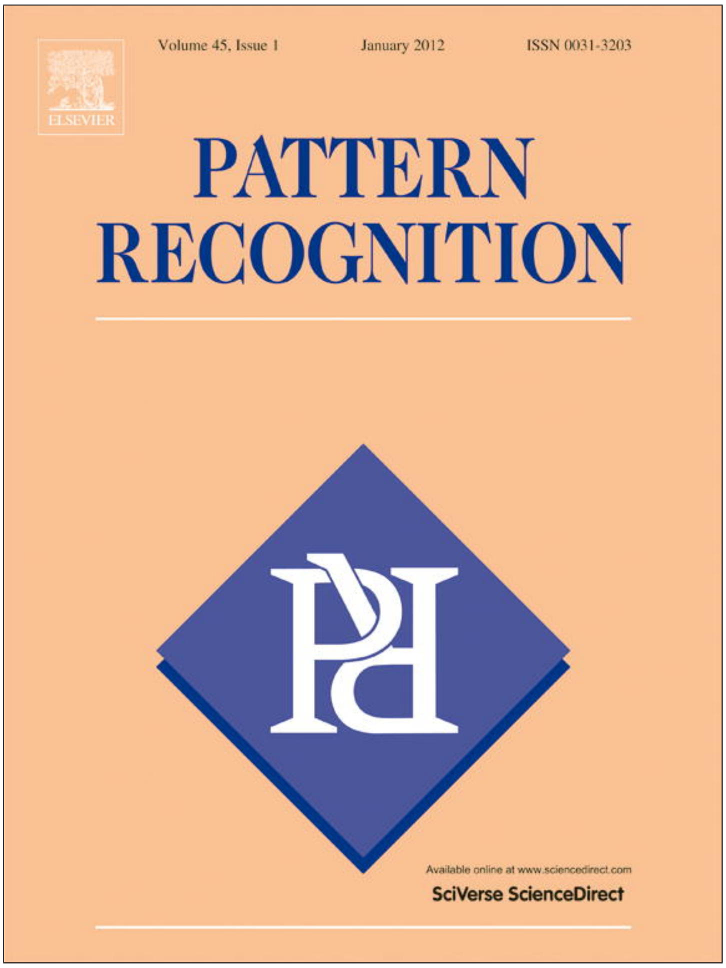
Volume 45, Issue 1
January 2012
ISSN 0031-3203
ELSEVIER
PATTERN
RECOGNITION
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Pattern Recognition 45 (2012) 162–171

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Writer identification using directional ink-trace width measurements

A.A. Brink [a,*], J. Smit [b], M.L. Bulacu [a], L.R.B. Schomaker [a]

[a] *Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands*
[b] *Department of Medieval History, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 12 October 2010
Received in revised form
11 February 2011
Accepted 5 July 2011
Available online 18 July 2011

Keywords:
Behavioral biometrics
Feature extraction
Handwriting analysis
Ink trace width
Medieval handwriting
Off-line
Paleography
Quill
Writer identification

ABSTRACT

As suggested by modern paleography, the width of ink traces is a powerful source of information for off-line writer identification, particularly if combined with its direction. Such measurements can be computed using simple, fast and accurate methods based on pixel contours, the combination of which forms a powerful feature for writer identification: the *Quill* feature. It is a probability distribution of the relation between the ink direction and the ink width. It was tested in writer identification experiments on two datasets of challenging medieval handwriting and two datasets of modern handwriting. The feature achieved a nearest-neighbor accuracy in the range of 63–95%, which even approaches the performance of two state-of-the-art features in contemporary-writer identification (*Hinge* and *Fraglets*). The feature is intuitive and explainable and its principle is supported by a model of trace production by a quill. It illustrates that ink width patterns are valuable. A slightly more complex variant of *Quill*, *QuillHinge*, scored 70–97% writer identification accuracy. The features are already being used by domain experts using a graphical interface.

1. Introduction

Historical handwriting written with a quill has a salient calligraphic appearance. The variability of the trace width, as illustrated in Fig. 1, is caused by physical properties of the writing instrument and the individual writing style of the writer. In *quantitative paleography* [1], a recent methodology in the manual study of such historical documents, writing hands are discerned based on measurable characteristics. Two of the characteristics form the motivation for this paper: *contrast*, which is the difference of width between the thinnest and thickest traces, and the *angle of writing*, which describes the habitual orientation of the pen tip, determined by the angle between the thinnest ink traces and the base line. These characteristics suggest that trace width is an important feature for writer identification and that it is relevant to relate trace width to the trace direction. Trace width and direction can both be determined automatically using simple contour-based image processing operations, and in this paper we will show that the combination of the two yields a powerful feature for automatic writer identification. It is not limited to historical handwriting since modern handwriting contains trace subtle width variations as well.

The value of directionality measurements for writer identification is known [2–5], but the added value of width measurements is new. Only one remotely related feature for writer identification has been evaluated that is based on *run lengths* of black pixels [6,7]. Apart from this approach, to the best of our knowledge, ink trace-width measurements have not been used for writer identification. However, such measurements have been evaluated for a few other applications. One study used width measurements for stroke detection and structural analysis [8]. A second experiment included a distribution of coarse trace-width measurements in a signature verification experiment [9]. In another study on signature verification, the trace width was used to express the mismatch between two signatures at corresponding locations [10].

Our approach is different in that width measurements are combined with direction measurements, and used for writer identification. The resulting statistic will be used as a text-independent writer identification feature, called *Quill*. It consists of a combination of simple methods, including trace-width computation using a method based on Bresenham's well-known line-drawing algorithm [11]. *Quill* depends on pen properties and individual movement style. In this paper, the power of such computational measurements of ink trace widths relative to the writing direction as a feature for writer identification is explored on various handwriting datasets. We will show that its power to

* Corresponding author. Tel.: +31 6 14855404.
E-mail addresses: a.a.brink@ai.rug.nl (A.A. Brink), J.Smit1@uva.nl (J. Smit), bulacu@ai.rug.nl (M.L. Bulacu), schomaker@ai.rug.nl (L.R.B. Schomaker).

0031-3203/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2011.07.005

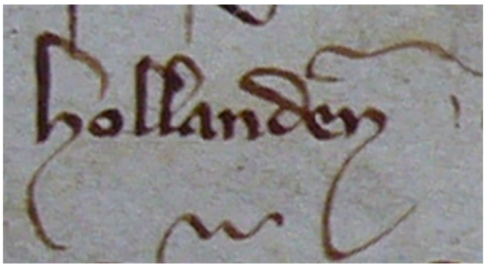

Fig. 1. The word "hollanden" in Dutch medieval handwriting. The width of the ink traces varies: vertical, southeast and northwest-bound strokes are thick, while southwest and northeast-bound strokes are thin.

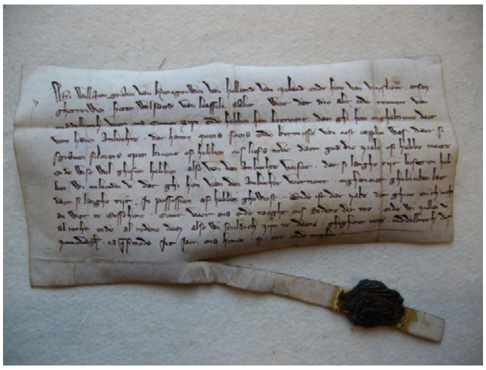

Fig. 2. Document in the *Dutch charter dataset*: medieval handwriting in a charter, a legal administrative document (1309). Other charters written by the same writing hand can be found using writer identification.

discern writers is comparable to that of *Hinge* [12] and *Fraglets* [13], which are among the world's best features.

2. Datasets

The *Quill* feature will be evaluated on two datasets of medieval handwriting, the *Dutch charter dataset* and the *English diverse dataset*, and on two datasets of contemporary handwriting, *Firemaker* and *IAM*. A description of these datasets follows.

2.1. Dutch charter dataset

The *Dutch charter dataset* is a new dataset of 118 early 14th-century Dutch charters (1299–1328): administrative documents which served as evidence, written with quill (goose feather) on parchment. See Fig. 2 for a sample photo. The charters are part of a collection that is studied at the University of Amsterdam in a research project called "Charters and Chancery of the Counts of Holland/Hainault, (1280) 1299–1345". The project is funded by NWO (the Netherlands Organisation for Scientific Research) and part of the VNC (the Flemish-Dutch Committee for Dutch Language and Culture) programme. As studies of medieval administrations rely heavily on writer identification, the different writing hands in this dataset were distinguished and code-named by the paleographers Jinna Smit and Jan Burgers. This was done independently and consistently, with the use of Burgers method [14], combining elements out of traditional and quantitative paleography.

The material in this dataset is graphically challenging in two ways. First, the originals contain difficulties such as pictorial letters, wrinkles, wax seals, and irregularly shaped parchment. Sometimes, parts of the ink are so faint that they are hard to distinguish. Many charters are in bad shape due to aging, tears or even fire damage. Second, the photos were taken using a 6 megapixel digital camera from different freehand positions and in different illumination conditions. Because of the free camera position, the images suffer from perspective distortions, non-uniform scaling, non-uniform illumination and occasionally blur. The photos also show document border shadows and sometimes document placement holders.

Most charters have been photographed several times; the dataset contains 248 photos in total. From this dataset, a variant was created, the *Dutch⁎ charters dataset*, which does not include the "duplicate" photos. The duplicate photos of each original were removed by randomly selecting one photo to keep and removing the others. This procedure was repeated 25 times, averaging the results.

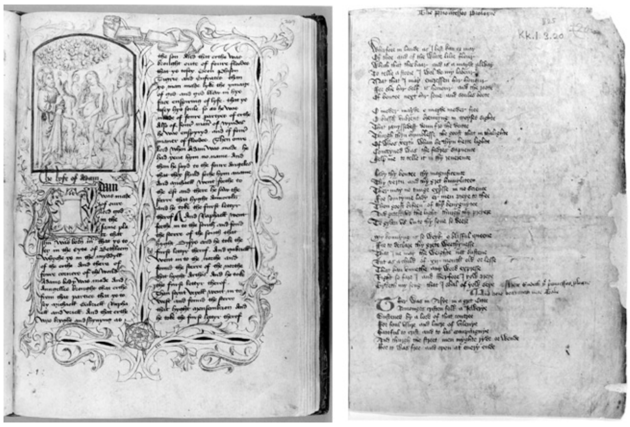

Fig. 3. Example documents from the *English diverse dataset*: the left image shows part of a manuscript produced by the Trinity Anthologies Scribe, Cambridge – Trinity College R.3.21, folio 249 Middle English verse and prose.

2.2. English diverse dataset

The *English diverse dataset* [15,16] is a collection of 70 grayscale images of various late medieval texts (1375–1525), written by 10 scribes. The dataset was collected from library and archival resources by Professor Linne Mooney, expert codicologist-paleographer at University of York, UK. She also ascertained the authorship of each manuscript. This dataset was kindly provided by Dr John Daugman, University of Cambridge, UK.

The documents are graphically complex: most of them contain decorative frames, other decorations or pictural letters. The layout also varies significantly. Frequently, the background is not uniform across the whole manuscript due to aging, stains and noise. They have been photographed from above, but probably from a variable distance thus the resolution of the digitization is not known and may not be constant. See Fig. 3 for sample photos.

2.3. Contemporary handwriting datasets

The feature was also tested on two datasets of contemporary handwriting: Firemaker [17] and IAM [18]. Firemaker contains 1004 pages written by 251 students; four pages each. Only pages 1 and 4 were used for this experiment, bringing the total to 502 pages. An example document was shown in Fig. 10(c). The IAM dataset contains handwritten English text written by 657 subjects, using different pens. The number of pages written by the subjects varies.

On each of the four datasets described above, the *Quill* feature will show to be effective for writer identification. The next section will discuss the theoretical background of this new feature.

3. Analysis of trace-width production

Since the *Quill* feature was inspired by paleographic methodology, it is instructive to understand how trace-width variations were produced in historical handwriting. Trace width is influenced by at least three factors: the writing instrument, the habitual tip angle and individual movement style. These factors are described in the next subsections. Following, the basic principle of production of historical handwriting is modeled, and support for trace-width variation in modern handwriting is presented.

3.1. Writing instrument

Quills are capillary-action writing instruments that were used until the 19th century [19]. A quill was made from the feather of a bird, usually a goose's, by hardening it [20, p. 163] and cutting a *nib* (writing end). The feather's pen was first cut twice to create a sharp point and then topped, creating an oblong *tip* (contact surface). Finally, the nib was incised from this tip, splitting it in two flexible parts, the *tines* [21, p. 5–7]. Fig. 4 shows such a quill nib. A few properties of this nib influence the trace width in a document [22, p. 72]:

- *Length of the tip*: This length depends on the radius of the pen and the position of the truncation on the nib. Scribes used to re-truncate the tip every so often during the process of writing, influencing the maxima in the ink trace width.
- *Flexibility of the tines*: The more flexible the tines are, the wider the ink traces could become. The flexibility is influenced by the length of the slit (incision) and degree of hardening. The flexibility also depends on properties of the used feather itself, as feathers vary naturally in pen stiffness, thickness and radius.
- *Angle of truncation*: The truncation was usually not exactly orthogonal, which had implications on the preferred pen grip.

3.2. Habitual tip angle

Scribes kept the orientation of the quill tip almost constant [23]. Since the tip of a quill is *oblong*, the effect on the produced ink trace is that it is thin where the tip was moved sideways and thick where it was dragged perpendicularly. Thus width variations were produced just by varying the writing direction. This principle is illustrated in Fig. 5. The tip angle can be measured in the handwriting: it is generally parallel to the thinnest traces and perpendicular to the widest. This tip orientation is an individual influence on the handwriting [23].

3.3. Individual movement style

Individual movement style emerges partly deliberate and partly unconsciously. The following individual methods to influence the trace width are known [24, p. 81–84]:

- *Force variation*: The incision in the quill tip allowed the scribe to vary the trace width by varying force on the pen. An example of unconscious force variation is the individual way of creating tapered trace endings. This is visible in Fig. 6 where the "z", "e" and "d" have tapered trace endings, created by gradually decreasing the force on the nib. The force was also manipulated deliberately, for example when writing north or west-bound, which requires a "pushing" pen movement. In this case, to avoid damaging the parchment, only a very light force could be applied. The effect is possibly visible in Fig. 6: in this example, the loops of the "l" and "d" are thinner than expected; these could have been drawn northwest-bound. It is also an established fact that the force variation is very writer-specific in contemporary handwriting [25].
- *Elevation* (*or pitch*) *variation*: A lower pen elevation lowers the force needed to bend the nib's tines and create a wider trace.
- *Pen orientation* (*azimuth*) *variation*: Even scribes with a very regular handwriting varied the orientation of the pen 2–3° in different documents [23]; it is conceivable that they also subconsciously varied the orientation within a document.
- *Rotation* (*or roll*): The quill could be rotated around its axis such that only one side of its surface touches the writing support; this enables creating curved thin lines. A slight pen rotation can be another explanation for the thin loops of the "l" and "d" in Fig. 6. In some cases scribes *reversed* the pen (a rotation of 180°) to create thin traces.

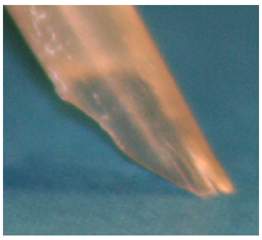

Fig. 4. A quill nib. The slit (incision) partly splits the nib into two tines; this enables width variation by changing pressure.

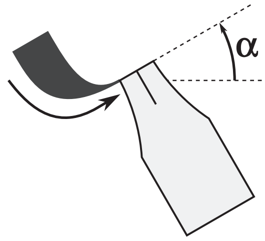

Fig. 5. A near-fixed quill tip angle α caused by a near-fixed pen orientation determines the pattern of width variation.

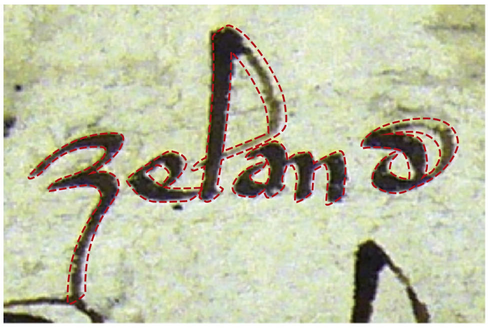

Fig. 6. Width variation in the medieval Dutch word "zeland". The superimposed contour (dashed) indicates the trace width to be expected when using a rigid oval pen tip. It illustrates individual influence on the trace width.

The first three subsections of this section motivates that many pen-specific and individual behavioral properties have effects on the width of the ink traces. In the next subsection, these influences will be temporarily abandoned to study the basic principle of width production.

3.4. Trace-width production models

To aid the interpretation of the upcoming measurements, it is instructive to disregard individual movement style for the moment and make a model for the basic calligraphic effect: ink traces are thin where the tip is moved sideways and wide where it is dragged perpendicularly, as illustrated in Fig. 5.

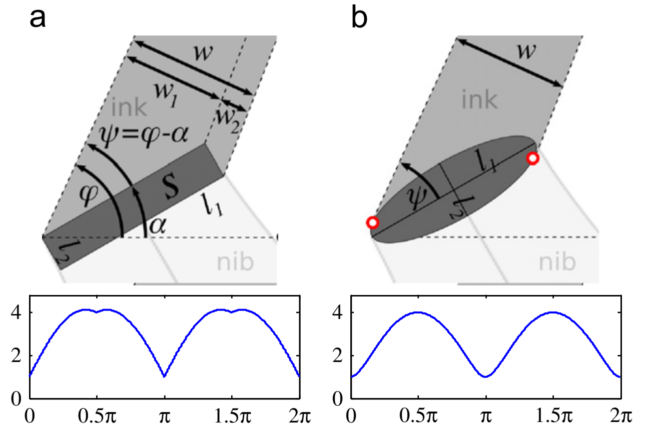

Fig. 7. Trace-width production in simplified conditions: the local ink trace width w depends on the tip shape and the relative orientation $\psi=\phi-\alpha$, where the tip is a rigid box or oval. Horizontal axis: ψ (radians); vertical axis: w (mm). w was plotted according Eqs. (1) (left) and 2 (right). Tip dimensions $(l_1,l_2)=(4\text{mm},1\text{mm})$ are assumed. (a) Box model. (b) Oval model.

This calligraphic effect is caused by the oblong tip shape of historical pens. The exact effective shape of the contact surface during writing is not known, therefore the effect of two basic oblong shapes will be modeled here: a *box* and an *oval*. Both shapes will be modeled to be rigid and fixed at a rotation angle α, similar to the usage of a *brush* in a graphics program. The oval shape may be more accurate than the box shape since cohesion and adhesion effects must make the ink at the quill tip round. The top row of Fig. 7 shows these shapes plus the parameters that will be used for the models.

In these simplified conditions, the trace width w only depends on the *relative* orientation $\psi=\phi-\alpha$, where ϕ denotes the local trace direction and α denotes the tip orientation. In the bottom row of Fig. 7, this relation is visualized in a graph for the two simple pen tip shapes, for any relative orientation. The graph for the box model was computed using Eq. (1), which computes the predicted trace width w for any relative orientation ψ, assuming a box shape of dimensions (l_1,l_2). The formula is easily derived from Fig. 7a. The graph for the oval model was computed using Eq. (2), which determines the height of a parametrized oval after rotation, where $t^*=\arctan l_2 \cos\psi/l_1 \sin\psi$ represents the parameter value at the top. Although the formulas for computing the graphs are not used for the *Quill* feature, these are mentioned for completeness and future reference.

$$w(\psi,l_1,l_2)=|l_1 \sin\psi|+|l_2 \cos\psi| \tag{1}$$

$$w(\psi,l_1,l_2)=|l_1 \cos t^* \sin\psi+l_2 \sin t^* \cos\psi| \tag{2}$$

The graphs demonstrate what the result of the *Quill* feature should look like if a rigid oblong tip shape can be assumed. Any deviation from these graphs will show the presence of individual influence on the trace width. Fig. 6 shows a piece of real data and illustrates that the oval model does not explain all width variation: the ink does not follow the expected contour exactly, proving the existence of influence on the trace width by the personal writing habits of the scribes. This individual influence makes the trace width a valuable source of writer-specific features.

3.5. Trace-width production in modern handwriting

Modern handwriting contains trace-width variations as well, although not as pronounced as in historical handwriting. At least two types of individual writing style play a role. First, the trace width is directly influenced by the force applied to the pen tip [8,10,26], and different writers apply this force differently [25]. In particular, downstrokes are usually wider than upstrokes [8,10]. Second, the width is altered by local retracings [8].

4. Quill feature

The *Quill* feature $p(\phi,w)$ captures the relation between the local width w and direction ϕ of ink traces in a probability distribution. It expresses properties of the used pen and the writer's unique way of producing variations in the width of the ink trace. The feature consists of a few simple parts that together form a powerful method for writer identification: contour tracing, angle measurements, width measurements, and calculation of a probability distribution. These parts are further explained in the next paragraphs. Fig. 9 illustrates the angle and width measurements, which form the heart of the *Quill* feature. The algorithm is summarized in pseudocode in Algorithm 1.

4.1. Contour tracing

After thresholding, the measurements are performed while traversing *contours*. An alternative approach would be to traverse the center line of the ink, but traversing contours is simpler and more robust [8]. Contours are here considered to be 8-connected circular trajectories of black pixels that are adjacent to white pixels. A fast method was designed that constructs these by following the *crack-edge contours*: contours consisting of the imaginary edges between foreground (black) and background (white) pixels. See Fig. 8. The crack-edge contours are followed counterclockwise, keeping the ink on the left-hand side, yielding an 8-connected pixel trajectory consisting of the foreground pixels that touch the crack edges. This method is fast and robust. It also ensures consistent measurements on both sides of a stroke, even on strokes that are only one pixel wide.

4.2. Angle measurements

Given the ink contours, ϕ is computed at all pixels on those contours. ϕ denotes the local ink direction, measured in a systematical way. Due to ambiguity in off-line handwriting, the actual writing direction can be in two opposite directions. It is possible to reconstruct the actual direction [8], although not reliably. Instead, ϕ is defined to be the angle of a local tangent line with respect to the horizontal, where the ink is to the left-hand side (in image space) of the tangent line. This is illustrated in Fig. 9. This systematic approach ensures robust measurements, rendering the actual writing direction unimportant: since the contours are circular, every measurement will generally be accompanied by an opposite measurement on the other side of the trace; both measurements counterbalance each other.

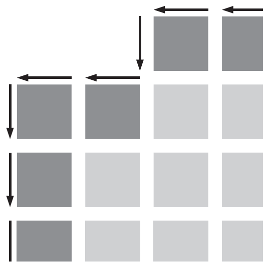

Fig. 8. Contour tracing by following crack-edge contours, shown as arrows. Foreground pixels are shown as blocks; the pixels in the resulting trajectory are shaded dark.

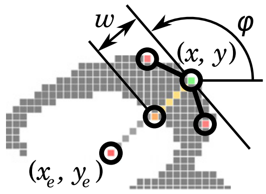

Fig. 9. ϕ and w are determined at each contour pixel (x,y). ϕ (trace direction) is measured by averaging the angles with two neighboring contour pixels at distance r. w (trace width) is computed using the so-called *Bresenham width*: the distance to the first background pixel that is hit when following a Bresenham path, perpendicular to ϕ, towards (x_e,y_e).

At contour pixel C_i, the ith pixel of contour C, ϕ can be estimated using two nearby contour pixels C_{i-r} and C_{i+r} which are the endpoints of imaginary "legs" originating from C_i at a distance of r contour pixels. A straightforward approach to estimate the local orientation ϕ would be to simply calculate the angle between C_{i-r} and C_{i+r}. This approach has the disadvantage that it is not accurate at positions of greatly varying curvature, such as stroke endings. Therefore, a slightly more elaborate method was used.

The leg from C_{i-r} to C_i defines an inbound angle ϕ_1; the leg from C_i to C_{i+r} defines an outbound angle ϕ_2. Since the pixel C_i is in the middle of these legs, ϕ can be estimated as the angle *between* ϕ_1 and ϕ_2. Because of the periodicity of angles, special care has to be taken to ensure that the resulting direction keeps the ink on the left side (in image space). When the difference between ϕ_1 and ϕ_2 is smaller than π radians, by definition the ink must be in the region between these angles. But when this difference exceeds π, by definition the ink must be on the other side, thus the resulting angle flips. Summarizing, ϕ is computed as follows:

$$\phi = \begin{cases} (\phi_1+\phi_2)/2 & \text{if } |\phi_1-\phi_2| < \pi \\ (\phi_1+\phi_2)/2+\pi & \text{otherwise} \end{cases} \tag{3}$$

4.3. Width measurements

After computing the local angle ϕ at contour pixel C_i, the local width of the ink trace w is computed. Any robust method could be used as part of the *Quill* feature. A few methods to compute trace width have been proposed before, designed for application on signatures [8,10]. These methods are based on traversing a line from C_i through the ink, perpendicular to ϕ, until the background is hit.

In this study, a variant of this principle was used because of its simplicity: a method based on Bresenham's algorithm [11], which constructs an approximated (quantized) linear path of pixels between two given pixel positions in an image. In this case, the starting pixel of this path is $C_i=(x,y)$. The end pixel (x_e,y_e) is a pixel that is on the line perpendicular to ϕ. This is illustrated in Fig. 9. The precise position of (x_e,y_e) is determined by a parameter, m, which signifies the maximum measurable width, as follows:

$$x_e = x+m*\cos(\phi+1\tfrac{1}{2}\pi) \tag{4}$$

$$y_e = y+m*\sin(\phi+1\tfrac{1}{2}\pi) \tag{5}$$

The pixels on the Bresenham path are traversed from C_i to (x_e,y_e) and checked for color: the algorithm stops if a background (white) pixel is hit. The trace width w is then computed as the distance from C_i to this background pixel (x_b,y_b) using a simple Euclidean measure:

$$w = \sqrt{(x-x_b)^2+(y-y_b)^2} \tag{6}$$

In the following, this method to compute the trace width will be called *Bresenham width*. In the next section, this method will be used together with angle measurements to form the *Quill* feature.

4.4. Probability distribution

The locally measured direction ϕ and width w are agglomerated in a probability distribution $p(\phi,w)$. It will be referred to as the *Quill probability distribution* (QPD). It is created as follows. Every measurement (ϕ,w) is agglomerated in an interpolated $p \times q$ histogram, where p is the number of bins into which the measured width is quantized; q is the number of angle bins. In the following, p is equal to m (maximum width) for simplicity, so each width bin corresponds to one pixel of width in the ink. The histogram was built using bilinear interpolation, updating four bins at once for every measurement, to avoid distortions caused by measurements close to bin boundaries. This is relevant because ϕ is discrete and because of delicate rounding errors due to angle periodicity. The resulting histogram was converted into a probability distribution by normalization, which makes it independent of the amount of text and usable as a writer-specific feature vector.

Algorithm 1. *Quill* feature. INPUT: binary image I, leg length r, number of width bins p, number of angle bins q. OUTPUT: 2D probability distribution P.

```
H ← empty_histogram(q,p)
Cs ← contourtrace(I)        {Cs is a list of contours}
for all C in Cs do
  n ← len(C)        {n is the current contour's length}
  for i in [0 to n−1] do
    φ1 ← angle(C[(i−r) mod n],C[i])
    φ2 ← angle(C[i],C[(i+r) mod n])
    φ ← { (φ1+φ2)/2        if |φ1−φ2| < π
        { (φ1+φ2)/2+π      otherwise
    (x,y) ← C[i]        {(x,y) is the current contour pixel}
    xe ← x+p*cos(φ+1.5π)
    ye ← y+p*sin(φ+1.5π)
    w ← bresenham_width(x,y,xe,ye)        {Compute width}
    H.update(φ,w)        {Update histogram, interpolated}
  end for
end for
P ← normalize(H)        {Make probability distribution}
```

4.5. Interpretation

By visually inspecting the QPD a variety of properties of the handwriting can be derived. See Fig. 10 for binarized example images and their QPD. Modern handwriting written with a ballpoint pen results in a near-horizontal structure, as the trace width is nearly the same for all ink directions. Dark (high-frequency) regions indicate frequently used angles, which are closely related to the handwriting's slant. For historical handwriting, the QPD reveals other properties of the handwriting as well:

- The most salient property of a QPD of historical handwriting is that it shows a wave shape, as predicted by the models in Section 3.4. The shape repeats itself after a period of π radians (180°) because generally every measurement on one side of a trace has a counterpart on the other side, in the opposite direction. The presence of peaks and valleys indicates that a writing instrument with an oblong tip was used, held at a near-fixed orientation: the ink width depends on the direction.
- The valleys correspond to the thinnest traces. The w value of the valleys corresponds to the width of the thinnest strokes; the ϕ value of the valleys reveals the near-fixed pen-tip angle: $\alpha \approx \phi \pm \pi$. By fitting a model described in Section 3.4 to the raw measurements, α could be estimated automatically.

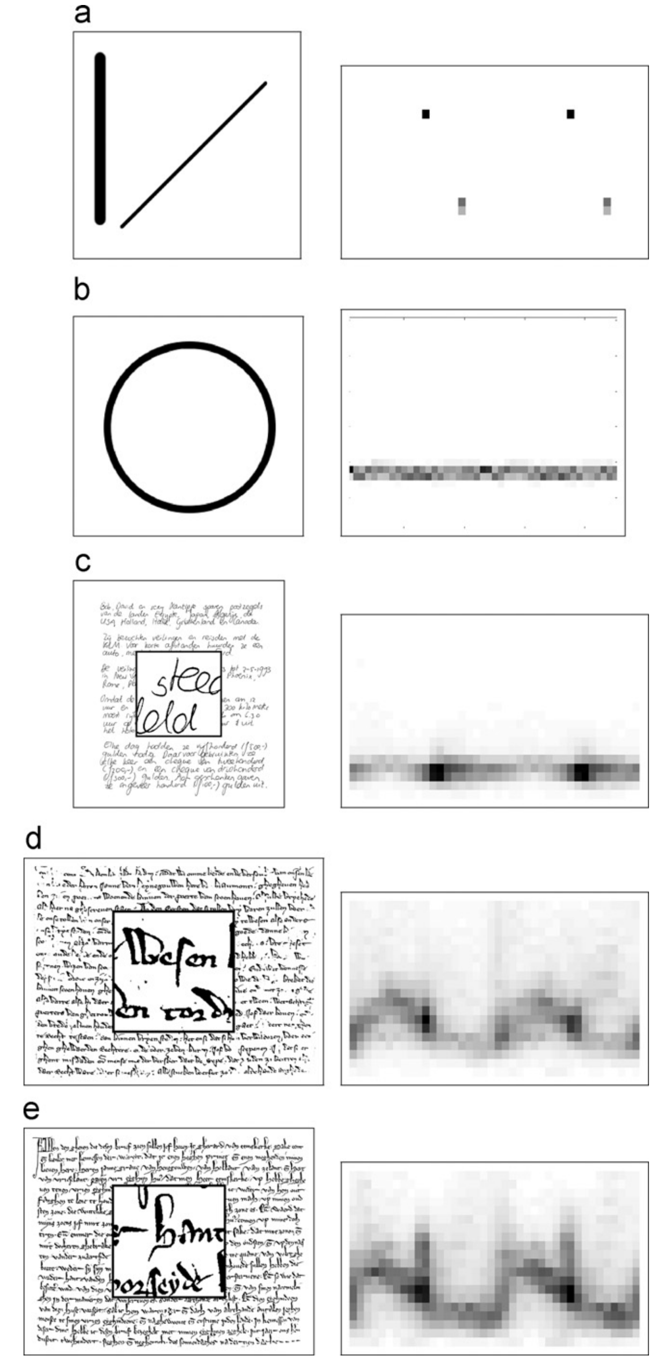

Fig. 10. Binary images (left) and their *Quill* probability distribution (QPD, right): a distribution of (ϕ,w) combinations. Dark regions indicate frequent combinations. The horizontal axis represents the trace direction ϕ $(0..2\pi$ radians), quantized in $q=40$ bins; the vertical axis represents the trace width w (1..20 pixels) in $p=20$ bins. Used parameters: $q=40, p=20, r=10$. (a) Artificial lines. The upper dark spots correspond to the vertical stroke; the lower spots correspond to the diagonal line. (b) Artificial circle. It has a constant trace width in all directions; it could have been produced using a stylus with a round tip and homogeneous ink deposition, possibly a fineliner. (c) Contemporary handwriting in the *Firemaker* dataset. Text produced with a ballpoint; width variation is limited. (d) Medieval Dutch text (1322) in the *Dutch* dataset. The QPD shows a wave shape, as predicted by the models in Fig. 7, thus this is quite regular medieval writing. (e) Medieval Dutch text (1300) in the *Dutch* dataset. The writer produced near-horizontal and near-vertical traces in a variety of widths, which show as vertical bands in the QPD.

- Similarly, the peaks correspond to the widest traces. The w value of the peaks corresponds to the width of the thickest strokes and the ϕ value of the peaks reveals the near-fixed pen-tip angle: $\alpha \approx \phi + \pi/2 \pm \pi$.
- The deviation from the models in Section 3.4 indicates the influence of physical pen properties and individual movement style, as described in Section 3.
- The ϕ value of the cell with the highest intensity (dark region in the figure) reveals the dominant stroke direction, which is closely related to the dominant *slant angle*. The corresponding w value reveals the dominant stroke width.

4.6. Variant: Quill–Hinge feature

Inspired by the success of the Hinge feature a modification of the *Quill* feature was developed: the *Quill–Hinge* feature $p(\phi_1,\phi_2,w)$. It records w in conjunction with ϕ_1 and ϕ_2 instead of ϕ, making the feature three-dimensional.

5. Performance experiment

The performance of the *Quill* feature was tested in a writer identification experiment on the four datasets that were introduced in Section 2. Its performance was also compared with the performance of other features. This section discusses the experiment.

5.1. Preprocessing

The images in the datasets of modern handwriting were preprocessed using text region cropping based on known fixed coordinates followed by Otsu thresholding. The described medieval documents are graphically more challenging and required additional steps:

- *Manual text region selection*: A region of interest (ROI) was manually selected in a graphical interface (GIWIS) by placing a four-sided polygon, which enables the careful selection of a text area that is rectangular in reality but subjected to perspective transformation. This was essential for the *Dutch charter dataset*, where the camera position was free.
- *Perspective correction*: The perspective distortion in the *Dutch* and *Dutch*∗ charter datasets was corrected by a reverse perspective projection. The parameters were derived from the positions of the four vertices of the ROI. The result was a rectangular image containing a version of the ROI, stretched using bilinear interpolation.
- *Automatic scaling*: The scale was estimated from the height of the text lines in the images, assuming that the true height of the text lines is equal in all documents. The text line height was determined by measuring the median width of the peaks in the smoothed horizontal projection profile of dark pixels. The images were then scaled to match a standard text line height of 50 pixels.
- *Highpass filtering*: Gradual intensity variations were canceled by applying a highpass filter, after grayscale conversion. This was implemented by a straightforward approach: blurring the image and subtracting that from the original image.
- *Otsu thresholding*: The last preprocessing step was to binarize the image using Otsu thresholding [27], which is widely recognized as a good general-purpose binarization method.

5.2. Writer identification experiment

The preprocessed datasets were used to test the power of the feature in a writer identification experiment. Writer identification is the recognition of the writer of a query document by yielding a *hit list*: a list of database documents that are similar to the query document in feature space. Typical hit list sizes are 1 or 10, therefore the feature's performance will be expressed as its *top-1* and *top-10 writer identification performance*, which means that a hit list will be counted as correct if at least one document of the query writer appears in it. Top-1 performance is also called nearest-neighbor accuracy.

In this experiment, nearest-neighbor (instance-based) classification is used. Any classifier could be used, but there are a few reasons for choosing nearest-neighbor classification. It is simple, effective and it does not require training or human intervention. Previously, several other popular classifiers such as MLPs and SVMs have been tried on similar classification tasks in unpublished studies at ALICE, but the nearest-neighbor classifier was never surpassed. Whereas trained classifiers allow for generalization of observations into a broad class, biometrics is concerned with the identification of individual samples. Nearest-neighbor search fits that goal very well and does not require evolved training. It is only necessary to find a proper distance function. Furthermore, all common parametric classifiers require large amounts of training examples, a requirement that cannot be easily fulfilled in practical settings.

The top-x performance is simply computed by treating each dataset document as a query, sorting the other documents by similarity, and counting how often another document from the same writer appears among the x most similar documents. Similarity is based on two documents' feature vectors and a distance measure. The following features were evaluated:

- The *Ink width* ($p(w)$) feature is a probability distribution (p.d.) of trace-width occurrence. It is a modified version of the *Quill* feature, where the number of angle bins is set to one ($q=1$). The effect is that this feature only measures the distribution of ink widths, without regarding the corresponding direction. This feature was included to roughly evaluate the importance of the ink width in the *Quill* feature.
- The *Directions* feature [2] ($p(\phi)$) is a p.d. of ink directions at the contours. This encodes slant and direction usage.
- The *Brush* feature [7] is a p.d. of ink intensities at stroke endings. It encodes pen landing and lifting habits.
- The *White runs* feature [6,7] ($p(rl)$, also named "HrunW") is a p.d. of horizontal white run-lengths. This encodes within- and between-letter spacing.
- The *Hinge* feature [12] ($p(\phi_1,\phi_2)$) is a p.d. of angle combinations that are measured on the boundaries of the ink. This encodes slant and curvature. *Hinge* is among the world's best features. It is partly similar to *Quill*, although it does not compute w, while it stores two angles at every measurement and not one.
- The *Fraglets* feature [13] ($p(g)$, also named "fCO3") is a p.d. of usage of graphemes (fragments of handwriting) from a codebook. This encodes allograph usage. In our experiment, the codebook was pre-computed using modern handwriting: all four pages of the first 100 subjects of the *Firemaker* dataset. The performance of *Fraglets* is comparable to *Hinge*'s.

The distance measure to determine the dissimilarity of two documents' corresponding feature values was the χ^2 distance [28], which is defined as:

$$d_{\chi^2}(\mathbf{v},\mathbf{w}) = \sum_{i=1}^{|\mathbf{v}|} \frac{(\mathbf{v}_i-\mathbf{w}_i)^2}{\mathbf{v}_i+\mathbf{w}_i} \qquad (7)$$

where i is an index to the elements of the feature vectors $\mathbf{v}$ and $\mathbf{w}$. This distance measure emphasizes differences in small feature values and has been shown to be effective on feature vectors of Hinge and Fraglets [29].

Four combinations of features were also tried: *Hinge & Fraglets*, *Quill & Hinge*, *Quill & Fraglets*, *Quill–Hinge & Fraglets*. These combinations were made by simply averaging the distance values.

For the *Dutch∗ charter dataset* the experiment was repeated 25 times; in each experiment one photo of each original was selected randomly. The results were averaged over the 25 runs.

5.3. Training

Instance-based classification does not require any training other than storing feature values, but the *Quill* feature contains three parameters (p, q and r) that need to be optimized. This was done using simple methods as will be described below. It will also be shown that the feature is not very sensitive to the actual parameter settings.

Optimizing the parameter values for p, q and r was done by evaluating 896 parameter combinations from a regularly spaced three-dimensional lattice on the *Dutch∗ charter dataset* using a 4×4-core PC with 128 GB of memory. It was not the maximum performance score in this evaluation that determined the final choice of the parameters. Instead, good parameter values were determined by visual inspection of the data plotted in figures such as Figs. 11 and 12. This procedure minimizes the effect of overtraining that may exist since no separate training set was used.

Figs. 11 and 12 show the sensitivity of *Quill* for its parameters. The three-dimensional evaluation cannot be fully visualized, therefore two "slices" of the results are shown. Fig. 11 shows how the performance relates to p and q, given a fixed $r=20$.

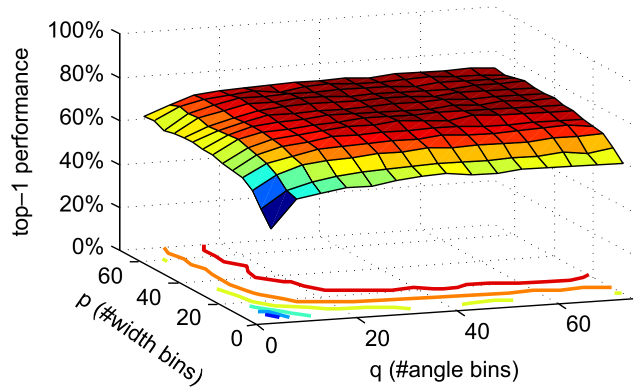

Fig. 11. Sensitivity of Quill's top-1 performance for the parameters q (number of angle bins) and p (number of width bins) in the *Dutch∗ charter* dataset for $r=20$. The graph shows that the choice of the parameters is almost arbitrary, as long as $q>30, p>30$.

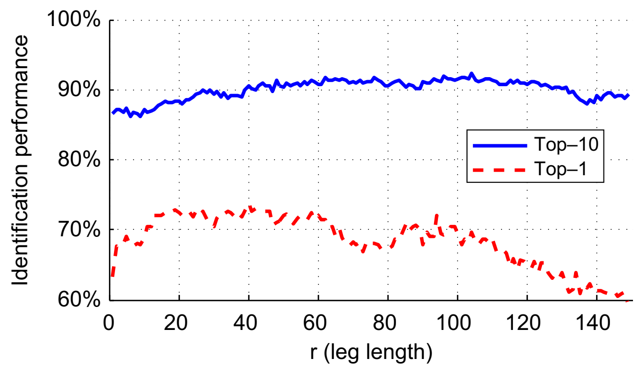

Fig. 12. Sensitivity of *Quill's* top-1 and top-10 performance for the parameter r (leg length) on the *Dutch∗ charter* dataset; $q=40, p=40$. The graph shows that the influence of this parameter is marginal, as long as the value is between about 10 and 100.

It shows that the performance is insensitive to the values of p and q, as long as they are at least about 30. Increasing the values much further does not increase the performance, but has a negative effect on memory usage and speed. The values $p=40$ and $q=40$ were chosen as safe values. The effect of the parameter r (leg length) on the performance is shown in Fig. 12. It shows that this performance is hardly affected by this parameter, as long as the value is between about 10 and 100. This is the same order of magnitude of the corpus height of the text (50 pixels). $r=20$ was chosen because higher values are less suitable for small characters and less intuitive. Summarizing, the found parameters are $p=40$, $q=40$ and $r=20$.

6. Results

Table 1 shows the results of all performance experiments. Top-1 performances of the well-known *Directions* feature are in the range of 48–74%. The *Ink width* feature, on the other hand, performs 22–73%. It coarsely describes the informational value of ink trace width: it is low in *Firemaker*, where ballpoints were used, intermediary on the medieval datasets and high in *IAM*, where different types of pens were used. The contribution of ink width in the Firemaker dataset seems small, but the *Ink width* feature does not take structural dependency on the trace direction into account while *Quill* does.

Ink width awareness partly explains *Quill*'s performance on the datasets, most notably *IAM*, in which different types of pens were used. The *Quill* feature performs much better than *Directions* or *Ink width* alone: top-1 performances are in the range of 63–95%. This proves that the combination of directionality measurements with trace-width measurements is fruitful. It also supports the foundation of characteristics in quantitative paleography that are based on trace angle and width. The value of trace directions for writer identification was known, but the added value of width measurements for writer identification is new.

Even when compared to the other features, *Quill* proves to perform very well: it amply outperforms *Brush* and *White runs*, and on average, it performs as well as *Fraglets* and *Hinge*, which are among the world's best features. The modified version of the feature, *Quill–Hinge*, seems to perform even better than *Quill*. Slightly better performances are achieved by combining features. Top-1 performances of *Quill & Fraglets* and *Quill–Hinge & Fraglets* are in the range of 77–97%. The results were obtained by averaging the distance scores for both features in the combination. Table 2 lists the performance of *Quill–Hinge* in the context of leading results obtained by others. Only results on modern handwriting are shown since little similar work has been done on historical handwriting.

Note that the reported performances may be several percentage points off (except in the column 'Dutch⁎') due to the relatively low number of queries and the influence of randomness. For example, the 95% confidence interval for a recognition score of 92 with 245 queries (*Quill* in column Dutch) is the range of 88–96 (based on the Binomial distribution).

Still, these performance figures suggest that *Quill* and *Quill–Hinge* can be used in a production environment. Researchers in application fields are currently evaluating the features [34–36] in our GIWIS software program, a user-friendly graphical user interface. A recognition score of 100% is not necessary: based on a top-10 hitlist, the domain specialist can make the final decision. However, some improvements are still possible, as we will describe in the following.

7. Future work

A simple optimization approach is to try other distance measures. An excellent overview of available distance measures

Table 2
Performance of *Quill–Hinge* on modern handwriting compared to leading results obtained by others. Note that the numbers cannot be well compared because of differences in dataset material, required level of human interference, and number of writers.

Publication	writers	top1 (%)	top10 (%)
Bensefia et al. [30]	150	87	99
Bulacu [31]	900	87	96
Garain et al. [32]	422	62	96
Schlapbach et al. [33]	100	97	98
Schomaker et al. [13]	150	97	100
Siddiqi et al. [4]	650	86	97
Srihari et al. [5]	900	88	
Quill–Hinge	**251**	**86**	**97**
Quill–Hinge	**657**	**97**	**98**

Table 1
Writer identification performance of several features (top, middle) and feature combinations (bottom) on datasets of medieval and contemporary handwriting. The numbers represent recognition percentages.

Feature	Medieval handwriting						Contemporary handwriting			
	Dutch⁎ 18 writers 118 images 25 × 112 queries		**Dutch** 18 writers 248 images 245 queries		**English** 10 writers 70 images 70 queries		**Firemaker** 251 writers 502 images 502 queries		**IAM** 657 writers 1539 images 1183 queries	
	Top1	Top10	Top1	Top10	Top1	Top10	Top1	Top10	Top1	Top10
Directions	57	86	69	93	56	93	48	79	74	90
Ink width	47	79	66	89	40	87	22	52	73	88
Quill	**75**	**90**	**92**	**98**	**63**	**96**	**71**	**89**	**95**	**97**
Quill–Hinge	**75**	**89**	**92**	**98**	**70**	**96**	**86**	**97**	**97**	**98**
Comparison features										
Brush	38	79	53	85	40	86	37	77	78	89
White runs	43	86	67	94	37	83	22	57	43	76
Hinge	71	92	86	97	76	93	83	92	94	96
Fraglets	73	94	92	98	89	100	72	90	97	98
Combinations										
Hinge & Fraglets	75	93	94	98	90	99	76	93	97	98
Quill & Hinge	73	91	93	98	74	96	83	95	96	97
Quill & Fraglets	77	94	94	98	86	99	78	95	97	98
Quill–Hinge & Fraglets	77	94	94	98	86	99	81	96	97	98

for probability distributions exists [28] which can serve as a starting point. However, the currently used χ^2 distance measure has proven to be effective in previous experiments [29].

The method of combining features could be improved as well. Simply averaging could be replaced by feature weighting, however little gain is expected since it is known that *Fraglets* and *Hinge* are best weighted by plain averaging [12]. This was confirmed by a small pilot experiment with *Quill* and *Fraglets*. The features could also be combined in different ways, for example, using Borda ranking.

Preprocessing could be improved. Although the preprocessing used in the performance experiment generally works quite well, it breaks the thinnest faint ink traces. The result is that measurements on those traces are underrepresented in the QPD, resulting in suboptimal performance. Preserving this weak signal is still a hard image processing challenge.

Quill and *Quill–Hinge* are pen-dependent features. This can be advantageous but it may not be so if a significant number of writers each use multiple pens. To partly cancel the pen dependence as described in Section 3.1, one of the models presented in Section 3.4 could be fitted to the measurement data, followed by a rescaling of the data based on the width range. This could also compensate for scaling differences due to varying camera distance or resolution; these are now dealt with in an explicit preprocessing step. However, this model fitting is not trivial since the data contains structural noise.

For performance analysis, better medieval datasets could be collected or constructed. The results on the used datasets are not fully reliable since some writer labels may be wrong: the labels were manually determined based on skill and experience, not facts. Furthermore, the currently used datasets are relatively small. Larger datasets with known writer labels will make the results more reliable.

The presented methods can be used for other applications as well, including pen and script type estimation, by analyzing the QPD, and estimation of the modal pen-tip orientation α by fitting the models from Section 3.4 to the measurement data.

8. Conclusion

As suggested by modern paleographic methodology, the width of the ink trace is a powerful source of information for writer identification, particularly in combination with the trace direction. This is not only true for off-line writer identification on historical handwriting, which shows salient width differences because of usage of quill pens, but for modern handwriting as well. This was found in a series of writer identification experiments using a newly introduced feature: *Quill*. It is a 2D joint probability distribution of ink trace width and direction. The feature consists of simple, fast and accurate methods based on pixel contours. The feature was tested on two datasets of modern handwriting and two datasets of medieval handwriting: writer identification scores (nearest-neighbor classification accuracy) scores are in the range of 63–95%. This is much higher than the individual performances of features based on either the ink width or direction. It even approaches the performances of *Hinge* (71–94%) and *Fraglets* (72–97%), which are among the world's best features. A slightly more complex version of the feature involving curve measurements, *Quill–Hinge*, seems to perform even better. The performance of *Quill* and *Quill–Hinge* strengthens the foundation of related measurements in quantitative paleography. The features can be used as general-purpose writer identification features. Slightly higher performance can be achieved by combining the features with other features (77–97% with *Fraglets*). In GIWIS, a user-friendly user interface, the features are already helping historians fruitfully.

Acknowledgments

This research was made possible thanks to NWO grant 634.000.434 (ToKeN/TriGraph).

Fig. 2 is: Charter of 1309 April 13th, written by writing hand KaAA. (*Zeeuws Archief*, *Onze Lieve Vrouwe abdij te Middelburg*, access number 27, item number 80, *regest* number 119). Photo: Jinna Smit.

References

[1] A. Derolez, Possibilités et limites d'une paléographie quantitative, in: P. Defosse (Ed.), Hommages à Carl Deroux, vol. 279, Editions Latomus, Brussels, 2003, pp. 98–102.
[2] M. Bulacu, L. Schomaker, Writer style from oriented edge fragments, in: CAIP 2003, LNCS, Springer, 2003, pp. 460–469.
[3] J.-P. Crettez, A set of handwriting families: style recognition, in: Proceedings of the 3rd ICDAR, 1995, pp. 489–494.
[4] I. Siddiqi, N. Vincent, A set of chain code based features for writer recognition, in: Tenth International Conference on Document Analysis and Recognition (ICDAR), 2009, pp. 981–985.
[5] S.N. Srihari, S.-H. Cha, H. Arora, S. Lee, Individuality of handwriting, Journal of Forensic Sciences 47 (2002) 856–872.
[6] B. Arazi, Handwriting identification by means of run-length measurements, SMC 7 (1977) 878–881.
[7] L. Schomaker, M. Bulacu, M. van Erp, Sparse-parametric writer identification using heterogeneous feature groups, ICIP, vol. 1, 2003, pp. 545–548.
[8] D.S. Doermann, A. Rosenfeld, Recovery of temporal information from static images of handwriting, International Journal of Computer Vision 15 (1995) 143–164.
[9] H. Lv, W. Wang, C. Wang, Q. Zhuo, Off-line chinese signature verification based on support vector machines, Pattern Recognition Letters 26 (15) (2005) 2390–2399 doi:http://dx.doi.org/10.1016/j.patrec.2005.04.013.
[10] J.K. Guo, Forgery detection by local correspondence, Ph.D. Thesis, Center for Automation Research, University of Maryland, 2000.
[11] D. Hearn, M.P. Baker, Computer Graphics, C version, Prentice-Hall, 1997.
[12] M. Bulacu, L. Schomaker, Text-independent writer identification and verification using textural and allographic features, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 29 (4) (2007) 701–717.
[13] L. Schomaker, M. Bulacu, K. Franke, Automatic writer identification using fragmented connected-component contours, in: F. Kimura, H. Fujisawa (Eds.), Ninth IWFHR, Tokyo, Japan, 2004, pp. 185–190.
[14] J. Burgers, De paleografie van de documentaire bronnen in Holland en Zeeland in de dertiende eeuw, Peeters, Leuven, 1995.
[15] J. Daugman, Identification of medieval scribal handwriting, Suggested student project, website consulted in March 2009, 〈http://www.cl.cam.ac.uk/~jgd1000/scribes.html〉.
[16] M. Bulacu, L. Schomaker, Automatic handwriting identification on medieval documents, in: Proceedings of 14th International Conference on Image Analysis and Processing (ICIAP 2007), IEEE Computer Society, Modena, Italy, 2007, pp. 279–284.
[17] L. Schomaker, L. Vuurpijl, Forensic writer identification: a benchmark data set and a comparison of two systems, Technical Report, NICI, Nijmegen, 2000.
[18] U. Marti, H. Bunke, A full english sentence database for off-line handwriting recognition, in: Proceedings of the 5th ICDAR, 1999, pp. 705–708.
[19] S. Howard, The steel pen and the modern line of beauty, Technology and Culture 26 (4) (1985) 785–798.
[20] A.M. Smith, Printing and Writing Materials: Their Evolution, Smith, Philadelphia, 1904.
[21] J. Nickell, Pen, Ink, and Evidence: A Study of Writing and Writing Materials for the Penman, Collector, and Document Detective, University Press of Kentucky, 1990.
[22] D.H. Wilson, The Encyclopedia of Calligraphy Techniques, New Burlington Books, London, 1991.
[23] L. Gilissen, L'expertise des écritures médiévales: recherche dúne méthode avec application à un manuscrit du XIème siècle: le lectionnaire de Lobbes (Codex Bruxellensis 18018), Story-Scientia, Ghent, 1973.
[24] M. Drogin, Medieval Calligraphy: Its History and Technique, Dover Publications, 1989.
[25] L.R.B. Schomaker, R. Plamondon, The relation between pen force and pen-point kinematics in handwriting, Biological Cybernetics 63 (1990) 277–289.
[26] K. Franke, The influence of physical and biomechanical processes on the ink trace – methodological foundations for the forensic analysis of signatures, Ph.D. Thesis, University of Groningen, The Netherlands, 2005.
[27] N. Otsu, A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man and Cybernetics 9 (1) (1979) 62–66.
[28] S.-H. Cha, Comprehensive survey on distance/similarity measures between probability density functions, International Journal of Mathematical Models and Methods in Applied Sciences 1 (2007) 300–307.
[29] L. Schomaker, M. Bulacu, Automatic writer identification using connected-component contours and edge-based features of uppercase western script, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (6) (2004) 787–798 doi:http://dx.doi.org/10.1109/TPAMI.2004.18.

[30] A. Bensefia, T. Paquet, L. Heutte, A writer identification and verification system, Pattern Recognition Letters 26 (13) (2005) 2080–2092 doi:http://dx.doi.org/10.1016/j.patrec.2005.03.024.
[31] M.L. Bulacu, Statistical pattern recognition for automatic writer identification and verification, Ph.D. thesis, University of Groningen (2007).
[32] U. Garain, T. Paquet, Off-line multi-script writer identification using AR coefficients, in: Tenth International Conference on Document Analysis and Recognition (ICDAR), 2009, pp. 991–995.
[33] A. Schlapbach, H. Bunke, A writer identification and verification system using HMM based recognizers, Pattern Analysis and Applications 10 (1) (2007) 33–43 http://10.1007/s10044-006-0047-5.
[34] M. Aussems, A. Brink, Digital palaeography, in: M. Rehbein, P. Sahle, T. Schassan (Eds.), Kodikologie und Paläographie im Digitalen Zeitalter/Codicology and Palaeography in the Digital Age, Schriftenreihe des Instituts für Dokumentologie und Editorik, Books on Demand, vol. 2, Norderstedt, 2009, pp. 293–308.
[35] J. Nobels, M. van der Wal, Tackling the writer-sender problem: the newly developed Leiden identification procedure (LIP), Historical Sociolinguistics and Sociohistorical Linguistics 9, 〈http://www.let.leidenuniv.nl/hsl_shl/index.html〉.
[36] J. Smit, Meten is weten? De toepassing van het Groningen Intelligent Writer Identification System (GIWIS) op Hollandse kanselarijoorkonden, 1299–1345, Handelingen van de Koninklijke Commissie voor Geschiedenis/Bulletin de la Commission Royale d'Histoire 176 (2) (2010) 343–359.

Axel Brink was born in 1979 in the Netherlands. He received the M.Sc. degree in Computing Science on the topic of Scientific computing and visualization from the University of Groningen in 2004. He completed one year of Communication and Information Sciences at the same university in 2005. Since 2005, he has been a Ph.D. student at the Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen. Since 2010, he is also employed as a knowledge analyst at a software company named Be Informed.