Image Processing Based Signature Verification Technique to Reduce Fraud in Financial Institutions

Walid Hussein, Mostafa A. Salama and Osman Ibrahim

Faculty of Informatics and Computer Science, The British University in Egypt, Cairo, Egypt

Abstract. Handwritten signature is broadly utilized as personal verification in financial institutions ensures the necessity for a robust automatic signature verification tool. This tool aims to reduce fraud in all related financial transactions' sectors. This paper proposes an online, robust, and automatic signature verification technique using the recent advances in image processing and machine learning. Once the image of a handwritten signature for a customer is captured, several pre-processing steps are performed on it including filtration and detection of the signature edges. Afterwards, a feature extraction process is applied on the image to extract Speeded up Robust Features (SURF) and Scale-Invariant Feature Transform (SIFT) features. Finally, a verification process is developed and applied to compare the extracted image features with those stored in the database for the specified customer. Results indicate high accuracy, simplicity, and rapidity of the developed technique, which are the main criteria to judge a signature verification tool in banking and other financial institutions.

1 Introduction

A signature is a gained behavioural biometric of a user to declare his/her unique identity on printed documents. The demand of authorization based on signature is increased including credit card validation, security systems, banking system, checks, contracts, etc., as shown in figure 1. It is widely used as proof of identity and a socially accepted authentication method in daily life. The system stakeholders are person, organization or banks that need to verify signatures [1, 2]. The stakeholders are Bank's customers who must write their signature, and bank's employees have to verify if the sample signature is the original signature in database, to complete any transaction required on that account. Another customers are organizations' employees: any organization that still depend on paper works, employees must take supervisor's signature [3, 4].

Automatic signature verification system compete the current visual verification that depends mainly on the experience, mood and working environment of the verifier. Moreover, it is difficult for the eyes of any experts to precisely verify the ratios between lines and angles of a genuine signature to a fraud signature [5]. One reason is that signature is just a special way of handwriting that contains complex geometric patterns and often unreadable plots [6]. A signature forgery is replicating the genuine signature by the forger after careful practice. This type of forgeries is called the skilled forgery which harden the signature verification task. The other two types are random forgery where the forger does not know the shape of the original signature, and the simple forgery where the forger knows the shape of the original signature but does not practice enough to increase the similarity value between the fraud and the genuine signatures.

Fig. 1. An example of a handwritten signature to prove user identity on a banking check. Authentication of this signature is the key process to provide the user with an access to his/her bank account.

Automatic verification systems that authenticates the person's signature can be categorized as two types, an online (dynamic) and an offline (static) system [6, 7]. In online systems, dynamic data can be obtained from an online user display suchlike electronic tablet with an instructed pen and in this case, the input is a sequence of dynamic features about the user writing activity such as

Corresponding author: walid.hussein@bue.edu.eg

the applied pressure, speed of writing, etc. On the other hand, in the offline systems, signatures are written in a paper which is processed as two dimensional image and has been converted to the system with the aid of scanner or camera [8, 9]. The signature verification architecture usually starts by extracting the features of the genuine signatures followed by classification of a set of genuine and skilled test signatures. The offline features can be categorized as Global and local features, the Global features describe the image as whole the image size, while local features are commonly extracted by partitioning the image into a grid and extract the features in each of its parts [10, 11]. Lately, Interest points are picked using SIFT (Scale-Invariant Feature transform) [SIFT] and/or SURF (Speed-up Robust Features) [SIFT] to perform the signature verification task [12, 13]. These models extract the interest points in each image, then extract the features/descriptors for each interest point to verify the matching between signatures. The accuracy of the solution of these methods depends on the number of matches of the genuine signatures to other genuine and forged signatures [14, 15].

The work proposed in this paper uses a different approach in utilizing the SIFT/SURF extractors, where the matching depends on the summing up of the Euclidian distance between the interest points in the two images. The work applied here is based on the database of offline genuine and skilled forged signatures extracted in the work in [16] and in [17]. The results shows 95% classification accuracy which is higher than that of current research.

The rest of this paper is organized as follows: Section 2 presents the previous work of applying machine learning tools to perform the signature verification task. Section 3 describes the proposed approach, while the experimental work and discussion appear in section 4 and finally the conclusion is in section 5.

2 Materials and Methods

The signature image is initially pre-processed to facilitate the job of the feature extractor technique. The pre-processing includes the cropping of the signature area, removal of the noise, banalization of coloured image to grayscale one and finally edge detection or smoothing of the signature lines. Features are categorized into two types, global and grid features. Global features define the entire structure of the signature like the height, height-to-width ratio, and the image projection. Maximum horizontal and vertical projections represent the row and column that contain the maximum number of black pixels respectively. The second type of features is the grid features where a virtual grid is created of 12x8 segments for describing the detailed parts of the image. Pixels density, pixels distribution, and predominant axial slant are examples of the extracted grid features.

The training is applied on a set of genuine signatures using either a simple distance calculation method or a complex machine learning techniques. This approach is considered a unique method that multiple features are extracted which have global features such as statistical features, image gradient came from distribution of pixels in a signature image and descriptors, the classification contains variation between signature of the same user and done a distribution in distance space. For any tested signature the method gains a distribution which is compared with the saved distributions and a similarity between them is obtained. This method does not utilize set of forgery signature in the training. The approach utilizes the geometric center for feature extraction. This center gains both horizontal and vertical of the signature. The classification and description is done by Euclidean classifier which defines vectors between two signatures. This method tested by database of 20 writer, 10 signatures for genuine and 10 for forgeries. The method achieved 11.4% AER. The use of complex machine learning techniques like Hidden Markov model, support vector machine and neural network show better results in the domain of Hand-written signature verification. Lately, SIFT and SURF detectors and extractors show better results in the field of offline signature verification.

2.1 SIFT/SURF algorithms

Scale Invariant Feature Transform (SIFT) is powerful and successful approach to do feature detection. Speeded Up Robust Features (SURF) is based on the same steps and principle but it uses different schema and provides better and faster result. The SURF algorithm is divided into main two steps: firstly, interest points are detected. Secondly, interest point description is performed. Both of these steps depend on a scale space representation.

SIFT algorithm: SIFT Generates for the image, another images of 1/2, 1/4 and 1/8 of the original sizes. Then for each image apply the Gaussian blurring operator on the image of intensity function I(x, y) as shown in equation 1.

$$L(x, y, kn * \sigma) = G(x, y, kn * \sigma) * I(x, y) \quad (1)$$

Equation 1 is applied for different scales for all image sizes (octave), such that $\sigma = kn*\sigma$ values where $k = 1.2$ and $n = 0, 1, 2, 3, 4$. Then apply the lablacian of Gaussian approximation, DOG, for each two successive layer in each octave as shown in equation 2.

$$DoG(x,y) = (L(x,y, kn*\sigma) - L(x,y, kn-1*\sigma) \quad (2)$$

Then extract the potential interest points, by detecting the local extrema for each point in 3x3x3 neighbourhood window. Remove all points smaller than a specific threshold, and accept only points that match with the Harris corners. Finally, for each point of interest, 128 descriptor is extracted. It shows the orientation and magnitude of the 4x4 grids surrounding the interest point in the eight 45-angled regions.

SURF algorithm: Meanwhile, SURF excludes the step of creating different images in its processing. It uses Hessian matrix that expresses the local changes in area of each point in x and y direction as shown in equation 3.

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \quad (3)$$

Where, $L_{xx}(p, \sigma)$ is the convolution of the image with the second derivative of the Gaussian as displayed in equation 4.

$$L_{xx}(p,\sigma) = I(p) * \frac{\partial^2 g(\sigma)}{\partial x^2} \quad (4)$$

A variable sized filter of a corresponding scale σ is used to form a scale space images. Afterwards, based on the sum of Haar wavelet responses, construct a square window centered around the interest point. This is to select dominant orientation of interest points (x,y) weighted Haar-wavelet responses (dx,dy). This allows the selection of only a single π/3 region along each of the selected orientation, out of the 8 regions in SIFT. The 4 values of orientation dx,dy and magnitude |dx|,|dy| of each point (x,y) is considered for the 4x4 grids in the 16x16 window around the interest point, which will form, 4x4x4, 64 descriptors vector for each feature.

3 The proposed signature verification model

The first step in the proposed verification system is the pre-processing of the image. In image pre-processing step, training and testing signature images are passed into this step. The main purpose of this phase is to make signature image ready for extracting features. The pre-processing phase includes grey scale conversion and binarization.

The image is converted into grey scale: Binary image is converted into grey scale to make process easier and feature extraction more accurate.

Binarizing image: In this step, grey scale image is converted to black and white image using a reasonable threshold. This will make signature clearer as it will be in black and the background will be in white. Threshold filtering is considered the simplest way to binarize an image. Threshold filtering does image binarization by using particular threshold value and intestines of image pixel. All pixels in the signature image with intestines equal all higher than specified threshold value will be converted to white otherwise, all pixels with intensities lower than threshold value will be converted to black which will make signature image black and white. The filter accepts grey scale images for processing, and the optimal threshold values differ from image to image and are obtained after deep investigation of the trained image.

The next step is applying the SURF / SIFT feature extraction method as described above. The final step is the verification/matching method. This is done by calculating the Euclidean distance between the descriptors. For every key point in the descriptor in the original image, two closed neighbours are computed and the distance will be taken to match between them and find similarity. The matching of each two signatures is applied based on the number of key points in each one as shown in figure 2.

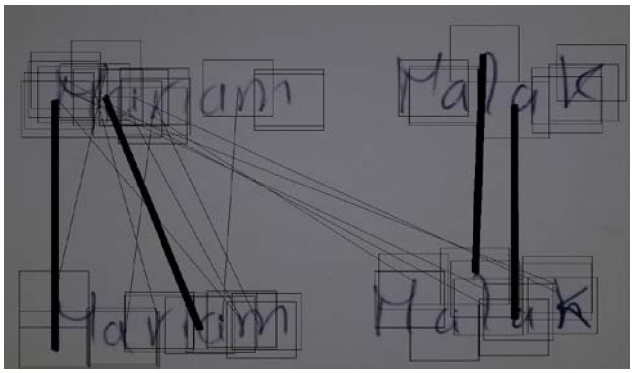

Fig. 2. Matching of features between two genuine signature images

4 Experimental Results

4.1 Accuracy of the proposed model

The proposed model is applied on two benchmark datasets available in [16] and [17]. For the first dataset [16], the results in figure 3 and 4 show that, the first 4 points are genuine-genuine matchings, while the rest of the points are genuine-forged matchings. In this test, 30 forged signatures are tested for verification of signatures of user one and user two. The defined threshold here is applied as the two points out of four which have the lowest Euclidian distance. Accordingly, the results shows nearly 85% classification accuracy of proving the forgery of user one and 92% classification accuracy for user two, as displayed in figure 3 and figure 4; respectively. Note that these results are based on the Euclidian distance values, not based on the number of matchings.

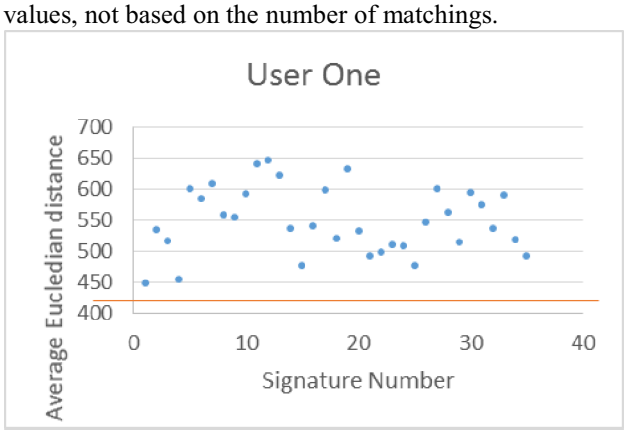

Fig. 3. Genuine-Genuine vs Genuine-Forged signatures for user one.

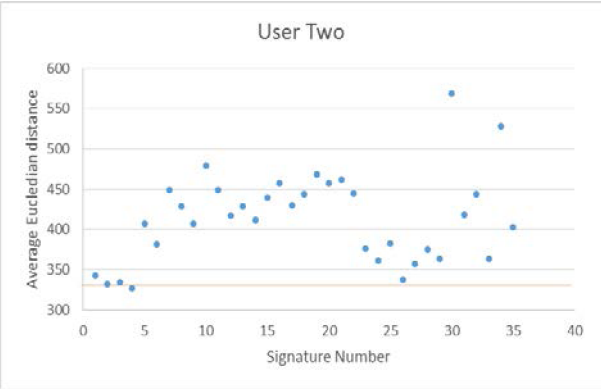

Fig. 4. Genuine-Genuine vs Genuine-Forged signatures for user two.

The same test is applied on the second dataset [17], but is applied by comparing between the first two genuine signatures and between the first genuine signature and another 3 forged signatures, the results are 100% verification accuracy of the forged signatures.

4.2 Parametric investigation of the model

A parametric study is applied in this work by changing threshold value in SURF detector. SURF detector depends on the Hessian matrix to find all interest points in a particular image. SURF divides the signature image into using second order Gaussian kernel and computes these kernels with box filter. The main function in EmguCV library is to detect interest points using SURF with Hessian threshold value. Experimental results show that best value for hessian Threshold could be from 300 to 500.

The SURF detector considers those features in the signature image whose hessian is larger than a specified hessian threshold. Therefore, as high specified threshold value, as less key points in the image will be taken by detector but with more accuracy. A low specified threshold, high key point will be taken.

When the Hessian threshold is 500, key points are 329 and when it is 100, key points are 1004. 1004 key points means many unnecessary key points is extracted from the feature , many calculation will be done which will increase the computational time of the algorithm and feature extracting will be not accurate. Therefore, after testing multiple signatures with different hessian threshold values best value for hessian threshold could be from 300 to 500.

The chart below in figure 5 shows the relation between hessian threshold value with the accuracy and the number of the detected points. Therefore, the relation of hessian threshold value and number of interest point is inversely proportional.as shown in fig.5, higher hessian threshold value will result in less key points and lower hessian threshold will result more key points.

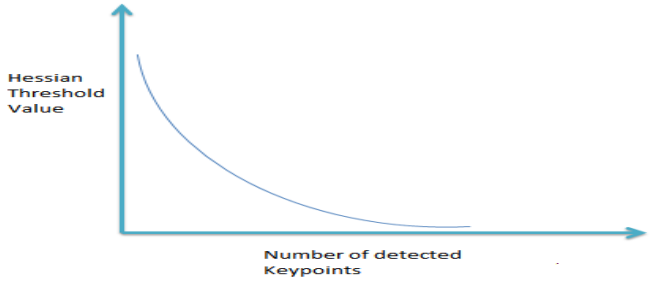

Fig. 5. Relation between hessian threshold value and Number of interest points.

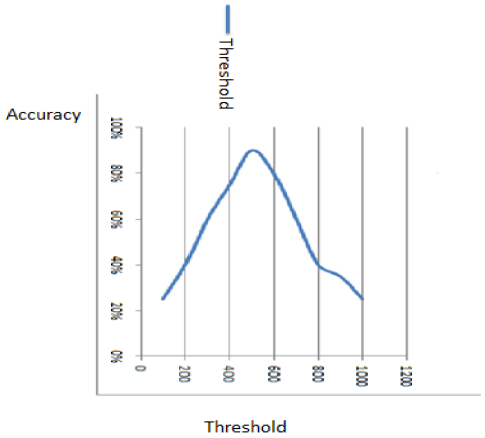

Fig. 6. Relation between hessian threshold value and Accuracy of SURF detector.

As shown in fig.6, after testing multiple signatures with different hessian threshold value shows that the best accuracy of the software is when hessian threshold value between 300 and 500. As shown in the figure 5 above, higher hessian threshold value high accuracy and lower threshold value low accuracy until 500 after hessian threshold; value bigger than 500 , gives less detected points in the signature image which will make accuracy of the software less so the best value of hessian threshold is 500.

5 Conclusion

Although the existence of an automatic signature verification tool is necessary, it is not yet applied in most of the financial institutions. The reason is that most of the currently available tools work with a highest accuracy of ca. 80%, which makes them not reliable in the verification task. For many years, researchers are trying to develop more robust signature verification tools using the advances in image processing algorithms.

The main objective of this work is to offer an economically and efficient offline handwritten signature verification system, in order to achieve the objective multiple methods have been reviewed and surf features algorithm is used in this paper as strong image descriptor. Databases of signatures were collected and saved containing known writer's signatures. The proposed model was successfully verified signatures of users with a

lowest accuracy of 85%, indicating its promising implementation and making a room for more improvements to be researched and investigated.

The future work of the current study is to enhance the feature extraction step of the algorithm by adopting features related to cross correlation, and signature energy and skewness. Finally, an automatic feature extraction tool may be developed to predict the relevant features, which define each signature and reduce the verification effort.

References

1. J. Chambers, W. Yan, A. Garhwal, and M. Kankanhalli (2015) Currency security and forensics: a survey, Multimed Tools Appl vol. 74, pp. 4013-4043.
2. Bhanu Priya Taneja, Navdeep Kaur, (2015) Biometric System Based on Off-Line Signatures, International Journal of Advanced Research in Computer and Communication Engineering, vol. 4(5) , May 2015.
3. Maykin Warasart and Pramote Kuacharoen (2012) Paper-based Document Authentication using Digital Signature and QR Code, 2012 4TH International Conference on Computer Engineering and Technology (ICCET 2012).
4. Vijaypal Singh Dhaka, Mukta Rao, Manu Pratap Singh (2009) Signature Verification on Bank Checks Using Hopfield Neural Network, KARPAGAM Journal of Computer Science, vol 3(4).
5. Ilkhan Cüceloğlu,Hasan Ogul (2014) Detecting handwritten signatures in scanned documents, 19th Computer Vision Winter Workshop, Krtiny, Czech Republic, February 3–5, 2014.
6. Biswajit Halder, Rajkumar Darbar, Utpal Garain, Abhoy Ch. Mondal (2014) Analysis of Fluorescent Paper Pulps for Detecting Counterfeit Indian Paper Money, Information Systems Security, series Lecture Notes in Computer Science, vol. 8880, pp 411-424.
7. Jain, U.A.; Patil, N.N. (2014) A comparative study of various methods for offline signature verification, in Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference, pp.760-764, 7-8 Feb. 2014.
8. Anu Rathi, Divya Rathi, Parmanand Astya (2012) Offline Handwritten Signature Verification By Using Pixel Based Method, International Journal of Engineering Research & Technology, vol. 1(7), September - 2012.
9. Saroj Ramadas, Geethu P.C (2015) Comparative Study On Offline Handwritten Signature Verification Schemes, International Journal of Advanced Research Trends in Engineering and Technology (IJARTET), vol. 2(10), March 2015.
10. Karrar Neamah, Dzulkifli Mohamad, Tanzila Saba, Amjad Rehman (2014) Discriminative Features Mining for Offline Handwritten Signature Verification, 3D Res (2014) 5:2
11. Roy, Sayantan, Maheshkar Sushila (2014) Offline Signature Verification using Grid based and Centroid based Approach, International Journal of Computer Applications, vol. 86(8), pp. 0975–8887, January 2014.
12. Javier Ruiz-del-Solar, Christ Devia, Patricio Loncomilla, Felipe Concha (2012) Offline Signature Verification Using Local Interest Points and Descriptors, Progress in Pattern Recognition, Image Analysis and Applications, of the series Lecture Notes in Computer Science, vol. 5197, pp. 22-29.
13. Malik, M.I.; Liwicki, M.; Dengel, A.; Uchida, S.; Frinken, V. (2014) Automatic Signature Stability Analysis and Verification Using Local Features, in Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference, pp. 621-626, 1-4 Sept. 2014.
14. Rajpal Kaur, Pooja Choudhary (2015) Offline Signature Verification in Punjabi based on SURF Features and Critical Point Matching using HMM, International Journal of Computer Applications , vol. 111(16), pp. 0975–8887, February 2015.
15. J. Ruiz-Shulcloper and W.G. Kropatsch (2008) Signature Verification Using Local Interest Points and Descriptors, CIARP 2008, LNCS 5197, pp. 22–29, 2008.
16. Javier Galbally, Moises Diaz-Cabrera, Miguel A. Ferrer, Marta Gomez-Barrero, Aythami Morales, Julian Fierrez (2015) On-line signature recognition through the combination of real dynamic data and synthetically generated static data, Pattern Recognition, vol. 48(9), September 2015, pp. 2921–2934.
17. Marcus Liwicki, Michael Blumenstein, Elisa van den Heuvel, Charles E.H. Berger, Reinoud D. Stoel, Bryan Found, Xiaohong Chen, Muhammad Imran Malik (2011) SigComp11: Signature Verification Competition for On- and Offline Skilled Forgeries, in proceedings of 11th Int. Conference on Document Analysis and Recognition.