Jurnal
Teknologi

Full Paper

FEATURE SELECTION METHOD FOR OFFLINE SIGNATURE VERIFICATION

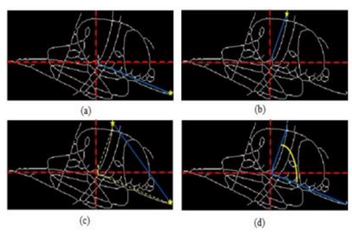
Zuraidasahana Zulkarnain, Mohd Shafry Mohd Rahim*, Nur Zuraifah Syazrah Othman

Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM johor Bahru, Johor, Malaysia

Article history
Received
3 December 2013
Received in revised form
2 July 2014
Accepted
25 November 2014

*Corresponding author
shafry@utm.my

Graphical abstract



Abstract

Signature verification is defined as one of the biometric identification method using a person's signature characteristics. The task of verifying the genuineness of a person signature is a complex problem due to the inconsistencies in the person signatures such as slant, strokes, alignment, etc. Too many features may decrease the False Rejection Rate (FRR) but also increases the False Acceptance Rate (FAR). A low value of FAR and FRR are required to obtain accurate verification result. There is a need to select the best features set of the signatures attributes among them. A combination of the current global features with four new features will be proposed such as horizontal distance, vertical distance, hypotenuse distance and angle. However, the value of FAR may increase if too many features are used which result a slow verification performance. In order to select the best features, the difference between the mean of the standard deviation ratio of each feature will be used. The main objective is to increase the accuracy of verification rate. This can be determined using best features set selected during the features selection process. A selection of signature set with strong feature sets will be used as a control parameter. The parameter is then used to validate the results.

Keywords: Feature extraction, verification, approaches

1.0 INTRODUCTION

Biometrics refers to automated techniques of recognizing an individual based on physiological or behavioral characteristic. To obtain data on a physiological biometric trait, some part of the human body is measured, such as fingerprint, face, retina or palm print. On the other hand, to obtain data on a behavioral biometric trait, a person's resulting action is measured such as his/her signature [1]. Since the biometric identifiers are inherent to an individual, it is difficult to be modified, shared or forgotten. Therefore, a strong and reasonable linkage between a person and his/her identity is formed from these biometric traits.

The subject of interest in this research is signature verification. Signature verification is defined as a biometric identification method using the characteristics of a person's signature. Signature verification is dissimilar from word character identification; a signature is regularly illegible, and it is often a representation with several particular curves [2] that correspond to the writing style of the person. Basically, a signature is just a unique case of handwriting, and regularly is just a symbol [3]. Therefore, it is essential to deal with a signature as a complete image that signifies a particular writing style and not as a compilation of letters and words. And because each individual often have an inimitable handwritten signature, a person's identity can be validated easily by referring to their signatures and writing style [4].

There are basically two types of system in signature verification; off-line system (use static features - the signature image) and on-line system (use dynamic features - time series data). Signatures taken by using pressure-sensitive tablets in order to extract information about that signature such as pressure applied on pen, speed of writing of the signature, etc. is defined as online signature verification. On the other hand, an offline method uses a simpler technique where data of the signature is captured by using an optical scanner [5]. However, offline signature verification is more difficult than online since the number of information presented is limited as dynamic information is not available.

2.0 RELATED WORK

Feature selection method is commonly used to select the best features from the large number of input features set. This consequently increases the performance of the verification system. Several features selection method for signature verification is briefly explained in this paper.

In [6], they proposed a method to find the effectiveness of some frequently used global features in offline signature verification. The sample size of 15 extracted global features of offline signatures is altered and varied using a wrapper method. Hence, the best features for different type of datasets were obtained. Results show that a recognition rate of 94% with 6% of FRR and 0% of FAR value was achieved. However, the wrapper method has some disadvantages as it needs a specific dataset in order to function properly, and it also lacks generality.

The work of [7] aim to obtain a compact set of features from a writer-independent offline signature verification. These features were derived from the surroundedness features extracted. An evaluation using different feature selection methods were performed in order to get the compact set of features. From the results obtained, their proposed features achieved an accuracy of 91.67% with a percentage value of 8.33% for both FAR and FRR by using CEDAR signature database, whereby for GPDS corpus, 86.65% of accuracy with FAR and FRR value of 13.76% each was attained. A plan for new features set and usage of other classifiers could definitely be a future scope for this proposed method.

A new approach for features selection was proposed by [8] for writer-independent offline signature verification system. The proposed approach involved combination of multiple feature extraction, dichotomy transformation and boosting feature selection (BFS). From the combination, a system with numerous users and a limited number of reference signatures was produced. Results show that the proposed approach was a success, where a single reference signature per writer was allowed for verification purposes. For future research purpose, a broader range of feature extraction techniques, resolutions and data sets will be inserted for measuring the performance of this method.

An approach for online feature set selection phase was proposed by [9] by implementing the selection method of deterioration variables based on Mallows Cp criterion. In which case, best feature subsets of different dimensions will be identified for every user. Then, the best subset that have Cp value nearest to p (number of regression coefficient) were selected. Lastly, in order to validate the elimination of the remaining features from the preliminary feature set, general features between best feature subsets will be checked. By using this approach, the usage of feature set for a certain user can be lessen before verifying their signature. The focus of their future work will be the ability to choose larger feature set by using the proposed selection method.

On the other hand, in a research conducted by [10], they proposed a new decision making technique that can be applied on multiple-sets of features (MSF) for automatic signature verification. For this purpose, the similarity between the input signatures and the reference ones was measured by using a distance measure (DM). Later, Euclidean distance was then used for measuring the accuracy of the system. With MSF, overall performance of the system had shown better results where the best feature set that could not be recovered by other method can be successfully captured by this method.

In a different study by [11], they also applied the same method in [10] for decision making technique which is the Multi-Sets of Features (MSF), that provides better forgery detection than best feature set (bfs). Results show that a greater achievement than that acquired by using the bfs was reached. Besides, some lost effectiveness was also recovered by the proposed method.

In conclusion, this features selection method can be categorized as a good method in determining the best features set for signature verification. In this research, four new features are proposed for features enrichment. So, by using the features selection method, we can determine whether the proposed features are appropriate or inappropiate for signature verification.

3.0 DATABASE

For our research, we use the Grupo de Senales (GPDS-960) database. This database consist a total of 23049 genuine and 28800 forgeries signatures obtained from 960 individuals. For each person, 24 genuine signatures and 30 forgeries signatures were stored. To retrieve genuine signatures, each person signed a form of 24 different sizes of boxes in just one session. To obtain the forgeries signatures, 1920 individuals were involved. Each person filled up 15 boxes of a given 5 genuine signatures that randomly chosen, where those signatures needed to be imitated three times each. Since the forgers are not an expert, they may take as much time as needed to imitate the signatures [12]. In this paper, only 200 dataset will be used which consist of 100 genuine signatures and 100 forgeries. These dataset was obtained from 10 different persons, where each person produces 10 genuine signatures and 10 forgeries.

4.0 SIGNATURE VERIFICATION

4.1 Feature Extraction

Features extraction can be defined as the characteristics of signature that are derived from that signature itself. These extracted features play an important role in developing the robust system as all other phases are based on these features.

Global Feature

A global feature is a feature extracted from the whole signature [13]. Based on the style of the signature, different types of global features are extracted. Figure 1 shows global features found in literatures:

i. **Signature area**: amount of pixels which belong to the signature
ii. **Signature height to width ratio (Aspect ratio)**: divide signature height to signature width
iii. **Orientation**: orientation of signature so that the image is positioned in line with the x-axis
iv. **Pure width**: width of the image after removal of horizontal blank spaces
v. **Pure height**: height of the signature after removing the vertical blank spaces
vi. **Maximum horizontal histogram**: the row with the maximum value
vii. **Maximum vertical histogram**: the column with the maximum value
viii. **Image area**: number of white (foreground) pixels in the signature image
ix. **Number of objects**: number of objects counted in the signature image

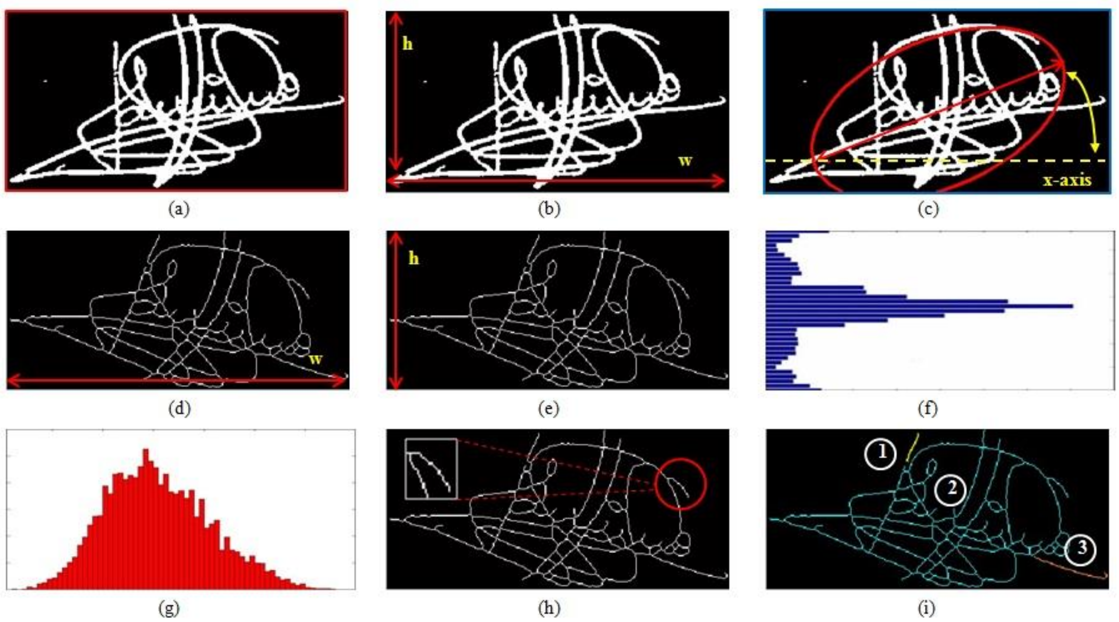


Figure 1 Global features: (a)Signature area (b)Aspect ratio (c)Orientation (d)Pure width (e)Pure height (f)Max. horizontal histogram (g)Max. vertical histogram (h)Image area (i)Number of obects [a]

New Proposed Feature

Four new features are proposed in this paper, as shown in Figure 2; Horizontal distance, Vertical distance, Hipotenuse distance, and Angle. The idea to propose these new features was derived from the center of gravity (COG). From Figure 2, the COG of signature image is shown by the cross point of the two dotted lines. From previous researchers, they stated that every person's signature has a unique COG. Therefore, we conclude that these new features have the potential to be selected as strong features that might increase the accuracy of signature verification.

i. **Horizontal distance**: extracted from MHP and COG. It is the distance between the last horizontal point from gravity center of signature image
ii. **Vertical distance**: extracted from MVP and COG. It is the distance between the first vertical point from gravity center of signature image
iii. **Hipotenuse distance**: the distance between MHP and MVP
iv. **Angle**: the angle between the line created by joining the two centers of gravity and the horizontal axis

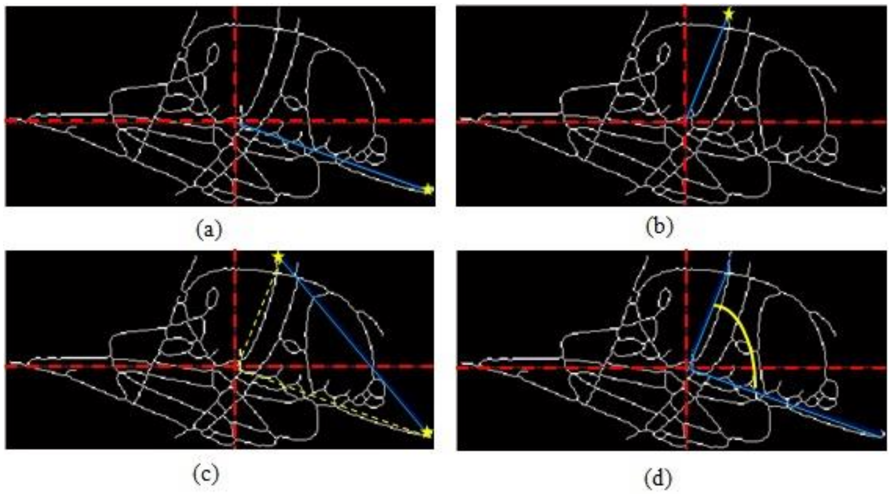


Figure 2 New proposed features: (a)Horizontal distance (b)Vertical distance (c)Hipotenuse distance (d) Angle [a]

4.2 Best Feature Selection

Based on [14], we know that a large number of features may decrease the value of FRR (overall amount of genuine signatures discarded by the system) but at the same time it will increase the value of FAR (number of forged signatures accepted by the system). However, little work has been done in measuring the consistency of these features. This consistency measurement is important to determine the effectiveness of the system. In order to measure the consistencies of these features, there is a need to select the best features set among them. Hence, a difference between mean to standard deviation ratio of each feature from the genuine feature vector and the forgeries features vector set is proposed.

$$D = \sqrt{\left(\left(\frac{Mean_g}{STD_g}\right) - \left(\frac{Mean_f}{STD_f}\right)\right)^2}$$

Where,
D = difference between mean/standard deviation ratio
Mg/STDg = mean/standard deviation ratio of genuine signatures
Mf/STDf = mean/standard deviation ratio of forgery signatures

By applying the proposed best features selection method, the best features sets among the 13 proposed features will be selected and used in verification process later. The selected features will help to improve the performance of signature verification rate. This features selection method which calculated using combinations of mean and standard deviation were used in the past for online signature verification. However, there was no research regarding offline signature verification that implements this kind of technique. Hence, by using the same technique, this research is conducted to select the best features in offline signature verification.

5.0 RESULT

In the equation above, the features with large value of mean/standard-deviation difference as compared to others were taken as best features selection. In order to choose between the best of them, we choose six features that have greater D than the others, as shown in Table 1.

Table 1 Results of best features selection

Features	Meang/ STDg	Meanf/ STDf	D
Signature area	2.4974	2.6912	0.1938
Aspect ratio	2.1798	3.1501	0.9703
Orientation	1.3333	1.4167	0.0434
Pure width	4.2602	4.2683	0.0081
Pure height	3.6865	4.5067	0.8203
Max. Horizontal histogram	1.6680	1.0033	0.6641
Max. Vertical histogram	1.7281	1.7900	0.0619
Image area	2.6029	2.8079	0.2051
Number of objects	1.5262	1.0847	0.4415
Horizontal distance	3.1625	4.1513	0.9889
Vertical distance	4.5857	3.8906	0.6951
Hipotenuse distance	3.0274	4.0998	1.0724
Angle	3.8240	4.1855	0.3615

During verification process, the test signature is compared to all the reference set signatures, resulting in a range of dissimilarity values. If the dissimilarity value is below a certain threshold value, the signature is detected as genuine, otherwise forgery.

A comparative study between existing techniques and proposed method is shown through Table 2, which shows combination of features, with variation of methods used and the accuracy of verification techniques.

Table 2 Comparison between existing techniques and proposed method

Combination of Features	Methods	Accuracy
Aspect ratio, Pure height, Max. Horizontal histogram, Horizontal distance, Vertical distance, Hipotenuse distance	Proposed Method	87.5%
Normalized area of signature, Aspect ratio, Maximum histograms, Centroid, Trisurface, Sixfold surface, Number of edge points, Transition	Euclidean distance	84.1%
Maximum histogram and vertical histogram, Center of mass, Normalized area of signature, Aspect ratio, Tri surface feature, Six fold surface feature, Transition feature	Neural Network (NN)	82.66%
Depth, Vertical splitting, Horizontal splitting	Euclidean distance	79.2%
Geometric center, Vertical splitting, Horizontal splitting	Neural Network (NN)	71.3%
Maximum horizontal and vertical histogram, Center of mass, Normalized area, Aspect ratio, Three surface features, Six fold surface features, Transition feature	Neural Network (NN)	65.3%
Outer and inner contour, Slope of different strokes, Angle between two consecutive strokes	Mathematical Morphology	58.0%
Height, Width, Diagonal distance, Aspect ratio, Center of gravity, Area of black pixel, Middle zone, Energy features	Correlation technique	56.66%

6.0 DISCUSSION

Based on the results, it shows that six features; **aspect ratio, pure height, max. horizontal histogram, horizontal distance, vertical distance, hipotenuse distance** have greater D than the other seven features. For the achievement, three out of four of the proposed new features were included as best features selection which are **horizontal distance, vertical distance** and also **hipotenuse distance**.

This proposed method was compared with the other seven previous researches that do not apply the best features selection method. The comparison was done in order to show the accuracy of verification result by using features selection method and non-feature selection method. By referring to the comparison made, results show that the proposed method gives 87.5% of accuracy by the selected features considered in verification process which also make it as the highest accuracy achievement among the other methods.

7.0 CONCLUSION

This paper presented a method for selecting the best features for offline signature verification by using difference between mean to standard deviation. Altogether 13 features including four new proposed features have been tested and as a result, six of them were selected as best features together with three new proposed features. As conclusion, the method of selecting the best features among a huge features will help to improve the performance of signature verification. For future work, a broader range of data sets and feature extraction techniques will be added for measuring the performance of this method.

Acknowledgement

This research was funded by the Ministry of Education through Fundamental Research Grant Scheme (FRGS) and manage by Universiti Teknologi Malaysia under Vot No. R.J130000.7828.4F627.

References

[1] Impedovo, D., G. Pirlo and R. Plamondom. 2012. Handwritten Signature Verification: New Advancements and Open Issues. *2012 International Conference on Frontiers in Handwritting Recognition (ICFHR)*. Bali. 18-20 Sept. 2012. 367-372.

[2] Saba, T., G. Sulong, S. Rahim and A. Rehman. 2010. On the Segmentation of Multiple Touched Cursive Characters: A Heuristic Approach. In V.V. Das and R. Vijaykumar (eds.). *Information and Communication Technologies*.

[3] Tiwari, D. and B. Sharma. 2012. Development of Intelligent Network for Offline Signature Verification using Pixel Density, Directional Method and Both Method Together. *International Journal of Computer Trends and Technology.* 31(3): 403-411.

[4] Arya, M. S. and V. S. Inamdar. 2010. A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches. *International Journal of Computer Applications.* 1(9): 55-60.

[5] Madhavi, M., M.R. Yaram and R.V. Krishnaiah. 2012. Effective Implementation Techniques in Offline Signature Verification. *IOSR Journal of Computer Engineering (IOSRJCE)*. 5(4): 25-30.

[6] Saikia, H. and K. C. Sarma. 2013. Selection of Some Global Features and Their Performance in Offline Signature Verification with Support Vector Machine. *International Journal of Advanced Research in Computer Science and Software Engineering.* 3(9): 880-884.

[7] Kumar, R., J. D. Sharma and B. Chanda. 2011. Writer-Independent Off-line Signature Verification using Surroundedness Feature. *Pattern Recognition Letters.* 33(3): 301-308.

[8] Rivard, D., E. Granger and R. Sabourin. 2011. Multi-Feature Extraction and Selection in Writer-Independent Off-line Signature Verification. *International Journal on Document Analysis and Recognition (IJDAR)*. 16(1): 83-103.

[9] Boyadzieva, D. and G. Gluhchev. 2011. Feature Set Selection for On-Line Signatures using Selection of

Regression Variables. *Pattern Recognition and Machine Intelligence*. 6744: 440-445.

[10] Ammar, M. 2010. Using Multi-Sets of Features to Improve the Performance of Automatic Signature Verification Systems. *Damascus University Journal.* 26(2): 7-16.

[11] Ammar, M., T. Watanabe and T. Fukumura. 2010. A New Decision Making Approach for Improving the Performance of Automatic Signature Verification using Multi-Sets of Features. *2010 International Conference on Frontiers in Handwriting Recognition (ICFHR)*. Kolkata. 16-18 Nov. 2010. 323-328.

[12] Vargas, J. F., M. A. Ferrer, C. M. Travieso and J. B. Alonso. 2007. Off-line Handwritten Signature GPDS-960 Corpus. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007).* Parana. 23-26 Sept. 2007. 764-768.

[13] Zulkarnain Z., M. S. M. Rahim and R. Kumoi. 2013. Review Paper on Offline Signature Verification. *International Journal on Interactive Digital Media (IJIDM 2013)*. 1(3): 54-57.

[14] Tariq S., S. Sarwar and W. Hussain. 2011. Classification of Features into Strong and Weak Features for An Intelligent Online Signature Verification System. *Proceedings of the 1st International Workshop on Automated Forensic Handwriting Analysis (AFHA)*. Beijing, China. 17-18 Sept. 2011. 11-16.