Document Writer Analysis with Rejection for Historical Arabic Manuscripts

Daniel Fecker*, Abedelkadir Asi†, Werner Pantke*, Volker Märgner*, Jihad El-Sana† and Tim Fingscheidt*

*Institute for Communications Technology
Technische Universität Braunschweig
Braunschweig, Germany
Email: {*fecker, pantke, maergner, fingscheidt*}@*ifn.ing.tu-bs.de*

†Ben-Gurion University of the Negev
Be'er-Sheva, Israel
Email: {*abedas, el-sana*}@*cs.bgu.ac.il*

***Abstract*—Determining the individuality of handwriting in ancient manuscripts is an important aspect of the manuscript analysis process. Automatic identification of writers in historical manuscripts can support historians to gain insights into manuscripts with missing metadata such as writer name, period, and origin. In this paper writer classification and retrieval approaches for multi-page documents in the context of historical manuscripts are presented. The main contribution is a learning-based rejection strategy which utilizes writer retrieval and support vector machines for rejecting a decision if no corresponding writer can be found for a query manuscript. Experiments using different feature extraction methods demonstrate the abilities of our proposed methods. A dedicated data set based on a publicly available database of historical Arabic manuscripts was used and the experiments show promising results.**

I. INTRODUCTION

Examining the individuality of handwriting is still a challenging task for law agencies. Addressing this problem in ancient manuscripts poses additional challenges due to the nature of these documents. Given a data set of known writers, the writer identification task aims to assign one of those writers to a query document. Writer retrieval tries to obtain all the documents, out of a set of documents, that are written by the same writer of the query document.

Writer-related methods can roughly be divided into two major categories: grapheme-based methods [1]–[3] and texture-based methods [4]–[6]. While grapheme-based methods represent writers by extracting features from characters, the texture-based methods use global features of the writing style. However, most of the suggested methods combine features from both types [4], [5]. A detailed overview about this topic is given in [7].

A particular problem is to decide if an identification of a document's writer was truly successful. Assigning the most similar writer from a training data set to a query document does not always seem to be the best solution, especially if no reference document of the query document's writer is included in the training set. In this case it is better to reject the decision, which means that no appropriate writer candidate was found. To our knowledge, this task is insufficiently covered in the literature. The Arabic writer identification competition at ICFHR 2012 included this task but most of the teams disregarded it [8]. In other research fields rejection strategies are explicitly considered, e. g., handwriting recognition [9], [10].

Fig. 1. Example pages of historical Arabic manuscripts [11].

We further present a data set for the writer-related tasks in Arabic historical documents. Manuscripts were collected from the Islamic heritage project (IHP) [11]. To the best of our knowledge, this is the first data set which is designated to writer identification and retrieval tasks in Arabic historical documents. Fig. 1 shows example pages from this data set.

The paper is organized as follows. In the next section the data set used in our experiments is presented, followed by the description of the employed methodology for feature extraction. The classification schemes, writer retrieval, and the rejection strategy are described in Section IV. In Section V we present our data set, the experimental setup, and the results of our experiments. At the end, conclusions are given.

(a) (b)

Fig. 2. Main-text region detection. (a) Original image and (b) binary image with masked main text region.

II. Data Set

We evaluate our experiments on a publicly available[1] data set of historical Arabic manuscripts. The manuscripts and their meta-data (ground truth) are part of the Islamic heritage project (IHP) [11]. The data set consists of $I = 60$ different manuscripts with common challenges of historical documents. For example, these manuscripts contain different layouts with decorations, notes and whole pages from other (anonymous) writers, and the usual degradation due to the aging of the ancient manuscripts. Fig. 1 and Fig. 2 are showing example pages of these manuscripts.

The data set has manuscripts from $M = 29$ different writers. 11 writers have multiple (2-10) manuscripts in the data set, in total 42 manuscripts consisting of 2313 pages are in the subset $\mathcal{S}_{\text{multi}}$. Out of the remaining 18 writers, 12 of them are known writers who have only one manuscript per writer, $\mathcal{S}_{\text{single}}$, with 2108 pages in total. The other 6 manuscripts have unknown writers and we do not know if there are multiple manuscripts of one writer among them. They are denoted as $\mathcal{S}_{\text{unknown}}$ and contain in total 174 pages. The number of pages in the whole data set is 4595.

III. Feature Extraction

In this section we present different feature extraction methods which we used for our experiments [12]. The similarity measure between two document images relies on the features' ability to distinguish between different writers. Due to the inherent noise in ancient manuscripts (see Fig. 2a) the feature extraction is applied to preprocessed document images.

A. Preprocessing

We apply a pre-processing step to detect the main-text region. The main text region imposes a unique texture and orientation. Following this observation, we employ the Gabor filter as it had been found to be particularly appropriate to distinguish between texture representations [13]. Hysteresis thresholding is used to generate a binary mask from the response of the filter. This mask is used to approximately determining the main text region, as shown in Fig. 2b.

[1]http://www.cs.bgu.ac.il/~abedas

B. Modified Contour-Based Features (CON)

Bulacu et al. suggested a contour-based feature to capture the slant of a writing style [5]. They capture the contour direction distribution by considering local contour fragments which are defined as two contour pixels with distance δ. The contour fragment direction with respect to the x-axis is defined as

$$\phi = \arctan\left(\frac{u_{k+\delta} - u_k}{v_{k+\delta} - v_k}\right) \tag{1}$$

where (u_k, v_k) denotes a pixel on the contour. Following the contour, a histogram with n bins, spanning the interval of angles $\phi \in [0°, 180°]$, can be collected. This histogram is normalized to a probability distribution $P(\phi)$, and the prevalent direction in $P(\phi)$ corresponds to the slant of the writing.

While a fixed δ was used in [5], we modify this feature by using variable values for δ allowing a more accurate description of the contour angles. In a first step, δ_0 is used as initialization and a reference angle ϕ_{δ_0} is computed. Afterwards, the distance is increased incrementally $\delta_{i+1} = \delta_i + 1$ and each time the angle $\phi_{\delta_{i+1}}$ is recomputed and compared to the reference angle. If

$$\phi_{\delta_{i+1}} \notin [\phi_{\delta_0} - \epsilon, \phi_{\delta_0} + \epsilon] \tag{2}$$

with ϵ being a predefined constant, we use ϕ_{δ_i} for building the histogram. The number of bins of the histogram corresponds to the amount of elements in the final feature vector. Commonly, twelve bins are used.

C. Oriented Basic Image Features (OBI)

The oriented basic image (OBI) features are based upon local symmetry and orientation and they were used, e.g., for character recognition in natural images [14]. Recently, these features were used in the winning approach of the Arabic writer identification competition at ICFHR 2012 [8]. To extract these features one needs to create a filter response space by applying Gaussian derivative filters with different orders and directions. From the filter responses seven features are constructed which approximate the local symmetry. Each image pixel is assigned to one of the seven features depending on the highest response. While orientations can be assigned to four of the features, the other three are rotationally invariant. Based on the symmetry features, orientations and scales, a histogram is created which, after a normalization, results in the final feature vector with 3969 elements.

D. Keypoint-based Features with Scale-Invariant Feature Transform (K-SIFT)

Scale-invariant feature transform (SIFT) descriptors are local features which are based on finding keypoints in an image [15]. In handwriting the keypoints lie on crossings, loops and peaks of the characters and consequently describe the slant and the curvature of the writer [3].

A keypoint descriptor vector $\mathbf{v}$ of $E = 128$ elements is generated by computing eight bins orientation histograms from a 4×4 region around the keypoint in different scales. Let O be the amount of descriptor vectors which are extracted from a document image. For each descriptor vector element $e \in \{1, ..., E\}$ a transformed vector $\tilde{\mathbf{v}}_e$ is computed which contains all e-th elements of all O descriptor vectors

$$\tilde{\mathbf{v}}_e = (\mathbf{v}_1(e), ..., \mathbf{v}_O(e)) \,. \tag{3}$$

We calculate the cosine distance between all the vectors $\tilde{\mathbf{v}}_e$, $e = 1, ..., E$, and use the distances to create the final feature vector with $\sum_{e=1}^{E}(E - e) = 8128$ elements. These distances quantify the relations among all the local orientations and magnitudes of the keypoints in the document. This relation describes a global characteristic of the examined writing style.

IV. Classification

This section provides an overview of all the classification related tasks, starting with writer identification and retrieval schemes in multi-page documents and followed by the description of the proposed rejection strategy.

A. Basic Classification Techniques

Classification of handwriting samples to different writers can be applied to multi-page manuscripts on various levels. We apply classification on two levels, i.e., manuscript (multiple pages) level and page level. Once the manuscript level classification is addressed, we apply the *averaging* classification scheme. The *voting* classification scheme is used when single pages are considered. In this section we briefly summarize the above mentioned classification schemes [12]. Independent of the classification level, we employ the nearest-neighbor classifier which can be used with an arbitrary distance metric $D(\mathbf{a}, \mathbf{b})$, measuring the distance between the vectors $\mathbf{a}$ and $\mathbf{b}$.

In general the task is to identify the writer of an unknown query manuscript Q by comparing it to a training data set $\mathcal{S}$ with I manuscripts of M already known writers. This is done by finding the manuscript which handwriting reveals highest similarity to the query manuscript. The data set $\mathcal{S}$ includes the feature vectors extracted for all pages of the data set

$$\mathcal{S} = \left\{ \mathbf{x}_1^{(1)}, ..., \mathbf{x}_{L_1}^{(1)}, ..., \mathbf{x}_1^{(I)}, ..., \mathbf{x}_{L_I}^{(I)} \right\} \tag{4}$$

with $L_i, i = 1, ..., I$ denoting the number of pages of the manuscript i.

Averaging. When addressing the classification problem on manuscript level, each manuscript i is represented by a single feature vector. This feature vector is the average over all $\mathbf{x}_j^{(i)}$ with $j = 1, ..., L_i$. Given a query manuscript represented by its feature vector, we look for the closest manuscript i^* in the data set in terms of a pre-defined distance metric D. The writer of manuscript i^* is assigned to the query manuscript.

Voting. For the voting scheme the distances are computed on page level. We maintain a histogram with the number of bins equaling the number of writers. Given a query manuscript, we consider every page in this manuscript by looking for the nearest neighbor in $\mathcal{S}$. Once it is found, we retrieve its corresponding writer and increment the histogram bin that represents this writer by one. At the end, the writer with the maximal number of votes is assigned to the query manuscript. Due to the fact that this approach works on page level, also a k-nearest neighbor approach can be employed.

Weighted Voting. In this scheme we refine the voting approach by making the votes proportional to the distances. In other words, the page with the smallest distance has the highest weight, while all the other pages have lower weights.

B. Retrieval

Writer identification can also be seen as a writer retrieval task. Here, the goal is to *retrieve* document pages of a given writer from $\mathcal{S}$. The writer of whom pages should be retrieved can be given in form of an example page or an entire query manuscript Q. The result is a ranked list of retrieved pages, sorted by the internal confidence of the retrieval system in the correctness of the respective page. In our case, the list is sorted by distances in ascending order.

For the evaluation of writer retrieval results, a ground truth is needed that defines *relevant* pages that are supposed to be retrieved, i. e., pages that belong to the queried writer. The performance is usually illustrated by a recall/precision curve and measured by its area under the curve, the average precision (AP) [16]. Recall denotes how many of the relevant pages in the dataset are retrieved, normalized by the total number of relevant pages in the dataset, whereas precision constitutes how many retrieved pages are actually relevant, normalized by the total number of retrieved pages. The curve is created as follows: Going through the ranked result list, each time a relevant page is found, the corresponding precision value up to this point in the result list is computed and drawn into the curve.

C. Rejection Strategy

The problem is to decide if for a query manuscript Q with an unknown writer, corresponding manuscripts of the same writer truly exist in $\mathcal{S}$. It is not enough to identify a writer by finding a manuscript which reveals the highest similarity to Q because even if it has more similarity than the other manuscripts in $\mathcal{S}$, it does not automatically imply that this identification is correct. In order to have successful identification of writers, we have to have a possibility to reject an identification decision if no corresponding writer exists in $\mathcal{S}$.

Concerning rejection, in the literature often a reject decision is determined in a consecutive stage after a classifier, e.g., utilizing the distance between the two first most probable classes [10]. A decision is rejected if this distance is below a certain threshold which can be set heuristically or be automatically computed with a learning-based approach [9]. However, we observed that this approach has no discriminative power for our problem due to the fact that the distances between rejection and accept cases are not distinctive.

Another straightforward approach could be to utilize one-class classification techniques such as support vector data description [17] trying to calculate a model for one class in the feature space. In our case we could either train a model for $\mathcal{S}$ and check whether Q is part of the model, or vice versa. Experiments showed that both ways do not work sufficiently. A reason can be that we have high-dimensional feature vectors and due to the curse of dimensionality we do not have enough data to generate an adequate model. Moreover, it seems that the feature vectors from different writer classes have only small dissimilarities and are therefore highly overlapping in the feature space.

We use a conceptional approach which uses writer retrieval and supervised learning to make an assumption about the reject decision. In a first step, for each page of Q the p best ranked pages and their corresponding writers in $\mathcal{S}$ are retrieved. From the retrieval results we infer features which are used in a subsequent two-class classification step which either assigns a page to an *accept* class or to a *reject* class. For the supervised training of the classifier we use $\mathcal{S}$ defined by

$$\mathcal{S} = \mathcal{S}_{\text{multi}} \cup \mathcal{S}_{\text{single}} \tag{5}$$

$$\mathcal{S}_{\text{multi}} \cap \mathcal{S}_{\text{single}} = \emptyset \tag{6}$$

with $\mathcal{S}_{\text{multi}}$ being a subset of S with multiple manuscripts per writer and $\mathcal{S}_{\text{single}}$ including only single manuscripts per writer.

We utilize leave-one-out cross validation (LOOCV) on $\mathcal{S}$ to obtain the retrieval results from which we infer features for training a classifier. As features we use the p smallest distances and the *stability* of the writer retrieval. The stability feature is a scalar value defined by the maximum number a certain writer is listed among the first p writers, normalized by p. If the first p retrieved writers are all the same writer this feature has the maximum value one and if the assigned writers all vary, it has the lowest value $\frac{1}{p}$. In total, the inferred feature vectors contain $p+1$ features which are normalized to have zero mean and a unit variance.

As classifier we employ support vector machines (SVMs) [18]. SVMs are a powerful classification technique which calculate a classification border by maximizing the margin between the data instances of the different classes. The resulting model is described by the so-called support vectors which are the closest data instances to the decision border.

For each page of Q the classifier decides if its writer identification result is accepted or rejected. The ratio between the accepted and total pages of a manuscript can be used as a confidence measure. For the final decision which specifies if a manuscript was written by a writer in $\mathcal{S}$ a threshold θ can be applied on this confidence measure. By varying θ, receiver operator characteristics (ROC) curves, which plot the true acceptance rate against the false acceptance rate, can be obtained for evaluation [16].

V. Experiments

In this section the experimental setup and the results are described. The results are based on the data set described in the following section by employing the presented feature extraction methods and classification schemes.

A. Experimental Setup

We employ leave-one-out cross validation in our experiments due to our limited data set size. The general strategy is that one of the manuscripts of the data set is regarded as a query manuscript Q which writer needs to be identified. The feature extraction and writer classification is carried out page-wise, excluding all the pages of the manuscript Q from the training set to obtain a realistic setting.

We conduct two major experiments. The first one focuses on the general ability of the presented features, basic classification and retrieval schemes from Section IV-A and IV-B to correctly identify a manuscript with an unknown writer from a set of *a priori* known writers. In this case we are only using $\mathcal{S}_{\text{multi}}$, thus we do not need a reject decision. To utilize the broader variability of the full data set, we also use $\mathcal{S}_{\text{single}}$ and $\mathcal{S}_{\text{unknown}}$ in the training data set.

The second experiment focuses on testing the presented rejection strategy. For this task the whole data set $\mathcal{S}$ is used for testing. Because the ground truth of $\mathcal{S}_{\text{unknown}}$ is not known, they are treated in a dedicated experiment.

For the nearest-neighbor classification schemes we measure the distance with the $\mathcal{X}^2$-distance which is commonly used in this field and defined as

$$\mathcal{X}^2(\mathbf{a}, \mathbf{b}) = \sum_{\Delta=1}^{d} \frac{(a_\Delta - b_\Delta)^2}{a_\Delta + b_\Delta} \tag{7}$$

between two feature vectors $\mathbf{a}$ and $\mathbf{b}$ of length d.

We use the presented features from Section III, the modified contour-based approach (dubbed CON), the oriented basic image (OBI) features, and our keypoint-based approach (K-SIFT).

For the rejection experiments we demonstrate the results with the $p = 5$ first retrieval results leading to the best results for inferring the features for the reject decision. Here, also a combination of the retrieval results using different features is utilized. We employ the LIBSVM library [19] as SVM implementation and utilize radial-basis kernels for learning. Optimal SVM parameters are obtained by cross-validation.

B. Results

Basic classification results. The results for the first experiment are listed in Table I. In addition to the accuracy for the correct assignment of single pages to a specific writer, the accuracy for whole manuscripts based on the three classification schemes averaging, voting and weighted voting (denoted as w-voting) are depicted.

The K-SIFT features achieve the best results on page level and with the weighted voting approach a perfect classification of the writers is obtained. In contrast the averaging classification yielded the worst results, which, we think, is due to the loss of information by averaging these large feature vectors. In this category the OBI-based feature extraction achieves the best results.

Retrieval results. Evaluating the retrieval task, we use each single page of the dataset as query. Following the LOOCV

TABLE I. WRITER CLASSIFICATION AND RETRIEVAL RESULTS

Features	Accuracy of Pages	Accuracy of Manuscripts Averaging	Voting	W-Voting	mAP of Page Queries
CON	0.686	0.810	0.881	0.929	0.422
OBI	0.876	**0.929**	0.929	0.929	0.627
K-SIFT	**0.925**	0.595	**0.976**	**1.000**	**0.727**

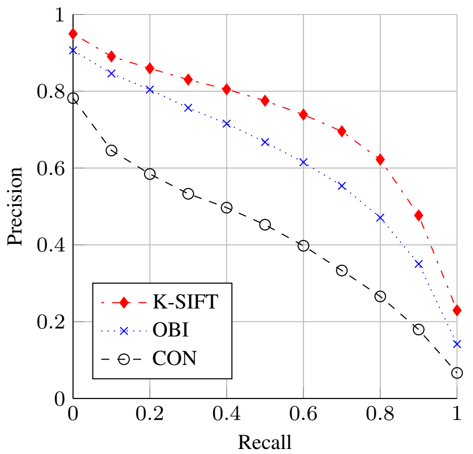

Fig. 3. Recall/precision curve using different feature sets

scheme, the whole manuscript that the respective query page is taken from is excluded. In Fig. 3, the standard 11-point interpolated recall/precision curve [16] averaged over all queries is shown. The K-SIFT feature set outperforms both the OBI and the CON feature sets. This can also be measured in terms of mean average precision (mAP) in Table I, which is the AP averaged over all queries. Due to the fact that the ranked result list contains every page of the dataset excluding the query manuscript, the recall reaches one in each query. However, the precision may still be underestimated because the ground truth lists only one writer per manuscript. It is likely that this may not be the case for each manuscript in the dataset [20].

Fig. 4 highlights the smallest distances obtained between each page of $\mathcal{Q}$ and $\mathcal{S}$ using LOOCV. The results for the two best features K-SIFT and OBI are depicted using $\mathcal{S}_{\text{multi}}$ and $\mathcal{S}_{\text{single}}$. The extreme and the zero values in these graphs correspond to empty pages or to pages with no text, i.e., figures or drawings. As can be seen in Fig. 4, the distances of sets $\mathcal{S}_{\text{multi}}$ and $\mathcal{S}_{\text{single}}$ are not very much different. However, separately averaging the distances over $\mathcal{S}_{\text{single}}$ and $\mathcal{S}_{\text{multi}}$ reveals a small but clear difference between the two subsets. The larger distances for $\mathcal{S}_{\text{single}}$ around page 3500 in Fig. 4(a) and around page 3000 in Fig. 4(b) furthermore demonstrate that a combination of both features might have benefits for the reject classification.

Rejection results. Table II shows the results for the rejection experiments in terms of accuracy on page level both for $\mathcal{S}_{\text{multi}}$ and $\mathcal{S}_{\text{single}}$, presenting the amount of correct accept and reject decisions. The best results are achieved using a combination of the OBI and K-SIFT retrieval results with a correct acceptance of nearly 80% of the pages from $\mathcal{S}_{\text{multi}}$ and a rejection of 85% of the pages from $\mathcal{S}_{\text{single}}$. The assumption that a combination has benefits proves to be true. A further combination with the CON features did not yield better results.

TABLE II. ACCEPTANCE AND REJECTION RESULTS

Features	Acc. of $\mathcal{S}_{\text{multi}}$	Acc. of $\mathcal{S}_{\text{single}}$
CON	0.797	0.521
OBI	0.776	0.805
K-SIFT	0.792	0.760
OBI & K-SIFT	**0.802**	**0.841**
CON & OBI & K-SIFT	0.801	0.821

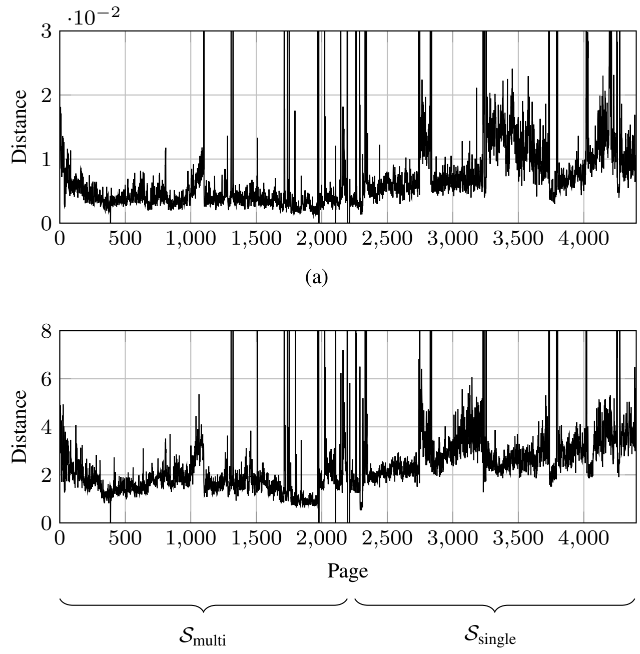

Fig. 4. Smallest distances between the pages of all $\mathcal{Q}$ and pages in $\mathcal{S}$ obtained with LOOCV. (a) OBI and (b) K-SIFT features.

The ratios between the accepted and the total number of pages in the manuscripts of $\mathcal{S}_{\text{multi}}$, $\mathcal{S}_{\text{single}}$, and $\mathcal{S}_{\text{unknown}}$ are shown in Fig. 5. One major observation is that the confidence for the first two manuscripts of $\mathcal{S}_{\text{multi}}$ equals zero. The metadata claims that both belong to the same writer but an examination of the manuscript's pages reveals different writing styles.

Another apparent observation is that in the results of $\mathcal{S}_{\text{unknown}}$ (Fig. 5(c)) for manuscript 2 and 3 the confidence for an acceptance is pretty high. Through the LOOCV procedure they are classified both to belong to the same writer. In the metadata no specific writer was named but both manuscripts have the same author[2]. Furthermore, both were written in the years 1745/46 and originate from the same country, Syria, which is an evidence that they are written by the same writer.

The overall performance for a decision if a manuscript writer identification decision is accepted or rejected is presented in Fig. 6. The ROC curves are shown for the different feature extraction approaches.The best performance in terms of area under curve is obtained again by combination of the OBI and K-SIFT features.

[2]Please note that in historical documents the author usually differs from the writer, also called scribe, who worked as a copyist.

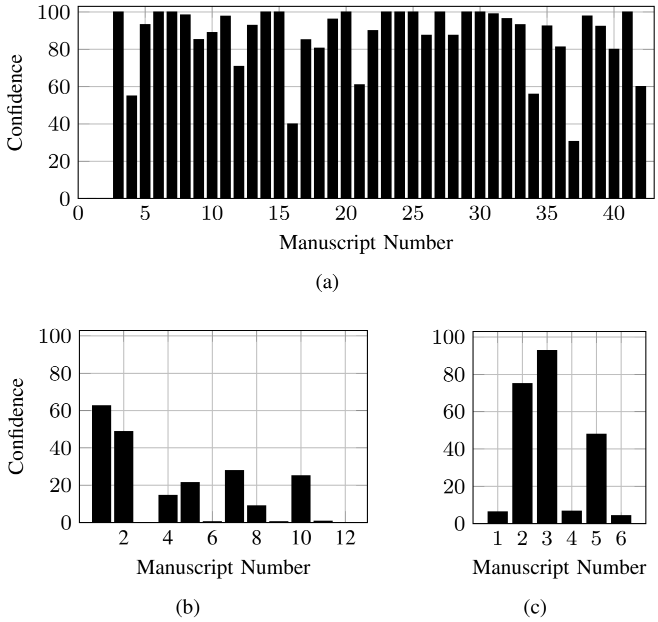

Fig. 5. Results for the number of accepted pages of Q which is regarded as a confidence measure. (a) $\mathcal{S}_{\text{multi}}$, (b) $\mathcal{S}_{\text{single}}$, and (c) $\mathcal{S}_{\text{unknown}}$.

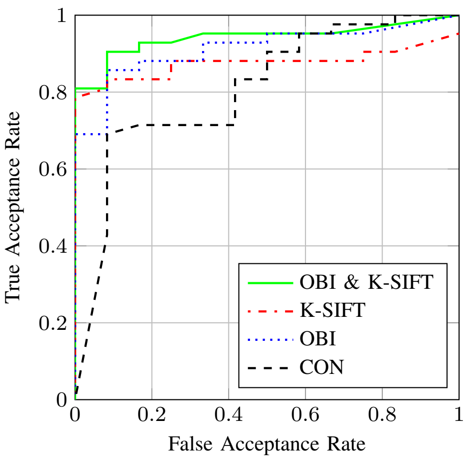

Fig. 6. ROC curves for the overall acceptance/rejection decision performance using the results in Fig. 5(a) and (b).

VI. Conclusions

Determining the individuality of handwriting in ancient manuscripts is an important aspect of the manuscript analysis process. In this paper we presented two main contributions in the context of writer identification for historical manuscripts, writer classification and retrieval approaches for multi-page documents and a rejection strategy. We have employed several feature extraction approaches which rely on contour-, textural- and keypoint-based principles and tested them with the classification schemes for the identification and retrieval of writers in multi-page documents. Our new rejection strategy is capable of rejecting decisions with a promising accuracy if no corresponding writer was found for a query manuscript. Furthermore, we presented a data set for the writer-related tasks in Arabic historical documents.

Acknowledgments

The German Research Foundation (project: FI 1494/3-2), the Ministry of Science and Technology of Israel, and the Council of Higher Education of Israel is gratefully acknowledged for financial support.

References

[1] S. Fiel and R. Sablatnig, "Writer Retrieval and Writer Identification Using Local Features." in *Proc. of Doc. Anal. Sys. (DAS)*, Queensland, Australia, March 2012, pp. 145–149.

[2] R. Jain and D. S. Doermann, "Writer Identification Using an Alphabet of Contour Gradient Descriptors," in *Proc. of Int. Conf. of Doc. Anal. and Rec. (ICDAR)*, Los Alamitos, CA, USA, August 2013, pp. 550–554.

[3] S. Fiehl and R. Sablatnig, "Writer Identification and Writer Retrieval Using the Fisher Vector on Visual Vocabularies," in *Proc. of Int. Conf. of Doc. Anal. and Rec. (ICDAR)*, Washington, DC, USA, August 2013, pp. 545–549.

[4] S. Srihari, M. J. Beal, K. Bandi, V. Shah, and P. Krishnamurthy, "A Statistical Model for Writer Verification," in *Proc. of Int. Conf. of Doc. Anal. and Rec. (ICDAR)*, Seoul, Korea, August 2005, pp. 1105–1109.

[5] M. Bulacu, L. Schomaker, and A. Brink, "Text-Independent Writer Identification and Verification on Offline Arabic Handwriting." in *Proc. of Int. Conf. of Doc. Anal. and Rec. (ICDAR)*, Curitiba, Brazil, 2007, pp. 769–773.

[6] A. Al-Dmour and R. A. Zitar, "Arabic Writer Identification Based on Hybrid Spectral-Statistical Measures." *J. of Exp. Theor. and Artif. Intell.*, vol. 19, no. 4, pp. 307–332, 2007.

[7] Sreeraj and S. M. Idicula, "A Survey on Writer Identification Schemes," *International Journal of Computer Applications*, vol. 26, no. 2, pp. 23–33, July 2011.

[8] A. Hassane and S. Al-Madeed, "ICFHR 2012 Competition on Writer Identification Challenge 2: Arabic Scripts." in *Proc. of. Int. Conf. in Front. of Handwr. Rec. (ICFHR)*, Bari, Italy, September 2012, pp. 835–840.

[9] M. Aksela, J. Laaksonen, E. Oja, and J. Kangas, "Rejection Methods for an Adaptive Committee Classifier," in *Proc. of Int. Conf. of Doc. Anal. and Rec. (ICDAR)*, St. Malo, France, 2001, pp. 982–986.

[10] D. Cakmakov, V. Radevski, Y. Bennani, and D. Gorgevik, "Decision Fusion and Reliability Control in Handwritten Digit Recognition System," *Jornal on Computing and Information Technology*, vol. 4, pp. 283–293, 2002.

[11] Harvard University. Islamic Heritage Project (IHP). [Online]. Available: http://ocp.hul.harvard.edu/ihp/

[12] D. Fecker, A. Asi, V. Märgner, J. El-Sana, and T. Fingscheidt, "Writer Identification for Historical Arabic Documents," in *Proc. of Int. Conf. on Pattern Recogn. (ICPR) (submitted)*, Stockholm, Sweden, August 2014.

[13] I. Fogel and D. Sagi, "Gabor Filters as Texture Discriminator," *Biological Cybernetics*, vol. 61, no. 2, pp. 103–113, June 1989.

[14] A. J. Newell and L. D. Griffin, "Natural Image Character Recognition Using Oriented Basic Image Features." in *Proc. of Int. Conf. Dig. Im. Comp.: Tech. and App. (DICTA)*, Queensland, Australia, December 2011, pp. 191–196.

[15] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[16] C. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge: Cambridge University Press, 2008, vol. 1.

[17] D. Tax and R. Duin, "Support Vector Data Description," *Machine Learning*, vol. 54, pp. 45–66, 2004.

[18] C. Bishop, *Pattern Recognition and Machine Learning*. Springer (Inf. Science and Statistics), 2006.

[19] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.

[20] R. Altman, "The Illusion of One Writer in Historical Documents and its Effect on Automating Writer Identification," in *Proc. of 14th Conf. of the Int. Graphonomics Society*, Dijon, France, September 2009.