SIGNATURE IDENTIFICATION AND VERIFICATION

[1]Somashekhar B M, [2]Medha Avadhani, [3]Sushma T S,[4] Rachitha H R, [5]Umamaheshwari M
[1,2,3,4,5]*Maharaja Institute of Technology Mysore*
(E-mail: somumtech@gmail.com,avadhanimedhamysore@gmail.com, sushmasrinivasan97@gmail.com, rachithahunsur97@gmail.com, umamallesh97@gmail.com)

Abstract— Several researches have recently appeared in the literature which propose pseudo-dynamic features for automatic static handwritten signature verification based on the use of gray level values from signature stroke pixels. The results have been obtained using local binary patterns (LBP) and statistical measures from various algorithm like Histogram of Templates etc. In these studies, the results observed was not accurate, so the proposed method was developed. In this paper, we have used GLCM method to measure gray level features robustness when it is distorted by a complex background and also to propose more stable features. The corpuses contain signatures written on a uniform white “non distorting” background or complex background such as a check or an invoice. The identification and verification step is performed by SVM. Experiments are conducted on a standard dataset which contains off-line signatures of 50 persons. The results obtained are very promising.

Keywords— *local binary patterns, gray level co-occurrences matrix (GLCM), offline signature verification, supporter vector machine (SVM).*

INTRODUCTION

Handwritten signatures occupy a very special place in the wide set of biometric traits. The main reason is tradition: handwritten signatures have long been established as the most widespread means of personal verification. Signatures are generally accepted by governments and financial institutions as a legal means of verifying identity. Moreover, verification by signature analysis requires no invasive measurements and people are used to this event in their day to day activities. The handwritten signature is recognized as one of the most widely accepted personal attributes for identity verification. This signature is a symbol of consent and authorization, especially in the credit card and bank checks environment, and has been an attractive target for fraud for a long time. A comparison of signature verification with other recognition technologies (fingerprint, face, voice, retina, and iris scanning) reveals that signature verification has several advantages as an identity verification mechanism. Firstly, signature analysis can only be applied when the person is or was conscious and willing to write in the usual manner, although it is possible that individuals may be forced to submit the handwriting sample.

Two methods of signature verification stand out. One is an offline method that uses an optical scanner to obtain handwriting data from a signature written on paper. The other, which is generally more successful, is an online method, which, with a special device, measures the sequential data, such as handwriting speed and pen pressure. Although less successful than the online method, offline systems do have a significant advantage because they do not require access to special processing systems when the signatures are produced. There are two main approaches for offline signature verification static approaches and pseudo-dynamic approaches.

SIGNATURE VERIFICATION

A. *Offline Signature verification based on static approach*

In off-line systems, a signature is digitized using a flat bed scanner and then stored as an image. These images are called statistical or off-line signatures. The static one involves geometric measures of the signature. Off-line data is a 2-D image of the signature. Off-line signature verification is considered as a behavioral characteristic based biometric trait in the field of security and the prevention of fraud. Off-line systems are of interest especially in scenarios where only hard copies of a signature are available, e.g., where a large number of documents need to be authenticated. Research in the field of off-line signature verification is relatively unexplored; this apathy can be attributed to the inherent limitation of available features from a statistical image of signatures.

B. *Offline Signature Verification based on pseudo-dynamic feature*

Dynamic information cannot be derived directly from static signature images. Instead, some features can be derived that partly represent dynamic information. These special characteristics are referred to as pseudo-dynamic information. The term ‘‘pseudo-dynamic’’ is used to distinguish real dynamic data, recorded during the writing process, from information, which can be reconstructed from the static image. There are different approaches to the reconstruction of dynamic information from static handwriting records. Techniques from the field of forensic document examination are mainly based on the microscopic inspection of the writing trace and assumptions about the underlying writing process. The pseudo-dynamic feature is the influence of physical and bio-mechanical processes on the ink trace and aims at providing a solid foundation for enhanced signature analysis procedures. Simulated human handwriting movements are considered by means of a writing robot to study the relationship between writing process characteristics and ink deposit on paper.

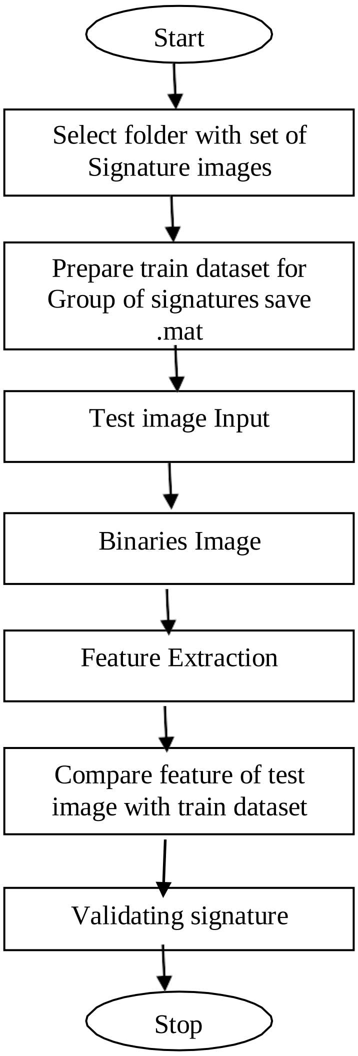

Fig 1. A flow chart of the proposed method

PROPOSED METHODOLOGY

A flowchart of the proposed method is shown in Fig1. The methods contain four steps: inputting signature, preprocessing, feature extraction, validation.

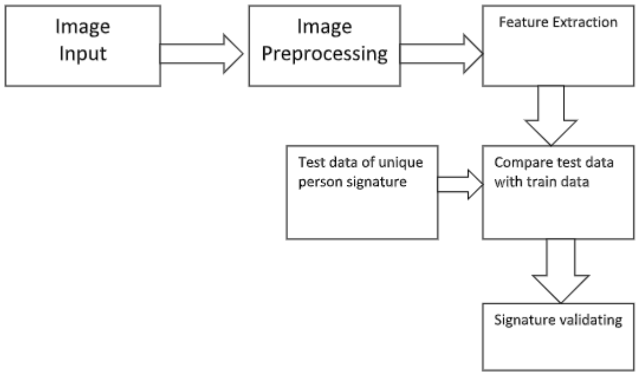

Fig 2: Block Diagram

A. Inputting Signature

The signatures are collected during ten days so as to account for the variations in the signatures with interval of time. The A4 size sheet of paper is given to the signers and each signer was requested to give their signature (30 signature images from each signer). Therefore, three-hundred genuine signatures were collected. Also, two hundred forged signatures were collected in this work. The genuine signatures were shown to a new signer and were requested to repeat that genuine signature ten times after going through a practice session, for collecting the forged signatures. Overall, database of 500 signatures has been created. This database consists of 300 signatures (30 signatures each of the 10 signers) and 200 forged signatures (20 signatures each of the 10 users). This database of signatures has further been divided randomly into training database and testing database.

B. Pre-processing

The background of the scanned signatures in MCYT and GPDS databases is well contrasted with the darker signature strokes so the signature images were binarized by pasteurization. Let be a 256-level, gray scale signature image of the database, the gray level pasteurized image is then defined. The pasteurization procedure is not useful with the databases of the signatures blended with the check images, because the background does not contain uniform character, lines and gray level textures. In this case, we use background removal algorithm i.e., Gray Level Distortion. Then feature extraction process takes by GLCM algorithm. The grey level co-occurrence matrix (GLCM) method is a way of extracting second order statistical texture features from the image. By using this algorithm, we are able to get many second level Statistical features like energy, homogeneity, correlation, contrast, shape etc.

Fig 3: Preprocessing and Background Subtraction

C. Feature Extraction

The classical feature measures extracted from the GLCM matrix and Connors are as follows:

$$Energy = \sum_{i,j=0}^{N-1} \left(P_{ij}\right)^2 \quad \text{--------(1)}$$

This measure of local intensity variation will favor contributions from P(i, j) away from the diagonal ,i.e i!=j.

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}\left(i-j\right)^2 \quad \text{--------(2)}$$

Correlation is a measure of grey level linear dependence between pixels at the specified positions relative to each other the equation for the correlation is as shown bellow

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad \text{--------(3)}$$

A homogeneous scene will contain only a few grey levels, giving a GLCM with only a few but relatively high values of P(i, j). Thus, the sum of squares will be high

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+\left(i-j\right)^2} \quad \text{---------(4)}$$

The mean and standard deviation of P(i, j) rows and the mean and standard deviation of P(i, j) columns, are calculated respectively.

D. Signature Validation

Signature Verification is done by SVM classifier. As there is unique(different) features as well as there is an important variation in the feature elements of each signature, so in order to match a particular signature with the database, the structural parameters of the signatures along with the local variations in the signature characteristics are used. We train the SVM classifier using the 2-dimensional feature vectors to separate the genuine and forgery samples in this normalized feature space as shown in Fig. 2. Then, a linear classification is made by picking a threshold value separating the two classes within the training set. The threshold is fixed and later used in the verification process.

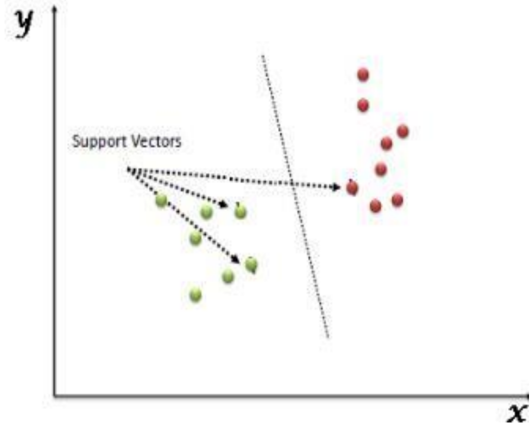

Fig. 4. Training of SVM classifier with respect to the 2- dimensional normalized distance vector.

SOFTWARE DEVELOPMENT PLATFORM

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. The name MATLAB stands for matrix laboratory. MATLAB was originally written to provide easy access to matrix software developed by the LINPACK and EISPACK projects, which together represent the state-of-the-art in software for matrix computation.

A. The MATLAB language

This is a high-level matrix/array language with control flow statements, functions, data structures, input/output, and object-oriented programming features. It allows both "programming in the small" to rapidly create quick and dirty throw-away programs, and "programming in the large" to create complete large and complex application programs.

B. The MATLAB working environment

This is the set of tools and facilities that you work with as the MATLAB user or programmer. It includes facilities for managing the variables in your workspace and importing and exporting data. It also includes tools for developing, managing, debugging, and profiling M-files, MATLAB's applications.

C. ATLAB Image Processing Toolbox

We have used MATLAB Image Processing Toolbox for the development of this software. Image processing involves changing the nature of an image in order to improve pictorial information of the image for human interpretation for autonomous human perception. The Image Processing Toolbox is a collection of functions that extend the capability of the MATLAB numeric computing environment. The toolbox supports a wide range of operations on the image.

SIGNATURE DATABASE

The signature database contains both genuine and skilled forgeries of ten different people. Genuine signature means the original/ authentic signature of the owner/signer whereas a skilled forgery is done by closely imitating the genuine signature of the signer by practicing through various time sessions. For correctly categorizing the signature class, testing is done with both genuine and forged signature. Fifty genuine signatures from ten different signers (given identity codes as) have been collected in this study. These signers are of different age groups and from different fields. The signatures are collected during ten days so as to account for the variations in the signatures with interval of time. The A4 size sheet of paper is given to the signers and each signer was requested to give their signature (10 signature images from each signer). Therefore, five-hundred genuine signatures were collected. Also, two hundred forged signatures were collected in this work. The genuine signatures were shown to a new signer and were requested to repeat that genuine signature ten times after going through a practice session.

Overall, database of 700 signatures has been created. This database consists of 500 signatures (50 signatures each of the 10 signers) and 200 forged signatures (20 signatures each of the 10 users)

Table 1: Training and Testing DB

	Genuine Signatures	Forged Signatures	Total
Training Set	240	160	400
Testing Set	60	40	100
Total	500	200	700

ACKNOWLEDGMENT

Maharaja Institute of Technology Mysore, supports this project work. Appreciate the support of MITM. We would also like to give my sincere gratitude to my parents, lecturers and friends.

REFERENCES

[1] Mohitkumar A. Joshi, Mukesh M. Goswami, Hardik H. Adesara, "Offline Handwritten Signature Verification Using Low Level Stroke Features" 978-1-4799-8792-4/15/$31.00 c 2015 IEEE

[2] Pritam Mondal, Nirmal Kundu "An Automated Handwritten Signature Detection Approach for e-Security Purposes", 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM).

[3] Ms. Pallavi Patil, Mr. Bryan Almeida, Ms. Niketa Chettiar, Mr.Joyal Babu, "Offline Signature Recognition System using Histogram of Oriented Gradients",

[4] Anastasia Beresneva, Anna Epishkina, Darina Shingalova, "Handwritten Signature Attributes for its Verification," 978-1-5386-4340-2/18/$31.00©2018 IEEE

[5] Subhash Chandra, Sushila Maheshkar "Offline signature verification based on geometric feature extraction using artificial neural network," 3rd Int'l Conf. on Recent Advances in Information Technology | RAIT-2016

[6] Mary Jane C. Samonte Roxanne Michelle G. Eullo Alan I. Misa, "Offline Handwritten Signature Verification Using OC-SVM And BC-SVM Classifier", 978-1-5386-0912-5/17/$31.00 ©2017 IEEE

[7] Anjan Dutta, Umapada Pal,Josep Llados," Compact Correlated Features for Writer Independent Signature Verification", 978-1-5090-4847-2/16/$31.00 ©2016 IEEE