Arabic Manuscripts Identification Based on Feature Relation Graph

Oleg Redkin, Olga Bernikova, Dmitry Shalymov, Vladislav Pavlov
Saint-Petersburg State University
Saint-Petersburg, Russia
oleg_redkin@mail.ru, o.bernikova@spbu.ru, dmitry.shalymov@gmail.com, vlad.pavlov24@gmail.com

***Abstract*—We investigate a new metric based on the Feature Relation Graph (FRG). This metric has proved to be effective for the text independent Persian writer identification. Since Persian script is based on Arabic writing similar principles of analysis may be also applied to the Arabic manuscripts. We have investigated the FRG for Arabic handwritten texts. Pattern based features are extracted from handwritten texts using Gabor and XGabor filters. The extracted features are represented for each author based on the FRG that plays a role of a feature vector in the classification problems. We have also investigated different parameters of the FRG.**

I. Introduction

Mediaeval Arabic civilization developed foundations of mathematics and astronomy, chemistry and philosophy, geography and jurisprudence, medicine, etc., which is vital element of the cultural heritage of mankind. It is not surprising that old Arabic manuscripts are in the focus of interest of researches in the East and in the West alike. Until recently manuscript studies as a rule were based on the textual analysis alone and only few scholars along with the content analysis, took into consideration quantitative parameters of the manuscript [1]. It has become clear that conventional methods of research often can not give answers to all questions about manuscript authorship, it classification, dating, and origin etc.

Meanwhile computational approaches may be rather effective in comparative analysis of manuscript texts, and provide authorship attribution and dating. As a rule, digitization technologies based on linear and vertical segmentation of written text are less effective for Arabic than for other languages. All this makes it necessary to find other ways of digitization as well as means of the manuscript layout analysis as well as to define the most important markers, which are relevant to the text distinction and singularity, and on this basis to carry out comparative analysis of various texts. The effectiveness of computational methods for such kind of analysis is often restricted by the quality of the written material its suitability for computational analysis and the level of noises - tears of paper, lacunas in text, contaminations, later additions to original text, etc. Many ancient writers did not mention themselves in their texts.

This fact creates a problem for scientists to identify the authorship of several manuscripts. Investigation of algorithms for processing and classification of such manuscripts becomes very important in such cases. On the basis of these algorithms it is possible to create a software system that is able to identify author with a certain degree of probability. For the last two decades such algorithms have been intensively investigated and significant results have been obtained.

Most of the modern systems for writer identification can be divided into online and offline systems. Online systems use information obtained from the mere process of writing while offline systems use the information extracted from the text only. Also such systems can be divided into the text-dependent and text-independent systems. The first ones are good for a fixed set of written texts, while text-independent systems are insensitive to the texts being processed. We investigate offline text-independent system for Arabic handwritten texts authorship identification. This system can be used to process a set of scanned documents with unknown authorship and perform a clustering. Based on the system it is possible to identify the subset of documents with the same handwriting, select the number of different writing styles and determine a set of potential authors of a manuscript.

II. Materials

Arabographic handwritten text classification is of high importance for manuscript libraries worldwide. Currently there are a lot of digital collections of manuscript, among them the most famous are: The British Librarys collection [2], Princeton Digital Library of Islamic Manuscripts [3], Islamic Manuscripts Collection at the University of Michigan [4], Juma Al Majid Center for Culture and Heritage[5] in Dubai, etc. Each of these collections has its own principles of functioning in terms of methods of digitalizing, cataloguing, search tools. In our previous research [6] we analyzed advantages and disadvantages of technology solutions used in different kinds of digital manuscript collections and came into conclusion about necessity of developing of better approach to Arabic manuscript processing. Such kind of interdisciplinary work started in St Petersburg University within the scope of the project Determination of the Formal Characteristics of the Arabic Manuscripts and their Digital Processing. The project has practical significance for St Petersburg State University, because it possesses one of the best manuscript collection in the world. Digitalization of manuscripts from this collection is considered to be very important and urgent task. That is why manuscript scanning, using special soft and hardware is accompanied by developing of technology of automatic manuscript classification, relying upon manuscripts metadata. For the purpose of the present research, we scanned and processed several fragments of four manuscripts from the collection of Saint-Petersburg State University. These fragments, written in one kind of Arabic handwriting (Naskh), belong to

ISBN 978-952-68397-0-7 (paperback), ISBN 978-952-68397-1-4 (PDF)

different authors and different period, they are:

- Commentary on the art of dispute Al-Idji; 1667-1668.
- Commentary on the logic works. Author: Al-Armawii; 1750-1751. Scribe: Muhammad Azam ibn Haafiz Samiir Kujaalii.
- Sayings of the Caliph Ali ibn Abi Talib Compiler Abd Al-Waahid ibn Muhammad ibn Abd Al-Waahid Al-Amdii (died 1144).
- The book of life. Commentary on the prayer rules. 1721, Egypt. Author: Mustafa ibn Hamza ibn Ibraaghiim ibn Walii Ad-Diin.

A fragment of manuscript taken from the manuscript collection of Saint-Petersburg State University is presented in Fig. 1

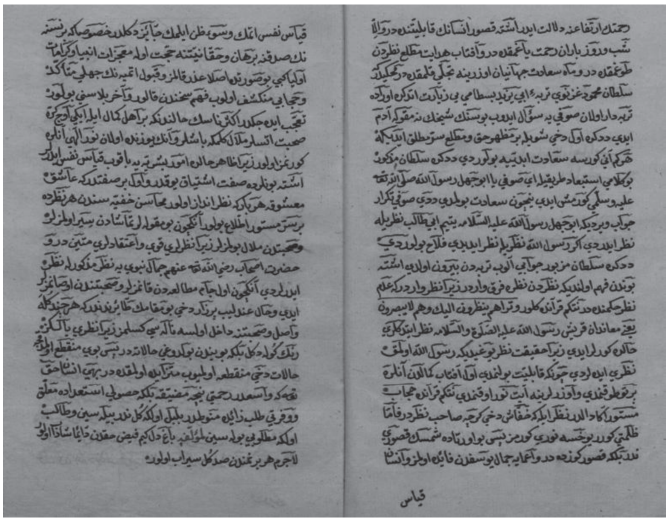


Fig. 1. A fragment of the manuscript from the collection of Saint-Petersburg State University. Commentary on the art of dispute Al-Idji; 1667-1668

The results of classification of these manuscript fragments are described in the further sections of the paper.

III. Related Works

The problem of handwritten documents clustering is often investigated together with a writer identification problem. In [7] a promising identification system was proposed where identification accuracy reached 96 percent. Several experiments were performed for 40 different writer classes. Feature extraction was organized with Gabor filters and grey-scale co-occurrence matrix technique. Classifier based on weighted Euclidean distance and the K-nearest neighbours classifier were compared.

An alternative method for automatic writer identification is proposed in [8]. The method is based on the presence of certain patterns in manuscript and orientation and curvature information of a handwritten text. The system was tested for multiple languages. Accuracy for Arabic handwritten texts was up to 92 percent for 100 writers.

Another writer identification system which extracts several types of features is proposed in [9]. Nearest neighbor classifiers were used for Euclidean distance measure. A database of handwritten paragraphs written by 250 writers was used. Experiments were performed for only 2 images for each writer. Accuracy was demonstrated up to 95.4 percent.

A text-independent Arabic writer identification method that showed significant results was presented by Bulacu et al. in [10]. For the feature extraction phase a combination of textural and allographic features were used. The nearest neighborhood classifier with χ^2 as a distance measure was used. Experiments were conducted with use of a database of manuscripts which consists of 350 writers with 5 samples per writer. The method provided an accuracy of 88 percent in top-1 and 99 percent in top-10 series of experiments.

Behzad Helli and Ebrahimi Moghaddam proposed a system of writer identification for Persian language in [11]. The system used pattern-based features to construct a special graph that represented the handwriting style of the author. Features were extracted based on Gabor and XGabor filters. Acquired feature values were compared using a fuzzy method to construct the Feature Relation Graph (FRG). A special algorithm of comparing graphs was proposed for classification. The proposed system demonstrated nearly 100 percent accuracy for 80 writers having train/test ratio equal to 3/2. These significant results inspired us to use this approach for Arabic texts. And to use an algorithm of comparing graphs for clustering problem in the FRG space.

IV. FRG Classification Algorithm

To solve the problem of classification, clusterization or writer identification a distance measure has to be determined first. We have to define a rule for calculation a similarity measure between two scanned texts (images). The algorithm proposed for Persian texts [11] can be used as the mentioned rule. The main concepts of the algorithm will be briefly described below.

A. Preprocessing

For efficient work of the algorithm each image has to be preprocessed.

1) Noise Removal: Since we are dealing with the scanned documents an external noise (random points, interference that usually occur when scanning) can dramatically affect the final result. Therefore the quality of the system depends on the quality of the input images. To eliminate these unnecessary effects we first apply Gaussian filter to gray-scale image for removing additive noise and apply Adaptive Image Thresholding. After that the line segmentation is performed.

2) Text-line Segmentation: For text-line segmentation there is a standard method of projection onto a vertical axis when local extrema are searched on the histogram. If a threshold value is passed then the corresponding y-coordinate is considered as the center of text line. One disadvantage of this method is that the algorithm can determine only straight text lines and also can fault with diacritic symbols. Therefore, we have decided to apply Block Covering (BC) algorithm [12]. In contrast to the standard methods, BC algorithm is able to process texts written with both horizontal and changing the angle (curve) lines. It is also resistant to the row intersections which substantially increases the accuracy of the system.

The basic idea of this method is to divide the input image into several vertical strips with fixed width r. Thereafter, in each of the strips all maximal blocks containing only text fragments have to be found. We should select all connected components in each of the vertical strips. By implementation we simply iterate through the rows of pixels and consider the sum S of them. If S becomes zero then we compare it with the previous sum S_p. If S_p is greater than 0 then we have found the end of the current text fragment. Also if the previous sum S_p is 0 and the current S is greater than 0 then we have found the beginning of a new fragment. The approach is based on the blocks which cover the entire document image without overlapping. BC method [12] is based on selecting the optimal width of the vertical strips, statistical analysis of block heights and analysis of blocks neighboring. This algorithm depends only on the width of the vertical stripes r and consists of the three steps:

1) Fractal analysis. It leads to the document feature vector extraction and classification (r can vary) into one of two classes: TSD (tightly spaced documents) and WSD (widely spaced documents).
2) Statistical analysis (SA) of the block heights (r is fixed). It is needed to determine an optimal value for the parameter r. SA is based on the three classes for TSD documents: small blocks (diacritics symbols), average blocks (main stream of writing) and large blocks (overlapping and touching lines). Two classes are used for WSD documents: small blocks and average blocks. For TSD documents each large block is divided into several average blocks.
3) Blocks neighboring analysis. It aims to build resulting text-lines by analyzing blocks in neighboring strips and revealing confident patterns which determine if blocks are in one text line or not.

B. Gabor Filter and Feature Extraction

For each input image a set of features is extracted using two-dimensional Gabor filters tuned for different orientations. Gabor filter is widely used in various problems of pattern recognition. Such popularity appeared after similarities between human visual system and Gabor filters have been found. Gabor filter is a remarkable tool for image filtering in the frequency and space domains. The two-dimensional Gabor function is known to be defined as a product of sinusoidal wave with a Gaussian function. It can be set by the following equation:

$$g(x, y, \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(\frac{-(u^2+\gamma^2 v^2)}{2\sigma^2}\right)\exp\left(\left(2\pi\frac{u}{\lambda}+\psi\right)i\right), \quad (1)$$

where $u = x\cos\theta + y\sin\theta$ and $v = -x\sin\theta + y\cos\theta$. In Eq.(1) λ represents the wavelength of the sinusoidal factor. For our case it was related to the thickness of the pen with which a text was written. Parameter θ represents the orientation of the normal to the stripes of a Gabor function. It controls the orientation of patterns for which the filter gives a stronger response. Parameter ψ is the phase offset. For our experiments it was always taken equal to 0. Parameter σ is the standard deviation of the Gaussian envelope, γ specifies the ellipticity of the Gabor function, in our experiments it was taken as 1 that made the filter circular. The parameters λ and σ are actually dependent.

The ratio $\frac{\sigma}{\lambda}$ determines the spatial frequency bandwidth and thus the number of observable parallel excitatory and inhibitory stripe zones. The half-response spatial frequency bandwidth b and the ratio $\frac{\sigma}{\lambda}$ are related as follows: $\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln(2)}{2}}\frac{2^b+1}{2^b-1}$. Hence, when the values of λ and b are given then the σ parameter can be calculated.

Let us focus on how Gabor filters work. As it was said Gabor filters enable us to detect specially-oriented local patterns in handwritings with particular wavelength. If a thickness of of a line in a handwritting is known and we would like to detect vertically-oriented patterns in an image a kernel for Gabor filter parametrized with $\theta = 0$ should be used. The result of convolving such a kernel is presented in 2.

Fig. 2. The result of usage of a Gabor filter for detection of vertical patterns

If, on the contrary, we would like to have stronher response of the filter to horizontal patterns we set the λ parameter to $\pi/2$. The result of convolving a kernel parametrized with $\lambda = \pi/2$ is presented

Fig. 3. The result of usage of a Gabor filter for detection of horizontal patterns

C. Generating the Feature Relation Graph

As the FRG is one of the key concepts of the work it should be described more precisely. When all l features are extracted for all lines of m images a maximum difference M between all features is computed.

$$M = \max_{s\in\{1..m\}, t\in\{1..t\}}(pr(v_s, t)) - \min_{s\in\{1..m\}, t\in\{1..t\}}(pr(v_s, t), \quad (2)$$

where $pr(v, i)$ is an i_{th} coordinate of a vector v.

To calculate the level of inequality between pair of features a five global variables are generated based on M. Let us put $d = \frac{M}{3}$. The value M_k can be calculated using the equation 3

$$M_k = kd - M, \quad (3)$$

where $d = \frac{M}{3}$

In [11] a set of variables r_1, r_2, r_3, r_4, r_5 is proposed to calculate for all pairs of features. Each r-variable showed value of inequality between pair of features. The behaviour of the variables based on difference of input values x and y is showed in Fig. 4.

These values were stored in matrices R_1, R_2, R_3, R_4, R_5 for each vector v_s where $R_k(i, j, s) = r_k(pr(v_s, i), pr(v_s, j))$.

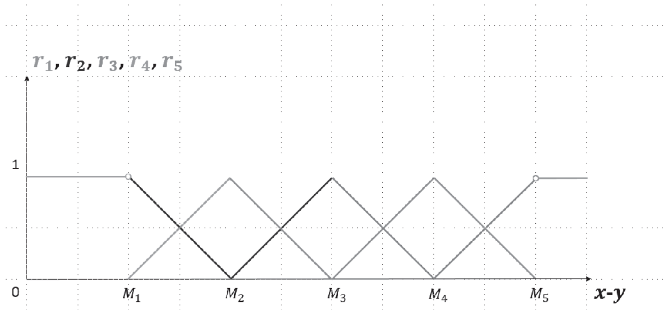


Fig. 4. r-variables

After R-matrices are obtained a final matrices Π_1, Π_2, Π_3, Π_4, Π_5 are calculated as:

$$\Pi_k(i,j) = \frac{\sum_{s=1}^{l} R_k(i,j,s)}{l} \tag{4}$$

Now the FRG can be calculated. The FRG is an unweighted directed graph that can contain not more than l vertices where each vertex denotes a feature. A set of edges of the FRG G is defined by the following rule:

$$(x,y) \in E(G) \leftrightarrow \sum_{k=1}^{5} (k-3)\Pi_k(x,y) \geq 1 \tag{5}$$

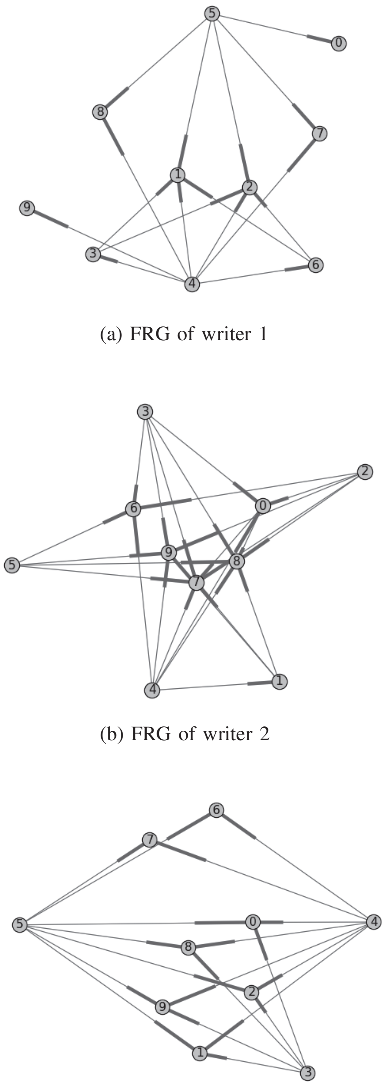
As explained in [11] a presence of an edge from vertex 1 to vertex 2 represents a tendency of feature 1 to be greater than feature 2 in all lines of input images set. From definition of the FRG it follows that it cannot contain a cycle. Otherwise transitivity of inequality is broken. Process described above enables us to interpret a set of images as the FRG.

D. Classification Problem

Let us denote the FRG of a test set by U and a set of created after the previous stage graphs by Γ. For each graph G_i in Γ values $S(U, G_i)$ are calculated as a measures of similarity between U and G_i. The $S(G_1, G_2)$ is calculated as the total number of common paths in G_1 and G_2. Special algorithm for calculating $S(G_1, G_2)$ is proposed in [11]. For all vertices of input graphs the length of the longest path from the vertex to a leaf is calculated. Based on the computed values a special sorting of the edges is proposed. $S(G_1, G_2)$ can be calculated by iterating over all edges in the graph. It should be noticed that $S(G_1, G_2)$ is a similarity measure (i.e. integer non-negative value) which increases when two graphs become more similar. Graph B is chosen so that $S(U, B) = \max_{G_i \in \Gamma} S(U, G_i)$. We can build B for each training set and hence define a 'center' of a corresponding class for each author. By this way a classification problem can be solved.

V. Numerical Experiments

According to [11] the algorithm briefly described above has showed a good results for Persian texts. We have investigated this algorithm on Arabic handwritten texts. We have conducted several experiments using different sources of arabic handwritten documents. KHATT database[13] was taken as one of the sources. Over a thousand persons from different countries, of different sex and age participated in creation of this database. The database stores images of 40000 texts where line segmentation has already been prepared. The examples of line images of different authors from KHATT database are presented at Fig 6.



(a) FRG of writer 1

(b) FRG of writer 2

(c) FRG of writer 3

Fig. 5. Examples of the FRGs for three authors

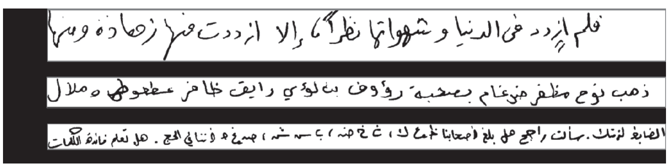


Fig. 6. Line images of three different authors from KHATT database

The second source was supplied by Faculty of Asian and African Studies of Saint-Petersburg State University. This source consists of handwritings of four different authors. Manuscript of each author had four images, each of which was a scan of two pages of the author's handwritten text. Each page of text contains approximately 20 lines. We had to segment texts of authors manually as we did not solve the problem of segmentation. The text on pages was well aligned so we used a basic algorithm to get lines of an author: an image was divided in several blocks and valid lines were taken for experiments. Luckily, such an approach gave us more than 70 lines for each writer. The examples of line images of different authors are presented at Fig 7.

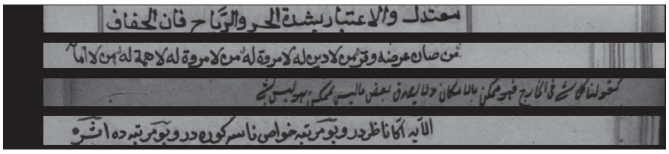


Fig. 7. Line images of three different authors from materials supplied by Faculty of Asian and African Studies of Saint-Petersburg State University

Experiments where we used these sources are described further.

A. *Classification of Arabic Handwritten Texts*

We have performed a set of experiments to test the classification algorithm [11]. Each experiment is specified by the following parameters:

- a - a number of authors participating in experiment
- p - a ratio of test images and train images number for the writer
- l - a number of totally extracted features

The accuracy for each of experiments is denoted as c. It is calculated as a ratio of correctly classified texts and total texts amount. Corresponding experiments are described below.

For the *first set of experiments* there were used 4000 files - 4 files for each writer. We took images from KHATT Database. Files were segmented manually as we do not resolve the problem of text segmentation here. Each file contains from 2 to 5 segmented text lines. Authors were selected randomly, therefore a is equal to 3. For each author more than 12 text lines was selected. p was set to 8/4 and 4/8. l was set to 8 and 24.

The *second set of experiments* contained over 13 thousand line images of different authors from KHATT Database. Each writer had approximately more than 12 line images.

The *third and forth set of experiments* are based on a text lines from the second set of experiments.

The most considerable results of experiments described above are presented in Table I.

Experiments with text data supplied by The East Faculty of Saint-Petersburg State University are presented in Table II. The *first set of experiments* were conducted using text lines supplied by East Faculty of Saint-Petersburg State University. In this set of experiments and following a was put to 4, as we were supplied with manuscripts written only by 4 different authors. In the fifth set of experiments we set l to 8 and changed the value of p to 8/20, 20/8, 20/10, 10/20.

The *second and third set of experiments* are based on the first set of experiments. The only difference is that the number of features was increased.

The parameters of Gabor filters were changed during the experiments. During these experiments the most optimal parameters for the filter were calculated. Experiments with parameters of Gabor filters will be described further.

TABLE I. CLASSIFICATION OF TEXTS FROM THE KHATT DATABASE

Experiment	a	l	p	Precision c (%)
1	3	8	8/4	87
1	3	8	4/8	96
1	3	24	8/4	93
1	3	24	4/8	100
2	10	16	8/4	76
2	10	16	4/8	80
2	10	64	8/4	72
2	10	64	4/8	86
3	40	16	8/4	62
3	40	16	4/8	70
3	40	64	8/4	63
3	40	64	4/8	74
4	100	24	8/4	54
4	100	24	4/8	63
4	100	76	8/4	70
4	100	76	4/8	72

TABLE II. CLASSIFICATION OF TEXTS FROM THE SAINT-PETERSBURG STATE UNIVERSITY COLLECTION

Experiment	a	l	p	Precision c (%)
5	4	8	8/20	95
5	4	8	20/8	93
5	4	8	10/20	100
5	4	8	20/10	100
6	4	10	8/20	98
6	4	10	20/8	98
6	4	10	10/20	100
6	4	10	20/10	100
7	4	24	8/20	100
7	4	24	20/8	100
7	4	24	10/20	100
7	4	24	20/10	100

It can be seen that the algorithm mentioned above can be successfully used for Arabic texts also. The accuracy of algorithm increases with the increase of training data set capacity. When number of features is significant and training and test sets capacity is considerable the algorithm can show up to 100 percent precision of classification on small amount of authors.

B. *Different Parameters of Gabor Filter*

The popularity of Gabor filters (1) can be explained by the fact that visual systems of humans and some animals can be approximated with these filters. Neurophysiological research has shown that the half-response spatial-frequency bandwidths of simple cells vary in the range of 0.5 to 2.5 octaves in the visual system of a cat and 0.4 to 2.6 octaves in the system of a macaque monkey. While there is a considerable spread, usually the bulk of cells have bandwidths in the range of 1.0-1.8 octaves. According to these facts we have checked several values for b (Eq.1). In our experiments we changed the parameter λ, the amount of Gabor features f and number of authors a. The identification process was assessed by the precision value c which determines the ratio of correctly identified authors to the total number of experiments in percents. Several results for Gabor filter are presented in Table III.

TABLE III. PARAMETERS OF GABOR FILTER

a	f	λ	b	Precision c (%)
5	5	4	1.6	60
5	5	7	1.6	80
8	16	7	1.4	25
8	16	4	1.6	37.5
8	16	7	1.8	37.5
8	16	7	1.6	50
8	32	4	1.6	50
8	32	4	1.8	50
20	16	4	1.6	25
20	32	4	1.6	35
20	32	7	1.6	50
20	64	7	1.6	60
30	64	7	1.6	50

It can be seen that the more the number of Gabor features f for the given number of authors a, the better precision value c.

As it is shown in [11] the usage of only Gabor filters demonstrates up to 60 percent of precision. It corresponds to our experiments. To improve the results the XGabor filter [11] can be proposed in addition to Gabor filter. In this case the length of a feature vector is calculated as $l = l_1 + l_2$, where l_1 and l_2 are the numbers of extracted Gabor and XGabor features respectively. Experiments shows that the best results is obtained when l_1 and l_2 are approximately equal.

VI. CONCLUSION

In this paper we have used an algorithm for handwritten texts classification proposed by B.Helli and M.E. Moghaddam [11] for Persian language to solve the problem of Arabic handwritten texts clustering. We have investigated performance of text-independent FRG approach for classification of Arabic texts. The FRG is a graph based on fuzzy variables that is created for a handwritten text lines. To compare the FRG graphs a special graph similarity measure is used. This approach has shown appropriate results especially for about four authors.

VII. FURTHER RESEARCH

Although the results of the algorithm is considerable for Arabic handwritten texts, we think that further research may be conducted. Now we are examining how the FRG algorithm can be used for solving a problem of handwritten texts clustering. We are planning to have a plenty of clustering algorithms and metrics for assesment of the results of the clustering process implemented. These may give us additional information about more precise tuning of the filters. We believe that the algorithm is applicable not only to Arabic handwritten texts. By now our list of plans includes collecting a Russian dataset of handwritten texts and evaluating the algorithm performance on it.

ACKNOWLEDGMENT

The authors acknowledge Saint-Petersburg State University for a research grants 6.37.181.2014 and 2.37.175.2014.

REFERENCES

[1] Polosin V. To the Method of Describing Illuminated Arabic Manuscripts Muslim Manuscripts: Quantitative Methods of Research. Selected articles. Saint-Petersburg: MAE RAS- Manuscripta Orientalia, 2015, Vol.5, p. 21.

[2] The British Librarys collection arabic materials, Web: bl.uk/reshelp/findhelplang/arabic/arabicsection/arabicmanuscripts.

[3] Princeton Digital Library of Islamic Manuscripts, Web: library.princeton.edu/projects/islamic.

[4] Islamic Manuscripts Collection at the University of Michigan, Web: www.lib.umich.edu/islamic.

[5] Juma Al Majid Center for Culture and Heritage, Web: www.almajidcenter.org.

[6] Redkin O., Bernikova O. Methods of the Arabic Manuscripts Digitization. In Proceedings of the 2014 Conference on Artificial Intelligence, 21-24 July, 2014, Las Vegas, Nevada, USA, pp. 511-516.

[7] H.E.S. Said, T.N. Tan, K.D. Baker *Personal identification based on handwriting*, Pattern Recognition, Vol. 33(1), 2000, pp. 149-160.

[8] Imran Siddiqi , Nicole Vincent *Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features*, Pattern Recognition, Vol 43(11), 2010, pp. 3853-3865.

[9] Sameh M. Awaida, Sabri A. Mahmoud *Writer identification of Arabic text using statistical and structural features*, Cybernetics and Systems: An International Journal, Vol. 44(1), 2013, pp. 57-76.

[10] M.Bulacu, L.Schomaker, A.Brink *Text-independent writer identification and verification on offline arabic handwriting* Ninth International Conference on Document Analysis and Recognition (ICDAR), 2007.

[11] B.Helli, M.E. Moghaddam *A text-independent Persian writer identification based on feature relation graph (FRG)*, Pattern Recognition, Vol. 43(6), 2010, pp. 2199-2209.

[12] A. Zahour, B. Taconet, L. Likforman-Sulem, W. Boussellaa *Overlapping and multi-touching text-line segmentation by Block Covering analysis*, Pattern Analysis and Applications, Vol. 12(4), 2009, pp. 335-351.

[13] S. A. Mahmoud, I. Ahmad, M. Alshayeb, W. G. Al-Khatib, M. T. Parvez, G. A. Fink, V. Margner, and H. EL Abed *KHATT: Arabic Offline Handwritten Text Database*, In Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (ICFHR 2012), Bari, Italy, 2012, pp. 447-452, IEEE Computer Society.

[14] V. Shiv Naga Prasad, Justin Domke *Gabor filter visualization*, TechnicalReport, University of Maryland, 2005.