Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International 164 (2006) 211–220
Forensic Science International
www.elsevier.com/locate/forsciint

Quantitative characterization of morphological polymorphism of handwritten characters loops

R. Marquis [a,*], F. Taroni [a], S. Bozza [b], M. Schmittbuhl [c]

[a] *Institut de Police Scientifique, School of Criminal Sciences, BCH, University of Lausanne, CH-1015 Lausanne-Dorigny, Switzerland*
[b] *Team Research EA 3428, Dental Faculty, F-67085 Strasbourg, France*
[c] *Department of Statistics, University Ca' Foscari of Venice, San Giobbe, Cannaregio 873, I-30121 Venezia, Italy*

Received 23 May 2005; accepted 1 February 2006
Available online 9 March 2006

Abstract

A methodology based on Fourier descriptors that was previously validated has been applied to 13 writers in order to quantify the polymorphism degree of the shape of the loops of the handwritten characters *a*, *d*, *o* and *q*. In a first step, the discriminating power of the parameters extracted from these letters was investigated. The loop of the letter *d* appeared to be the most discriminant with a correct classification rate of 82.4%, whereas the least discriminant one was the loop of the letter *o* (69.7%). The second aim of the study was to extract grouping characteristics which make it possible to discriminate between writer sets, whatever the letter. Trends in the writing of loops could effectively be shown: the 13 writers of the study were separated into five main groups according to the shape and surface of their loops. The most discriminating features between the writer groups were the importance of the loops elongation and the surface of the loops. Finally, the differences between writers belonging to distinct groups could be characterized more precisely, and differences between writers belonging to the same group were revealed; the individual writings were distinguished by the variability of the parameters of shape and surface of their loops and the morphological distances between its different letters. The correct classification rates reached in this study suggest that carrying out an expertise of fragmentary samples of handwriting comprising only some loops is completely possible.

Keywords: Handwriting; Polymorphism; Shape analysis; Fourier descriptors

1. Introduction

Handwriting examination consists in describing handwriting features, such as elements of style or elements of execution, and studying their range of variation in order to individualize a particular writer through comparison. Characterization of these writing habits as well as the evaluation of the extent of their variability is essentially subjective. Few studies on Roman handwriting were concerned with this lack of objectivity and suggested solutions to provide an objective and quantitative description of writing habits from a forensic point of view. The feature vectors obtained from handwriting documents in previous studies were related to global (based on the handwriting image) [1–4], local (based on zones of interest of the handwriting image, such as lines, words or allographs) [5–12] or both [13,14] aspects of handwriting, but they did not reflect precise visual aspects of handwriting. These studies were merely focused on the development of techniques providing the most accurate identification rates possible, rather than the precise description of handwriting features as they are observed by examiners during the comparison process of handwriting samples.

In a first paper [15], a methodology based on Fourier descriptors was developed, validated and used to precisely characterize and objectively express the within-variability and the between-variability of the parameters of the shape of the loops of handwritten characters *o* in a population of three writers. This procedure was completely new, since the variability of the shape of loops had hitherto been described only in a subjective or partial way [8].

In this further part of the study, the developed methodology has been applied on a larger population of writers to quantify the morphological polymorphism of the loops of the

* Corresponding author. Tel.: +41 21 692 46 00; fax: +41 21 692 46 05.
E-mail address: raymond.marquis@unil.ch (R. Marquis).

0379-0738/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2006.02.008

handwritten characters *a*, *d*, *o* and *q*. Three main steps were accomplished to describe this polymorphism:

- the discriminating power of the parameters of shape and surface of the loops of the letters *a*, *d*, *o* and *q* was investigated and compared;
- then, the similarity of the loops shape between the writers was evaluated, in order to extract grouping characteristics which make it possible to discriminate between writer sets, whatever the letter;
- finally, the distinctive characteristics of each writing could be described according to the morphological distances between its different letters and the variability of their loops shape and surface parameters.

2. Materials and methods

2.1. Sampling

Approximately 100 individuals of the *Institut de Police Scientifique*, University of Lausanne, filled out five documents, where each document had to be written on a different day. On each one of these documents, they had to write 10 times a series of alphabet letters, in their usual way. Paper (standard blank paper of format A4) and pen (ball point pen Bic® Cristal™ with blue color ink) were provided to each participant. Among the collected samples, only the 13 writers showing closed loops for their characters *a*, *d*, *o* and *q* were retained. The total number of observations was 2325 (591 *a* loops, 547 *d* loops, 596 *o* loops and 591 *q* loops).

2.2. Image analysis procedure/size normalization/Fourier analysis

The extraction of the skeletons of the handwritten loops, as well as the size normalization of these skeletons and the Fourier analysis of their shape, were carried out according to the methodology described in detail in reference [15]. In addition, before normalizing the size of the loops, the surface enclosed in the loops was automatically calculated for each character by means of the Visilog 6.0® software.

2.3. Statistical analysis

S-plus® 2000 (Mathsoft Inc.) and SPSS® 12.0 (SPSS Inc.) were used to analyse the numerical data obtained.

Table 1
Surface and Fourier analysis of the handwritten loops *a* of the writers W1–W13: summary statistics[a] of the surface and the first four pairs of Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4)[b]

Writer	Statistics	Surface	A_1	A_2	A_3	A_4	θ_1	θ_2	θ_3	θ_4
W1	*X*	0.065	0.06	1.18	0.25	0.25	337.86	64.51	93.38	70.41
	S.D.	0.012	0.01	0.33	0.12	0.12	29.40	11.82	20.02	14.91
W2	*X*	0.072	0.07	0.43	0.23	0.11	337.38	57.51	93.87	72.63
	S.D.	0.012	0.01	0.23	0.12	0.06	25.03	24.14	22.64	17.81
W3	*X*	0.049	0.06	0.57	0.33	0.21	280.07	34.87	73.96	49.12
	S.D.	0.010	0.03	0.30	0.17	0.08	92.56	35.65	19.23	19.04
W4	*X*	0.058	0.11	1.16	0.31	0.27	350.13	66.85	124.56	73.30
	S.D.	0.018	0.04	0.50	0.14	0.18	45.75	24.07	11.18	20.32
W5	*X*	0.037	0.06	2.02	0.18	0.56	308.59	43.18	128.77	44.68
	S.D.	0.006	0.02	0.27	0.08	0.15	78.71	5.86	13.96	5.66
W6	*X*	0.031	0.13	1.92	0.23	0.46	365.94	51.53	115.93	53.33
	S.D.	0.008	0.05	0.41	0.10	0.18	30.28	7.90	14.03	10.84
W7	*X*	0.025	0.13	1.99	0.21	0.51	341.72	51.90	117.78	53.28
	S.D.	0.007	0.05	0.28	0.08	0.13	18.37	6.56	12.51	8.33
W8	*X*	0.028	0.10	1.93	0.21	0.56	358.56	56.45	123.93	57.91
	S.D.	0.005	0.05	0.23	0.10	0.13	37.15	5.92	15.50	5.83
W9	*X*	0.049	0.05	0.55	0.29	0.17	291.06	67.73	76.47	55.49
	S.D.	0.007	0.02	0.23	0.11	0.07	72.84	20.55	13.95	11.26
W10	*X*	0.147	0.08	1.08	0.21	0.20	339.80	83.50	100.65	80.11
	S.D.	0.054	0.03	0.32	0.08	0.09	35.88	8.23	13.83	8.12
W11	*X*	0.078	0.08	0.98	0.19	0.13	351.69	29.29	77.63	38.88
	S.D.	0.021	0.02	0.23	0.08	0.08	33.93	11.30	11.85	21.79
W12	*X*	0.025	0.19	1.70	0.38	0.50	340.52	70.08	97.05	66.98
	S.D.	0.005	0.05	0.26	0.19	0.10	22.81	6.53	9.83	4.75
W13	*X*	0.016	0.14	1.40	0.22	0.28	339.94	72.00	124.35	75.87
	S.D.	0.004	0.04	0.30	0.09	0.11	18.18	8.37	17.78	8.85

[a] *X*, mean; S.D, standard deviation.
[b] Surface is given in cm^2 and phases are given in degrees.

For each writer and for each letter, the mean and the standard deviation were calculated for the surface and for each pair of Fourier descriptors. All the analyses described below were performed on the surface and the pairs of Fourier descriptors of the contours.

To compare the discriminating power of the four letters *a*, *d*, *o* and *q*, four quadratic discriminant analyses were performed on the 13 writers. The estimation of the discriminating power was given by the rates of correct classification within the observations retained for validation, which represented 20% of the overall data.

Then, a quadratic discriminant analysis was applied on the 13 writers considering the four letters simultaneously. In order to estimate the between-writers variability, Mahalanobis distances between each pair of writers – a writer being characterized by his set of *a*, *d*, *o* and *q* loops – were calculated. Then, the Hotelling's *T* squared test was used to test differences in means between each possible pair of writers.

Finally, a quadratic discriminant analysis was separately applied on each writer considering the four letters simultaneously. Mahalanobis distances were calculated between each pair of letter sets for each writer. Euclidean distances were calculated between pairs of observations within each letter set for each writer; the within-writer variability was estimated through the mean and the standard deviation of these distances.

3. Results

The statistics of the Fourier descriptors (amplitudes and phases) of the handwritten characters *a*, *d*, *o* and *q* contours of each writer are summarised in Tables 1–4. Only the first four pairs of Fourier descriptors were retained, since the global shape of each character contour was practically reconstructed on the basis of these four Fourier harmonics, and because from the fifth harmonic, for the majority of the writers, the phase values were randomly distributed and were not specific of the writing shape of the characters loops.

3.1. Discriminating power of the shape of handwritten characters loops

An attempt to discriminate between the 13 writers of the study was conducted in applying a discriminant analysis on the data of each one of the letters *a*, *d*, *o* and *q*. The corresponding correct classification rates calculated with the observations retained for validation were 74.8%, 82.4%, 69.7% and 81.4%,

Table 2
Surface and Fourier analysis of the handwritten loops *d* of the writers W1–W13: summary statistics[a] of the surface and the first four pairs of Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4)[b]

Writer	Statistics	Surface	A_1	A_2	A_3	A_4	θ_1	θ_2	θ_3	θ_4
W1	*X*	0.055	0.12	1.03	0.24	0.22	339.35	62.33	97.81	67.92
	S.D.	0.012	0.05	0.40	0.10	0.13	27.36	11.71	18.51	17.03
W2	*X*	0.064	0.06	0.33	0.34	0.12	370.05	122.05	106.00	80.33
	S.D.	0.013	0.02	0.20	0.14	0.06	50.13	48.26	10.07	13.78
W3	*X*	0.038	0.10	0.77	0.16	0.17	308.66	18.13	59.52	31.41
	S.D.	0.010	0.03	0.30	0.11	0.09	27.57	18.96	41.22	28.93
W4	*X*	0.052	0.11	1.47	0.35	0.25	365.69	66.06	133.31	71.90
	S.D.	0.013	0.05	0.39	0.13	0.14	39.65	10.65	8.42	17.35
W5	*X*	0.040	0.06	1.90	0.31	0.46	345.52	45.89	131.34	45.66
	S.D.	0.007	0.03	0.24	0.08	0.14	64.77	4.62	5.95	4.89
W6	*X*	0.026	0.09	1.84	0.26	0.39	339.99	42.61	123.04	42.10
	S.D.	0.007	0.05	0.42	0.09	0.20	56.62	5.78	10.36	10.04
W7	*X*	0.017	0.13	2.27	0.29	0.62	314.75	49.85	129.76	49.27
	S.D.	0.005	0.07	0.34	0.09	0.19	22.17	4.17	13.25	5.29
W8	*X*	0.021	0.10	2.12	0.33	0.48	360.47	55.13	135.51	55.53
	S.D.	0.004	0.06	0.29	0.09	0.16	55.43	3.94	7.24	6.01
W9	*X*	0.059	0.07	0.49	0.15	0.11	301.35	49.42	62.72	48.04
	S.D.	0.007	0.02	0.18	0.08	0.06	30.51	18.78	29.53	22.70
W10	*X*	0.117	0.05	0.63	0.25	0.12	374.46	64.46	90.84	62.99
	S.D.	0.034	0.03	0.28	0.11	0.07	62.56	19.01	14.06	17.74
W11	*X*	0.071	0.08	1.76	0.20	0.45	400.00	−3.78	77.40	84.22
	S.D.	0.021	0.04	0.35	0.08	0.16	64.50	15.60	20.42	15.74
W12	*X*	0.025	0.12	1.21	0.36	0.18	343.00	63.50	130.98	77.55
	S.D.	0.004	0.05	0.32	0.12	0.09	39.43	6.67	13.25	15.34
W13	*X*	0.013	0.11	1.62	0.31	0.27	303.66	64.97	142.36	71.93
	S.D.	0.004	0.06	0.39	0.13	0.15	49.63	5.55	12.59	14.32

[a] *X*, mean; S.D, standard deviation.
[b] Surface is given in cm^2 and phases are given in degrees.

Table 3
Surface and Fourier analysis of the handwritten loops *o* of the writers W1–W13: summary statistics[a] of the surface and the first four pairs of Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4)[b]

Writer	Statistics	Surface	A_1	A_2	A_3	A_4	θ_1	θ_2	θ_3	θ_4
W1	*X*	0.059	0.13	1.36	0.25	0.31	329.36	57.98	87.07	64.25
	S.D.	0.017	0.05	0.34	0.13	0.12	26.64	8.96	17.46	9.65
W2	*X*	0.083	0.06	0.53	0.29	0.11	368.12	12.97	85.69	83.70
	S.D.	0.013	0.02	0.22	0.13	0.06	48.22	22.40	14.85	21.12
W3	*X*	0.045	0.07	0.62	0.38	0.18	309.04	126.88	80.23	58.38
	S.D.	0.010	0.03	0.29	0.16	0.07	83.68	39.44	14.86	29.21
W4	*X*	0.071	0.11	1.03	0.19	0.23	329.21	76.78	96.19	75.57
	S.D.	0.018	0.06	0.44	0.12	0.12	23.77	19.20	29.33	15.22
W5	*X*	0.043	0.07	1.18	0.13	0.29	296.85	54.72	122.88	53.12
	S.D.	0.007	0.02	0.26	0.07	0.08	33.61	6.22	41.32	6.31
W6	*X*	0.028	0.12	1.03	0.17	0.16	333.05	50.64	114.23	71.04
	S.D.	0.010	0.05	0.39	0.10	0.09	21.28	18.20	35.43	19.26
W7	*X*	0.030	0.14	1.70	0.23	0.35	350.32	51.81	100.50	52.68
	S.D.	0.004	0.04	0.36	0.10	0.16	23.67	10.04	8.48	15.04
W8	*X*	0.026	0.12	1.48	0.18	0.30	338.09	62.15	120.92	64.72
	S.D.	0.006	0.05	0.37	0.09	0.14	28.06	8.92	22.43	10.90
W9	*X*	0.065	0.04	0.36	0.21	0.10	313.96	119.00	78.21	61.30
	S.D.	0.009	0.02	0.18	0.11	0.05	61.34	39.73	20.25	17.52
W10	*X*	0.114	0.08	0.85	0.19	0.14	334.31	71.75	92.48	70.46
	S.D.	0.039	0.03	0.30	0.09	0.08	23.91	16.10	26.32	15.91
W11	*X*	0.074	0.08	1.00	0.30	0.24	398.71	−2.05	97.03	86.39
	S.D.	0.021	0.03	0.37	0.12	0.13	49.90	13.06	14.31	7.84
W12	*X*	0.027	0.12	0.75	0.24	0.13	328.60	53.73	126.58	66.61
	S.D.	0.006	0.04	0.26	0.11	0.09	19.51	33.48	30.44	14.60
W13	*X*	0.022	0.09	0.99	0.17	0.15	307.82	83.47	109.11	83.23
	S.D.	0.005	0.03	0.31	0.09	0.09	23.03	11.55	40.52	16.72

[a] *X*, mean; S.D, standard deviation.
[b] Surface is given in cm^2 and phases are given in degrees.

respectively. Furthermore, the multivariate means of the loops parameters based on each one of the four letters were different in a highly significant way according to the Hotelling's *T* squared test ($p < 0.01$) between each possible pair of writers.

3.2. Morphological characterization of groups of loops writings

The main groups of writers formed when applying a discriminant analysis on the features pertaining to one letter or another one were practically the same. Consequently, a discriminant analysis was performed on the four letters considered simultaneously in order to minimise the differences between the writers which were specific to one letter, thus allowing for the extraction of shape characteristics shared by different writers whatever the letter.

In this case, the first discriminant function accounted for 58.6% of the total variance, the second one 17.9%. The first two functions together explained 76.5% of the total variance of the data set. Within the observations retained for validation, 66.8% were correctly classified, i.e. allocated to the adequate writer. Once again, according to the Hotelling's *T* squared test (at $p < 0.01$), the differences in the multivariate means of the loop parameters between the writers were highly significant.

The graphic representation of the first two axes of the discriminant analysis (see Fig. 1) and the comparison of the Mahalanobis distances between the writers (see Table 5) suggested the constitution of five main groups among the 13 writers of the study: one group of one writer (W10), three groups of two writers (W1-W4, W2-W11 and W3-W9) and one group of six writers (W5-W6-W7-W8-W12-W13). The writing groups sharing shape characteristics are illustrated by the reconstructions of any *d* loops in Fig. 2, on the basis of the first four pairs of Fourier descriptors.

The first two discriminant functions were principally correlated with the same variables, which were the surface, the amplitude of the second harmonic (A_2) and the phase of the third harmonic (θ_3) (see Table 6).

The group formed by the writers W5, W6, W7, W8, W12 and W13 was distinguished from the other groups by the fact that these writers presented, in average, a very high value of the second amplitude of the loops of all the letters analysed. That was especially marked for the letter *a*: 2.02, 1.92, 1.99, 1.93, 1.70 and 1.40, respectively (see A_2 in Table 1). These high values of the second amplitude indicated that the loops of these writers presented an important elongation. In addition, the loops of these writers presented very small surfaces, especially for the letter *q*: 0.031, 0.025, 0.020, 0.020, 0.022 and 0.015 cm^2, respectively (see surface in Table 4). Finally, these six writers were also characterized by high values of the third phase, more than 100° in average (see θ_3 in Tables 1–4), this means that the triangular contribution of their loops was slightly backward oriented.

The writers group W3–W9 was characterized by very low values of the second amplitude of the letters *a*, *d*, *o* and *q*. This amplitude was particularly reduced for the letter *q*: 0.50 and 0.47, respectively (see A_2 in Table 4). Thus, the loops of these writers presented a weak elongation. In addition, in average, the loops of these writers presented the lowest values of the third phase; one leaf of the triangular contribution of their loops was oriented at 68° in average (see θ_3 in Tables 1–4).

The writers W2–W11 were discriminated from the other writers groups by the high values of the surface of their loops. This characteristic led to associate these writers even if the values for the second amplitude were quite different between them; elongation was rather marked in the loops of the writer W11, contrary to those of writer W2 (see A_2 in Tables 1–4).

The group W1–W4 was characterized by moderate values of all the loops parameters which played an important role in the discrimination between-writers groups, namely the surface, the second amplitude and the third phase (see surface, A_2 and θ_3 in Tables 1–4).

The writer W10 constituted a group by himself, being characterized by the highest values of the surface of all his loops, in a very marked way for the letter *a*: 0.147 cm^2 (see surface in Table 1). Furthermore, the elongation of the loops of the writer W10 was not very marked, as indicated by the rather low value of the second amplitude of the loops of this writer: *a*: 1.08; *d*: 0.63; *o*: 0.85; *q*: 0.94 (see A_2 in Tables 1–4).

3.3. Morphological characterization of individual loops writings

Precise characterization of individual writings was then investigated, on the basis of the relationship between the shape and surface parameters of the loops of the different letters in each writing, as well as the extent of their variability. Four

Table 4
Surface and Fourier analysis of the handwritten loops *q* of the writers W1–W13: summary statistics[a] of the surface and the first four pairs of Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4)[b]

Writer	Statistics	Surface	A_1	A_2	A_3	A_4	θ_1	θ_2	θ_3	θ_4
W1	*X*	0.044	0.10	1.14	0.22	0.24	331.39	54.16	85.81	61.00
	S.D.	0.009	0.04	0.44	0.09	0.13	30.23	12.60	17.64	10.60
W2	*X*	0.054	0.06	0.88	0.24	0.14	396.57	31.60	96.23	83.77
	S.D.	0.012	0.02	0.33	0.12	0.08	41.46	48.85	17.20	20.97
W3	*X*	0.037	0.05	0.50	0.26	0.12	303.56	26.33	59.75	41.61
	S.D.	0.010	0.03	0.29	0.13	0.08	77.30	33.68	19.75	24.73
W4	*X*	0.038	0.10	1.52	0.27	0.33	355.53	54.68	105.03	78.92
	S.D.	0.011	0.04	0.49	0.13	0.20	38.17	15.05	17.33	24.26
W5	*X*	0.031	0.06	1.38	0.13	0.37	340.73	38.69	122.97	40.74
	S.D.	0.006	0.02	0.27	0.08	0.13	40.74	5.79	29.46	3.46
W6	*X*	0.025	0.10	1.82	0.25	0.35	402.74	32.39	105.88	33.55
	S.D.	0.005	0.04	0.44	0.10	0.18	32.22	9.46	9.68	15.73
W7	*X*	0.020	0.10	1.94	0.20	0.59	358.71	41.22	92.18	40.79
	S.D.	0.004	0.03	0.39	0.07	0.24	23.49	7.55	15.19	6.63
W8	*X*	0.020	0.13	2.01	0.30	0.47	389.13	41.45	118.80	43.73
	S.D.	0.004	0.05	0.34	0.12	0.19	30.92	5.18	10.03	6.63
W9	*X*	0.057	0.03	0.47	0.23	0.15	294.90	70.43	55.21	38.96
	S.D.	0.007	0.03	0.25	0.10	0.07	113.49	21.76	18.07	13.86
W10	*X*	0.102	0.06	0.94	0.21	0.16	311.47	74.91	87.34	71.90
	S.D.	0.028	0.03	0.36	0.11	0.08	59.66	13.12	22.47	13.84
W11	*X*	0.064	0.08	1.32	0.14	0.26	330.07	12.60	57.01	102.01
	S.D.	0.016	0.03	0.31	0.07	0.10	38.02	12.11	25.30	13.16
W12	*X*	0.022	0.19	1.73	0.46	0.52	345.57	63.72	91.38	61.22
	S.D.	0.003	0.04	0.28	0.16	0.16	27.25	7.25	6.36	5.33
W13	*X*	0.015	0.13	1.20	0.21	0.25	337.25	61.95	120.57	71.83
	S.D.	0.004	0.04	0.34	0.12	0.11	28.67	9.11	21.91	11.07

[a] *X*, mean; S.D, standard deviation.
[b] Surface is given in cm^2 and phases are given in degrees.

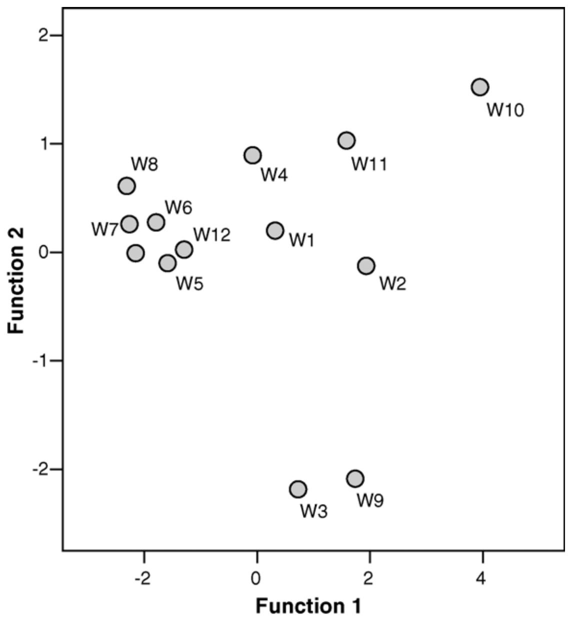

Fig. 1. Results of the discriminant analysis performed on the first four pairs of Fourier descriptors (A_1–A_4, θ_1–θ_4) and the surface of the contours of the loops *a*, *d*, *o* and *q* of the writers W1–W13: representation of the centroid of the loops of each writer on the first two discriminant functions.

writers were chosen within the 13, belonging to the same or different writers groups: W1, W5, W11 and W12. These writers were selected because they presented morphological characteristics appearing clearly different between them through the visualisation of their corresponding discriminant analysis (see Fig. 3). For each one of these writers and for each letter, Euclidean distances calculated between each possible pair of observations are illustrated in Fig. 4, while the Mahalanobis distances between the letters of each writer separately are given in Table 7 for these writers. Table 8 informs about the variables responsible for the discrimination between the letters in each writer.

In writer W1, the within-variability of the parameters of surface and shape within the set of observations for a letter was rather weak for all four sets of letters (see Fig. 4). However, the four letters clouds of points represented on the discriminant graph overlapped (see Fig. 3), since the distances between these sets were very low. The loops of the different letters were morphologically very similar, in the most marked way for the letters *a* and *d*, as indicated by the particularly small distance separating them (see Table 7). Nevertheless, the loops of the letter *q* could be slightly distinguished from the other letters by their smaller surface and their lower second phase (see surface and θ_2 in Tables 1–4); this means that elongation of *q* loops was closer to the horizontal axis compared to that of the other letters. Furthermore, W1 is the only writer in which the letter *o* is the most elongated letter, whereas this letter is the least elongated in the majority of the writers (see A_2 in Tables 1–4).

In writer W5, the variability of the loop parameters was the most extended for the letter *a* (see Fig. 4). The loops of letters *a* and *d* showed almost the same morphology, as demonstrated by the superimposition of their corresponding clouds of points on the discriminant graph, and the small distance separating them. On the contrary, the letters *o* and *q* were well discriminated and could be distinguished from the letters *a* and *d* (see Fig. 3 and Table 7). Letters *a* and *d* were characterized by a more pronounced elongation, as well as a more marked quadrangularity (see A_2 and A_4 in Tables 1–4). The triangularity of the letter *d* was two times more marked than that of the other letters (see θ_3 in Tables 1–4). Moreover, the letters *o* and *q* were mainly separated by their difference in orientation of their elongation (see θ_2 in Tables 3 and 4) and their quadrangular contribution (see θ_4 in Tables 3 and 4).

In writer W11, the greatest extent of variability of the loop parameters was observed in the letter *d* (see Fig. 4). The groups of the letters *d* and *q* were very close and practically superimposed. The distance between the group of the letter *a* and each one of the other letters was extremely high (see Fig. 3 and Table 7). The letters *d* and *q* presented both a distinctive orientation of the elongation which was almost horizontal (see θ_2 in Tables 2 and 4). The loops of the letter *a* were separated from the other letters by the low values of their second and fourth phases (see θ_2 and θ_4 in Tables 1–4) as well as the low value of their fourth amplitude (see A_4 in Tables 1–4). The letter *o* is distinguished from the other letters by its higher

Table 5
Discriminant analysis of the surface and the first four pairs of Fourier descriptors of the handwritten loops *a*, *d*, *o* and *q* of the writers W1–W13: Mahalanobis distances between each pair of writers

	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13
W1	0	6.32	7.96	1.67	8.86	5.75	8.48	8.90	8.72	16.45	7.57	5.06	8.53
W2		0	9.24	6.53	17.31	17.17	23.05	21.59	6.84	11.73	6.55	15.21	19.13
W3			0	12.64	13.82	14.17	17.26	19.46	2.41	26.63	14.84	13.23	18.14
W4				0	7.26	4.99	8.23	6.55	12.92	17.40	9.03	4.43	6.41
W5					0	3.52	4.06	3.64	16.95	32.03	17.68	8.57	8.31
W6						0	1.56	0.96	19.13	35.59	16.21	3.24	3.71
W7							0	0.92	23.01	41.35	20.03	4.62	6.48
W8								0	24.51	40.61	20.46	4.06	4.41
W9									0	18.32	15.88	18.32	21.91
W10										0	14.65	36.61	42.30
W11											0	18.55	23.27
W12												0	2.98
W13													0

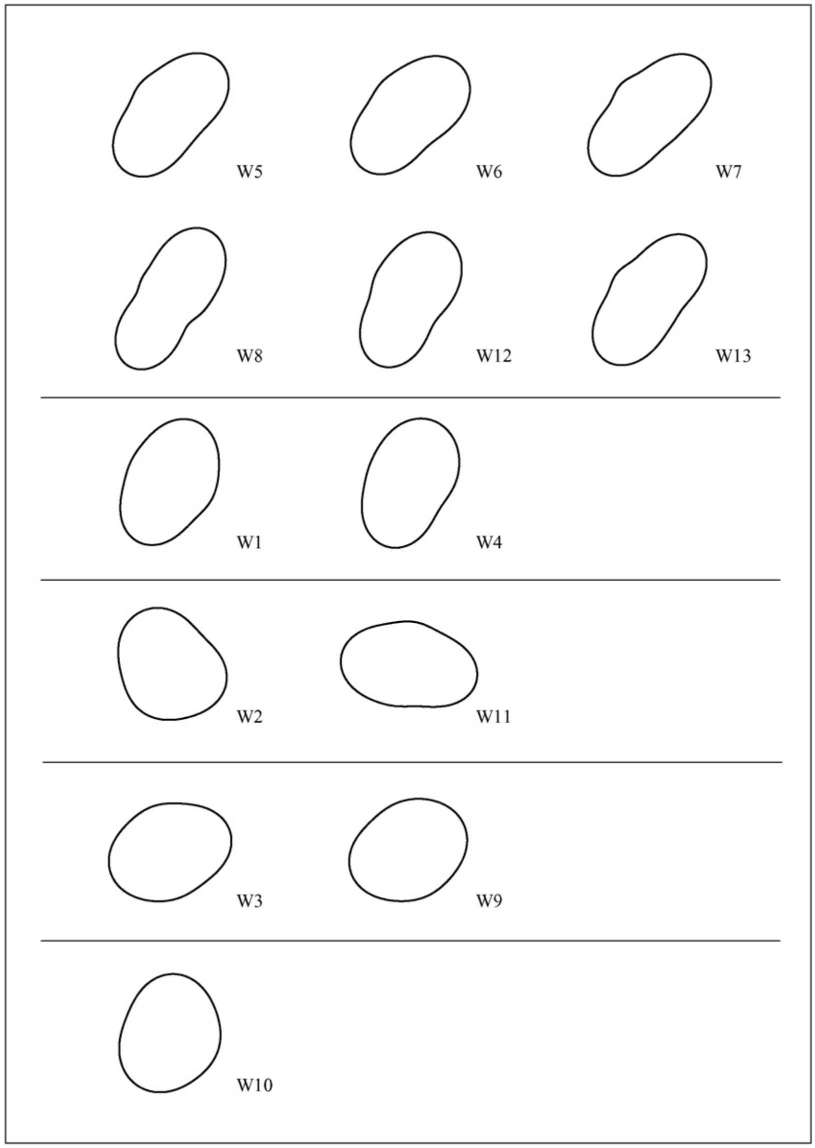

Fig. 2. Examples of reconstructions of d loops on the basis of the first four pairs of Fourier descriptors, one character being shown for each writer (W1–W13), to illustrate the five main groups of shape tendencies in loops writings (size is not illustrated here since it was normalized for the Fourier analysis).

Table 6
Discriminant analysis of the surface and the first four pairs of Fourier descriptors of the handwritten loops a, d, o and q of the writers W1–W13: correlation coefficient of the surface (S) and the Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4) with the first (r_1) and the second (r_2) discriminant functions

	r_1	r_2
S	0.796	0.466
A_1	−0.268	0.231
A_2	−0.438	0.548
A_3	−0.029	−0.022
A_4	−0.311	0.251
θ_1	−0.001	0.313
θ_2	0.024	−0.122
θ_3	−0.320	0.413
θ_4	0.111	0.372

third phase and its pronounced triangularity (see θ_3 and A_3 in Tables 1–4). The writer W11 is the only one in which the letter o is the most triangular.

In writer W12, the within-variability of the loops parameters of each letter was weak, especially for a and q (see Fig. 4). In this writer, observations of these letters presented a similar morphology, as demonstrated by the very low distance between these two letters groups. The letters d and o were distinctively separated from the other letter groups (see Fig. 3 and Table 7). The loops of the letter o were discriminated by their weaker elongation (see A_2 in Tables 1–4) and their less pronounced quadrangularity (see A_4 in Tables 1–4). Additionally, the loops of the letter d were characterized by a slightly less pronounced triangularity, as demonstrated by the lower value of their third amplitude compared to that of the letters a, o and q (see A_3 in Tables 1–4). In comparison with the other writers, W12 is all the same the writer whose loops are the most triangular on average.

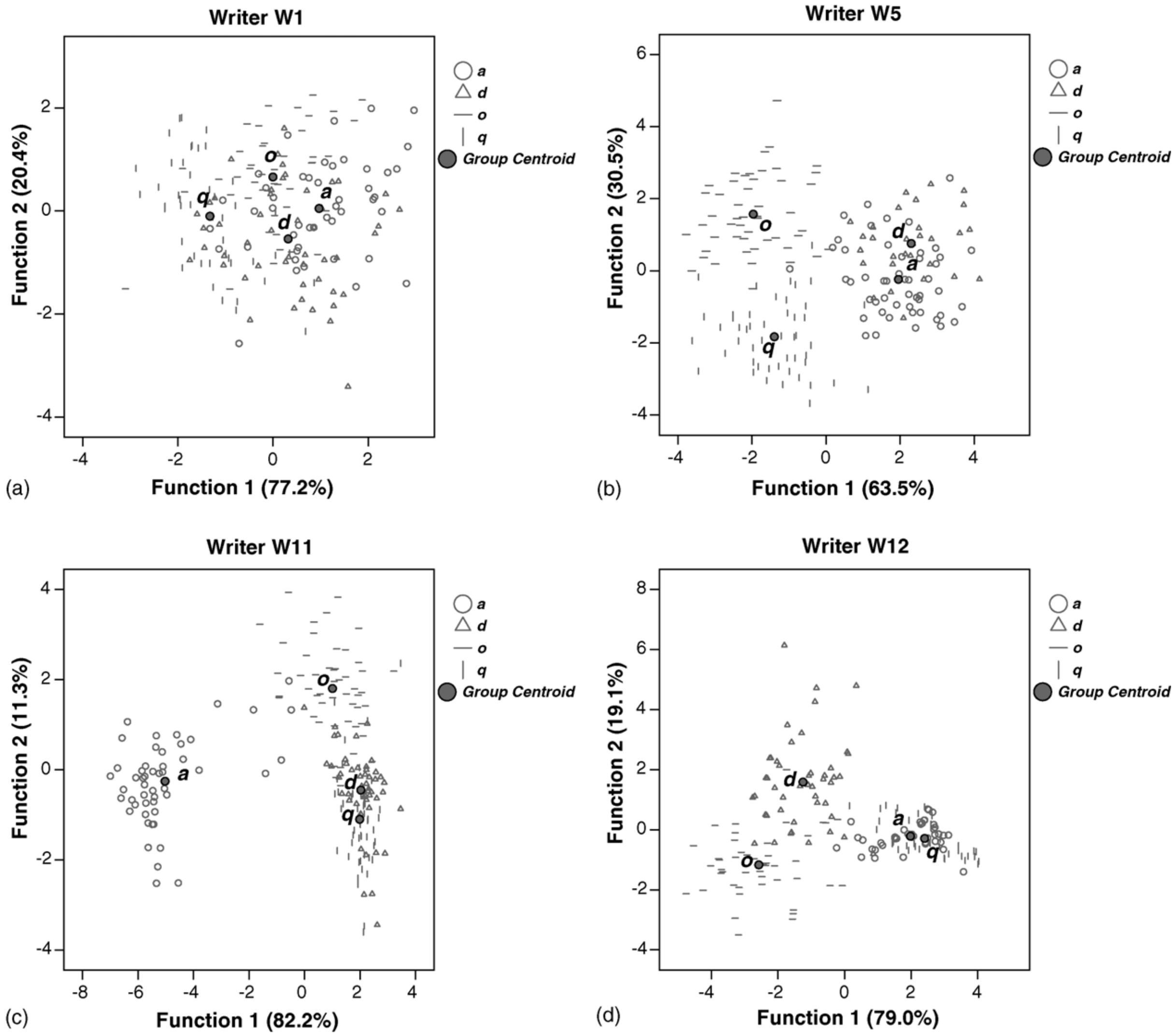

Fig. 3. Results of the discriminant analyses performed on the first four pairs of Fourier descriptors (A_1–A_4, θ_1–θ_4) and the surface of the contours of the loops *a*, *d*, *o* and *q* of the writer W1 (a), W5 (b), W11 (c) and W12 (d). Percentages given in parenthesis are the rates of variance explained by the corresponding functions.

4. Discussion

The methodology developed in a previous study to characterize objectively the global shape of the loops of handwritten characters was successfully validated. This procedure has now been applied to the loops of letters of other writers who were selected according to their habit to close their loops, and not on the basis of visual and subjective differences in the shape of their loops. The loops of the letters *a*, *d*, *o* and *q* were retained, as they all are circular bowls [16] and the first construction element of the corresponding letter.

This approach is new since it describes the shape of handwritten loops by decomposing it into specific contributions being visually understandable, and since variability of these parameters – within and between the writers, as well as within and between letters of a given writer – were detailed and quantified, allowing for an objective characterization of handwriting features, as well as their variability. The shape of loops of the handwritten characters *a*, *d*, *o* and *q* was described by using the Fourier descriptors. The decision criteria for determining the number of pairs of Fourier descriptors to be retained were based on the will to characterize the global shape of the loops without introducing random information assimilable to noise.

The discriminating power of the loops within the 13 writers of the study was different according to the letter. The loop of the letter *o* appeared to be the least discriminant one, as demonstrated by its lowest correct classification rate (69.7%). On the contrary, the loop of the letter *d* showed more individual characteristics of shape and size, as demonstrated by its better efficiency in the discrimination between the writers (correct classification rate of 82.4%). The order of these results is in agreement with the order of identification performances obtained by using micro-features of these letters (gradient, structural and concavity features extracted on the characters images) [17]. From these differences in discriminating power according to the letter, we deduce that all the individuals of the study did not form the loops of various letters in the same way; in a writer, shape parameters of various letters loops could be different.

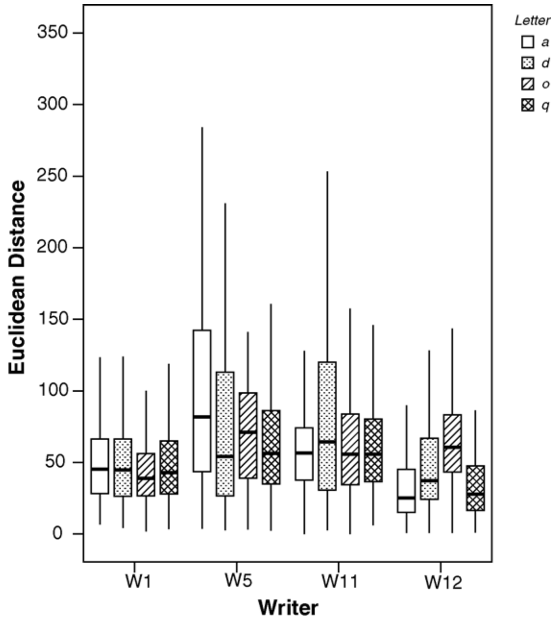

Fig. 4. Boxplots of Euclidean distances between all possible pairs of observations for each letter and each one of the writers W1, W5, W11 and W12. For a purpose of clarity, outliers were omitted from the figure.

Table 7
Variability of the surface and the first four pairs of Fourier descriptors of the loops *a*, *d*, *o* and *q* between each letter group of writers W1, W5, W11 and W12: Mahalanobis distances between each pair of letters groups (n = number of observations)

	n		a	d	o	q
W1	47	a	0	0.89	1.41	5.29
	49	d		0	1.55	2.96
	44	o			0	2.39
	46	q				0
W5	50	a	0	4.11	18.97	14.66
	30	d		0	20.25	21.02
	50	o			0	12.03
	49	q				0
W11	50	a	0	51.21	40.73	50.98
	50	d		0	8.52	5.56
	50	o			0	9.89
	50	q				0
W12	41	a	0	13.91	21.88	1.09
	47	d		0	9.35	17.04
	46	o			0	25.74
	40	q				0

Categories of loop writings presenting similar parameters of shape and surface of the loops *a*, *d*, *o* and *q* were established on the basis of discriminant analysis and between-writers distances. Then, even if all the writers did not form all the loops in the same way whatever the letter, main tendencies in size and shape could be highlighted and allowed for the characterization of groups of writers. The degree of dissimilitude between the groups of writers, as well as the degree of similitude between writers belonging to the same group, were quantified by the Mahalanobis distances calculated between the writers (see Table 5). The between-writers variability of the parameters of shape and surface of the loops of the letters *a*, *d*, *o* and *q* could thus be estimated in an objective way. The most discriminative features could be extracted from the discriminant analysis. In particular, the surface and the amplitude of the second harmonic were the variables contributing most to the separation or lack thereof between the writers. In other words, among the most discriminative characteristics were the size of the loops and the importance of their elongation.

The characteristics of the loops writing of each writer could be highlighted. The differences between writers belonging to distinct groups could be detailed, but it was also possible to reveal individual differences between writers belonging to the same group (see Fig. 3). For each writing, the variability of the size and shape parameters of the loops within each letter could be evaluated and compared (see Fig. 4). This variability, which is a quantitative representation of the natural variation (or consistency) of some aspects of handwriting, was different according to the writer and the letter. These results are in agreement with the hypothesis of the individuality of variations ranges of handwriting features [16]. Another major writing characteristic was the degree of morphological proximity between the loops of the different letters (see Table 7); the letters showing the strongest similitude were not always the same ones between the writers,

Table 8
Discriminant analyses of the surface and the first four pairs of Fourier descriptors of the handwritten loops *a*, *d*, *o* and *q* of the writers W1, W5, W11 and W12: correlation coefficient of the surface (S) and the Fourier amplitudes (A_1–A_4) and phases (θ_1–θ_4) with the first (r_1) and the second (r_2) discriminant functions

Variables	Writer W1		Writer W5		Writer W11		Writer W12	
	r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2
S	0.680	0.409	0.713	−0.148	−0.066	0.129	−0.129	−0.040
A_1	0.342	0.231	−0.068	0.177	0.028	0.028	0.353	−0.157
A_2	0.004	0.722	0.713	−0.148	0.209	−0.450	0.666	0.284
A_3	0.132	0.100	0.341	0.147	0.000	0.582	0.226	0.185
A_4	0.009	0.612	0.415	−0.171	0.236	−0.149	0.723	−0.195
θ_1	0.118	−0.293	0.042	−0.170	0.062	0.382	0.080	0.144
θ_2	0.405	−0.207	−0.173	0.774	−0.293	−0.376	0.122	0.137
θ_3	0.225	−0.441	0.065	0.013	−0.032	0.615	−0.435	0.298
θ_4	0.309	−0.162	−0.147	0.646	0.500	−0.056	−0.128	0.409

and the between-letters distances were different according to the writer. These differences in values and variability of the parameters of the loops between different letters, which were specific for a writer, suggest that it is of interest to compare the loops of the same letter when examining handwritten documents, and that each letter can provide additional information for writer identification. Indeed, the shape of a letter does not allow for the prediction of the shape of another letter, at least not for every writer.

A polymorphism of size and shape of loops of letters *a*, *d*, *o* and *q* could be shown in a quantitative and objective way. These parameters were useful to characterize peculiarities of any handwriting and to show general trends in the shape of handwritten loops; the most discriminative contributions could be highlighted, explaining the differentiation between-writers groups or between letters of a writer. Moreover, the high correct classification rates obtained showed that the retained parameters are also useful to associate loops with their corresponding writer. Then, in average, the variability of the shape and surface parameters were higher between the writers than within the writers.

Amplitudes and phases of the Fourier harmonics, directly related to precise contributions to the shape of loops, as well as size, were very relevant to the understanding of the differences between the writers groups or between the letters of each writer. Furthermore, from a practical point of view related to handwriting examination, one can conclude from the differences in discriminating power between the letters that it would be advisable to grant less weight to a correspondence of the shape of letters *o* than to that of letters *d* when comparing handwritten documents for writer identification.

In addition, characterization of local features, measured on the allographic level, is very relevant in order to examine documents containing little handwritten material. On the contrary, methods based on global features, such as texture analysis [1], require a sufficient amount of handwritten material that allows for a reliable description of the style of the questioned handwriting [18]. The correct classification rates reached in this study suggest that carrying out an expertise of fragmentary samples of handwriting comprising only some loops is completely possible.

Acknowledgement

This work was supported by a research grant (11–66787.01) from the Swiss National Science Foundation.

References

[1] H. Said, T. Tan, K. Baker, Personal identification based on handwriting, Pattern Recognit. 33 (2000) 149–160.

[2] M. Bulacu, L. Schomaker, Writer style from oriented edge fragments, Proceedings of the Tenth International Conference on Computer Analysis of Images and Patterns (CAIP 2003), Groningen, The Netherlands, 2003, pp. 460–469.

[3] A. Seropian, M. Grimaldi, N. Vincent, Writer identification based on the fractal construction of a reference base, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 1163–1167.

[4] M. Wirotius, A. Seropian, N. Vincent, Writer identification from gray level distribution, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 1168–1172.

[5] E.N. Zois, V. Anastassopoulos, Fusion of correlated decisions for writer verification, Pattern Recognit. (2001) 47–61.

[6] A. Bensefia, A. Nosary, T. Paquet, L. Heutte, Writer identification by writer's invariants, in: Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR 2002), Niagara-on-the-Lake, Canada, 2002, pp. 274–279.

[7] C. Hertel, H. Bunke, A set of novel features for writer identification, Lecture Notes Comput. Sci. 2688 (2003) 679–687.

[8] G. Leedham, S. Chachra, Writer identification using innovative binarised features of handwritten numerals, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 413–417.

[9] B. Zhang, S.N. Srihari, Analysis of handwriting individuality using word features, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 1142–1146.

[10] L. Schomaker, M. Bulacu, Automatic writer identification using connected-component contours and edge-based features of uppercase western script, IEEE Trans. Pattern Anal. Mach. Intell. 26 (2004) 787–798.

[11] M. Tapiador, J.A. Sigüenza, Writer identification method based on forensic knowledge, in: Proceedings of the International Conference on Biometric Authentication (ICBA 2004), Hong Kong, China, 2004, pp. 555–561.

[12] C.-K. Li, N.-L. Poon, W.-K. Fung, C.-T. Yang, Individuality of handwritten numerals in local population, J. Forensic Sci. 50 (2005) 185–191.

[13] U.-V. Marti, R. Messerli, H. Bunke, Writer identification using text line based features, in: Proceedings of the Sixth International Conference on Document Analysis and Recognition (ICDAR 2001), Seattle, Washington, 2001, pp. 101–105.

[14] S.N. Srihari, S.-H. Cha, H. Arora, S. Lee, Individuality of handwriting, J. Forensic Sci. 47 (2002) 857–872.

[15] R. Marquis, M. Schmittbuhl, W.D. Mazzella, F. Taroni, Quantification of the shape of handwritten characters: a step to objective disrimination between writers based on the study of the capital character O, Forensic Sci. Int. 150 (2005) 23–32.

[16] R. Huber, A. Headrick, Handwriting Identification Facts and Fundamentals, CRC Press, Boca Raton, 1999, pp. 100 and 133–134.

[17] B. Zhang, S.N. Srihari, S. Lee, Individuality of handwritten characters, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 1086–1090.

[18] A. Bensefia, T. Paquet, L. Heutte, Information retrieval based writer identification, in: Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), Edinburgh, Scotland, 2003, pp. 946–950.