Writer Identification and Verification

Chapter · January 2008

DOI: 10.1007/978-1-84628-921-7_13

1 author:

Lambert Schomaker
University of Groningen

Writer identification and verification

Lambert Schomaker

Artificial Intelligence Dept., University of Groningen,
Grote Kruisstr. 2/1, 9712 TS, Groningen, The Netherlands,
schomaker@ai.rug.nl

Summary. The behavioral biometrics methods of writer identification and verification are currently enjoying renewed interest, with very promising results. This chapter presents a general background and basis for handwriting biometrics. A range of current methods and applications is given, also addressing the issue of performance evaluation. Results on a number of methods are summarized and a more in-depth example of two combined approaches is presented. By combining textural, allographic and placement features, modern systems are starting to display useful performance levels. However, user acceptance will be largely determined by explainability of system results and the integration of system decisions within a Bayesian framework of reasoning that is currently becoming forensic practice.

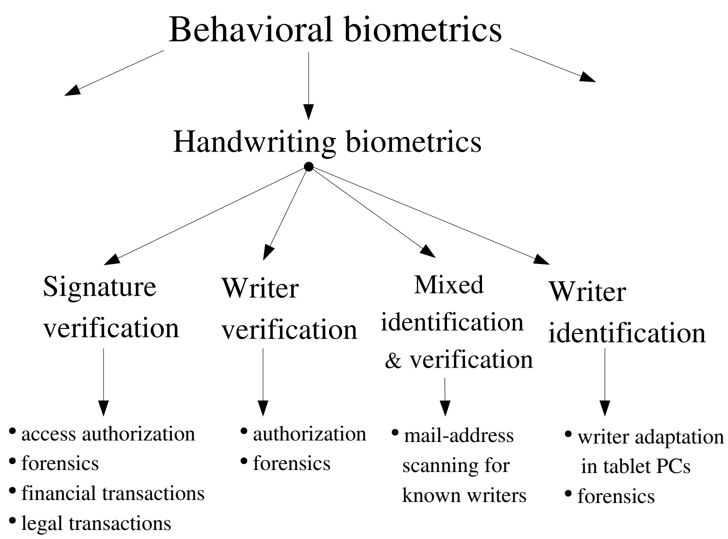
1 Introduction

Writer identification and verification belong to the group of behavioral methods in biometrics. Contrary to biometrics with a purely physical or biophysical basis, the biometric analysis of handwriting requires a very broad knowledge at multiple levels of observation. For the identification of a writer in a large collection of known samples on the basis of a small snippet of handwriting, multi-level knowledge must be taken into account. In forensic practice, many facets are considered, ranging from the physics of ink deposition (Franke and Rose, 2004) to knowledge on the cultural influences on the letter shapes in a writer population (Schomaker and Bulacu, 2004). Recent advances have shown that it is possible to use current methods in image processing, handwriting recognition and machine learning to support forensic experts. Due to the difficulty of the problem in terms of (a) variability and variation of handwritten patterns, (b) the limited amount of image data and (c) the presence of noise patterns, this application domain still heavily relies on human expertise, i.e., with a limited role for the computer. In this chapter, basic properties of the human writing process will be introduced in order to explore the possible basis for the use of handwritten patterns in biometrics. Subsequently, current algorithms will be described. A distinction is made between fully interactive, semi-automatic and autonomous methods. Within these three major groups, different types of information and shape features are being used. From bottom to top, the computed features may involve ink deposition patterns, low-level image features and texture, stroke-order information, character-shape styles (allographic variation) and layout, up to spelling and interpunction peculiarities in script samples. A challenge for any computer-based method in forensic handwriting analysis is the degree to which search and comparison results can be conveyed in a manner which is comprehensible to the human users, i.e., the detectives, lawyers and judges. Not all algorithms are equally suitable for the derivation of a verbal account of the quantitative analyses. Fortunately, the use of Bayesian reasoning, which is essential in current pattern recognition, is also becoming acceptable in the application domain. The chapter concludes with an outlook on the application of methods in forensic practice.

2 A special case of behavioral biometrics

Ratha et al. (2001) make a distinction between three person-authorization methods: (a) authorization by possession of a physical token); (b) possession of knowledge and (c) biometrics, i.e., the science of verifying the identity of a person based on physiological of behavioral characteristics. In behavoural biometrics, features of an action sequence are computed such that the identity of an individual can be verified (1:1 comparison) or correctly established from a given list (1:N). Whereas other forms of biometrics concentrate on *what you know* (password), *what you possess* (key), *how you are* (fingerprint), behavioral biometrics concentrates on *how you behave*. Typical behaviours that lend themselves to individual characterization are speech (Markowitz, 2000), eye movements (Kasprowski and Ober, 2004; Andrews and Coppola, 1999), gait (BenAbdelkader et al., 2004), key-stroke statistics in computer use (Gutiérrez et al., 2002), 'mouse'-pointing behavior during internet use (Gamboa and Fred, 2003), signatures (Jain et al., 2002) and handwritten text patterns (Srihari et al., 2002; Schomaker and Bulacu, 2004). Furthermore, individuals differ in spelling idiosyncrasies and interpunction statistics when producing texts. This behavioral characteristic is exploited in forensic linguistic-stylistic analysis (McMenamin, 2001). The stochastic variation of human behavioral features is evidently much larger than is the case in DNA analysis or iris-based identification (Daugman, 2003). The degree to which two instances of a behavioral pattern by the same actor are similar is subject to the intrinsic noise in the human motor system (Van Galen et al., 1989). Additionally, the observed behaviour, e.g., the signing of a document, is heavily influenced by a wide range of context factors (Franke, 2005). Still, behavioral biometrics can be very useful in a number of applications. behavioral biometrics are usually not invasive, and may provide intrinsic evidence of the presence of a living individual, as is the case in on-line signature verification from a digitizer pad. Table 1 gives an overview on types of **handwriting biometrics**, i.e., methods which use characteristics of shape, trajectory (kinematics), force (kinetics) and/or pen-tilt changes of handwriting to model an individual writer as a unique signal source.

Table 1. Types of handwriting biometrics and application areas



In this chapter, we will mainly focus on writer identification and verification. Here, unlike the case of the handwritten signature, it is not guaranteed that the content of a questioned sample and another sample of known identity are identical. Furthermore, in writer identification and verification, there may exist large differences in the amount of written ink over samples, ranging from a few isolated words in a threat message to complete diaries. For some applications, the distinction between identification and verification may be blurred. As an example, consider the case where a stream of mail items is scanned for the presence of a limited number of suspect writers. In this case, one would like an algorithm both to come up with the most likely identity in case of a match and at the same time provide a measure of verification, such that unreliable matches can be ignored and the false-alarm rate can be limited.

3 A basis for handwriting biometrics

A good biometric (Jain et al., 2000) is universal, unique, permanent and collectable: Each person should possess the characteristics (universality); No two persons should share the characteristics (uniqueness); The characteristics should not change over time (permanence); and it should be easily presentable to a sensor and be quantifiable (collectability). Handwriting is not fully universal, since there exists still a considerable proportion of non-writing individuals in the population. Uniqueness can be only determined empirically: assuming a natural writing attitude, there is more individual information in handwriting than is generally assumed, especially if one has enough text per sample. Handwriting does change over time: both for signature verification and handwriting identification and verification, regular enrollment procedures over the years will be required. Samples can be collected on-line, using a special digitizer which samples the pen-tip position and possibly other movement-related signals in time. Alternatively, existing samples of handwriting (off-line) can be scanned. In the forensic application domain, samples are sometimes obtained at the moment when an individual becomes a suspect in criminal investigation. The subject may be asked to produce a given sample of text. Although 'off line', such a procedure can be considered as interactive. Obviously, the most critical issue is uniqueness, to which the next sections will be devoted. Fig. 1 shows four factors causing variability in handwriting Schomaker (1998).

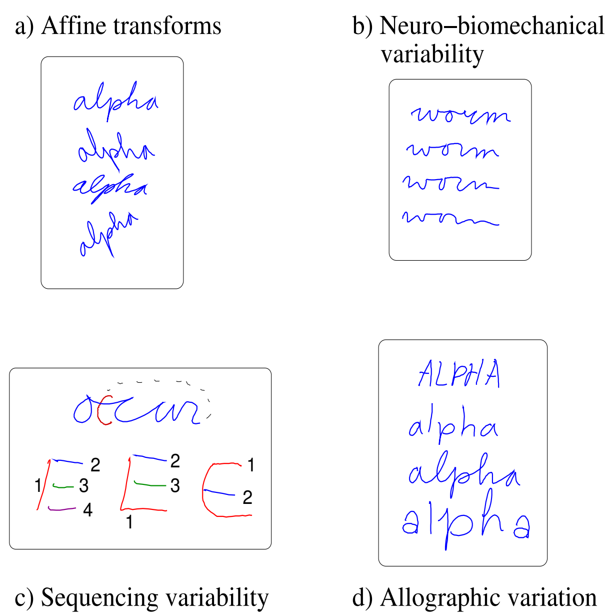


Fig. 1. Factors causing handwriting variability: (a) Affine transforms are under voluntary control. However, writing **slant** constitutes a habitual parameter which may be exploited in writer identification; (b) neuro-biomechanical variability refers to the amount of effort which is spent on overcoming the low-pass characteristics of the biomechanical limb by conscious cognitive motor control; (c) sequencing variability becomes evident from stochastic variations in the production of the strokes in a capital *E* or of strokes in Chinese characters, as well as stroke variations due to slips of the pen; (d) allographic variation refers to individual use of character shapes. Factors b) and c) represent system state more than system identity. In particular, **allographic variation** (d), is a most useful source of information in forensic writer identification.

The first factor concerns the **affine transforms** (Fig. 1a), which are under voluntary control by the writer. Transformations of size, translation, rotation and shear are a nuisance but not a fundamental stumbling block in handwriting recognition or writer identification. In particular *slant* (shear) constitutes a habitual parameter determined by pen grip and orientation of the wrist subsystem versus the fingers (Dooijes, 1983).

The second factor concerns the **neuro-biomechanical variability** (Fig. 1b) which is sometimes referred to as ”sloppiness space”: the local context and physiological state determines the amount of effort which is spent on character-shape formation and determines the legibility of the written sample. In realizing the intended shape, a writer must send motor-control patterns which compensate for the low-pass filtering effects of the biomechanical end-effector. This category of variability sources also contains tremors and effects of psychotropic substances on motor-control processes in writing. As such, this factor is more related to system state than system identity.

The third factor is also highly dependent on the instantaneous system state during the handwriting process and is represented by **sequencing variability** (Fig. 1c): the stroke order may vary stochastically, as in the production of a capital *E*. A four-stroked *E* can be produced in $4! * 2^4 = 384$ permutations. In the production of some Asian scripts, such as Hanzi, stochastic stroke-order permutation are a well-known problem in handwriting recognition (even though the training of stroke order at schools is rather strict). Finally, spelling errors may occur and lead to additional editing strokes in the writing sequence, post hoc, e.g., after finishing a sentence. Although sequencing variability is generally assumed to pose a problem only for handwriting recognition based on temporal (on-line) signals, the example of post-hoc editing (Fig. 1c) shows that static, optical effects are an expected consequence of this source of variability.

The fourth factor, **allographic variation** (Fig. 1d), refers to the phenomenon of writer-specific character shapes, which produces most of the problems in automatic script recognition but at the same time provides the information for automatic writer identification.

3.1 Nature and Nurture

There exist two fundamental factors contributing to the individuality of script, i.e., allographic variation: *genetic* (biological) and *memetic* (cultural) factors.

The first fundamental factor consists of the *genetic* make up of the writer. Genetic factors are known or may be hypothesized to contribute to handwriting style individuality:

- The biomechanical structure of the hand, i.e., the relative sizes of the carpal bones of wrist and fingers and their influence on pen grip;
- The left or right handedness Francks et al. (2003);
- Muscular strength, fatiguability, peripheral motor disorders (Gulcher et al., 1997);
- Central-nervous system (CNS) properties, i.e., aptitude for fine motor control and the neural stability in performing motor tasks (Van Galen et al., 1993).

The second factor consists of *memetic* or culturally transferred influences (Moritz, 1990) on pen-grip style and the character shapes (allographs) which are trained during education or are learned from observation of the writings of other persons. Although the term *memetic* is often used to describe the evolution of ideas and knowledge, there does not seem to be a fundamental objection to view the evolution and spreading of character shapes as a memetic process: the fitness function of a character shape depends on the conflicting influences of (a) legibility and (b) ease of production with the writing tools (Jean, 1997) which are available within a culture and society. The distribution of allographs over a writer population is heavily influenced by writing methods taught at school, which in turn depend on factors such as geographic distribution, religion and school types. For example, in the Netherlands, allographic differences may be expected between protestant and catholic writers, writers of different generations, and immigrant writers.

Together, the genetic and memetic factors determine a habitual writing process, with recognizable shape elements at the local level in the writing trace, at the level of the character shape as a whole and at the level of character placement and page layout. In this paper, we will focus on the local level in the handwritten trace and on the character level.

The writer produces a pen-tip trajectory on the writing surface in two dimensions (x,y), modulating the height of the pen tip above the surface by vertical movement (z). Displacement control is replaced by force control (F) at the moment of landing. The pen-tip trajectory in the air between two pen-down components contains valuable writer-specific information, but its shape is not known in the case of off-line scanned handwritten samples. Similarly, pen-force information is highly informative of a writer's identity, but is not directly known from off-line scans (Schomaker and Plamondon, 1990). Finally, an important theoretical basis for the usage of handwritten shapes for writer identification is the fact that handwriting is not a feed-back process which is largely governed by peripheral factors in the environment. Due to neural and neuro-mechanical propagation delays, a handwriting process based upon a continuous feed-back mechanism alone would evolve too slowly (Schomaker, 1991). Hence, the brain is continuously planning series of ballistic movements ahead in time, i.e., in a feed-forward manner. A character is assumed to be produced by a "motor program" (Schmidt, 1975), i.e., a configurable movement-pattern generator which requires a number of parameter values to be specified before being triggered to produce a pen-tip movement yielding the character shape (Schomaker et al., 1989; Plamondon and Maarse, 1989; Plamondon and Guerfali, 1998) by means of the ink deposits (Doermann and Rosenfeld, 1992; Franke and Grube, 1998). Although the process described thus far is concerned with continuous variables such as displacement, velocity and force control, the linguistic basis of handwriting allows for postulating a discrete symbol from an alphabet to which a given character shape refers.

4 Methods and applications

Table 2. Methods in handwriting biometrics

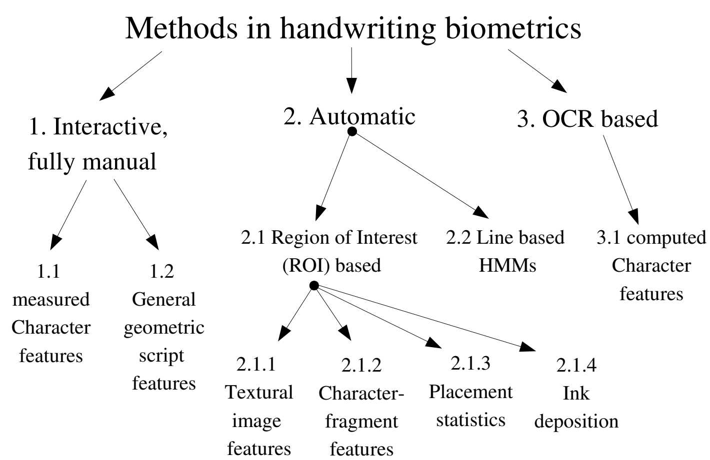


Table 2 provides an overview over current methods. Current systems in forensic practice (Fish/BKA, Script/TNO, Cedar-Fox) rely heavily on human-expert interaction (Table 2, **1.**). The user selects relevant, small image elements and/or provides a manual measurement on these. The measured features are either character based (**1.1**) or concern general geometric script features (**1.2**) such as word and line spacing as well as the overall slant angle of handwriting. Manual methods are very costly and there is a risk of subjective influences on the measurement process, which is usually realized with a mouse on a graphical user interface or using a digitizer tablet. From the point of view of computer science, it seems appropriate to focus on fully automatic methods (**2.**), at least at the level of feature measurement itself. The user may coarsely select a region of interest (ROI) with a sufficient amount of handwritten ink, from which handwriting features are the automatically computed and compared between samples (**2.1**). Alternatively, a system may try to automatically segment a handwritten text into lines, using Hidden-Markov Models (Schlapbach and Bunke, 2004) for comparison between samples (**2.2**). A third group (Table 2, **3.**) consists of systems which are 'semi-automatic' in the sense that characters need to be selected in a manual process. Subsequently, character-based features are computed which were specifically designed for use in handwritten OCR systems (Srihari and Shi, 2004). For the image-based (ROI) methods (**2.1**), a number of features haven been described in literature. A group of features is designed to focus on textural aspects of the script (**2.1.1**), i.e. angular histograms (Maarse et al., 1988; Schomaker et al., 2003) for slant-related characteristics and angular-combinations histograms (Bulacu and Schomaker, 2007) for both slant and curvature aspects of the script texture. Here, also the use of Gabor features (Said et al., 1998) has been proposed. The features in this group capture, indirectly, pen-grip and pen attitude preferences. In order to automatically get at learned or preferred use of allographic elements, an alternative group of featurs focuses on character or character-fragment shapes (**2.1.2**). Bensefia et al. (2003) uses size-normalized connected-component images. Alternatively, contour-based features have been proposed for upper case script (Schomaker and Bulacu, 2004), as well as for fragmented connected-component contours in miscellaneous free styles (Schomaker et al., 2004). The third group of features (**2.1.3**) is focused on the writer's placement (layout) preferences in the process of placing ink elements horizontally or vertically across the page. A classical paper proposes to use white-pixel run-length histograms (Arazi, 1977). Other features in this group are autocorrelation and Fourier descriptors. Although less performant than texture or allographic features, the placement-related features have been shown to boost overall system performance if used in combination with other features. At the micro level (**2.1.4**), features can be derived from the results of the ink-deposition process. Such stroke-morphology features (Franke and Rose, 2004) are especially useful for script samples written with a ball-point pen. A brush feature (Schomaker et al., 2003)) was introduced to characterize pen-lift and pen-down stroke tails. In order to obtain satisfactory results, feature or classifier-combination methods (Kittler et al., 1998) are required, due to the complexity of the problem, the usually limited image quality and the multi-faceted nature of handwritten patterns. In order to prevent cascaded training for both classification and combination stages, it is advisable to use simple, sparse-parametric combination or voting schemes. Untrained and/or population-based methods are more convenient than schemes which require writer-specific training.

4.1 Performance evaluation

In the case of writer identification, performance evaluation is usually done on the basis of cumulative ”Top-N” correct-writer identification proportions, i.e., Top-1 performance indicating the proportion of correctly identified writers at the top of a returned hit list, Top-10 performance indicating the probability of finding the correct writer within the best ten returned identification matches, and so on. For the case of many samples (instances) in a database, also the traditional performance measures of from *Information Retrieval* can be used, i.e., precision and recall. Precision concerns the proportion of correct-identity instances in the hit list. Recall concerns the proportion of correct-identity instances relative to the total number of instances in the data base from the sought writer.

In case of writer verification, the common measures in binary-decision problems are used: false-acceptance rate (FAR), false-reject rate (FRR) and receiver-operating curve (ROC). For a distance (or similarity) measure x, the distributions $P_{genuine}(x)$ for same-writer distances and $P_{impostor}(x)$ for different-writer distances can be counted, evaluating large number of pairwise verifications. Subsequently, by varying a threshold

θ for accepting or rejecting claimed identities, the following cumulative probability distributions can be obtained, for the case of a distance measure x:

$$FAR(\theta) = \int_0^{\theta} P_{impostor}(x)dx \quad (1)$$

and

$$FRR(\theta) = \int_{\theta}^{\infty} P_{genuine}(x)dx \quad (2)$$

Thus, for each value of θ the probability of falsely accepting an impostor or the probability of falsely rejecting a genuine sample of handwriting from the claimed identity can be determined. The overall performance of a handwriting-verification system can be characterized by the equal-error rate ($EER = FAR = FRR$). A more robust quality measure is the area under the ROC curve (AUC). Varying the threshold θ, the ROC curve represents the series pairwise points where the y value equals the proportion of correctly verified sample pairs and the x value equals the corresponding false-acceptance rate. The estimation of FAR, FRR and AUC is not trivial in the case of jagged probability distributions of distances in $P_{genuine}$ and $P_{impostor}$. In such a case it is advisable to obtain more reference data and/or apply a different distance measure. There is a fundamental relation between the non-parametric Wilcoxon test for pairwise data and the AUC. Recent work presents a comparison of traditional to newly identified confidence bounds for the AUC (Agarwal et al., 2005) as a function of the number of samples.

5 Two facets of handwriting: texture and allographs

In this section examples will be given of two recent approaches which have proved fruitful. The first method is based on directionality and curvature of patterns in handwriting, constituting a textural feature of the handwritten image. Angles and curvature in handwriting are to a large extent determined by the degrees of freedom in wrist and finger movement, which in turn depend on the pen grip attitude and the applied grip forces. A stiff hand-pen system will allow for less variation in directions than a compliant agile pen grip. The second method is based on a convenient way of breaking up handwritten patterns and comparing them to a 'code book' of known shapes. In this approach, the fraglets of ink composing together the preferred allographs for a writer are counted and their histogram is taken as characteristic for the writer as a stochastic selector of shapes during the writing process.

5.1 Texture

In order to capture the curvature of the ink trace, which is very typical for different writers,it is informative to observe the local angles along the edges. In its simplest form, the histogram of angles along the edges of a handwritten image may be considered (Maarse et al., 1988). However, while such a feature will capture the variations around the average slant, it will miss the curvature information. Therefore, a related feature with some additional complexity was proposed. The central idea is to consider the **two** edge fragments emerging from a central pixel and, subsequently, compute the joint probability distribution of the orientations of the two fragments of this 'hinge'. The final normalized histogram gives the joint probability distribution $p(\phi_1, \phi_2)$ quantifying the chance of finding in the image two "hinged" edge fragments oriented at the angles ϕ_1 and ϕ_2 respectively. The orientation is quantized in 16 directions for a single angle, as before. From the total number of combinations of two angles we will consider only the non-redundant ones ($\phi_2 > \phi_1$) and we will also eliminate the cases when the ending pixels have a common side. Therefore the final number of combinations is $C(2n, 2) - 2n = n(2n - 3)$. The edge-hinge feature vector will have 464 dimensions (16 directions considered). A more detailed description of the method can be found elsewhere (Bulacu et al., 2003). A further refinement can be obtained by computing the hinge feature separately for the upper and the lower half of handwritten lines ('splitHinge').

5.2 Allographs

In the application domain of forensic handwriting analysis, there is a strong focus on allography, i.e., the phenomenon that for each letter of the alphabet, shape variants or allographs exist. Individuals display a preference to produce a particular subset of all the allographs emanating from the population of writers, due to schooling and personal preferences. Figure 2 illustrates that allographic style does not concern isolated characters, per se. It is very common that writers emit families of comparable shapes for components of handwriting such as the ascenders and descenders in the given example. Not all writers are equally consistent in their use of allographs, but given a few lines of text there may be enough information to estimate occurrences of shapes.

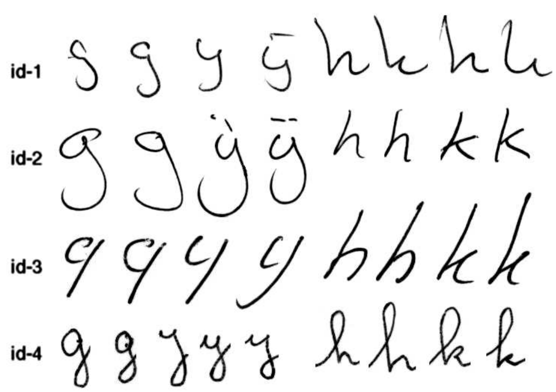


Fig. 2. The letters *g,y,h,k* produced by four Dutch anonymous subjects (id1-id4) from a forensic data collection. The development of handwriting style will often entail the choice for homogeneous style elements, as illustrated by the corresponding ascender and descender shapes within each of these writers.

In order to replace the common human manual character segmentation by automatic algorithms one would seem to need a system which is able to segment allographs out of the image and compute their histogram of occurrence for a questioned sample. However, there exists no exhaustive and world-wide accepted list of allographs in, e.g., Western handwriting. The problem then, is to generate automatically a codebook, which sufficiently captures allographic information in samples of handwriting, given a histogram of the usage of its elements. Since automatic segmentation into characters is an unsolved problem, we would need, additionally, a reliable method to segment handwritten samples to yield components for such a codebook. It was demonstrated that the use of the shape of connected components of upper-case Western handwriting (i.e., not using allographs but the contours of their constituting connected components) as the basis for codebook construction can yield high writer-identification performance (Schomaker and Bulacu, 2004). This approach was dubbed $p(CO^3)$ for 'probability of COnnected-COmponent COntours'. A later study showed that this success could be replicated for the case of mixed-styles connected cursive by segmenting handwritten patterns into appropriate fraglets and tracing the resulting fragmented connected-component contours to obtain a histogram $p(FCO^3)$ for each sample of handwriting. The resulting fraglets will usually be of character size or smaller. Sometimes a fraglet will contain more than one letter. In the next section a number of results will be presented on the 'hinge' feature, and the connected-component codebook approach, alongside with results on other systems and feature groups. Although comparability is a difficult topic, the performances give a good indication of what is achievable today.

6 Recent results

Table 3 gives an overview on a number of recent studies on writer identification and verification. Comparison between methods is very difficult. There are large differences as regards the number of parameters and the effort spent in system training as well as regards the required amount of text. Furthermore, some systems just require the selection of a region of interest (ROI) whereas other systems require detailed manual character or word segmentation and/or manual measurements. The most convenient methods would require no training at the level of individual writers and just a few lines of written text of which a rectangle is cut out from the context in a graphical user interface.

Table 3. Writer-identification performances for a number of systems. Training modes are: w=writer model, p=population model, n=no training.

Method/Feature	*Nwriters*	*Top-1* (%)	*Top-10* (%)	EER (%)	Train mode	Match	Style	Reference
Misc. line features	20	91	-	-	w	MLP	misc	Marti et al. (2001)
'SysA'	100	34	90	-	w	-	misc	Schomaker and Bulacu (2004)
'SysB'	100	65	90	-	w	LDA	misc	Schomaker and Bulacu (2004)
character models	100	-	-	0.9	w	HMM	misc	Schlapbach and Bunke (2006)
co3	150	72	93	-	p	1NN	UPP	Schomaker and Bulacu (2004)
brush	250	53	81	-	n	1NN	misc	Schomaker et al. (2003)
splitEdge	250	29	69	-	n	1NN	misc	Bulacu and Schomaker (2003)
splitAla	250	64	86	-	n	1NN	misc	Bulacu and Schomaker (2003)
splitHinge	250	79	96	-	n	1NN	misc	Bulacu and Schomaker (2003)
fco3	900	76	92	5.8	p	1NN	misc	Bulacu and Schomaker (2006)
hinge+fco3+runl	900	79	96	3.3	p	1NN	misc	Bulacu and Schomaker (2006)
misc. features	1500	96	-	3.5	w	1NN	misc	Srihari et al. (2002)

7 Explainable results

There are a number of stumbling blocks in user acceptation of automatic methods for writer identification and verification. In forensic handwriting research, experts have developed a personal and individual knowledge on handwriting. This expertise is partly perceptual and is difficult to verbalize, partly cognitive and explainable to others: colleagues in the forensic domain, criminal investigators, lawyers and judges. However, there is no common language on shapes and styles of handwriting. Fortunately, in many countries, forensic handwriting examiners are tested on a regular basis. Still, little is known on human versus machine performance in identification and verification, as the results of these tests are often confidential. In the United States, a famous legal case (US Supreme Court, 1993) provides criteria for the admission of handwriting-related expertise in court. These non-exclusive criteria include: testing and peer review of experts and methods, determination of the rates of error, the existence of standards and reference data and general acceptance in the relevant scientific field. Evaluation of these criteria will assist in the determination of whether a particular form of evidence is reliable. In its current form and practice, handwriting biometrics does not fulfill these criteria, similar to the field of voice identification. However, research in computer-based handwriting biometrics will help to improve on this matter (Srihari et al., 2002).

Current technologies can be separated into black-box versus explicit, knowledge-based approaches. Typical black-box systems, use textural, global image features and distance or similarity-based matching. Sometimes, multi-layer perceptrons, support-vector machines or hidden-Markov models are used 'under the hood' of such a system. Performance evaluation occurs on the basis of calibrated probability distributions for false-acceptance rate (FAR) and false-reject rate (FRR). The output of such a system consists of a yes/no decision reported together with a confidence measure (verification) or a hit list with similarity scores and/or confidence measures. Although, technically speaking, black-box systems may be very advanced they cannot answer important questions such as: "Why is the decision: *similar*?" or "Why does individual X appear at the top of the hit list?". Such questions may be answerable in the near future. The advantage of allographic-matching approaches (Srihari and Shi, 2004; Niels et al.) is that in the coming years, matching results may be translated to a verbal report, referring to the set of allographs found for a writer, their probability of occurrence in the population and their conditional probabilities occurring within one writer. Since the judicial system is now getting used to Bayesian reasoning, it will be possible in the near future to automatically generate human-legible reports including likelihood ratios for the decisions made by the system. Apart from using standardized allograph lists, other types of knowledge concerning style clusters, the text content and writer-related facts (Franke et al., 2004) may be integrated in a Bayesian framework. It is of utmost importance for the field that reports include the margin of error, both for human and machine expertise. Even if binary answers can only be given with some margin of error, handwriting biometrics may provide valuable information which would be unwise to ignore in criminal investigation and unwise to withhold in legal cases. Finally, it must be noted that the amount of broad expertise which human experts can spend on a given questioned handwritten document is unsurpassable with current techniques: There is more to forensic handwriting analysis than verification and identification.

References

S. Agarwal, T. Graepel, R. Herbrich, S. Har-Peled, and D. Roth.
Generalization bounds for the area under the roc curve.
Journal Machine Learning Research, 6:393–425, 2005.
ISSN 1533-7928.

T. Andrews and D. Coppola.
Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments.
Vision Research, 39:2947–2953, 1999.

B. Arazi.
Handwriting identification by means of run-length measurements.
IEEE Trans. Syst., Man and Cybernetics, SMC-7(12):878–881, 1977.

C. BenAbdelkader, R. Cutler, and L. Davis.
Gait recognition using image self-similarity.
EURASIP Journal on Applied Signal Processing, 2004(4):572–585, 2004.

A. Bensefia, T. Paquet, and L. Heutte.
Information retrieval-based writer identification.
In *7th International Conference on Document Analysis and Recognition (ICDAR 2003), 3-6 August 2003, Edinburgh, Scotland, UK*, pages 946–950. IEEE Computer Society, 2003.
ISBN 0-7695-1960-1.

M. Bulacu and L. Schomaker.
Writer style from oriented edge fragments.
In *Proc. of the 10th Int. Conference on Computer Analysis of Images and Patterns*, pages 460–469, 2003.

M. Bulacu and L. Schomaker.
Combining multiple features for text-independent writer identification and verification.
In *Proc. of 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2006)*, pages 281–286, 2006.
23 - 26 October, La Baule, France.

M. Bulacu and L. Schomaker.
Text-independent writer identification and verification using textural and allographic features.
IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 29(4):701–717, April 2007.
Special Issue - Biometrics: Progress and Directions.

M. Bulacu, L. Schomaker, and L. Vuurpijl.
Writer identification using edge-based directional features.
In *Proc. of ICDAR 2003*, pages 937–941. IEEE Computer Society, 2003.

J. Daugman.
The importance of being random: statistical principles of iris recognition.
Pattern Recognition, 36(2):279–291, 2003.

D. Doermann and A. Rosenfeld.
Recovery of temporal information from static images of handwriting.
In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 162–168, 1992.

E. Dooijes.
Analysis of handwriting movements.
Acta Psychologica, 54:99–114, 1983.

C. Francks, L. DeLisi, S. Fisher, S. Laval, J. Rue, J. Stein, and A. Monaco.
Confirmatory evidence for linkage of relative hand skill to 2p12-q11.
American Journal of Human Genetics, 72:499–502, 2003.

K. Franke.
The influence of physical and biomechanical processes on the ink trace - Methodological foundations for the forensic analysis of signatures.
PhD thesis, University of Groningen: Artificial Intelligence Institute, November 2005.

K. Franke and G. Grube.
The automatic extraction of pseudo-dynamic information from static images of handwriting based on marked gray value segmentation.
Journal of Forensic Document Examination, 11:17–38, 1998.

K. Franke and S. Rose.
Ink-deposition model: The relation of writing and ink deposition processes.
In *IWFHR '04: Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04)*, pages 173–178. IEEE Computer Society, 2004.
ISBN 0-7695-2187-8.

K. Franke, L. Schomaker, C. Veenhuis, C. Taubenheim, I. Guyon, L. Vuurpijl, M. van Erp, and G. Zwarts.

Wanda: A generic framework applied in forensic handwriting analysis and writer identification.
In *Proceedings of the 9th IWFHR, Tokyo, Japan*. IEEE Computer Society, 2004.
H. Gamboa and A. Fred.
A user authentication technic using a web interaction monitoring system.
In *Pattern Recognition and Image Analysis*, LNCS, pages 246–254. Springer Verlag, 2003.
ISBN 978-3-540-40217-6.
J. Gulcher, P. Jonsson, A. Kong, and et al.
Mapping of a familial essential tremor gene, fet1, to chromosome 3q13.
Nature Genetics, 17(1):84–87, 1997.
F. Gutiérrez, M. Lerma-Rascón, S.-G. L.R., and F. Cantu.
Biometrics and data mining: Comparison of data mining-based keystroke dynamics methods for identity verification.
In *MICAI '02: Proceedings of the Second Mexican International Conference on Artificial Intelligence*, pages 460–469, London, UK, 2002. Springer-Verlag.
ISBN 3-540-43475-5.
A. Jain, F. Griess, and S. Connell.
On-line signature verification.
Pattern Recognition, 35:2963–2972, Dec 2002.
A. Jain, L. Hong, and S. Pankanti.
Biometric identification.
Commun. ACM, 43(2):90–98, 2000.
ISSN 0001-0782.
G. Jean.
Writing: The story of alphabets and scripts.
Thames and Hudson Ltd., 1997.
P. Kasprowski and J. Ober.
Eye movements in biometrics.
In *European Conference on Computer Vision*, LNCS, pages 248–258. Springer Verlag, 2004.
J. Kittler, M. Hatef, R. Duin, and J. Matas.
On combining classifiers.
IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 20(3):226–239, March 1998.
F. Maarse, L. Schomaker, and H.-L. Teulings.
Automatic identification of writers.
In G. van der Veer and G. Mulder, editors, *Human-Computer Interaction: Psychonomic Aspects*, pages 353–360. Springer, New York, 1988.
J. Markowitz.
Voice biometrics.
Communications of the ACM, 43(9):66–73, 2000.
ISSN 0001-0782.
U. Marti, R. Messerli, and H. Bunke.
Writer identification using text line based features.

In *Proc. 6th ICDAR*, pages 101–105. IEEE Computer Society, 2001.
G. McMenamin.
Style markers in authorship studies.
Forensic Linguistics, 8(2):93–97, 2001.
E. Moritz.
Replicator-based knowledge representation and spread dynamics.
In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 256–259. The Institution of Electrical Engineers, 1990.
R. Niels, L. Vuurpijl, and L. Schomaker.
Automatic allograph matching in forensic writer identification.
Int. Journal of Pattern Recognition and Artificial Intelligence.
R. Plamondon and W. Guerfali.
The generation of handwriting with delta-lognormal synergies.
Biological Cybernetics, 78:119–132, 1998.
R. Plamondon and F. Maarse.
An evaluation of motor models of handwriting.
IEEE Trans. Syst. Man Cybern, 19:1060–1072, 1989.
N. Ratha, A. Senior, and R. Bolle.
Automated biometrics.
In *ICAPR '01: Proceedings of the Second International Conference on Advances in Pattern Recognition*, pages 445–474, London, UK, 2001. Springer Verlag.
ISBN 3-540-41767-2.
H. Said, G. Peake, T. Tan, and K. Baker.
Writer identification from non-uniformly skewed handwriting images.
In *British Machine Vision Conference*, pages 478–487, 1998.
A. Schlapbach and H. Bunke.
Using HMM based recognizers for writer identification and verification.
In *IWFHR '04: Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04)*, pages 167–172, Washington, DC, USA, 2004. IEEE Computer Society.
A. Schlapbach and H. Bunke.
Off-line writer identification using gaussian mixture models.
In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, pages 992–995, Washington, DC, USA, 2006. IEEE Computer Society.
R. Schmidt.
A schema theory of discrete motor skill learning.
Psychological Review, 82:225–260, 1975.
L. Schomaker.
Simulation and recognition of handwriting movements: A vertical approach to modeling human motor behavior.
PhD thesis, University of Nijmegen, NICI, The Netherlands, 1991.
L. Schomaker.
From handwriting analysis to pen-computer applications.
IEE Electronics Communication Engineering Journal, 10(3):93–102, 1998.

L. Schomaker and M. Bulacu.
Automatic writer identification using connected-component contours and edge-based features of upper-case western script.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(6):787–798, 2004.

L. Schomaker, M. Bulacu, and K. Franke.
Automatic writer identification using fragmented connected-component contours.
In *Proceedings of the 9th IWFHR, Tokyo, Japan*, pages 185–190. IEEE Computer Society, 2004.

L. Schomaker, M. Bulacu, and M. Van Erp.
Sparse-parametric writer identification using heterogeneous feature groups.
In *IEEE International Conference on Image Processing*, page (in press), 2003.

L. Schomaker, A. Thomassen, and H.-L. Teulings.
A computational model of cursive handwriting.
In . M. S. R. PLamondon, C.Y. Suen, editor, *Computer Recognition and Human Production of Handwriting*, pages 153–177. World Scientific, 1989.

L. R. B. Schomaker and R. Plamondon.
The Relation between Pen Force and Pen-Point Kinematics in Handwriting.
Biological Cybernetics, 63:277–289, 1990.

S. Srihari, S. Cha, H. Arora, and S. Lee.
Individuality of handwriting.
Journal of Forensic Sciences, 47(4):1–17, July 2002.

S. Srihari and Z. Shi.
Forensic handwritten document retrieval system.
In *Proc. of the First International Workshop on Document Image Analysis for Libraries (DIAL'04)*, page 188, Washington, DC, USA, 2004. IEEE Computer Society.
ISBN 0-7695-2088-X.

US Supreme Court.
Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579.
125 L.Ed.2d 469, 1993.

G. Van Galen, R. Van Doorn, and L. Schomaker.
Effects of motor programming on the power spectral density function of writing movements.
Journal of Experimental Psychology, 16:755–765, 1989.

G. P. Van Galen, J. Portier, B. C. M. Smits-Engelsman, and L. Schomaker.
Neuromotor noise and poor handwriting in children.
Acta Psychologica, 82:161–178, 1993.