Offline Handwriting Acquisition under Controlled and Uncontrolled Conditions

Linda Alewijnse
Netherlands Forensic Institute
Department of Digital Technology and Biometrics
The Hague, The Netherlands
l.alewijnse@nfi.minvenj.nl

***Abstract*—This paper gives a description of offline handwriting acquisition under controlled and uncontrolled conditions for research purposes. The data collection task is an underestimated part in the process of developing signature verification or handwriting identification systems. There is a continuous need for new, unpublished data to train and evaluate new algorithms. Handwriting samples that make up the current publicly available databases have all been collected under controlled conditions. However, good quality data is still limited.**

On the contrary, research databases constituted of case related biometric data in general are scarce. To suit forensic purposes, it is preferred to start building databases with forensically relevant data. When verification and identification systems are trained on this type of material, the output will be more suited for forensic examination purposes. The challenges in this area are considered.

***Keywords*—offline data, data collection, signature verification, forensic handwriting examiner**

I. Introduction

Signature verification is a biometric technique with promising results for the near future for implementation within the forensic handwriting examination. In the past 10 years rapid developments are made within the pattern recognition discipline [1]. Implementing analysis tools in the forensic practice is the next challenge. Before an automated signature verification or handwriting identification system can be implemented, the forensic community must be ascertained that the systems are trained, evaluated and validated by correct environmental conditions.

Collecting and selecting handwriting samples for research purposes is often an underestimated task. The number of publicly available databases with handwriting is limited, so new data must be collected regularly. Data are primarily collected to provide information regarding a specific topic. Therefore, data must be in accordance with the objective of the study. The overall performance of a biometric technology is eventually influenced by the quality of the input data.

A. Learning from the past

The following example illustrates the importance of sample design and sample selection to suit the purpose of the study. In 2002, Srihari and colleagues [2] conducted a study to test the principle of individuality of handwriting. Handwriting samples were collected from 1500 individuals. The dataset was representative for the US population with respect to gender, age, ethnicity, handedness, etc. The automated system CEDAR-FOX was used to evaluate the handwriting, and could identify the writer of a particular sample with 98 percent confidence. Inferring these statistics over the entire U.S. population, writer identification can be established with 96 percent confidence.

Saks [3] commented on this study by arguing that to test individuality, a better sampling design would have been to gather a representative sample of clusters of writers, with each cluster composed of highly similar writers. Only then, the data would have been discriminative of highly similar handwriting. And it would have been repeatable if the same effect was observed between the clusters. The choice of data by Srihari and colleagues was not adequate for testing the hypothesis that handwriting is individual.

In a response to this, Durina and colleagues [4] conducted a study in which samples of writing were obtained from 52 writers and their teachers who were taught the same copybook style at the same Catholic elementary school approximately 4 decades ago. The research addressed the criticisms that earlier studies on the individuality of handwriting did not include populations from homogeneous writing communities. It demonstrated that there is a high degree of inter-writer variation among writers, even in populations where the driving forces for variation are low. In spite of the size of the dataset, it was better fit for purpose to investigate the uniqueness of handwriting.

B. Learning from each other

In the past years, from 2009 until 2013, different datasets with signatures as well as handwriting are collected by the Netherlands Forensic Institute for the Signature Competition (SigComp) [5]. This competition allows researchers and practitioners from academia and industries to compare performance on signature verification on new and unpublished datasets. Because all participating parties in the competition are provided with the same data, results are comparable. While the competition provides an overview of involved parties and shows the performance of the available systems to the forensic community, the pattern recognition researchers are more concerned about which features are most discriminative. The SigComp provides a platform to bridge the gap between the two communities.

Two years ago, in 2011, a group of researchers from different fields of expertise started the discussion about how to bridge the gap between the two communities and to signal the challenges. Computer programmers learned how a forensic handwriting examination is carried out and examples of real casework are described. Forensic scientists got an overview of state-of-the-art automatic verification systems. Recent advances are comparing the performance with Minimum Cost of Log Likelihood Ratios [6], the task of reporting a probabilistic output score, and the addition of disguised signatures in new datasets. Nevertheless, much work needs still to be done in order of bringing together researchers in the field of automated handwriting analysis and signature verification and experts from the forensic handwriting examination community.

The scope of the competition changes each year. In the end, when automated systems are meant to aid the FHE in the examination or as an objective tool. The first competition was focused on skilled forgeries. After that, disguised signatures were added to the questioned signatures. Last year we've provided different scripts, i.e. Dutch and Chinese signatures. The consequence of the changing focus of the competition allows the developers to improve their algorithms and benefit from new and unpublished handwriting data.

II. OBJECTIVE

Three scenarios for handwriting data collection can be distinguished: 1) The samples are collected under controlled conditions, e.g. let the participants write on the same make of paper, with the same writing instrument, in similar writing position, etc., 2) spontaneous writings are collected from participants by gathering their writings from the past, and 3) forensic handwriting samples from casework are shared, either anonymously or by an online evaluation platform.

Topics that are covered in this paper are:

- offline and online data
- requirements of the dataset
- controlled versus uncontrolled conditions
- research data versus forensic data

The first part of the paper describes the most favorable and pragmatic approach for offline handwriting sample collection. The second part stresses the importance of data collection under uncontrolled conditions. Furthermore, this paper calls for exploring the possibilities of using forensic datasets to further develop automated systems.

III. METHOD

Two categories capturing a person's handwriting can be distinguished, namely, offline and online. The online modality is discussed here very shortly, because this data is not available to the forensic handwriting examiner. It is useful for biometric identification and finding the new features or feature combinations that are most discriminative. Handwriting examiners will in particular be interest in offline systems and therefore offline data acquisition is described more in detail.

A. *Online data*

Online data collection requires an electronic writing tablet and recording software. Most often WACOM tablets are used to collect handwriting samples, but since pen-input devices getting more widespread this might change on short term. The online handwriting is captured with an electronic writing tablet and stored digitally in x, y, and z-positions as a function of time.

B. *Offline data*

Offline handwriting data is a representation of the handwriting in as a scanned image. It has been demonstrated [7] the FHE's can infer dynamic information, such as writing velocity and pen pressure, from the static trace. Writing velocity is reflected in line quality, pen pressure differences and blunt beginnings and endings of stroke. The pen pressure is not useful for the examiner as an absolute measure, since it is not only writer specific but strongly depends on extrinsic factors. It is only writer specific if other conditions such as writing surface and writer instrument are constant. The indentation of the paper shows the handwriting examiner if the ink was deposited by a natural course of writing or by forced writing.

For offline data collection all that is needed is a pen, a piece of paper and a scanner. To aid the writer, a guiding line or box can be used. The easiest and practical solution is to use an underlying sheet of paper with the lineation or boxes printed with a black, bold line. No lineation or bounding boxes must strike trough the writings. In this way, the data is kept 'clean' and less effort for data preparation is needed.

Fig. 1 Offline specimen signatures collected under controlled conditions.

C. *Data requirements*

The requirements for a high-quality offline dataset of handwritten samples are summed up below. A formal data collection process is necessary as it ensures that gathered data are both defined and accurate and that decisions based on arguments embodied in the findings are valid [7].

The first list proposed shows which requirements of the dataset are advised for training and evaluating automated systems. Additionally there is a list of extra requirements which are important for forensic handwriting researchers. The summed information is necessary for forensic handwriting examiners to get a better understanding of the data used in experiments. In general, the data must reflect the variation of handwriting in the relevant population, and intra-writer variation must represent reality.

Pattern recognition data requirements:

- Substantial number of specimen writers
- Substantial number of simulators
- High resolution scans of the written samples, preferably 400 dpi.
- Suitable format (PNG format would be preferable. This lossless format will retain information from images when re-opened and re-saved. The PNG format also creates smaller file size but without the quality loss of a GIF-file)
- Cropping of the image
- Assign an identification code as filename
- Compatibility with earlier collections

Additional forensic requirements:

- Writer sex, age, handedness, level of education, and profession
- Cultural origin (for signatures) or copybook system (for handwritten text)
- Substantial amount of questioned writing (e.g. half a page of text)
- Substantial amount of reference writing (number of reference signatures or number of lines of text)
- Specification of conditions of forgery and/or disguised
- Time span over which the data was collected

IV. Forensic Handwriting Data

A. *Collecting existing specimens*

One way of acquiring relevant data is to collect existing writings. Such handwriting can consist signatures on agreements, receipts, cheques, passports, etcetera. In short, it can comprise handwriting, which is comparable to the reference material in casework. All factors that are considered by forensic handwriting examiners are in the dataset: natural variation in the writings, different surfaces, different writing instruments, different time period and the samples are written under different mental circumstances. Both intrinsic and extrinsic factors are represented. Participants are not approached to write something, but provide the researcher with their previously written material.

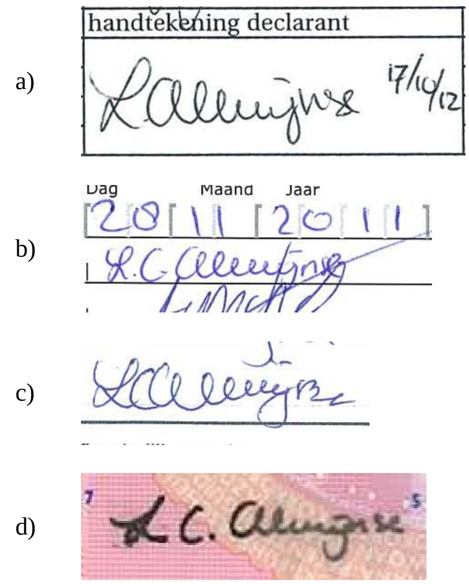

Fig. 2 Examples of collected specimen signatures written under uncontrolled conditions: a) A signature that was written under a declaration form, b) two overlapping signatures with restricted space for signing, c) signature on a receipt that was written in a standing writing position, and d) signature on an ID-document, dating from 5 years ago.

B. *Case related data*

The best would be using forensic casework data to evaluate and validate automated systems, but legal aspects regarding privacy form an obstacle. One possible solution for sharing forensic samples is to facilitate access at an online evaluation platform. BEAT [8] is a project that is funded by the European Commission, under the Seventh Framework Programme and is offering such an approach. The goal of the project is to propose a framework of standard operational evaluations for biometric technologies. Unfortunately, it is not available for forensic biometrics yet.

Simulated data can be used in the training phase of system development, because the ground truth of the origin is known. The evaluation phase should at least contain case related data. However, the validation of the system should completely be performed with real casework samples.

V. Conclusion and Discussion

Where biometric systems usually have access to high quality and uniform data, in forensic practice the trace under investigation is often characterized by poor quality. This is not represented by the currently existing handwriting databases.

Since input data determines the overall performance of the automated system, a next step in bridging the gap between the pattern recognition community and forensic handwriting examiners should logically involve the use of samples that were written under uncontrolled circumstances. The condition of the dataset has its effect on the systems' performance on that trace and accordingly influences the strength of the evidence.

REFERENCES

[1] M. Caligiuri and L. Mohammed, "The Neuroscience of Handwriting: Applications for Forensic Document Examination," CRC Press, 2012.

[2] S.N. Srihari, S-H Cha, H. Arora, and S. Lee, "Individuality of handwriting", J Forensic Sci, vol. 47(4), pp. 856—872, 2002.

[3] M. Saks, Authors' Response in the J Forensic Sci, vol. 48(4), July 2003.

[4] M.E. Durina and M.P. Caligiuri, "The Determination of Authorship from a Homogenous Group of Writers," Journal of ASUDE, vol. 12, nr. 2, 2010.

[5] M.I. Malik, M. Liwicki, L. Alewijnse, and W. Ohyama, "ICDAR2013 Competitions on Signature Verification and Writer Identification for On- and Offline Skilled Forgeries (SigWiComp2013)," in press.

[6] M. Liwicki et al., "Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011)," in Document Analysis and Recognition (ICDAR), 2011 International Conference on, pp. 1480-1484, 2011.

[7] D. Meuwly and R.N.J. Veldhuis, "Forensic biometrics: From two communities to one discipline," IEEE Conference publications BIOSIG 2012, Darmstadt Germany, pp. 1-12, Sep 2012.

[8] Information available at www.beat-eu.org