



GRAPHLOG – COMPUTER SYSTEM SUPPORTING HANDWRITING ANALYSIS*

Ewa FABIAŃSKA¹, Marcin KUNICKI¹, Grzegorz ZADORA¹, Tomasz DZIEDZIC¹, Dariusz BUŁKA²

¹ *Institute of Forensic Research, Krakow*

² *Cyborg Idea S.C., Krakow*

Abstract

The Institute of Forensic Research has finalised creation of software entitled Graphlog. It was designed as a supportive tool in handwriting analysis for taking and processing measurements of a variety of features. In addition to taking measurements and image editing, Graphlog also allows storage and descriptive statistics analysis of a great number of results (the number of measurements, minimum, maximum, mean and the standard deviation are automatically displayed in a special window). The collected data can be then exported to other advanced statistical software. Graphlog's measuring functions have been divided into universal and particular. The former encompass general measurements such as distance, angle, baseline alignment angle, size (oval diameter) and proportions. Particular measurement functions compute specific handwriting features. Each particular measurement is coupled with one of the six universal measurements. The software is equipped with a default list of measurable features, which can be further expanded by the user to include other features – vital for carrying out research. Accuracy, repeatability and reproducibility of the measurements computed by Graphlog were confirmed by measurements of standardised manuscripts written by ten people. Each sample included twenty-two characters in three categories: central, upper, and lower zone letters, with each character written separately. The documents were scanned and the graphic files were given to seven people, who measured each manuscript five times. Repeatability was checked on the basis of ten samples of ordinary handwriting.

Key words

Handwriting; Document examinations; Computer programme; Graphic features.

Received 15 December 2006; accepted 12 January 2007

1. Introduction

Handwriting is an individual graphic trace of a person, determined by his or her intellectual and physical potential, and developed and consolidated through the process of learning to write. The individual appearance of handwriting is determined by inborn talents, primer models, and the influence of other people's graphic models. Personal identification based on handwriting analyses is significantly more problematic

than (personal) identification based on examination of other forensic traces such as fingerprints or DNA, due to, amongst other things, the natural variability of handwriting – which stems from its development as well as the state of health and age of the writer. It is for this reason, amongst others, that many attempts have been made over the years to systematise graphic features and objectify forensic handwriting examination.

Nowadays, the graphic-comparative method is used in handwriting identification. It evolved out of previous research methods and constitutes their logical continuation. „The graphic-comparative method is based

* This work was supported by the research grant (No. 2 HO2A 024 24) from Polish Ministry of Science.

upon the observation that handwriting is a manifestation of the author's psychophysical potential. The focus in handwriting analysis should therefore be on all features allowing determination and comparison of this potential" [7]. This statement implies the necessity to analyse all information present in the handwriting, both graphic and non-graphic, such as language traits (spelling, grammar, language). The graphic-comparative method is also based on the assumption that text or signature identification is the result of the natural phenomenon of graphic consistency of each sample in relation to all the others of the same writer, only in terms of a defined number of features, but never all of them (identicalness is evident proof of forgery). Consequently, every comparative analysis of the handwriting sample with evidentiary material is carried out on an individual basis, depending on the specific example [2, 6].

Document examination focuses on the analysis of various handwriting features, further divided into general features (for example, handwriting type, degree of naturalness, stages of handwriting development, level of writing skills), topographical features (for example, shape and width of margins, the number and width of indentations, character, word, and line layout in relation to each other), motor features (for example, writing speed, impulse, pressure and shading), measurable features (for example, word's area, character size, character width, slant angle, and proportions of various elements), and constructive features (for example, character construction, their variations, connections and repetitive forms). A handwriting expert analyses all the above-mentioned features using objective and subjective criteria. Unfortunately, subjectivity may lead to gross errors in recognising similarities or differences, and eventually result in erroneous conclusions. Thus, the fundamental task is to optimise and standardise the objective methods of examination of handwriting features: first and foremost, the measurable features, followed by topographic features, as they can also be measured [3, 7].

The Institute of Forensic Research has finalised a research project financed by the Polish Ministry of Science, with the aim of creating software for precise measuring analysis of handwriting features that would be reliable, objective, but, at the same time, simple-to-use on a daily basis. The Graphlog program makes it possible to avoid mistakes incurred through the use of imprecise measuring scales, and allows computation of measurements of many features, which, as a consequence, minimises the probability of committing unsystematic errors, and also permits extensive statistical processing of obtained results. Graphlog, in contrast to advanced systems for hand-

writing recognition and comparison such as: FISH, NIFO-TNO and WANDA [4, 5, 8, 10], is a simple-to-use tool, which is designed to help experts with their daily work.

Accuracy, repeatability and reproducibility of measurements computed with Graphlog were validated on the basis of measurements from standardised and ordinary handwriting samples, written by ten people.

2. Equipment – Graphlog software

2.1. Preparation for computations

Graphlog software has been designed to work in the Windows environment [11, 12] (Figure 1 and 2). The software was designed with the aim of being able to perform complex measurements on all graphic files and to be compatible with functions in simple graphic software. After the programme has been installed, the first step is to import the content of a graphic file (jpg, bmp, gif, tiff etc.), created by a different application through scanning or digital photography. As the software cannot recognise the resolution of the imported file, it is vital that when "fixing" the picture (e.g. by scanning), a linear scale of known length should be positioned. The imported image will be stored in a special *.grl format, which permits storage not only of graphic information but also of all measurements and additional objects imposed on the image.

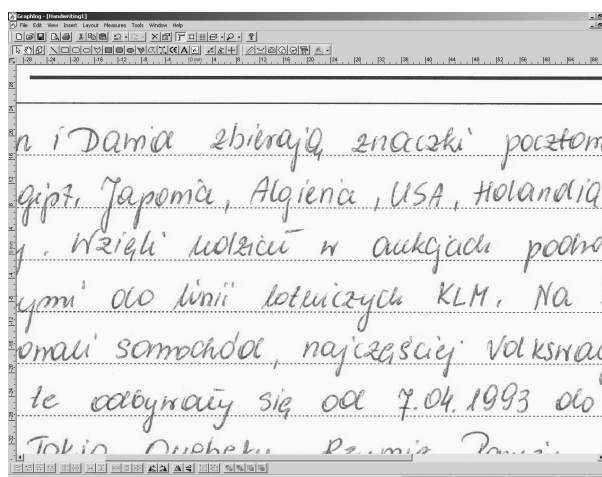


Fig. 1. Graphlog window.

Before starting measurements, it is necessary to scale the image, so that the image size projected on the monitor reflects its real size. In order to accomplish this, in the "Tools" menu, choose the function "Scale Tool", draw a line between two scale points, and store

the value of the actual distance between them in the appropriate field. This is to guarantee that all computations will reflect the actual size (Figure 3).

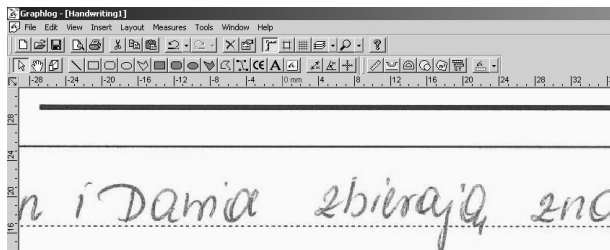


Fig. 2. Main menu.

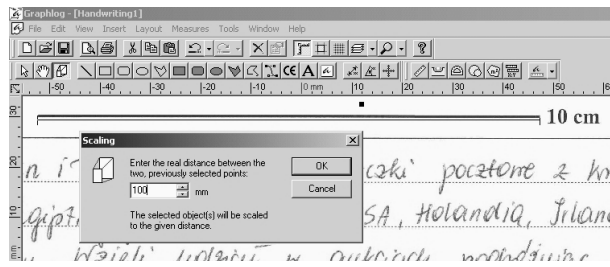


Fig. 3. Scaling window.

Additionally, a short description of the computed document, personal data of the examiner, and necessary annotations (e.g. case number) can be inserted in the “Document Properties” window (Figure 4), which can be helpful in filing documents and computations.

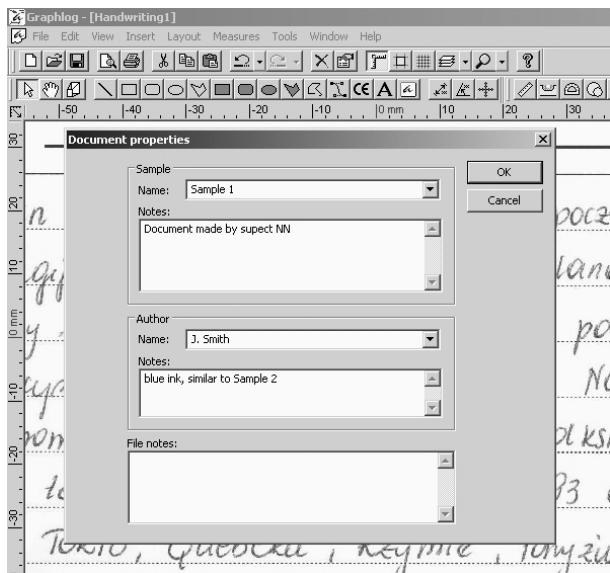


Fig. 4. Document properties window.

2.2. Universal computations

Graphlog’s basic tools are universal measure functions that permit individual computations of various dimensions in a picture. After each computation, the text-coded result is stored by the software on the clipboard, from where it may be exported to any other software. Universal computations include the following (Figure 5):

- distance computation: measures the distance between two points;
- slant computation: measures the angle of any line in relation to the horizontal line;
- angle computation: measures the angle between any two lines;
- size computation: measures the diameter of the chosen element or area in the picture;
- surface area computation: measures the surface area of the chosen field;
- proportion computation: automatically measures the ratio of two lengths.

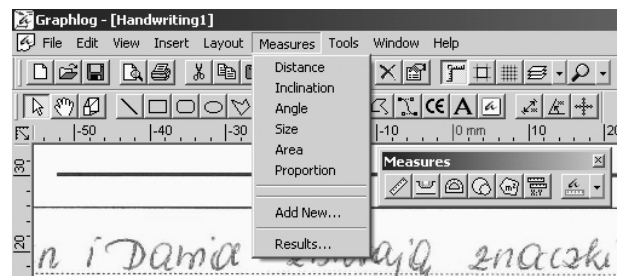


Fig. 5. Universal measurements menu.

2.3. Particular computations

Particular computations measure all significant measurable dimensions in handwriting and document analysis. The list of measurements available in Graphlog in the “Measures” menu includes the following features: heights of lower, central, and upper zone letters, mean slant and slant of individual characters, word widths, oval sizes, character and word spacing, and proportions of specific characters. The list can be expanded by the user to include any other variables vital for the examination. One of six universal measurements is assigned to each particular measurement. For example, if “The Height of Central Zone Letter” function is chosen, the distance between two points is measured. However, if “The Slant of the Letter” function is chosen, the angle between the selected line and the baseline is measured (Figure 6).

The results for each of the measurements are displayed in the appropriate window for the specific par-

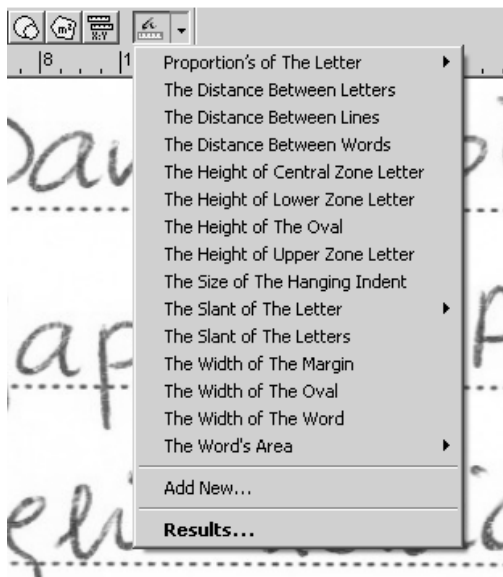


Fig. 6. Particular measurements menu.

particular measurement. This allows the obtained results to be monitored in real time, and if a mistake is made, to correct it immediately. Moreover, this window allows all the results to be copied onto the clipboard, automatically displays the number of measurements taken, their minimal, maximal, mean and standard deviation. The results of all the particular measurements are stored together with the *.grl graphic file and can be exported to data bases (Figure 7).

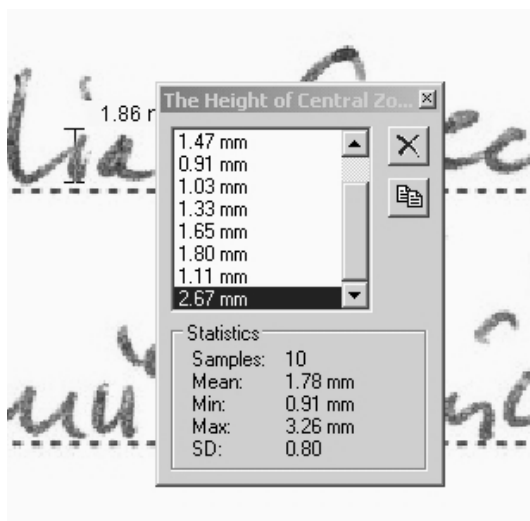


Fig. 7. Results of measurements window.

2.4. Graphic functions

Graphlog software has at its disposal a number of graphic elements, which may be used to mark chosen features of an examined handwriting sample. The fol-

lowing objects can be selected from the “Insert” menu: line, rectangle, ellipsis, polygon, broken line, Bezier curve, any symbol or text (Figure 8). Each of them can be modified in terms of size, line thickness, colour, slant etc. It is also possible to introduce the tools: linear value and angle value, which illustrate the appropriate measurements and provide the calculated values (Figure 9).

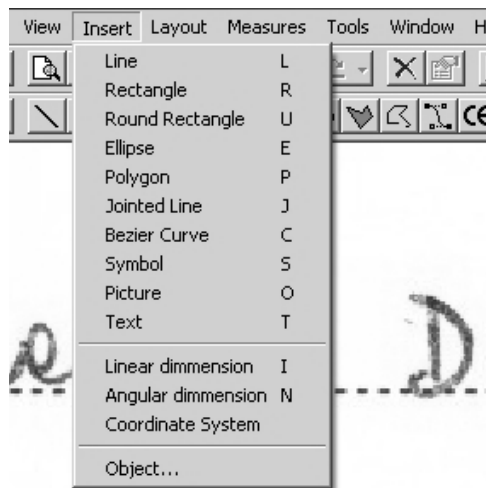


Fig. 8. List of graphical objects.

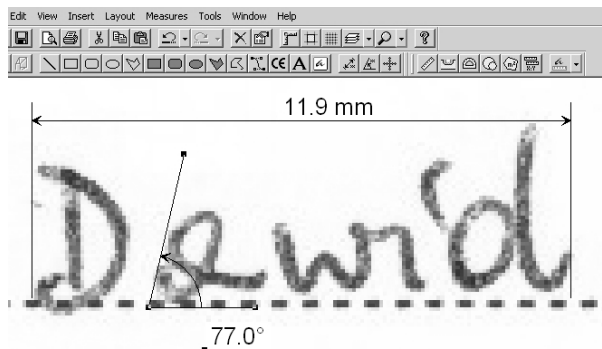


Fig. 9. Illustrating measurements.

Additionally, Graphlog software permits editing of pictures – change their shape, rotate and group them, set display order, magnify as desired, and impose reference lines.

3. Experimental

3.1. Graphlog software validation postulates

In practice, in order to examine measurable features in a handwriting sample, variables defined as the following are utilised:

- heights of the central zone letters (e.g. *a, c, e*), upper zone letters (e.g. *b, d, k*) and lower zone letters (e.g. *g, j, y*);
- upper zone to central zone letters heights ratio;
- lower zone to central zone letters heights ratio;
- slant angle of characters.

For these defined variables to be useful, letter heights and slant angle measurements computed by Graphlog have to be reliable. That is why the repeatability and reproducibility of measurements were thoroughly checked. The two were defined as:

- repeatability – descriptions of a given script sample in terms of single-category letter heights or character slant angles carried out by one examiner in several independent series of tests are the same,
- reproducibility – descriptions of a script sample in terms of single-category letter heights or character slant angles carried out by several examiners are the same.

Moreover, it would be advisable that all Graphlog computed measurements be accurate – for example, reflect the real height values for specific letters.

3.2. Sample preparation

In order to verify accuracy, repeatability, and reproducibility of measurements computed by Graphlog, standardised manuscripts written by ten different people were prepared. Each sample included characters falling into one of the three groups (central, upper and lower zone) – 22 characters from each group (Figure 10).

Nr 11	W tym polu proszę podać wiek i płeć M., 28
----------	---

(Proszę przepisać poniższy tekst pismem zwykłym)



Fig. 10. Standardised sample.

The documents were scanned and their graphic files handed over to seven people responsible for taking measurements. Each person individually scaled

the sample scans (graphic files) before examination. Additionally, repeatability was examined on the basis of samples of ordinary handwriting comprising several hundred characters written by ten people.

3.3. Manual measurements

Each of the ten standardised documents (handwriting samples taken from ten people) was measured using the classical method, that is using a ruler to measure the heights of 22 upper zone, 22 lower zone and 22 central zone letters, with accuracy to 0.1 mm. These measurements were taken by three experts (designated as A, B, and C) from the Institute of Forensic Research in Krakow, Poland. Taking into account the difficulty in obtaining accurate and repeatable angle measurement values when using the classical method (manual measurements), it was decided not to use them in validation, with character height values as the sole basis for validation of Graphlog's computational accuracy.

3.4. Measurements taken with the use of Graphlog software

3.4.1. Standardised samples

In each of ten standardised handwriting samples, Graphlog computed the heights of 22 upper zone, 22 lower zone, and 22 central zone letters with an accuracy to 0.1 mm, and the slant angle of 44 upper zone and lower zone letters with an accuracy to 0.1°. The measurements were taken by seven examiners, designated as A to G, at three different locations, namely, the Institute of Forensic Research, Krakow (examiners designated as A, B, C), the Department of Criminalistics at the University of Silesia, Katowice (examiners designated as D and E), and Department of Criminalistics at Adam Mickiewicz University, Poznań (examiners designated as F and G). Each person measured each sample five times. Since one of the examiners did not comply with the rules concerning the character slant angle measurements, the results obtained for this category examination were rejected.

3.4.2. Ordinary samples

Ordinary samples, similarly to standardised samples, were examined by seven people. They consisted of ten specially-designed forms filled in with natural handwriting. All lower zone letters were measured together with a corresponding number of other letters from the other two categories. In the case of ordinary

samples, character slant angles were not measured, due to varying legibility of the scripts leading to problems in establishing a precise procedure for such calculations.

4. Results and discussion

4.1. Descriptive statistics

As an example, Tables 1–2 present the descriptive statistics of results obtained by person A, when letter

heights were measured manually and computed by Graphlog in ten standardised and ordinary handwriting samples (Table I). These may also give some insight into the accuracy of measurement performed manually and computed by Graphlog. The results presented in Table II illustrate the repeatability of letter heights and slant angle computed by Graphlog. The reproducibility of these measurements is represented by the data in Table III. These show the descriptive statistics of computations by Graphlog Central Zone Letter heights for ten ordinary handwriting samples taken from 7 people.

TABLE I. DESCRIPTIVE STATISTICS OF LETTER HEIGHTS MEASURED MANUALLY AND COMPUTED BY GRAPHLOG IN TEN STANDARDISED AND ORDINARY HANDWRITING SAMPLES – FOR EXAMPLE THE RESULTS OBTAINED FROM PERSON A

Sample	Statistics	Manually measured			Measured by Graphlog					
		Standardised handwriting			Standardised handwriting			Ordinary handwriting		
		Letter type			Letter type			Letter type		
		CZ	UZ	LZ	CZ	UZ	LZ	CZ	UZ	LZ
1	Mean	2.34	3.55	4.00	2.28	3.49	3.97	2.59	3.64	5.51
	SD	0.33	0.54	0.48	0.37	0.6	0.46	0.27	0.37	0.62
2	Mean	3.17	5.06	5.73	3.12	5.02	5.74	2.09	4.71	6.37
	SD	0.46	0.49	0.89	0.44	0.49	0.87	0.49	0.74	1.16
3	Mean	2.70	4.14	4.30	2.74	4.13	4.30	1.97	4.03	5.25
	SD	0.30	0.51	0.48	0.29	0.52	0.48	0.57	0.76	0.90
4	Mean	2.19	3.30	4.34	2.20	3.26	4.36	1.87	3.33	4.02
	SD	0.44	0.42	0.61	0.45	0.42	0.64	0.36	0.29	0.50
5	Mean	2.25	4.70	4.75	2.24	4.73	4.75	2.20	3.41	5.40
	SD	0.34	0.45	0.61	0.32	0.44	0.62	0.33	0.43	0.59
6	Mean	2.25	3.78	4.03	2.21	3.75	4.01	1.74	3.35	4.88
	SD	0.34	0.38	0.53	0.34	0.35	0.52	0.42	0.64	0.74
7	Mean	3.68	5.09	5.58	3.69	5.10	5.52	3.04	6.15	7.92
	SD	0.82	0.59	0.82	0.81	0.59	0.83	0.61	1.09	1.48
8	Mean	2.50	3.56	3.71	2.52	3.50	3.71	2.83	5.21	7.39
	SD	0.40	0.54	0.66	0.40	0.56	0.65	0.38	0.64	1.02
9	Mean	3.11	4.50	5.35	3.13	4.50	5.39	3.22	5.78	7.10
	SD	0.50	0.65	1.32	0.51	0.65	1.35	0.47	0.67	0.83
10	Mean	2.38	3.57	5.12	2.37	3.58	5.14	1.48	4.49	4.81
	SD	0.34	0.54	0.70	0.35	0.54	0.70	0.41	0.78	0.83

CZ – central zone letters, UZ – upper zone letters, LZ – lower zone letters.

TABLE II. DESCRIPTIVE STATISTICS OF LETTER HEIGHTS AND SLANT ANGLE COMPUTED BY GRAPHLOG USING A STANDARDISED SAMPLE (No. 1) BY PERSON A FIVE TIMES

Run	Statistics	Type of measurement			
		CZ	UZ	LZ	SL
I	Mean	2.28	2.25	2.3	2.36
	SD	0.37	0.38	0.37	0.36
II	Mean	2.28	2.25	2.29	2.38
	SD	0.37	0.37	0.36	0.35
II	Mean	2.28	2.26	2.28	2.34
	SD	0.37	0.37	0.36	0.34
IV	Mean	2.3	2.27	2.3	2.34
	SD	0.37	0.35	0.37	0.37
V	Mean	2.3	2.26	2.29	2.32
	SD	0.36	0.37	0.37	0.36

CZ – central zone letters, UZ – upper zone letters, LZ – lower zone letters, SL – slant angle.

4.2. Analysis of accuracy of measurements

It can be postulated that Graphlog measurements precisely reflect real letter height values if the letter height measurements of the i -th letter obtained classically with a ruler and computed by the software are the same. It can be transcribed more formally as:

$$d = x_{hi} - x_{Gi} = 0 \quad \{1\},$$

where: x_{hi} – i -th letter height (e.g. central zone) measured manually; x_{Gi} – i -th letter height (e.g. central zone) computed by Graphlog.

Ideally, in every case of letter height measured manually and computed by Graphlog the same values are obtained, i.e. $d = 0$ for each measured letter. The mean d value for all measured letters will also equal 0 ($\bar{d} = 0$). Thus, it may be hypothesised that measurements computed by Graphlog are accurate if $\bar{d} = 0$. In reality, such a situation is very unlikely to occur, which is further supported by the results provided in the Table IV.

That is why it was checked whether $\bar{d} = 0.064$ is statistically different from zero. In order to do that, the t-Student test [9] was used, in which the null hypothesis was tested in the form of $H_0: \bar{d} = 0$.

The measurements were computed with the routines written with the R software [1, 13]. The decision to accept or to reject the null hypothesis was taken on

the basis of calculated value of significance probability p , compared then with the presumed significance level α . The α value, for the test duration, was set at the $\alpha = 0.05$ level; most often used in type t-Student test analysis. Values written in italics (Table V) indicate cases, where H_0 had to be rejected due to substantial differences between manual measurements and those computed by Graphlog.

However, in the majority of cases, the H_0 hypothesis was accepted, which indicates a lack of substantial differences between manual measurements and Graphlog computations. It can thus be postulated that, in the majority of cases, measurements computed by Graphlog reflect the real values of letter heights. Cases where the null hypothesis had to be rejected are probably related to the inaccuracy of the manual measurements with accuracy to 0.1 mm, at $16 \times$ microscopic magnification, which, in consequence, may have led to errors in taking of letter measurements, especially upper zone and lower zone letters with a considerable slant.

4.3. Measurements repeatability analysis

As mentioned above, every examiner taking part in the experiment measured each sample ten times, i.e. calculated the height of 22 central zone, 22 upper zone, and 22 lower zone letters, as well as the slant of upper zone and lower zone letters (44 letters in total). Measurements repeatability is achieved if every person in each examination run obtains the same data for a particular handwriting sample.

Assuming that mean height of letters in a given category of characters describes the results obtained by a specific person in a particular examination run, the following hypothesis for results repeatability can be formulated – means of letters heights, in one category of characters, calculated in each of the examinations runs for a particular handwriting sample are the same, e.g.:

$$H_0: \bar{x}_{A1,I} = \bar{x}_{A1,II} = \bar{x}_{A1,III} = \bar{x}_{A1,IV} = \bar{x}_{A1,V} \quad \{2\},$$

where: $\bar{x}_{A1,I}$ – mean height of central zone letters (s), calculated on the basis of the heights of 22 letters, measured in sample No. 1 by person A, in examination run I.

A null hypothesis formulated in this way was tested with the F test [9]. Similarly to measurements accuracy analysis, the decision to accept or reject the null hypothesis was taken on the basis of calculated probability of significance p , compared with the assumed significance level, $\alpha = 0.05$. Similar hypotheses were formulated and tested for:

a) each of the three categories of characters (central zone, upper zone, and lower zone letters) and each of the seven examiners;

b) angle slant measurements results obtained by six people.

All comparisons showed probabilities of significance p greater than 0.05, and in most cases, they were close to 1. Therefore, the obtained p values allow acceptance of all proposed null hypotheses (H_0). It can thus be stated that compared means are statistically equal, and consequently, that measurements obtained by one person using Graphlog allow obtaining of repeatable descriptions for the analysed samples.

4.4. Reproducibility of measurements analysis

It was assumed that Graphlog computations can be accepted as reproducible if the sample descriptions (mean letter heights for each category of characters, or

mean slant angle) obtained for a specific handwriting sample by various examiners are identical.

In this case the F test was also used to evaluate reproducibility; this tested the hypothesis, for example, for central zone letters, that mean heights of central zone letters in a specific standardised handwriting sample (e.g. the first) measured by all seven people are statistically equal, which may be transcribed as the following:

$$H_0: \bar{x}_{s_{AI}} \quad \bar{x}_{s_{BI}} \quad \bar{x}_{s_{CI}} \quad \bar{x}_{s_{DI}} \quad \bar{x}_{s_{EI}} \quad \bar{x}_{s_{FI}} \quad \bar{x}_{s_{GI}} \quad \{3\},$$

where: $\bar{x}_{s_{AI}}$ – mean height for central zone letters (s), calculated from the height measurements of 22 letters, measured for sample No. 1 by person A in examination run I.

Similar hypotheses were formulated for:

- a) letter heights in each of three analysed categories of characters, for each of the 10 analysed standardised handwriting samples and as part of each examination run;

TABLE III. DESCRIPTIVE STATISTICS OF CENTRAL ZONE LETTER HEIGHTS COMPUTATIONS BY GRAPHLOG FOR TEN ORDINARY HANDWRITING SAMPLES TAKEN FROM 7 PEOPLE

Sample	Statistics	Person						
		A	B	C	D	E	F	G
1	Mean	2.59	2.59	2.59	2.58	2.65	2.59	2.66
	SD	0.27	0.45	0.31	0.37	0.45	0.26	0.34
2	Mean	2.09	1.99	2.25	2.17	2.26	2.22	2.38
	SD	0.49	0.54	0.57	0.58	0.61	0.64	0.61
3	Mean	1.97	1.95	1.98	1.97	2.03	1.97	2.17
	SD	0.57	0.77	0.62	0.64	0.75	0.57	0.86
4	Mean	1.87	1.74	1.86	1.83	1.91	1.86	1.97
	SD	0.36	0.39	0.39	0.37	0.51	0.39	0.55
5	Mean	2.20	2.11	2.21	2.16	2.31	2.27	2.30
	SD	0.33	0.45	0.39	0.38	0.46	0.45	0.45
6	Mean	1.74	1.72	1.75	1.70	1.84	1.75	1.84
	SD	0.42	0.56	0.38	0.40	0.55	0.42	0.57
7	Mean	3.04	2.91	2.98	2.93	3.03	3.04	2.98
	SD	0.61	0.65	0.54	0.52	0.61	0.61	0.54
8	Mean	2.83	2.93	2.84	2.81	2.94	2.93	2.84
	SD	0.38	0.61	0.38	0.37	0.62	0.61	0.38
9	Mean	3.22	3.22	3.17	3.14	3.23	3.22	3.17
	SD	0.47	0.63	0.52	0.52	0.62	0.63	0.52
10	Mean	1.48	1.51	1.46	1.41	1.53	1.51	1.54
	SD	0.41	0.56	0.38	0.37	0.56	0.56	0.57

- b) letter slant angle of lower and upper zone letters values obtained by six people in each of the 10 analysed standardised handwriting samples;
- c) letter heights in each of three analysed categories, for each of the 10 analysed ordinary handwriting samples.

TABLE IV. CENTRAL ZONE LETTER HEIGHTS MEASURED MANUALLY BY PERSON A IN THE FIRST SAMPLE EXAMINATION AND COMPUTED BY GRAPHLOG (FIRST MEASUREMENTS SERIES)

Manual x_{hi}	Graphlog x_{Gi}	d
3.0	2.9	0.1
2.3	2.3	0.0
2.3	2.4	0.1
2.2	2.1	0.1
2.1	2.1	0.0
2.6	2.6	0.0
1.9	1.9	0.0
2.2	2.3	0.1
1.8	1.8	0.0
2.4	2.3	0.1
2.6	2.6	0.0
2.7	2.6	0.1
2.8	2.8	0.0
2.4	2.3	0.1
2.6	2.5	0.1
2.2	1.6	0.6
1.8	1.7	0.1
2.2	2.1	0.1
2.1	2.0	0.1
2.3	2.2	0.1
2.1	2.1	0.0
2.9	2.9	0.0
		$\bar{d} = 0.064$

Analysis of the reproducibility of letter height values computed by Graphlog from standardised samples showed all obtained p values to be greater than the assumed significance level $\alpha = 0.05$, and, in the majority of cases, these were close to 1. Letter heights measurements computed from ordinary samples by Graphlog also showed, in the majority of cases, p values greater

than 0.05. Values lower than the assumed significance level were obtained for the central zone letters in samples 2 ($p = 0.002$) and 5 ($p = 0.039$), and for the upper zone letters in the case of sample 4 ($p = 0.013$), these may be the result of accidental measurement errors, or carelessness of the examiner.

Nevertheless, it may be stated that the compared means are statistically equal, and, consequently, that the measurements computed by various people with the use of Graphlog software enable obtaining of reproducible sample descriptions for heights of analysed letters.

Reproducibility of letter slant angle values in the upper and lower zone letters was achieved (i.e. null hypotheses were accepted) in 7 out of 10 examined samples, at the level $\alpha = 0.05$, and in 9 samples at the level $\alpha = 0.01$. The lack of perfect reproducibility is related to accuracy problems in establishing the axis in characters while measuring slant angles. Even with this drawback, Graphlog software provides better angle measurement results than the classical method.

5. Conclusions

Graphlog software, designed and developed at the Institute of Forensic Research in Krakow, is a precise, reliable, and easy-to-use tool for analysis of measurable features of handwriting. It permits quick and easy obtaining of numerous measurements of a diverse nature as well as their preliminary analysis (the software automatically generates the number of measurements, minimum and maximum values, mean and standard deviation). An additional advantage can be found in its potential to minimise mistakes, which are a consequence of inaccurate measuring scales, and the standardisation of measuring procedures. Thanks to archiving of all measurement results and the software's high integration with the database, all data can be easily exported to other software environments, which provide more advanced statistical analysis.

Graphlog software was validated through accuracy, repeatability, and reproducibility verification. For this purpose, two types of handwriting samples were prepared (standardised and ordinary), which were then examined by seven people using the classical method and with the use of the software. Statistical analysis encompassing ten to twenty thousand individual results, showed not only accuracy in letter height measurements, but also repeatability of measurements as far as central, upper, and lower zone letters and character slant angles were concerned. Furthermore, ideal reproducibility of letter height values was ob-

TABLE V. MEASUREMENT ACCURACY ANALYSIS – PROBABILITIES OF SIGNIFICANCE (p) OBTAINED IN t -STUDENT TEST TESTING

Sample	Central zone letters			Upper zone letters			Lower zone letters		
	Person A	Person B	Person C	Person A	Person B	Person C	Person A	Person B	Person C
1	0.040	0.576	0.576	0.216	0.657	0.747	0.083	0.163	0.162
2	0.298	0.279	0.540	0.000	0.576	0.072	0.329	0.197	0.747
3	0.073	0.009	0.286	0.605	0.001	0.162	0.747	0.031	0.266
4	0.378	1.000	0.345	0.000	0.030	0.009	0.446	0.192	0.001
5	0.213	0.186	0.313	0.016	0.612	0.261	0.747	0.104	0.540
6	0.176	0.623	0.427	0.162	0.234	0.261	0.171	0.116	0.031
7	0.329	0.201	0.186	1.000	1.000	0.162	0.000	0.329	0.000
8	0.104	0.107	0.329	0.034	0.030	0.220	1.000	0.715	0.162
9	0.043	0.329	0.747	0.747	0.104	0.715	0.131	0.229	0.771
10	0.896	0.665	1.000	0.427	0.225	1.000	0.162	0.030	0.540

served in standardised samples, and close to ideal in ordinary samples. Due to accuracy problems in the procedure for establishing the axis of characters during slant angle measurements, the reproducibility for this value was achieved in 7 out of 10 standardised samples (at the level of $\alpha = 0.05$). Nonetheless, Graphlog enables examiners to obtain decidedly better results for angle measurements than does the classical method.

Summing up, Graphlog is a valuable tool for handwriting examination, and may be a useful research tool in analysis of measurable features of handwriting for population tests. It may also contribute to standardisation of various expert laboratory examination procedures.

References

1. Crawley M. J., Statistics – An introduction using R, John Wiley & Sons Ltd, Chichester 2005.
2. Ellen D., The scientific examination of documents, Taylor & Francis Ltd, London 1997.
3. Fabiańska E., Propozycja układu kodowania cech identyfikacyjnych pisma ręcznego, *Z zagadnień kryminalistyki* 1988, 20, 71–78.
4. Franke K., Schomaker L., Veenhuis C. [et al.], WANDA: A common ground for forensic handwriting examination and writer identification, *Bulletin of the European Network of Forensic Handwriting Experts* 2004, 1, 23–47.
5. Fund B., Rogers D., Schmittat R., A computer program designed to compare the spatial elements of handwriting, *Forensic Science International* 1994, 68, 195–203.
6. Harrison W. R., Suspect documents, Their scientific examination, Sweet & Maxwell Limited, London 1966.
7. Koziczak A., Metody pomiarowe w badaniach pismoznawczych, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków 1997.
8. Kroon L. N., de Jong W. C., Schmidt R. P., Computer aided analyses of handwriting, the NIFO-TNO approach, Proceedings of 4th European Conference for Police and Government Handwriting Experts, London 1994.
9. Lucy D., Introduction to statistics for forensic scientists, John Wiley & Sons Ltd, Chichester 2005.
10. Philipp M., Expected future developments in the Forensic Information System Handwriting (FISH), Proceedings of 4th European Conference for Police and Government Handwriting Experts, London 1994.
11. Tadeusiewicz R., Korohoda P., Komputerowa analiza i przetwarzanie obrazu, Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997.
12. Wojnar L., Majorek M., Komputerowa analiza obrazu, Fotobit Design, Kraków 1994.
13. www.r-project.org, The R foundation for statistical computing version 2.0.1, ISBN 3-900051-07-0.

Corresponding author

Ewa Fabiańska
Instytut Ekspertyz Sądowych
ul. Westerplatte 9
PL 31-033 Kraków
e-mail: efabianska@ics.krakow.pl

GRAPHLOG – KOMPUTEROWY SYSTEM WSPOMAGAJĄCY BADANIE CECH PISMA RĘCZNEGO*

1. Wstęp

Pismo ręczne jest indywidualnym śladem graficznym człowieka, zależnym od jego możliwości intelektualnych i fizycznych, a wykształconym i utrwalonym w procesie nauki pisania. O indywidualnym obrazie pisma decydują m.in. umiejętności wrodzone, wzorce elementarzone oraz wpływ wzorców graficznych innych osób. Identyfikacja człowieka na podstawie jego pisma ręcznego w porównaniu do innych śladów kryminalistycznych o indywidualnym charakterze, np. odcisków palców czy DNA, następuje znacznie większych trudności, m.in. ze względu na naturalną zmienność pisma wynikającą z jego rozwoju, a także stanu zdrowia lub wieku osoby piszącej. Dlatego też od wielu lat podejmowane są próby usystematyzowania cech graficznych oraz zobiektywizowania badań identyfikacyjnych pisma ręcznego.

Obecnie w badaniach identyfikacyjnych pisma stosowana jest metoda graficzno-porównawcza będąca efektem ewolucji poprzedzających ją sposobów badań i stanowiąca ich logiczną kontynuację. „Metoda graficzno-porównawcza opiera się na spostrzeżeniu, że pismo jest zewnętrzznym wyznacznikiem możliwości psychofizycznych wykonawcy. Przedmiotem badań w piśmie powinny być zatem wszystkie cechy pozwalające owe możliwości ustalić i porównać” [7]. Implikacją tego stwierdzenia jest konieczność analizy wszystkich informacji obecnych w piśmie, zarówno graficznych, jak i pozagraficznych, takich jak zawartość treściowa oraz cechy językowe (ortografia, gramatyka, język). Podstawę metody graficzno-porównawczej stanowi także założenie, iż identyfikacja zapisu bądź podpisu jest wynikiem całkowicie naturalnego zjawiska, którym jest zgodność graficzna każdego zapisu w stosunku do wszystkich pozostałych tylko pod względem pewnej liczby cech, a nigdy wszystkich (identyczność jest ewidentnym dowodem fałszerstwa zapisu). Z tego powodu każde badanie komparatystyczne materiału dowodowego i porównawczego oceniane jest indywidualnie w zależności od konkretnego przypadku [2, 6].

W trakcie przeprowadzania badań analizuje się zespoły cech pisma, na które składają się cechy ogólne (np. typ pisma, stopień naturalności, etapy rozwoju pisma, klasa pisma), cechy topograficzne (np. kształt marginesów i ich przebieg, liczba i głębokość akapitów, układ znaków, wyrazów i wierszy względem siebie), cechy motoryczne (np. tempo pisania, impuls pisma, nacisk i cieniowanie linii graficznej), cechy mierzalne (np. pole pisma, wielkość pisma, szerokość znaków, nachylenie pis-

ma i proporcje wielkościowe różnych elementów) oraz cechy konstrukcyjne (np. budowa znaków, ich odmiany, budowa wiązań, a także rodzaje form powtarzalnych). Ekspert pisma, analizując wszystkie wyżej wymienione cechy, stosuje kryteria zarówno obiektywne, jak i subiektywne. Niestety te ostatnie mogą stanowić źródło istotnego błędu w ocenie podobieństw bądź różnic pomiędzy materiałem kwestionowanym i porównawczym, a w efekcie decydować o podjęciu błędnych wniosków końcowych. Tak więc podstawowym zadaniem jest jak najlepsze dopracowanie i standaryzacja obiektywnych metod oceny własności pisma, do których należą przede wszystkim cechy mierzalne, a w następnej kolejności topograficzne, ze względu na ich wymierny charakter [3, 7].

W Instytucie Ekspertyz Sądowych zrealizowano projekt badawczy finansowany przez Ministerstwo Nauki i Informatyzacji mający na celu opracowanie wiarygodnego, obiektywnego, a zarazem prostego w codziennym użyciu programu komputerowego służącego do precyzyjnej analizy cech mierzalnych pisma. Program ten, nazwany Graphlog, umożliwi uniknięcie błędów będących skutkiem stosowania niedokładnych skal pomiarowych i pozwala na wykonanie dużej liczby pomiarów, co redukuje do minimum możliwość popełnienia niesystematycznych błędów, a także daje możliwość szerokiej obróbki statystycznej otrzymanych wyników. Graphlog, w przeciwieństwie do zaawansowanych systemów rozpoznawania i porównywania pisma ręcznego, takich jak FISH, NIFO-TNO i WANDA [4, 5, 8, 10], jest prostym w użyciu narzędziem, które ma być pomocne w codziennej pracy eksperta.

Dokładność, powtarzalność i odtwarzalność pomiarów wykonanych za pomocą programu Graphlog została poddana procedurze walidacji. Oceny wymienionych cech dokonano na podstawie pomiarów uzyskanych dla zestandaryzowanych i rzeczywistych próbek pisma nakreślonych przez 10 osób.

2. Narzędzia – program Graphlog

2.1. Przygotowanie do pomiarów

Program Graphlog został zaprojektowany do pracy w systemie Windows [11, 12] (rycina 1 i 2). Głównym założeniem w trakcie projektowania programu było zastosowanie jego funkcji do wykonywania zróżnicowanych pomiarów wielkościowych na dowolnych plikach graficznych oraz możliwość zastosowania funkcji spotykanych w prostych programach graficznych. Po zains-

* Projekt został sfinansowany przez Ministerstwo Nauki i Informatyzacji w ramach grantu nr 2 HO2A 024 24.

talowaniu programu pracę rozpoczyna się od importu zawartości pliku graficznego (jpg, bmp, gif, tiff itd.) powstałego za pomocą innej aplikacji poprzez skanowanie lub fotografię cyfrową. Ze względu na to, że program nie rozpoznaje rozdzielczości importowanego pliku, podczas utrwalania (np. skanowania) obrazu konieczne jest umieszczenie podziałki liniowej o znanej długości. Zaimportowany obraz zostaje zapisany w specjalnym formacie *.grl pozwalającym na zachowanie nie tylko informacji graficznej, ale także wszystkich wykonanych pomiarów oraz nałożonych na obraz dodatkowych obiektów.

Przed przystąpieniem do wykonywania pomiarów konieczne jest odpowiednie wyskalowanie zdjęcia, tak, by widoczne na ekranie monitora odległości odpowiadały rzeczywistym wymiarom. W tym celu z menu „Narzędzia” należy wybrać funkcję „Skalowanie”, a następnie przeprowadzić linię pomiędzy dwoma punktami podziałki i wprowadzić wartość rzeczywistej odległości między nimi w odpowiednie pole. W ten sposób uzyskuje się pewność, że wszystkie wykonywane pomiary będą odpowiadały rzeczywistym (rycina 3). Dodatkowo w oknie „Właściwości dokumentu” (rycina 4) możliwe jest zamieszczenie krótkiego opisu mierzonego dokumentu, danych osoby wykonującej pomiary oraz koniecznych uwag (np. numeru sprawy), co przydatne jest w archiwizacji dokumentów i pomiarów.

2.2. Pomiary uniwersalne

Podstawowymi narzędziami programu Graphlog są funkcje zaliczane do pomiarów uniwersalnych, które umożliwiają wykonywanie pojedynczych pomiarów różnych wielkości na zdjęciu. Po wykonaniu każdego z pomiarów, uzyskany wynik w postaci tekstowej umieszczany jest przez program w schowku, z którego może być kopiowany do dowolnego innego programu. Do pomiarów uniwersalnych należą (rycina 5):

- pomiar odległości służący do wyznaczania odległości między dwoma punktami;
- pomiar nachylenia pozwalający zmierzyć kąt nachylenia dowolnej linii względem linii poziomej;
- pomiar kąta wyznaczający wartość kąta pomiędzy dwiema dowolnymi liniami;
- pomiar rozmiaru, który podaje wartość średnicy wskazanego elementu lub obszaru na rysunku;
- pomiar pola powierzchni, mierzący pole powierzchni wskazanego obszaru;
- pomiar proporcji, który automatycznie oblicza proporcję pomiędzy dwiema długościami.

2.3. Pomiary specjalistyczne

Pomiary specjalistyczne służą do oznaczania wszystkich istotnych wielkości mierzalnych w badaniach pisma i dokumentów. Lista tych pomiarów znajdująca się

w programie Graphlog w menu „Pomiary”, a obejmująca m.in. takie cechy, jak wysokość liter śród-, nad- i podlinijnych, kąt nachylenia wszystkich oraz poszczególnych znaków, szerokości wyrazów, owali, odstępów między literami czy wyrazami i proporcje poszczególnych liter, może być uzupełniana przez użytkownika o dodatkowe, konieczne do wykonania badań zmienne. Do każdego z pomiarów specjalistycznych przypisany jest jeden z sześciu pomiarów uniwersalnych. Na przykład po wybraniu funkcji „Wysokość litery śródlinijnej” oznacza się pomiar pomiędzy dwoma punktami, zaś po wybraniu funkcji „Kąt nachylenia znaku” mierzy się nachylenie wyznaczonej linii względem linii poziomej (rycina 6).

Wyniki każdego z wykonanych pomiarów widoczne są w odpowiednim oknie danego pomiaru specjalistycznego. Pozwala to śledzić na bieżąco uzyskiwane wyniki, a w przypadku popełnienia błędu na jego szybkie skorygowanie. Ponadto okno to umożliwia skopiowanie do schowka wszystkich wyników oraz automatycznie pokazuje liczbę wykonanych pomiarów, ich średnią, wartość minimalną i maksymalną, a także odchylenie standardowe. Rezultaty wszystkich wykonanych pomiarów specjalistycznych zostają zapisane wraz z plikiem graficznym *.grl i mogą być eksportowane do baz danych (rycina 7).

2.4. Funkcje graficzne

Program Graphlog dysponuje szeregiem elementów graficznych, które mogą służyć do zaznaczenia wybranych cech mierzonej próbki pisma. I tak, z menu „Wstaw” wprowadza się następujące obiekty: linię, prostokąt, elipsę, wielokąt, łamaną, krzywą Beziera, dowolny symbol i tekst (rycina 8). Każdy z nich może być dowolnie edytowany pod względem wielkości, grubości linii, koloru, nachylenia itd. Ponadto możliwe jest wprowadzenie narzędzi: wymiaru liniowego i wymiaru kąтового, służących do zilustrowania odpowiednich pomiarów łącznie z podaniem ich wartości (rycina 9).

Dodatkowo program Graphlog pozwala pracować na warstwach obiektów, zmieniać ich kształt, obracać, grupować, decydować o kolejności wyświetlania, dowolnie powiększać oraz wprowadzać linie odniesienia.

3. Część eksperymentalna

3.1. Walidacja programu Graphlog – założenia

W praktyce do oceny cech mierzalnych próbki pisma ręcznego najczęściej stosuje się zmienne zdefiniowane jako:

- a) wysokość liter śródlinijnych (np. *a*, *c*, *e*), nadlinijnych (np. *b*, *d*, *k*) i podlinijnych (np. *g*, *j*, *y*);
- b) stosunek wysokości liter nadlinijnych do wysokości liter śródlinijnych;

- c) stosunek wysokości liter podlinijnych do wysokości liter śródlinijnych;
- d) kąt nachylenia znaków.

Aby tak definiowane zmienne były użyteczne, to pomiary wysokości liter, jak również kątów ich nachylenia wykonane za pomocą programu Graphlog, powinny być rzetelne. Dlatego też sprawdzono powtarzalność i odtwarzalność pomiarów, które zdefiniowano następująco:

- powtarzalność – opisy danej próbki pisma, w formie wysokości liter z danej kategorii lub kąta nachylenia znaków, uzyskane przez jedną osobę w kilku niezależnych seriach pomiarowych, są takie same;
- odtwarzalność – opisy danej próbki pisma, w formie wysokości liter z danej kategorii lub kąta nachylenia znaków, uzyskane przez różne osoby, są takie same.

Ponadto wskazane byłoby, aby wyniki pomiarów wykonanych za pomocą programu Graphlog były dokładne, np. odzwierciedlały rzeczywiste wartości wysokości poszczególnych liter.

3.2. Przygotowanie próbek

W celu weryfikacji dokładności, powtarzalności i odtwarzalności danych uzyskiwanych za pomocą programu Graphlog, przygotowano zestandaryzowane rękopisy nakreślone przez dziesięć różnych osób. Każdy z nich obejmował trzy grupy znaków (śródlinijne, nadlinijne i podlinijne) – po 22 znaki z każdej grupy (rycina 10).

Tak przygotowane dokumenty zeskanowano, a utworzone pliki graficzne przekazano siedmiu osobom, które miały wykonać pomiary. Przed przystąpieniem do pomiarów każda z osób skalowała próbki graficzne we własnym zakresie. Ponadto powtarzalność przeanalizowano również na próbkach pisma rzeczywistego w postaci tekstu w języku polskim, składającego się z kilkuset znaków, nakreślonego przez dziesięć osób.

3.3. Pomiary wykonane ręcznie

W każdym z dziesięciu dokumentów (próbek pisma pobranych od 10 osób) zmierzono metodą klasyczną, tj. za pomocą linijki, wysokość 22 liter nadlinijnych, 22 podlinijnych i 22 śródlinijnych z dokładnością do 0,1 mm. Pomiary takie wykonały trzy osoby będące biegłymi w Instytucie Ekspertyz Sądowych w Krakowie (oznaczone jako A, B, C). Ze względu na trudności w uzyskaniu wiarygodnych i powtarzalnych wyników pomiarów kątów metodą klasyczną (pomiary ręczne), postanowiono nie uwzględniać ich w walidacji, a ocenę dokładności pomiarów wykonanych za pomocą programu Graphlog oprzeć jedynie na wysokości znaków.

3.4. Pomiary wykonane za pomocą programu Graphlog

3.4.1. Próbkki zestandaryzowane

W każdej z 10 zestandaryzowanych próbek pisma zmierzono za pomocą programu Graphlog wysokość 22 liter nadlinijnych, 22 podlinijnych i 22 śródlinijnych z dokładnością do 0,1 mm oraz kąt nachylenia 44 liter nadlinijnych i podlinijnych z dokładnością do 0,1°. Pomiary takie wykonało 7 osób, oznaczonych literami od A do G, w trzech różnych ośrodkach tj. w Instytucie Ekspertyz Sądowych w Krakowie (osoby oznaczone jako A, B i C), w Katedrze Kryminalistyki Uniwersytetu Śląskiego w Katowicach (osoby oznaczone jako D i E) oraz w Katedrze Kryminalistyki Uniwersytetu im. Adama Mickiewicza w Poznaniu (osoby oznaczone jako F i G). Każda z tych osób pięciokrotnie zmierzyła daną próbkę pisma. Ze względu na to, że jedna z osób mierzących próbki nie zastosowała się do określonych zasad pomiaru kątów nachylenia liter, uzyskane przez nią wyniki tej kategorii pomiarów zostały odrzucone.

3.4.2. Próbkki rzeczywiste

Próbki rzeczywiste, które obejmowały dziesięć rękopisów tekstowych nakreślonych naturalnym piśmem zwykłym na specjalnie przygotowanych do tego formularzach, podobnie jak próbki zestandaryzowane, zostały zmierzone przez siedem osób. Zmierzono wszystkie litery podlinijne i odpowiadającą im liczbę liter z pozostałych dwóch kategorii. W przypadku próbek rzeczywistych zrezygnowano z pomiarów kąta nachylenia znaków ze względu na różny poziom czytelności zapisów i związane z tym trudności w precyzyjnym określeniu procedury takiego pomiaru.

4. Rezultaty i dyskusja

4.1. Statystyka opisowa

Przykładem rezultatów uzyskanych w trakcie wszystkich przeprowadzonych pomiarów są wyniki w formie wartości średnich i ich odchyłeń standardowych uzyskanych przez osobę A i przedstawione w tabelach I i II. W tabeli I znajdują się rezultaty pomiarów wysokości liter, które zmierzono ręcznie i za pomocą programu Graphlog w 10 zestandaryzowanych i 10 rzeczywistych próbkach pisma. Dane zebrane w tej tabeli pozwalają również na ocenę dokładności pomiarów wykonywanych za pomocą programu Graphlog, tj. ustalenie, czy wysokości liter uzyskane za pomocą programu odpowiadają ich rzeczywistym wartościom. Wyniki zamieszczone w tabeli II umożliwiają ocenę powtarzalności pomiarów

wysokości i kątów nachylenia liter, które wykonano za pomocą programu Graphlog. Ilustracją odtwarzalności pomiarów są dane zamieszczone w tabeli III, które dotyczą pomiarów wysokości liter śródliniowych wykonanych przez 7 osób w 10 rzeczywistych próbkach pisma za pomocą programu Graphlog.

4.2. Analiza dokładności pomiarów

Można przyjąć, że pomiary wykonane za pomocą programu Graphlog dokładnie odzwierciedlają rzeczywiste wysokości mierzonych liter, gdy wysokość konkretnej i -tej litery uzyskana w trakcie pomiarów wykonanych klasycznie za pomocą linijki i za pomocą programu, są sobie równe. Można to zapisać bardziej formalnie jako:

$$d = x_{hi} - x_{Gi} = 0 \quad \{1\},$$

gdzie: x_{hi} – wysokość i -tej litery (np. śródliniowej) w próbce zmierzona ręcznie; x_{Gi} – wysokość i -tej litery (np. śródliniowej) w próbce zmierzona za pomocą programu Graphlog.

W idealnym przypadku dla każdej litery jej wysokość zmierzona ręcznie i za pomocą programu Graphlog są równe, tzn. uzyskane zostanie $d = 0$ dla każdej z mierzonych liter. Średnia wartość d dla wszystkich zmierzonych liter również będzie równa 0 ($\bar{d} = 0$). Tym samym można postawić hipotezę, że pomiary wykonane za pomocą programu Graphlog są dokładne w przypadku, gdy prawdziwa jest hipoteza, że $\bar{d} = 0$. W realnym przypadku taka sytuacja raczej się nie zdarzy, co obrazują przykładowe dane zaprezentowane w tabeli IV.

Dlatego też sprawdzono, czy wartość $\bar{d} = 0,064$ jest statystycznie różna od zera. W tym celu zastosowano test t-Studenta [9], którym testowano hipotezę zerową w postaci $H_0: \bar{d} = 0$.

Obliczenia wykonano za pomocą funkcji napisanych w programie R [1, 13]. Decyzję o przyjęciu bądź odrzuceniu hipotezy zerowej podjęto w oparciu o obliczone wartości prawdopodobieństw istotności p i porównując je z założonym poziomem istotności α . W badaniach przyjęto $\alpha = 0,05$; poziom ten jest najczęściej używany w analizie za pomocą testów typu test t-Studenta. Wartości p napisane kursywą (tabela V) odpowiadają przypadkom, gdy H_0 należało odrzucić, co oznacza występowanie istotnych statystycznie różnic pomiędzy pomiarami wykonanymi ręcznie i za pomocą programu Graphlog.

Niemniej jednak w większości przypadków przyjęto hipotezę H_0 , co sugeruje, że brak było statystycznie istotnych różnic pomiędzy pomiarami wykonanymi ręcznie i za pomocą programu Graphlog. Można więc uznać, że pomiary wykonane za pomocą programu Graphlog w większości przypadków odzwierciedlały rzeczywiste wysokości mierzonych liter. Przypadki, w których hipotezę

zerową należało odrzucić, związane były najprawdopodobniej z nieprecyzyznością pomiarów dokonanych metodą klasyczną. Wykonano je z dokładnością do 0,1 mm przy szesnastokrotnym powiększeniu mikroskopowym, co mogło spowodować błędne odczytanie wyniku pomiaru liter, zwłaszcza nadliniowych i podliniowych, o znacznym stopniu nachylenia osi.

4.3. Analiza powtarzalności pomiarów

Jak już wspomniano, każda z osób biorących udział w eksperymencie mierzyła pięciokrotnie każdą z 10 próbek, tzn. wyznaczała wysokość 22 liter śródliniowych, 22 liter nadliniowych i 22 liter podliniowych, a także kąt nachylenia liter nadliniowych i podliniowych (łącznie 44 litery). Pomiary można było uznać za powtarzalne gdy dana osoba w każdej z serii pomiarowych uzyskała nie różniące się statystycznie dane.

Przyjmując, że średnia wysokość liter z danej kategorii opisuje wyniki uzyskane przez konkretną osobę w danej serii pomiarowej, to wówczas do oceny powtarzalności wyników można zdefiniować następującą hipotezę – średnie wysokości liter z danej kategorii uzyskane w każdej z pięciu serii pomiarowych są jednakowe:

$$H_0: \bar{x}_{A1,sl} = \bar{x}_{A1,dl} = \bar{x}_{A1,ml} = \bar{x}_{A1,sw} = \bar{x}_{A1,sw} = \{2\},$$

gdzie: $\bar{x}_{A1,sl}$ – średnia wysokość liter śródliniowych obliczona na podstawie wysokości 22 liter zmierzonych przez osobę A w I serii pomiarowej.

Tak skonstruowaną hipotezę zerową testowano testem F Snedecora [9]. Podobnie jak w przypadku analizy dokładności pomiarów, decyzję o przyjęciu bądź odrzuceniu hipotezy zerowej podjęto w oparciu o obliczone wartości prawdopodobieństw istotności p , a następnie porównując je z założonym poziomem istotności $\alpha = 0,05$. Podobne hipotezy skonstruowano i testowano dla:

- każdej z trzech kategorii analizowanych liter (litery śródliniowe, nadliniowe i podliniowe) i każdej z siedmiu osób;
- wyników pomiarów kątów uzyskanych przez sześć osób.

We wszystkich porównaniach uzyskano wartości prawdopodobieństw istotności p większe od 0,05, a w większości przypadków bliskie 1. Tym samym uzyskane wartości p pozwoliły na przyjęcie wszystkich założonych hipotez zerowych (H_0). Można więc stwierdzić, że porównywane średnie były statystycznie równe, a tym samym, że pomiary wykonywane przez jedną osobę za pomocą programu Graphlog umożliwiają uzyskanie powtarzalnych opisów analizowanych próbek.

4.4. Analiza odtwarzalności pomiarów

Przyjęto, że pomiary wykonane za pomocą programu Graphlog można uznać za odtwarzalne, gdy opisy próbki

(średnie wysokości liter w danej kategorii lub średnia wartość kątów nachylenia liter) uzyskane dla konkretnej próbki pisma przez różne osoby są identyczne.

Również w tym przypadku do oceny odtwarzalności zastosowano test F Snedecora, którym testowano hipotezę np. dla liter śródliniowych, że średnie wysokości liter śródliniowych w danej zestandaryzowanej próbie pisma (np. pierwszej) uzyskane przez wszystkie 7 osób są statystycznie równe, co można zapisać w następujący sposób:

$$H_0: \bar{x}_{s_{AI}} \quad \bar{x}_{s_{BI}} \quad \bar{x}_{s_{CI}} \quad \bar{x}_{s_{DI}} \quad \bar{x}_{s_{EI}} \quad \bar{x}_{s_{FI}} \quad \bar{x}_{s_{GI}} \quad \{3\},$$

gdzie: $\bar{x}_{s_{AI}}$ – średnia wysokość liter śródliniowych obliczona ze zmierzonych przez osobę A wysokości 22 liter dla pierwszej próbki pisma w I serii pomiarowej.

Analogiczne hipotezy skonstruowano dla:

- wysokości liter każdej z trzech analizowanych kategorii dla każdej z 10 analizowanych zestandaryzowanych próbek pisma i w ramach każdej z serii pomiarowych;
- kątów nachylenia liter uzyskanych przez 6 osób dla każdej z 10 analizowanych zestandaryzowanych próbek pisma;
- wysokości liter każdej z trzech analizowanych kategorii dla każdej z 10 analizowanych rzeczywistych próbek pisma.

W przypadku analizy odtwarzalności pomiarów wysokości liter za pomocą programu Graphlog dla próbek zestandaryzowanych uzyskane wartości p były większe od założonego poziomu istotności $\alpha = 0,05$, a w większości przypadków bliskie 1. Takie rezultaty uzyskano we wszystkich seriach pomiarowych. Również w przypadku pomiarów wysokości liter wykonanych na próbkach rzeczywistych za pomocą programu Graphlog stwierdzono w większości przypadków wartości p większe od 0,05. Wartości mniejsze od założonego poziomu istotności uzyskano dla liter śródliniowych w próbie 2 ($p = 0,002$) i 5 ($p = 0,039$), dla liter nadliniowych w przypadku próbki 4 ($p = 0,013$), co związane jest najprawdopodobniej z pojedynczymi, przypadkowymi błędami pomiaru, spowodowanymi niedokładnością osoby mierzącej.

Niemniej jednak można powiedzieć, że porównywane średnie są statystycznie równe, a tym samym, że pomiaru wykonywane przez różne osoby za pomocą programu Graphlog umożliwiają uzyskanie odtwarzalnych opisów próbek opartych na wysokościach analizowanych liter.

Odtwarzalność pomiarów kątów nachylenia liter z kategorii liter nad- i podliniowych uzyskana została (tj. przyjęto hipotezy zerowe) dla 7 z 10 przebadanych próbek przy poziomie $\alpha = 0,05$, a dla 9 próbek przy poziomie $\alpha = 0,01$. Brak doskonałej odtwarzalności związany jest z trudnością w precyzyjnym ustaleniu osi znaków w pomiarach kątów nachylenia liter. Mimo to program Graph-

log pozwala uzyskać zdecydowanie lepsze wyniki pomiarów kątów niż metoda klasyczna.

5. Wnioski

Zaprojektowany i wykonany w Instytucie Ekspertyz Sądowych program komputerowy Graphlog jest precyzyjnym, wiarygodnym i prostym w użyciu narzędziem, służącym do analizy cech mierzalnych pisma ręcznego. Pozwala on na uzyskanie w szybki i łatwy sposób dużej liczby różnego rodzaju pomiarów i ich wstępną analizę statystyczną (program automatycznie podaje liczbę wykonanych pomiarów, wartość minimalną i maksymalną, średnią i odchylenie standardowe). Jego dodatkową zaletą jest zmniejszenie liczby błędów, będących skutkiem niedokładnych skal pomiarowych oraz standaryzacji procesu wykonywania pomiarów. Dzięki archiwizowaniu wyników wszystkich pomiarów oraz współpracy programu z bazą danych istnieje możliwość łatwego eksportu danych do innych programów służących bardziej zaawansowanej obróbce statystycznej.

Poprawność działania programu Graphlog sprawdzono, weryfikując dokładność, powtarzalność i odtwarzalność pomiarów. W tym celu przygotowano dwa rodzaje próbek pisma ręcznego (zestandaryzowaną i rzeczywistą), które zmierzone zostały przez siedem osób metodą tradycyjną i przy użyciu programu. Analiza statystyczna kilkunastu tysięcy jednostkowych wyników wykazała dokładność pomiarów wysokości liter, a także powtarzalność pomiarów wysokości liter śródliniowych, nadliniowych i podliniowych oraz kątów nachylenia znaków. Stwierdzono ponadto idealną odtwarzalność pomiarów wysokości liter w próbkach zestandaryzowanych i niemal idealną w próbkach rzeczywistych. Na skutek trudności w precyzyjnym określeniu procedury wyznaczania osi znaków w pomiarach kątów nachylenia liter, odtwarzalność tego parametru uzyskano jedynie w 7 z 10 zestandaryzowanych próbek (przy poziomie $\alpha = 0,05$). Mimo to program Graphlog pozwala uzyskać zdecydowanie lepsze wyniki pomiarów kątów niż metoda klasyczna.

Podsumowując, program Graphlog jest cennym narzędziem przydatnym w pracy ekspertów pisma ręcznego oraz może stanowić dobre narzędzie badawcze przydatne do analizy własności mierzalnych pisma ręcznego w badaniach populacyjnych. Może także przyczynić się do standaryzacji ekspertyz wykonywanych w różnych laboratoriach.