www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 6 Issue 4 April 2017, Page No. 21192-21195
Index Copernicus value (2015): 58.10 DOI: 10.18535/ijecs/v6i4.68

Signature Verification for Banking Sector Based on SIFT

Mary Jolve. J[1],Jisna Joy[2]
[1]Assistant Professor, Department of computer Application,
St. Thomas 'College (Autonomous), Thrissur -680001, Kerala, India
jolvejoyanthikadan@gmail.com
Department of computer Application,
St. Thomas 'College (Autonomous), Thrissur -680001, Kerala, India
jisnajoy111@gmail.com

Abstract: ***A person can be uniquely identified by evaluating one or more distinguishing biological traits. One such trait is scanned images of hand writing. It is useful biometric modality with application in forensic analysis, banking sector, document verification etc. This study focus on text-independent writer identification method based on scale invariant feature transform (SIFT). It consists of three phase: training, enrolment and identification phase. In all of these three phases an isotropic LoG filter is applied first. It is used to segment the handwriting image into word regions (WRs) and then, the SIFT descriptors (SDs) of word region and the corresponding scales and orientations (SOs) are extracted. In the first phase, an SD codebook is constructed by gathering the SDs training samples. In the second phase, SD signatures (SDS) are generated by looking up the SD codebook and SOs is used to generate a scale and orientation histogram (SOH). And finally in last phase, the SDS and SOH of the input handwriting are extracted and matched with the enrolled ones for identification. Here instead of analysing handwritten document we experiment on signature of an individual. Using SIFT algorithm we demonstrate that he/she is genuine or not.***

Keywords: SIFT, segmentation, SIFT descriptor signature, scale and orientation histogram, signature verification.

1. Introduction

The offline hand writer identification is to determine the writer of a text among a number of known writers using their handwriting images. It is very important for document authorization forensic analysis etc. The existing approach of offline text-independent writer identification can be categorized into two: texture based approaches and structure based approaches

For identifying the writer, texture based approach uses handwriting as a special texture image and extracts the textural feature. Structural based approach is more significant, secure and intuitionist. The existing approaches of the structural features are based on contours or the allograph fragments of handwriting. While extracting the feature from the allograph, they fail to derive the structural features between the allograph in the same words and also it may easily affect the aspect ratio of the character in handwriting. When writing a document, the structures of the whole word are stable and have a strong distinguishable for different writers. This approach is very important for characterizing writers' individuality. To handle these problem, the proposed system deals with new algorithm known as scale invariant feature transform (SIFT). This method extracts the key points based structural features at word level from handwriting images. It contains the structural information of whole words and it is inconsiderate to the aspect ratio and slant of the characters.

Signature verification is a behavioural biometric method. The basic goal of the handwritten signatures is to provide an accurate method in order to verify a person's identity. The signature is based on the way in which he/she signs his/her name. Hence for this reason, the handwritten signatures are widely accepted, socially and legally throughout the world.

The signature verification is required to verify the identity of the authorised signatory who attested the document like POR, telephone bill, bank documents etc. Signature is based on many features like pressure, time, distances, pen ups, pen downs, velocity etc.

Signatures can be operated in two different ways: offline mode and online mode. In off-line mode, users write their signature on paper, digitize it through an optical scanner or a camera, and the biometric system recognizes the signature analysing its shape. In on-line mode, users write their signature in a digitizing tablet, which acquires the signature in real time. . Some systems also operate on smart-phones or tablets with a capacitive screen, where users can sign using a finger or an appropriate pen. These information are stored in the computer and used for necessary action.

Accurate signature verification is very difficult since forgery and fraud can cost organizations money, time and their reputation. Forgeries including random (signature does not match the name of the authorized signee); blind (correct name but incorrect writing style) and even skilled forgeries (signature closely resembles the signature on file). Proposed system helps protect these forgeries.

2. Methodology

2.1 The structure of the proposed method

The proposed method consists of three stages: training, enrolment and identification, as shown in Fig. 1. In all of these three stages, the handwriting image is firstly segmented into word regions (WRs). Then SIFT is used to detect the key points and derive their SIFT descriptors (SDs), and the corresponding scales and orientations (SOs) from the word regions. The SDs and SOs will be used in different stages.

In the training stage, a codebook is generated for enrolment and identification by extracting SDs from training dataset.

In the enrolment stage, SD signature (SDS) and SO histogram (SOH) are stored for identification. They are extracted from SDs and SOs Of word regions of enrolling handwriting image.

Finally in the identification stage, SDS and SOH are extracted from the input handwriting image and they are matched with enrolled one to get two matching distance respectively.

There are four main parts in the structure of proposed system. They are word segmentation, codebook generation, feature extraction and feature matching and fusion.

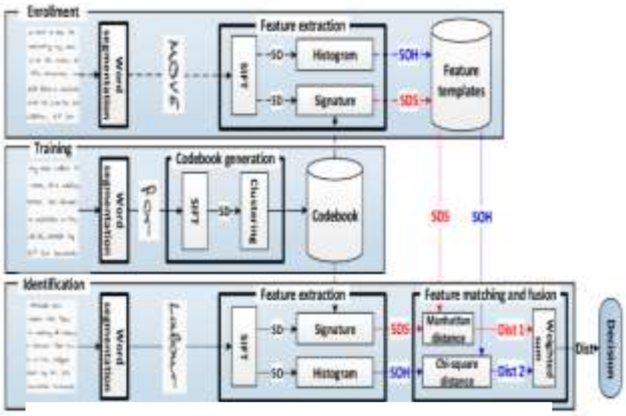

Fig. 1 The structure of the proposed method

2.2 Word Segmentation

Word segmentation is very important for handwriting image analysis. To detect the word-level structural features of handwriting image, we should segment the handwriting image into word regions (WRs). Earlier the segmentation is done manually, which is very difficult and time consuming. Even though many automatic word segmentation techniques have been derived, they fail to segment some skew handwriting images. To avoid these problem and reduce the effect of the direction of text lines, an isotropic LoG filter is employed.

In order to extract word level structural feature, we should segment the handwriting image into word regions (WRs). The word segmentation process is simply described as below:

1. Given a handwriting image *I*. Convert I to binary image I_b by using Otsu's algorithm.

2. Collect all connected – components (CCs) in Ib and then compute their average height ha.

3. Filter Ib with an isotropic LoG filter to get the filtered image If. Compute the variance σ of the filter as σ = 2.5 × ha.

4. Binarizing If to get a binary image Ifb by using the threshold obtained by Otsu's algorithm.

5. Assign each connected-component in Ib to the nearest connected region of Ifb to form semi-word regions (SWRs).

6. Based on the distances between the adjacent SWRs, merge the SWRs to get the word regions (WRs).

7. Split the overlapping connected-components (OCCs) in multiple text lines from the middle line of these OCCs' boundary box.

After word segmentation, the handwriting image is divided into many WRs, which will be used for feature extraction.

2.3 Scale Invariant Feature Transform

SIFT is presented by D. Lowe for distinctive scale- invariant feature extraction from images. This has been widely and successfully applied in many fields. The SIFT algorithm has four main parts: (1) scale-space construction (2)key point localization (3)orientation assignment (4) key point descriptor extraction.

Scale-Space Construction: In this stage the original images are decomposed into a Gaussian pyramid, and each level of the pyramid is called an octave, which is further decomposed into several sub-levels by convolving the initial image at the corresponding pyramid level with LoG filters with different variances.

Key Point Localization and Orientation Assignment: In both these stages many secured key points are detected and the locations of the points, scales and orientations of these key points are calculated.

Key Point Descriptor Extraction: In this final stage, a SIFT descriptor for each key point is generated.We use SIFT, in order to find out the key points of handwriting images, their SIFT descriptors (SDs), and the corresponding scales and orientations (SOs). The SDs are scale and rotation invariant. They can reflect the structures of the image regions centered at the key points. The SOs can prevent the scale and orientation information of these structures. SD and SO are very important information of handwriting to distinguish different writers. SIFT information will be used to extract features of handwriting for writer identification.

2.4 Codebook Generation

For a given handwriting image, word regions (WRs) are generated after word segmentation and for each word region we detect key points and extract their descriptors, scales, and orientations using SIFT algorithm. The result generates a varying amount of key points from different handwriting images. It is very difficult to store all these key points, SDs and SOs as features for writer identification. So in order to make the no. of the feature limited and fixed, we group the SDs of the key points extracted from the input samples into N categories. Each category with its center represents a code. All of the N codes form a SD codebook with size N and then we will compute a histogram with limited and fixed dimension for writer identification based on the codebook.

2.5 Feature Extraction

The text in the identifying handwriting image may be totally different with the text in the enrolled handwriting image. So the key points are totally different in the different handwriting images, even if they are written by the same person. Here we concentrate on extracting two features: one is SIFT Descriptor Signature (SDS) Extraction and the other is Scale and Orientation Histogram (SOH) Extraction.

SIFT Descriptor Signature (SDS) Extraction:

In order to obtain SDs, first we compute the Euclidean distance as follows:

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

Let SD ={d1, d2, . . . , dn} denote n SIFT descriptors (SDs),which are extracted from an offline handwriting image I, and let C ={c1, c2, . . . , cN} denote a SD codebook with size N. Sort the components of EDV in ascending order and obtain the top t components' index in EDV. For each idx∈IDX, update the SDS feature vector as follows:

$$SDSi_{dx} = SDS_{idx} + \delta(EDV_{idx}) \quad (2)$$

Where $\delta(x)$ is a non- increasing function. Repeat these step until all SDs are processed and compute the final SDS vector as follows:

$$SDS_i = \frac{SDS_i}{\sum_{j=1}^{N} SDS_j} \quad (3)$$

Scale and Orientation Histogram (SOH) Extraction:

In the second feature the images are decomposed into X octaves and Y sub-levels in each octave, i.e. Z= (X × Y) scales, by using SIFT. Compute the index idx in SOH feature vector as follows:

$$bin = [oi / \phi]$$

$$idx = Obin \times (s_i - 1) + bin \quad (4)$$

Where O= {o1, o2, o3… on) denotes the corresponding orientations of the SIFT key points, S = {s1, s2, s3… s_n} denote n SIFT key points' scales. Given an angle step ϕ, the orientation [0, 360] can be quantized to Obin = [360/ϕ] intervals. Update the SOH feature vector as Follows:

$$SOHidx = SOHidx + 1 \quad (5)$$

Repeat these steps unit all key points' are processed. Finally compute SOH feature vector as follows:

$$SOHi = \frac{SOHi}{\sum_{j=1}^{M} SOHj} \quad (6)$$

2.6 Feature Matching and Fusion

Let $I1$ and $I2$ are two handwriting images, and let u = ($u1$,$u2$, . . . , u_N) and v = ($v1$, $v2$, . . . ,v_N) are their SDSs, and x = ($x1$, $x2$, . . . , x_M) and y = ($y1$,$y2$, . . . ,y_M)are their SOHs.

Manhattan distance is used to measure the dissimilarity between two SDSs u and v because of its simplicity and high efficiency.

$$D_1(u,v) = \sum_{i=1}^{N} | u_i - v_i | \quad (7)$$

Chi-square distance is employed to measure the dissimilarity between SOH x and y, because it is used to improve the importance of the small value components by giving them more weight. It is calculated as follows:

$$D_2(x,y) = \sum_{j=1}^{M} \frac{(x_j - y_j)2}{(x_j + y_j)2} \quad (8)$$

D_1 and D_2 are normalized into interval [0, 1]. These two distances are then combined to form a new distance to measure the dissimilarity between I_1 and I_2 as below:

$$D(I_1, I_2) = w \times D_1(u,v) + (1-w) \times D_2(x,y) \quad (9)$$

Where $0 \le w \le 1$ is a weight. The parameter w is database dependent and can be determined by cross – validation on the training dataset.

3. Experimental Results and Analyzes

Our experiments are conducted on handwritten signature instead of whole document. Signatures are extracted from bank document when they fill the form to open an account. We usually sign the document at least three times and scan the images through scanner or using digital camera. Here we scan all three different images and store those images in the computer. These images are taken as input to the proposed system. Using SIFT algorithm, first we detect the key points' on the images.

DOI: 10.18535/ijecs/v6i4.68

These key points are saved in the system for further action and then extract their SIFT descriptors (SDs) and the corresponding scales and orientation (SOs).

As described above, proposed system consists of three stages: training, enrolment and identification. The key points', SDs and SOs generated from the images are used in these stages. In the training stages, SDs extracted from scanned image are used to generate a codebook. The Codebook consist of a no. of key points' their descriptors, scales and orientations. Based on this codebook, we will compute a histogram with limited and fixed dimensions for writer identification.

In the next stage, we generate two features known as SD signature (SDS) and SO histogram (SOH) are extracted from SDs and SOs of the enrolling signature image. These values are stored for identifying the individual person. Finally in the identification stage, the SDS and SOH are extracted from the input handwriting images. These images are matched with enrolled ones to get two matching distance. The distance are calculated using Manhattan distance formula and Chi-square distance formula. These equation are described in the above section. From these distance we fuse to form the final matching distance which is a weighted sum for decision.

By applying SIFT algorithm we calculate the dimensions like sectors, pixels, and projection values. Based on these values we decide whether the account holder is genuine or not. Value of projections and pixels may vary on many features like pressure, time, distances, pen ups, pen downs, velocity etc. Fig 2 shows the final result after applying SIFT algorithm.

Fig. 2 Final result after applying SIFT

4. Conclusion

Handwritten signatures are very much dependent on the user's psychology and has great difference in different surroundings and time Dynamic features of signature biometrics can be used in the following applications like finance, insurance, real estate, health, telecom etc. To determine whether the handwriting is genuine, and whether the person signing the check is authorized to use the account this type of algorithm may be used. Many algorithms can be applied to parameters and be calculated in mathematical way. The first step is to obtain the sample signature from the users through the signing pad or tablet. Since signature recognition is a behavioural biometric method, it is dependent on the user's mental, physical presence of mind.

SIFT algorithm extracts the key points, their descriptors, scales, and orientation for a given handwritten signature. From these values we generate a SIFT codebook and bas4ed on this we compute histogram with limited and fixed dimensions for signature verification. This algorithm generates two SIFT features, i.e.SDS and SOH, are extracted from handwriting images to characterize the writer's individuality. Using the two features we calcute two distance, and fuse to form the final matching distance for decision.

5. Acknowledgement

I express my deep sense of gratitude to the University Grants Commission, WRO, and Bangalore for financial assistance to this minor research project.

6. References

[1]. Xiangqian Wu, Youbao Tang, and Wei Bu "Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform", IEEE Transactions on Information Forensics and Security, Vol 9, NO.3, MARCH 2014.
[2]. Thasneem .P, Febina .P, "Offline Hand Writer Identification Based on Scale Invariant Feature Transform", Vol 4, no 5, pp 3194-3197 May 2015.
[3]. Syed Faraz Ali Zaidi Shahzaan Mohammed, "Biometric Handwritten Signature Recognition ", Information Security Course Linköpings university, Sweden.
[4]. D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int. J. Comput. Vis., Vol. 60, no.2, pp. 91-110, Nov. 2004.
[5]. A. Gordo, A. Fornés, and E. Valveny, "Writer Identification in Handwritten Musical Scores With Bags of Notes," Pattern Recognition, vol. 46, no. 5, pp. 1337– 1345, May 2013.
[6]. Y. Zhu, T. Tan, and Y. Wang, "Biometric Personal Identification Based on Handwriting", in Proc. Int. Conf. Pattern Recognition, Barcelona, Spain, pp. 797-800, 2000.
[7]. M. Bulacu and L. Schomaker, "Text-Independent Writer Identification and Verification Using Textural And Allographic Features," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 701–717, Apr. 2007.
[8]. E. Zois and V. Anastassopoulos, "Morphological Waveform Coding For Writer Identification," Pattern Recognition, Vol. 33, no. 3, pp. 385-398, March 2004
[9]. V. Pervouchine and G. Leedham, "Extraction And Analysis of Forensic Document Examiner Features Used For Writer Identification," Pattern Recognition, vol. 40, no. 3, pp. 1004–1013, Mar. 2007.
[10]. https://www.parascript.com/signature-verification