PILOT STUDY
A New Experiment on Signature Recognition

S. Mukherjee[1] *and S. K. Mondal*[2]

Abstract. *The complexity in signature recognition problems lies in the fact that a signature usually comprises a small number of handwritten letters that may have limited identifying features and, at the same time, contain natural variations from one signature to the next. Even though it is a frequently encountered problem in forensic sciences, the document examiner's common method of comparing a questioned signature with a group of control signatures depends upon human perceptual and cognitive processes that are often subjective. To reduce the subjective element in signature comparisons, the authors have experimented with an interactive signature recognition system called the Matching Index (MI) that allows a forensic document examiner (FDE) to utilize his/her expertise to select comparable characteristic features in both questioned and control signatures as well as introduce greater objectivity in the comparison process. An evaluation and assessment of the MI developed by the authors of the questioned signature with the group of control signatures has been implemented by considering numerically assessable features and quantitatively accounting for the information on natural variations manifested in the group of control signatures. Such a comparison provides a more objective expert opinion to present to the court. Successful results of a preliminary test, suggest that the present experiment is potentially promising to provide a more reliable and less subjective approach to signature identification.*

Reference: Mukerjee, S., Mondal, S.K. (2016). Pilot Study: A New Experiment on Signature Recognition. J. Forensic Document Examination, Vol. 25, pp. 36-45.

Key Words: Signature matching, Signature identification, Signature similarity, Feature ambiguity, Feature similarity, Handwriting recognition.

1. Introduction

The legal importance of signature recognition in criminal cases cannot be overemphasized as such cases are not merely confined to verification of a disputed signature, but also have a great significance in the sense that establishment of authorship identifies the specific author. In many criminal cases, an FDE is supplied with one questioned signature and a several sample signatures from the suspect to establish a match or a mismatch. As compared to handwriting samples, the number of letters and handwriting features available in a signature are limited and thus, the problem of signature recognition can be more challenging. Since the court's decision in *U.S.A. v. Starzecpyzel* (1995) in which a federal judge determined that handwriting identification was not a science but a technical skill, engineers in pattern recognition and FDEs worldwide have investigated ways to increase the accuracy and to reduce the bias and subjectivity that can enter the examination process.

The different methods followed by pattern recognition engineers are largely centered around artificial neural networks or these methods resort to general techniques of pattern recognition that

1. State Forensic Science Laboratory, 37/1/2 Belgachhia Road, Kolkata 700037, India. Email: Susantamukhopadhyay9@gmail.com
2. Email: skmandalfsl@gmail.com

are most suitable for classification of a large set of signatures rather than providing FDEs with a basis for an opinion in casework. The neural network approach demands a large control signature database for a successful training, while the same database is not available in practical case situations, as already noted. Furthermore, a totally automated opinion may not be accepted until the courts determine that this method is completely reliable. Thus, even though the published research carried out by pattern recognition engineers has enriched the literature with varied approaches on feature extraction and feature matching based on different algorithms, FDEs still must resort to their traditional methods of examining handwriting features in signature comparison cases, examinations carried out by human perceptual and cognitive processes.

To assist FDEs provide more accurate and objective opinions to the courts, the authors have developed an interactive signature recognition system that considers numerically assessable characteristic features of the signatures as well as quantitatively accounts for natural variation in the control signatures. In section 2 that follows, numerically assessable characteristic features that have been considered for this research have been introduced followed by the proposed procedure for comparing the questioned and control feature data. Section 3 will provide the results of present study and recommendations for further research.

2. Material and methods

2.1 Measurable signature features

In general, the control signatures were collected on plain paper and the signatory had ample freedom for signing his/her signature. The questioned signature appeared on an important document that may or may not have had a prefixed allotted space for placing the signature. As a result, the sizes of the signatures may have differed, but the basic characteristics were still present. The measurable features that were selected had to be defined as to make them scale invariant. To measure these features, the signatures were scanned to produce digital images. Consideration was given to the paper alignment while placing the signature in the scanner to not produce skewed images especially when the slant of letters was to be measured. To add to the accuracy of the features to be measured, a baseline connecting suitable lowermost points of the first and the last letters was drawn by interactively marking those two points. The end points of the baseline need not strictly be the lowermost points of the first and last letters since these points may vary depending on the alignment of the scanned signature. The end points of the baseline in signatures having a long beginning/ending trail may also be difficult to identify. In such cases the end points for this research were chosen to be as close to the specified location as possible and at a geometrically well defined characteristic point of the signature as can be seen in Figure 1 where the encircled end points have been connected to form the base line.

In fact, such consideration has always been made during all sorts of point marking on the signature during feature selection processes, as it may be noted that the different feature values considered below are derived only from interactive selection of some well defined points on the signatures. Thus, all sorts of inputs fed to the proposed system are from marking points on scanned signatures by placing a graphic cursor on it and the possible input error needs to be checked.

In a validity test on accuracy of interactive point marking, 10 points on a signature were selected and were marked 20 times each and the X and Y co-ordinates were recorded in terms of pixel value. Next, the standard deviations for errors along both X and Y axes were computed for each of the 10 points. The maximum standard deviation along the X axis was found to be 1.39 times a pixel width and the same was 1.32 times a pixel width along the Y axis, while the standard deviations were below one pixel width for seven points along the X axis and for six points along the Y axis out of 10 points considered. It makes the maximum possible error, $(\mathrm{MaxXerror}^2 + \mathrm{MaxYerror}^2)^{1/2}$, in point marking as 1.9 times of one pixel width that amounts to only 0.006 of an inch, since the signatures have presently been scanned with a resolution of 300 dpi for digital measurements on them. Thus the input error for interactive point marking is significantly low and reliable. The length of the base line has been set as the unit of all linear measurements making them scale invariant, and

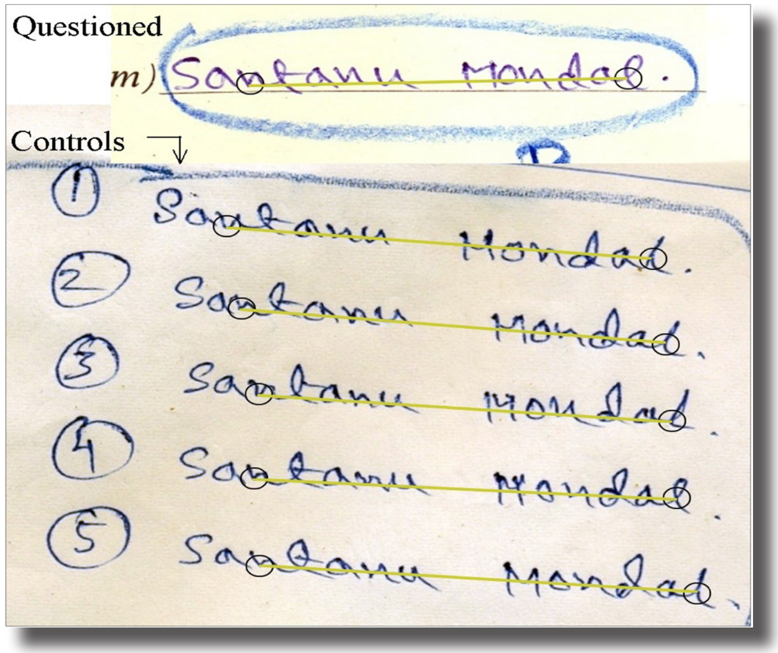

Figure 1. Example of end point selection for formation of the base line.

angular measurements have been carried out with respect to the baseline to deal with the above noted alignment problem. The different features that have been considered in the present study are defined below.

1. **Slant:** This feature is applicable to any inclined linear or nearly linear segment of any letter in the signature having well defined end points (upper-lower or left-right) that are marked interactively with the aid of a graphic cursor. Subsequently, an alignment invariant feature, slant, is computed as the angle subtended by the line joining the marked points with the base line that has been expressed in degrees. It is quite clear that this definition of slant is applicable to all sorts of inclined line segments produced either by an upstroke or by a downstroke.

2. **Gap:** The linear distance between the end point and the starting point of two consecutive disconnected letters, measured in units of the length of base line, serves as a scale invariant feature gap.

3. **Height Ratio (HR):** The heights of the first and third letters of the signature from the base line are evaluated by interactively marking the respective top most points of the letters and then computing their perpendicular distances from the base line. HR is defined as the ratio - smaller height divided by larger height.

4. **Line Length (LL) (Figure 1):** The total line length of any continuous segment of the signature having well defined end points measured in units of base line has been considered as the feature LL. Such a line being non-linear, in general, is fragmented into small linear segments by marking points at close intervals and the total length is computed by adding up the lengths of those small linear segments. In case an unbroken line does not have geometrically well defined end points, its largest fragment with well defined end points can also be considered. This feature is applicable to wavy parts of a signature where the letters are not explicitly legible or in case of garlanded letters or even for the whole signature in case the same is unbroken and thus continuous.

5. **Proportion (PPN):** This feature defines the shape of a letter in terms of the ratio of maximum width and maximum height. PPN is suitable for many letters like *a, b, d, f, g, h, k, o, s, t* and their capital forms as well.

6. **Arc**: Curvature of letters like, *m, n, u,* etc., can be utilized as a measurable characteristic feature by marking the point where the letter is most curved as well as two free end points on

its two sides. The angle between the two lines from the end points to the most curved point is evaluated and considered as the feature ARC. This feature can also be measured for letters like *V, W, Z and L* while those occur in a signature.

7. **Arc Direction (AD):** Document experts examine the direction of arc in case of letters that have rounded loop-like structure. Presently, AD is assessed as a two valued feature by visually examining the point of start of the letter and the direction followed to complete the letter formation. The feature AD is assigned a value "+1" for a clockwise direction and for a counter clockwise direction "-1".

8. **Dot Inclination (DI):** The position of dot above the letters *i* or *j* can be a significant feature in case such a letter belongs to a signature. The inclination of the line joining the dot and the top point of the corresponding letter with respect to the base line expressed in degrees has been considered as the feature DI. This angle would be acute in case the dot appears to the right of the letter *i* or *j*. Otherwise it would be obtuse.

2.2 Comparison criteria

While comparing a questioned signature with a group of control signatures, natural variations associated with the control signatures had to be quantitatively assessed before making a decision whether the questioned signature was a match or mismatch. To do so, the authors considered only those features that could be assigned with numerical values, the features detailed above. As an example, looking at the control feature F_C, the measured values of F_C for the group of control signatures manifest the natural variations of the suspected signatory's handwriting. The corresponding questioned feature F_Q has a single measurable value that is unique, while F_C is ambiguous as noted. Now, it is necessary to formulate a procedure that can yield quantitative information regarding the similarity of unique F_Q with ambiguous values of F_C. This quantitative information will help make a decision regarding matching/mismatching. An implementation of such a procedure needs some statistical considerations as follows.

The mean (M) and the standard deviation (D) of measured values of F_C for all the signatures in the control group are first computed. Now, this feature (F_C) may be regarded as an ambiguous feature for the control group, defined as the set $\{M - D < F_C < M + D\}$ that manifests its ambiguities due to natural variations. Such ambiguities vary from one person's signatures to that of another, and the same is expected to be more or less similar when evaluated for different sets of statistically significant number of signatures of an individual. Notably, a similar observation on selected features in cursive handwriting under controlled conditions has already been made elsewhere in another context.

For comparing the single and thus unique questioned feature F_Q with the ambiguous control feature set F_C, the authors assumed that the questioned feature F_Q has the same ambiguity around its unique value as that of the control feature set F_C, i.e., it may also be considered as an ambiguous feature set $\{M - D < F_Q < M + D\}$. Such an assumption amounts to considering that the questioned and control signatures were of the same individual in an ideal situation where the total number of control signatures was statistically significant. In case this assumption is true, the questioned and control feature sets must have a definite intersection. Thus, the authors defined the ratio of intersection and union of the two questioned and control feature sets F_Q and F_C as a suitable measure of their similarity and define the same as similarity factor for the feature F, as follows.

$$S_F = (F_Q \cap F_C) / (F_Q \cup F_C) .$$

Similarity factor for any particular feature, as expressed by the above ratio, has been converted to a percentage value and it is 100% only in a case of total overlapping of the ambiguous questioned and control feature sets, while for all other cases, in general, the larger the overlapping region, (i.e., intersection is larger is the value of similarity factor). But, a mere existence/non-existence of an intersection for a particular feature should not be taken for a sufficient condition to decide on match/mismatch of the whole signature since the number of control signatures defining the ambiguity is not statistically significant in practical

Signature category	Number of cases examined by present method	Agreement with results of conventional method	Non- agreement with results of conventional method
Normal hand forgery	2	2	Nil
Simulation or Free hand forgery	40	40	Nil
Trace forgery	2	2	Nil
Known Signature	20	20	Nil

Table1. Types of cases on signature recognition considered in the present experiment.

cases. Further, natural variations for different features are different and similarity factor is controlled by such natural variations that define the ambiguity of features due to natural variations. In fact, the nature of natural variations of a person's signature depends on his/her writing habits and handwriting training, as well as on many other personal and environmental factors during signing. All such factors cannot be framed by any specific mathematical rule. As a result, for a final decision on match/mismatch, the authors had to simultaneously consider as many features as possible and look for a suitable quantity that accounts for all such feature-similarities in an integrated manner. The simplest and straight forward candidate for the final matching index (MI) for the questioned signature is the average value of similarity factors for all features in the signature that have been considered. In fact, it is the total number of features considered that is important in the present analysis – even if these features appear infrequently. In fact, a signature being a collection of only a small number of letters, a great number of features are not expected to be present in a single signature.

The authors' next task was to ascertain whether or not the observed magnitude of such a MI for a questioned signature, MI_Q, is adequate enough to decide on match/mismatch. This MI as derived from all the similarity factors for different features should depend on the inherent natural variations in the control signatures on which the similarity factor depends. So, a decision on the adequacy of observed MI_Q requires a similar controlled analysis with known signatures of the same individual that can provide us with acceptable values for MI_Q. To find such values, the authors considered the group of control signatures only and isolated one signature from the control group (say C1). Now, C1 is treated as the questioned signature and the matching index for C1 with the rest of control signatures, MI_{C1}, is computed in a similar way as noted above. Such a computation is repeated for other signatures of the control group as well in order to get the values of all other control matching indices, e.g., MI_{C2}, MI_{C3}, etc. This controlled test on matching indices will serve as the guide for the range of acceptable values of MI_Q for a case of matching and the final inference can be made on that basis.

3. Results and discussions

Forged signatures are commonly classified under four categories according to their modes of reproduction: simply spurious; freehand simulations; traced simulations; and auto forgeries (Huber & Headrick, 1999). This experiment included each of these types of forgeries as well as the control signatures. However, the authors could not conduct the present experiment with a larger number of signatures because of a lack of access to a larger sample. The authors have tested the proposed method on no less than 60 cases belonging to different categories as noted above. All the signature samples were examined independently by the two authors using both the existing conventional methods of examination and by the proposed method in this paper.

To avoid any type of bias, each of the authors was unaware about the outcome of the other's examination results until the final comparison of all the results. The final decision regarding matching/mismatching of each questioned signature with corresponding

Feature name	Questioned feature	Control feature 1	Control feature 2	Control feature 3	Control feature 4	Control feature se FC	Similarity factor (%)
SLANT	61.453	58.797	60.309	53.89	57.971	55.365 < FC <60.119	12.3
SLANT	56.979	61.389	53.81	58.371	52.862	53.154 < FC <59.516	89.9
SLANT	59.663	66.814	60.86	59.036	57.673	57.606< FC <64.586	65.9
GAP	0.247	0.36	0.438	0.351	0.365	0.344< FC < 0.414	0
HR	0.453	0.435	0.402	0.442	0.381	0.39< FC <0.44	13.6
HR	0.458	0.605	0.638	0.592	0.489	0.525< FC <0.637	0
PPN	0.166	0.152	0.179	0.178	0.153	0.153< FC < 0.179	100
PPN	0.71	0.959	0.979	0.719	0.807	0.758< FC < 0.974	16.1
PPN	0.283	0.142	0.191	0.189	0.158	0.149< FC <0.191	0
ARC	58.776	69.902	68.422	66.718	87.58	64.752< FC < 81.56	7.8
DI	26.252	13.797	17.225	15.832	16.079	14.498< FC <16.968	0
Matching Index of questioned signature, MIQ = 27.79 %							

Table 2. Observed feature values for computation of matching index of questioned signature with respect to control signatures as displayed in figures 2 and 3.

control samples as inferred from the present method was compared with the findings based on conventional document examiners' procedures.

The cases of different categories taken up in this experiment have been listed in Table 1 along with results of comparison of present conclusion with that of a conventional document examiner. A total agreement of the two sets of findings is a bit surprising and appears like a chance coincidence as the sample size that the authors considered was not statistically significant. Therefore, the authors are not in a position to claim the method as 100% successful. However, even with a small number of signature samples, the agreement could not be ignored.

The first step in the present experiment for comparing a questioned signature with a group of control signatures is searching all available common features among the signatures under consideration by a visual inspection of the questioned and the group of control signatures. The second and final step is interactive selection of the associated features appearing in both the questioned as well as the control signatures by placing a graphic cursor over a few points on the specific digitized signature followed by a mouse click as has already been noted in sub-section 2.1. The rest of the experiment is a programmed computation of both questioned and control matching indices that leads to the final decision on match/mismatch.

To give the readers an idea about the feature values in the experiment, the authors cite a practical case example in which the suspected signatory denied that the questioned signature (Figure 2) was executed by him. A set of his sample signatures was collected by the police as displayed in Figure 3. Relevant feature data for computing the MI of the questioned signature, MI_Q, have been furnished in Table 2, the final computed value being given in the last row. The control matching indices, MI_{C1} etc., were computed in a similar fashion and those were found to be 27.41%, 27.08%, 35.45% and 33.87% respectively, indicating an acceptable range for MI_Q as, 27.08% – 35.45%. This result leads to a conclusion that the questioned signature and the control signatures were of the same individual. As noted already in Table 1, the outcome of conventional FDEs' examination was also the same and that led to a personal identification of the suspected signatory. It may be noted that, even though it was a case of matching, the similarity factor was

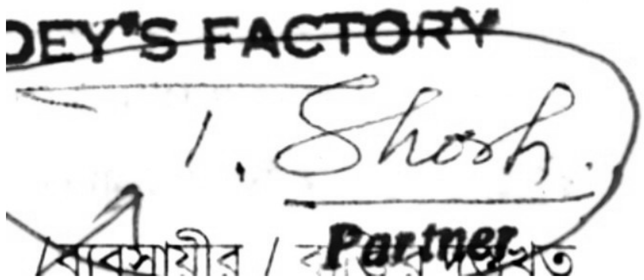

Figure 2. Questioned signature.

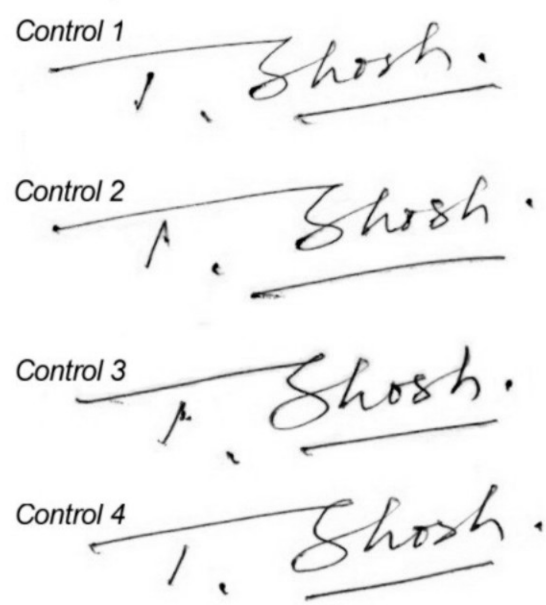

Figure 3. Control Signatures.

found to be vanishing for four features out of a total number of eleven. This is due to the ambiguity in the control feature set as defined in sub-section 2.2 that is not strictly true since the total number of control signatures is not statistically significant. In fact, while computing control matching indices, MI_{C1} etc., the authors also encountered such occurrences for a few features for the same reason. Thus, a few vanishing similarity indices for questioned signature should not be taken for a confirmed case of non-matching. Similarly, mere existence of non-vanishing similarity factors in some features cannot be accepted as a matching case. In fact, the larger the number of control signatures, the smaller the number of such occurrences, while a simultaneous consideration of a good total number of features in an integrated fashion led to a faithful inference as observed in all the cases examined.

Some typical values of questioned MIs and the corresponding ranges of control matching indices for signature forgery of different categories have been displayed in Table 3 that shows the results for both matching and mismatching cases.

During feature selection, often, all the features introduced in sub-section 2,1 or even a good variety of such features-types may not be available in a signature because the number of letters in a signature is too limited unlike that in a handwriting sample. But, this does not necessarily obscure the comparison process. As in such cases, repeated appearances of only a few types of features can facilitate considering a good total number of features in a signature, especially in cases of short initials and such cases were also dealt with by the present method and agreed with the outcome of conventional document examiner's approach.

Type of example	Matching Index of questioned signature (MIQ)	Range of MIC amongst control signatures	Present Inference	Conventional expert opinion
Normal hand forgery	15.2%	23.87 – 26.67%	Dissimilar	Dissimilar
Recorded simulation	13.78%	19.31 – 30.45%	Dissimilar	Dissimilar
Recorded trace forgery	12.15%	17.05 – 30.85%	Dissimilar	Dissimilar
Signature by known person	31.11%	27.33 – 45.97%	Similar	Similar
Questioned signature denied by signatory	26.77%	16.61 – 32.71	Similar	Similar
Questioned signature denied by signatory	27.79%	27.08 – 35.45	Similar	Similar
Simulated forgery	12.81%	22.85 – 27.89%	Dissimilar	Dissimilar
Qstnd. signature compared with First suspected signatory	17.47%	10.44 – 41.56%	Similar	Similar
Same signature compared with Second suspected signatory	6.44%	22.34 – 31.16%	Dissimilar	Dissimilar
Signature by known person	22.19%	17.87 – 27.46%	Similar	Similar
Initial	49.09%	42.79 – 68.38%	Similar	Similar
Initial	48.24%	36.73 – 50.76%	Similar	Similar

Table 3. Typical values of questioned matching indices and their acceptable ranges in cases of different types of forged signatures,

The main limitation of the present experiment is that the authors were unable to consider a large collection of signature samples. However, altogether 10 difficult cases of initials were considered successfully dealt with were considered in the present experiment. A larger sample of signatures especially with cases of highly skilled forgery will assist in determining the effectiveness and repeatability of the described method. As such, a detailed account of the present experimental procedure has been reported so that any programmer can easily develop the software suited to accomplish the simple tasks like, feature selection and result computation.

A further and straightforward exploration of the present method is applying it to signatures in other languages as well. To date, the authors' forensic laboratory has successfully conducted a few controlled tests on signatures in Bengali with promising results.

Framing new numerically measurable features is also a possible way of improving the present method and while doing so one should take care that the newly introduced features must be scale invariant as well as alignment invariant. A prospective new feature, that has not been considered here, may be derived from variation of pixel densities along notable signature-segments, which is a static manifestation of a dynamical aspect of signature characteristic that can hardly be imitated. Consideration of such features is expected to be effective for cases of highly skilled forgery or simulation that are also quite difficult to deal with by document examiners' conventional methods. An evaluation of such a density gradient demands an ingenious algorithm that is worth considering. Further, extending the applicability of this method to handwriting examination is expected to increase its versatility to a greater extent.

While the research presented is not totally automated, it is a simple and easily semi-automated system in which human expertise of selecting and locating appropriate characteristic features has been complemented with a data-based assessment of those features leading to an objective conclusion on signature comparison.

Finally, the authors hope that the method reported in this paper will stimulate researchers to formulate new ideas and innovations leading to a user-friendly and effective working tool for the document examiners.

References

Eldridge, M.A., Nimmo-Smith, I., Wing, A.M., Totty, R.N. (1984).The Variability of Selected Features in Cursive Handwriting: Categorical Measures, Journal of the Forensic Science Society. 24, 179-219.

Ellen, D. (2005). Scientific Examination of Documents: Methods and Techniques, Third Ed. CRC Press, Boca Raton, Florida.

Fang,B., Leung, C.H., Tang, Y.Y. Tseb, P.C.K. Kwok, K.W, Wong Y.K. (2003) Off-line signature verification by the tracking of feature and stroke positions, Pattern Recognition. 36, 91-101.

Hilton, O. (1982) Scientific Examination of Questioned Documents, Elsevier Science Publishing Co., New York, NY.

Hanmandlu, M., Hafizuddin, M., Yusof, M., Madasu, V.K. (2005). Off-line signature verification and forgery detection using fuzzy modeling, Pattern Recognition. 38, 341-356.

Huber, R.A. & Headrick, A.M. (1999). Handwriting Identification: Facts and Fundamentals, CRC Press, Boca Raton, Florida.

Impedovo, D. & Pirlo, G. (2008). Automatic signature verification: the state of the art, IEEE Trans. Syst., Man Cybern. C, Appl. Rev. 38, 609 -635.

McCabe, A., Trevathan, J., Read, W. (2008). Neural Network-based Handwritten Signature Verification. Journal of Computers. 39, 22.

U.S. District Court Ruling, United States v. Starzecpyzel, 880 F. Supp. 1027 (S.D.N.Y 1995).

Srihari S. N. & Leedham, G. (2003). A survey of computer methods in forensic document examination, Proc. 11th International Graphonomics Society Conf., Scottsdale, Arizona. 270-281.

Acknowledgement

The authors express their gratitude to Dr. P. Sinha for his continued interest, enlightening suggestions and also for a critical reading of the manuscript.