IJE TRANSACTIONS A: Basics Vol. 28, No. 7, (July 2015) 984-994

International Journal of Engineering

Journal Homepage: www.ije.ir

Offline Language-free Writer Identification Based on Speeded-up Robust Features

M. K. Sharma*, V. P. Dhaka

Department of Computer Engineering, Jaipur National University, Jagatpura Jaipur Rajasthan, Jaipur, India

PAPER INFO

Paper history:
Received 15 January 2015
Received in revised form 07 March 2015
Accepted 11 June 2015

Keywords:
Speeded-up Robust Features Descriptors
Codebank
Scale Orientation
Word Regions

ABSTRACT

This article proposes offline language-free writer identification based on speeded-up robust features (SURF), going through training, enrollment, and identification stages. In all stages, an isotropic Box filter is first used to segment the handwritten text image into word regions (WRs). Then, the SURF descriptors (SUDs) of word region and the corresponding scales and orientations (SOs) are extracted. In the training stage, an SUD codebank is constructed by clustering the SUDs of training samples. In the enrollment stage, the SUDs of the input handwriting adopted to form an SUD signature (SUDS) by looking up the SUD codebank and the SOs are utilized to generate a scale and orientation histogram (H_{SO}). In the identification stage, the SUDS and H_{SO} of the input handwriting are extracted and matched with the enrolled ones for identification. Experimental results on eight public data sets demonstrate that the proposed method outperforms the state-of-the-art algorithms.

***doi**: 10.5829/idosi.ije.2015.28.07a.04*

1. INTRODUCTION

Auto offline language-free writer identification is very important for digitalization of handwritten documents, automation of postal codes identification, bank cheque's signature identification and forensic(s) analysis, etc. Wide research work has been conducted in this field [1-12]. Number of national, international handwriting identification contests [13-17] have been organized successfully. Debashis Ghosh et al. [18] presented a comprehensive review of earlier research literatures with respect to writer identification. Therefore, the existing offline language-free writer identification approaches can be classified in texture and structural based approaches.

Texture related approaches take offline handwritten texts as a texture-image and extract the textural features for writer identification [2-7, 13].

The structural features of offline handwritten texts are much more prominent and committed for handwriting writer identification compared to the textural features based identification. Therefore, recently huge numbers of researches are focused on the structure based approaches for handwriting identification [14-22].

The existing structural approaches are contours driven, so they are easily affected by the slants and aspect ratio of the characters in handwritten texts. To deal with the discriminabilities, we propose a speeded-up robust features (SURF) based key point extraction at word level from handwritten texts, containing the structural features information of words.

The rest of the work is organized as follows: Section II gives a description of the proposed method in detail. Section III reports the experimental results and analyses. Finally, in Section IV the conclusions are presented.

2. METHODOLOGY

2. 1. The Framework of the Proposed Method

The proposed method consists of three stages: training, registration, and identification, as shown in Figure 1.

As shown in Figure 1, there are five main sections in the framework, i.e., line segmentation, word segmentation, codebank generation, feature extraction, and feature matching and fusion.

*Corresponding Author's Email: manoj186@yahoo.co.in (Manoj Kumar Sharma)

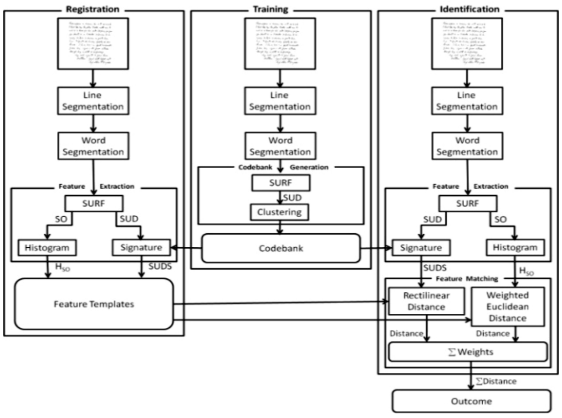


Figure 1. The structure of the proposed method.

In the above three stages, the handwritten text image is firstly segmented into word regions (WRs). Then the SURF is used to extract their SURF descriptors (SUDs) and to detect the keypoints and orientations (SOs) from the LRs and WRs. The SURF descriptors and orientations will be used in different stages.

In the training stage, SURF descriptors are used to generate a codebank for the use of enrollment and identification. In the registration, features SURF descriptors signature (SUDS) and SURF orientations histogram (SOH), are extracted from SUDs and SURF orientations of WRs. In the identification stage, the SURF descriptors signature and SURF orientations histogram are extracted from the input handwritten text images and respectively harmonized with the registered ones to get two harmonized distances, which are then fused to form the final harmonized distance for decision.

2. 2. Word Segmentation The text image is segmented into word regions (WRs) and their structural features are extracted. In earlier literatures, handwritten text images are manually segmented [23], which is very tedious and time taking. Many automatic word segmentation techniques have been devised in recent years, and are mostly based on line segmentation [24, 25]. To overcome the problems of line segmentation, we have used an isotropic Box filter to segment words from handwritten text images.

Algorithm 1: Word_Segmentation

Step 1: Input an offline handwritten text image I (Figure 2.(a)). The word segmentation process is as follows:

Step 2: Convert text image I to gray scale image and then into a binary image I_b using Multi-Otsu's method [See Figure 2(b)].

Step 3: Compute average height h_a of all connected components (CCs) in I_b.

Step 4: Filter binary image I_b with an isotropic Box filter and get the filtered image I_f [See Figure 2(c)].

(a) Find space between CCs using Binary Image Euclidean Distance Transform.

i. Intra-word Space $(S_{sw}) = bwdist(WRs)$.

ii. Inter-word Space $(S_{dw}) = bwdist(WRs)$.

(b) If $((S_{sw}) < (S_{dw}))$ than

Concatenate the CCs of Intra-word.

Else

Separate the CCs of Inter-word.

(c) If $((S_{sw}) \equiv h_a)$ for all CCs in I_b than

Variance of filter $\sigma = 2.5 \times h_a$

Step 4: Get filtered binary image I_{fb} by Binarizing filtered image I_f using the threshold obtained by Multi-Otsu's method.

Step 5: Form Semi-word regions (SWRs) by assigning connected component in I_b to the nearest connected region of I_{fb}.

Step 6: Extract the word-regions (WRs) according to the distances between the adjacent SWRs by merging the SWRs.

Step 7: Split overlapping connected components (OCCs).

However, the word segmentation divides text image into many word regions. These segmented word regions will be used for feature extraction.

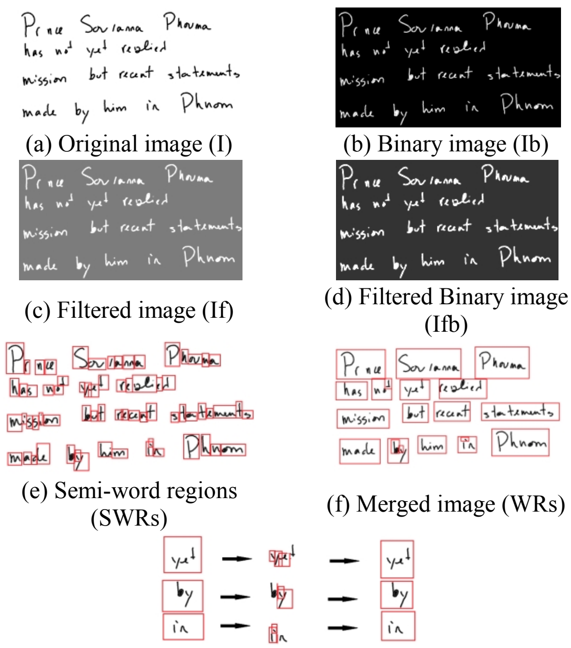


Figure 2. Word segmentation process.

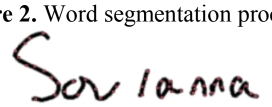


Figure 3. The keypoints detected in a WR by SURF.

2. 3. SURF Speeded-up Robust Features (SURF), was presented by Herbert [26] for a novel scale and rotation-invariant detection. The SURF algorithm generates keypoints and descriptors efficiently by employing efficient scale-space construction. The SURF algorithm has four stages of computation:(1) employing integral images (2) scale-space construction, keypoints generation, (3) keypoints descriptor, (4) reproducible orientation. In the first stage, the integral images allow the speeded-up computation of approximate Laplacian of Gaussian images using box filter. Due to integral image representation, computation overheads applying the Box filter which is independent from the size of the filter. In the second stage, the Hessian matrix is used to detect the keypoints, and computes the locations, reproducible orientations and scale-space of these keypoints, by keeping the constant image size and variable filter size.

The first stage results in invariance to location and scale and in the final stage, detected keypoints are first assigned to a reproducible orientation. A Haar-wavelet responses for orientation in x and y, which are calculated for a number of pixels within range of 6σ, where σ denotes keypoints scale. Then SURF descriptor for detected keypoints is generated by imposing a window around the keypoint and oriented as per obtained orientation. This window imposed around the keypoint further divides in 4 × 4 sub-windows and calculates Haar-wavelets of size 2σ in each sub-window. In the construction of 64D descriptor vectors, each sub-window contributes 4 values and then normalized to the unit length. The SURF resulting descriptor is invariant to scale and rotation and partially invariant to other transformations.

The SURF algorithm is used in this work to detect the keypoints of handwritten scripts, their SURF descriptors (SUDs), and the corresponding space-scales and reproducible orientations (SROs). The SURF's descriptors are not only faster but also scale and in-plane rotation-invariant and can reflect the skeleton of the image regions centered at the keypoints and the SROs can store the space-scale and reproducible orientation related information of these skeletons.

2. 4. Codebank Generation The SURF algorithm is used to find keypoints (shown in Figure 3) for every word region and extracts their scales, descriptors and orientations. To resolve the storage and complexity issues, we consider the limited number of features, and cluster the SUDs of the key points into N categories and center of each category, called code, forms a SUD codebank with size N. On behalf of the codebank, a histogram will be computed with fixed dimension as feature vector for script identification in following subsection. For codebank generation we used the hierarchical EM clustering algorithm, which has been successfully used for clustering dynamic textures [27], and size N of SUD codebank is empirically selected as 450.

2. 5. Feature Extraction For feature extraction and matching, we have employed the occurrence frequency of SUD and SO in the handwritten text images in place of position of keypoints, because each text image may have different keypoints layout.

1) SURF Descriptor Signature (SUDS) Extraction: Let $SUD = \{d_1, d_2, ..., d_n\}$ denote n SUDs, and let $CB = \{c_1, c_2, ..., c_N\}$ denote a SUD codebank with size N. The SUDS feature extraction process is as follows.

1. SUDS feature vector initialization $SUDS = (0, 0, ..., 0)$

2. Compute the Euclidean distance between $SUDS$ and each code word $c_j \in CB$ for each $d_i \in SUD$

$$ED_{ij} = \sqrt{\sum_{k=1}^{L}\left(d_{ik} - c_{jk}\right)^2} \tag{1}$$

where $L = 128$ is the size of the SUDs, and a Euclidean distance vector EDV for d_i is obtained as below:

$$EDV = \left(ED_{i1}, ED_{i2}, ..., ED_{iN}\right) \tag{2}$$

3. Obtain top t component's index in EDV by sorting EDV components.

$$INDX = \left(indx_1, indx_2, ..., indx_t\right) \tag{3}$$

4. Update the SUDS feature vector for each $indx \in INDX$, as follows:

$$SUDS_{indx} = SUDS_{indx} + \delta\left(EDV_{indx}\right) \tag{4}$$

where $\delta(x)$ is a non-increasing function.

5. Repeat step 2 to 4 until all $SUDs$ are processed.

6. Compute the final $SUDS$ vector as follows:

$$SUDS_i = \frac{SUDS_i}{\sum_{j=1}^{N} SUDS_j} \tag{5}$$

The database dependent parameter t is determined by cross-validation of the training dataset and the function $\delta(x)$ is a decreasing function [28]. Therefore, we employ the constant function $\delta(x) = 1$ to extract SUDS features.

2) Histogram Extraction: The text images are decomposed into X octaves and Y sub-levels in each octave, i.e., $Z = (X \times Y)$ scales, using SURF. Let $SU = \{su_1, su_2, ..., su_n\}$ denote n SURF keypoints scales, $1 \le su_i \le Z$, and let $SO = \{so_1, so_2, ..., so_n\}$ denote the orientations of SURF keypoints. Using an angle ϕ, the

orientation $[0,360]$ can be quantized to $UO_{bin} = \left\lceil 360/\phi \right\rceil$ intervals, and get the nearest integer using $\lfloor x \rfloor$ operator for every $\lfloor x \rfloor \geq x$. The process of Scale and Orientation Histogram (H_{SO}) feature extraction is as follows.

1. Initialize the H_{SO} feature vector with size $H_M = Z \times UO_{bin}$ by $H_{SO} = (0,0,...,0)$;
2. Compute index $indx$ in H_{SO} for each key point's scale and orientation, $su_i \in SU$ and $so_i \in SO$

$$bin = \lceil uo_i / \phi \rceil$$
$$indx = UO_{bin} \times (us_i - 1) + bin \tag{6}$$

3. Update the H_{SO} feature vector

$$H_{SO}(indx) = H_{SO}(indx) + 1 \tag{7}$$

4. Repeat steps 2 and 3 until all keypoints are processed.
5. Compute the final H_{SO} feature vector as follows:

$$H_{SO_i} = \frac{H_{SO_i}}{\sum_{j=1}^{H_M} H_{SO_i}} \tag{8}$$

The parameter X and Y are selected empirically by extensive experiments, and angle ϕ is database-dependent which is determined by performing cross-validation on the training dataset. Figure 4 shows the average absolute differences between ten positive and negative pairs for each component of SUDS and SOH. The figure predicts that the difference between inter-handwriting is much larger than the difference between intra-handwriting for both SUDS and SOH, which means that both SUDS and H_{SO} have much discriminability to different handwritings.

Figure 4(b) predicts that the components differences with large index in H_{SO} decrease with respect to index increase. The large indexes correspond to large scales according to the H_{SO}. Therefore, the SURF keypoints detection become much less at large scales and becomes much large at small scales, which devise the decrement of the values of components with large scales with the increase of scale index.

2. 6. Feature Matching and Fusion

I_1 and I_2 are two handwritten text images, $u = (u_1, u_2, ..., u_N)$ and $v = (v_1, v_2, ..., v_N)$ are their SUDSs, and $x = (x_1, x_2, ..., x_M)$ and $y = (y_1, y_2, ..., y_M)$ are their H_{SO}. In this work, the Rectilinear distance which is successfully used in minimax location problem on the plane [29] is adopted to measure the dissimilarity between two SUDSs u and v:

$$D_1(u,v) = \sum_{i=1}^{N} |u_i - v_i| \tag{9}$$

As discussed, the differences of the components with large indexes in H_{SO} are much smaller. Therefore, with Rectilinear distance, the contribution of large index components is very less to measure the dissimilarities between H_{SO}, than the ones with small indexes. The Weighted Euclidean distance, which successfully applied in clustering [30] is used to measure the dissimilarity between H_{SO} x and y:

$$D_2(x,y) = \sum_{j=1}^{M} \frac{(x_j - y_j)^2}{(x_j + y_j)} \tag{10}$$

The fusion of both D_1 and D_2 form a new distance to measure the dissimilarity between I_1 and I_2.

$$D(I_1, I_2) = w \times D_1(u,v) + (1-w) \times D_2(x,y) \tag{11}$$

where $0 \leq w \leq 1$ is a weight which is database dependent and can be determined by cross-validation on the training dataset.

3. RESULTS AND ANALYSIS

(a) Datasets Eight datasets are used: Five of them are English-datasets, i.e., MNIST CD-1 [31], MNIST CD-2 [31], IAM [32], Firemaker [33], and Unipen [34], one Chinese dataset HIT-MW [35], and two hybrid-language datasets, i.e., ICDAR 2011 [36-38] and ICFHR 2012[39, 40]. A brief overview of datasets used in experiments is given in Table 1.

The MNIST CD-1 and MNIST CD-2 datasets [31] contain 700 and 500 sample images from 300 and 250 writers, respectively. There are 168 writers owning 4 or more handwritten samples, but we consider only [31] D-2 datasets are used for training and testing respectively. After fusion of both data ets modified MNIST dataset is renamed as MNIST' dataset.

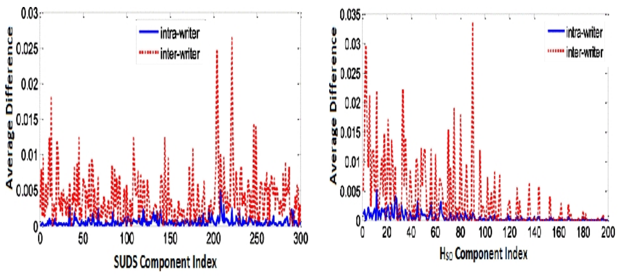
The IAM dataset [32] contains 1839 English sample images from 850 writers. There are 148 writers owning 4 or more handwritten samples, but we consider only two samples from each. The modified IAM dataset is renamed as IAM' dataset. The Firemaker dataset [33] contains 900 handwriting sample pages from 350 writers and three pages per writer. Pages 1 and 2 have a text of four paragraphs in normal handwriting and Page 3 has the content of cartoons written in their own words and Page 1 and Page 3 are used for handwritten text identification. The Unipen dataset contains handwritten from 250 writers and two samples per writer [34]. We modify the database and select 600 samples from 130 writers. The HIT-MW dataset [35] consists of 1000 handwritten Chinese samples, from 250 writers. Most of the writers have two pages of handwritten sample text. We modify the HIT-MW dataset as done elsewhere [41], only one page from each writer is used.

TABLE 1. Datasets

Dataset	Writers	Language
MNIST CD-1	300	English
MNIST CD-2	250	English
IAM	850	English
Firemaker	300	English
Unipen	250	English
HIT-MW	250	Chinese
ICDAR2011	80+240	English, German, French, Greek
ICFHR2012	70+80	German, English, Greek

TABLE 2. Soft Top-N performance of the features extracted from IAM' dataset

Level	SUDS (%)		H_{SO} (%)	
	Top-1	Top-15	Top-1	Top-15
Page	77.5	93.3	67	89
Paragraph	85.3	92.2	82.3	91.5
Word	**95.2**	**99.1**	**79.4**	**94.7**
Alphabets	74	89.1	69.3	89.5



(a) SUDS feature differences. (b) HSO feature differences.

Figure 4. SUDS and HSO b/w intra-writer and inter-writer.

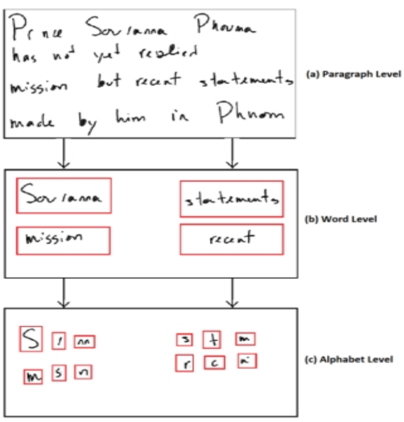


Figure 5. Different levels feature extraction.

The ICDAR2011 dataset [36] consists of 80 writers for training dataset and 240 writers for testing dataset. Each writer copied ten pages with three language text (English, German and French) and each language has four pages. A new modified dataset, called MICDAR2011 dataset, is built by considering only the first four text lines from each sample page. The ICFHR2012 dataset [39] consists of 70 writers for training dataset and 80 writers for testing dataset. Each writer copied ten pages and each page contains text in two languages (German and English) and each page contains of text lines in range of four to eight.

(b) Criterions We considered "leave-one-out", soft Top-N and hard Top-N criterion for each dataset. In "leave-one-out" strategy, the distances to all other samples of datasets are computed for each handwritten sample [34]. In soft Top-N criterion only the correct hits are considered, as at least one sample document of the same writer is considered in the N similar sample documents [34-36]. In hard Top-N only the correct hits are considered for all N similar sample documents written by the same writer [36] and when at least N reference sample documents exist for a writer then the hard Top-N criterion is used.

3. 1. Experiments with Different Level Features

The feature extraction is proposed at multiple levels i.e. page-level, paragraph-level, word-level, and alphabet-level. Some examples of different levels are shown in Figure 5.

We adopt MNIST, Unipen and IAM' datasets for training and testing respectively to investigate the discriminability between SUDS and H_{SO} at different levels, and for experiments we employed soft Top-N and "leave-one-out" strategy. The different levels of experimental results are presented in Table 2. SUDS and H_{SO} produce better experimental performance at the word-level; it means that these features are more powerful at word-level to represent the handwritten text than other levels.

Thus, the words are detectable by either leading or trailing space, and their features extraction is stable. At page-level features extraction, multiple features like paragraph-level features, word-level features and the features at line-level are to be extracted which are not stable and hence not suitable for offline handwritten identification. At the paragraph-level features extraction, multiple features like word-level features and the features at line-level are to be extracted which are not stable and hence not suitable for offline handwritten identification. Alphabet-level feature which are subset of word-level features misses various features between alphabets in the same words. These missing features have strong discriminability for different offline handwritten text. For word-level features extraction with IAM and HIT-MW datasets, the word regions is being used with the datasets and for the rest of the datasets the word regions are segmented for word-level feature extraction through proposed word segmentation algorithm.

TABLE 3. Soft Top-N enactment on the IAM' dataset (%)

Methods	Writers	Top-1	Top-5	Top-10	Top-15
LPQ [8]	675	N/A	95.5	93.4	N/A
Line Fragment [15]	675	92.3	N/A	96.9	98.7
Grapheme emission [29]	675	81.2	89.3	93.4	94.4
GE+GH [29]	675	89.2	N/A	96.7	97.3
Contour-hinge [29]	675	81.3	N/A	92.5	93.2
GMF[36]	677	90.6	95.6	95.3	95.3
Quill-Hinge[37]	677	96	N/A	97	98
Contour-hinge[37]	677	95	96.3	96	95
Siddiqi[38]	659	92	N/A	98	96
SDS	659	93.2	97.6	97.9	97.9
H_{SO}	659	79.4	92.3	94.7	94.7
SDS+H_{SO}	**659**	**99.5**	**99.4**	**99.8**	**99.7**

TABLE 4. Soft Top-N enactment on the NMIST CD-2 dataset (%)

Methods	Top-1	Top-5	Top-10	Top-15
Line Fragment [15]	92.8	93.4	99.6	92.8
GH [26]	83	N/A	93	82
Grapheme emission [29]	77	92	91	76
GH+GE [29]	84	N/A	95	84
GMF [36]	79	91.4	93.4	79
Quill-Hinge [37]	85	N/A	98	87
SDS	91.6	96.2	98.4	92.6
HSO	74.8	91.2	93.8	74.8
SDS+H_{SO}	**96.4**	**97.2**	**99.8**	**98.4**

3. 2. Performances on Public Datasets

3. 2. 1. Experiments on English Datasets

In the earlier research works, [20 , 31 , 34], MNIST CD-1 dataset and Unipen dataset was used for training, and Firemaker, MNIST CD-2 and IAM' datasets were used for testing and evaluation, and analysis of the proposed methodology is done through the "leave-one-out" and the soft Top-N. Table 3 lists the Soft Top-N enactment on the IAM' dataset of different approaches and predicts that the presented results in a previous work [42] are better than the results of another study [34] for the same Contour-hinge feature. Most of the writers were referred to 2 to 58 samples during recognition [42], while each writer was referred to only one sample in the other work [34], the same strategy is adopted for the proposed method, even though the proposed method outperforms all of the state-of-the-art approaches, including the one proposed in the literature [42] with SUDS and H_{SO}. Thus, all 850 images of IAM' dataset are used in testing while some earlier approaches [20, 34 , 43] only use 650 samples.

Table 4 lists the handwritten text recognition results of different approaches using NMIST CD-2 dataset. According to the results presented in Table 3 and Tables 4 it is concluded that the Top-1 performances of different approaches in IAM' dataset are better than those in NMIST CD-2 dataset. The possible reason is that since the average amount of handwritten text images of IAM' dataset are more than those in NMIST CD-2 dataset, the extracted features from handwritten text images of IAM' dataset are more stable than those extracted from NMIST CD-2 dataset.

Table 5 lists the handwritten text recognition results of different approaches using Firemaker dataset. According to the results presented in Tables 3, 4 and 5, it is concluded that the Top-1 performance of different approaches in IAM' dataset are better than those in Firemaker dataset. The results presented in Tables 3, 4 5 also predict that the fusion of SUDs or H_{SO} can outperform some state-of-the-art approaches, meaning that SUDS and H_{SO} characterize the different aspects of the handwritten text and can complement each other.

3. 2. 2. Experiments on Chinese Datasets

In Chineses dataset, U-HIT-MW dataset is used for training and L-HIT-MW dataset is used for testing purposes. As presented in the literature [41], one handwritten text image in L-HIT-MW is used for the query and the other one is used for reference. Query handwritten text images are compared with corresponding reference images. Table 6 lists the handwritten text recognition results of different approaches on the dataset.

TABLE 5. Soft Top-N performance of different approaches on the firemaker dataset (%)

Methods	Top-1	Top-5	Top-10	Top-15
Contour-hinge(GH)[29]	82	N/A	92	81
Grapheme emission(GE) [29]	76	93	92	72
GH+GE[29]	83	N/A	94	83
GMF[21]	79.3	92.4	94.4	77
Quill-Hinge[22]	84.5	94.4	97	88
Line Fragment [24]	92.3	N/A	98.6	91.8
SDS	91.2	95.2	99.4	93.6
HSO	74.3	92.2	94.8	75.8
SDS+H_{SO}	**95.4**	**96.3**	**98.2**	**97.3**

TABLE 6. Soft Top-N performance of different approaches for writer identification on the L-Hit-Mw dataset (%)

Methods	Top-1	Top-5	Top-10	Top-15
Contour-hinge(GH)[21]	84.6	95.4	96.7	95.4
Multi-scale Contour-hinge[21]	92.5	97.1	97.5	97.1
GMF[21]	95	98.3	98.8	98.3
SDS	85.8	95.4	98.3	95.4
H_{SO}	85	93.8	95.4	93.8
SDS+H_{SO}	95.4	98.8	99.2	98.8

TABLE 7. Soft and hard Top-N performance on ICDAR2011 dataset and micdar2011dataset (CICDAR2011) (%)

Method	Soft evaluation						Hard evaluation					
	Entire dataset			Entire Cropped dataset			Entire dataset			Entire Cropped dataset		
	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15	Top-2	Top-7	Top-10	Top-2	Top-7	Top-10
TEBESSA	97.6	99	99	89.5	96.6	99.7	97.51	82.3	80	75	54.1	74.4
QUQA-A	91.9	97.1	99	78	92.8	96.4	77.4	62.3	70.2	62.4	65.9	63.4
QUQA-B	97.1	98.5	100	69.3	92.8	94.6	91.3	79.4	61.4	57.6	52.6	66.3
TSINGHUA	98.5	99	100	92.9	97.6	98.5	96.2	85.1	71.4	89.8	68.6	72.5
ECNU	85.6	98	89.9	68.9	87.7	86.9	61	86.9	60	59.4	52.9	60
GWU	92.8	99.1	98	78	95.4	94.2	82.3	74.2	60.2	61.4	70.2	66.3
CS-UMD	98.5	98.5	98.5	67.8	86.7	88.9	90.8	79.9	52.1	71.9	72.1	63.4
MCS-NUST	99.4	98.8	99.5	84.2	97.6	96.6	92.3	88.9	78.9	81.6	75.6	51.1
SUDS+H_{SO}	**99.7**	**100**	**100**	**96.2**	**99.6**	**100**	**99.6**	**92.4**	**88.9**	**95.5**	**83**	**82.3**

TABLE 8. Soft Top-N performance on ICDAR2011 different languages sub-dataset (%)

Method	English			French			Greek			German		
	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15
TEBESSA	21.5	29.2	43.1	23.5	25.4	29.2	56.2	56.2	66.2	41.2	41.1	46.9
QUQA-A	54.2	63.1	80.4	65.8	85.0	84.7	61.9	85.5	95.3	81.2	96.5	92.2
QUQA-B	54.6	86.9	90.8	73.5	94.4	86.2	68.1	82.6	95.5	64.2	94.6	94.3
TSINGHUA	61.9	88.1	88.1	86.9	92.2	100	90.8	86.2	94.2	94.6	95.2	100
ECNU	52.3	75.4	86.9	60.0	79.2	92.77	77.7	96.5	94.3	79.2	98.5	91.3
GWU	60.4	77.3	66.9	54.2	79.2	92.7	79.6	94.6	92.4	83.1	92.4	94.2
CS-UMD	62.3	90.8	82.3	79.2	86.5	100	73.5	84.4	92.2	81.2	89.5	98.1
MCS-NUST	75.8	94.6	96.2	77.3	85.5	94.3	85.4	88.3	95.2	88.9	92.2	99.1
SUDS+H_{SO}	**80.8**	**96.2**	**100**	**88.5**	**98.1**	**100**	**98.1**	**100**	**100**	**92.3**	**100**	**100**

TABLE 9. Soft Top-N performance on MICDAR2011 different languages sub-dataset(%)

Method	English			French			Greek			German		
	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15	Top-1	Top-5	Top-15
TEBESSA	21.5	29.2	33.1	23.5	25.4	29.2	56.2	56.2	56.2	31.2	31.1	36.9
QUQA-A	54.2	83.1	92.4	65.8	85.0	85.7	61.9	87.5	94.3	81.2	85.5	96.2
QUQA-B	44.6	86.9	88.8	68.5	97.4	96.2	78.1	84.6	86.5	74.2	88.6	97.3
TSINGHUA	61.9	98.7	98.8	75.9	95.2	100	88.8	94.2	94.2	94.6	95.2	100
ECNU	52.3	75.4	86.9	70.0	79.2	87.77	67.7	88.5	95.3	79.2	89.5	95.3
GWU	60.4	77.3	86.9	64.2	79.2	85.7	79.6	74.6	92.4	83.1	92.4	94.2
CS-UMD	62.3	86.8	95.3	79.2	89.5	100	73.5	80.4	92.2	81.2	87.5	94.1
MCS-NUST	65.8	86.6	96.2	77.3	86.5	95.3	75.4	94.3	94.2	88.9	93.2	97.1
SUDS+H_{SO}	**88.88**	**97.2**	**100**	**98.5**	**99.1**	**100**	**98.7**	**100**	**100**	**96.3**	**100**	**100**

3. 2. 3. Experiments on Hybrid Datasets For hybrid dataset, ICDAR2011 dataset is used for training and ICFHR2012 dataset is used for testing purposes. As there are ten text documents for a given handwritten text with ICDAR2011and ICFHR2012 dataset, it is possible to do hard evaluation with Top-8 criterions: one page for query and other eight as the references. The soft and hard Top-N performances on ICDAR2011 and ICFHR2012 datasets are given in Tables 7 and 10, respectively and the results predicts that all the approaches outperform all of the state-of-the-art approaches with soft Top-N criterion instead of those with hard Top-N criterion. The performance degradation of the proposed methodology is less when we switched the criterion from the soft Top-N to the hard Top-N, it means the proposed methodology is much stable than other approaches. The performances of different language sub-datasets of ICDAR2011 and MICDAR2011 are presented in Tables 8 and 9. The results presented in these tables predict that proposed methodology outperforms in all tests, except the soft Top-5 on Greek cropped sub-dataset. All the approaches outperform with the ICDAR2011 dataset and its sub-datasets instead of those in MICDAR2011and this performance is due to the less number of handwritten lines in MICDAR2011. As the results in these tables show, the proposed methodology is more robust to the amount of handwritten text documents than other earlier presented approaches.

To further investigate the robustness of the proposed methodology, the testing is again done on the ICDAR2011 different languages sub-datasets. Two images for each writer from the testing dataset are used for reference sample and the rest is used as a query. With the fixed reference samples a different number of segmented words of the querying image are randomly selected and the Top-1 performance of the proposed methodology is evaluated for 140 times. Figure 6 shows the average performances of Top-1 with different number of handwritten words; it is reflected from the graph that the raising of performance of different datasets is directly proportional to the increase of the handwritten word, when the handwritten word increases to 16, 23 and 64, the Top-1 performance for German, English and French increases accordingly and for Greek, the Top-1 performance converges to 97.45% when the handwritten word increases to 26. As the results Figure 6 show, the performance of Top-1 for German, English and French exceeded 99.4% when the images contain only 6, 8 and 10 handwritten words and for Greek it exceeded 95.4% when the handwritten text image contains only 16 words, which means the proposed methodology can better perform on the less amount of handwritten text images. The soft Top-N performance of ICFHR2012 entire dataset and sub-datasets is shown in Table 10, which predicts that the proposed methodology outperforms the state-of-the-art approaches. The soft Top-N performance of ICFHR2012 English and Greek sub-datasets is shown in Table 11, which predicts that the proposed methodology outperforms the state-of-the-art approaches. Therefore, it is observed that the proposed methodology performs well with English, Chinese, French, German, Greek and hybrid-languages and it also came in notice from Table 3 to 11 that the performance of the proposed methodology slightly varies with dataset of different languages and it is because of distinct word structures in different languages.

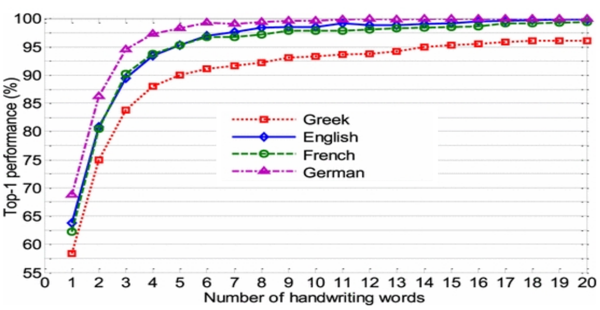


Figure 6. Top-1 performance of the proposed method with different number of handwriting words on ICDAR2011 dataset.

TABLE 10. Soft and hard Top-N performance on ICFHR2012 entire dataset (%)

Method	Soft evaluation				Hard evaluation	
	Top-1	Top-2	Top-5	Top-15	Top-5	Top-7
QATAR-A	80.3	84.8	92.8	96.3	34.3	21.3
QATAR-B	82.0	83.3	96.0	97.0	35.0	25.3
TSINGHUA	94.8	97.8	95.8	97.3	57.5	37.3
TEBESSA-A	95.3	97.5	99.1	98.0	67.5	49.0
TEBESSA-B	88.8	95.3	98.8	97.8	67.5	49.3
TEBESSA-C	77.5	98.3	99.2	98.3	75.0	47.8
HANNOVER	88.5	95.3	96.3	98.3	61.5	32.8
SUDS+H_{SO}	**98.7**	**99.3**	**99.8**	**99.8**	**69.4**	**48.0**

TABLE 11. Soft Top-N performance on ICFHR2012 sub-dataset (%)

Method	Soft evaluation				Hard evaluation			
	Top-1	Top-2	Top-5	Top-10	Top-1	Top-2	Top-5	Top-10
TEBESSA-a	90.0	97.0	98.0	97.5	93.0	96.0	99.8	99.2
TEBESSA-b	84.0	91.0	97.0	98.0	86.5	94.5	96.5	99.3
TEBESSA-c	92.5	96.5	98.5	97.0	94.5	98.0	98.5	99.8
QATAR-a	54.5	67.5	86.0	93.0	77.0	97.0	98.5	97.5
QATAR-b	73.5	83.5	93.5	97.5	86.5	91.0	97.0	98.8
TSINGHUA	95.0	95.5	96.5	99.0	92.0	95.0	98.9	99.1
HANNOVER	83.0	89.0	92.5	96.0	88.5	94.0	98.8	99.6
SUDS+H_{SO}	**96.5**	**97.0**	**98.0**	**99.5**	**99.4**	**99.8**	**100**	**100**

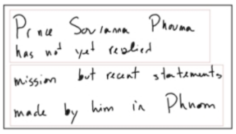


Figure 7. Correctly identified same writer samples and incorrectly identified by the state-of-the-art approaches.

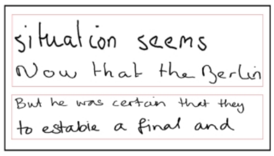


Figure 8. Correctly identified different writers samples and incorrectly identified by the state-of-the-art approaches.

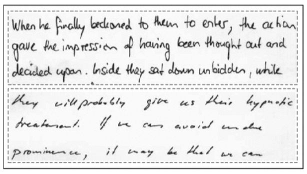


Figure 9. False rejection and false acceptance of the text samples from the same writer.

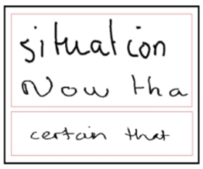


Figure 10. Incorrect identification.

3. 3. Analysis In this section, we re-analyze the performances of Top-1 criterion for the proposed methodology. A handwritten text sample correctly recognized by the proposed methodology and incorrectly recognized by the state-of-the-art approaches as done in Contour-hinge [34], Grapheme emission [34], and GMF [41] is shown in Figure 7. From the results, it is observed that the gap between most of the alphabets is very similar, but their height and width are different. The heights of the characters in the top one are more than those in the bottom, while the widths are similar, with the different aspect ratios between sample characters. Second, the bottom sample is more slant than the top one. There is no slant normalization with the contour based approaches (Contour-hinge, Grapheme emission and GMF) that why they are more sensitive to the slant of the handwritten texts. The size variance problem with the handwritten text may overcome with the size normalization feature of contour based approach. That's why sampled text shown in Figure 7 is incorrectly recognized by the state-of-the-art approaches. The proposed methodology is based on SURF keypoints and hence insensitive to the aspect ratio and slant of handwritten text, that's why the proposed methodology can correctly recognize these handwritten text images.

Figure 8 shows correctly recognized samples by different writers and correctly recognized ones by the proposed method and incorrectly recognized ones by the state-of-the-art approaches, i.e., Contour-hinge[34], Grapheme emission [34], and GMF [41]. As shown in Figure 8, the two handwritings are very similar in their slant, orientation, and the shape, and hence these state-of-the-art approaches based on the contour or alphabets fragments incorrectly identified them from the same writer. However, the alphabets in the bottom sample are more compact than the top one, which indicates that the text is written by different writers.

Figure 9 shows false rejection and acceptance of the text samples written by same writer in different styles. The text shown in the above figure cannot be judged by the experts that they are from the same writer. Due to the heavy distortion in handwriting even with the same writer cannot properly identify by the proposed methodology. Figure 10 shows the wrong identification done by the proposed methodology. In the extracted features, the occurrence frequency of the local structure features is reflected and the amount of handwritten text is too less, the proposed methodology cannot extract stable features and fails to identify correct writer.

4. CONCLUSION

In this research article we have proposed an automatic offline language-free writer identification based on SURF, in which two SURF features, i.e., SUDS and H_{SO}, are extracted from offline handwritten text samples. The experiments on eight public datasets demonstrate that the proposed methodology outrun the state-of-the-art approaches. This method is based on SURF keypoints, which is insensitive to the aspect ratio and slant of the characters. SUDS and H_{SO} are very stable and can reflect the structures around the SURF keypoints. The word-level features of handwritten text are much more suitable to recognize the text. The proposed methodology is language-free and can perform well with different and hybrid languages with their complex structures. The proposed methodology outperforms the state-of-the-art approaches. With the less amount of handwritten text, the performance of the proposed methodology is little bit less, unless otherwise the proposed methodology devises very promising results. In further research, we will improve the performance of the proposed methodology with the small amount of handwritten text samples.

5. REFERENCES

1. Zois, E.N. and Anastassopoulos, V., "Morphological waveform coding for writer identification", ***Pattern Recognition***, Vol. 33, No. 3, (2000), 385-398.
2. Said, H.E., Tan, T.N. and Baker, K.D., "Personal identification based on handwriting", ***Pattern Recognition***, Vol. 33, No. 1, (2000), 149-160.
3. Hanusiak, R., Oliveira, L.S., Justino, E. and Sabourin, R., "Writer verification using texture-based features", ***International Journal on Document Analysis and Recognition (IJDAR)***, Vol. 15, No. 3, (2012), 213-226.
4. Wshah, S., Kumar, G. and Govindaraju, V., "Script independent word spotting in offline handwritten documents based on hidden markov models", in Frontiers in Handwriting Recognition (ICFHR), International Conference on, (2012), 14-19.
5. Helli, B. and Moghaddam, M.E., "A text-independent persian writer identification based on feature relation graph (FRG)", ***Pattern Recognition***, Vol. 43, No. 6, (2010), 2199-2209.
6. Espana-Boquera, S., Castro-Bleda, M.J., Gorbe-Moya, J. and Zamora-Martinez, F., "Improving offline handwritten text recognition with hybrid hmm/ann models", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 33, No. 4, (2011), 767-779.
7. Ryu, J., Koo, H.I. and Cho, N.I., "Language-independent text-line extraction algorithm for handwritten documents", ***Signal Processing Letters, IEEE***, Vol. 21, No. 9, (2014), 1115-1119.
8. Ghosh, D., Dube, T. and Shivaprasad, A.P., "Script recognition—a review", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 32, No. 12, (2010), 2142-2161.
9. Lowe, D.G., "Distinctive image features from scale-invariant keypoints", ***International journal of computer vision***, Vol. 60, No. 2, (2004), 91-110.
10. Nister, D. and Stewenius, H., "Scalable recognition with a vocabulary tree", in Computer Vision and Pattern Recognition, IEEE Computer Society Conference on,. Vol. 2, (2006), 2161-2168.

11. Dhaka, V.P. and Sharma, M.K., "An efficient segmentation technique for devanagari offline handwritten scripts using the feedforward neural network", ***Neural Computing and Applications***, (2011) 1-13.

12. Dhaka, V.P. and Sharma, M.K., "Classification of image using a genetic general neural decision tree", ***International Journal of Applied Pattern Recognition***, Vol. 2, No. 1, (2015), 76-95.

13. Bertolini, D., Oliveira, L.S., Justino, E. and Sabourin, R., "Texture-based descriptors for writer identification and verification", ***Expert Systems with Applications***, Vol. 40, No. 6, (2013), 2069-2080.

14. Bulacu, M. and Schomaker, L., "Text-independent writer identification and verification using textural and allographic features", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 29, No. 4, (2007), 701-717.

15. Bianne-Bernard, A.-L., Menasri, F., Mohamad, R.A.-H., Mokbel, C., Kermorvant, C. and Likforman-Sulem, L., "Dynamic and contextual information in hmm modeling for handwritten word recognition", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 33, No. 10, (2011), 2066-2080.

16. Van der Zant, T., Schomaker, L. and Haak, K., "Handwritten-word spotting using biologically inspired features", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 30, No. 11, (2008), 1945-1957.

17. Graves, A., Liwicki, M., Fernández, S., Bertolami, R., Bunke, H. and Schmidhuber, J., "A novel connectionist system for unconstrained handwriting recognition", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 31, No. 5, (2009), 855-868.

18. Jiang, Y., Ding, X., Fu, Q. and Ren, Z., *Context driven chinese string segmentation and recognition*, in *Structural, syntactic, and statistical pattern recognition*., Springer (2006), 127-135.

19. Fink, G. and Plotz, T., "On the use of context-dependent modeling units for hmm-based offline handwriting recognition", in Document Analysis and Recognition, ICDAR. Ninth International Conference on,. Vol. 2, (2007), 729-733.

20. Ghiasi, G. and Safabakhsh, R., "Offline text-independent writer identification using codebook and efficient code extraction methods", ***Image and Vision Computing***, Vol. 31, No. 5, (2013), 379-391.

21. Jain, R. and Doermann, D., "Offline writer identification using k-adjacent segments", in Document Analysis and Recognition (ICDAR), International Conference on, IEEE. (2011), 769-773.

22. Djeddi, C., Siddiqi, I., Souici-Meslati, L. and Ennaji, A., "Text-independent writer recognition using multi-script handwritten texts", ***Pattern Recognition Letters***, Vol. 34, No. 10, (2013), 1196-1202.

23. Srihari, S.N., Cha, S.-H., Arora, H. and Lee, S., "Individuality of handwriting", ***Journal of Forensic Sciences***, Vol. 47, No. 4, (2002), 856-872.

24. Louloudis, G., Gatos, B., Pratikakis, I. and Halatsis, C., "Text line and word segmentation of handwritten documents", ***Pattern Recognition***, Vol. 42, No. 12, (2009), 3169-3183.

25. Papavassiliou, V., Stafylakis, T., Katsouros, V. and Carayannis, G., "Handwritten document image segmentation into text lines and words", ***Pattern Recognition***, Vol. 43, No. 1, (2010), 369-377.

26. Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., "Speeded-up robust features (SURF)", ***Computer vision and image understanding***, Vol. 110, No. 3, (2008), 346-359.

27. Mumtaz, A., Coviello, E., Lanckriet, G.R. and Chan, A.B., "Clustering dynamic textures with the hierarchical em algorithm for modeling video", ***Pattern Analysis and Machine Intelligence, IEEE Transactions on***, Vol. 35, No. 7, (2013), 1606-1621.

28. Tan, G.X., Viard-Gaudin, C. and Kot, A.C., "Automatic writer identification framework for online handwritten documents using character prototypes", ***Pattern Recognition***, Vol. 42, No. 12, (2009), 3313-3323.

29. Krivulin, N., "Algebraic solution to a constrained rectilinear minimax location problem on the plane", in Multimedia Technology (ICMT), International Conference on, IEEE. (2011), 6216-6220.

30. Karayiannis, N.B. and Randolph-Gips, M.M., "Soft learning vector quantization and clustering algorithms based on non-euclidean norms: Multinorm algorithms", ***Neural Networks, IEEE Transactions on***, Vol. 14, No. 1, (2003), 89-102.

31. Honkela, T., Duch, w., girolami, M. and Samuel Kaski, "Artificial neural networks and machine learning-icann ", in 21th Int. Conf. on Artificial Neural Networks, Espoo, Finland. (2011).

32. Marti, U.-V. and Bunke, H., "The iam-database: An english sentence database for offline handwriting recognition", ***International Journal on Document Analysis and Recognition***, Vol. 5, No. 1, (2002), 39-46.

33. Bulacu, M., Schomaker, L. and Vuurpijl, L., "Writer identification using edge-based directional features", in null, IEEE. (2003), 937-945.

34. Guyon, I., Schomaker, L., Plamondon, R., Liberman, M. and Janet, S., "Unipen project of on-line data exchange and recognizer benchmarks", in Pattern Recognition, Vol. 2-Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on, IEEE. Vol. 2, (1994), 29-33.

35. Su, T., Zhang, T. and Guan, D., "Corpus-based hit-mw database for offline recognition of general-purpose chinese handwritten text", ***International Journal of Document Analysis and Recognition (IJDAR)***, Vol. 10, No. 1, (2007), 27-38.

36. Louloudis, G., Gatos, B. and Stamatopoulos, N., "Icfhr 2012 competition on writer identification challenge 1: Latin/greek documents", in Frontiers in Handwriting Recognition (ICFHR), International Conference on, IEEE. (2012), 829-834.

37. Louloudis, G., Stamatopoulos, N. and Gatos, B., "ICDAR writer identification contest", in Document Analysis and Recognition (ICDAR), International Conference on, IEEE. (2011), 1475-1479.

38. Hassaïne, A., Al-Maadeed, S., Alja, J.M., Jaoua, A. and Bouridane, A., "The ICDAR arabic writer identification contest", in Document Analysis and Recognition (ICDAR), International Conference on, IEEE. (2011), 1470-1474.

39. Fornes, A., Dutta, A., Gordo, A. and Llados, J., "The icdar 2011 music scores competition: Staff removal and writer identification", in Document Analysis and Recognition (ICDAR), International Conference on, IEEE. (2011), 1511-1515.

40. Hassaine, A. and Maadeed, S.A., "ICFHR competition on writer identification challenge 2: Arabic scripts", in Frontiers in Handwriting Recognition (ICFHR), International Conference on, IEEE. (2012), 835-840.

41. Li, X. and Ding, X., Writer identification of chinese handwriting using grid microstructure feature, in Advances in biometrics., Springer. (2009), 1230-1239.

42. Brink, A., Smit, J., Bulacu, M. and Schomaker, L., "Writer identification using directional ink-trace width measurements", ***Pattern Recognition***, Vol. 45, No. 1, (2012), 162-171.

43. Siddiqi, I. and Vincent, N., "Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features", ***Pattern Recognition***, Vol. 43, No. 11, (2010), 3853-3865.

Offline Language-free Writer Identification Based on Speeded-up Robust Features

M.K.Sharma, V.P.Dhaka

Department of Computer Engineering, Jaipur National University, Jagatpura Jaipur Rajasthan, Jaipur, India

PAPER INFO

Paper history:
Received 15 January 2015
Received in revised form 07 March 2015
Accepted 11 June 2015

Keywords:
Speeded-up Robust Features Descriptors
Codebank
Scale Orientation
Word Regions

چکیده

در این مقاله روشی برای شناسایی نویسنده زبان آزاد به صورت آفلاین پیشنهاد می شود که بر اساس الگوهای سرعت بالای قوی (SURF) و گذراندن مراحل آموزش، ثبت نام، و شناسایی است. در تمام مراحل، از جعبه فیلتر ایزوتروپیک برای تقسیم کردن تصویر متن دست نوشته به نواحی کلمه (WRS) استفاده می شود .پس از آن، توصیف کننده های SURF ناحیه کلمه و مقیاس های مربوطه (SOs) استخراج می شود. در مرحله آموزش، کدبانک SUD از طریق دسته بندی SUD های نمونه های آموزشی ساخته می شود. در مرحله ثبت نام، SUD های دست خط ورودی برای تشکیل امضای SUD استفاده شد که با نگاهی به کدبانک SUD بود و SOها برای تولید مقیاس و جهت گیری هیستوگرام (H_{SO}) مورد استفاده قرار گرفتند. در مرحله شناسایی، SUDها و H_{SO} های دست خط ورودی استخراج و با آنهایی که برای شناسایی ثبت نام کرده بودند انطباق داده شدند. نتایج تجربی هشت مجموعه داده های عمومی نشان می دهد که روش پیشنهادی بهتر از الگوریتم های مدرن و پیشرفته است.

.***doi****: 10.5829/idosi.ije.2015.28.07a.04*