Analyses of different methods of writing using SVM classifier

Akhil V M
Department of Mechanical Engineering
The National Institute of Engineering
Mysuru, India
akhil@nie.ac.in

Chandan K J
Department of Mechanical Engineering
The National Institute of Engineering
Mysuru, India
chandankj44@gmail.com

Deepa Itagi
Department of Mechanical Engineering
The National Institute of Engineering
Mysuru, India
deepaitagi07@gmail.com

K R Prakash
Department of Mechanical Engineering
The National Institute of Engineering
Mysuru, India
krp@nie.ac.in

2021 2nd International Conference on Communication, Computing and Industry 4.0 (C2I4) | 978-1-6654-2013-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/C2I454156.2021.9689248

***Abstract*—A person's handwriting reveals his personality and emotion, feeling like fear, happiness, honest or depression. The computer may find it difficult to detect a human motion because it is unaware of human emotion. To predict human emotions, various methods like facial recognition, eye tracking and analysis of handwriting are used. The scientific name of handwriting analysis is graphology; each person writing style is unique when compared to others. Analyzing each person handwriting in the offline method is difficult, costly and takes time. So, this paper is aimed at analyzing the different types of handwriting and extraction of various features. The features extracted will be trained for further classification. In this paper, a machine learning technique support vector machine is being implemented and discussed. The average time taken for writing a sentence was 17 seconds. This proposed system has accuracy varying from 92.33% to 62.5% while predicting the signal.**

Keywords—Handwriting, Graphology, Feature Extracting, Support Vector Machine (SVM).

I. INTRODUCTION

For the means of expanding human memory and also for communication the handwriting was developed long ago. More importantly, the handwritten document was considered an art of communication until the end of the 20th century [1].

Nowadays, there are so many ways to expand human memory and also to establish communication. As the world moved on and so many technologies were introduced like printer type machine where it formatted the handwriting into a printed document. Later in the digital era where we used fax machines, e-mail, and so many other technologies have impacted literacy and handwriting. It is also considered a highly skilled activity. To write a sentence or word requires lower arm, wrist, and finger muscles [2].

In an attempt to identify the human personality, graphology is used. Graphology can analyse the behaviour of the pattern and physical characteristics of individual handwriting. There are many areas where analysis of handwriting has been used as evidence to make sure of writers document and this analysis have helped in the police investigation, banking, e-commerce, lawsuit. But there is still much research to be done and to automate the handwriting analysis [3].

Each person handwriting is different and each person has his/her strokes and pattern. But in recent times handwriting analysis and pattern recognition has been an extensive field of research interest by scientists for the last two decades [4].

Handwriting reveals the true personality and other personal attributes like emotional outlay, fears, honesty, defences. The offline analysis of writing depends on how the analyst is skilled and accurate [5].

Handwriting analysis helps in revealing the behavioural prediction which is done with the aid of the computer. The analyses done by a computer is fast, accuracy and it identifies the handwriting better than offline inspection. Moreover, the analysis of writing done by the computer is free from human error [6].

The speed of handwriting varies from person to person, and it has attracted other investigators, and they think that the handwriting system is a fundamental characteristic for modelling the system[7]. In the existing system i.e., visual/offline analysis, which is difficult to analyse stroke, and the pattern of a person [8].

Handwriting analysis has concentrated on the identification or authentication of individuals, and this is the area of study where little researches are going on. This will help in the field of Forensics, mainly in police investigation [9].

In this research different types of handwriting are being analysed where a person is going to write a character or sentence on paper and it is captured/recorded by a camera with high resolution. Later, the captured video is analysed by software where individual trajectory, speed and acceleration is obtained.

II. METHODOLOGY

This research focuses mainly on the analysis of handwriting which can interpret the individual features like speed, acceleration and trajectory without the intervention of humans.

The flow diagram of this analysis is shown in Figure 1. It has mainly 3 steps Capturing, Time-domain Feature extraction, Classification. For this analysis, 15 subjects

handwriting samples were taken which belongs to both genders for better results.

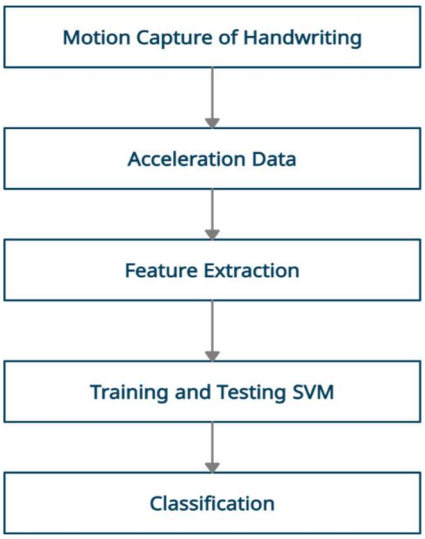


Figure-1: Flow Diagram of the proposed method

The experiment setup consists of a flat table with a paper lock, blue/black pen with a marker point at the top, and a 60 Fps Canon CMOS Camera. Firstly, subjects are advised to sit freely in a chair asked to write something on paper. After five free trials, subjects got comfort in writing and ask to write a sentence (*Artificial Neural Network Analysis*) in A4(21 x 29.7 cm) size paper under four different conditions; simultaneously video is recorded using Canon CMOS Camera. The camera is located in such a way that, focus on the top of the pen so that the path of the pen is being tracked.

In this analysis, the person is going to write a sentence in four conditions as shown in the figure.

The conditions are,

1. Closed Eyes Wrist in Air.
2. Closed Eyes Wrist Down.
3. Opened Eyes Wrist in Air.
4. Opened Eyes Wrist Down

.

Figure 2: The cropped image of Wrist in Air

Once the motion is captured by the camera, the video is uploaded to Kinovea software for analysis. After tracking the path of the pen, speed and acceleration results are obtained. After obtaining the acceleration results and these results are used for further feature extraction like the time-domain feature and also used for SVM classification.

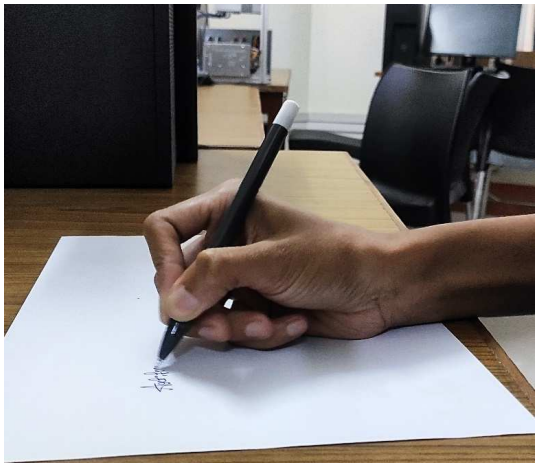


Figure 3: The cropped image of Wrist Down

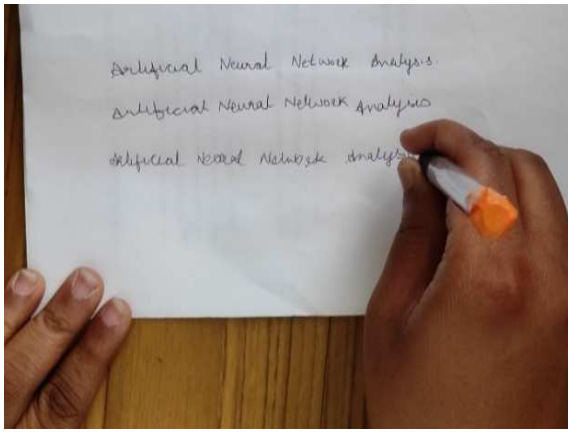


Figure-4: The cropped image of handwriting

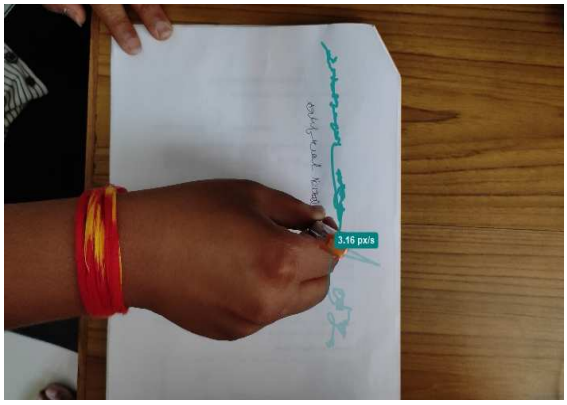


Figure-5: Path trajectory of a handwriting Sample

III. Time Domain Features

To analyse the physical function and mathematical function with respect to time use time-domain analysis. These time-domain features have been mainly used in the engineering field [10]. The below are the five major time-domain features used in this analysis.

Root Mean Square (RMS): The RMS value is the square root of the mean square.

$$RMS=\sqrt{1/n}\ \sum_{i=1}^{n} x_i^2 \quad (1)$$

Logarithmic root mean square (LRMS): It is the ratio between actual values and predicted values and takes the log of the predictions and actual values.

$$LR = \log(\sqrt{1/n}\ \sum_{i=1}^{n} x_i^2) \quad (2)$$

Variance (VAR): It is the average of the squared deviation of a variable. The mathematical definition of VAR is expressed as

$$VAR= \frac{1}{n-1}\ \sum_{i=1}^{n} x_i^2 \quad (3)$$

Mean absolute value (MAV): It is the average absolute value of the acceleration signal.

$$MAV = \frac{1}{n}\ \sum_{i=1}^{n} |x_i| \quad (4)$$

Standard deviation: The standard deviation is a measure of the amount of variation or dispersion of a set of values.

$$SD = \frac{1}{n}\ \sum_{i=1}^{n} (x_i - \bar{x})^2 \quad (5)$$

IV. Support Vector Machine

SVM or Support Vector Machine is a set of supervised machine learning used for classification and regression detection. SVM is used to solve many practical problems like linear and non-linear problems. The reason we are using the SVM classifier is that it is accurate and time-efficient[11].

The unknown data are classified into correct data categories using SVM. Suppose that a training data set are given[12].

$$\{X_i, Y_i\}\ i\ =\ 1,2,3,\ldots,n \quad (6)$$

Where X_i is a vector of input variables and Y_i Will be the corresponding output value. Let φx and φy be their sample spaces. Now, based on the training data, the SVM will develop a function f(x) which can predict the output based on the training data. The feature function f(x) is used to map a pair of (X1, Y1) to a computationally simple form

$$f(x) = \varphi_x * \varphi_y \quad (7)$$

The vector ϕ predicts the points by calculating a class where the feature function values are high, thus restricting f(x).

$$min_{\phi}\|\phi\|^2 + c\sum_n max_y(\Delta(y_n, y) + \phi, (f(x)(x_n, y)) - \phi(f(x)(x_n, y_n)) \quad (8)$$

Where (x1, y1), (x2, y2), (x3, y3), …, (xn, yn) is the positive training data set.

After extracting the data from the time domain feature like RMS, Log RMS, Standard Deviation, MAV and Variance. Once we get the values of each feature now, we train this value for the accuracy and error percentage of the system.

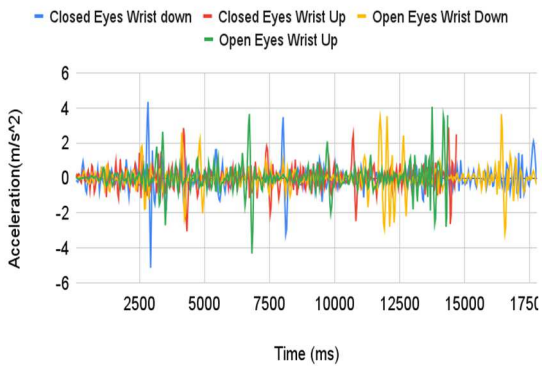
V. Results and Discussion

The main aim of this paper is to analyse the different types of handwriting and to classification using SVM. For these 15 healthy subjects, handwriting samples were taken and analysed each sample under four conditions.

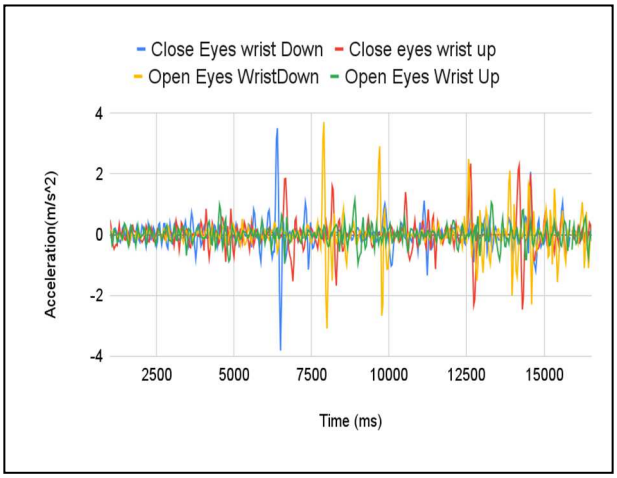
Table 1: Four Condition used in this Analysis

Sl. No	Condition
1	Closed Eyes Wrist in Air
2	Closed Eyes Wrist Down
3	Opened Eyes Wrist in Air
4	Opened Eyes Wrist Down

(a)



(b)



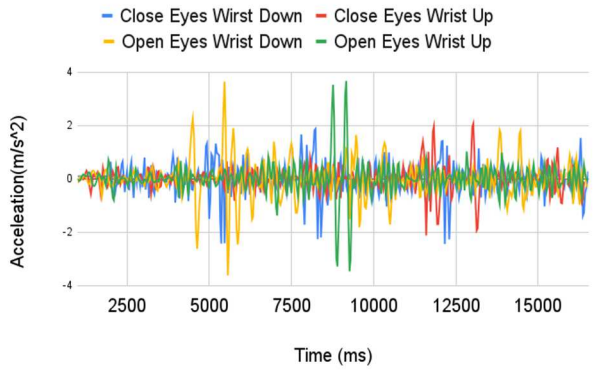


Figure 4: (a)(b)(c) Acceleration v/s Time of 3 different person

Condition	Close eyes Wrist Down	Close Eyes Wrist Up	Open Eyes Wrist Down	Open Eyes Wrist Up
Acceleration	0.00618	0.003226	0.001685	0.0011454
Avg Time (Sec)	14.3	14.9	16.6	19.0

Table 2: Average time and Average acceleration

The above table contains the average acceleration of a person's handwriting and the average time taken to write a sentence. From the above-average acceleration values, Closed Eyes Wrist Down condition has high acceleration with less time and also in Opened Eyes Wrist Up condition have low acceleration with increased time.

After analysing each person's acceleration under each condition, nearly 5200 values are obtained. Using MATLAB Vector Conversion, Standard Deviation, Mean Absolute Value, RMS, Log RMS, Variance values are obtained. The obtained values are converted into 60 data sets, and each data set consists of above-mentioned time-domain feature values.

Table 3: Detailed Description of sample values being taken and obtaining error percentage based on time domain feature

Total Data Set	Values being taken	Error Percentage	Correct	Incorrect	Time
5	5	20	4	1	0.13
10	9	11.11	8	1	0.10
15	13	7.67	12	1	0.09
20	17	17.65	14	3	0.10
25	22	13.64	19	3	0.05
30	25	32	17	8	0.08
35	30	20	24	6	0.08
40	32	15.63	27	5	0.10
45	35	22.86	27	8	0.10
50	31	25.81	23	8	0.10
55	40	37.50	25	15	0.01
60	36	30.56	25	11	0.05

The system will take random values from the data set and it will calculate the accuracy and error percentage. For Example, the system will take 5 random values from 5 random data set and it will give the results. Accuracy of this data set is 80% and error is 20%. So, based on the accuracy percentage the system will predict that this signal belongs to this particular person. From this analysis, the accuracy is varying from 92.33% to 62.5%. The error percentage is high in some conditions, it can reduce by incorporating more number of sets or features.

VI. Future Work

Currently, this experiment has been conducted with the same age group of people of both gender for the handwriting analysis but in the future, we will include all the age groups of people thus making the analysis richer and accurate.

VII. Conclusion

In this analysis, the subjects are advised to write a sentence in A4 size white paper simultaneously the video is recorded. After recording the sentence acceleration, speed features are been extracted. From this acceleration data the time domain features like Standard Deviation, Variance etc are obtained.

After calculating the average acceleration value open eyes wrist up condition has low acceleration i.e., 0.00114 m/s^2 and the time taken for this condition is 19 seconds. Close eyes wrist down condition has high acceleration i.e., 0.00618 m/s^2 and time taken for this condition is 14.3 seconds.

The time-domain features are being trained and we use SVM classification for identification and this system has accuracy varying from 92.33% to 62% while predicting the signal.

References

[1] Kavya N, Madhushree N, Maheshwari S, Manasa H C, Anandhi G, 2019, Handwriting Recognition Using SVM, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCRACES – 2019 (Volume 7, Issue 10),

[2] R. Plamondon and S. N. Srihari, "Online and offline handwriting recognition: a comprehensive survey", IEEE Transactions on pattern analysis and machine intelligence, vol. 22(1), pp. 63-84, 2000.

[3] H.-L. Teulings, J.L. Contreras-Vidal, G.E. Stelmach, C.H. Adler,Parkinsonism reduces coordination of fingers, wrist, andarm in fine motor control, Exp. Neurol. 146 (1) (1997) 159–170. Identifying Human Behavior Characteristics using Handwriting Analysis Devesh Sony, Rakshita Sawant

[4] Vikram Kamath, Nikhil Ramaswamy, P. Navin Karanth, Vijay Desai and S. M. Kulkarni " Development Of An Automated Handwriting Analysis System" VOL. 6, NO. 9, September 2011 ISSN 1819Journal of Engineering and Applied Sciences.

[5] Franke Katrin, L.R.B. Schomaker, L.G. Vuurpijl & S. Giesler (2003) A common ground for computer-based forensic writer identification, Forensic science international, 136 (supp): 84

[6] Character Revealing Handwriting Analysis based on Segmentation method using Support Vector Machine Ankur M. Bobade Prof. Nikkoo. N. Khalsa

[7] Chihi, Ines & Abdelkrim, Afef & Benrejeb, Mohamed. (2012). Analysis of Handwriting Velocity to Identify Handwriting Process from Electromyographic Signals. American Journal of Applied Sciences. 9. 1742-1756. 10.3844/ajassp.2012.1742.1756.

[8] "Identifying Human Behavior Characteristics using Handwriting Analysis" Devesh Sony, Rakshita Sawant, nternational Research Journal of Engineering and Technology (Volume: 06 Issue: 04)

[9] Franke Katrin & Mario Köppen (2001) " A computer-based system to support forensic studies on handwritten documents", International Journal on Document Analysis and Recognition, 3 (4): 218–3

[10] V. M. Akhil, J. Varghese, P. K. Rajendrakumar and K. S. Sivanandan, "Effect of VMD decomposition of soleus muscle EMG in SVM classification," 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), 2019, pp. 1-5, doi: 10.1109/ICACCP.2019.8882900.

[11] Prasad, Shitala & Singh, Vivek Kumar & Akshay, Sapre. (2010). Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine. International Journal of Computer Applications. 8. 10.5120/1256-1758.

[12] Bobade, A. M., & Khalsa, N. N. (2015). Character revealing handwriting analysis based on segmentation method using support vector machine. International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE), 4, 203-207.