



A QUANTITATIVE METHOD OF EXAMINATION OF SHORT AND SIMPLE SIGNATURES (INITIALS)

Dmytro MELNYK¹, Alexey SMIRNOV²

¹ *State Scientific Research Forensic Centre of the Ministry of Internal Affairs of Ukraine, Kyiv, Ukraine*

² *Russian Federal Centre of Forensic Examination of the Ministry of Justice, Moscow, Russia*

Abstract

This article is dedicated to the experimental development of an automated quantitative method for differentiating genuine and skilfully simulated short and simple signatures (initials) without conspicuous motor variations. 263 known freehand simulated signatures and 186 samples were processed. The method uses the Bayesian approach to assess the nature of overshoots in the questioned signatures compared to the samples. The decision rules for conclusions were determined, as were the reliability and efficiency of the method.

Key words

Short and simple signatures; Quantitative method; Optimization; Frequency; Significance; Decision function; Decision rule; Reliability and efficiency.

Received 8 May 2014; accepted 30 September 2014

1. Introduction

The increase in flow of documents is leading to a gradual increase in the number of short and simple signatures (initials). It is evident that it is easier to reproduce elements of short and simple models than elements of large and complex signatures. When simulating a signature, the forger is faced with unnatural writing conditions, which may lead to subtle variations in motor features. In the case of continuous training, these motor variations gradually disappear: hence, a smooth writing line can be achieved or one with only local irregularities. Eventually, after long term training, simulated signatures can be achieved with the overall appearance of the model, without conspicuous symptoms of forgery. Such signatures may contain no signs of conscious control over the writing instruments and are thus indistinguishable from natural samples.

For these reasons, it is frequently impossible, particularly in the case of short and simple signatures, to individualize them visually. However, according to

experimental studies of short and simple signatures, even expert penmen could not accurately reproduce more subtle features. Today, examinations supported by statistical analysis of handwriting are (remain) amongst the most effective ways of examining handwriting. Each particular writer has their own statistical values for handwriting variations, which are a basis for individualizing such a writer.

A comprehensive *Methodology* for differentiating genuine from skilfully simulated short and simple signatures (*Methodology*) was developed in the All-union Scientific Research Forensic Institute of the Ministry of Justice of the Union of Soviet Socialist Republic (MJ USSR) in 1987 (Dutova, Li, Orlova, et al., 1987). 43 employees from 25 different Scientific Research Forensic Institutes of MJ USSR were involved in developing the *Methodology*.

That *Methodology* included a number of quantitative methods to research static (structural and geometrical parameters) and dynamic (pen pressure) impressions in signatures.

Mathematical calculations in the *Methodology* were really difficult for handwriting examiners without a special mathematical education. The measurements of signature parameters were executed manually using rulers, protractors and templates, which added to the complexity and hence made it more difficult to use the *Methodology* effectively. The aforementioned difficulties necessitated the development of special software for simplification of the procedure of researching signatures. Only in 2000 in the Russian Federal Centre of Forensic Examination of the Ministry of Justice was the *Methodology* automated on the basis of modern computers and scanners with the use of specially designed software *Oko-1* (Smirnov, 2006).

However, the automated *Methodology* used obsolete empirical data for the following reasons:

1. the measurement procedures had been executed manually up till that time, which meant they were quite different from modern measurement procedures carried out using a computer;
2. the significant social and educational changes which have occurred over the last 25 years may have affected modern signatures;
3. the difference in writing between Russian and Ukrainian.

Hence the main objective of the current research was to develop an automated method of examination of short and simple signatures for Ukraine and to

verify this method on new empirical data. However, it should be noted that the signatures used in the current handwriting examinations were not entirely suitable for verification, since we did not know during the study which signatures were genuine and which were forged. Nevertheless, the results of the experiment can be used in the future for developing a quantitative methodology in the case of multi-object examinations. Such a problem has become common in civil and arbitration proceedings in 2013. Reliable grouping (clustering) of signatures using quantitative methods can significantly improve reliability and/or efficiency of the solution.

2. Materials and methods

2.1 Experiment

Specimens signed by 18 executors and forged signatures by 17 known imitators, including handwriting examiners were studied. Short, simple and highly mature signatures were used, executed in the cursive style. Each imitator individually determined the duration and principles of training (after which they felt ready to reproduce signatures). 52 sets were prepared and processed by *Oko-1* for further statistical analysis. Each set consisted of 5 known forged signatures (Figure 1) and 10–11 specimens (Figure 2).



Fig. 1. Signatures simulated by expert penmen after preliminary training.



Fig. 2. Set of specimens.

Three different forgers imitated the same signature model. Hence, the same specimen sets were used 3 times for comparison. Overall, 263 known forged signatures and 186 specimens were processed (a year later the executors of these specimens executed 5 or more simultaneous signatures for the purpose of developing a multi-object method, but this issue has not been described in the present paper).

The quantitative study of structural and geometrical parameters (integral parameters and curvature) of signatures turned out to be of the most practical interest. These parameters proved their usefulness and ef-

iciency. Hence, 12 types of integral parameters (*IP*) were measured in signatures, such as: the difference in distance between midpoints and the base line (*DDMP*, Figure 3), the difference in orientation (*DO*, Figure 4), the ratio of horizontal to vertical movement length (*H/V*, Figure 5), the ratio of flexor to extensor movement length (*F/E*, Figure 6). These four basic types of parameters were analysed in three sections of the signatures – at the beginning (*B*), in the middle part (*M*) and at the ending (*E*; Dutova, Li, Orlova, et al., 1987) (Figure 7). The procedure of measurement of integral parameters with *Okoi-1* software is shown in Figure 8.

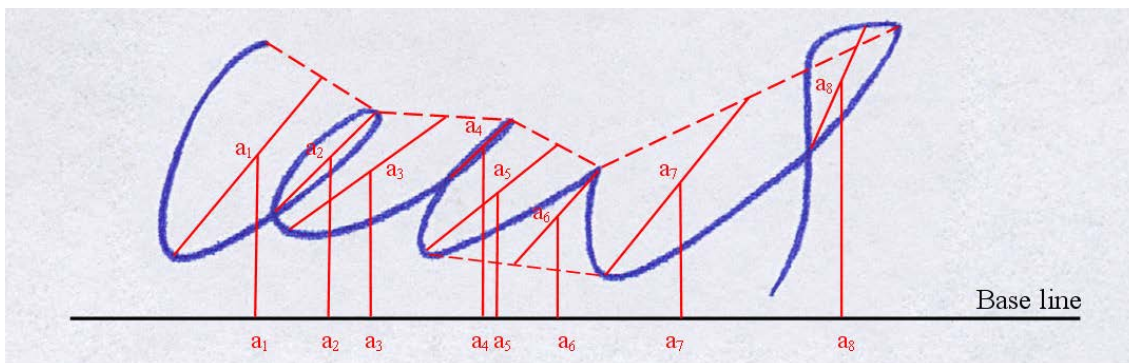


Fig. 3. Difference in distances between midpoints and the base line (*DDMP*).

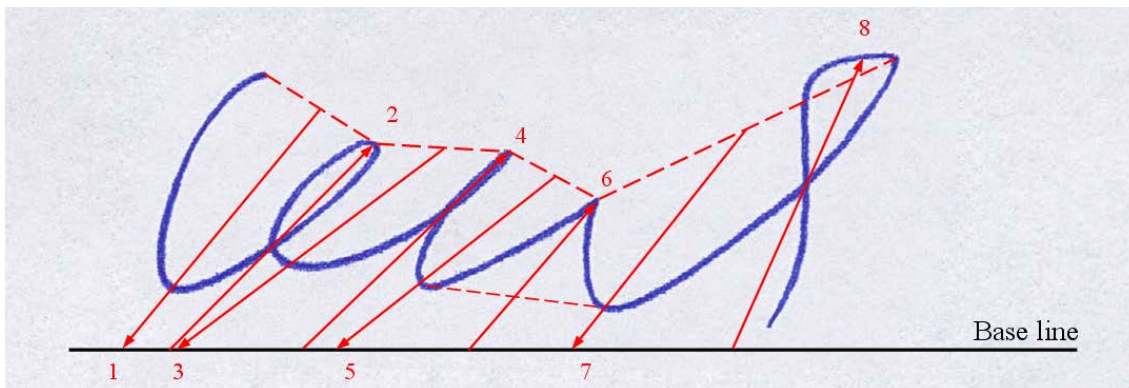


Fig. 4. Difference of orientation (*DO*).

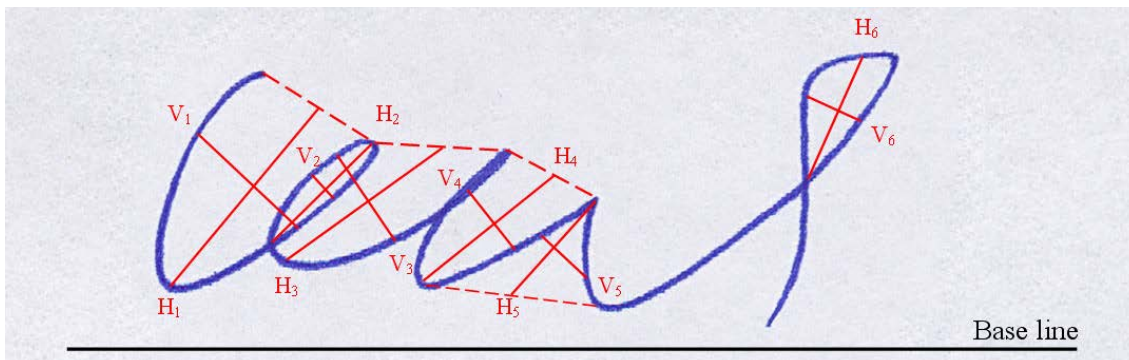


Fig. 5. Horizontal/vertical movement length ratio (*H/V*).

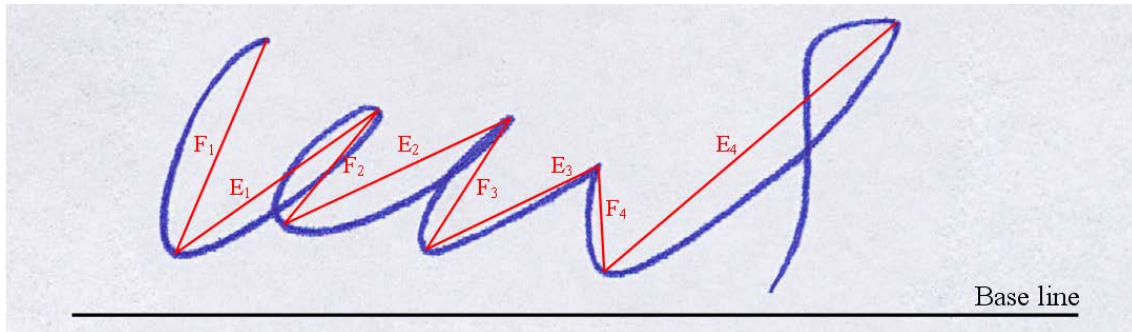


Fig. 6. Flexor/extensor movement length ratio (F/E).

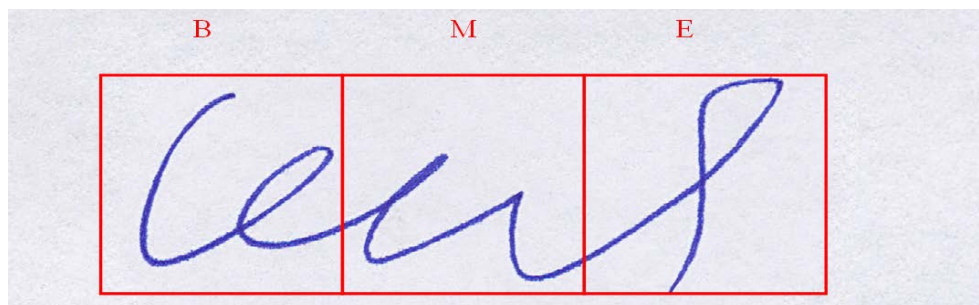


Fig. 7. Beginning (B), middle part (M) and ending (E) of the signature.

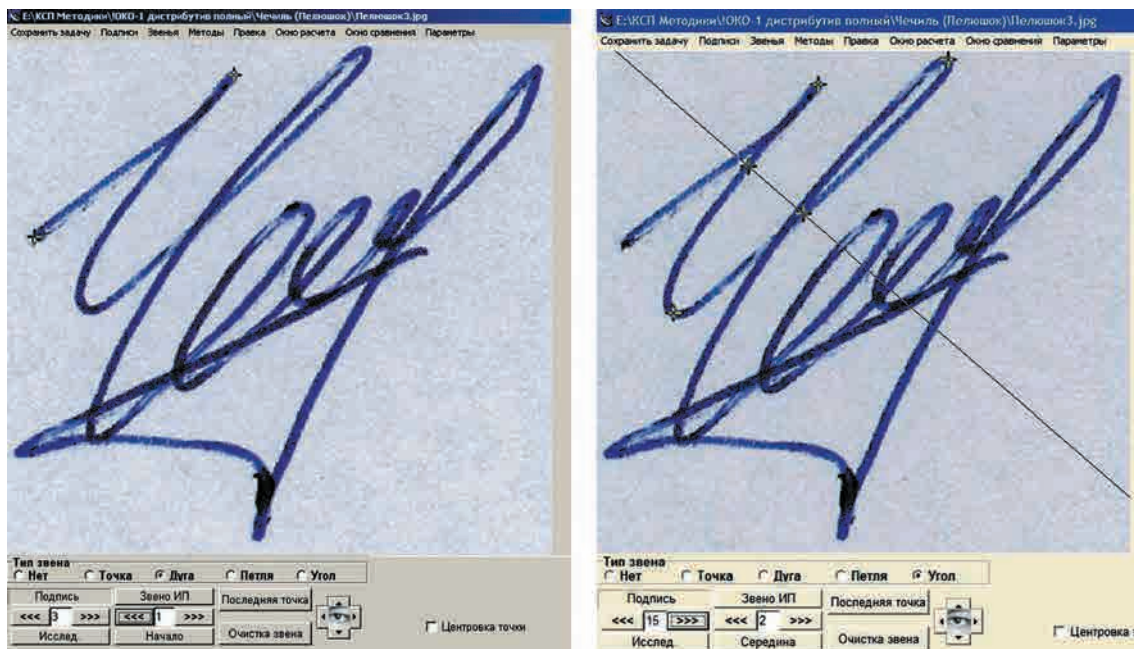


Fig. 8. Measurements of integral parameters with *Oko-1* software.



Fig. 9. FL, EX, LE, UE, C, A.

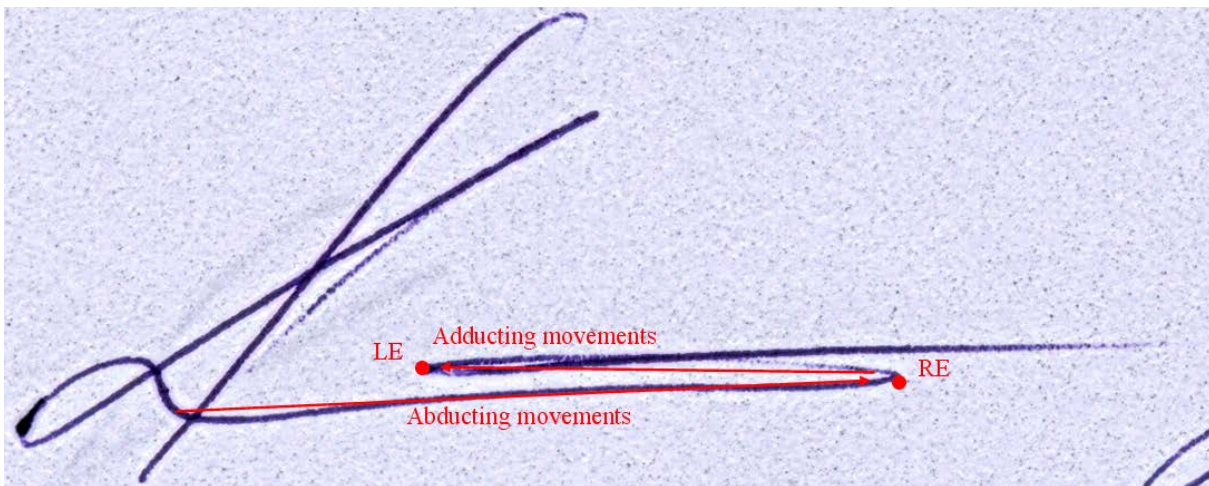


Fig. 10. RLE and AA.

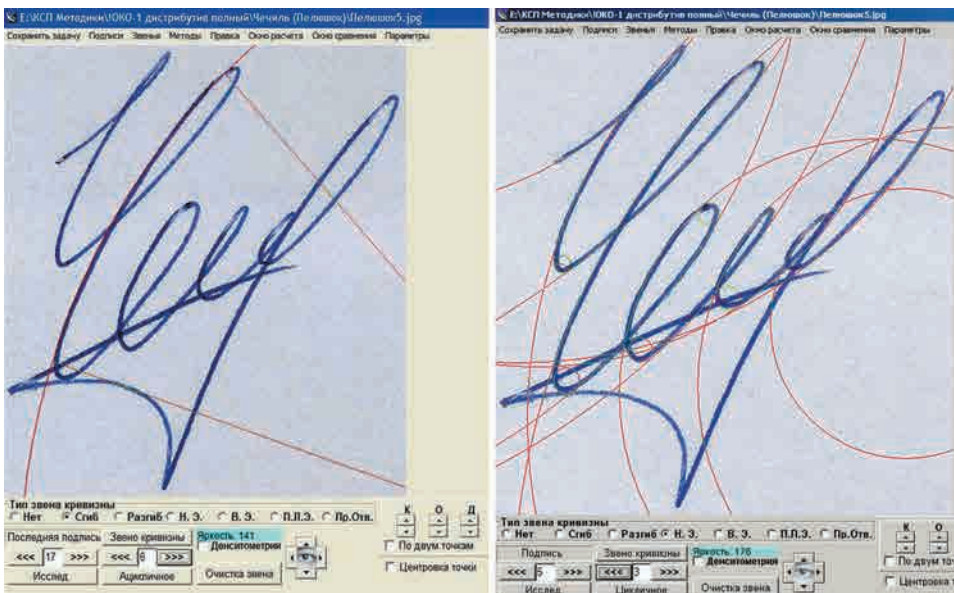


Fig. 11. Measurements of curvature with *Oko-1* software.

For analysis of curvature, we introduced 12 types of links. The measurements encompassed flexor movement (*FL*), extensor movement (*EX*), lower extremes (*LE*), upper extremes (*UE*), right and left extremes (*RLE*), and adducting and abducting movements (*AA*). These six types were measured using cyclic (*C*) and acyclic (*A*) links (Dutova, Li, Orlova, et al., 1987) (Figures 9, 10). The procedure of measurement of curvature with *Oko-1* software is shown in Figure 11.

2.2. Mathematical model

The statistical model for pattern recognition applied in the study is described in Dutova, Li, Orlova, et al. (1987). This model is a development and natural extension of the quantitative identification methods reported in Bohachkina, Vul, Orlova, et al. (1976). The model uses the Bayesian approach, simple additional decision functions (*DF*) and information about the significance of binary features – the existence or absence of parameter overshoots in the investigated signature relative to variability limits in specimens. More precisely, we introduced a generalized parameter relating to the task or the “study situation”: it is the ratio of the deviation of the measured parameter in the investigated signature from the mean value in specimens to the mean deviation of the parameter in specimens. If the absolute value of the parameter is greater than or equal to $B = 2$, we call it “overshoot”. The significance of the overshoot Z_i is described by the equation:

$$Z_i = 50 \lg\left(\frac{P_{1i}}{P_{2i}}\right),$$

where: P_{1i} – is the probability of overshoot of parameter i in a genuine signature (compared with specimens), P_{2i} – is the overshoot probability of parameter i in a forged signature (compared with specimens). The significance of overshoot absence is the value:

$$\bar{Z}_i = 50 \lg\left(\frac{1 - P_{1i}}{1 - P_{2i}}\right).$$

The sum of significances of the occurrence and absence of overshoot in the signature in the case of independence of features (variables) in the signature is estimated using the log-likelihood ratio. Hence this serves as a decision function. Since we used results of comparison between questioned signatures and authentic signatures for recognition of an authentic signature, examination of samples from the training collection in our model reflected a real examining situation. The vector of features of the authentic signature is the first class object. The vector of features of the forged signature is the second class object. This approach al-

lows us to solve the problem of signature examining as a standard task of differentiating two classes in pattern recognition theory.

Such an approach linked with introduction of parameterization and methods of comparison is optimal in the case of statistically independent variables. However, the decision rules of the method were estimated according to the experimental distribution of values of decision functions for the collection of training samples. This allows us to define appropriate limits for the decision rules and to evaluate the reliability and efficiency of the method, regardless of any theoretical assumptions.

This simple signature recognition method was developed for use without advanced computer software. This purpose also defined the choice of the logarithm base and factor. Further investigations showed that the developed method is preferable to other more “refined” approaches by virtue of its robustness. For more details on the statistical sample recognition methods, see: Chernick (1999), Gnedenko (1988), Kulbak (1967) and Zaks (1976).

2.3. Optimization criteria

The value range below the minimum value of the *DF* for the genuine signatures is a zone of error-free negative conclusions for the training collection. The value range above the maximum value of the *DF* for the forged signatures is a zone of error-free positive conclusions. The average of the relative frequencies of these two error-free types of conclusions will be the basis for evaluating the efficiency of the *DF* or method. We have introduced such a measure of efficiency because in forensic examinations, reliable conclusions are of the utmost importance.

2.4. Optimization of data processing

At first sight, the results of the initial calculations performed on the experimental data were paradoxical. The significances (absolute values) of overshoot occurrence or absence for almost all types of parameters were higher than in the *Methodology* (Dutova, Li, Orlova, et al., 1987), but the efficiency of the model based on the accepted value of *DF* in the presented approach was less than in the previous method. Presumably, this is related to the inadequacy of the model rather than the non-optimality of the *DF* for the training sample. In the optimal case, the distribution of probability densities in Figure 12 should have been close to the bell-shaped normal distribution. Obviously, deviation from this normal distribution is observed

foremost for the distribution densities of forged signatures. The significant differences in the quality of imitation in the forged signatures in the sample executed by various penmen are the main reason for this. This phenomenon has been observed in similar research on one-word samples (Orlova, Yakovleva, Smirnov, 2010). Theoretically, if the distribution of the *DF* and *IP* (integral parameters) for genuine signatures is close to normal, the distribution for forged signatures should be asymmetric or even unimodal. In order to improve the correspondence of the sample to the model and enhance the efficiency of the *DF*, the most unsuccessful imitations were removed from the training sample of forged signatures in the phase of calculating significances. Imitations of the *IP* in 21 tasks and the curvature in 22 tasks (in forged signatures) for which average overshoot frequency was higher than 0.4 were excluded from calculations of significances. The choice of this limit was based on maximization of the efficiency. Exclusion of forged signatures with a large number of overshoots from the calculation of significances obviously did not increase the number of incorrect answers after applying the new decision rules. A confirmation of this assumption was the fact that this optimization reduced the absolute values of significances and at the same time increased the efficiency of the developed method (within the limits of statistical error) in relation to the efficiency of the method described by Dutova, Li, Orlova et al. (1987). The accuracy of estimates of complex summary statistics (efficiency and decision errors) was evaluated by bootstrap methods (Chernick, 1999).

3. Results

Tables 1–8 show overshoot statistics of integral parameters.

Table 1
Number of measurements in forged signatures

| Kind of link | Location | | |
|--------------|------------------|---------------|------------|
| | <i>Beginning</i> | <i>Middle</i> | <i>End</i> |
| <i>DDMP</i> | 435 | 565 | 245 |
| <i>DO</i> | 435 | 565 | 245 |
| <i>H/V</i> | 125 | 285 | 50 |
| <i>F/E</i> | 110 | 265 | 50 |

Table 2
Number of measurements in genuine signatures

| Kind of link | Location | | |
|--------------|----------------------|-------------------|----------------|
| | <i>B (beginning)</i> | <i>M (Middle)</i> | <i>E (End)</i> |
| <i>DDMP</i> | 1546 | 2170 | 962 |
| <i>DO</i> | 1546 | 2170 | 962 |
| <i>H/V</i> | 325 | 1026 | 131 |
| <i>F/E</i> | 292 | 930 | 131 |

Table 3
Absolute *IP* overshoot frequency in forged signatures

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | 89 | 136 | 81 |
| <i>DO</i> | 101 | 154 | 85 |
| <i>H/V</i> | 43 | 103 | 17 |
| <i>F/E</i> | 24 | 83 | 14 |

Table 4
Absolute *IP* overshoot frequency in genuine signatures

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | 139 | 184 | 71 |
| <i>DO</i> | 133 | 183 | 87 |
| <i>H/V</i> | 28 | 80 | 15 |
| <i>F/E</i> | 24 | 88 | 9 |

Table 5
Relative *IP* overshoot frequency in forged signatures

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | 0.2059 | 0.2416 | 0.332 |
| <i>DO</i> | 0.2334 | 0.2734 | 0.3482 |
| <i>H/V</i> | 0.3465 | 0.3624 | 0.3462 |
| <i>F/E</i> | 0.2232 | 0.3146 | 0.2885 |

Table 6
Relative IP overshoot frequency in genuine signatures

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | 0.0904 | 0.0852 | 0.0747 |
| <i>DO</i> | 0.0866 | 0.0847 | 0.0913 |
| <i>H/V</i> | 0.0887 | 0.0788 | 0.1203 |
| <i>F/E</i> | 0.085 | 0.0955 | 0.0752 |

Table 7
Significance of IP overshoot

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | -17.9 | -22.6 | -32.4 |
| <i>DO</i> | -21.5 | -25.4 | -29.1 |
| <i>H/V</i> | -29.6 | -33.1 | -23.0 |
| <i>F/E</i> | -21.0 | -25.9 | -29.2 |

Table 8
Significance of IP overshoot absence

| Kind of link | Location | | |
|--------------|----------|----------|----------|
| | <i>B</i> | <i>M</i> | <i>E</i> |
| <i>DDMP</i> | 2.9 | 4.1 | 7.1 |
| <i>DO</i> | 3.8 | 5.0 | 7.2 |
| <i>H/V</i> | 7.2 | 8.0 | 6.4 |
| <i>F/E</i> | 3.6 | 6.0 | 5.7 |

An analysis of significance shows that diligent imitators of IP copied signatures better at the beginning part than in the middle, and in the middle better than at the end. Tables 9–16 show the overshoot statistics for curvature.

Table 9
Number of measurements in forged signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 190 | 350 | 260 | 250 | 110 | 90 |
| <i>C</i> | 40 | 115 | 120 | 65 | 0 | 0 |

Table 10
Number of measurements in genuine signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 628 | 1225 | 982 | 846 | 464 | 370 |
| <i>C</i> | 374 | 583 | 514 | 313 | 0 | 0 |

Table 11
Absolute curvature overshoot frequency in forged signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 91 | 90 | 75 | 70 | 34 | 39 |
| <i>C</i> | 5 | 31 | 29 | 25 | 0 | 0 |

Table 12
Absolute curvature overshoot frequency in genuine signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 49 | 90 | 76 | 56 | 36 | 34 |
| <i>C</i> | 23 | 45 | 35 | 15 | 0 | 0 |

Table 13
Relative curvature overshoot frequency in forged signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 0.4792 | 0.2585 | 0.2901 | 0.2817 | 0.3125 | 0.4348 |
| <i>C</i> | 0.1429 | 0.2735 | 0.2459 | 0.3881 | 0 | 0 |

Table 14
Relative curvature overshoot frequency in genuine signatures

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 0.0794 | 0.0742 | 0.0783 | 0.0672 | 0.0794 | 0.0941 |
| <i>C</i> | 0.0638 | 0.0786 | 0.0698 | 0.0508 | 0 | 0 |

Table 15
Significance of curvature overshoot

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | -39 | -27.1 | -28.5 | -31.1 | -29.8 | -33.2 |
| <i>C</i> | -17.5 | -27.1 | -27.4 | -44.2 | 0 | 0 |

Table 16
Significance of curvature overshoot absence

| Type of link | <i>FL</i> | <i>EX</i> | <i>LE</i> | <i>UE</i> | <i>RLE</i> | <i>AA</i> |
|--------------|-----------|-----------|-----------|-----------|------------|-----------|
| <i>A</i> | 12.4 | 4.8 | 5.7 | 5.7 | 6.3 | 10.2 |
| <i>C</i> | 1.9 | 5.2 | 4.6 | 9.5 | 0 | 0 |

On the basis of significances obtained for genuine and forged signature samples, additional decision functions were calculated for *IP*, curvature and for these two factors combined. Histograms of relative frequencies of the *DF* (densities of probability distributions) are shown in Figure 12. The frequencies for forged signature samples are shown in black and for genuine signature samples in white.

The histograms show that we can build decision rules with a rather high percentage of reliable solutions primarily for negative conclusions.

The decision rules for conclusions in the form: definite positive, probable positive, definite negative, and probable negative, were drawn up on the basis of *DF* distributions and are shown in Tables 17–19.

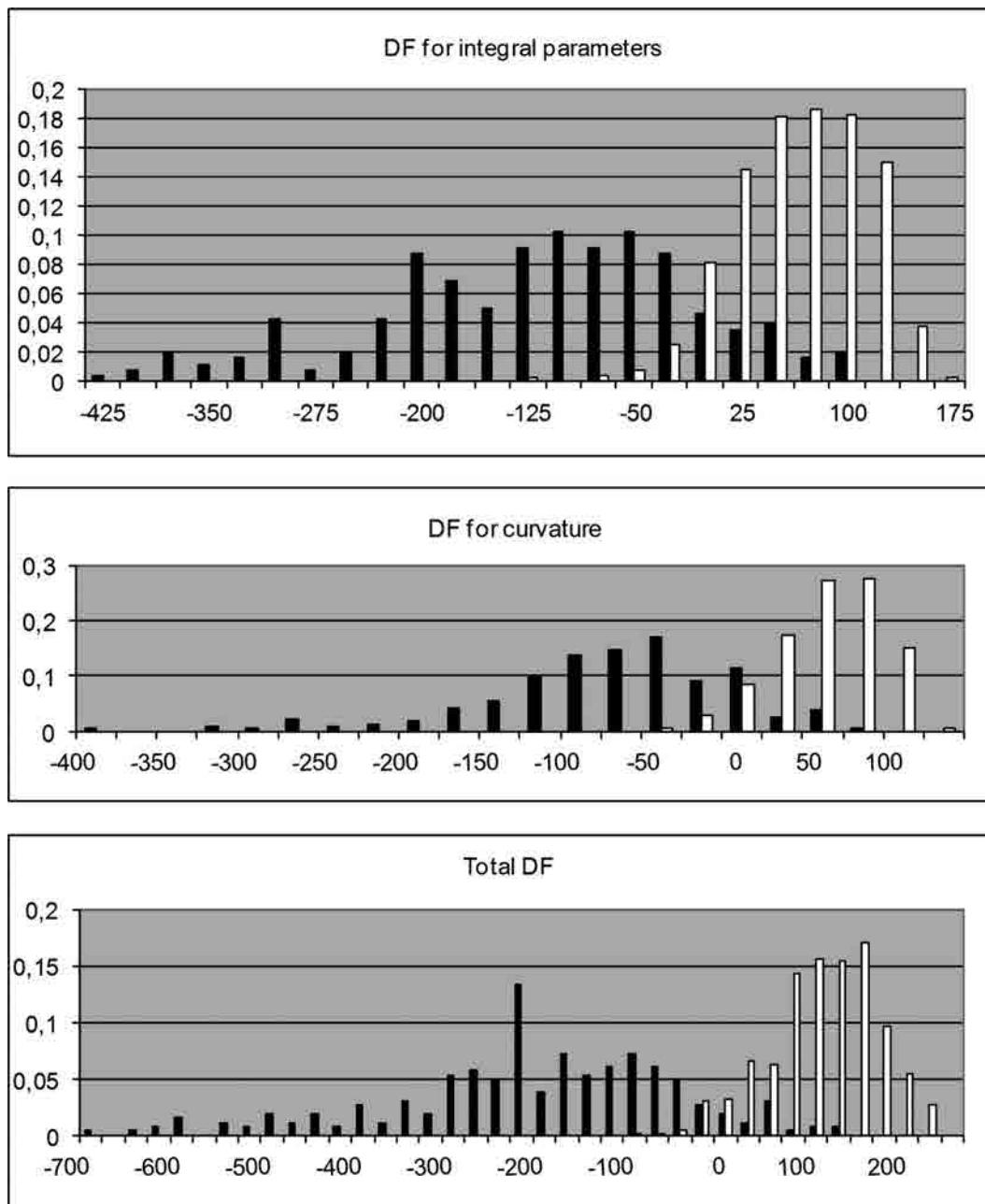


Fig. 12. Histograms of relative frequencies for *DFs*.

Table 17
Decision rule for integral parameters

| Total significance | Cell frequency for forged signatures | Cell frequency for genuine signatures | Reliability of decision | Conclusion |
|---------------------|--------------------------------------|---------------------------------------|-------------------------|-------------------|
| $S < -100$ | 0.567 | 0.00176 | 99% | Definite negative |
| $-100 \leq S < -65$ | 0.141 | 0.00879 | 90% | Probable negative |
| $70 < S \leq 100$ | 0.0190 | 0.206 | 90% | Probable positive |
| $100 < S$ | 0 | 0.188 | 99% | Definite positive |

Table 18
Decision rule for curvature

| Total significance | Cell frequency for forged signatures | Cell frequency for forged signatures | Reliability of decision | Conclusion |
|---------------------|--------------------------------------|--------------------------------------|-------------------------|-------------------|
| $S < -100$ | 0.407 | 0 | 99% | Definite negative |
| $-100 \leq S < -65$ | 0.198 | 0.00352 | 95% | Probable negative |
| $65 < S \leq 100$ | 0.00380 | 0.267 | 95% | Probable positive |
| $100 < S$ | 0 | 0.00703 | 99% | Definite positive |

Table 19
Total (combined) decision rule

| Total significance | Cell frequency for forged signatures | Cell frequency for forged signatures | Reliability of decision | Conclusion |
|----------------------|--------------------------------------|--------------------------------------|-------------------------|-------------------|
| $S < -140$ | 0.673 | 0 | 99% | Definite negative |
| $-140 \leq S < -100$ | 0.110 | 0.00176 | 95% | Probable negative |
| $100 < S \leq 140$ | 0.00761 | 0.253 | 95% | Probable positive |
| $140 < S$ | 0 | 0.250 | 99% | Definite positive |

To solve the problem (of whether an analysed signature is authentic/forged) an examiner should use the total (combined) decision rule. The results obtained for decision rules calculated separately for integral parameters and curvature only show what part of the signature is the most changeable. The *IP* decision rule can be used if the examiner cannot find at least 10 curvature links to measure in a signature. The new method has been automated using specially designed software: *Oko-1 Ukr*.

4. Discussion

Tables 7 and 8 show that the absolute values of significances of *IP* increase from the beginning to the end of the signature. An exception is the parameter *H/V* at the end of the signature. However, Tables 1 and 2 show that the number of measurements of this param-

eter and therefore the accuracy of estimation of the significance is considerably lower than for the other parameters. This means that pseudo-closed links with these parameters are rare at the end of the signature. A similar trend was observed for significances of *IP* in the method described by Dutova, Li, Orlova, et al. (1987). The significance of *DDMP* at the beginning of the signature in that experiment was rounded down to zero in view of the rather small value. In the present experiment, the absolute value of significance of this parameter at the beginning of the signature was also smaller than that of all the others. It is interesting that according to the experiment: Orlova, Yakovleva, Smirnov (2010), significances of *IP* (limited data) are not linked with location. We can explain this by the following hypothesis. The imitator perceives and imitates a very short and simple signature – with non-letter strokes, monograms and conventionally readable letters – as a whole unit. Therefore, in training, the

imitator achieves more success in imitating the first part and less in imitating the other parts. But in the case of analysis of legible individual words, the imitator distinctly divides the word into letters, and then the imitator tries to reproduce each of the letters separately regardless of its position in the word.

According to Tables 15 and 16, significances of curvature for the present experiment and that of Dutova, Li, Orlova, et al. (1987) and for *RLE* and *AA* coincide well. For other parameters of curvature, it is difficult to establish any conformity, except for the average value. In both experiments, the average significances of the curvature parameter are slightly higher than for *IP*. The reasons for better conformity of *IP* in the experiments may be twofold – measurement of curvature is a more difficult and at the same time more subjective process. In the first experiment, as we said, all measurements were performed manually; in the second, almost all measurements were carried out by one person, with the help of computer software.

The effectiveness of both methods (in the present study) was 63%. The effectiveness of the methods was also calculated for different thresholds of overshoot: $B = 1.75$ and $B = 2.25$. In both cases, the efficiency was 59%. Thus the threshold of overshoot $B = 2$ was chosen as close to optimum, but this change will not tangibly increase efficiency.

5. Conclusions

The performed experiment verified the method described by Dutova, Li, Orlova, et al. (1987) successfully, and its localized version can be introduced in examiner institutions of the Interior Ministry of Ukraine.

Some increase in efficiency of the method could be obtained in the case of dividing non-negative half-interval into three half-intervals optimally and using three significances instead of two for each type of the parameter. For this, we need a larger experiment. In addition, the representativeness of the signature sample and the balance of experimental conditions are important factors. Clearly, eliminating more than a third of measured signatures at the calculation stage is not a good solution. We should select a group of more motivated imitators with flexible skills. On the other hand, if imitators with professional knowledge and flexible skills are too well motivated, we may obtain a method with low effectiveness, which does not correspond to the reality encountered by forensic examiners on a day-to-day basis.

Another possibility is to develop a multi-object method, enabling analysis of several signatures simultaneously. However, this requires the development and selection of the best mathematical model for such experimental work.

References

1. Bohachkina, H. F., Vul, S. M., Orlova, V. F., et al. (1976). *Primenenie metodov issledovaniya, osnovannykh na veroyatnom modelirovanii v sudebno-pocherkovedcheskoy ekspertize*. Moskva: Metodicheskoe posobie.
2. Chernick, M. R. (1999). *Bootstrap methods, a practitioner's guide*. Hoboken: Wiley Interscience.
3. Dutova, N. V., Li, L. E., Orlova, V. F., et al. (1987). *Kompleksnaya metodika ustanovleniya podlinnosti (nepodlinnosti) kratkih i prostykh podpisov*. Moskva: VNIIE.
4. Gnedenko, B. V. (1988). *Kurs teorii veroyatnostey*. Moskva: Nauka.
5. Kulbak, S. (1967). *Teoriya informatsii i statistika*. Moskva: Nauka.
6. Orlova, V. F., Yakovleva, E. V., Smirnov, A. V. (2010). Metodika issledovaniya odnoslovnnykh kratkih zapisey. *Teoriya i praktika sudebnoy ekspertizy*, 2, 151–159.
7. Smirnov, A. V. (2006). Programma *Oko-1* dlya issledovaniya kratkih i prostykh pocherkovykh objectov. *Teoriya i praktika sudebnoy ekspertizy. Nauchno-prakticheskiy zhurnal*, 1, 121–126.
8. Tu, J., Gonsales, R. (1978). *Printsipy raspoznavaniya obrazov*. Moskva: Mir.
9. Zaks, L. (1976). *Statisticheskoe otsenivanie*. Moskva: Statistika.

Corresponding author

Dr. Dmyto Melnyk
The State Scientific Research Forensic Centre
of the Ministry of Internal Affairs of Ukraine
4, Velyka Okruzhna str.
03680 Kyiv, Ukraine
e-mail: meln.dmitryi@gmail.com

METODA ILOŚCIOWEJ ANALIZY PARAF

1. Wprowadzenie

Zwiększony przepływ dokumentów sprawia, iż podpisujący się coraz częściej parafują dokumenty. Jednak krótkie i proste modele parafy są zdecydowanie łatwiejsze do sfalszowania niż pełne, skomplikowane podpisy. W trakcie podrabiania parafy fałszerz początkowo konfrontuje się z nienaturalnymi dla niego warunkami pisanja, które mogą powodować powstanie subtelnych zmian w cechach motorycznych pisma. W wyniku długotrwałego uczenia się podrabiania parafy zmiany te stopniowo zanikają – następuje wygładzenie linii pisma, a ewentualne jej nieregularności obserwowane są jedynie miejscami. Długotrwały trening pozwala odtworzyć ogólny wygląd parafy tak, iż nie zdradza ona żadnych symptomów fałszerstwa. Sfalszowane w ten sposób parafy mogą nie wykazywać znaków świadomej kontroli nad narzędziem pisarskim i dlatego też stają się nierozróżnialne od paraf autentycznych.

Wyżej wymienione powody sprawiają, iż wizualne zindywidualizowanie, zwłaszcza bardzo prostej i krótkiej parafy, staje się niemożliwe. Niemniej jednak badania paraf dowodzą, iż nawet eksperci z dziedziny pismoznawstwa próbujący sfalszować autentyczną parafę mogą nie być w stanie dokładnie odtworzyć bardziej subtelnych cech pisma. Dziś jedną z najbardziej skutecznych sposobów rozwoju badań pisma są badania wsparte analizą statystyczną. Okazuje się bowiem, iż każdą osobę charakteryzuje pewna stała statystyczna wartość zmienności pisma, która jest jej cechą indywidualną.

Wszechstronna metodyka odróżniania autentycznych i umiejętnie podrobionych paraf nazwana *Metodologią* została opracowana przez naukowy Instytut Badań Kryminalistycznych Ministerstwa Sprawiedliwości Związku Socjalistycznych Republik radzieckich (MJ USSR) w 1987 roku (Dutova, Li, Orlova, i in., 1987). W rozwój *Metodologii* zaangażowanych było 43 pracowników z 25 różnych instytutów kryminalistycznych MJ USSR. W programie wykorzystano kilka ilościowych metod służących do analizy statycznych (strukturalnych i geometrycznych) oraz dynamicznych (nacisk narzędzia pisarskiego) cech parafy.

Matematyczna obróbka danych w programie *Metodologia* była jednak wyjątkowo trudna, zwłaszcza dla biegłych z zakresu badania pisma, którzy nie posiadali wiedzy z zakresu matematyki. Pomiary parametrów parafy były wykonywane ręcznie za pomocą linijek, kątomierzy i szablonów, co również komplikowało wykorzystanie programu.

Wymienione utrudnienia spowodowały konieczność opracowania oprogramowania, które uprościłoby

badanie paraf. W 2000 roku w Rosyjskim Federalnym Centrum Analiz Kryminalistycznych Ministerstwa Sprawiedliwości program *Metodologia* zautomatyzowano, wprowadzając nowoczesne komputery i skanery oraz specjalnie zaprojektowane w tym celu oprogramowanie *Oko-1* (Smirmov, 2006).

Niemniej jednak zautomatyzowany program *Metodologia* wykorzystywał przestarzałe dane empiryczne z następujących powodów:

1. dotychczasowe procedury pomiarów były wykonywane manualnie, co czyniło je całkowicie różnymi od nowoczesnej procedury pomiaru z wykorzystaniem komputera;
2. znaczące socjalne i edukacyjne zmiany, które zaszły w ciągu ostatnich 25 lat, mogły wpłynąć na unowocześnienie paraf;
3. występujących różnic w pisowni pomiędzy Rosjanami a Ukraincami.

Dlatego też głównym celem projektu był rozwój automatycznej metody badania paraf ludności ukraińskiej oraz sprawdzenie wspomnianej metody na nowych danych empirycznych. Należy zaznaczyć na wstępie, że parafy badane podczas tego eksperymentu nie były w pełni odpowiednie do weryfikacji proponowanej metody, gdyż autorzy artykułu nie wiedzieli podczas badania, która z badanych paraf jest autentyczna, a która sfalszowana. Niemniej jednak wyniki przeprowadzonych eksperymentów mogą posłużyć w przyszłości do rozwoju obiektywnej metody opartej na obliczaniu stosownych parametrów statystycznych do badania paraf w przypadku badań wielu paraf jednocześnie. W 2013 roku problem ten często pojawiał się w postępowaniu cywilnym czy arbitrażowym. W takim przypadku możliwość np. pogrupowania paraf z wykorzystaniem stosownych metod statystycznych pozwoliłoby znacząco przyspieszyć proces analizy ich autentyczności.

2. Materiały i metody

2.1 Opis eksperymentu

Przedmiotem badań były autentyczne parafy nakreślone przez 18 osób oraz parafy sfalszowane wykonane przez 17 „fałszerzy naśladowców”, w tym biegłych z zakresu badania pisma. W badaniach wykorzystano krótkie, proste i dojrzałe modele paraf kreślone stylem pochyłym. Każdy z „fałszerzy” sam określał czas i zasady treningu, po którym był gotów odwzorować parafę. Do obliczeń statystycznych paraf wykorzystano oprogramowanie *Oko-1*, w którym przygotowano i przetworzono 52 zesta-

wy danych. Każdy zestaw składał się z 5 sfalszowanych paraf (rysunek 1) oraz 10–11 paraf autentycznych (rysunek 2). Ponieważ trzech różnych fałszerzy odtwarzało ten sam model parafy, to podczas porównania identyczny zestaw paraf autentycznych był wykorzystywany trzykrotnie. Generalnie w programie *Oko-1* przetworzono 263 sfalszowane parafy oraz 186 paraf autentycznych. Rok później każdy z „fałszerzy” kreślił 5 lub więcej paraf jednocześnie w celu rozwinięcia metody pozwalającej na porównanie wielu obiektów, ale w niniejszym artykule problem ten nie został opisany.

W badaniach mierzono strukturalne i geometryczne parametry parafy (parametry integralne i krzywiznę). Uznano, iż te cechy pisma pozwolą ocenić użyteczność i skuteczność metody.

Dlatego też do analizy wykorzystano 12 typów parametrów integralnych (*IP*), takich jak: różnicę w odległości pomiędzy środkiem odcinka a linią podstawową pisma (*DDMP*, rysunek 3), różnicę w położeniu parafy w stosunku do linii podstawowej (*DO*, rysunek 4), proporcję długości ruchu w kierunku poziomym do ruchu w kierunku pionowym (*H/V*, rysunek 5), proporcje długości ruchu zstępującego do długości ruchu wstępującego (*F/E*, rysunek 6). Te cztery podstawowe typy parametrów analizowano dla trzech fragmentów parafy – początku (*B*), środka (*M*) i końca parafy (*E*; Dutova, Li, Orlova, i in., 1987; rysunek 7). Procedurę pomiaru integralnych parametrów w programie *Oko-1* pokazano na rysunku 8.

Do analizy krzywizny wprowadzono 12 typów połączeń. Pomiar obejmowały ruch zstępujący (*FL*), ruch wstępujący (*EX*), dolne ekstrema (*LE*), górne ekstrema (*UE*), połączenia między prawym i lewym ekstrema (*RLE*) oraz ruch przywodzący i odwodzący (*AA*). Te sześć parametrów mierzono za pomocą cyklicznych (*C*) i acyklicznych połączeń (*A*; Dutova, Li, Orlova i in., 1987; rysunek 9, 10). Procedura pomiaru krzywizn wykonywanych w oprogramowaniu *Oko-1* pokazano na rysunku 11.

2.2 Model matematyczny

Zastosowany w badaniach model statystyczny rozpoznawania obrazów opisano w Dutova, Li, Orlova, i in. (1987). Model ten stanowi rozwinięcie ilościowych metod identyfikacji pisma opisanych przez Bohachkina, Vul, Orlovą i in. (1976). Nowszy model wykorzystuje podejście bayesowskie, jak też dodatkowo funkcje decyzyjne (*DF*) oraz informacje o istotności cech binarnych, tj. występowanie lub brak w mierzonej parafie wartości tej cechy (sygnału), która przekracza względną granicę zmienności danego parametru w próbkach. Dokładniej mówiąc, wprowadzono parametr będący stosunkiem różnicy wartości mierzonego parametru w analizowanej parafie od średniej jego wartości w innych parafach do średniego odchylenia tego parametru w parafach. War-

tość bezwzględna tego parametru większa bądź równa $B = 2$ jest nazywana sygnałem przekroczonym. Istotność sygnału przekroczonego opisuje wartość Z_i wyznaczona ze wzoru:

$$Z_i = 50 \lg\left(\frac{P_{1i}}{P_{2i}}\right),$$

gdzie P_{1i} oznacza prawdopodobieństwo wystąpienia sygnału przekroczonego dla parametru i oszacowane w wyniku badań autentycznej parafy, a P_{2i} oznacza prawdopodobieństwo wystąpienia sygnału przekroczonego dla parametru i gdy parafa jest sfalszowana.

Istotność, dla której uważa się iż sygnał przeregulowany nie występuje określa wartość:

$$\bar{Z}_i = 50 \lg\left(\frac{1 - P_{1i}}{1 - P_{2i}}\right).$$

Suma istotności, dla której obserwuje się sygnał przekroczony lub braku tego sygnału w parafie w przypadku zmiennych niezależnych, jest oszacowywana za pomocą funkcji decyzyjnej zdefiniowanej jako logarytm z ilorazu wiarygodności. Ponieważ do rozpoznawania parafy autentycznej wykorzystywane są wyniki porównania parafy kwestionowanej i parafy autentycznej, to próbki ze zbioru uczącego w przedstawionym modelu odzwierciedlały rzeczywistą sytuację badania paraf. Takie podejście pozwala odpowiedzieć na pytanie, czy mamy do czynienia z parafą autentyczną, czy sfalszowaną i jest typowym zadaniem rozróżniania dwóch klas próbek w teorii rozpoznawania wzorów. Obiektem pierwszej klasy jest wektor cech autentycznej parafy. Obiektem drugiej klasy jest wektor cech parafy sfalszowanej.

Należy zaznaczyć, że takie założenie związane z wprowadzeniem parametryzacji i zastosowaniem metody porównania paraf, jest właściwe w przypadku statystycznie niezależnych zmiennych. Niemniej jednak funkcje decyzyjne metody zostały oszacowane na podstawie eksperymentalnego rozkładu wartości funkcji decyzyjnych dla próbek zbioru uczącego. Pozwala to na zdefiniowanie odpowiednich granic reguł decyzyjnych oraz oszacowanie wiarygodności i skuteczności metody, bez względu na teoretyczne założenia.

Ta prosta metoda rozpoznawania paraf oparta o obliczanie i interpretację logarytmów stosowanych parametrów została opracowana po to, aby móc jej używać bez pomocy zaawansowanych obliczeń komputerowych. Dalsze badania pokazały jednak, że jest preferowana bardziej niż inne, bardziej skomplikowane podejścia ze względu na jej wiarygodność. Więcej szczegółów dotyczących statystycznych metod rozpoznawania próbek można odnaleźć w: Chernick (1999), Gnedenko (1988), Kulbak (1967) i Zaks (1976).

2.3 Kryteria optymalizacji

Wartości poniżej minimalnej wartości DF uzyskane dla autentycznych paraf przyczyniają się do tego, że odpowiedzi fałszywie negatywne nie występują w przypadku analiz paraf ze zbioru uczącego. Natomiast wartości powyżej maksymalnej wartości DF uzyskane dla autentycznych paraf powodują sytuację, że odpowiedzi fałszywie pozytywne nie występują w przypadku analiz paraf ze zbioru uczącego. Częstość wystąpienia tych błędów jest podstawą oceny efektywności modelu opartego na przyjętej wartości DF . Taki parametr oceny efektywności metody jest niezwykle istotny, gdyż w badaniach prowadzonych do celów sądowych przedstawienie miarodajnych wniosków stanowi sprawę nadrzędną.

2.4 Optymalizacja procesu przetwarzania danych

Na pierwszy rzut oka wynik początkowych obliczeń wykonanych dla danych eksperymentalnych był paradoksalny. Bezwzględne wartości, dla których obserwowano sygnał przekroczony lub jego brak, dla prawie wszystkich typów parametrów były większe niż w *Metodologii* (Dutova, Li, Orlova i in., 1987). Z drugiej jednak strony efektywność modelu opartego na przyjętej wartości DF w prezentowanym podejściu była mniejsza niż w poprzedniej metodzie. Prawdopodobnie efekt ten związany jest raczej z niedoskonałością modelu niż z nieodpowiednio dobranym parametrem DF dla próbek ze zbioru uczącego. W sytuacji optymalnej rozkład gęstości prawdopodobieństwa (rysunek 12) powinien być zbliżony do rozkładu normalnego. Odstępstwo od tego rozkładu gęstości prawdopodobieństwa zaobserwowano przede wszystkim dla sfalszowanych paraf. Głównym powodem mogą być tu znaczące różnice w jakości podrobionych paraf, które wynikają z faktu, iż były wykonane przez kilku fałszerzy. Zjawisko to zaobserwowano również w trakcie prowadzenia podobnych badań związanych z analizą jednowyrazowych zapisów (Orlova, Yakovleva, Smirnov, 2010). Teoretycznie, jeśli rozkład DF i IP dla autentycznych paraf był bliski rozkładowi normalnemu, to rozkład dla podrobionych paraf powinien być asymetryczny lub nawet unimodalny. Dlatego też, aby opracować lepiej dopasowany model, który lepiej oszacowywałby wartości DF i obliczał istotność, usunięto ze zbioru uczącego najmniej umiejętnie sfalszowane parafy. Sfalszowane parafy, dla których średnia częstotliwość sygnału przekroczonego dla IP w 21 zadaniach, a dla krzywizny w 22 zadaniach była większa niż 0,4, zostały odrzucone. Wybór takiego limitu podyktowany został uzyskaniem maksymalnie efektywnego modelu. Sfalszowane parafy charakteryzujące się dużą liczbą sygnałów przekroczonych, które wykluczono z obliczeń, nie wpłynęły na zwiększenie się liczby błędnych odpowiedzi po zastosowaniu nowych zasad decyzyjnych. Potwierdzeniem poprawności tego

założenia dla procesu optymalizacji był również fakt, iż uzyskano mniejszą bezwzględną wartość istotności sygnału, jednocześnie zwiększając (w granicy błędu statystycznego) efektywność metody w stosunku do efektywności metody opisanej w: Dutova, Li, Orlova i in. (1987). Dokładność oszacowania poszczególnych parametrów dokonano poprzez zastosowanie metody *bootstraps* (Chernick, 1999).

3. Wyniki

Statystykę opisową sygnału przekroczonego wyznaczonego dla parametrów integralnych zamieszczono w tabelach 1–8. Analiza istotności poszczególnych parametrów wykazała, iż parafa jest najwierniej odwzorowywana na jej początku, mniej dokładnie w środkowej części, natomiast najslabiej w końcowym fragmencie.

Statystykę opisową sygnału przekroczonego dla krzywizn zamieszczono w tabelach 9–16. Na podstawie uzyskanych istotności dla obu typów próbek, tj. autentycznych, jak i sfalszowanych paraf, policzono dodatkowe funkcje decyzyjne. Funkcje te wyliczono zarówno dla parametrów integralnych, krzywizny, jak i łącznie dla obu tych wielkości. Histogramy przedstawiające względną częstotliwość DF (rozkład gęstości prawdopodobieństwa) pokazano na rysunku 12. Na histogramach czarne słupki odnoszą się do danych uzyskanych dla paraf sfalszowanych, a białe dla paraf autentycznych.

Przedstawione histogramy jednoznacznie pokazują, iż możliwe jest wprowadzenie reguł decyzyjnych, które pozwolą na niezawodne rozwiązanie zadania polegające na określeniu, czy mamy do czynienia z parafą autentyczną, czy sfalszowaną. Wiarygodny i rzetelny wynik uzyskuje się zwłaszcza dla wyników negatywnych.

Reguły decyzyjne związane z rodzajem stawianych wniosków zostały zaproponowane na podstawie rozkładu DF (tabele 17–19) i przyjmują formę: definitywnie pozytywny, prawdopodobnie pozytywny, definitywnie negatywny i prawdopodobnie negatywny. Wyniki uzyskane dla reguł decyzyjnych wyliczonych osobno zarówno dla parametrów integralnych, jak i krzywizny, pokazują jedynie, który fragment parafy jest najbardziej zmienny. Dlatego też, aby rozwiązać zadanie i odpowiedzieć na pytanie, czy analizowana parafa jest autentyczna, czy też sfalszowana, biegły powinien wykorzystać jednocześnie reguły decyzyjne ustalone dla parametrów integralnych oraz krzywizny. Funkcje decyzyjne zbudowane dla parametrów integralnych mogą być użyte, jeśli biegły nie znajdzie w analizowanej parafie przynajmniej 10 połączeń krzywizny, które mógłby zmierzyć. Nowa metoda została zautomatyzowana dzięki opracowanemu programowi *Oko-1 UKr*.

4. Dyskusja wyników

Bezwzględne wartości istotności parametrów *IP* wykazały tendencję rosnącą od początku do końca długości parafy (tabela 7 i 8). Wyjątek stanowił parametr *H/V* wyznaczony dla końcowej części parafy. Jednakże liczba pomiarów tego parametru (tabela 1 i 2) okazała się mniejsza niż w przypadku innych parametrów i dlatego dokładność oszacowania istotności była znacznie niższa. Podobny trend obserwuje się dla istotności *IP* w metodzie opisanej w: Dutova, Li, Orlova i in. (1987). W tym eksperymencie istotność parametru *DDMP* wyznaczona dla początkowej części parafy przyjmowała bardzo małe wartości, dlatego zastąpiono ją zerami. W nowym eksperymencie bezwzględna wartość istotności tego parametru na początku parafy była również mniejsza niż w przypadku innych parametrów. Interesujące jest to, że według eksperymentu opisanego w: Orlova, Yakovleva, Smirnov (2010), istotności *IP* nie wykazywały żadnych tendencji związanych z tym, jaki fragment parafy był analizowany. Autorzy tłumaczą to zjawisko, stawiając następującą hipotezę: bardzo krótka i prosta parafa pozbawiona krętych, wijących się linii oraz w ogólnym rozumieniu czytelnych liter jest postrzegana przez fałszerza jako całość i tak też jest przez niego odwzorowywana. Tym samym początek parafy zostaje lepiej odwzorowany niż pozostałe jej części. W przypadku analizy pojedynczych czytelnych słów fałszerz rozkłada słowo na pojedyncze litery, starając się odwzorować rzeczywisty wygląd każdej litery z osobna bez względu na jej miejsce występowania w słowie.

Zgodnie z tabelami 15 i 16 istotności krzywizny w niniejszym eksperymencie i w opisanym przez Dutova, Li, Orlovą i in. (1987) dla *RLE* i *AA* wyszły zaskakująco dobrze. Dla innych pomiarów krzywizny bardzo trudno ustalić jakikolwiek parametr za wyjątkiem wartości średniej. W obu eksperymentach średnie wartości istotności dla krzywizny są nieznacznie większe niż dla *IP*. Powód lepszego dopasowania *IP* w eksperymentach można tłumaczyć tym, iż pomiar krzywizny jest zdecydowanie trudniejszym procesem, a ponadto subiektywnym. W pierwszym eksperymencie wszystkie pomiary wykonywane były manualnie, a w niniejszym praktycznie wykonywała je jedna osoba za pomocą programu komputerowego.

Skuteczność obu metod wyniosła 63%. Natomiast skuteczność metody policzona przy dwóch różnych progach, powyżej których sygnał uznawano za przekroczone, a mianowicie dla $B = 1,75$ oraz $B = 2,25$, wyniosła 59%. Dlatego też, optymalizując sygnał, uznano, iż za sygnał przeregulowany będzie uznany sygnał powyżej $B = 2$, co okazało się najlepszym rozstrzygnięciem, przy czym zmiana ta nie wpłynęła istotnie na wzrost efektywności metody.

5. Podsumowanie

Przeprowadzone eksperymenty pozwoliły zweryfikować metodę opisaną w: Dutova, Li, Orlova, i in. (1987) i ocenić ją jako skuteczną w przypadku języka ukraińskiego. Tym samym zmodyfikowana wersja tej metody może zostać przedstawiona biegłym pracującym w instytucjach wymiaru sprawiedliwości na Ukrainie.

Poprawność efektywności modelu może być zapewne podwyższona poprzez użycie funkcji decyzyjnej opartej na trzech regułach decyzyjnych zamiast dwóch, które zastosowano w opisywanych badaniach. Niemniej jednak należałoby przeprowadzić więcej eksperymentów, zwłaszcza że duży wpływ na ich wyniki ma fakt, czy stosowana w badaniach baza danych zawiera odpowiednio reprezentatywne próbki paraf (wspomniano tu o konieczności usunięcia więcej niż jednej trzeciej paraf z bazy danych). Tym samym, do eksperymentu powinno się wybrać grupę bardziej zmotywowanych „fałszerzy naśladowców” potrafiących starannie odwzorować parafę. Z drugiej jednak strony, jeśli fałszerze posiadający wiedzę profesjonalną i świetne zdolności byłiby zbyt mocno zmotywowani, to wydajność metody okazałaby się niska i z pewnością nie odzwierciedlałaby sytuacji, z jaką na co dzień spotykają się biegli.

Inną możliwością jest opracowanie metody umożliwiającej analizę kilku paraf jednocześnie. Wymaga to jednak opracowania innego modelu statystycznego.