Journal of Forensic Research
ISSN: 2157-7145
Brugnatelli, J Forensic Res 2013, 4:5
DOI: 10.4172/2157-7145.1000206
Research Article
Open Access

The Variability Measurement of Shape Independent Features to Establish a New Method to Differentiate Genuine Signatures from Simulated Ones

Fausto Brugnatelli*

Member of National Association of Document Examiners, Italy

Abstract

A preliminary step to the examination of signatures is the objective assessment of the variability range that is peculiar to the writer; in document examination it is quite rare to come across standardized procedures allowing an effective discrimination between genuine and simulated signatures.

This study took into consideration 104 questioned signatures (coming from 20 real and Judge-ruled cases) that were inspected to evaluate non-shape-related indicators in their ability to assess correctly genuineness or simulation; the research particularly focused on the evaluation of letter heights - with special reference to the positions of the individual letters within the signature - and on the subsequent height ratios between extenders (upper and lower) and the other small letters. The combination of the information gained from the statistical survey, assisted by calculation of confidence intervals, of these height indicators with the observation within the signature of the spots where between-letter pen lifts occur led to an effective discriminating method for medium/long signatures.

Data processing also yielded useful indications on the letters that most frequently trigger off height errors in forgery and provided other interesting results, such as the positions within the signature where simulators most frequently fail.

Keywords: Document examination; Signature; Simulation; Indicators; Letter height; Extenders; Connections; Pen lifts; Method; Rhythm; Variability; Confidence interval; Standard deviation; Forensic science

Introduction

In document examination - especially in signature assessment - two of the main issues are the definition of the handwriting variability peculiar to a writer and the application of step by step methodological standards capable to allow a successful differentiation between genuine and simulated signatures.

The definition of the individual variability range is essential for a successful distinction between what is really part of a writer's repertoire and what comes instead from the hand of another person [1,2]; a preliminary definition of the variability range of a given writer definitely proves advantageous for an expert in establishing whether the writing under examination is attributable to that writer.

Another fundamental aspect is the development and the application of reliable, repeatable standards to the signature verification process. In this respect, several researches were carried out taking into consideration the statistical investigation of some of the many numerical features measurable in writings: the use of statistical/computerized measurements [3,4] and the critical analysis of data are in fact expected to allow a more precise profiling of the writer's typical graphic habits, thus facilitating the whole verification process. Still in this measurement quest more than one research favoured the statistical study of features mainly associated with the shape of the letters - therefore focusing on "morphology", intended here in its etymology of 'related to the *shape* (the contour or outline) of the object' [5,6]. This shape-dependent approach entails the risk that when fortuitous or transient morphological changes occur in the signature, the mathematical ('graphometrical') data previously collected and classified as characteristic of that specific writer can accordingly result modified [7], consequently increasing the chances of an incorrect opinion being released as to the signatory. By way of an example one can think of the substantial changes in the shape of ovals ("a", "o", etc.) brought about when the signature is signed by the same person through an increase (or decrease) in the degree of the letters' axial slant or through boosted speed [8]. It is worth remembering here that a change of slant may not be such an uncommon writing occurrence, as it can happen for instance when a sheet of paper gets shifted, to the right or to the left side, away from the perpendicular to the writer's trunk (a significant rightward shift of the paper tendentially increasing the forward slant of the signature, while a leftward shift usually increases the backward inclination of the longitudinal axes, thus inevitably changing the original shape of the ovals and their geometrical/mathematical descriptions).

This matter adds to the known fact that it is quite natural for an imitator to be focusing his/her best imitating efforts on the production of graphic forms meant to be morphologically closest to the target sample; in other words, the forger's efforts are usually mainly concentrated on a plausible or even optimal formal rendition of the letters [9], also because shape-related features generally capture most of one's eye attention when anyone is occasionally called to check at a glance on the genuineness of a signature [10].

***Corresponding author:** Fausto Brugnatelli, Member of NADE (National Association of Document Examiners), 100% accuracy and best document examiner at the 2009 ST²AR Test, Via Giotto 24, 20145 Milano, Italy, E-mail: scripta@examina.net

Received July 26, 2013; **Accepted** November 19, 2013; **Published** November 22, 2013

Citation: Brugnatelli F (2013) The Variability Measurement of Shape Independent Features to Establish a New Method to Differentiate Genuine Signatures from Simulated Ones. J Forensic Res 4: 206. doi:10.4172/2157-7145.1000206

For the above mentioned reasons this study tried to shift its focus from the shape (the outline) of a writing to the less conspicuous, more stable and strongly characteristic indicators/graphical rhythms hidden within a signature [11]; ideally the best indicators would be those that are really peculiar to the specific writer, that are reasonably independent of accidental graphic events and that are hard to be noticed or reproduced even by skilled imitators.

This line of reasoning led to the investigation of shape-independent (non-morphologic) parameters as a way to determine the graphic variability of different writers so that, once a writer-specific variability range is defined, the expert can proceed to a sound comparison with the questioned signatures through a repeatable practice.

By no means this study is advocating the renouncement of the evaluation of all the information inferable by an unbiased analysis of the morphological features possessed by a signature (no factual information should ever be discarded while conducting a document examination); what is being proposed here is a change of perspective, which might prove useful if at last an expert would like to tackle the question of signature genuineness (or non-genuineness) by a route fundamentally different from the outline-dominated visual path usually followed by an imitator.

The main parameter investigated is the height of the small letters making up the signature; capital letters have been intentionally neglected because, being showy, are usually the favourite graphic targets attracting most of the imitator's care and effort [10,12].

The other main non-shape-related indicator considered here is the between-letter personal connection rhythm along the signature that is the study of the recurrent localizations of between-grapheme disconnections (or connections): in other words, the detection of the spots where interruptions or pen lifts typically occur while writing the various letters composing a signature.

Materials and Methods

The programs used in this study were an image editing software (Photoshop 7.0 or Gimp 2.6) and a spreadsheet (Excel 2007 or Open Office 3.2).

The examined signatures came from real Judge-ruled cases in the Courts of Italy. In details there were 104 questioned signatures concerning 20 different subjects (men and women aged between 35 and 78); in various occasions - 5 out of 20 instances - questioned signatures pertaining to a single subject simultaneously included genuine and simulated signatures.

The cases and their related signatures were recruited by a single prerequisite: the availability within the known group of signatures having a minimum of 10 measurable small letters.

The initial hypothesis was that, in long or average-sized signatures, the likelihood of errors made by imitators in the rendition of the typical height of the letters is enhanced, and this due to the fact that through a long imitating process simulated letters can turn out to be too big or too small compared to the usual genuine letter range in that specific position within the signature. The same hypothesis applies to the likelihood of mistakes concerning the spots where between-letter pen lifts preferentially occur.

Capital letters were excluded from the evaluation of letter heights and only the small letters were taken into consideration, with a total number of measurable small letters equal or greater than 10 per signature. For obvious reasons, the heights of completely unrecognizable threadlike letters in the middle or at the end of signatures were not considered.

The number of 10 or more measurable letters per signature included the letters having upper or lower extenders. From the data concerning the heights of the small letters and the measurement of the length of the upper extenders ("b", "d", "l", "t") or lower extenders ("g", "p", "q") the ratios between mean lower (or upper) extenders and the mean of the remaining small letters (with no extensions) were obtained, thus supplying another variability defined indicator to be compared with the respective ratio values in the questioned signatures.

As to the questioned signatures (Q), this study considered the binary case of genuineness or imitation; the possibilities of self disguise and traced forgery were not specifically addressed, even though among the 104 examined signatures oneself disguise case and one traced forgery were present, both of them correctly detected by the method.

Known signatures used for the comparison

For each questioned case 10 known signatures (K) were used. The number of 10 was suggested by the fact that in document examination practice a quantity of ten known signatures is in step with the usual actual availability of known exemplars. Moreover, the measurement of the longitudinal letter heights of 10 known signatures is not too time consuming (about two hours) and allows easy and readily done evaluations of the means [13]. If, among the ten known exemplars, a specific small letter happened to be always or frequently absent (or stretched in such a threadlike way that it was hardly possible to carry out a reliable assessment of its height), that letter was excluded from the evaluation grid. Conversely, if a specific letter turned out to be always and unequivocally present in all 10 known exemplars but not in the questioned signature, that occurrence was marked up as a potential error signal by the signatory (the orange colour was used here to distinguish this anomaly from the yellow colour usually utilized to mark up the potential height error signals on the spreadsheet).

If the available known samples were more than 10, the array of the actually examined known group was reduced to 10 using a strictly chronological criterion, that is selecting only those ten signatures which were closest in time to the date of the questioned signature. When a request writing bearing many signatures was available, a maximum of five random signatures was used so as to avoid the concentration of all ten known exemplars from the same sheet of paper and on the same single date.

By the above criteria the maximum time gap from the questioned signature date to the least contemporary signature in the pool of the ten known exemplars was 20 years. Known signatures usually fell in a much shorter time interval, that is between 5 years before and 3 years after the writing of the questioned signature. In a single case, the 10 known signatures were all written on 10 different documents in a time interval not exceeding one year from the date on the questioned document [case# 9, of which it is shown the table describing the distribution of the letter heights (Table 1)]. In this case the known (K) and questioned (Q) signatures neatly matched throughout all the letter heights, as well as in the ratio between mean upper extenders and mean small letter heights, and also in the connection/disconnection pattern along the signature (that is throughout all the indicators examined in this research).

Once the 10 known signatures were selected - adopting the criterion of the closest date preceding or following the date on the questioned document - the codes relevant to the ten known signatures (1, 2, 3…, 10) were entered into the rows of the worksheet, whereas in the columns were singularly entered the small letters to be measured for their longitudinal size (Table 2).

The heights were always measured following exactly the slant (forward or backward) of the specific questioned letter; therefore, by "letter height" here it is meant the top to bottom length of the letter measured not perpendicularly to the writing line but along the peculiar axial arrangement of that specific letter within the examined signature. Letter height values were measured in tenths of millimetre. The average questioned signature length was 12.4 measurable small letters, with most signatures being made of 11 measurements.

Confidence intervals and dimensional evaluation of the signatures

To assess the dispersion - thus the variability - of letter heights, the standard deviation of the measured values was calculated. Two standard deviations (2 sigma=2σ) on the left and on the right side of the mean include 95.45% of the dispersed values, leaving outside only the two farthest tails covering 4.55% of the population values [14].

In this study a double standard deviation (±2σ or, in a more statistically appropriate way, ±2SD) was used as confidence interval; in the questioned signatures statistically significant values exceeding this 2σ interval above or under the mean were considered as hypothetical forger's errors.

The calculation of the confidence interval was carried out for

	09 - % avelli §aria												
	Small letters height									Upper extensions height			
	a	v	e	i	a	r	i	a	Mean	l	l	Mean	HR
K1	38	24	26	23	25	20	23	27	25,75	42	41	41.50	1,61
K2	22	25	32	37	30	34	20	24	28,00	67	58	62,50	2,23
K3	27	26	26	23	28	28	28	27	26,63	39	41	40,00	1,50
K4	29	25	29	27	30	31	19	27	27,13	59	58	58.50	2,16
K5	30	26	25	27	23	31	23	22	25,88	50	42	46.00	1.78
K6	33	32	34	23	32	33	21	25	29,13	50	48	49.00	1,68
K7	31	26	25	27	33	37	25	39	30,38	44	44	44.00	1,45
K8	29	29	30	16	25	24	26	30	26,13	41	45	43.00	1,65
K9	26	20		18	35	38	25	31	27,57	38	43	40,50	1,47
K10	28	26	22	20	29	33	24	40	27,75	48	33	40,50	1,46
Mean	29,30	25,90	27,67	24,10	29,00	30,90	23,40	29,20	27,43	47.80	45,30	46.55	1,70
SD	4,27	3,11	3,84	5,92	3,83	5,59	2,80	6,03	1,48	9,26	7.72	7.92	0.28
Confidence interval from	20,76	19,69	19,99	12,27	21,34	19,73	17,81	17,13	24,47	29.28	29,86	30.71	1,13
to	37,84	32,11	35,35	35,93	36,66	42,07	28,99	41,27	30,39	66.32	60,74	62.39	2,27
Q genuine	33	29	28	20	36	23	23	29	27,63	42	38	40,00	1,45

SD=Standard deviation
H R=Height ratio between upper extensions and small letters

Table 1: Diagram to calculate letter heights and the ratio-1.

	*ertola £abio													
	Small letters height									Upper extensions height				
	e	r	o	a	a	b	i	o	Mean	t	l	b	Mean	HR
K1	58	33	27	38	17	46	17	23	32,38	103	141	60	101,33	3,13
K2	50	24	24	29	23	19	9	16	24,25	73	84	70	75.67	3,12
K3	57	21	20	31	15	16	7	x	23,86	70	76	62	69,33	2,91
K4	66	27	26	25	15	24	x	11	27,71	88	103	83	91,33	3,30
K5	83	34	26	19	16	22	x	11	30,14	105	107	90	100.67	3.34
K6	97	53	36	42	37	35	13	x	44,71	120	145	92	119,00	2,66
K7	58	18	25	19	17	27	x	26	27,14	105	95	75	91,67	3,38
K8	59	21	17	21	13	15	7	x	21,86	71	95	60	75,33	3,45
K9	72	14	26	21	16	27	x	18	27,71	89	103	76	89,33	3,22
K10	57	21	28	16	13	34	15	28	26,50	97	88	86	90.33	3.41
Mean	65,70	26,60	25,50	26,10	18,20	26,50	11,33	19,00	28,59	92.10	103.70	75,40	90,40	3,19
SD	14,45	11,19	4,99	8,71	7,18	9,63	4,27	6,88	6,42	16,93	22,76	12,21	14,66	0,25
Confidence interval from	36,79	4,23	15,51	8,68	3,85	7,24	2,79	5,24	15,74	58,24	58,18	50,97	61,08	2,70
to	94,61	48,97	35,49	43,52	32,55	45,76	19,88	32,76	41,43	125.96	149.22	99.83	119.72	3,69
Q simulated	54	7	45	30	29	19	4	39	28,38	95	123	128	115,33	4,06
Q genuine	59	20	25	19	25	26	x	26	28,57	103	95	73	90,33	3,16

SD=Standard deviation
H R=Height ratio between upper extensions and small letters
=signal of potential error

Table 2: Another diagram with signals of errors only in the simulated signature.

each known letter in its specific position within the signature; as a consequence, by way of an example, one matter is the typical height of the oval "o" executed in the second position after the capital letter (e.g. "Paolo") and another question - possibly with a different height value - is the mean typical size (plus its variability, as well as its confidence interval) of the "o" written in the fourth position after the capital ("Paolo"). This distinction stems directly from the notion of graphical rhythm in the execution of a signature [15]. The peculiar dimensional micro-fluctuations all along the letter heights make the signature an exclusively personal production; since these height variations get often unperceived by the naked eye, in the long run any simulator is liable to be deceived in the imitation process by this quantitative oversight.

Recapitulating, in a signature composed of eleven measurable small letters this method calculates eleven mean height values in the eleven positions taken by those letters; moreover eleven different letter variability ranges are worked out for the same eleven letters (Table 2). This applies also for the small letters having upper or lower extensions; if in the signature there are no upper or lower extenders, the further indicator concerning the ratio between the mean size of upper/lower extenders and the mean size of all the other small letters is obviously not assessable.

For those questioned signatures where the mean of the small letters without extenders over-exceeds or under-exceeds the confidence interval limits of the height mean of all the small letters in the known signatures, it is necessary to work out a standardization of each single and average value: the entire line of the values referring to the questioned signature showing an "abnormal" letter size mean must be multiplied by a correction factor able to relate the mean value of the questioned small letters to the central mean value of all the known signatures. When all the letter height values in the questioned signature have been recalculated by this correction factor, then it is possible to reassess - along each column - if the height values of the relevant "corrected" questioned small letters (in that specific position of the signature) fall or do not fall within the variability range set by the confidence interval obtained by the 10 known signatures; if at this point the questioned value does not fit into the confidence interval, the spreadsheet box is marked with a yellow signal of potential height error.

This prudential correction procedure of the questioned values prevents that a genuine questioned signature be erroneously considered as simulated by the simple fact that the *whole* signature had been written - for merely fortuitous reasons - in very large or very small dimensions compared to the usual practice.

Once all the letter height measures have been gathered and processed as above described, the evaluation of the possible height inconsistencies in the questioned data is done through the study of:

- the ratio between the mean of letter heights having upper extensions and the mean of the small letters without extensions;
- the ratio between the mean of letter heights having lower extensions and the mean of the small letters without extensions;
- the comparisons between the known signature height and the questioned letter height carried out through each column with respect to every single constituent letter of the signature.

Any anomaly here is marked in yellow as signal of potential error.

Further inconsistencies - the orange coloured signals of potential error - have to be recorded if a letter that is always present in the 10 known signatures (and the height of which is always clearly measurable in that group), is altogether absent in the questioned signature or is made in such a threadlike manner that it surpasses any alphabetic recognizability/measurability.

The evaluation of connections and pen lifts between letters

The other main non- indicator under examination is the alternating sequence of connections and between-letter pen lifts. This survey relies on the notion that every subject, in his/her typical signature, tends to express a personal preference for pen lifts in specific spots [12,16]. In the questioned signature group the detection of between-letter interruptions in spots that are instead regularly and always continuously inked in all the 10 known signatures was considered an additional signal of error (that occurrence was marked in black in the spreadsheet).

Thus further signals of potential error in a questioned signature are scored every time an anomalous between-letter stop is recorded in a letter-to-letter sequence which is instead always carried out in graphical continuity in all the 10 available known signatures. For example (Table 3 and Figure 1) in the signature "**ertola*", the "*t-o*" transition is always executed by a between-letter connection in all the 10 known exemplars (K1, K2…K10), whereas in the questioned signature Q1 (simulated) an interruption between the "t" and the following "o" occurs.

The threshold of signature discrimination significance for the signals of potential error

Due to the adopted confidence interval (+2σ till -2σ) that leaves as outliers all the values falling in the farthest tails of the curves - that is about 4.55% of the actual population for every distribution curve of each examined letter - in order to prevent a massive score of false positives during the final discriminating assessment of the questioned signature it is necessary to make sure that the signal-of-error-based conclusive opinion does not pivot on the recording of the chance outlying value (probability of 4.55%), but stems from *repeated* observations of anomalies and inconsistencies in the letter height measurements and/or in the connection/pen lift rhythm. Simplifying the notion of the degree of freedom and mutual independence of each signal of error, a discriminating threshold of at least three cumulated signals of error per signature was assumed as precautionary against false positiveness, theoretically allowing only a chance of about 1 case out of 10000 for an actually genuine signature to possess simultaneously a sum of three misleading signals of error.

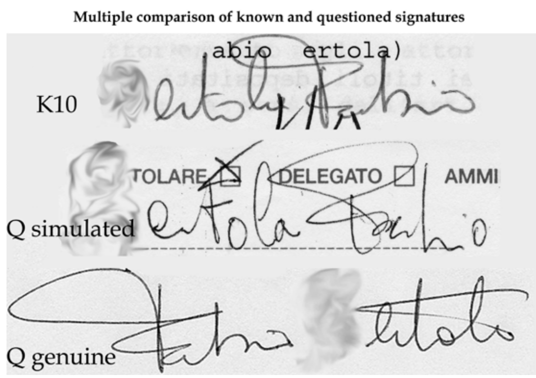

Figure 1: ertola multiple comparison.

Known exemplars	*		e		r		t		o		l		a	£		a		b		i		o
K1																						
K2																						
K3																						
K4																						
K5																						
K6																						
K7																						
K8																						
K9																						
K10																						
Pen lifts in knowns																						
Q simulated																						
Q genuine																						

Table 3: Diagram to show pen lift errors in simulated signature only.

Results

The 104 questioned signatures were made of minimum 10 to maximum 20 measurable letters and split into 57 simulated and 47 actually genuine signatures. For each of these questioned signatures the examination was first directed to the dimensional indicators - letter heights and heights ratios - and later on focused on the parameter dealing with connections/pen lifts.

Combining the results obtained by the above indicators, 103 out of 104 questioned signatures scored either at least 3 signals of error in the imitated group or 2 to 0 signals in the genuine group, thus being properly classified respectively as false or authentic, with a compound total of 99.04% correct assessments.

The only wrongly assessed signature was a simulated one (a false signature that resulted genuine according to this method).

The ratio between upper (or lower) extenders and the mean of the small letters heights

The data concerning the calculated ratio between mean upper extenders (or mean lower extenders) and the average height of all the other small letters showed high reliability [17-20]. When the questioned signatures were actually genuine, the ratio between upper/lower extenders and the height of the remaining small letters correctly fell, in 35 out of 35 cases, within the known exemplars variability range, without ever producing a misleading signal of error.

Out of the remaining 30 questioned signatures (this time the simulated ones) that chanced to have upper and/or lower extenders, 20 signatures showed a ratio which was outside the known variability range, therefore providing a correct signal of simulation.

This means that compounding genuine+simulated questioned signatures, in 85% of the cases (55 out of 65 cases) the indicator of the ratio between extenders and small letters height proved to be on its own definitely adequate and capable of correct discrimination both in the sense of genuineness - without any false positive - and in the sense of non-genuineness of the signatures. Such a high reliability level for this indicator alone is partly explained from a statistical point of view by the fact that the adopted parameter is based on the variability of values deriving from a ratio between means (i.e. the mean of upper and lower extenders compared to the mean of the heights of the remaining measured small letters) (Table 2); this fact gives rise to smaller standard deviation values, thus to a lower dispersion of the distribution curve around its mean and consequently a higher discriminating capacity of the curve itself.

The heights of the small letters in their respective positions along the signature

As statistically expected, lower levels of conclusive and unambiguous discrimination between simulated and genuine signatures were obtained from the indicator concerning the height of the various small letters in their respective positions along the signature [17,21]. Since the height of each letter in the questioned signatures has to be individually compared with the relevant height variability curve in the known group, it is natural to record a few misleading signals of error in one or a couple of letters composing a genuine questioned signature, especially so if the confidence interval ranges from +2σ to -2σ from the mean (more precisely ±2SD), thus accepting as statistically significant - in the pool of the potential signals of error - all the values falling in the two far tails of the curve and covering 4.55% of the theoretical population.

At 3 or more letter height errors per signature, the discriminating capacity of this identifying parameter proved to be fairly effective; as a matter of fact, in none of the genuine questioned signatures (that is 0 out of 47 cases) were simultaneously recorded 3 or more height errors per signature; this means that in the genuine questioned group a conclusive wrong signature assessment was never elicited by the sum of the recorded misleading height signals of error.

In the simulated group (total number: 57) 40 signatures showed 3 or more signals of errors from letter height anomalies. Combining this result with the fact that no wrong calls were produced in the genuine group, the letter height indicator alone allowed to identify correctly 83.7% of all the questioned signatures.

Between-letter pen lifts

In the genuine questioned signatures group no misleading error was found a propos of between-letter pen lifts (that is undue occurrences of pen lifts in places where these stops are never present in the known group); therefore, in 0 cases out of 47 genuine questioned signatures an extra pen lift was recorded in comparison with the pen lift pattern observed in the group of the known signatures.

On the contrary, whenever a disconnection signal of error was found, this always correctly indicated a simulated signature; the appropriate signal of error for the imitated signatures group was scored in 50 out of 57 simulated cases. Thus the aggregate data from the adequate/anomalous pattern of between-letter pen lifts amount to 93.2% correct calls.

It is to be noted, though, that this indicator, unlike the previous ones, is not based on a procedure whereby a numerical measure fits or does not fit into a given distribution curve of values (as for the ratios between extenders and small letters, or for the case of letter heights), but it is a simple observation in the questioned signature of its compliance (or lack of it) with the known exemplars pattern describing the places where between-letter interruptions usually occur. Consequently, this indicator can be useful to facilitate the identification of anomalies which can lead *synergically* to an opinion of imitation on the questioned signature, but clearly it would not be advisable to carry out genuineness calls on the basis of this parameter alone. Instead it appears a more reasonable approach that of using this letter connections/disconnections indicator to further integrate and perfect the already high discriminating rate obtained by the evaluation of the height ratios and the assessment of the letters' sizes.

It is finally to be noted that in 6 out of 47 genuine questioned signatures (12.8%) the evaluating procedure recorded an extra connection beside those already typically present in the 10 known signatures group; this sporadic fact bears no particular significance, being only a further between-letter linkage (that is a higher degree of complexity) added to the normal connection pattern observed in the known group; still, if in a questioned signature three or more extra connections were to be found beside the normal known signatures pattern, this score would need to be closely considered, because in this case the repeated addition of anomalous between-letter connections could indicate a slapdash forgery attempt, that is a simulation carried out without an in-depth preliminary study of the to-be-imitated signature or without any real knowledge of the target signature.

Overall combined results and summarization of the differentiation method

The above data combine into a completely non-morphologic analytical method for a rapid discrimination between genuine and simulated average to long signatures. Sequencing the three following summarized steps it was possible to identify correctly 56 simulations out of 57 imitated signatures, at the same time precisely assessing as genuine all of the 47 non-simulated signatures; this summed up to an accurate identification of 103 signatures out of 104, with a final rate of 99.04% correct calls.

This shape independent method is briefly summed up by the following steps:

1) The height of each letter composing the questioned signature is calculated and a comparison is made between every single questioned measure and the confidence interval obtained in the known group for that letter height in that specific position within the signature. If during this process on letter sizes (here including the possible anomalous absence of a letter in the questioned signature when the same letter is constantly present and measurable in all the known exemplars) 3 or more error signals emerge, the signature is classified as simulated; otherwise this score has to be added to those obtained in step #2 and step #3;

2) Assessment of the ratio between upper and/or lower extenders and the height of the remaining small letters; if in the questioned signature this ratio falls outside the confidence interval of the ratio curve for the 10 known signatures, one signal of error is recorded;

3) The last step is the inspection of the between-letter pen lifts pattern; if during this last stage, the questioned signature shows anomalies for the presence of extra between-letter *disconnections* with regards to the known exemplars pattern, these occurrences are recorded as signals of error. The pen lifts signals of error must be accumulated with the signals of error counted during the examination of the ratios (Step 2) and the analysis of the letter heights (Step 1): if the final sum of the various signals of error reaches or exceeds the overall count of 3, the questioned signature is considered simulated, whereas with an inferior sum of potential errors (equal or lower than 2) the signature is regarded as genuine.

The diagrams here show the trend of the error signals in the 47 genuine questioned signatures group (Figure 2) and in the 57 simulated signatures group (Figure 3). It is clear that, through a tolerance limit per signature of up to 2 potential signals of error for letter heights and/or between-letter pen lifts, a better selectivity is achieved thanks to the establishment of a prudential protection buffer towards those cases where the questioned signature is actually genuine despite having one (or two) uncommon values (outliers) in its letter height measurements.

By the joint observation of Figure 1 with Tables 2 and 3 stands out the usefulness of this method, which is capable of adding more discriminating evidences to the already discernible shape-related discrepancies (e.g. the lower outline of letter "b") by the objective assessment of statistically significant differences in the heights of both letters "o", in the height of the long stem characterizing letter "b" and especially in the ratio between the upper extenders and the other small letters.

Places and frequencies of height or pen-lifts errors

Through the measured data it was possible to investigate the subject of the specific positions inside the signatures where forgers' height errors tend to be more frequent. A count was made of the height errors frequencies in the various positions after the capitals, particularly focusing on the effective height errors frequencies, that is the frequencies of error refined from the share of the possible misleading error signals [that is those out-of-normal-range signals that chanced to crop up in the group of the genuine questioned signatures]. The height error frequencies in the various positions along the signatures are shown in the following histogram (Figure 4). From the available data it stands out that there are two loci where the imitators' height errors most frequently occur: in the fifth letter after the starting capital (29.33% of the cases) and in the first position just after the initial letter (27.28%); lower frequencies are instead noted in the other positions along the signature (Table 4).

The same procedure with data evaluation has been carried out to investigate the places within the signatures where pen lift errors more frequently occur (Table 5). The higher frequencies of error for this indicator were recorded in the 4th (38.46%) and 2nd (35%) letter-to-letter transitions; lower - but not dramatically so - frequencies were seen in the other positions along the signatures, with rates ranging from about 25% (1st and 3rd position) to about 18% (5th and 6th position).

Letters that most frequently trigger off height errors

Other valuable data emerged from the study of the small letters most frequently triggering off height errors during the signature imitation process. [NB: the pie chart on Figure 5 shows by groups the frequency percentages of the various letters making up all the examined signatures].

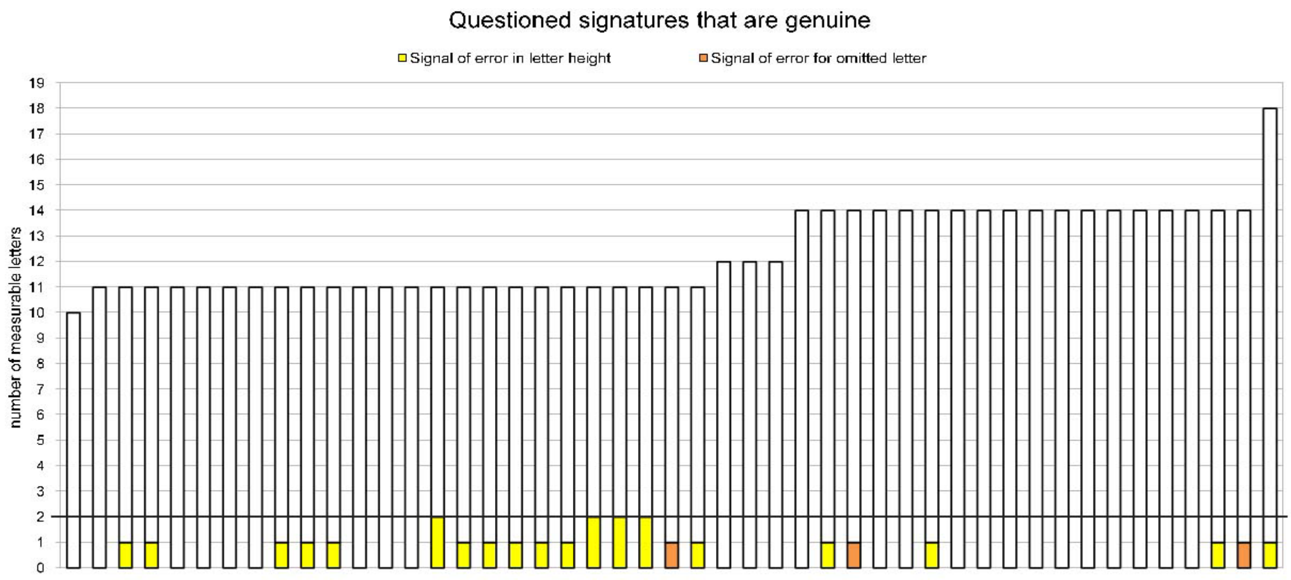

Figure 2: Signals of error in all genuine questioned signatures.

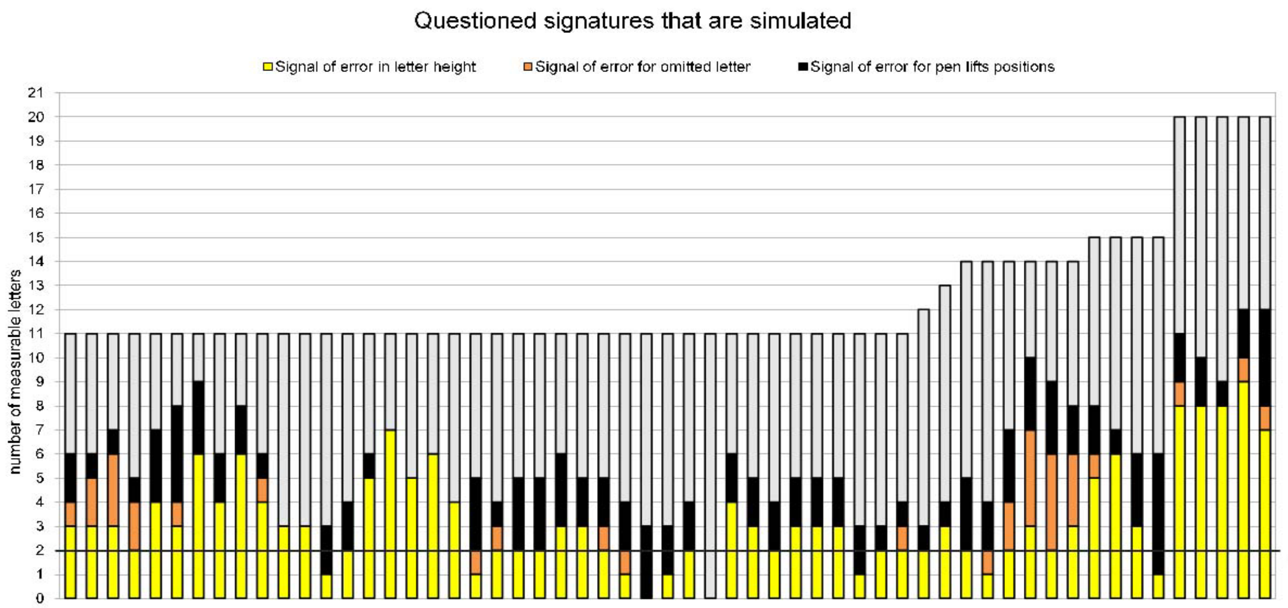

Figure 3: Signals of error in all simulated questioned signatures-1.

In this case too the effective frequencies of error were considered, that is the frequencies refined from the possible misleading signals of error (for that specific letter) recorded in the genuine questioned signatures. The letters that most frequently turned out to trigger off appropriate height error signals are the "t" (49.09%) and "l" (42.80%), followed by the "n" (26.61%) and the "i" (23.16%) (Table 6).

It has to be remembered that in the Italian alphabet, therefore in the signatures here examined, letters like "j", "y", "w" and "x" are usually missing.

Signal to noise ratio for the examined letters

The signal-to-noise ratio is a measure used to compare the level of a desired signal to the level of background noise; it is defined as the ratio of signal power to noise power that is the amount of useful information versus false or irrelevant data. A ratio higher than 1:1 indicates more signal than noise, and the higher the value the more informative the variable under examination [22,23].

In this study the signal to noise ratio is calculated as the ratio of the mean (μ) to the standard deviation (σ) for each letter found in

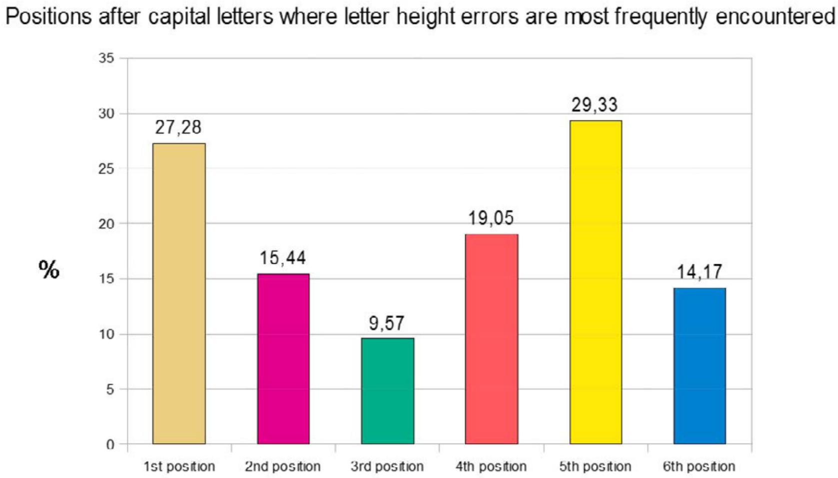

Figure 4: Positions where height errors are more frequently encountered.

	No. of height measures carried out in the specific position within the signature (name + surname)	% of correct signals of height error	Positions in the signature that most frequently trigger off a correct signal of height error
1st position	215	27,28	High frequency
2nd position	218	15,44	
3rd position	217	9,57	
4th position	196	19,05	
5th position	192	29,33	High frequency
6th position	158	14,17	
7th position	74	22,05	Non-reliable percentages due to too few data
8th position	15	10,00	
9th position	8	83,33	
10th position	5	80,00	

Table 4: Percentages of positions having frequent height errors.

	Total possibilities of pen lift errors	No. of pen lift errors	% of pen lift errors	Between letter sites that most frequently trigger off a pen lift error
1st position	55	14	25,45	
2nd position	60	21	35,00	High frequency
3rd position	90	22	24,44	
4th position	52	20	38,46	High frequency
5th position	56	10	17,86	
6th position	67	12	17,91	
7th position	24	6	25,00	Non-reliable percentages due to too few data
8th position	0	0	0,00	
9th position	6	1	16,67	
10th position	5	0	0,00	

Table of pen lift error percentages in false Qs according to the positions of the errors along the signature

Table 5: Positions where pen lift errors are more frequently encountered.

the 200 known exemplars used as comparison signatures; the ratio concisely describes how potentially powerful a specific known letter is to elicit a height error from the imitator, in other words indicating how narrow is the opening for the forger to fit the imitated letter size into the acceptable confidence interval set by the known exemplars for that specific alphabetical character. Basically the signal to noise ratio, with its gradient of values, here can help to test if there are specific letters in the alphabet that can be considered intrinsically more susceptible than others to trigger off height error signals.

The graph in Figure 6 shows the signal to noise ratio for each extender-free small letter encountered in the 200 known exemplars: the signal/noise confidence interval for the small letters ranges from 3.68 to 7.46 - with the letters “i” and “n” having low or average ratios as a consequence of normal (or non-particularly small) standard deviation values.

A definitely higher - actually the highest - signal to noise ratio has been calculated for the letter “t” (=7.11) (Table 7), an outcome that is in agreement with the top percentage of appropriate height error signals (49.09%) scored by that specific grapheme; this result documents the outstanding capacity of the letter “t” to trigger off size errors from the imitators.

Discussion

The evaluation of the overall available data, also in the light of the graphs displayed in Figures 2 and 3, is in substantial agreement with the initial hypothesis - that is the assumption that as the number of the letters composing a signature grows, a significant increase in the number of error signals should be expected in the simulated signatures

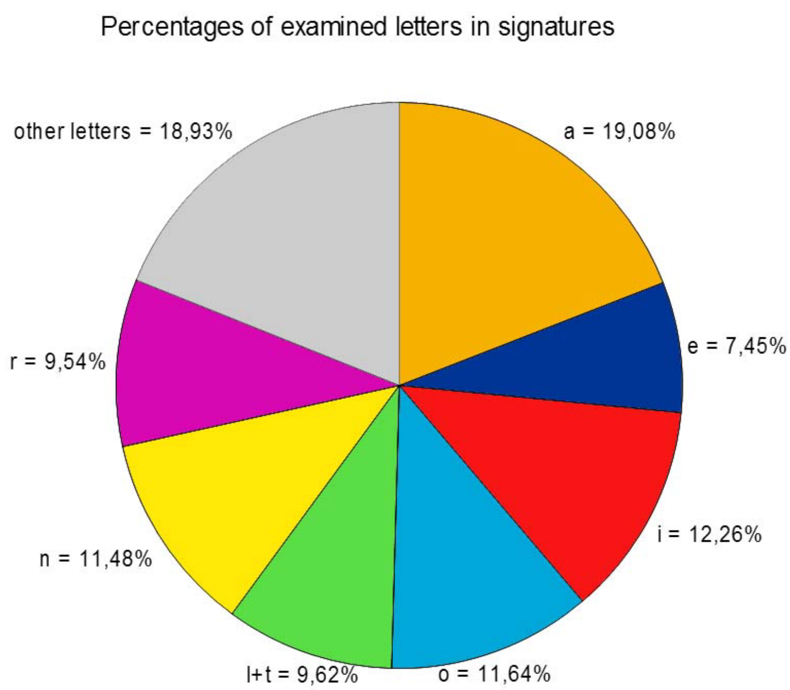

Figure 5: Pie chart with percentages of examined letters in signatures.

	No of letters examined in the questioned group	% on the total of letters examined	% of correct signals of error for letter heights	letters resulting most suited to trigger a correct signal of error
a	246	19,08	16,13	
b	16	1,24	-1,82	
c	50	3,88	15,79	
d	32	2,48	0,00	
e	96	7,45	15,36	
g	25	1,94	20,00	
i	158	12,26	23,16	4th
l	57	4,42	42,80	2nd best
m	5	0,39	0,00	
n	148	11,48	26,61	3rd
o	150	11,64	14,61	
r	123	9,54	0,30	
s	21	1,63	12,50	
t	67	5,20	49,09	Best
U	13	1,01	27,27	
V	17	1,32	8,33	
Z	65	5,04	12,22	
Total	1289	100		

Table 6: Letters most suited to trigger height errors.

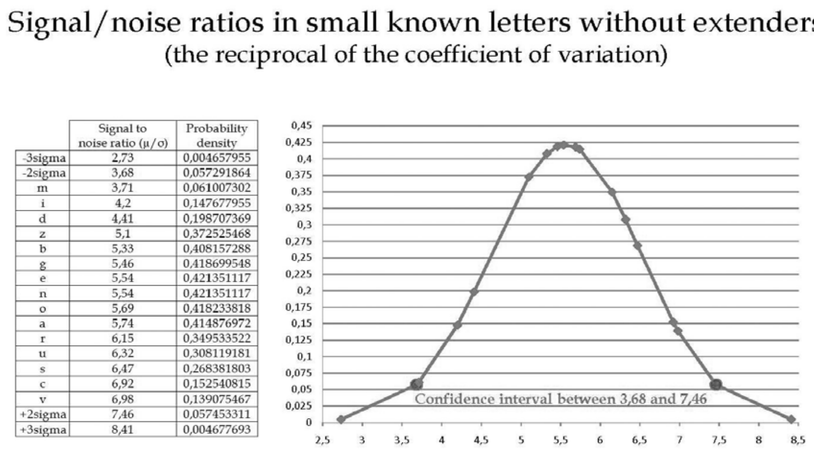

Figure 6: Signal to noise ratio curve for all small letters with no extenders.

	Signal to noise ratio (μ/σ)	% of the specific letter among all Ks	∑ of μ/σ in all K letters	% of correct signals of error for letter heights
a	5,74	21,96	269,78	16,13
b	5,33	1,40	15,99	-1,82
c	6,92	4,21	62.28	15,79
d	4,41	1,40	13,23	0,00
e	5,54	10,75	127,42	15,36
g	5,46	1,40	16,38	20,00
i	4,2	13,55	121,8	23,16
m	3,71	1,40	11,13	0,00
n	5,54	10,28	121,88	26,61
o	5,69	11,68	142,25	14,61
r	6,15	11,21	147,6	0,30
s	6,47	3,27	45,29	12,50
u	6,32	2,34	31,6	27,27
v	6,98	2,34	34,9	8,33
z	5,1	2,80	30,6	12,22
Mean (μ)	5,571			
SD (σ)	0,946			
Confidence interval	3,678			
	7,463			
l	5,89	7,79	111,91	42,80
t	7,11	4,51	78,21	49,09

Signal to noise ratios in all K signatures

Table 7: Signal to noise ratios for all letters in known exemplars.

alone, with no or little change in the genuine signatures' misleading signals. In fact, in the simulated questioned signatures composed of 11 measurable small letters 4.62 error signals were counted on average, whereas only 0.83 signals of error per signature were detected in equal length genuine signatures. As the overall length of the questioned signature grows, this gap definitely tends to widen, so much so that in signatures having 14 measurable small letters the average number of signals of error in the simulated category rises to 7.1, against 0.28 in the genuine equal length subgroup. For even longer imitated signatures a further increase in error signals was recorded, with an average of 9.8 signals in the simulated signatures made of 20 measurable letters, compared to only 1 signal of potential error in the genuine signature made of 18 measurable letters (Figure 2).

Moreover, what was statistically expected in the genuine questioned group as to the frequency of occurrence of the yellow/orange signals of height potential error turned out to be coherent with the actual data; within the 47 genuine questioned signature group occasional data outside the known authentic variability range were recorded: the total number of these misleading signals of potential height errors was 26, with an average per signature of 0.55. Since 580 were the letter height measures carried out in all the genuine questioned signatures, a total of 26 misleading error signals translates into 4.48% (=the percentage of values that turned out to be anomalous - that is outlying - in comparison with the reference curves of the known exemplars and their calculated confidence intervals). This is in line with what was theoretically expected from the concept of confidence limit, a statistical frontier set in this study at ±2σ from the mean that is a significance boundary cutting off tails of "unlikely" occurrences amounting to 4.55% of the population values.

The data concerning the *letters most frequently triggering off height error signals* (Table 6) offer interesting cues for the following considerations: in the first place, it is clear that the imitators meet serious difficulties in reproducing with sufficient precision the size of the letters having upper extenders (letters "l" and "t"): in almost half of the cases, these letters' height values are disproportionate in comparison with the known exemplars' sizes. [As far as the letter "d" is concerned, the available data are not sufficient for a sound statistical evaluation, since that grapheme only rarely appears in the imitated signatures=2.48% of the cases].

According to the data it seems rather complicated for almost any simulator to succeed in making an abduction (upward) movement with just sufficient - but not excessive - dynamic energy so that the written imitated production be kept within the variability range of the genuine letter height; such difficulty in the proportionate reproduction of the (*upper*) extenders is partially confirmed by the 20% error rate recorded for letter "g" (that contains a *lower* extender). Still a considerable difference remains attested, with a higher level of difficulty in the execution of correctly sized upper extenders in comparison with the still complicated - but less challenging - execution of the lower extenders, and this outcome might stem from the fact that the energy to be applied while swaying the pen along the sheet of paper is more easily modulated/controlled in the adduction gestures rather than during an abduction movement. In this respect it would be interesting to evaluate in the future the frequency of error for the two height components of the cursive letter "f", which simultaneously contains both the upper and the lower extender (in this research the letter "f" was unfortunately not present throughout the signatures).

Interesting but lower frequencies of error were also noted in the letters "n" (26.61%) and "i" (23.16%), with a rather high error frequency also for the grapheme "u" (27.27%), for which very few samples were available - so that the calculated percentage is not numerically significant.

As to the letters "n" and "u", the fairly good frequency of height errors (almost one quarter of the cases) can be referred to the special features of these alphabetical characters, made of multiple short downstrokes whose heights must be exactly reproduced by the forger in order to obtain size values in step with the genuine exemplars.

With regards to the frequent occurrence of height error signals in the execution of the letter "i", this outcome most probably stems from the frequent relative miniaturization of that alphabetic character during the execution of genuine signatures, thus from the consequent difficulty for the imitator - if this rhythmic tendency to reduce that character is not physiologically ingrained in his/her own writing style - to control and master a suddenly shortened movement in its longitudinal size.

As a matter of fact, the letter "i" in almost two third of the cases (19 out of 29 occurrences=65.5%) turns out to be either the smallest or the second smallest letter among all characters composing the signatures (Table 8); these are unparalleled data in comparison with any other letter in the questioned signatures, "n" included (which scored a 25% rate of tininess, resulting - in 5 out 20 cases - the smallest or second smallest letter). As shown in Figure 6 and Table 7, the remarkable frequency of size error signals triggered off by the "i" is not related to a superior signal-to-noise ratio for that letter (its ratio has instead a quite unexceptional value of only 4.2), so that the reason for this outstanding incidence has to be looked for elsewhere - most probably, as seen above, in the repeated relative miniaturization of the grapheme "i" within the signatures, a fact that entails a higher level of difficulty for the congruent reproduction of that sudden letter size rhythm change.

The recurrent *height error frequencies* in the 1st and 5th positions (Table 4) corroborate, from a numerical point of view, what empirically observed in the document examination practice: the frequent height error in the fifth position after the capital confirms that in the letter by letter development of the simulation, the imitator inevitably tends to lose control over the multiple graphic aspects that need to be simultaneously reproduced, thus slowly accumulating a concentration overload that in turn generates, in almost one third of the cases under examination, the fatal height error.

As to the high frequency of error in the letters immediately adjacent to the capital (1st position after the capital: 27.28%), once again the fact can be explained from what empirically experienced: the simulator quite naturally tends to concentrate on the most showy graphical aspects - that is on the correct imitation of the capital letter, which is usually the most conspicuous, eye-catching ornate element of the signature. After having spent most energies on the exact duplication of the capital letter shape and structure, it stands to reason that the imitator cannot just find the time to reorganize instantly the focus of his attention on the precise size of the immediately following letter, thus often falling into the trap of misreproducing that longitudinal dimension.

On the other side, the higher *frequencies of pen lift errors* in the 4th and 2nd letter-to-letter transitions after the capital (Table 5) document the complementarity of this indicator with respect to the letter height parameter: the data show in fact that either the imitator manages to focus his attention on the letter height rendition process - thus being more liable to unwarranted stops along the signatures, hence to pen lift errors - or he/she concentrates more on the signature fluency and rhythmic development, minding the correct connections along the imitated signature but inevitably falling into misreproductions of the suitable heights of this or that letter.

Least square linear regressions were calculated on the collected data. In particular, a suitable fit was observed when plotting the total number of errors made by the imitators (that is extenders/small letters ratio errors+letter height errors+pen lift errors) against the number of the measurable small letters contained in the signatures. As Figure

About the letter "i" It shows up in 29 occurrences out of the 20 types of signatures examined: in 13 cases it is the smallest letter of all letters composing the signature; in 6 cases it is the second smallest.
About the letter "n" It shows up in 20 occurrences out of the 20 types of signatures examined: in 3 cases it is the smallest letter of all letters composing the signature; in 2 cases it is the second smallest.

Relative height of the letters "i" and "n" in comparison with the height of the other letters composing the given known signature

Table 8: Relative height of the letters i and n.

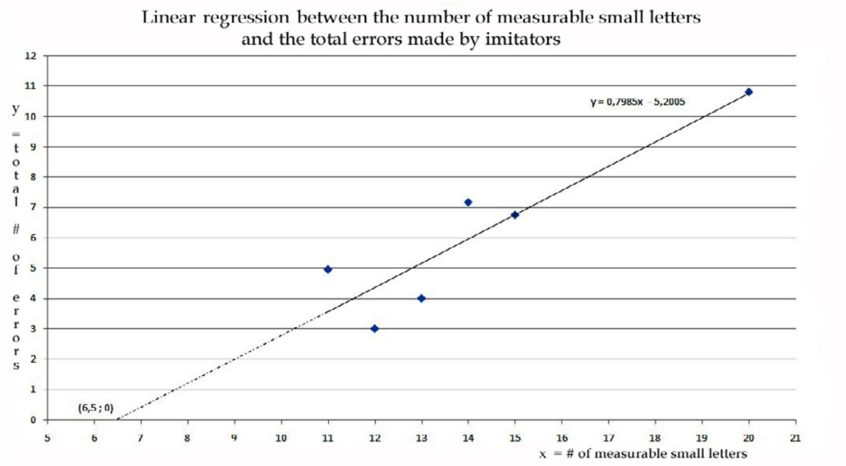

Figure 7: Linear regression between n of small letters and total n of errors by imitators.

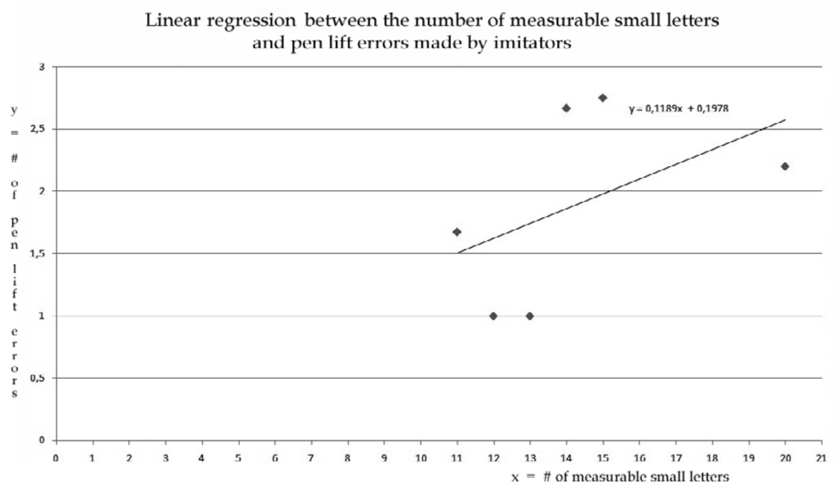

Figure 8: Linear regression between n of small letters and n of only pen lifts errors.

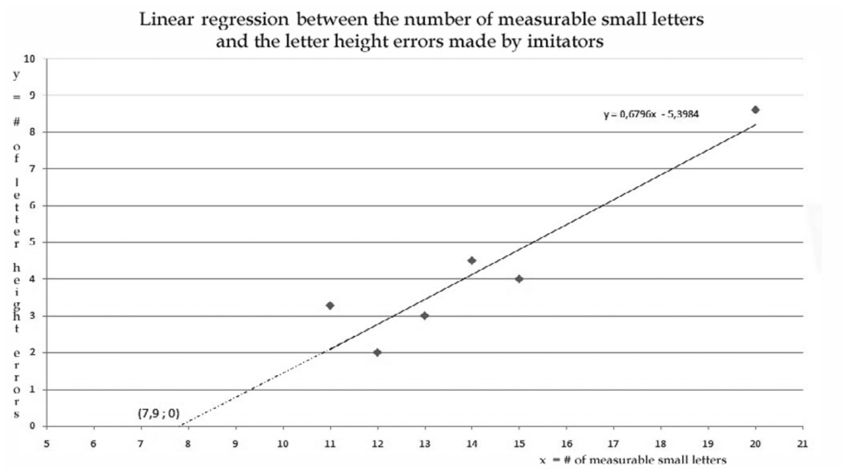

Figure 9: Linear regression between n of small letters and n of only height errors.

7 shows, when signatures contain 10 small measurable letters, on average three errors are expected by the imitators; as signatures get shorter (less letters to be measured), the total number of errors made by the imitators (ratios+letter heights+pen lifts) is expected to dwindle, becoming about two for 9 small letter long signatures and about one error (1.2) when the signature is made of 8 measurable letters.

An adequate linear correlation is also seen in the graph (Figure 9) concerning the increase of only letter height errors with respect to signatures' length. Here for an 11 small letter long signature the expected number of only letter height errors is about two, in projection decreasing to about 1.5 errors when signatures shrink to 10 measurable letters and to slightly less than one when signatures are made of only 9 measurable letters.

Much more erratic data are registered when observing only the imitators' pen lift errors plotted against signatures' length (Figure 8). On average 1.5 pen lift errors are to be expected when signatures are made of 11 measurable letters, but the linear regression coefficient in this particular graph is weak (far from 1), with error data scattered unevenly above and below the line while signature lengths increase.

All in all it is definitely interesting to note (Figures 7 and 9) that when the signature length (in terms of number of measurable small letters) decreases down to 9 and especially to 8 letters, other graphic indicators should be probably recruited and measured in order to expect the attainment of the threshold of three errors per signature which is statistically advisable (with a $\pm 2\sigma$ confidence interval) to make a reliable assessment of non-genuineness for the questioned signature.

Conclusions

The study addressed the subject of the discrimination between genuine and simulated signatures through the evaluation of a few shape-independent indicators, the morphologic approach being sometimes misleading especially when the imitated signature is executed with pictorial skill and naturalness; in particular the focus was set on a restricted number of non-morphologic parameters (ratios between extenders and letter heights, small letter longitudinal sizes, places where between-letter disconnections occur) and these non-shape-related data resulted sufficient to yield correct opinions on 103 out of 104 questioned signatures.

The mutual synergy of these few shape-independent indicators led to a valid method that might prove helpful for the solution of cases concerning signatures of average or above average length (10 or more measurable small letters).

By the data obtained it was possible to determine the positions within the questioned signatures that most frequently trigger off height errors in the imitated letters, as well as the alphabetic characters most frequently fooling the imitator's efforts (this on the ground of the intrinsic difficulties found in the exact reproduction of those letters' genuine sizes).

This line of research has undoubtedly wide margins of refinement ahead and could be useful in supplying new impulse to quantitative studies [24-26] aimed on the one hand at cutting down the margins of error or inconclusiveness in document examination practice, and on the other at developing or standardizing reliable and applicable procedures for the discrimination between genuine and false signatures.

References

1. Wing AM, Nimmo-Smith I (1987) The variability of cursive handwriting measure defined along a continuum: letter specificity. Journal of the Forensic Science Society 27: 297-306.
2. Justino EJR, Bortolozzi F, Saborin R (2002) The Interpersonal and Intrapersonal Variability Influences on Off-Line Signature Verification. Proceedings of 15th Brazilian Symposium on Computer Graphics and Image Processing - IEEE Computer Society 1: 197-201.
3. Srihari SN (2010) Computational Methods for Handwritten Questioned Document Examination. Report submitted to U.S. Department of Justice: 1-46.
4. Found B, Rogers D, Schmittat R (1994) A computer program designed to compare spatial elements of handwriting. Forensic Science International 68: 195-203.
5. Marquis R, Schmittbuhl M, Mazzella WD, Taroni F (2005) Quantification of the shape of handwritten characters: a step to objective discrimination between writers based on the study of the capital character O. Forensic Sci Int 150: 23-32.
6. Marquis R, Taroni F, Bozza S, Schmittbuhl M (2006) Quantitative characterization of morphological polymorphism of handwritten characters loops. Forensic Sci Int 164: 211-220.
7. Marquis R, Taroni F, Bozza S, Schmittbuhl M (2007) Size influence on shape of handwritten characters loops. Forensic Sci Int172: 10-16.
8. Halder-Sinn P, Funsch K (1998) Intra-individual changes in handwriting features depending on handwriting velocity. Journal of Forensic Document Examination 11: 1-16.
9. Found B, Rogers D, Metz H (1999) The objective static analysis of spatial errors in simulations. Journal of Forensic Document Examination 12: 81-99.
10. Nezos R (1994) Judicial Graphology – authentication of handwritten documents. Scriptor Books 3: 25-128.
11. Hanlen R, Manzollilo P, Muehlberger R, Sperry G (1999) Survey of handwriting habit areas used by forensic document examiners: degree of use and discriminatory power. Journal of the American Society of Questioned Document Examiners 2: 45-50.
12. Robertson EW (1991) Fundamentals of Document Examination. Nelson Hall Publishers, 108.
13. Ling S (2002) A Preliminary Investigation into handwriting examination by multiple measurements of letters and spacing. Forensic Sci Int 126: 145-149.
14. Shiffler RE (1988) Maximum Z Scores and Outliers. The American Statistician 42: 79-80.
15. Wieser R (1973) Rhythmus und Polaritaet in der Handschrift - Ein Beitrag zur Rhythmusforschung. Ernst Reinhardt Verlag.
16. Koppenhaver KM (2007) Forensic Document Examination – principles and practice. Humana Press.
17. Ellen D (1997) Scientific Examination of Documents – methods and techniques. Taylor & Francis Group, 47 and 25.
18. Robertson EW (1991) Fundamentals of Document Examination. Nelson Hall Publishers 100-101.
19. Harrison WR (1981) Suspect Documents – their scientific examination. Nelson Hall Publishers 313.
20. Spatz C (2001) Basic statistics: tales of distributions (7thedn.), Wadsworth: Belmont, CA, USA.
21. Evett IW, Totty RN (1985) A Study of the variation in the dimensions of genuine signatures. Journal of the Forensic Science Society 25: 207-215.
22. Bushberg JT, Seibert JA, Leidholdt EM, Boone JM (2006) The essential physics of medical imaging (2ndedn.), Lippincott Williams & Wilkins, 280.
23. Leung SC, Cheng YS, Fung HT, Poon N (1993) Forgery 1 - simulation. Journal of Forensic Sciences, 38: 402-412.
24. Srihari SN, Tomai C, Lee S (2003) Quantitative assessment of handwriting individuality. Proceedings of the fifth International Conference on Advances in Pattern Recognition.
25. Pu D, Srihari SN (2010) A probabilistic measure for signatures verification based on Bayesian learning. International Conference on Pattern Recognition 1188-1191.
26. Al-Musa Alkahtani A, Platt AWG (2010) A statistical study of the relative difficulty of freehand simulation of form, proportion and line quality in Arabic signatures. Sci Justice 50: 72-76.