is increasingly used in many data analysis fields such as handwriting recognition [3,4], speech recognition [5,6], activity recognition [7,8], and gesture recognition [2]. In this approach, sequences of information are utilized in decision-making [9]. Sequential pattern recognition in time-series information matches the input data with a pattern, both of which being sequences of information. However, inequality in the length of sequences between the input data and the pattern is a typical problem. Studies in 1960s showed that an appropriate pattern recognition tool should be developed to overcome this problem in speech recognition [10]. One of the main solutions to this problem is the “time-wrapping” idea, presented by Vintsyk [11]. He showed how dynamic programming can be used to find the best assignment between two sequences. In 1970s and 1980s, many researchers proposed various models that mostly included modeling of acoustic information in a network with some finite states, deriving the stochastic information of acoustic features and comparing them with the input signal using the dynamic time-wrapping method [12]. Some examples of these works can be seen in [13–16]. The most well-known method in this field is Hidden Markov Model (HMM) [17]. The capabilities of the HMM such as the ability to match with stochastic sequential time series and suitable training algorithms have made it a pervasive tool in speech recognition. Since mid-1990s, the HMM has been a dominant method in handwriting and speech recognition [18]. Besides, many researchers have demonstrated its superiority to short time classifiers such as MLP [19] and SVM [8].

Despite the powerful mathematical modeling of the HMMs, the main criticisms about them fall into two categories: 1) limitations due to the selected mathematical modeling assumptions. One of these assumptions is that the state duration in the HMM is

* Corresponding author.
E-mail address: sina.shah71@gmail.com (S. Shahmoradi).

https://doi.org/10.1016/j.asoc.2017.10.036
1568-4946/

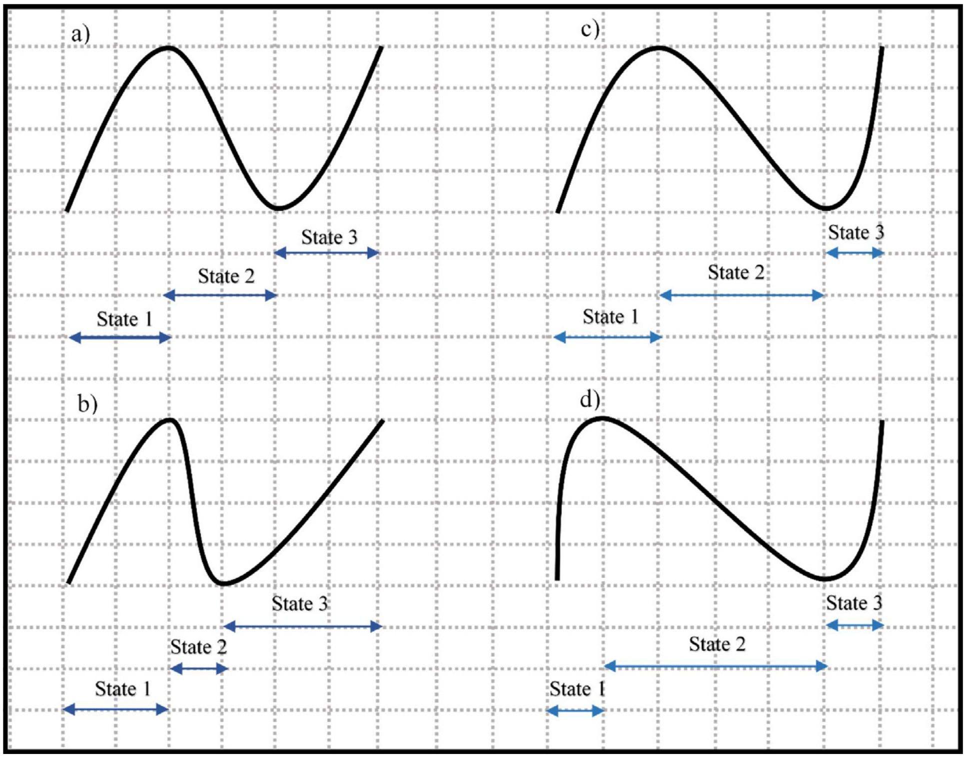

Fig. 1. Four samples that show a pattern. In the sequential approach, each sample is segmented into three segments. Each segment has specific features with an elastic nature. An appropriate sequential data model should describe the pattern according to these samples. Also, it should measure the similarity between the input data and this pattern.

implicitly a geometric distribution [20]. Consequently, the HMM is incapable of accommodating many important facets of the speech pattern structure [21]; 2) its inflexible mathematical structure, which is inconsistent with the imprecise nature of handwritten and speech data.

The main solution to the first problem is modeling the state duration explicitly using a state duration probability density. These works are referred to in the literature as the "Segmental Hidden Markov Model" [22] or "Hidden Semi-Markov Model" [20]. In addition, state duration modeling has been proven to be effective in the extensions of the HMMs such as Bayesian non-parametric HSMM [23] or explicit duration switch HMMs [24]. Although modeling the state duration has mostly yielded better recognition rates, more computational complexity is the main problem of standard HSMMs [25,26].

The major solution to the second problem is combining the HMM with fuzzy modeling in order to reduce its sensitivity to changes (e.g. different handwriting/speech styles) and to improve recognition performance in the presence of noise [27,28]. Tseng combined fuzzy vector quantization with HMMs (FVQ/HMM), thereby improving the recognition performance [29]. Besides, some researchers have attempted to express the elements of the HMM and extend Baum-Welch and Viterbi algorithms with Type1 [30] or Type2 [28] Fuzzy sets. In the same manner, Cheok proposed the generalized fuzzy HMM and showed that, compared to the HMM, this method can achieve the same recognition rate faster [31]. Due to the success of fuzzy HMM, it has been implemented in many applications [32–34]. Nevertheless, elasticity should be considered in fuzzy-based methods in order to deal with large and long datasets of imprecise speech and handwritten data. Elasticity is an inherent property of speech and handwritten data, which means that a part of the input data can stretch or skew based on the speaker/writer's style [35]. According to the authors' knowledge, no fuzzy sequential pattern recognition tool that can successfully handle the elasticity property in speech or handwritten data has been presented yet.

The goal of this paper was to present a fuzzy sequential pattern recognition tool which considers the elasticity property of input data. Unlike most studies that tried to improve the performance of the HMM by changing its structure or combining it with other methods [20–31], we focused on the definition of the basic element of a sequential pattern recognition tool using fuzzy elastic pattern. Since this tool is designed to compare and match input data with the fuzzy elastic pattern, it is called "Fuzzy Elastic Matching Machine" (FEMM).

The rest of this paper is organized as follows: In the second section, fuzzy elastic pattern is presented and its adequacy as an appropriate sequential data model is investigated. The third section is devoted to the introduction of FEMM and its basic elements. To make FEMMs applicable in pattern recognition problems, three basic problems with their solutions are presented for FEMMs in the fourth section. The fifth section discusses the results of implementing FEMM for speech and Persian handwriting recognition. The results and efficiency of FEMM are discussed in the sixth section. Finally, in the seventh section, we draw conclusions and future works are explained briefly.

2. Definition of fuzzy elastic pattern

Fig. 1 demonstrates four samples of a pattern that we modeled using the sequential approach. In this approach, the samples are segmented into three segments that come in a sequence, as shown in Fig. 1. With regard to these samples, the requirements of an appropriate sequential data model are enumerated below. Also, for recognition, the effective elements in measuring the degree of similarity between the input data and the pattern are indicated. First, the features of each segment should be described and the similarity