Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 84 (2016) 177 – 185
Procedia Computer Science

7th International conference on Intelligent Human Computer Interaction, IHCI 2015

# Perception Based Decision Support System for Handwriting Behaviour Analysis

Asok Bandyopadhyay[a], Bhaswati Mukherjee[b],Abhisek Hazra[b]

[a]*Principal Engineer, Centre for Development of Advanced Computing, Plot-E2/1, Block-GP, Sector-V, Salt Lake City, Kolkata-700 091, India*
[b]*Project Engineer-I, Centre for Development of Advanced Computing, Plot-E2/1, Block-GP,Sector-V, Salt Lake City, Kolkata-700 091, India*

**Abstract**

Handwritten text is potentially the most powerful and conventional means of personal authentication in Human Computer Interaction, with applications to be found in document analysis, deception detection, banking and many other areas. Handwriting is a complex perceptual motor task generating linguistic information. Characters reflect shape distinction needed to perceive different phonetic information of words. In this paper, we have tried to emphasize the role of perception and cognition in identifying unique characteristics of handwriting of any person to screen out deceptive and true statements as a computational model in the areas of Pattern Recognition and Human Computer Interaction. The paper reports the prototype development of a decision support system based on handwriting behavior analysis.


Peer-review under responsibility of the Organizing Committee of IHCI 2015

*Keywords:*Perception Engineering, Pattern Recognition; Online Handwriting; Handwriting Analysis; Deception Detection

## 1. Introduction

Human tops the evolutionary pyramid because of their sophisticated and refined perceptual system. The exact processes lying behind animate perceptual system are yet to be fully explored but scientists do agree that there are certain factors, which form the fundamental part. Handwriting, although apparently simple, is a complex perceptual motor activity essentially involving factors like emotion, motivation, motor movement, attention, cultural background etc.

* Mr. Asok Bandyopadhyay Tel.: +91-033-23579846; fax: +91-33-23575141
*E-mail address:asok.bandyopadhyay@cdac.in*


Peer-review under responsibility of the Organizing Committee of IHCI 2015
doi:10.1016/j.procs.2016.04.084

Human perception is not the same as artificial perception, because the human organisms are concerned with the problems of life, of development, of survival. Therefore Cognitive Scientists claim that there must be a regulatory system that interacts with the perceptual component. And it may well be said that it is the perceptual component that is subservient framework behind all our complex activities including handwriting. Law Enforcement Agencies frequently use handwriting as potential evidence, and the evaluator of that evidence is often described as a handwriting expert. There are people who study and research on handwriting to discover the behavioral traits of the writer. To do this on a scientific basis it is necessary to build up background knowledge by study of what is found in handwriting in many different circumstances. Thus, to identify the handwriting of an individual person it is necessary to know how the writing of one person differs from that of other, and how the writing of one individual varies within itself under different circumstances.

### *1.1. Brief Outline*

The goal of this study is to compare and analyze the handwriting behaviour of true and false or distorted writing. Based on the cognitive load known to be experienced while communicating a deceptive message, we will hypothesized a difference (in temporal and spatial, pressure measures and peak velocities) between the handwriting of true vs. distorted messages written on a digitizer. We would try to evaluate brain-hand performance, as manifested through handwriting behaviour which is a valid measure for detecting the dis-automaticity that is indicative of detecting deception.

The present study will focus on cognitive responses and detection. The cognitive approaches explaining detection assume that encoding a deceptive message requires a greater cognitive effort than telling the truth because of higher processing capacity demands particularly when the lie involves a report about a complex event.

In case of Lie detection, suspected person may be identified by asking him to write a statement about a particular incident in which his involvement is suspected. A typical pattern will emerge through the analysis of perceptual behavioral pattern of the suspect which will reflect the stressed condition of the person and it may be used as an aid to detect whether a person is lying or not.

The computerized system makes it possible to compare handwriting under different conditions; therefore we will compare the handwriting of the same individuals when asked to write truthful and deceptive sentences. Our research hypothesis is that differences will be found between writing of truthful sentences and writing of distorted sentences in pressure, temporal (stroke duration on digitizer and in air) and spatial measures (strokes path length, height and width) obtained by the computerized system. Based on the finding of the studies above we can predict that in deceptive writing, the mean and standard deviations of handwriting measures of each participant will be varied. Thus while writing deceptive sentences, higher pressure will be implemented, longer duration time per stroke (on paper and in air) will be required, and letter strokes will be larger in comparison to truthful writing.

### *1.2. Handwriting Analysis*

The principle of individuality, also known as the principle of uniqueness, forms the basis for handwriting analysis. That is, no two writers share the same combination of handwriting characteristics given sufficient quantity and quality of writing to compare. Albert S. Osborn (1929) [1] detailed in great length the principle of individuality in the second edition of Questioned Documents.

A computer software program has been developed to extract macro-features (slant; word proportion; and measures of pen pressure, writing movement, and stroke formation) from the entire document, from a paragraph in the document, and from a word in the document. It was also used to extract micro-features (gradient, structural, and concavity features) at the character level of the document. Based on only a few macro- and micro-features, Srihari et al. [2] established that the writer of a particular sample can be identified with 98 percent confidence.

### *1.3. Variation*

No two people write exactly the same way, even the writings vary for a person. This is known as natural variation, or intra-writer variation, and represents the second principle of handwriting analysis. Human beings are not capable of machine-like precision and repetition. As a result of the neuromuscular process, some variation in style (formation) is expected. Variation is an integral part of an individual's writing. It describes the changes and deviations, often minute that are found in repeated samples of one person's writing. More specifically, variation refers to the different way(s) that a writer makes each letter or character. This variation is normal and serves as an added factor to personalize and individualize writing.

### *1.4. Writing Skill*

Every writer has a writing skill that cannot be dramatically improved in a short time frame while maintaining all appearances of natural writing. For this reason, the third principle of handwriting analysis is skill level, or the writer's ability to physically reproduce the letter formations they visualize.

## 2. Online Handwriting Analysis

Originally handwriting was a medium of communication primarily restricted to pen and paper but with the technological revolution the digitized means of capturing handwriting has evolved. Online handwriting refers to the methods and techniques dealing with the automatic processing of a handwritten input produced using a digitizer and a stylus [3]. The digitizer and the stylus capture the information of x-y coordinates, velocity, acceleration, pressure of a handwritten content thereby making it more useful.

The block diagram of the system depicts that Handwriting involves a complex perceptual human motor activity and can be used as a deception detection system by analyzing the same.

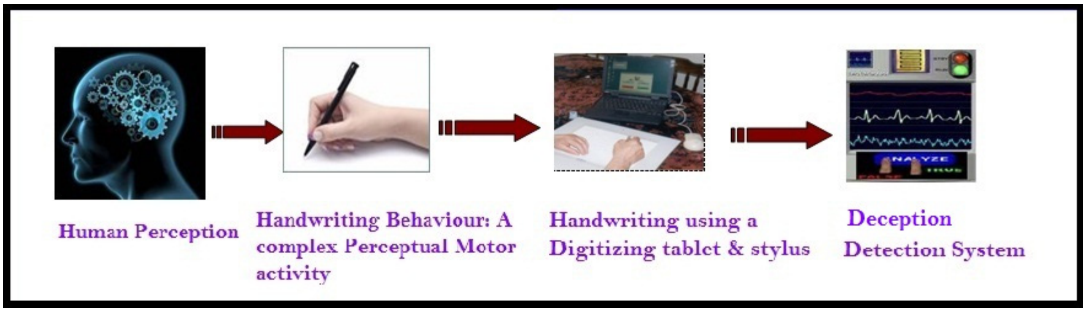


Fig.1. Block Diagram of the Handwriting Behaviour Analysis System

## 3. Data Acquisition

The participants need to write two short paragraphs in sequence describing autobiographical events and memories (say, for example), one about a true event and the other a false description of the same event. The subjects may be requested to write the true and distorted paragraphs in a language on a digitizing tablet. The computerized system enables the collection and analysis of spatial, temporal [4], and pressure handwriting data while the subject is writing on a digitizer (an electronic tablet).All writing tasks are performed on the surface of a WACOM Intuos pro digitizing tablet, using a wireless electronic pen with a pressure-sensitive tip. Displacement, pressure and pen-tip angle are sampled at 200 points per second. The digitizer provides accurate temporal measures throughout the writing, both when the pen is touching the tablet (On-paper time) and when it is raised (In-air time) [5]. It also provides accurate spatial measures when the pen is touching the tablet and/or when it is lifted above the digitizer up to 6 mm. Beyond 6 mm, the spatial measurement is not reliable.

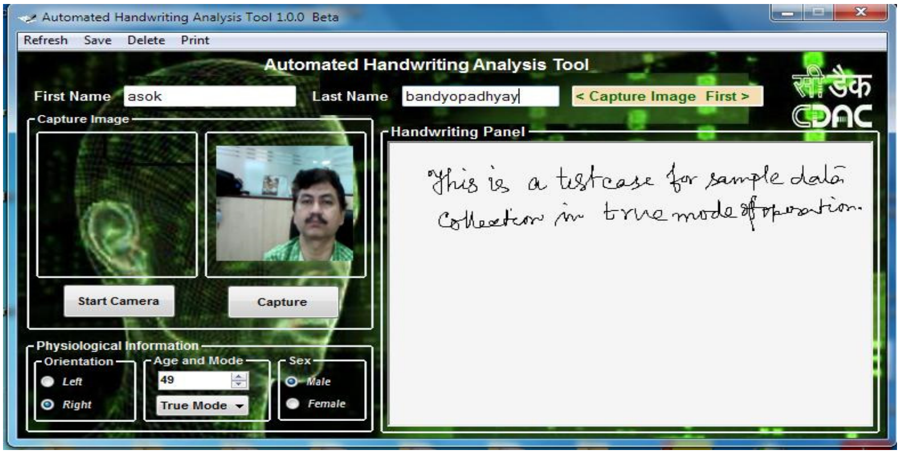


Fig. 2. Front end of the data acquisition software

Handwriting paragraphs of each participant are collected in order to illustrate the differences between the true and distorted paragraphs. The differences are also presented for one specific stroke, which is chosen in both paragraphs in the same location (two examples in which the writer wrote the letter in one stroke and not in several strokes). The analysis software points to the number of the strokes and the designated letter in both paragraphs as a particular stroke.

Perceptual experiments for collection of handwritten data were done at Don Bosco College of Engineering & Technology at Guwahati on 114 Subjects in association with CID, Guwahati. A video with high emotional content along with English subtitles was shown to the subjects. They were asked to express their individual perceptual information in the form of their handwritten contents, firstly representing the true situation as seen in the video and then were asked to write the distorted information of the same in terms of persons, actions and events shown in the video. Here, Person refers to the people present in the video, actions are their activities shown in the video and events represent the overall experience of the happenings as perceived by the subject. In the online mode, our software is defined with the three different categories of Person, Action and Events under true and distorted cases for facilitating students to response in the appropriate category.

## 4. Feature Extraction

Different sets of features were computed from the handwritten raw data obtained from the tablet which are listed below

i. Local Feature- These features are extracted at each sampling point of the handwritten raw data. The features are $\Delta x$, $\Delta y$, $\Delta t$, $\Delta p$, $\Delta p/\Delta y$, $\Delta p/\Delta x$, $\Delta v$ and $\Delta a$. These features represent change in x-y coordinates, change in time, pressure, velocity and acceleration respectively and computed using the following equations(1) to (7)

$$\Delta \mathrm{x} = x(t) - x(t-1) \tag{1}$$
$$\Delta \mathrm{y} = y(t) - \mathrm{y}(\mathrm{t}-1) \tag{2}$$
$$\Delta \mathrm{p} = p(t) - \mathrm{p}(\mathrm{t}-1) \tag{3}$$
$$\Delta \mathrm{v} = \sqrt{[(\Delta \mathrm{x}/\Delta \mathrm{t})^2 + (\Delta \mathrm{y}/\Delta \mathrm{t})^2]} \tag{4}$$
$$\Delta \mathrm{a} = \Delta \mathrm{v}/\Delta \mathrm{t} \tag{5}$$
$$\Delta \mathrm{p}/\Delta \mathrm{y} = \{\mathrm{p}(\mathrm{t}) - \mathrm{p}(\mathrm{t}-1)\}/\{\mathrm{y}(\mathrm{t}) - \mathrm{y}(\mathrm{t}-1)\} \tag{6}$$
$$\Delta \mathrm{p}/\Delta \mathrm{x} = \{\mathrm{p}(\mathrm{t}) - \mathrm{p}(\mathrm{t}-1)\}/\{\mathrm{x}(\mathrm{t}) - \mathrm{x}(\mathrm{t}-1)\} \tag{7}$$

ii. Global Feature -A 49 feature set has been calculated [6]. Each of the features has a single value for a whole handwriting curve. Examples are maximum writing speed, total writing duration, aspect ratio of the handwriting, total pen down time on the tablet surface etc. given in Appendix A.

iii. Spatial &Temporal Feature - A 12 feature set computed based on time and space information provided in Appendix B.

## 5. Analysis & Results

Based on previously described handwritten data collection and features, the coefficient of variance for the stroke duration, path length, height and width is analyzed as a measure of the consistency of handwriting performance. Descriptive statistics of the dependent variables are tabulated and examined. The number of strokes for the true and distorted paragraphs is compared by paired sample t-test.

Following the finding that there are differences between the groups for the number of strokes, a measure of the difference between number of strokes at the truth task and number of strokes at the distorted task is computed.

A total of 49 global features have been taken into account. The MANOVA was performed using IBM-SPSS ver.22.0 on Global and spatio-temporal features extracted which showed significant differences in true and distorted cases against the three categories of Person, Event and Action. The MANOVA performed on the 49 derived features yielded 25 significant features (described in Appendix C)

Twelve spatio-temporal features were extracted, out of which six (described in Appendix D) have shown significant difference. This set of data is fed to the SVM classifier which yielded good classification results as indicated below.

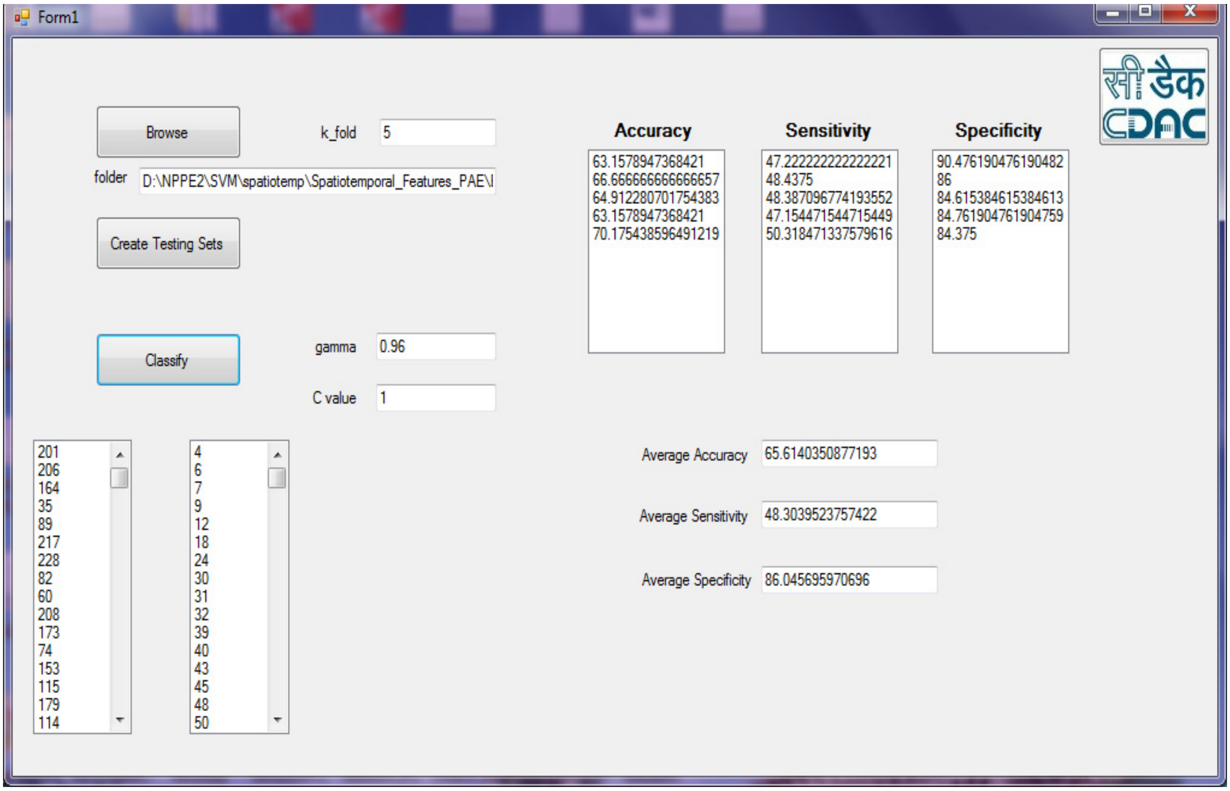


Fig. 3. SVM classification output of spatio-temporal features

The Accuracy, Specificity and Sensitivity percentage for spatio-temporal features are 65.51%, 73.16% and 62.82% respectively. Similarly, for global features, the accuracy, specificity and sensitivity are reported to be 66.42%, 76.20% and 56.94% respectively. Results reflect the fivefold cross validation output of both spatio-temporal and global features having value of c as 1 and γ as 0.96 for SVM classifier.

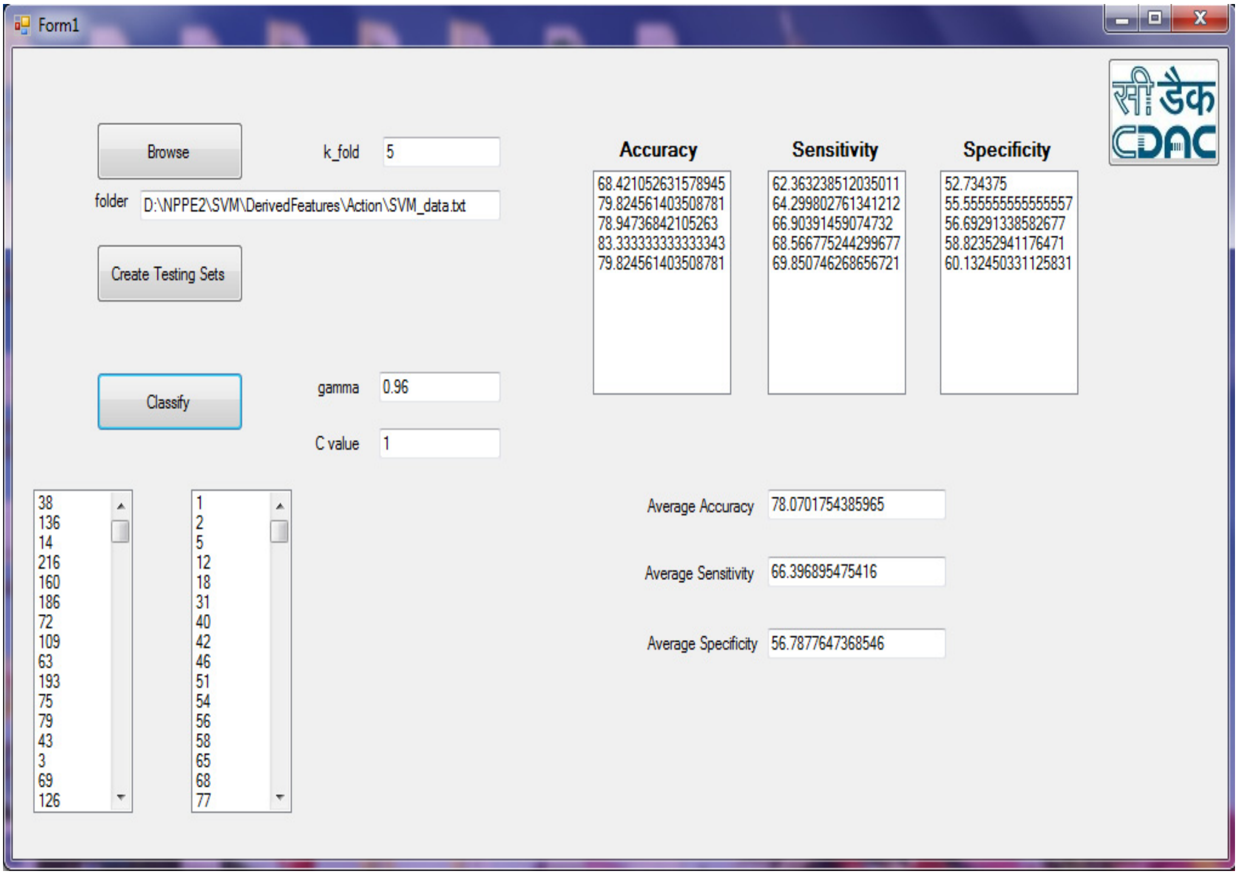


Fig. 4. SVM Classification output of global features

## Conclusion

The system developed by C-DAC, Kolkata consist of Perception based handwritten data collection, feature extraction and SVM based classification modules. The same has been utilized for field data collection and analysis. A set of 114 student's data have been collected in both true and distorted mode under three categories, namely, person, action and event. The spatio-temporal and global features have shown significant difference in true and distorted cases amongst all the features that have been taken into consideration. The classification output is showing consistent differentiability between true and false or distorted handwritten contents. First version of the integrated decision making system for deception detection based on perceptual cues of handwritten inputs has been developed.

So far the experiments have been done on student's data under simulated stressed environment for observing the effect of distortion in handwritten content. We have also collected some data where subjects were asked to imagine some events which they have not actually faced. To differentiate between handwritten content of a person writing a poem or imaginary story and a criminal's false statement we are to compare handwritten data taken from three conditions; at first, the subject is to be asked to imagine a situation and write about it, then after showing an eventful video true and false/distorted handwritten content is to be collected under the same condition. Finally subject will be asked to underline the intentional distorted words of the handwritten content. We have started working in that direction.

**Acknowledgements**

We sincerely acknowledge Dr. Amit Chaudhuri, Joint Director, C-DAC, Kolkata and Col. (Retd.) A.K.Nath, Executive Director, C-DAC, Kolkata for their constant support and help in executing the project. We sincerely thank Dr. Subhranshu Aditya, School of Cognitive Science ,JU for his valuable inputs in designing perceptual experimental set up. We also thank CID Guwahati, Assam and Don Bosco College and Engineering, Guwahati, Assam for providing support for Data Collection. The project is funded by DeitY (Department of Electronics and Information Technology), Govt. of India.

**Appendix A. 49 GLOBAL FEATURES**

1. dx= x ($1^{st}$ pen down) − x (last pen down) [length of the handwriting content]
2. dy= y ($1^{st}$ pen down) – y (last pen down) [width of the handwriting content]
3. Aspect Ratio= dx/dy
4. Average writing speed
5. Maximum writing speed
6. Total writing duration($T_w$)
7. $x_{max}$
8. $y_{max}$
9. $x_{min}$
10. $y_{min}$
11. $(x_{max} – x_{min}) \times (y_{max} – y_{min}) = A_{min}$
12. dx / $A_{min}$
13. Range Ratio
14. Standard deviation of x
15. Standard deviation of y
16. $x_{first} – x_{min}$
17. $x_{end} – x_{max}$
18. $x_{end} – x_{min}$
19. Maximum $V_x$ – Average $V_x$
20. Maximum $V_x$– Minimum $V_x$
21. Maximum $V_y$ – Average $V_y$
22. Maximum $V_x$– Minimum $V_y$
23. Maximum $V_y$– Minimum $V_y$
24. Minimum horizontal velocity
25. Average positive $V_x$
26. Average negative $V_x$
27. Average positive $V_y$
28. Average negative $V_y$
29. $V_x$ zero count [Total horizontal velocity components where value become 0]
30. Vy zero count [Total vertical velocity components where value become 0]
31. Pen down time($T_s$)
32. $T_w$ / $T_s$
33. Average writing velocity – Maximum velocity
34. Mean velocity – Maximum $V_x$

35. Mean velocity – Maximum $V_y$
36. Minimum $V_x$ – Average $V_x$
37. Minimum $V_y$ – Average $V_y$
38. $A_{min}$/ ($dx_{Pen Down}$ × $dy_{Pen Down}$) [ $dx_{Pen Down}$ = Difference in x coordinates where stylus is in contact with the tablet, $dy_{Pen Down}$ = differences in y coordinates where stylus is in contact with the tablet ]
39. $(x_0 - x_{max}) / dx_{Pen Down}$
40. $(x_0 - x_{min}) / dx_{Pen Down}$
41. $(x_{end} - x_{max}) / dx_{Pen Down}$
42. $(x_{end} - x_{min}) / dx_{Pen Down}$
43. $(y_0 - y_{max}) / dy_{Pen Down}$
44. $(y_0 - y_{min}) / dy_{Pen Down}$
45. $(y_{end} - y_{max}) / dy_{Pen Down}$
46. $(y_0 - y_{min}) / dy_{Pen Down}$
47. Factor $[\{(x_{max} - x_{min}) / (y_{max} - y_{min})\}/\{(dx_{Pen Down}/dy_{Pen Down})\}]$
48. Standard deviation of x / dx
49. Standard deviation of y / dy

## Appendix B. SPATIAL AND TEMPORAL FEATURES

1. Mean Pressure
2. Standard deviation of Pressure
3. Mean stroke duration on tablet surface
4. Mean stroke duration in air
5. Standard deviation of the stroke durations on tablet surface
6. Standard deviation of the stroke duration in air
7. Mean stroke length
8. Mean stroke width
9. Mean stroke height
10. Standard deviation of the stroke lengths
11. Standard deviation of the stroke widths
12. Standard deviation of the stroke heights

## Appendix C. SIGNIFICANT GLOBAL FEATURES FROM MANOVA ANALYSIS

1. dx
2. dy
3. Aspect Ratio
4. Average writing velocity
5. Range Ratio
6. Maximum writing velocity
7. Total writing duration
8. $x_{max}$
9. $y_{max}$
10. $A_{min}$
11. Writing length / $A_{min}$
12. $x_{end} - x_{min}$
13. Maximum $V_x$– Average $V_x$
14. Maximum $V_x$– Minimum $V_x$
15. Maximum $V_y$– Average $V_y$
16. Maximum $V_y$– Minimum $V_y$
17. $V_x$ zero count
18. $V_y$ zero count

19. Average velocity – Maximum velocity
20. Mean velocity – Maximum$V_x$
21. Mean velocity – Maximum $V_y$
22. $(y_0 – y_{max})$ / $dy_{Pen Down}$
23. $(y_0 – y_{min})$ / $dy_{Pen Down}$
24. $(y_{end} – y_{min})$ / $dy_{Pen Down}$
25. Standard deviation of y / dy

## Appendix D. SIGNIFICANT SPATIO-TEMPORAL FEATURES

1. Mean stroke duration on tablet surface
2. Mean stroke duration in air
3. Standard deviation of the stroke durations on tablet surface
4. Standard deviation of the stroke duration in air
5. Mean stroke height
6. Standard deviation of the stroke heights

## References


1. Osborn, A. S. "Questioned Documents". 2nd ed. Nelson-Hall, Chicago, 1929, pp. 205–216, 226–233, 247–248, 363–376
2. Srihari, S. N., Cha, S.-H., Arora, H., and Lee, S. "Individuality of handwriting", Journal of Forensic Sciences (2002) 47:856– 872
3. R. Plamondon and Sargur N. Srihari, "On-line and Off-line Handwriting Recognition: A Comprehensive Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, January 2000.
4. Sara Rosenblum, Shula Parush, Patrice L. Weiss. "Computerized Temporal Handwriting Characteristics of Proficient and Non-Proficient Handwriters", American Journal of occupational therapy March/April 2003, Volume 57, Number 2
5. Sara Rosenblum, Shula Parush, Patrice L. Weiss, "The In Air phenomenon: Temporal and spatial correlates of the handwriting process Perceptual and Motor Skills", 96(3), 933-954
6. Luan L. Lee, Toby Berger, and Erez Aviczer,"Reliable On-Line Human Signature Verification Systems", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 6, June 1996.