Pattern Recognition Letters 103 (2018) 1–7

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A lexicon-free approach for 3D handwriting recognition using classifier combination

Pradeep Kumar[a,*], Rajkumar Saini[a], Partha Pratim Roy[a], Umapada Pal[b]

[a] *Department of Computer Science and Engineering, Indian Institute of Technology, Roorkee, India*
[b] *Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India*

ARTICLE INFO

Article history:
Received 30 June 2017
Available online 13 December 2017

Keywords:
3D Handwriting
Classifier combination
Leap motion
Devanagari script
ROVER
BLSTM-NN

ABSTRACT

Recent developments in depth sensing technology such as Leap Motion have opened novel directions in Human-Computer-Interaction (HCI) research domain. The sensor extends the way of writing from traditional method to a gesture based writing in the 3D space. The online text written in 3D space over the sensorâs viewing field is different from traditional 2D handwriting in several ways. The 3D handwriting does not consist any stroke information, since all characters are connected by a single stroke. Moreover, non-uniform text styles and jitters during writing in 3D space create additional challenge for the recognition task. Because of these challenges in 3D handwriting, recognition of cursive words is not satisfied using a single classifier. In this paper, we present a lexicon free approach for the recognition of 3D handwritten words in Latin and Devanagari scripts by combining multiple classifiers. The individual recognition systems are computed using Bidirectional Long-Short Term Memory Neural Network (BLSTM-NN) classifier with the help of different features. The combination of multiple classifier is performed by aligning the output word sequence of each classifier using the Recognizer Output Voting Error Reduction (ROVER) framework. Accuracies of 72.25% and 71.86% are recorded using the proposed methodology for Latin and Devanagari scripts, respectively.

1. Introduction

In recent years, the emergence of 3D interaction technologies such as Leap Motion [32] and Microsoft Kinect [11] sensors have opened-up various research directions in the field of Human-Computer-Interaction (HCI). These devices provide bare-hand interaction where a user can provide input to the system with the help of gestures and postures in the 3D space with no devices or wires attached [19]. The devices are able to provide 3D point cloud of the observed scene that appears in their viewing field. Researchers have successfully used these depth sensing technology in various real life applications including gesture recognition in [14], 3D air-writing in [15], interactive gaming in [26] and rehabilitation systems in [29].

Leap motion sensor is specifically designed to detect and track the positions and orientations of the finger and hand movements with high level of precision as discussed in [19]. The sensor uses infrared imaging technique to determine the hand and finger positions. It consists of three infrared LED emitters and two high precision infrared cameras to capture hand information within the sensor's viewing field [32]. The field of view of the sensor is an inverted pyramid of 8 cubic feet and the effective range extends from 1 inch to 2 feet above the device. The sensor is successfully used by researchers in various HCI applications. For example, [22] have proposed a sign language recognition (SLR) system for Australian sign language alphabets using Leap motion sensor. The authors have used device's Application Programming Interface (API) to extract the finger and hand positions. The recognition of the sign symbols was performed with the help of Artificial Neural Network (ANN) classifier. Similarly, the authors in [14] have proposed a SLR framework for Indian sign language using Leap motion in conjunction with the Kinect sensor. They have developed a calibration of both the sensors to acquire dynamic sign gestures. The recognition of the sign gestures was performed using Coupled Hidden Markov Model (CHMM) with an accuracy of 90.8%. The sensor is also used to develop various rehabilitation systems. Vamsikrishna et al. [29] have developed an assistive system for palm rehabilitation using Leap motion. The system was trained on 8 basic gestures that were related to palm and finger rehabilitation. The identification of the gestures was performed using discriminant analysis and Support Vector Machine (SVM) classifiers, whereas the recognition of gesture sequences was performed using HMM during rehabilitation.

* Corresponding author.
E-mail address: pra14.dcs2014@iitr.ac.in (P. Kumar).

https://doi.org/10.1016/j.patrec.2017.12.014
0167-8655/

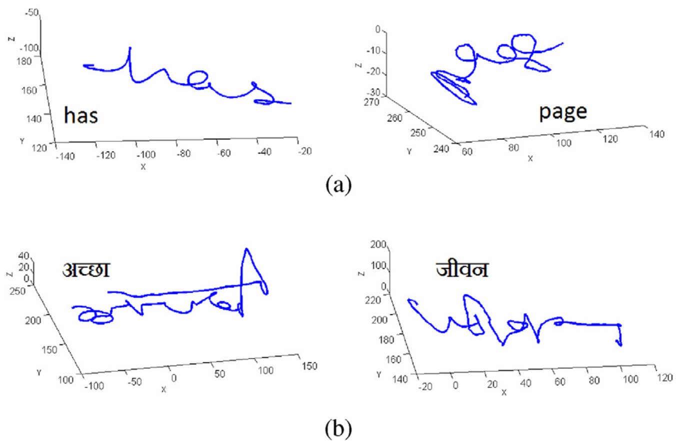

Fig. 1. 3D handwriting examples written in: (a) Latin script (b) Devanagari script.

Leap motion sensor extends the way of writing from the traditional pen-paper based style to a novel 3D environment that facilitates a user to write in air without pen-paper. The technique has multiple benefits which allows users to write notes, searching name from directory, internet surfing, entering security passwords, etc. with the help of touch less interfaces [14]. However, 3D handwriting is different from traditional 2D online handwriting in multiple ways. (i) There is no pen up/down stroke information to determine the beginning and end of the characters, thus, all characters in the words are connected by a single stroke. (ii) There is absence of visible reference point for aligning the fingertip in the 3D space while writing, thus, large variations are recorded even with similar words. (iii) There is no predefined plane for writing i.e. the user can write in any plane or in any direction in the 3D coordinate geometry. Few examples of such variations are depicted in Fig. 1, where two different words in Latin and Devanagari scripts are written by two different writers.

Kumar et al. [16] have proposed a methodology for 3D text segmentation and recognition using Leap motion. The authors have recorded 320 text lines of air-writing in Latin language from multiple users. Word segmentation was performed using a 3D window based analysis and the recognition of these segmented words was performed using HMM classifier with lexicon where an accuracy of 92.73% was recorded. Likewise, Xu et al. [34] have proposed a 3D Chinese character recognition system using Leap motion. The authors have extracted 8-directional and direction change features from the text trajectory and combined them for recognition purpose where an accuracy of 69.67% was recorded using Linear Discriminant Analysis (LDA) classifier. In [30], the authors have developed a 3D handwriting recognition system using Leap motion sensor. They fed the input 3D text trajectory into a Dynamic Time Warping (DTW) based recognition algorithm.

In literature, multiple feature and classification methods have been proposed by researchers for both online and offline handwriting recognition systems. However, none of them is able to reach a totally satisfactory solution to the problem due to the cursive text and complications in character segmentation [31]. Therefore, classifier combination techniques can be used to improve the handwritten character recognition performance as discussed by researchers [13,18,33]. Different classifier combination methodologies are available that combine results from individual classifiers. The process is also known as decision fusion as it combines the decisions of multiple classifiers. The combination system rescores each decision and the decision with maximum score is selected as the final result. Existing approaches of classifier combination include majority voting, Borda count, max rule, sum rule, weighted sum rule, etc. However, these approaches fail to combine the decisions if the classifier generates a string of character symbols with not only substitution errors but also with insertion and deletion errors [31]. Moreover, due to segmentation errors, all recognizers can not generate sequences of same length [2]. Therefore, to deal with these problems, we have used Recognizer Output Voting Error Reduction (ROVER) algorithm given by Fiscus of NIST [6] for the recognition of 3D text. Originally, ROVER was developed to reduce word error rates in automatic speech recognition systems by combining multiple recognizers. Liwicki and Bunke [17] have proposed a multiple classifier system for handwritten text line recognition using ROVER. Different recognizers have been obtained using online and offline features and later the text sequences were incrementally aligned using ROVER for final recognition. An accuracy of 83.64% has been recorded using the ensemble method. Similarly, Wang et al. [31] have applied ROVER algorithm for the handwritten word recognition. They have obtained multiple classifiers using discrete HMMs using different feature vectors. Finally, the results have been combined using ROVER and relative error reduction of 13% and 35% have been recorded at word and character level in comparison to single classifier, respectively.

Since the characters of the 3D words are connected to each other through a continuous stroke, we have used Bidirectional Long Short Term Memory Neural Networks (BLSTM-NN) classifier for recognition purpose. The classifier uses a Connectionist Temporal Classification (CTC) layer that allows the network to be trained and tested on unsegmented data. The classifier has been successfully used by researchers in various handwriting and scene text recognition problems [25,27]. Su and Lu [28] have proposed a scene text recognition technique by integrating multiple Recurrent Neural Networks (RNNs). The recognition was performed on word level without character segmentation by converting the word image to a temporal signal. Likewise, Kumar et al. [15], the authors have used BLSTM-NN classifier for the recognition of the 3D text written in air using Leap motion sensor. Graves et al. [9] have used BLSTM-NN to label sequences when the data is hard to segment and has multi-directional inter-dependencies. They have shown that the network performs better than HMM classifier for speech and handwritten text recognition. In this work, we present a lexicon free approach for the recognition of 3D text of Devanagari and Latin scripts written in air using Leap motion sensor. Multiple BLSTM-NN classifiers are trained using different feature sets. Next, individual recognition systems are combined using ROVER algorithm. The main contributions of the work are as follows:

- Firstly, we present a lexicon free touch less handwriting recognition system for Latin and Devanagari scripts using Leap motion sensor.
- Secondly, multiple classifiers are learned on different 3D features using BLSTM-NN classifier with CTC layer. Additionally, more recognizers are obtained by changing the initial weights of the networks.
- Thirdly, we combine the output of multiple recognizers by incrementally aligning the sequences using ROVER algorithm.
- Fourthly, this is the first work on 3D text recognition in Devanagari script.

Rest of the paper is organized as follows. The proposed methodology of word recognition is described in Section 2 along with the details of the classification method and combination strategy. The results are presented in Section 3. Finally, we conclude in Section 4 by discussing some future extension of the work.

2. Proposed methodology

In this section, we present the details of our lexicon free 3D words recognition framework. The acquisition of 3D handwriting text trajectories is accomplished using the Leap motion sensor API.

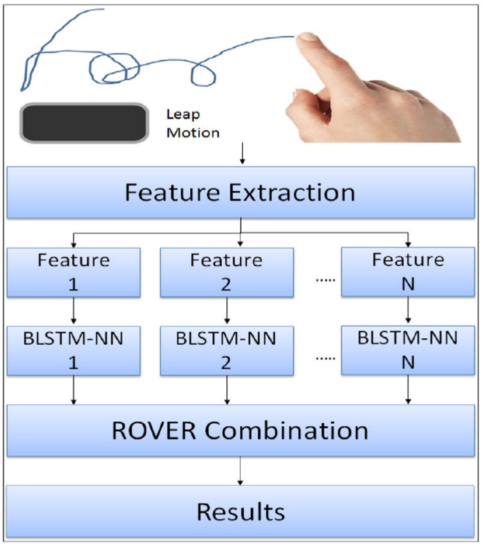

Fig. 2. Proposed framework for lexicon free 3D handwriting recognition using classifier combination through ROVER.

Next, the raw text trajectories are preprocessed using scale normalization and aligned to one of the coordinate axes to remove writer specific properties. Different features are then extracted from the trajectories and training is performed separately on each feature using BLSTM-NN classifier. Thus, multiple classification models are obtained which are later combined using ROVER framework. An overview of the proposed framework is depicted in Fig. 2, where a writer is writing text in Latin script in-air over the Leap motion sensor. Next, preprocessing of the acquired 3D signal is performed to make the text consistent after removing writer specific properties. Multiple 3D features such as writing direction, curvature, convex-hull etc. are extracted from the preprocessed data and learned on different BLSTM-NN classifiers. Finally, the transcriptions from multiple classifiers are aligned using dynamic algorithm and combined using ROVER framework that results into best word recognition.

In this work, to show the robustness of the framework, experiments are conducted in two different scripts, namely, Latin and Devanagari. Latin script is the most widely adopted writing system in the world and commonly used by about 70% of the world's population whereas Devanagari is the most popular script in India and the Indian national language Hindi is written in Devanagari script. Hindi is the third most popular language in the world [20,21]. Devanagari script is composed of 13 vowels, and 34 consonants. Sometimes two or more characters combine and generate a new complex shape in Devanagari which is called as compound character (or clusters). Likewise, a vowel following a consonant may take a modified shape, and depending on the vowel it is placed to the left, right, top, or bottom of the consonant, and are referred them as modifiers or *'matras'* [12]. As a result, in Devanagari, there are more than 200 character shapes. Most of the Devanagari characters have a horizontal line at the upper part called *headline* and the characters of words are connected using such headline. There is no concept of upper or lowercase characters in Devanagari. In contrast, Latin script has 52 characters (26 uppercase and 26 lowercase). In Latin, there is no compound character. The structure of the Latin alphabet consists of more vertical and slant strokes [4]. Despite of these differences in both the scripts, we propose a general framework that utilize the 3D stroke information and extract various features which are learned separately on multiple recognition systems. The benefit of such framework is that if one feature

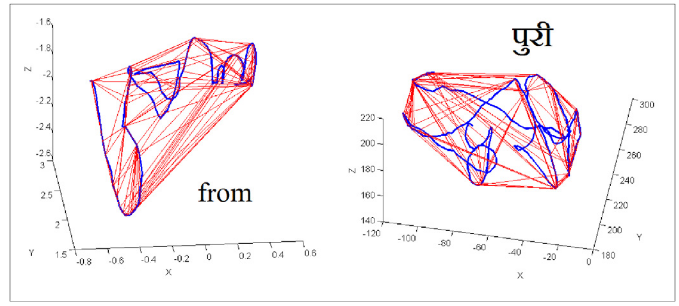

Fig. 3. Example of 3D convex hull features for both Latin and Devanagari scripts words: (a) 'from' (b) 'puri'.

does not work for a script then learning of other features dominate in the final recognition during combination using ROVER.

2.1. Preprocessing and feature extraction

Since there is no predefined plane and reference point for the writers to draw text in 3D space, preprocessing techniques (Normalization and Alignment) are applied on 3D raw text trajectories to make them of uniform size to remove writer specific characteristics and align on one of the coordinate axes [15]. In this work, normalization is carried out for each axis by computing the maximum value of the text along the axis. Doing this, the size of the 3D text trajectories lies between [1,−1]. Next, the 3D text trajectories are aligned along the z-axis over the xz-plane by computing regression line for each 3D word. Finally, the trajectories are rotated about z-axis using regression line for final alignment. More details about the preprocessing can be found in [15,16]. Next, three different features are extracted, namely, writing direction, curvature and convex hull from the raw text. The details of 3D writing direction and curvature features can be found in [15,16] whereas 3D convex hull based features are discussed here.

2.1.1. 3D convex Hull

Convex hull based feature is successfully used by researchers in handwriting recognition systems [3,23,24]. Convex hull is generally defined as the smallest convex polygon that contains all the points of the object [7]. The topological structure of the 3D point set ($N \in R^3$) with p points can be represented using a small set of vertices that belong to the convex hull defined in (1), where x is a point in N and is considered the vertex of $V(N)$ if x is a vertex of convex hull and removing it will result into smaller convex hull [5]. We have calculated the 3D convex hull features for both Latin and Devanagari scripts using the triangulation method [8]. A pictorial representation of 3D convex hull for the words 'from' and 'puri' written in Latin and Devanagari scripts are depicted in Fig. 3.

$$V(N) = \{x | x \in N \ \ \& \ \ x \notin V(N - \{x\})\} \tag{1}$$

2.2. 3D text recognition using BLSTM-NN classifier

BLSTM-NN is a sequence modeling classifier and popularly used in various gesture and handwriting recognition problems [10]. The classifier processes the input sequences in forward and backward directions with the help of two hidden layers. Both the layers are connected to a common output layer. We have trained the network with CTC network which is designed to overcome the inherent problem of Recurrent Neural Networks (RNNs) that require pre-segmented data for training. CTC uses the network to define a probability distribution over a fixed set of labels and a special node 'ϵ' to denote 'no label'. CTC objective function (O) is defined as the negative log probability of correct labeling of the entire training set

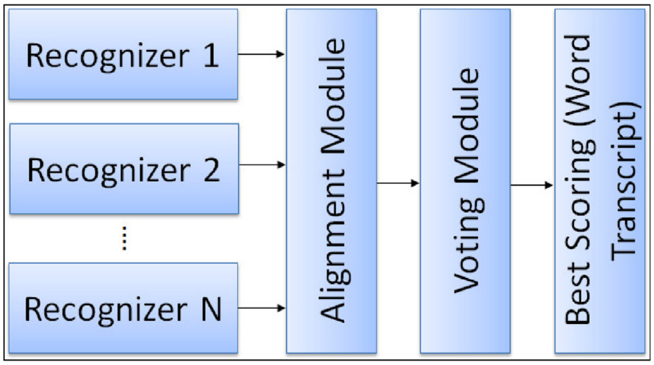

Fig. 4. Architecture of the ROVER framework to combine the results of multiple recognizers.

and can be computed using (2) for the training set T, where (x, z) represents a pair of input and target sequence.

$$O = -ln\left(\prod_{(x,z)\epsilon T} p(z|x)\right) = -\sum_{(x,z)\epsilon T} ln(p(z|x)) \quad (2)$$

O models the label sequence with the given inputs directly. In [10], BLSTM-NN based approach has been proposed for the recognition of online and offline handwriting recognition by providing long range context in directions. They have performed the recognition at word level instead of character level to resolve the issues related to segmentation and overlapping characters. We have adopted the concept for the recognition of Latin and Devanagari scripts by performing the training of the network using different feature vectors. Moreover, different classifiers can be obtained from each feature set by initializing the network with different random seeds.

2.3. ROVER Framework

ROVER was initially developed by Fiscus of NIST [6]. Its aim was to reduce word error rates for automatic speech recognition systems by combining multiple speech recognizers. ROVER is successfully used by researchers in handwriting recognition systems [1,2,17]. ROVER framework consists of two modules, namely, alignment and scoring module. The alignment module builds a Word Transition Network (WTN) by aligning more than two strings with the help of an iterative process. This module treats the first WTN as the base and aligns second WTN with the first using dynamic programming. The process iterates until all inputs are combined into a single WTN. After forming the composite WTN from the initial system outputs, the scoring module rescores each word at each node of the WTN with the help of a voting scheme, using frequency, confidence and time information and selects the best scoring word sequence as the final decision. The overall architecture of the ROVER framework is shown in Fig. 4.

The trade off between the frequency of occurrence and the confidence score is weighted using γ. If c represents the set of WTN with n word classes, namely, $w_1, w_2, \ldots w_n$. Then the score $(S(w_i))$ of word w_i can be calculated using (3), where N_i and $C(w_i)$ are the number of occurrences and combined confidence scores of word class w_i [2].

$$S(w_i) = \gamma * \frac{N_i}{\sum_{j=1}^{n} N_j} + (1-\gamma) * C(w_i) \quad (3)$$

The value of $C(w_i)$ can be set as the average of all confidence scores of occurrences for w_i, or alternating the maximum confidence score among all can also be chosen. The confidence score of

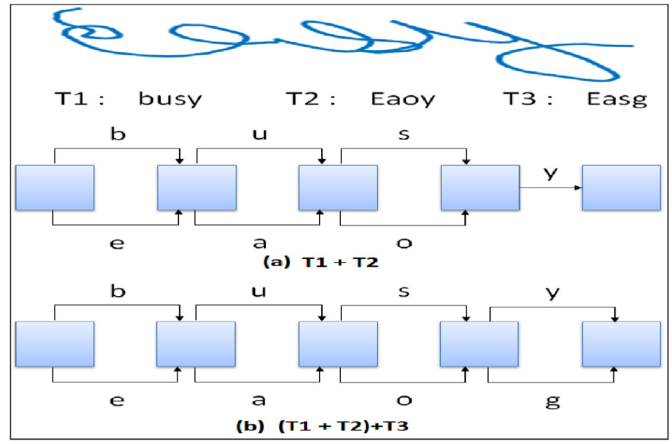

Fig. 5. Example showing iterative alignment of multiple recognition results for a Latin word 'easy': (a) transcriptions T1 and T2 are aligned (b) Transcription T3 is aligned with (T1+T2) to make a single WTN.

null-transition arcs C_ϵ is another parameter that is required during combination. Low C_ϵ results into long output transcriptions with more insertion errors whereas high C_ϵ results into more deletion errors [17]. If the value of γ is set to 1 in Eq. (3) then, $C(w_i)$ has no impact and $S(w_i)$ results into simple plurality voting. In this work, instead of word WTN, we have character WTNs with multiple classes belonging to Latin and Devanagari scripts. The 3D word recognition has been performed for both scripts using character models with the help of BLSTM-NN classifier. Multiple recognizers for the same 3D word sequence '*easy*' of Latin script results into three different transcriptions (T1, T2 and T3) as depicted in Fig. 5. The alignment of these transcription using ROVER is shown in Fig. 5 (a) and (b), where results of T1 and T2 are aligned first in a single WTN and subsequently T3 is aligned with this network.

3. Results

Here, we present the recognition results on 3D words of Latin and Devanagari scripts using proposed classifier combination framework. The results are computed by dividing the dataset into training, validation and testing sets with a share of 70%, 10% and 20%, respectively.

3.1. Dataset description

The dataset consists of two different scripts, namely, Latin and Devanagari. For Latin script, we have analyzed the 3D handwriting dataset proposed in [15], where dataset consist of 1600 Latin words with multiple attempts of 40 isolated words drawn by multiple users over the Leap motion sensor. We have now increased the size of this dataset by recording 10 additional words from different writers. Now, it has 2000 samples of 50 Latin words. In addition to this, a 3D handwriting dataset of 50 Devanagari words is proposed in this work with the help of 15 writers. Each word is repeated 5 times by every user, thus, a total of 3750 words are recorded. The length of the Devanagari words in the dataset varies from 2 to 5 characters. Since, there is no predefined reference plane for writing in the 3D space, high variation can be seen within the same words when drawn by different users as depicted in Fig. 6. The dataset is made online[1] for the research community.

[1] https://sites.google.com/site/iitrcsepradeep7/.

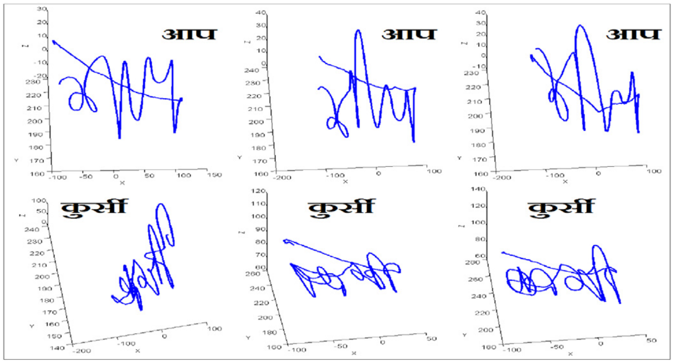

Fig. 6. Variations in the 3D handwriting for two different words of Devanagari script written by three different users.

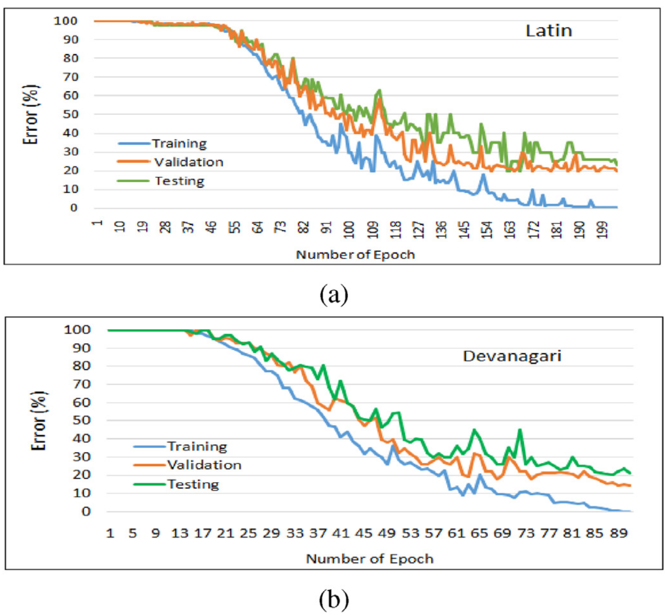

Fig. 7. Learning curves for individual recognizers of: (a) Latin script (b) Devanagari script.

3.2. Results

Recognition of 3D words has been performed by training the BLSTM-NN classifier on different feature sequences separately. Therefore, for each feature, we have obtained four individual recognition systems corresponding to four features of each script, i.e., Latin and Devanagari. Moreover, five recognizers have been obtained for each feature by changing the initial weights of the network during training which ultimately leads to better recognition. Thus, a total 20 recognizers have been obtained for each script. The network has been trained by keeping the learning rate $1e-4$ and a momentum of 0.9. The training of CTC network has been performed for character based recognition present in the 3D text sequences. The learning curves of both Latin and Devanagari scripts corresponding to a feature sequence are depicted in Fig. 7. The learning curves show the decay in training, validation and test classification errors.

The recognition of the 3D words has been performed at character level without lexicon information. The best character recognition performances of individual classifiers in each feature of both scripts are presented in Table 1, where maximum accuracies of 67.90% and 68.10% have been recorded using 3D curvature features in Latin and Devanagari scripts, respectively.

Table 1
Best character recognition performance of individual classifiers.

Feature	Devanagari (%)	Latin (%)
Raw	50	52.40
Convex	68.10	55
Curvature	58.40	67.90
Writing direction	63.80	65.20

Next, all the 20 BLSTM-NN recognizers of each script have been ensembled using the ROVER framework defined in Section 2.3. The framework first aligns the word sequences in a WTN and then, a voting scheme has been used to select the best transcription. The recognition results of the ROVER framework are depicted in Fig. 8, where average accuracies of 72.25% and 71.86% have been recorded in Latin and Devanagari scripts, respectively. It can be seen from the Fig. 8 that minimum and maximum accuracies are varies from 33.33% to 100% for different 3D words of both scripts. Such variations in accuracies are discussed in Section 3.3.

3.3. Error analysis

In this section, we provide the details of the 3D words that are not correctly recognized by the proposed framework in both Latin and Devanagari scripts. It can be seen from the Fig. 8(a) and (b) that recognition rates of various words are below the average performance. It is because of the noisy data recorded by different writers while drawing text in air over the Leap motion sensor. Therefore, writer specific properties such as slow writing, distorted signal, etc. make recognition performance low. Moreover, due to continuous writing with single stroke, many character in the words look similar that ultimately results into wrong recognition. Examples of such noisy words are depicted in Fig. 9, where two confusing and noisy words are shown that are not recognized correctly by the proposed system,

3.4. Comparative study

Here, we perform the comparative analysis between the best recognition performance of individual classifiers and the proposed classifier combination framework for both scripts. The comparison is shown in Fig. 10, where the proposed framework in both Latin and Devanagari scripts outperform the individual recognizers by a margin of 4.35% and 3.76%, respectively.

In this work, we have used an additional 3D convex hull based feature for both scripts. Therefore, to get the impact of such features, recognition results are computed without including them in ROVER framework. A loss of 3.12% is recorded in the results of Devanagari script. However, not much loss in accuracy is recorded in the recognition of Latin script when recognition is performed without convex hull features using ROVER. To have an idea of comparison results obtained using ROVER scheme and lexicon based approach, we have tested the system by providing lexicon to the BLSTM-NN classifier trained ones on all features. The results are depicted in Fig. 11, where accuracies of 83.50% and 79.46% were recorded for Latin and Devanagari scripts, respectively. It can be seen from the Fig. 11 that lexicon based approach results in better recognition performance.

In addition, a comparison with state of the art techniques is also performed for 3D word recognition in Latin script. The authors in [15] have proposed the 3D handwriting recognition system using Leap motion sensor. The recognition was performed using multiple features, where an accuracy of 69.30% was recorded without lexicon using BLSTM classifier for 40 Latin words. The comparison is presented in Table 2, where the recognition performance of

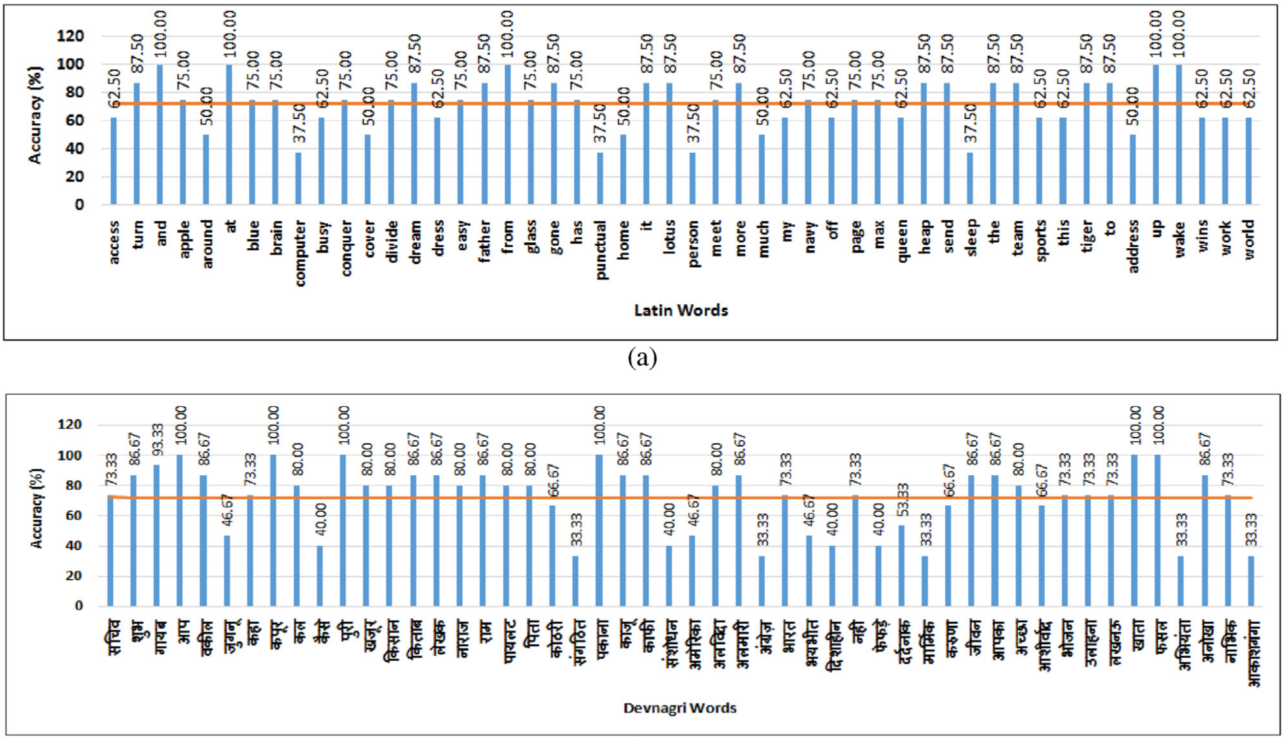

Fig. 8. 3D word recognition rates using classifier combination with the help of ROVER framework for: (a) Latin script (b) Devanagari script. Note: The average accuracy of both the scripts is marked with the red line in the figures. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

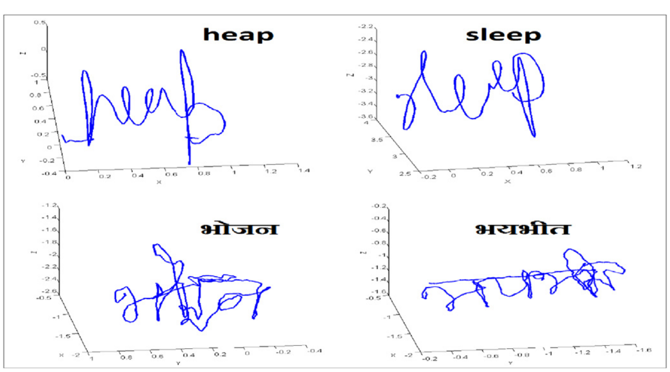

Fig. 9. 3D plots of two different words that are noisy and share similar characteristics that leads to wrong recognition.

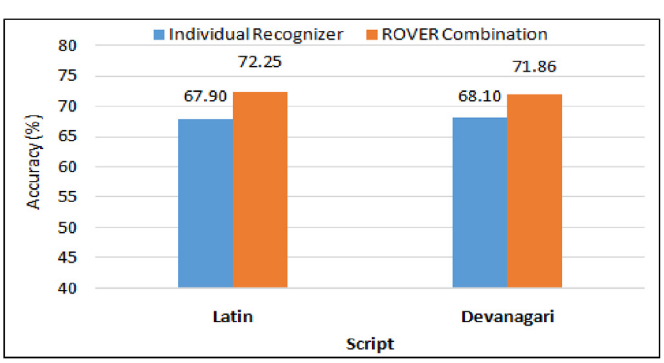

Fig. 10. Comparative performance analysis between the proposed classifier combination scheme and the best individual recognizer for both Latin and Devanagari scripts.

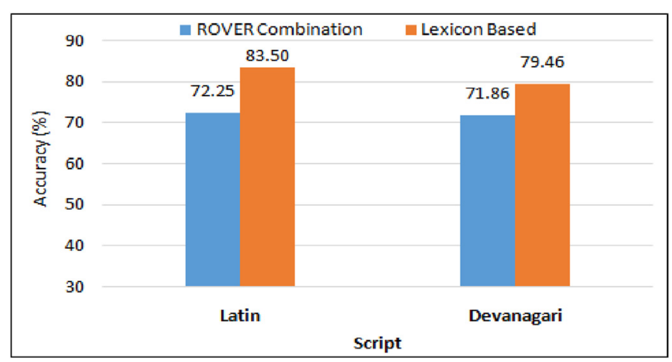

Fig. 11. Comparative performance analysis of recognition rates obtained using ROVER and Lexicon based approach for both Latin and Devanagari scripts.

Table 2
Comparison of proposed methodology with existing technique.

Ref.	Approach	Dataset	Accuracy (%)
Kumar et al. [15]	3D handwriting, writing direction, curvature, HMM, BLSTM-NN	40 Latin words	69.30%
Proposed methodology	3D fingertips, 3D convex-hull, writing direction, curvature, BLSTM-NN, ROVER	50 Latin, 50 Devanagari words	72.25% (Latin) 71.86% (Devanagari)

proposed classifier combination framework outperforms the existing work with a margin of 2.95% when tested for 50 Latin words. The results were also computed on the dataset proposed by Kumar et al. [15] consisting 40 Latin words, where we got an accuracy of 73.33% using the proposed ROVER framework.

4. Conclusion

In this paper, we have proposed a lexicon free approach for 3D handwriting recognition for Latin and Devanagari scripts using Leap motion. Multiple BLSTM-NN classifiers have been learned on different features extracted from the raw text. Moreover, various recognizers have been obtained by initializing the network with different weights. Finally, ROVER framework has been used to combine the transcriptions of the different classifiers. The framework has been tested on Latin and Devanagari scripts and obtained 72.25% and 71.86% accuracies, respectively. The proposed classifier combination methodology outperforms the recognition performance of individual classifiers which demonstrate the robustness of our work. In future, recognition rates can be improved by exploring more robust features that better represent the words. Moreover, other sequential classifiers such as HMM can be used to improve the overall recognition performance.

References

[1] R. Bertolami, H. Bunke, Multiple classifier methods for offline handwritten text line recognition, in: International Workshop on Multiple Classifier Systems, 2007, pp. 72–81.
[2] R. Bertolami, H. Bunke, Hidden markov model-based ensemble methods for offline handwritten text line recognition, Pattern Recognit. 41 (2008) 3452–3460.
[3] N. Das, R. Sarkar, S. Basu, P.K. Saha, M. Kundu, M. Nasipuri, Handwritten bangla character recognition using a soft computing paradigm embedded in two pass approach, Pattern Recognit. 48 (2015) 2054–2071.
[4] D. Dhanya, A. Ramakrishnan, P.B. Pati, Script identification in printed bilingual documents, Sadhana 27 (2002) 73–82.
[5] J. Fan, J. Yang, D. Ai, L. Xia, Y. Zhao, X. Gao, Y. Wang, Convex hull indexed gaussian mixture model (ch-gmm) for 3d point set registration, Pattern Recognit. 59 (2016) 126–141.
[6] J.G. Fiscus, A post-processing system to yield reduced word error rates: recognizer output voting error reduction (rover), in: Workshop on Automatic Speech Recognition and Understanding, 1997, pp. 347–354.
[7] M. Gao, T.T. Cao, A. Nanjappa, T.S. Tan, Z. Huang, ghull: a gpu algorithm for 3d convex hull, ACM Trans. Math. Software 40 (2013) 3.
[8] M. Gao, T.T. Cao, T.S. Tan, Flip to regular triangulation and convex hull, IEEE Trans. Vis. Comput. Graph. 23 (2017) 1056–1069.
[9] A. Graves, S. Fernández, F. Gomez, J. Schmidhuber, Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks, in: 23rd International Conference on Machine Learning, 2006, pp. 369–376.
[10] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, J. Schmidhuber, A novel connectionist system for unconstrained handwriting recognition, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 855–868.
[11] J. Han, L. Shao, D. Xu, J. Shotton, Enhanced computer vision with microsoft kinect sensor: a review, IEEE Trans. Cybern. 43 (2013) 1318–1334.
[12] R. Jayadevan, S.R. Kolhe, P.M. Patil, U. Pal, Offline recognition of devanagari script: a survey, IEEE Trans. Syst. Man Cybern. Part C 41 (2011) 782–796.
[13] Y. Kessentini, T. Burger, T. Paquet, A dempster–shafer theory based combination of handwriting recognition systems with multiple rejection strategies, Pattern Recognit. 48 (2015) 534–544.
[14] P. Kumar, H. Gauba, P.P. Roy, D.P. Dogra, Coupled hmm-based multi-sensor data fusion for sign language recognition, Pattern Recognit. Lett. 86 (2017) 1–8.
[15] P. Kumar, R. Saini, P. Roy, D. Dogra, Study of text segmentation and recognition using leap motion sensor, IEEE Sens. J. (2016).
[16] P. Kumar, R. Saini, P.P. Roy, D.P. Dogra, 3d text segmentation and recognition using leap motion, Multimed Tools Appl. (2016) 1–20.
[17] M. Liwicki, H. Bunke, Combining on-line and off-line bidirectional long short--term memory networks for handwritten text line recognition, in: 11th International Conference on Frontiers in Handwriting Recognition, 2008, pp. 31–36.
[18] R.A.H. Mohamad, L. Likforman-Sulem, C. Mokbel, Combining slanted-frame classifiers for improved hmm-based arabic handwriting recognition, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 1165–1177.
[19] M. Nabiyouni, B. Laha, D.A. Bowman, Poster: Designing effective travel techniques with bare-hand interaction, in: Symposium on 3D User Interfaces, 2014, pp. 139–140.
[20] U. Pal, B. Chaudhuri, Indian script character recognition: a survey, Pattern Recognit. 37 (2004) 1887–1899.
[21] U. Pal, N. Sharma, T. Wakabayashi, F. Kimura, Off-line handwritten character recognition of devnagari script, in: 9th International Conference on Document Analysis and Recognition, 2007, pp. 496–500.
[22] L.E. Potter, J. Araullo, L. Carter, The leap motion controller: a view on sign language, in: 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, 2013, pp. 175–178.
[23] P.P. Roy, U. Pal, J. Llados, M. Delalandre, Multi-oriented and multi-sized touching character segmentation using dynamic programming, in: 10th International Conference on Document Analysis and Recognition, 2009, pp. 11–15.
[24] P.P. Roy, U. Pal, J. Lladós, F. Kimura, Convex hull based approach for multi-oriented character recognition from graphical documents, in: 19th International Conference on Pattern Recognition, 2008, pp. 1–4.
[25] B. Shi, X. Bai, C. Yao, An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 2298–2304.
[26] T. Shinde, Gamer and robot-an interactive gaming and robot control tool using leap motion, Int. J. Emerging Technol. Comput. Sci. 2 (2017).
[27] B. Su, S. Lu, Accurate scene text recognition based on recurrent neural network, in: Asian Conference on Computer Vision, 2014, pp. 35–48.
[28] B. Su, S. Lu, Accurate recognition of words in scenes without character segmentation using recurrent neural network, Pattern Recognit. 63 (2017) 397–405.
[29] K. Vamsikrishna, D.P. Dogra, M.S. Desarkar, Computer-vision-assisted palm rehabilitation with supervised learning, IEEE Trans. Biomed. Eng. 63 (2016) 991–1001.
[30] S. Vikram, L. Li, S. Russell, Handwriting and gestures in the air, recognizing on the fly, in: Proceedings of the CHI, 2013, pp. 1179–1184.
[31] W. Wang, A. Brakensiek, G. Rigoll, Combination of multiple classifiers for handwritten word recognition, in: 8th International Workshop on Frontiers in Handwriting Recognition, IEEE, 2002, pp. 117–122.
[32] F. Weichert, D. Bachmann, B. Rudak, D. Fisseler, Analysis of the accuracy and robustness of the leap motion controller, Sensors 13 (2013) 6380–6393.
[33] L. Xu, A. Krzyzak, C.Y. Suen, Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Trans. Syst. Man Cybern. 22 (1992) 418–435.
[34] N. Xu, W. Wang, X. Qu, Recognition of in-air handwritten chinese character based on leap motion controller, in: International Conference on Image and Graphics, 2015, pp. 160–168.