Beyond OCR: Handwritten manuscript attribute understanding
He, Sheng

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
He, S. (2017). Beyond OCR: Handwritten manuscript attribute understanding [Groningen]: University of Groningen

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
This chapter presents a method for extracting rotation-invariant features from images of handwriting samples that can be used to perform writer identification. The proposed features are based on the Hinge feature (Bulacu and Schomaker, 2007), but incorporating the derivative between several points along ink contours. Finally, we concatenate the proposed features into one feature vector to characterize the writing style of the given handwritten document. The proposed method has been evaluated using the Firemaker and IAM datasets in writer identification, showing promising performance gains.

2.1 Introduction

In this chapter, we present a new set of features called ΔⁿHinge with different n based on the Hinge feature proposed in (Bulacu and Schomaker, 2007). Although the Hinge feature has been successfully used in writer identification, there is one obvious drawback: it is sensitive to rotation changes of document images, which can be easily introduced in poor scanning practices. To overcome this problem, we generalize the Hinge feature to the ΔⁿHinge feature, which has the rotation-invariant property when n > 0. On the other hand, when n = 0, Δ⁰Hinge is exactly the original Hinge feature. Therefore, the proposed ΔⁿHinge feature can be considered as the generalization of the Hinge feature.

The proposed ΔⁿHinge features with different n have several advantages: 1) They are rotation-invariant, which are, to our best knowledge, the first rotation-invariant features in identification of writers; 2) Although the proposed features are computed from off-line documents, they are indicative of temporal events. There is a lawful relation between curvature and pentip velocity that has been extensively studied (Morasso and Ivadi, 1982; Teulings and Maarse, 1984; Schomaker et al., 1989; Guerfali and Plamondon, 1998). The features proposed here, therefore, can also be directly applied to on-line handwriting.
2. Writer Identification Using Delta-n Hinge Feature

2.2 $\Delta^n$Hinge Feature

The Hinge feature captures the joint probability distribution of orientations of two legs of the obtained “contour-hinge” (Bulacu and Schomaker 2007) along the ink contours. Given an arbitrary starting point, a counter-clockwise evaluation follows. If we assume that points on the ink contour are generated one by one, like the on-line handwriting, with a writing direction $\varphi$, two legs of the hinge can be defined as “previous” orientation $\varphi_1$, which is opposite to the writing direction $\varphi$, and as “succeeding” orientation $\varphi_2$, which follows the writing direction $\varphi$. Here we denote one point $p_j$ associated with two orientations $\varphi_1\{p_j\}$ and $\varphi_2\{p_j\}$ as a “Hinge kernel” (see $\Delta^0$Hinge $\{p_3\}$ in Fig. 2.1).

The Hinge feature can be considered as a statistical descriptor of handwritten contours, which counts the probability of each pattern appeared in the considered contours. For each point $p_j$ which has pair angles $(\varphi_1\{p_j\}, \varphi_2\{p_j\})$, the probability of such pattern in a given document is calculated by:

$$p(\varphi_1, \varphi_2) = \frac{c(\varphi_1, \varphi_2)}{C}$$  \hspace{1cm} (2.1)

where $c(\varphi_1, \varphi_2)$ is the number of the pattern $(\varphi_1, \varphi_2)$ appeared in the given document image, and $C$ is the total number of patterns in all ink contours. $p(\varphi_1, \varphi_2)$ is a bivariate probability distribution capturing both the orientation and the curvature of handwriting contours (Bulacu and Schomaker 2007). Finally, the probability distribution is agglomerated in a $q \times q$ histogram, where $q$ is the number of angle bins. The histogram is built using the bilinear interpolation to avoid distortions caused by measures close to bin boundaries.

Based on the Hinge feature, we propose a new set of features for writer identification, which is called $\Delta^n$Hinge with different $n$. A sequence of pixels with a fixed interval of distance along the ink contours are considered simultaneously to construct the probability of angle derivative on the “previous” and “succeeding” directions. We denote such sequence with a fixed interval of Manhattan distance $\Delta l$ as $\{p_j, p_{j+1}, ..., p_{j+n-1}\}$, where $\Delta l = |p_i - ...
2.2. $\Delta^n$Hinge feature

$p_{i-1}, i = j + 1, j + 2, \ldots, j + n - 1$. The starting point of the sequence is $p_j$, and the end point is $p_{j+n-1}$. Given this sequence, the $(n-1)$-th derivative of the two orientations in Hinge kernel is denoted as:

$$j\Delta^{n-1} \phi_i = \phi_i\{p_j, p_{j+1}, p_{j+2}, \ldots, p_{j+n-1}\} \quad i = 1, 2$$  \hspace{1cm} (2.2)

where $\phi_1$ and $\phi_2$ are the two “previous” and “succeeding” orientations in the Hinge kernel respectively. $j\Delta^{n-1} \phi_i$ is the $(n-1)$-th derivation along the $\phi_i$ orientation with the starting point $p_j$.

When the $(n-1)$-th derivative of the two orientations is obtained, the $n$-th derivative is computed as:

$$j\Delta^n \phi_i = \frac{j+1\Delta^{n-1} \phi_i - j\Delta^{n-1} \phi_i}{\Delta l} \quad i = 1, 2$$  \hspace{1cm} (2.3)

Two sequences with different stating points $p_{j+1}$ and $p_j$ subjected to $|p_{j+1} - p_j| = \Delta l$ are involved in the computation of $n$-th derivation in two orientations of the Hinge kernel. From Eq. (2.3) we can find that the computation of $n$-th derivative relies on the $(n-1)$-th derivative. When $n - 1 = 0$, we can get the initial value of “previous” angle $j\Delta^0 \phi_1 = \phi_1\{p_j\}$ and “succeeding” angle $j\Delta^0 \phi_2 = \phi_2\{p_j\}$, which are the Hinge kernel on point $p_j$ (see $\Delta^0$Hinge on the point $p_3$ in Fig. 2.1).

Given handwritten contours, each pixel on the contour is considered as the $j$-th starting point and the pattern $(j\Delta^n \phi_1, j\Delta^n \phi_2)$ is obtained by Eq. (2.3). All patterns are quantized into a histogram, and finally the $\Delta^n$Hinge feature is given by:

$$\Delta^n\text{Hinge} = p(\Delta^n \phi_1, \Delta^n \phi_2) \quad n = 0, 1, 2, 3, ...$$  \hspace{1cm} (2.4)

where the $p(\Delta^n \phi_1, \Delta^n \phi_2)$ is defined as same way as Eq. (2.1). From Eq. (2.2), Eq. (2.3) and Eq. (2.4) we can find that the $\Delta^n$Hinge feature is built on the $\Delta^{n-1}$Hinge, which can be recursively computed by the $\Delta^{n-2}$Hinge and the $\Delta^{n-3}$Hinge and so on. The initial $\Delta^0$Hinge is the Hinge (Bulacu and Schomaker, 2007). Therefore, as we mentioned before, the proposed $\Delta^n$Hinge is the generalization of the Hinge feature, and the Hinge feature is the special case of the $\Delta^n$Hinge feature when $n = 0$.

Corollary 1: Properties of the $\Delta^n$Hinge feature:

1. When $n = 0$, $\Delta^0$Hinge is the Hinge feature (Bulacu and Schomaker 2007).
2. When $n = 1$, $\Delta^1$Hinge works similarly as the first derivative (alike to the angular velocity long the contours) of pen coordinates in signature verification (Kholmatov and Yanikoglu 2005; Richiardi et al., 2005).
3. When $n = 2$, $\Delta^2$Hinge works similarly as the second derivative (alike to accelerations) of pen coordinates in signature verification (Kholmatov and Yanikoglu 2005; Richiardi et al., 2005).
4. When $n > 2$, $\Delta^n$Hinge contains high order derivative information of handwritten contours in document images.
Corollary 2: The proposed $\Delta^n\text{Hinge}$ has the rotation-invariant property when $n > 0$. Assume that the document has a small rotation angle $\theta$, and the $\Delta^n\text{Hinge}$ probability of the rotated document is denoted as $p(\tilde{\Delta^n}\phi_1, \tilde{\Delta^n}\phi_2)$. Then we have

$$p(\tilde{\Delta^n}\phi_1, \tilde{\Delta^n}\phi_2) = p(\Delta^n\phi_1, \Delta^n\phi_2) \quad n = 1, 2, 3, ...$$ (2.5)

Proof: According to Eq. 2.3, if there is a small rotation angle $\theta$ on the whole document, when $n > 0$, the $n$-th derivative of the $\Delta^n\text{Hinge}$ kernel is computed as:

$$\tilde{j\Delta^n}\phi_i = \frac{(j\Delta^{n-1}\phi_i + \theta) - (j+1\Delta^{n-1}\phi_i + \theta)}{\Delta l} = \frac{j\Delta^{n-1}\phi_i - j+1\Delta^{n-1}\phi_i}{\Delta l} = j \Delta^n\phi_i$$ (2.6)

2.2.1 $Ho^2D^n$ feature

Previous studies have shown that the performance of combined different feature sets is better than individual features involved in the combination (Schomaker and Bulacu, 2004; Siddiqi and Vincent, 2010; Bulacu and Schomaker, 2007; Bulacu et al., 2006). Inspired by this observation, different components of the proposed $\Delta^n\text{Hinge}$ features with different $n$ are concatenated into one feature vector to form the Histograms of Hinge over Derivative with $n$ feature, dubbed $Ho^2\text{D}^n$, or $Ho^2\text{D}^n$, which is defined as:

$$Ho^2\text{D}^n = \{\Delta^0\text{Hinge}, \Delta^1\text{Hinge}, ..., \Delta^n\text{Hinge}\}$$ (2.7)

From this definition, the $Ho^2\text{D}^0$ feature is the original Hinge feature, which is sensitive to rotation changes. If the rotation-invariant feature is required, the $\Delta^0\text{Hinge}$ should be excluded from $Ho^2\text{D}^n$, denoted as $Ho^2\text{D}^{n+}$, which is a rotation-invariant feature.

2.3 Writer Identification

The nearest-neighbor classifier with a “leave-one-out” strategy is often used in writer identification system (Schomaker and Bulacu, 2004; Bulacu and Schomaker, 2007; Siddiqi and Vincent, 2010; Brink et al., 2012). Given a query document $Q$, the system sorts all documents in the training set based on a given distance function ($\chi^2$ distance in this chapter) to the query $Q$. Ideally, the sample with the minimum distance should be the pair produced by the same writer. Not only the nearest neighbor (Top-1), but also a longer list up to a given rank (Top-10) are used to measure the performance of the identification system, corresponding to the Top-1 and Top-10 performance.
2.4 Experiments

2.4.1 Data sets

In this chapter, two data sets are used to evaluate our proposed method: Firemaker (Schomaker and Vuurpijl [2000]) and IAM (Marti and Bunke [2002]). The Firemaker set contains handwriting collected from 250 Dutch subjects, who were required to write four different A4 pages. In this dataset, lowercase pages are commonly used to evaluate writer identification methods (Schomaker and Bulacu [2004]; Bulacu and Schomaker [2007]). In our experiments, we also perform searches/matches of page 1 versus page 4 (lowercase pages). The IAM data set is modified as (Bulacu and Schomaker [2007]): we randomly selected two samples for those writers who contributed more than two documents, and we roughly split the document in two parts for those writers with a unique page. Finally, the IAM data set used in the experiments contains lowercase handwriting from 650 people, two samples per writer.

2.4.2 Experimental setting

The images of the Firemaker and IAM datasets are binarized using Otsu thresholding (Otsu, 1975), which is widely used on modern handwritten documents. After thresholding, the ink contours are extracted by the tracing method proposed in (Brink et al., 2012). Given the extracted ink contours, the two orientations $\phi_1$ and $\phi_2$ of the Hinge kernel are computed at all pixels on those contours.

There are four parameters in the proposed method: the number of angle bins $q$, leg length $r$, Manhattan distance $\Delta l$, and the number of derivative $n$. It was shown in (Brink et al., 2012) that the performance is insensitive to the value of $q$, as long as it is at least about 30, and to value of $r$ as long as it is between 10 and 100. Therefore, in our experiments we set $q = 40$, $r = 15$. We empirically set the Manhattan distance $\Delta l = 7$. The experiment shows that the better choice for $n$ is $n = 2$ or $n = 3$, depending on the specific data set.

2.4.3 Rotation-invariant study

In this section, we perform a rotation-invariant study on the Firemaker and IAM datasets. In both datasets, each writer has two samples. Therefore, we keep the first one and rotate the second one with a small $\theta$ angle. In our experiments, we evaluate the rotation change angle $\theta \leq 10$. For those documents which have rotation angle greater than 10, some rotation operators can be used manually or automatically to adjust it to the normal ones. The experimental results on the Firemaker and IAM dataset are presented in Fig. 2.2 and Fig. 2.3 respectively. These figures show that, with the increase of rotation change angle $\theta$ from 0 to 10, the Top-1 performance of $\Delta l$Hinge decreases significantly from 89.2% to 25.6% in Firemaker, a drop of 63.6%, and from 91.6% to 17.1% in IAM, a drop of 74.5%. However, the performance
2. Writer Identification Using Delta-n Hinge Feature

Figure 2.2: Rotation study on the Firemaker dataset. The left figure shows the Top-1 identification rate with rotation angle (°), and the right one shows the Top-10 results with rotation angle (°) from 0 to 10 degree.

Figure 2.3: Rotation study on the IAM dataset. The left figure shows the Top-1 identification rates with rotation angle (°), and the right one shows the Top-10 results. Note that the Firemaker data set is based on a single type of ball point pen, whereas the IAM data set contains many writing instruments.

of $\Delta^1$Hinge, $\Delta^2$Hinge and $\Delta^3$Hinge decreases slightly, by 14.4%, 18.6% and 21.6% in Firemaker respectively, and by 4.5%, 6.6%, 11.5% in IAM respectively. The slight decrease is partly caused by quantization artifacts introduced by the rotation operator, since the image is defined on a discrete grid. The same trend can be found on the Top-10 performance on both Firemaker and IAM. Therefore, the proposed $\Delta^n$Hinge, $n > 0$ is less sensitive to rotation changes.
Table 2.1: The writer identification performance of the proposed $\Delta^n$Hinge feature with different values of $n$ from 0 to 10.

<table>
<thead>
<tr>
<th>$\Delta^n$Hinge</th>
<th>n</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firemaker</td>
<td>Top-1</td>
<td>89.2</td>
<td>84.4</td>
<td>79.8</td>
<td>72.6</td>
<td>75.0</td>
<td>60.2</td>
<td>65.0</td>
<td>57.6</td>
<td>57.0</td>
<td>45.6</td>
<td>40.1</td>
</tr>
<tr>
<td></td>
<td>Top-10</td>
<td>95.8</td>
<td>97.4</td>
<td>95.0</td>
<td>91.6</td>
<td>93.4</td>
<td>84.6</td>
<td>86.8</td>
<td>85.0</td>
<td>86.2</td>
<td>73.8</td>
<td>70.5</td>
</tr>
<tr>
<td>IAM</td>
<td>Top-1</td>
<td>91.6</td>
<td>84.8</td>
<td>83.5</td>
<td>66.8</td>
<td>67.3</td>
<td>49.9</td>
<td>50.8</td>
<td>38.6</td>
<td>43.0</td>
<td>30.3</td>
<td>35.5</td>
</tr>
<tr>
<td></td>
<td>Top-10</td>
<td>96.0</td>
<td>95.3</td>
<td>94.9</td>
<td>87.5</td>
<td>87.2</td>
<td>76.6</td>
<td>78.2</td>
<td>66.7</td>
<td>71.9</td>
<td>58.5</td>
<td>63.4</td>
</tr>
</tbody>
</table>

### 2.4.4 Performance of the $\Delta^n$Hinge feature

In this section, we evaluate the performance of each part of $\Delta^n$Hinge with different $n$. Table 2.1 shows experimental results with different $n$ from 0 to 10. From the table we can see that the performance is slightly different on two datasets. For Firemaker, the maximum identification rate of Top-10 is achieved when $n = 1$. When $n > 1$, the identification rate decreases gradually. However, the performance in IAM decreases gradually from $n = 0$. The main reason is that documents in IAM are pen-dependent. The writers used different writing instruments to create the handwriting text, which may cause a variation in the derivative along the ink trace. We can conclude from the table that $\Delta^n$Hinge contains less information with a high value of $n$. For example, when $n > 100$, the derivative of two orientations will be closed to zero. Another interesting observation is that, although the performance of the features with different $n$ varies in both two datasets, $\Delta^n$Hinge contains discriminative information when $n \leq 3$.

### 2.4.5 Performance of the $Ho^2D^n$ feature

In this section, the performance of the proposed $Ho^2D^n$ feature which concatenates the $\Delta^n$Hinge with different $n$ is evaluated. The results are presented in Fig. 2.4, where we can find that the maximum Top-1 identification rate is 90.4% on Firemaker when $n = 1$ and 97.2% on IAM when $n = 2$. The corresponding Top-10 identification rates are 98.2% ($n = 4$) on Firemaker and 97.2% ($n = 2$) on the IAM dataset. The results support our conclusion we mentioned before that the $\Delta^n$Hinge contains discriminative information when $0 \leq n \leq 4$.

### 2.4.6 Performance of the $Ho^2D^{n+}$ feature

In this section, the performance of the $Ho^2D^{n+}$ feature is evaluated. The results are shown in Table 2.2. Without the $\Delta^n$Hinge feature, the Top-1 performance decreases comparing to the performance of $Ho^2D^n$. However, the Top-10 performance is still comparable to $Ho^2D^n$. 
2. Writer Identification Using Delta-n Hinge Feature

Figure 2.4: The writer identification performance of different $n$ of the $Ho^2D^n$ feature. The left figure is the performance on the Firemaker dataset, and the right one is on the IAM dataset.

<table>
<thead>
<tr>
<th>$Ho^2D^n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firemaker</td>
<td>Top-1 84.0</td>
<td>84.0</td>
<td>81.4</td>
</tr>
<tr>
<td></td>
<td>Top-10 97.0</td>
<td>97.4</td>
<td>97.2</td>
</tr>
<tr>
<td>IAM</td>
<td>Top-1 85.8</td>
<td>86.4</td>
<td>84.8</td>
</tr>
<tr>
<td></td>
<td>Top-10 96.0</td>
<td>95.3</td>
<td>94.9</td>
</tr>
</tbody>
</table>

Table 2.3: Comparison of writer identification studies on the Firemaker database.

<table>
<thead>
<tr>
<th>Study</th>
<th>Top1(%)</th>
<th>Top10(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghiasi and Safabakhsh</td>
<td>89.2</td>
<td>98.6</td>
</tr>
<tr>
<td>Bulacu and Schomaker</td>
<td>83.0</td>
<td>95.0</td>
</tr>
<tr>
<td>Brink and Smit</td>
<td>86.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Proposed</td>
<td>90.4</td>
<td>98.2</td>
</tr>
</tbody>
</table>

2.4.7 Comparison with other studies

In this section, we present a performance comparison of our method with some recent studies. Table 2.3 and Table 2.4 show the performance of recent studies and our proposed method. The proposed feature performs better than others on the Firemaker data set, which achieves 90.4% (Top-1).

Comparing the performance on the IAM data set, we achieve an identification rate of 93.2% (Top 1) and 97.2% (Top 10), which is better than the results in (Bulacu and Schomaker, 2007; Siddiqi and Vincent, 2010), and comparable to the results in (Ghiasi and Safabakhsh, 2013).
2.5. Conclusion

Table 2.4: Comparison of writer identification studies on the IAM database.

<table>
<thead>
<tr>
<th>Study</th>
<th>Top1(%)</th>
<th>Top10(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siddiqi and Vincent (2010)</td>
<td>89.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Ghiasi and Safabakhsh (2013)</td>
<td>93.7</td>
<td>97.7</td>
</tr>
<tr>
<td>Bulacu and Schomaker (2007)</td>
<td>89.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Brink and Smit (2012)</td>
<td>97.0</td>
<td>98.0</td>
</tr>
<tr>
<td>Proposed</td>
<td>93.2</td>
<td>97.2</td>
</tr>
</tbody>
</table>

Table 2.5: Comparison of writer identification studies with the best results of the ICDAR2013 competition.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>State-of-the-art in ICDAR2013</th>
<th>Top-1</th>
<th>Top-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greek</td>
<td>95.6</td>
<td>99.2</td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>96.0</td>
<td>98.4</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>94.6</td>
<td>99.0</td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>93.4</td>
<td>97.8</td>
<td></td>
</tr>
</tbody>
</table>

Note that Top-1 performance of Quill-Hinge (Brink et al., 2012) is higher on the IAM data set due to the fact that the Quill-Hinge feature is designed for pen-dependent documents.

2.4.8 Comparison with best results of the ICDAR2013 competition

We evaluate the proposed method on the ICDAR2013 database (Louloudis et al., 2013) which is used for writer identification competition. This database consists of 250 writers with four documents per writer. Two documents were written in Greek, the other two in English. Ideally, the parameters of the proposed method should be learned from this dataset. However, in this experiment, we find that Manhattan distance $d = 15$ provides a better result. The results in Table 2.5 show that our proposed method is comparable to the best results of the ICDAR2013 competition.

2.5 Conclusion

We have proposed a new set of features which generalizes the Hinge feature for writer identification in a rotation-invariant manner. The results on two widely used data sets and a comparison with the best results on the ICDAR2013 benchmark show that the proposed method
is promising and comparable to state-of-the-art techniques. The implication of this finding is that not only the (absolute) slant angle distribution of handwriting is biometrically informative; also the distribution of relative angles along the ink trace provides the writer-specific information, capturing the curvature information of handwritten patterns.

The proposed feature in this chapter captures the curvature information of the ink traces. Next chapter will focus on extracting curvature-free features for writer identification, such as the statistical information of the space between words and the line information approximated from writing contours.