HAL
archives-ouvertes.fr

Training-and Segmentation-Free Intuitive Writer Identification with Task-Adapted Interest Points

Angelika Garz, Marcel Würsch, Rolf Ingold

▶ **To cite this version:**

Angelika Garz, Marcel Würsch, Rolf Ingold. Training-and Segmentation-Free Intuitive Writer Identification with Task-Adapted Interest Points. Céline Rémi; Lionel Prévost; Eric Anquetil. 17th Biennial Conference of the International Graphonomics Society, Jun 2015, Pointe-à-Pitre, Guadeloupe. 2015, Drawing, Handwriting Processing Analysis: New Advances and Challenges.

HAL Id: hal-01165881
https://hal.univ-antilles.fr/hal-01165881

Submitted on 20 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Training- and Segmentation-Free Intuitive Writer Identification with Task-Adapted Interest Points

Angelika GARZ[a], Marcel WÜRSCH[a] and Rolf INGOLD[a]
[a] *University of Fribourg, DIVA*
1700 Fribourg; Switzerland
[angelika.garz, marcel.wuersch, rolf.ingold]@unifr.ch

Abstract. Identifying the writer of a document establishes its authenticity or authorship and has several applications, notably in forensic and historical document analysis. Previous research has shown the potential of Interest Points (IP) for writer identification, but existing methods require segmentation or training. This paper evaluates the performance of intuitive features computed directly from IP properties rather than extracting descriptors at their locations; allowing for a training-free approach. Secondly, we show that adapting detectors to the specific task of writer identification is not only vital for performance but also allows for segmentation-free approaches. Experiments on widely-used datasets show the potential of the method applied self-contained and when combined with existing methods. Limitations of our method relate to the amount of data needed in order to obtain reliable models.

1. Introduction

Applications of Writer Identification (WI) are manifold. In handwriting recognition writer-dependent models tailored to the personal writing style allow for improved performance. Knowing a historical manuscript's scribe, its history, origin, and authenticity can be determined. In forensic investigations concerning fraud, homicide, suicide, or the execution of a last will, establishing the genuineness of a document is a task that often arises.

Current research in WI focuses on two problems: improving identification performance especially on large datasets (Louloudis et al., 2013), on the one hand; and explainability of system results, i.e., rendering the decision-making comprehensible to a human (Niels & Vuurpijl, 2005), on the other hand. Drawbacks of existing methods are the need for binarization or segmentation; open problems themselves (Pratikakis et al., 2013). IP-based methods offer the potential of circumventing the aforementioned drawbacks. Such methods detect salient points in handwriting in order to compute a descriptor at their locations. Using a so-called codebook of clustered descriptors from an independent training set, a probability distribution of descriptors is computed to characterize a writer. Literature focused on using various codebook clustering methods (Fiel & Sablatnig, 2012), or developing different descriptors and combinations (Jain & Doermann, 2014). A recent approach combines a codebook of IP descriptors and a histogram of scales and orientations computed from IP (Wu et al., 2014). For comprehensive reviews of WI methods relate to (M. Awaida, 2012; Schomaker, 2007; Sreeraj & Idicula, 2011).

In existing work great focus has been laid on the development of highly performant descriptors; however, parameters and properties of IP detectors themselves have not been regarded. This leads to potentially losing out on discriminant features due to both, IP not detected and disregarding information encoded in IP themselves. This paper addresses these two points. We directly compute features of IP properties rather than extracting descriptors and building a codebook; proposing a method that can be applied out of the box, and omitting a training phase. Our method is intuitive to understand for a human expert on the one hand, and fast to compute and performant on the other. We emphasize the importance of adapting an IP detector to the requirements of WI showing that it leads to improved results and moreover can render segmentation and binarization superfluous.

The remainder of this paper is structured as follows. Section 2 analyzes the features, followed by a description of the feature computation in Section 3. Results achieved and a comparison to existing work are given in Section 4, followed by a conclusion and outlook into future work.

2. The Scale - Dominant Orientation Histogram (SDO)

IP are defined as locations in an image with a two-dimensional signal change (Tuytelaars & Mikolajczyk, 2008); i.e., they are located at image structures such as corners, junctions, circles, or dots. We use an IP's dominant orientation, which describes the prevailing direction of gradients in its neighborhood, together with its spatial extent (scale) to compute a 2D histogram. An illustration of IP and their scales is shown in Figure 1 (the image is cropped for illustrative purposes - we extract IP from entire pages). While other IP detectors can be employed, as a proof of concept we chose the widely-used Difference-of-Gaussian (DoG) detector (Lowe, 2004).
Orientation and scale being properties intuitively understood, our feature allows for translation of the results comprehensible to human experts in document examination. A conceivable visualization that can be used as *visual fingerprint* of a handwriting sample is shown in Figure 2; and the information encoded in the features is summarized in the following. The slant is encoded as peaks across scales in the histogram. However, note that the dominant orientation captures several properties of the script; thus, the peaks might be deviated from the overall slant (Figure 2 d). Furthermore, the distribution of stroke orientations along with their scales allows for deducing character shapes, e.g. roundness, especially of loops and holes; and uniformity of the handwriting, i.e., how parallel the strokes are, or consistent the handwriting is. While round characters produce a higher variance in stroke orientations

Figure 1 Cropped example of a text line with the detected IP and their dominant orientations denoted as circles, where the size indicates the scale, and the line originating from the center the dominant orientation. Several orientations indicate multiple IP with different dominant orientations at the same location.

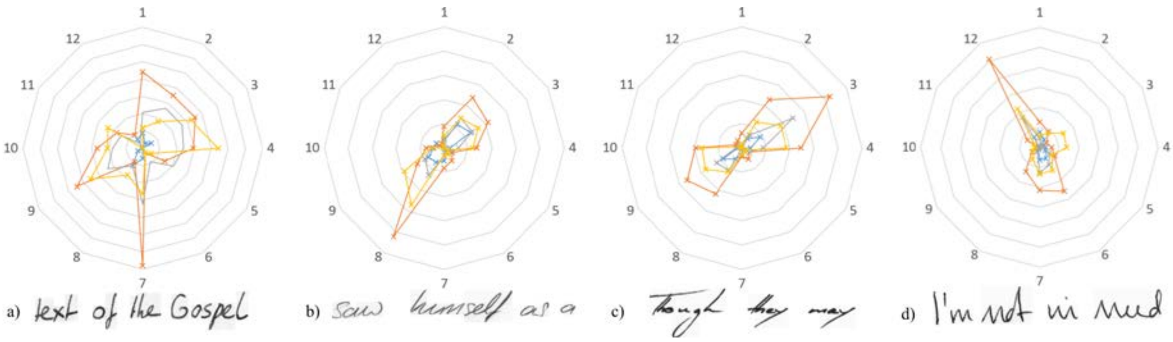

Figure 2 Samples (cropped for legibility) and their features as polar histogram. Angular coordinates denote quantized orientations, radial coordinates denote proportional frequency, and scales are denoted by color-coded markers. For legibility the 4 smallest scales only are shown and markers are connected. Round uniform writing causes high (a), long slanted strokes are reflected in two opposing distinct peaks (c), and angular writing one distinct peak (b, d).

(Figure 2 a), elongated narrow characters with angular shapes induce clear peaks in the orientation histogram (Figure 2 b, d). Furthermore, compact condensed writing has smaller structures compared to loose or uncondensed writing; thus producing a higher relative number of small IP. Continuous writing with long strokes generates a signature of small IPs different to one of intermittent writing produced by, e.g., little pressure put on the pen.

Scales of IPs are not discriminant enough by themselves but provide subsidiary cue when combined with other properties. Small-scale IP cover the width of a stroke, allowing for inference of information about, e.g., stroke widths including their variance. Large scales represent big loops of characters, or spaces between ascenders, descenders, characters, words, or lines.

A formal translation of the histogram into a verbal description as well as an interactive visualization an expert can use to explore aspects of the visual fingerprint described are beyond the scope of this paper.

(Wu et al., 2014) also propose a scale and orientation histogram, yet their method requires prior word segmentation. Since segmentation of text lines and words is an open research topic (Stamatopoulos et al., 2013) and potentially introduces errors in further processing, we omit it in our approach.

3. Feature extraction

As feature we compute the normalized probability density function of a 2D histogram size $\mathrm{X} \times \mathrm{Y}$, with X and Y being the total number of quantized scales and dominant orientations, respectively. Quantization inherently has an effect on the information captured (between-writer variability) and the invariance incorporated (within-writer variability). Using DoG, we decompose an image into a scale space[1] of $\mathrm{X} = \mathrm{M} \times \mathrm{N}$, with M octaves and N sub-levels. IP are then identified as local extrema of the scale space, i.e., a pixel is selected as IP if it is the minimum or maximum compared to its eight neighbors at the same scale and nine corresponding neighbors in the adjacent scales, and its magnitude exceeds the threshold th. The scale $\mathrm{s_i}$ of IP i is in the range $1 \leq \mathrm{s_i} \leq \mathrm{M} \times \mathrm{N} \times \sigma \times m_{\mathrm{i}}$, where σ is the standard deviation of the Gaussian kernel for creating the scale space, and m_{i} is the magnitude of the extrema. It is quantized using step size σ, which groups IP according to their location in the scale space. The dominant orientation is quantized from $[0°, 360°]$ with angle step α.

The best parameter combination for the DoG detector was determined on the *ICDAR2011 cropped dataset* (Louloudis et al., 2011) as $[M = 3, N = 6, \sigma = 1.3, th = 5, r = 0]$ with th being the detector sensitivity, and r the edge threshold, where 0 means that IP located at edges are not suppressed. With an angle step $\alpha = 30$ we create a feature vector of size 216 which we truncate to the first 108 elements (smallest scales) based on experiments. Note that changing the size of the scale space X and truncating are not the same operation since the actual scale of an IP additionally depends on the strength of extrema the IP is located at.

4. Evaluation

Performance Evaluation. We conducted our experimental study on full pages of the datasets listed in Table 1. The evaluation design follows the *ICDAR2011 competition* evaluation procedure, except that only the TOP-1 identification criterion is reported, i.e., the document ranked first has to be by the same writer as the query. We employ a naïve nearest neighbor approach in a leave-one-out manner for identification, and the χ^2 distance metric as dissimilarity measure between two documents. Significance is tested using a χ^2-Test ($p<0.05$).

Table 1 Overview of the datasets used

Dataset	# writers	# pages	# lines	Language
IAM[2] (Marti & Bunke, 2002)	657	2	3-14	English
ICDAR2013 (Louloudis et al., 2013)	250	4	4	English, Greek
ICDAR2011 full (Louloudis et al., 2011)	26	8	13-23	English, French, German, Greek
ICDAR2011 cropped	26	8	2	English, French, German, Greek

Table 2 Comparison of state-of-the-art methods (a, b) and our proposals (c-e) on the test datasets.

Method		IAM	ICDAR2013	ICDAR2011 full	ICDAR2011 cropped
(a)	(Wu et al., 2014)	98.5	94.8	99.5	95.2
(b)	(Jain & Doermann, 2014)	94.7	N/A	N/A	N/A
(c)	SDO-E T	81.9	81.4	98.6	87.5
(d)	Descriptor (Fiel & Sablatnig, 2012)	82.3	80.1	99.5	80.8
(e)	SDO-E T & Descriptor	86.9	87.3	98.6	88.0

Table 3 Evaluation of IP detector parameters on the *IAM dataset*. SDO is our proposal, with postfix "-E" for IP on edges permitted and "-NS" for IP on background suppressed. F is the full feature vector (size 216), T the vector truncated to 108.

Method		Settings	Result	
(1)	(Wu et al., 2014) Documents	Scale Orientation Histogram on full pages $[M = 6, N = 3]$	64.0	
(2)	(Wu et al., 2014) Words	Scale Orientation Histogram on segmented words $[M = 6, N = 3]$	78.4	
(3)	Our implementation of (1)[3]	$[M = 6, N = 3, \mathrm{r} = 0.1]$	65.5	
			F	**T**
(4)	SDO	$[M = 3, N = 6, r = 0.1]$	79.1	79.6
(5)	SDO-E	$[M = 3, N = 6, r = 0]$	81.3	**81.9**
(6)	SDO-E-NS	$[M = 3, N = 6, r = 0]$, no background IP	79.8	80.6
(7)	SDO-NS	$[M = 3, N = 6, r = 0.1]$, no background IP	75.3	76.3
(8)	SDO (Words)		65.2	

Results are summarized in Table 2. Our method (c) achieves competitive scores on the ICDAR2011 full dataset. The performance declines when having considerably less data: ICDAR2011 cropped only contains two text lines per page. Note that the performance difference between (a) and (c) on the ICDAR2011 sets is insignificant. In order to assess potential gains of feature combination, we additionally combined our method with a descriptor-codebook-based approach (d), which is our own implementation of (Fiel & Sablatnig, 2012)[4]. The codebook is computed on an independent dataset. On larger datasets such as *IAM* and *ICDAR2013*, feature combination (e) significantly boosts the performance, showing that our feature captures complementary information..

Parameter Evaluation. In the following we show that appropriate parameters for IP detectors are critical for the overall performance of an IP-based method. IP originate from object detection and recognition, where a homography (mapping between two projections of an object) is computed that requires stable and repeatable IP. However, the task of writer identification is of different nature (we need to describe the strokes present), and unreflected adopting of standard parameters suitable for one task is likely to being ill-suited for another.

For the evaluation of the DoG detector's parameters we used the IAM dataset for its size and variability in writing styles and amounts of data per page (3-14 lines). The TOP-1 identification results are shown in Table 3 . With the settings reported in (Wu et al., 2014) as baseline (3), we tested our feature with following alterations: improved selection of scale space parameters $[\mathrm{M}, \mathrm{N}]$ (4); permitting IP located on edges $[\mathrm{r}]$ (5), and excluding IP on the background (7), i.e., those IP corresponding to minima in the scale space, as means to avoid segmentation; and their combination (6). First and foremost, using the best combination and optimizing the feature vector (5 T) we can forego segmentation with significant increase in performance with respect to (2). Furthermore, comparing 1 and 1, we see that IP located on white space do encode valuable information. We want to stress that IP located on edges capture supplementary information as they describe prominent strokes otherwise not detected, see 1 and 1, and particularly 1 and 1. For comparison, we show the result of the basic SDO (4) on the segmented text line images of the IAM dataset (8) Figure 3 illustrates the effect of the parameter combinations evaluated.

[2] The dataset has been modified to contain two samples per writer according to the procedure described in (Wu et al., 2014).

[3] Lacking information about the original paper's parameters, we use the parameter combination determined in Section 3.

[4] Note that the performance reported on the *IAM dataset* (90.8) by the authors is not directly comparable to the literature since they used only a subset of documents for evaluation: writers with only one sample are not evaluated, and 2 to 58 reference samples are kept for identification, while for each writer we keep only one reference sample in our evaluation. It is inherent that fewer writers and more reference samples result in better performance. To assess whether our implementation (d) is comparable to the original, we evaluated it according to the strategy explained, achieving a slightly better identification rate of 92.4.

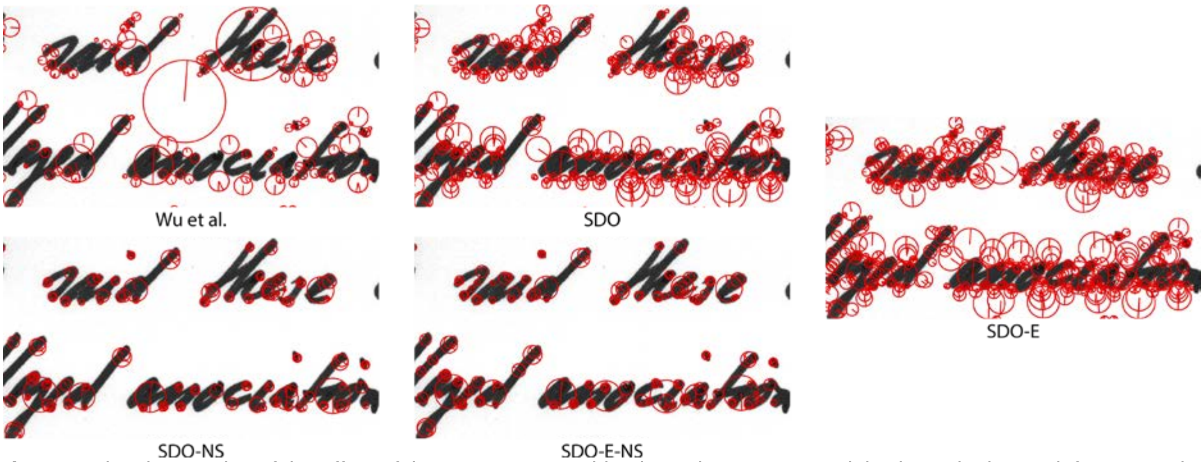

Figure 3 Visual examples of the effect of the parameter combinations chosen. Wu et al. is shown in the top-left corner. The best-performing parameter combination (SDO-E) is shown in the third column (refer to Table 3 for the exact parameters).

5. Conclusion

The writer identification method presented in this paper makes use of overlooked properties of IP as feature, rather than employing a codebook of descriptors as in existing methods. Our method is segmentation-free and does not require training, as we compute the feature directly as a 2D histogram of interest points' scales and dominant orientations. An additional contribution targets the need of careful adaption of methods originating from another field. We showed that adapting the detector to the task of writer identification, on the one hand, boosts performance since IP capturing additional information about a writer are detected, and facilitates a segmentation-free approach, on the other.

Opposing the inference of (Wu et al., 2014), who state that word segmentation is essential – for IP in inter-word and inter-line space being instable – we conclude that spaces encode valuable additional information and boost performance, rendering segmentation unnecessary. Our feature outperforms their proposition by a significant margin. Furthermore, we showed that using IP detected on the foreground only, are another alternative to word segmentation with the limitation of losing some performance with respect to a method including inter-space IP.

One limitation of the features proposed is the amount of data needed to create a reliable model of a writer; however, combined with a descriptor-based method performance can be boosted with respect to both features, especially for large datasets. We propose to incorporate our features into future IP-based methods for its simplicity, capability to capture complementary information, good performance, and minor expense to compute.

References

Fiel, S., & Sablatnig, R. (2012). Writer Retrieval and Writer Identification Using Local Features. In *Int. Wksp on Document Analysis Systems* (pp. 145–149).

Jain, R., & Doermann, D. (2014). Combining Local Features for Offline Writer Identification. In *Int. Conf. on Frontiers in Handwriting Recognition* (pp. 583–588).

Louloudis, G., Gatos, B., Stamatopoulos, N., & Papandreou, A. (2013). ICDAR 2013 Competition on Writer Identification. In *Int. Conf. on Document Analysis & Recognition* (pp. 1397–1401).

Louloudis, G., Stamatopoulos, N., & Gatos, B. (2011). ICDAR 2011 Writer Identification Contest. In *Int. Conf. on Document Analysis & Recognition* (pp. 1475–1479).

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. of Computer Vision*, *60*(2), 91–110.

M. Awaida, S. (2012). State of the Art in Off-line Writer Identification of Handwritten Text and Survey of Writer Identification of Arabic Text. *Educational Research and Reviews*, *7*(20).

Marti, U., & Bunke, H. (2002). The IAM-Database: An English Sentence Database for Offline Handwriting Recognition. *Int. J. on Document Analysis & Recognition*, *5*(1), 39–46.

Niels, R., & Vuurpijl, L. (2005). Using Dynamic Time Warping for Intuitive Handwriting Recognition. *Int. Graphonomics Soc.*, 217–221.

Pratikakis, I., Gatos, B., & Ntirogiannis, K. (2013). ICDAR 2013 Document Image Binarization Contest (DIBCO 2013). In *Int. Conf. on Document Analysis & Recognition* (pp. 1471–1476).

Schomaker, L. (2007). Advances in Writer Identification and Verification. In *Int. Conf. Document Analysis & Recognition* (pp. 1268–1273).

Sreeraj, M., & Idicula, S. M. (2011). A Survey on Writer Identification Schemes. *Int. Journal of Computer Applications*, *26*(2), 23–33.

Stamatopoulos, N., Gatos, B., Louloudis, G., Pal, U., & Alaei, A. (2013). ICDAR 2013 Handwriting Segmentation Contest. In *Int. Conf. on Document Analysis & Recognition* (pp. 1402–1406).

Tuytelaars, T., & Mikolajczyk, K. (2008). Local Invariant Feature Detectors: A Survey. *Foundation and Trends in Computer Graphics and Vision*, *3*(3), 177–280.

Wu, X., Tang, Y., & Bu, W. (2014). Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform. *IEEE Trans. on Information Forensics and Security*, *9*(3), 526–536.

Acknowledgement. This work has been supported by the Swiss National Science Foundation project 205120_150173.