Sparse Radial Sampling LBP for Writer Identification

Anguelos Nicolaou*, Andrew D. Bagdanov*, Marcus Liwicki†, and Dimosthenis Karatzas*

*Computer Vision Center, Edifici O, Universitad Autonoma de Barcelona,Bellaterra, Spain
†DIVA research group, Department of Informatics, University of Fribourg, Switzerland
Email: anguelos@cvc.uab.es, bagdanov@cvc.uab.es, marcus.liwicki@unifr.ch, dimos@cvc.uab.es

Abstract—In this paper we present the use of Sparse Radial Sampling Local Binary Patterns, a variant of Local Binary Patterns (LBP) for text-as-texture classification. By adapting and extending the standard LBP operator to the particularities of text we get a generic text-as-texture classification scheme and apply it to writer identification. In experiments on CVL and ICDAR 2013 datasets, the proposed feature-set demonstrates State-Of-the-Art (SOA) performance. Among the SOA, the proposed method is the only one that is based on dense extraction of a single local feature descriptor. This makes it fast and applicable at the earliest stages in a DIA pipeline without the need for segmentation, binarization, or extraction of multiple features.

I. INTRODUCTION

Writer identification is the problem of identifying the authorship of text samples based on an index of examples of text written by known authors [1]. It has a long tradition in forensics where it has been accepted by the court as evidence for more than a century. From a pattern recognition perspective, three variations of the problem of writer identity are defined: writer identification, writer verification, and writer retrieval. Writer identification is the most popular of these, and in most cases a method can be modified from solving one to solving another with little effort.

From a Document Image Analysis (DIA) perspective there are other applications such as scribe identification for historical documents. Handwriting has been considered as a behavioral biometric [2], [3], although experiments on disguised handwriting have proved to confuse both graphonomists, and automatic systems [4]. Despite the attention writer identification has received from DIA researchers in recent years, it remains a difficult problem due to variations in writing style and conditions, and the myriad problems arising from variations in image quality due to document degradation and other incidental factors.

LBP are dense local texture descriptors that can be used to describe the local structure of images [5]. They have been successfully applied to many of the major computer vision problems, as well also been applied to specific problems in Document Image Analysis, including optical font recognition and writer identification. LBP were originally designed for graylevel images, and despite their widespread application to bilevel document images, it remains unclear how LBP should be computed on such images in order to remain discriminative and robust to noise.

In this paper we introduce Sparse Radial Sampling LBP (SRS-LBP), a variant of LBP that is better suited for the task of writer identification and text-as-oriented-texture classification in bilevel images in general. Our main contribution is the introduction of sparse radial sampling of the circular patterns used for LBP construction. This allows sampling of patterns up to very large radii for each pixel at low computational cost, and we can also avoid vocabulary compression techniques such as rotation invariant or uniform patterns commonly applied to standard LBP representations. We show that using a single local descriptor, our SRS-LBP variant densely extracted and pooled over the entire image, results in a low-dimensional feature representation that yields SOA performance at a fraction of the cost of other techniques. Our representation is compact and extremely efficient to compute.

In the next section we describe work from the literature related to our application of LBP to writer identification. We describe the modified LBP we use and how we apply it to writer identification in Section III. In Section IV we report on experiments we performed comparing our approach with the SOA on two standard writer identification benchmarks, and in section V we conclude with a discussion of our contribution and ongoing work.

II. RELATED WORK

A. Local Binary Patterns

LBP characterize local image patches using binary codes that encode the relationship between a central pixel and its neighbors [5]. LBP feature extraction usually consists of computing LBP descriptors at each pixel of an image to create an image of integer valued codes, followed by pooling of these codes into a histogram [5]. LBP have been successfully applied to many of the major computer vision problems such as face recognition [6], and human detection [7].

In DIA, LBP have been used for text detection of text in television streams [8], for printed script detection [9], and in [10] LBP were compared to other features in a feature selection process for historical document layout analysis and were selected as best across all datasets. The authors used an earlier proponent of the method presented in this paper for Arabic font recognition [11].
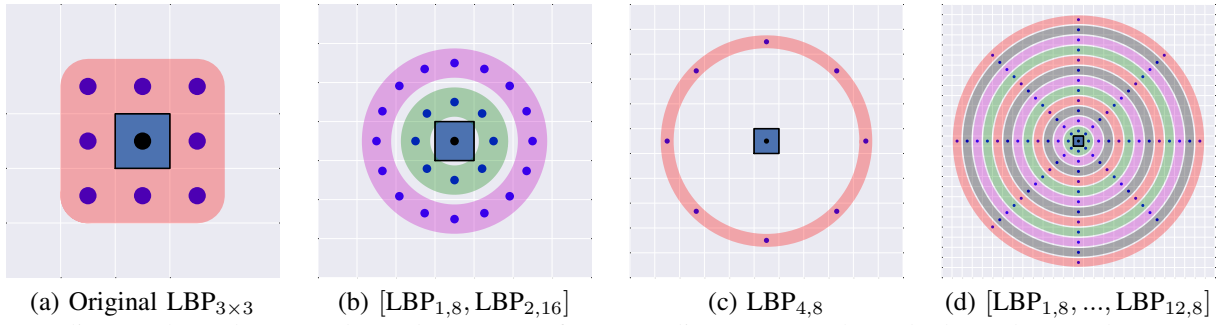
(a) Original LBP$_{3\times3}$ (b) [LBP$_{1,8}$, LBP$_{2,16}$]   (c) LBP$_{4,8}$ (d) [LBP$_{1,8}$, ..., LBP$_{12,8}$]

Fig. 1: Baseline (a),(b) and Proposed (c),(d) LBP transform sampling patterns. (b) and (d) mark several LBP used in in conjunction.

B. Writer Identification

Automatic writer identification has been researched for many decades. In 2007 Schomaker et al. propose allographic features for writer identification, and for an overview of previous work on writer identification, we refer to their excellent survey [1]. Chawki et al. applied run length features in 2010 for Arabic writer identification [12]. Jain et al. used clusters of pseudo-letters [13] and fusion of features [14]. LBP have also been used for writer identification: Du et al. extracted LBP features from the wavelet transform of Chinese hand-writing samples [15], and Bertolini et al. [3] used LBP in a comparative study with Local Phase Quantization and concluded that LPQ performed better than their LBP variant.

C. Our contribution with respect to the State-Of-The-Art

Our adaptation of the LBP consists of replacing the sign operator with a threshold statistically derived from each image, the use of sparse radial sampling at each radius, and the use of very large radii when computing LBP. Though we concatenate LBP histograms extracted at many radii, sparse radial sampling ensures that the final LBP features are compact.

We apply our SRS-LBP to writer identification using a standard LBP pipeline. SRS-LBP are computed at each location in an image, and these features are pooled over the entire page image. Our approach requires no character segmentation and is based on a single, compact feature that is extremely efficient to extract. In this sense it stands out with respect to SOA approaches based on complex character segmentation, clustering, and extraction of multiple feature descriptors [14].

III. WRITER IDENTIFICATION WITH LBP

In this section we describe our approach to LBP extraction based on sparse radial sampling. We follow the development and notation of [5].

A. The LBP transform

LBP feature extraction consists of two principal steps: the LBP transform, and the pooling of LBP into a histogram representation of an image. The LBP transform maps each pixel to an integer code representing the relationship between the center pixel and the pixels of its neighbourhood. It encapsulates the local geometry at each pixel by encoding binarized differences with pixels of its local neighbourhood:

$$\text{LBP}_{P,R,t} = \sum_{p=0}^{P-1} s_t(g_p - g_c) * 2^p, \tag{1}$$

where g_c is the central pixel being encoded, g_p are P symmetrically and uniformly sampled points on the periphery of a circular area of radius R around g_c, and s_t is a binarization function parametrized by t. The sampling of g_p is performed with bilinear interpolation. The use of local differences in (1) endows LBP with a degree of illumination invariance.

In our LBP definition, s is a simple threshold:

$$s_t(x) = \begin{cases} 1 & : x \geq t \\ 0 & : x < t \end{cases}, \tag{2}$$

where t, which in the standard definition is considered zero, is a parameter that determines when local differences are considered "big enough" for consideration.

$\text{LBP}_{P,R,t}$ can be seen as a transform from the graylevel domain to a domain of discrete labels encoded over a vocabulary of 2^P integers. In Fig. 1 sampling patterns of popular and proposed LBP can be seen where g_c is marked as a black dot and g_p are marked as blue dots.

B. Sparse sampling LBP on bilevel images

The original LBP was designed for graylevel images. Though text images are often fundamentally bilevel by nature, the bilinear interpolation used to extract neighbouring pixel values g_p renders pixels non-binary and standard LBP is (at least mathematically) applicable. However, images of a bilevel nature such as text even when they are acquired on the graylevel domain do not benefit much from illumination invariance. Another problem is that large $g_c - g_p$ differences are more rare than small ones and so treating both of them the same introduces noise.

Rather than use arbitrary or empirically derived threshold t to re-binarize differences in the computation of LBP in (1), we propose to apply Otsu's method to estimate optimal threshold $\hat{t}$ from the statistics of image differences themselves:

$$\hat{t} = \arg\min \omega(d_{1,t})\sigma^2(d_{1,t}) + \omega(d_{t,P})\sigma^2(d_{t,P}) \tag{3}$$

where $d_{1,t}$ is the set of $|g_c - g_p|$ less than threshold $\hat{t}$, $d_{t,P}$ is the set of $|g_c - g_p|$ greater than the threshold, ω is the probability

of $d_{...}$ and σ^2 its variance. The use of $\hat{t}$ yields a unified solution to both of these problems. The Otsu threshold of the differences effectively separates the significant differences from insignificant ones. Note that this formulation works for bilevel and graylevel source imagery.

Sparse radial sampling is integrated into our descriptor by holding constant the number of points sampled at each radius:

$$\text{SRS-LBP}_R = \text{LBP}_{8,R,\hat{t}}. \qquad (4)$$

Keeping $P = 8$ constant (i.e. sparse radial sampling) allows us to sample more radii while maintaining a compact code.

C. Processing pipeline

Our complete processing pipeline is comprised of the following steps:

1) **SRS-LBP transformation:** each image pixel is transformed to several SRS-LBP according to (1) and (4). This encodes the input image as several 8-bit images (one for each radius).
2) **SRS-LBP pooling:** a histogram of SRS-LBP codes is computed for each radius. We discard the zero pattern which corresponds to foreground- and background-only patterns, and then L1 normalize and concatenate all histograms. The result is a block-normalized descriptor of size $256 \times |R|$, where $|R|$ is the number of radii.
3) **PCA projection**: the block-normalized descriptor is projected onto the first N principal components computed through Principal Component Analysis (PCA).
4) **Normalization:** the Hellinger kernel is applied to the projected descriptor, followed L2 normalization. This combination has been shown to improve performance of a variety of image recognition techniques and specifically on writer identification [16].

Note that we use none of the standard vocabulary compression techniques such as rotation invariance or uniformity used in [5] to make possible the usage of larger radii. Avoiding these techniques is one of the main motivations for SRS-LBP because, in the case of textual images specifically, there is important information in the discarded patterns.

IV. Experimental results

In this section we report on a series of writer identification experiments we performed to evaluate the potential of SRS-LBP and to compare its performance with the SOA.

A. Datasets

We use a range of publicly available benchmark datasets for our experimental evaluation.
ICDAR 2013: The dataset from the ICDAR 2013 competition consists of 1,000 samples from 250 persons who each contributed two samples in English and two in Greek [17].
CVL: The CVL [18] dataset consists of 1,550 samples from 310 persons who contributed four samples in English and one in German. The samples were acquired in color with different pens.

ICHFR 2012: The ICHFR 2012 dataset consists of 400 samples from 100 subjects contributing two samples in English and two samples in Greek [19]. We use this dataset for baseline performance analysis.

B. Evaluation protocols

Our approach uses a nearest neighbor classifier, and evaluation is based on leave-one-out cross validation. For each sample represented in feature space, we rank all remaining samples by their distance to that sample. Two important performance measures are the top-n soft criterion, which means having any image of the same class as the query sample in the first n most ranked results, and the top-n hard criterion which means having *only* images of the same class as the query sample in first n samples [17], [19].

Comparison with the SOA is complicated by the wide variety of evaluation protocols for writer identification used by international benchmarks and contests. Just indicatively on the three most recent competitions in writer identification, methods had to be, a similarity measurement [19], a trainable classifier [4], and a feature extraction method accompanied by a metric [17].

In our evaluation we employ two evaluation protocols. One we refer to as **metric** and is a protocol totally compatible with measurements in [19], [17] (e.g. we only consider pairs of samples, never the entire dataset as a whole). The other we call **l1out** and is the average performance of a leave-one-out cross-validation in a trainable classifier sense which is compatible with [4]. In practice the difference between **l1out** and **metric** is that **l1out** allows access to all the samples in the evaluation dataset while **metric** restricts access to each sample alone.

For our approach, the difference between the two amounts to whether PCA analysis was done on the evaluation dataset (and thus learning from it) or an independent dataset. In all experiments other than comparison with the SOA, the evaluation protocol used is **l1out**. Since **metric** is a stricter protocol than **l1out**, all SOA performance numbers derived from a protocol that explicitly adheres to metric are marked with an asterisk (*).

C. Baseline performance analysis

Here we report on a number of baseline experiments we performed to quantify the performance of our approach and to estimate key parameters of our SRS-LBP. All experiments in this section were performed on the ICHFR 2012 dataset.
Comparison of SRS-LBP and standard LBP. We consider four standard LBP variants in these experiments: $\text{LBP}_{3\times3}$ [20], $\text{LBP}_{8,1}$ [5], $\text{LBP}_{16,2}$ [5], and a concatenation of LBP $[\text{LBP}_{8,1}, \text{LBP}_{16,2}]$ [5]. For SRS-LBP we consider two variations: a single radius $\text{SRS-LBP}_{8,4}$ and a multi-radius $[\text{SRS-LBP}_{8,1}, ..., \text{SRS-LBP}_{8,12}]$.

In Fig. 3 we report the error rates of top-1 accuracy leave-one-cross-validation for standard LBP and proposed SRS-LBP. In addition to the full LBP vocabulary, we also show results for each LBP variant with combinations of vocabulary compression commonly employed for standard LBP. From

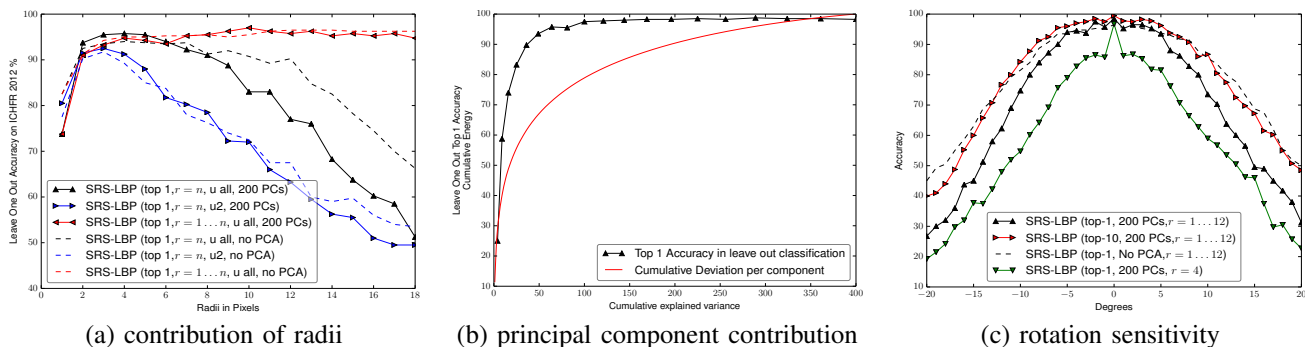| (a) contribution of radii | (b) principal component contribution | (c) rotation sensitivity |

Fig. 2: Parameter estimation and rotation sensitivity. (a) Individual and cumulative contribution of each radius to the top-1 accuracy. (b) The cumulative contribution principal components have in top-1 accuracy. (c) The tolerance of SRS-LBP to rotations of the image.
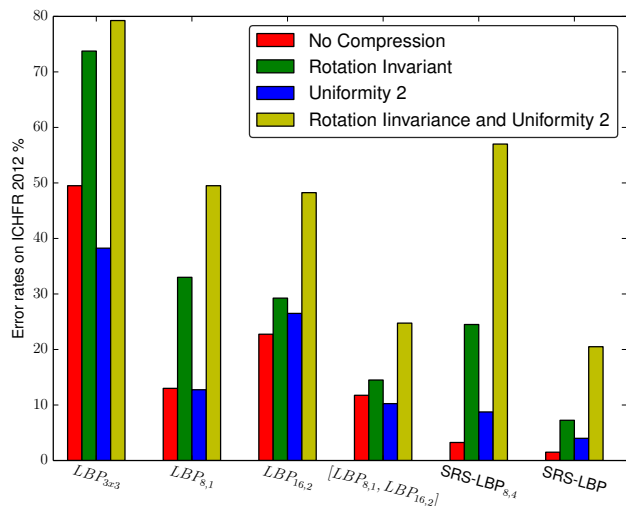

Fig. 3: Comparison of SRS-LBP with baseline LBP.

this figure, we see that SRS-LBP outperform standard LBP in all compression modalities. Note also that SRS-LBP are less sensitive to compression than standard LBP. Compression is usually applied to standard LBP in order to render vocabulary sizes tractable. Due to their sparse nature, SRS-LBP are already a compact and tractable without resorting to compression.

Radii Contribution. To better understand the contribution of each radius and principal component to the robustness of our descriptor, we decomposed performance as a function of each. In Fig. 2(a) we show the accuracy of each radius independently and cumulatively (i.e. by concatenation). From this, we see that even large radii continue to contribute to improved recognition performance. Note also how, as radii grow beyond 3, the performance of the uniform-compressed features (in blue), drops quite sharply compared to the non-compressed features (in black). This hints that uniformity compression doesn't scale to large radii. Since performance continues increasing until about twelve sparsely sampled radii, we use this configuration of SRS-LBP $= [\text{LBP}_{8,1}, \dots, \text{LBP}_{8,12}]$ for all subsequent experiments.

Contribution of principal components. PCA in the pipeline pipeline improves recognition accuracy in all cases.

TABLE I. Comparison with the SOA on ICDAR 2013

Method		Top 1	Top 2	Top 5	Top 10	Hard 2	Hard 3
Tebessa-C*	[21]	93.4	96.1	97.8	98.0	62.6	37.8
CS-UMD-a*	[13]	95.1	97.7	98.6	99.1	19.6	7.1
Super Vector	[16]	97.1	NA	NA	NA	42.8	23.8
SRS-LBP$_{8,4}$ **l1out**		97.2	98.2	98.9	99.2	52.9	29.2
SRS-LBP **metric***		96.9	98.5	99.0	99.5	54.5	32.9

TABLE II. State-of-the-art on ICDAR 2013 Greek

Method		Top 1	Top 2	Top 5	Top 10
Tebessa-C*	[21]	93.1	97.0	99.5	99.5
Delta-n Hinge	[2]	93.4	NA	NA	98.4
CS-UMD-a*	[13]	95.1	97.7	98.6	99.1
Multi Feature*[14]		**99.2**	**99.6**	**99.8**	**99.8**
SRS-LBP$_{8,4}$ **l1out**		96.6	98.0	99.6	99.8
SRS-LBP **metric***		96.6	97.8	98.8	99.4
SRS-LBP **l1out**		98.4	99.2	99.4	**99.8**

In Fig. 2(b) we plot of top-1 writer identification accuracy as a function of increasing PCs. Although performance quickly begins to saturate, it improves steadily until around 200 principal components. For all subsequent experiments we use 200 principal components, resulting in a very compact image descriptor of only 200 dimensions.

Rotation sensitivity. Tolerance to small rotations can be important for text documents, and in the case of handwritten text the exact orientation of text is probably unknown. To measure sensitivity to rotation, the ICHFR 2012 data-set was rotated from angles -20° to 20°. For each rotation, all samples in the non-rotated dataset were used as queries against the rotated dataset. In Fig. 2(c) we show the sensitivity to rotation of the proposed SRS-LBP pipeline. An interesting observation in Fig. 2(b) is that while the single radius SRS-LBP performs nearly equivalently to the multi-radius one, it is more sensitive to rotations.

D. Comparison with the State-Of-The-Art

In recent years the topic of writer identification has seen a lot of activity. Contests, datasets, as well as several top performing methods have been published. In this section we compare SRS-LBP with the SOA in writer identification.

Performance on ICDAR 2013. In Table I we compare the performance of SRS-LBP with the SOA on the ICDAR 2013 contest dataset. In Table II and Table III we compare

TABLE IV. State-of-the-art on CVL using only 4+1 samples per writer

Method	Soft Top 1	Soft Top 2	Soft Top 5	Soft Top 10	Hard Top 2	Hard Top 3	Hard Top 4
Tebessa-C* [21]	97.6	97.9	98.3	98.5	96.1	94.2	90.0
Multi-feature* [14]	**99.4**	**99.5**	**99.6**	**99.7**	98.3	94.8	82.9
Super Vector [16]	99.2	NA	NA	NA	98.1	95.8	88.7
SRS-LBP$_{8,4}$ l1out	99.0	99.2	99.4	99.5	97.7	95.2	86.0
SRS-LBP metric*	98.6	98.8	98.9	99.1	97.8	94.6	85.3
SRS-LBP l1out	**99.4**	99.4	99.5	99.6	**98.6**	**97.0**	**90.1**

TABLE III. State-of-the-art on ICDAR 2013 English

Method	Top 1	Top 2	Top 5	Top 10
Tebessa-C* [21]	91.5	95.5	97.5	98.0
Delta-n Hinge [2]	93.4	NA	NA	97.8
CS-UMD-a* [13]	95.2	98.2	98.8	99.2
Multi Feature* [14]	**97.4**	**97.8**	**98.6**	98.8
SRS-LBP$_{8,4}$ l1out	95.2	96.4	98.0	98.4
SRS-LBP metric*	95.6	96.8	98.4	**99.0**

our performance with the SOA on the Greek and English portions of the ICDAR 2013 dataset. As of this writing, the Multi Feature method represents the SOA in writer identification [14]. This technique is based on character segmentation and clustering (which is one reason they do not report results on the mixed-language dataset) and multiple features extracted from characters. It is interesting that our approach, which is based on dense extraction of a single feature, performs comparably to this more complicated technique.

Performance on CVL. Finally, in Table IV we compare the performance on SRS-LBP with the SOA on the CVL dataset. On this dataset we have the most complete comparison with SOA approaches for both hard and soft criteria. Our approach performs equivalently to the Multi Feature technique of [14] for top-1 evaluation criterion, and we outperform all others approaches for the hard recognition evaluation criterion across all ranks which is associated with writer retrieval.

V. CONCLUSION

In this paper we introduced a writer identification approach based on sparse radial sampling Local Binary Patters. Our approach achieves SOA performance on ICDAR 2013 and CVL datasets. Our proposed method has several advantages over other SOA techniques. Of the top performers, ours is the only one that is based on dense extraction of a single local feature descriptor. This makes it applicable at the earliest stages in a DIA pipeline without the need for segmentation, binarization, or extraction of multiple features. Its simplicity makes it a good candidate for fusion with other methods.

The SRS-LBP we propose is a generic LBP variant that should be applicable to many document image analysis tasks such as font recognition, handwritten and printed script detection, ant document page classification. Future and ongoing work will concentrate on experiments demonstrating the range of application the proposed method has and the improvements it can bring in end-to-end document recognition scenarios. While PCA was the sole learning component of the proposed method, it is unsupervised and feature space transformations taking full advantage of labelled training data should yield improved results.

REFERENCES

[1] M. Bulacu and L. Schomaker, "Text-independent writer identification and verification using textural and allographic features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 701–717, 2007.

[2] S. He and L. Schomaker, "Delta-n hinge: rotation-invariant features for writer identification," in *International Conference on Pattern Recognition*, 2014.

[3] D. Bertolini, L. S. Oliveira, E. Justino, and R. Sabourin, "Texture-based descriptors for writer identification and verification," *Expert Systems with Applications*, vol. 40, no. 6, pp. 2069–2080, 2013.

[4] M. I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found, "Icdar 2013 competitions on signature verification and writer identification for on-and offline skilled forgeries (sigwicomp 2013)," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 1477–1483.

[5] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[6] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2037–2041, 2006.

[7] Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, "Discriminative local binary patterns for human detection in personal album," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.

[8] M. Anthimopoulos, B. Gatos, and I. Pratikakis, "A hybrid system for text detection in video frames," in *Document Analysis Systems, 2008. DAS'08. The Eighth IAPR International Workshop on*. IEEE, 2008, pp. 286–292.

[9] M. A. Ferrer, A. Morales, and U. Pal, "Lbp based line-wise script identification," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 369–373.

[10] H. Wei, K. Chen, A. Nicolaou, M. Liwicki, and R. Ingold, "Investigation of feature selection for historical document layout analysis," in *Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on*. IEEE, 2014, pp. 1–6.

[11] A. Nicolaou, F. Slimane, V. Maergner, and M. Liwicki, "Local binary patterns for arabic optical font recognition," in *Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on*. IEEE, 2014, pp. 76–80.

[12] D. Chawki and S.-M. Labiba, "A texture based approach for arabic writer identification and verification," in *Machine and Web Intelligence (ICMWI), 2010 International Conference on*. IEEE, 2010, pp. 115–120.

[13] R. Jain and D. Doermann, "Writer identification using an alphabet of contour gradient descriptors," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 550–554.

[14] ——, "Combining local features for offline writer identification," in *2014 14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. IEEE, 2014.

[15] L. Du, X. You, H. Xu, Z. Gao, and Y. Tang, "Wavelet domain local binary pattern features for writer identification," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 3691–3694.

[16] V. Christlein, D. Bernecker, F. Hoenig, and E. Angelopoulou, "Writer identification and verification using gmm supervectors," in *IEEE Winter Conference on Applications of Computer Vision*. IEEE, 2014.

[17] G. Louloudis, B. Gatos, N. Stamatopoulos, and A. Papandreou, "Icdar 2013 competition on writer identification," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 1397–1401.

[18] F. Kleber, S. Fiel, M. Diem, and R. Sablatnig, "Cvl-database: An off-line database for writer retrieval, writer identification and word spotting," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 560–564.

[19] G. Louloudis, B. Gatos, and N. Stamatopoulos, "Icfhr 2012 competition on writer identification challenge 1: Latin/greek documents." in *ICFHR*.

Citeseer, 2012, pp. 829–834.

[20] M. Topi, O. Timo, P. Matti, and S. Maricor, "Robust texture classification by subsets of local binary patterns," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 3. IEEE, 2000, pp. 935–938.

[21] C. Djeddi, L. Souici-Meslati, and A. Ennaji, "Writer recognition on arabic handwritten documents," in *Image and Signal Processing*. Springer, 2012, pp. 493–501.