ONLINE SIGNATURE VERIFICATION AND HANDWRITING CLASSIFICATION

Summary of the Ph.D. thesis

Erika Griechisch

Supervisor:
Dr. János Csirik
professor emeritus

Doctoral School of Computer Science
University of Szeged
Institute of Informatics

2018

1. INTRODUCTION

Authentication has always played an important role in our world, since the identification of a person is essential. Generally different types of biometric authentication are applied to determine a person's identity. Figure 1.1 summarizes the most common ones. Some of them have been discovered to be unique in the last few decades, such as DNA and iris recognition and they provide more accurate results than the earlier methods did (e.g. fingerprint, signature). Hence they are more difficult to forge.

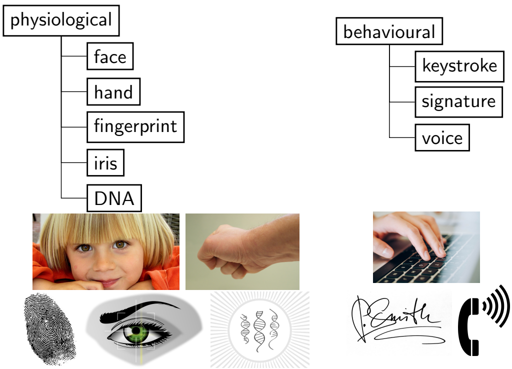


Figure 1.1. Biometrical traits

The uniqueness of handwriting is relative, thus the authentication and verification methods based on handwriting provide less accurate results. Therefore experts and researchers still try to develop better and more accurate methods and devices.

During verification process our main goal is to decide whether a given questioned handwriting or signature belongs to the same person we suspect. Reference handwritings or signatures are always given to compare with.

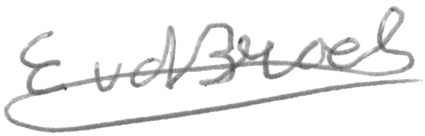
In spite of the fact that, among the forensic experts, forensic handwriting experts have the necessary qualifications and knowledge to verify documents for authenticity or to certify on the basis of a court's request, efforts have been made since the 1970s to examine handwriting by automated way.

In my dissertation, I summarized my research on verification and classification of online signatures and handwritings.

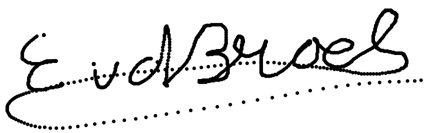
ONLINE HANDWRITING

The recent technological evolution brought several devices which can capture the dynamical aspects of the handwriting process in addition to the image information. Depending on the devices, besides timestamp the x, y coordinates, pressure, inclination are recorded. Handwritings and signatures recorded with such devices are called online handwriting or online signatures. The features applied in verification and classification are provided by the device itself (e.g. x, y coordinates, pressure, inclination of the pen relative to the writing surface). Sometimes accelerometer or gyroscope sensors attached to a pen measures acceleration or angular velocity directly. In addition to the feature provided by the device, we can calculate features such as velocity, acceleration, curvature.

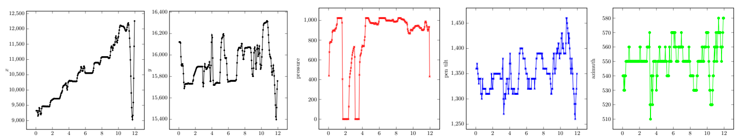
Figure 1.2.a shows an image of a scanned signature, Figure 1.2.b is a generated image, based on the online data recorded during the handwriting process. Below Figure 1.2.c shows the corresponding x, y, pressure and inclination angles.



(a) Offline signature



(b) Online signature



(c) x, y coordinates, pressure, pen tilt and azimuth time functions

Figure 1.2. A signature sample taken from SigComp2009 database

2. CONTRIBUTIONS

SIGNATURE DATASETS

Only few public online signature datasets are available and none of them contains acceleration and angular momentum data. We wanted to examine the usability of the easily accessible accelerometer and gyroscope sensors and create public signature databases for further examination.

AccSigDb

Here we used a ballpoint pen fitted with a three-axis accelerometer to follow the movements of handwriting sessions. We placed the accelerometer very close to the tip of the pen to track the movements as accurately as possible, see Figure 2.1.

The first version of the AccSigDb (AccSigDb1) was collected from 40 subjects between January and March of 2011 [1]. Each subject supplied 10 genuine signatures and 5 simple forgeries, so the dataset contains $40 \cdot 15 = 600$ signatures in total. The signature forgers were asked to produce 5 signatures of another person participating in the study. Each participant supplied forged samples and genuine samples as well.

Figure 2.1. The accelerometer is mounted close to the tip of the pen

Afterwards the dataset was extended between April 2011 and May 2011 [2] (AccSigDb2), and the extended set contained 300 additional signatures. We asked 20 authors who contributed to the first version of the database and repeated the same process with them (10 genuine, 5 forged signatures per person). This extension provided an opportunity to examine the similarities between signatures from the same author captured in two recording periods.

Figure 2.2 shows the scanned images and the reduced acceleration signals of 2 genuine signatures and 2 forged signatures. Subfigure a and b belong to the same author, and they appear quite similar. Subfigure c and d are corresponding forged signatures.

GyroSigDb

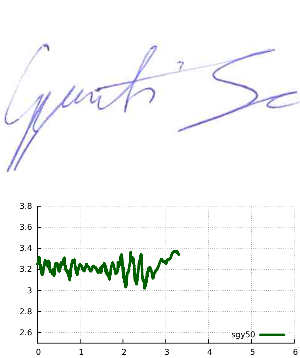
For comparison purposes we decided to try other sensor as well. Therefore we replaced the accelerometer with a 2-axis gyroscope to measure the angular momentum of the pen during the signing.

21 authors contributed to the GyroSigDb, each of them contributed ten signatures, except one, who gave 50 signature samples. Skilled forgeries were recorded from four of them as well.

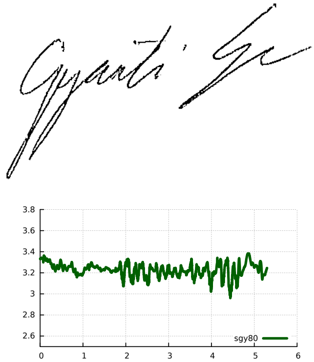
Figure 2.3 shows two signatures and the corresponding signals from the same author as mentioned above. It shows the output voltage of the gyroscope directly. Each row belongs to one signature, the first column (left) shows the signal along the x-axis, the second (right) shows the signal along the y-axis.

Verification [1]

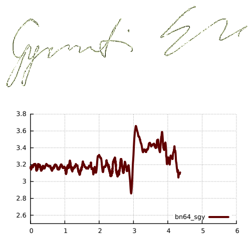
Verification was performed based on DTW distance on the AccSigDb1 dataset. For each writer, 5 genuine signatures were chosen randomly as references. All the other signatures of this writer and skilled forgeries of their signature were used as questioned signatures: 5 genuine and 5 skilled forged signatures for each writer. We compared minimal, maximal and average DTW distances on the test and train dataset. Table 2.1 shows EER, the percentage where the false acceptance and the false rejection rates are equal. The best result (the lowest EER) is achieved, when



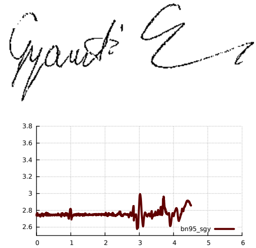
(a) Genuine - 1st time period



(b) Genuine - 2nd time period



(c) Forgery - 1st time period



(d) Forgery - 2nd time period

Figure 2.2. The images and corresponding acceleration signals of two genuine signatures and two forged signatures from AccSigDb1 and AccSigDb2

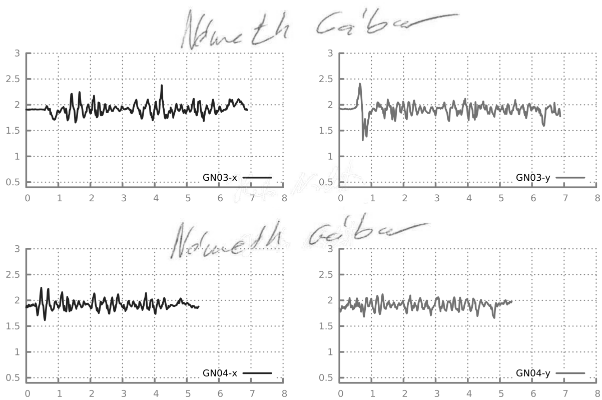


Figure 2.3. Two genuine signatures and the corresponding signals from writer NG (GyroSig, left/right: x/y axis)

we use minimum distance both for the reference and the questioned signatures.

		Test distance		
		average	**maximum**	**minimum**
Training	**average**	14.50% (1.36)	23.50% (0.56)	18.00% (3.34)
	maximum	17.25% (2.02)	29.50% (0.84)	23.25% (4.82)
	minimum	15.50% (0.98)	23.25% (0.38)	13.00% (2.28)

Table 2.1. EER depending on the chosen distance on the reference set and the chosen distance between references and the sample. The values in brackets are the corresponding c multipliers.

Effect of reference selection [2]

After AccSigDb was extended with 300 signatures, further examinations were carried out. We tested how the selection of reference signatures affected the accuracy and whether one can experience changes in signatures recorded with a few months difference.

The FRR for 27 authors (out of 40) and for each possible choice of the 5 reference signatures were 0%, thus the Type I error rate did not depend on the choice of reference signatures in 67.5% of the cases. The average FAR was 14.34%, with a standard deviation of 13.62%; the average FRR was 12.89%, with standard deviation of 24.33% on the whole AccSigDb1. The much higher standard deviation of FRR together with result that FRR was 0% for 27 author, shows that FRR much more depends on the author and has much greater variability.

We found that the performance of a verifier depends largely on the reference set and writer's signature may not vary much over a period of weeks or months, but it can vary more over longer periods.

Comparison of accelerometer and gyroscope [3, 4]

Hereinafter we compared dataset recorded with gyroscope and accelerometer. The gyroscope database contains only a few forged signatures, so instead of verification we performed classification on the available data.

Both dynamical data were represented by a fixed length vector, which were determined as the coefficients of the Legendre approximation and length was varied to examine which accuracy can be achieved.

300 signatures were analysed using SVM classifier and binary and multi-class classification. In all cases, we found that approximation with higher than 20 order, the accuracy does not increase. Multi-class classification with 10 classes provided 43%/52% accuracy (on accelerometric/-gyroscopic data, with order 13), binary classification provided 88%/80.44%. In binary classification accelerometric data clearly outperformed gyroscopic in terms of accuracy, however in multi-class classification gyroscopic data provided slightly higher accuracy.

STATISTICAL DISTANCE IN VERIFICATION [5]

The last contribution achieved in online signature verification was tested on the SigComp2011 Dutch dataset. Our method is based on the Kolmogorov-Smirnov distance (KSD) which is a statistical one. We performed verification and examined the achievable Type I and Type II error rates using KSD on single x, y coordinates, pressure (p) value and absolute velocity (v) features and their combinations. We also examined whether time constraints for the duration length of the signature can improve accuracy, since slower handwriting process often indicates counterfeiting.

Table 2.2 shows the best results. Italic letters show the lowest error rates using a single feature, bold error rates show the best result with combined feature. We observe if the time constraint is applied, the equal error rates decrease about 10%.

		Time constraint	
		with	**without**
p	average	*9.66% / 9.57%*	19.97% / 19.91%
$p \wedge v$		**7.86% / 8.02%**	15.55% / 15.74%
p	max	*9.66% / 9.88%*	23.73% / 23.61%
$p \wedge v$		**7.86% / 8.02%**	16.69% / 16.67%
p	min	*9.00% / 8.95%*	20.13% / 20.22%
$(x \wedge y) \wedge p$		**8.67% / 8.64%**	17.02% / 17.13%

Table 2.2. *FAR* and *FRR* values on the SigComp2011 Dutch online dataset

In several criminal cases it is necessary to narrow down the focus of the investigation, the demographic scope of suspects. If no sample is available for comparison, attempts may be made to narrow the possible suspects. If it is possible to deduce the age (children, young, middle-aged, elderly), gender, handedness, education level can help to find or identify the writer.

With the guidance of the Hungarian Institute for Forensic Science (NIFS) I examined online handwriting samples and classified them by handedness based on forensic-methodology.

In order to support forensic handwriting examination, we have also examined whether it is possible to support examination of handwriting using a statistical approach based on forensic knowledge. We examined whether there is a possibility to determine the gender of the writer based on a large representative database. We examined which revealed features may be typical (i.e. significantly different) among male or female handwritings.

Handedness detection [6]

In collaboration with the Hungarian Institute for Forensic Science (NIFS) we examined online handwritings based on the direction of short horizontal strokes. Earlier forensic studies showed that this feature (the direction of the short horizontal strokes) was a good indicator of the writer's handedness [7, 8] and in contrast to the offline handwriting data, the online handwriting data contains the exact coordinate information, therefore the direction is given.

During forensic handwriting examinations, samples are mostly given in offline format and with microscopic examination the direction of a stroke can not be always assessed. Thus forensic handwriting experts take into account several other features in handedness determination. In forensic examination more and more online handwriting samples have to be analysed, which makes it possible to examine directional features.

Figure 2.4.a represents a right-handed sample with 16 detected left-to-right horizontal strokes (thick lines marked in red). Figure 2.4.b is a

left-handed handwriting sample with 13 detected right-to-left horizontal strokes (thick lines marked in blue).

Based on the earlier studies which found that the direction of crossings (short horizontal lines) were reliable feature in handedness detection, we calculated the number of left-to-right and right-to-left short horizontal strokes. Based on the quantities we compared two approaches.

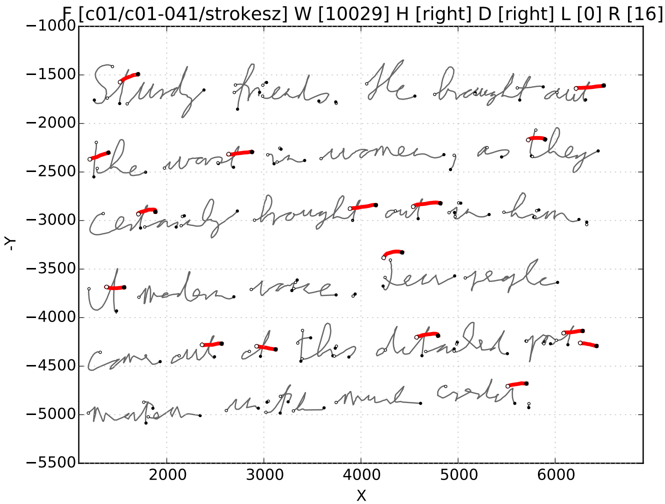
One approach is based on majority voting (MV): if more left-to-right horizontal strokes are present in the sample, than right-to-left, it is classified as right handed, if more right-to-left, it is left-handed. Equal number or less than 2 horizontal strokes resulted in inconclusive decision.

Based on the rigorous forensic observations, right-to-left horizontal strokes were very unlikely to occur if the writer was right-handed. Thus if a right-to-left stroke was found it belonged most likely to a left-handed writer. Using the IAM-onDB dataset which contains samples with 6-7 lines of text, our automatic detection method found 10-20 short horizontal strokes, but - due to the automation - our method detected falsely crossing lines which were sometimes caused by retouching. For this reason, we introduced a less strict criterion that we refer as LC_k: if k or more right-to-left horizontal strokes are detected, the sample is classified as it belongs to a left-handed writer, otherwise the decision is based on the above described majority voting.

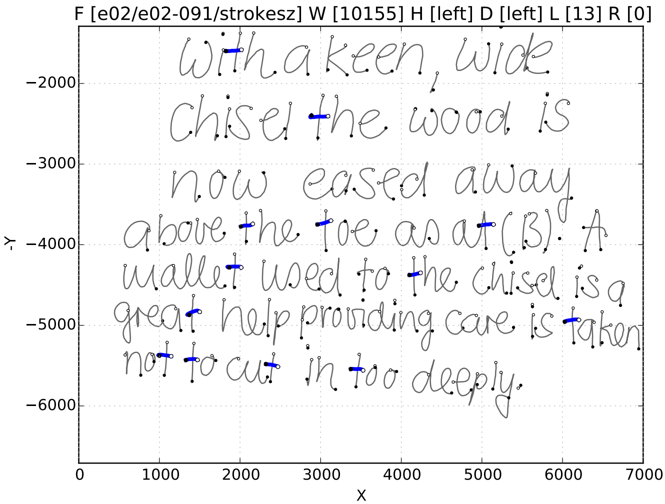
Comparison of male and female handwriting [9]

The biological gender of the writer can be determined with less certainty. The gender of the writer is affected psychologically, biologically, by social stereotypes, but primarily hormonally. Lots of psychological studies examine how reliably non-professionals can determine the gender of the writer [10, 11], which claim that non-professionals can determine the gender of the writer with 75% or higher accuracy.

According to the studies which focus on differences between male and female handwriting, general differences were found: female handwriting is more regular, more bounded, has more starting and ending lines and more rounded; the male handwriting has stronger pressure, their handwritings contain more narrow middle-zone letters, have larger upper zone according to Huber and Headrick [12] and especially their signatures are more illegible according to Mohammed et al. [13].

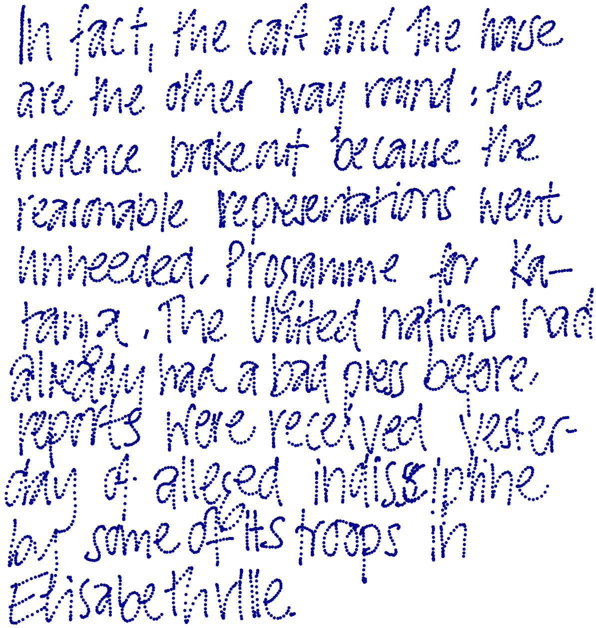


(a) Right-handed writer sample

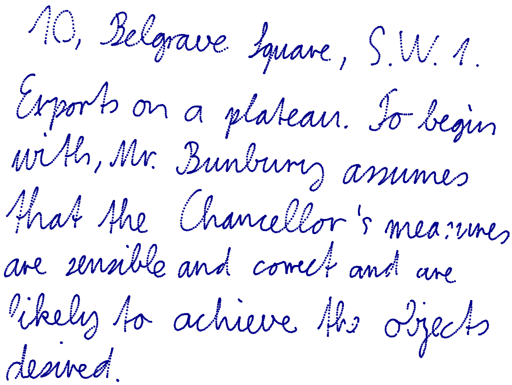


(b) Left-handed writer sample

Figure 2.4. Detection of horizontal strokes (sufficient number of strokes)



(a) Female (10191): Standard letters and even placement of the handwriting more frequently occur in female handwriting



(b) Male (10218): Narrow letters, the bigger size of the upper zone, larger capital letters more frequently occur in male

Figure 2.5. Handwriting samples from IAM-OnDB

We used the publicly available IAM-OnDB online handwriting dataset with writer genderness information and filtered the dataset on the basis of criteria of representative samples.100 samples were selected, from which 50 were written by female, and 50 by male. Based on the examination of the Hungarian Institute for Forensic Science, which revealed 2500 features (25 features on 100 samples), such as proportion of the capital letters, characteristics of starting and ending lines, form of the lines (see Figure 2.6) and the automatically measured continuous features (like velocity, left-margin, first line indent) we could carry out statistical tests. Discrete variables were compared with Fisher's exact tests, continuous variables with Welch's independent t-tests.

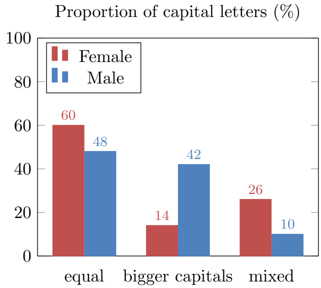
The main significant differences were in evenness of the placement of handwritings, form of lines, coordination, zone proportions, proportions of capital letters, structure of the letters, direction and form of starting and ending lines, consistency of the cross bar. Other special significant features were extracted, such as arcade in starting lines which did not occur in male, but occurred in 16% of the female handwriting. At the same time there was no female handwriting without ending line nevertheless its proportion in male writing was 8%.

Class	**RH**	**LH**	**Total**
correct	1085+173 (80.23%)	41+49 (5.74%)	1348 (85.97%)
incorrect	2+7 (0.57%)	10+12 (1.40%)	31 (1.98%)
inconclusive	149 (9.50%)	40 (2.55%)	189 (12.05%)
Total	1416 (90.31%)	152 (9.69%)	1568 (100.0%)

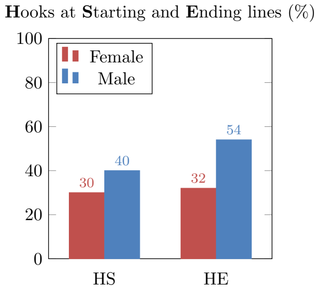
Table 2.3. Summary of results – MV

	RH	**LH**	**Total**
correct	1085+126 (77.23%)	41+58 (6.31%)	1310 (83.55%)
incorrect	2+54 (3.57%)	10+6 (1.02%)	72 (4.59%)
inconclusive	149 (9.50%)	37 (2.36%)	186 (11.86%)
Total	1416 (90.31%)	152 (9.69%)	1568 (100.00%)

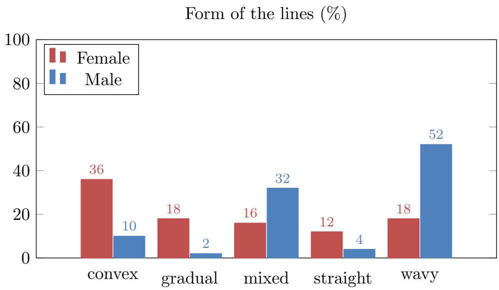
Table 2.4. Summary of results – LC_2



(a) Proportion of capital letters



(b) Hooks



(c) Form of the lines

Figure 2.6. Barplots from each feature grou

Comparing continuous features significant differences were found in velocity means and velocity medians ($p = 0.0339$ and $p = 0.01116$ respectively), but no significant differences were found between the means of first line indent and the left margin.

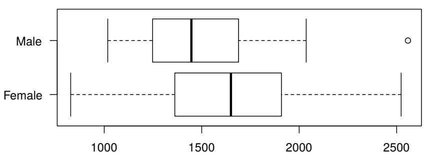


Figure 2.7. Distribution of velocity means

3. SUMMARY

My PhD dissertation includes several studies in online handwriting analysis. The first thesis group consists of studies in online signature verification and classification, while the second thesis group focuses on online handwriting classification and analysis based on features applied by Forensic Handwriting Experts.

My main contributions to these studies are the followings:

1. I recorded two online signature datasets based on data measured by 3D accelerometer and 2D gyroscope. I evaluated the applicability of the databases in verification using baseline verifiers and classification and examined how training data selection effects the results .

   Based on the results and an online signature verification method implementation which is based on the often used DTW, the accelerometric is applicable in online signature verification and analysis, however the accuracy does not reach the accuracy of the recently available systems on the market which are more expensive, but due to the more accurate measurement with advanced hardware, more sophisticated methods can be applied in the verifier able to detect forgeries.

   During the analysis of the database recorded with the accelerometer, I found that the selection of the test dataset greatly influences the verification results.

   Signature classification was performed on the signature samples recorded by accelerometer and gyroscope. Features extracted using Legendre approximation and classification carried out with SVM showed that approximation with higher than 20 order, the accuracy did not increase. Multi-class classification with 10 classes provided 43%/52% accuracy (on accelerometric/gyroscopic data, with order 13), binary classification provided 88%/80.44%. In binary classification accelerometric data clearly outperformed gyroscopic in terms of accuracy, however in multi-class classification gyroscopic data provided slightly higher accuracy.

2. I introduced an online signature verification method based on Kolmogorov-Smirnov statistical distance

   The method was tested on the SigComp2011 Dutch online dataset, thus x, y coordinates and pressure data were available. I added velocity feature and examined which single feature and feature combinations were most suitable for signature verification purposes. The developed method is competitive to the competing methods of the SigComp2011 competition and I found that the primarily pressure, secondarily the velocity features proved to be the most suitable features in online signature verification.

3. With the guidance of the Hungarian Institute for Forensic Science (NIFS) I examined online handwriting samples and classified them by handedness based on forensic-methodology.

   Previous results of forensic handwriting examinations showed that horizontal lines (e.g. crossings) were the most reliable feature to distinguish left-handed and right handed handwriting, I automatically detected horizontal strokes and by taken to account their directions, made conclusion about the handedness of the handwriting samples. The results obtained are competitive to the available methods. The actual error rates compared to the results of other methods was lower, with the newly introduced inconclusive class, which does not give conclusion about handedness if insufficient information is available to make decision. Thus, based on the majority voting (MV), we could achieve a 1.98% error rate for 1568 handwritings with 12.05% inconclusive cases.

4. I examined handwriting differences of several discrete and continuous features comparing the writers' biological gender using statistical tests

   Based on the examination of the Hungarian Institute for Forensic Science, which revealed 2500 features and automatic measurement of velocity-related features carried out by my software, I analyzed these extracted features using statistical tests (Fisher's exact test for discrete variables and Welch t-test for continuous), examining which features might be suitable to distinguish male and female handwriting.

Publications by the author, related to theses

Publication	Thesis point			
	I. group		II. group	
	1	2	3	4
[1]	•			
[2]	•			
[3]	•			
[4]	•			
[5]		•		
[6]			•	
[9]				•

BIBLIOGRAPHY

[1] Horst Bunke, János Csirik, Zoltán Gingl, and Erika Griechisch. Online signature verification method based on the acceleration signals of handwriting samples. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, CIARP*, pages 499–506. Springer, 2011. doi: 10.1007/978-3-642-25085-9_59.

[2] János Csirik, Zoltán Gingl, and Erika Griechisch. The effect of training data selection and sampling time intervals on signature verification. In *CEUR Workshop Proceedings*, volume 768, pages 6–10, Beijing, 2011.

[3] Erika Griechisch, Muhammad Imran Malik, and Marcus Liwicki. Online signature verification using accelerometer and gyroscope. In *Proceedings of the 16th Conference of the International Graphonomics Society*, IGS 2013, pages 143–146, 2013.

[4] Erika Griechisch, Muhammad Imran Malik, and Marcus Liwicki. Online signature analysis based on accelerometric and gyroscopic pens and Legendre series. In *Proceedings of the 12th International Conference on Document Analysis and Recognition*, ICDAR 2013, pages 374–378, Washington, DC, 2013. IEEE. doi: 10.1109/ICDAR.2013.82.

[5] Erika Griechisch, Muhammad Imran Malik, and Marcus Liwicki. Online signature verification based on Kolmogorov-Smirnov distribution distance. In *Proceedings of International Conference on Frontiers in Handwriting Recognition*, ICFHR 2014, pages 738–742. IEEE, 2014. ISBN 9781479943340. doi: 10.1109/ICFHR.2014.129.

[6] Erika Griechisch and Erika Bencsik. Handedness detection of online handwriting based on horizontal strokes. In *Proceedings of the 13th International Conference on Document Analysis and Recognition*, ICDAR 2015, pages 1272–1277. IEEE Computer Society, 2015. ISBN 9781479918058. doi: 10.1109/ICDAR.2015.7333953.

[7] Marianne Conrad. Left-hand writing vs. right-hand writing. *Journal of the American Society of Questioned Document Examiners*, 11(1):19–

27, 2011. Presented at ENFHEX Conference Modern Developments in Handwriting Examination, Vilnius, September 20–22, 2007.

[8] Vaibhav Saran, Suneet Kumar, AK Gupta, and Syeed Ahmad. Differentiation of handedness of writer based on their strokes and characteristic features. *Journal of Forensic Research*, 4(204):2, 2013.

[9] Erika Bencsik and Erika Griechisch. The frequency of occurence of handwriting features in online male and female handwriting. In *Proceedings of IGS2017 (18th International Graphonomics Society Conference)*, IGS 2017, pages 169–172, 2017.

[10] William N Hayes. Identifying sex from handwriting. *Perceptual and motor skills*, 83(3):791–800, 1996.

[11] Vivien Burr. Judging gender from samples of adult handwriting: Accuracy and use of cues. *The Journal of social psychology*, 142(6): 691–700, 2002.

[12] Roy A Huber and Alfred M Headrick. *Handwriting identification: facts and fundamentals*. CRC press Boca Raton, 1999.

[13] Linton Mohammed, Bryan James Found, Douglas Kelman Rogers, et al. Frequency of signatures styles in San Diego County. *Journal of the American Society of Questioned Document Examiners*, 2008.