Adaptive Document Image Skew Estimation

Sepideh Barekat Rezaei, Jamshid Shanbehzadeh, and Abdolhossein Sarrafzadeh

***Abstract*— The skew of the scanned document image is inevitable, and its correction improves the performance of document recognition systems. Skew specifies the text lines deviation from the horizontal or vertical axes. To date, skew estimation algorithms have employed specific features in a repetitive process. We can improve these algorithms by simply using an adaptive algorithm. This approach is suitable when we have large number of similar documents. This paper proposes adaptive document image skew estimation algorithm using the features of existing methods and supervised learning. This approach significantly improves the skew estimation time and accuracy. The time improvement comes from the training that need be performed only once on the training images rather than the repetitive process for each image of previous algorithms. The accuracy improvement comes from the appropriate selection of features, learning algorithm and image adaptively. This method works well in all skew ranges up to 0.1°.**

***Index Terms*— scanned document image, skew, skew estimation, supervised learning**

I. Introduction

IMAGE acquisition is the first step of document recognition and this may result in an output image with skew. This will complicate the recognition phase or degrade its performance. To date, several skew detection algorithms, including projection profile analysis, Hough Transform, nearest neighbour clustering, cross-correlation, piece-wise painting algorithm, piece-wise covering by parallelogram, transition counts, and morphology, have been proposed. All these algorithms employ specific features to estimate the skew and disregard prior document knowledge. In other words, these algorithms are non-adaptive. However, in situations where we have similar documents with known specifications; we can exploit this information. This happens in situations where an organization attempt to transmit from paper based into electronic based documentation. For example, it is very suitable for documents in a university (that the certifications, thesis and papers have the same text with different characteristics), bank cheque recognition (the cheques have the same form with different amount, date and name) or letter recognition in an organization.

This paper focuses on employing prior image knowledge in skew estimating, and it utilizes learning algorithm with features employed in previous algorithms. The conventional algorithms perform calculations for each image in a range of angles, while the proposed algorithm performs calculations just once for an image. The title of this paper comes out of employing features of an image set and estimating the skew of similar images by those features. We achieve this goal by a supervised learning algorithm. We syntactically rotate the training images with different angles and employ their extracted features in the training phase. Then, we perform calculations once for each new image and estimate the unknown skew by the trained system. The result shows improvements of processing time and accuracy.

The rest of this paper is as follows. Section II briefly presents the literature review of skew detection algorithms. Section III explains the proposed algorithm. Section IV introduces the data set and performance evaluation criteria. Section V shows the experimental results. Section VI compares the proposed method with the existing ones. Section VII presents the comparative analysis.

Manuscript received Nov 20, 2016. Manuscript revised Jan 9, 2017. This work is the outcome of Barekat Rezaei's M.Sc thesis at Kharazmi University.

Sepideh Barekat Rezaei is a Ph.D. student with the Department of Computer Engineering, Iran University of Science & Technology, Tehran, I.R. Iran (e-mail: s_barekat_r@yahoo.com).

Jamshid Shanbehzadeh is an Associate Professor with the Department of Electrical & Computer Engineering, Faculty of Engineering, Kharazmi University (Tarbiat Moallem University of Tehran), Tehran, I.R. Iran (phone: +98 26 34550002; fax: +98 26 34569555; e-mail: jamshid@khu.ac.ir).

Abdolhossein Sarrafzadeh is a Professor and Director, High Tech Research Network, Unitec Institute of Technology, Auckland, New Zealand (Phone +64 9 815 4321 ext. 6040; email: hsarrafzadeh@unitec.ac.nz).

II. Literature Review

Barekat Rezaei et al. [1] surveyed skew detection algorithms and presented eight groups of algorithms. Table I summarizes these algorithms. These algorithms consist of two consecutive steps: dimension reduction and skew estimation. Dimension reduction finds a criterion function by using an appropriate feature for skew. Skew estimation finds the maximum or minimum of the criterion function and introduces it as the skew. You can see a comprehensive description of the algorithms in [1] or in the related paper.

III. Adaptive Document Image Skew Estimation

This section describes Adaptive Document Image Skew Estimation (ADISE). Fig. 1 shows the six steps of ADISE. Steps 1 to 4 are the learning phase, step 5 estimates the skew and the final step finds the skew direction. Next, each step is presented in detail.

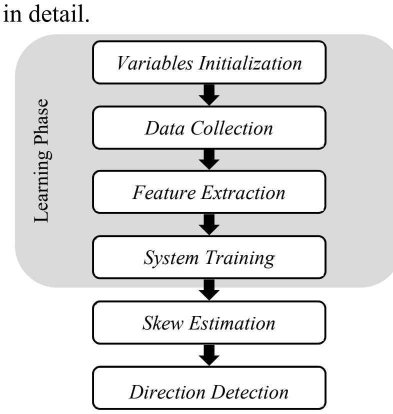


Fig. 1. Block diagram of ADISE algorithm

TABLE I. The review of skew detection methods

Algorithm	Author(s)	Year	Reference	Comments
projection profile analysis	Postl	1988	[2]	The input image is rotated to different angles. At any angle, the horizontal projection profile is obtained. The sum of squared differences between adjacent elements of the projection profile is used as the criterion function.
	Bloomberg and Kopec	1993	[3]	The variance of the number of black pixels in each row is employed as the criterion function.
	Bloomberg et al.	1995	[4]	To reduce the size of the search space, a small part of total search space containing the skew is selected and horizontal projection profile analysis is performed.
	Kavallieratou et al.	1999	[5]	The curve of maximum intensity corresponding to the Cohen distribution of each horizontal profile projection is employed as the criterion function.
	Li et al.	2007	[6]	A Two-dimensional discrete wavelet transform for improving accuracy was used.
	Sadri and Cheriet	2009	[7]	The number of local minimums and maximums of the projection profile is reduced, then the sum of local maximum heights and the sum of local minimum heights is calculated, and the difference considered as a criterion function.
	Papandreou and Gatos	2011	[8]	The sum of squares of the vertical projection profile is used for skew detection.
	Papandreou et al.	2014	[9]	The vertical and horizontal projection profiles are combined.
	Brodic et al.	2015	[10]	Entropy of the horizontal projection profiles is used.
Hough Transform	Srihari	1989	[11]	The rate of change in the accumulator array is used as the criterion function.
	Hinds	1990	[12]	The Run-length method for computational cost decreasing is used.
	Yu and Jain	1996	[13]	Authors proposed a Hierarchical Hough Transform.
	Amin and Fischer	2000	[14]	Connected components of image for computation decreasing grouped.
	Epshtein	2011	[15]	The author estimated the skew of interline regions instead of the text lines.
	Kumar and Singh	2012	[16]	The HT space is divided each time into one-tenth and a portion, including the final skew, is selected using the base HT method.
nearest neighbor clustering	Hashizume	1986	[17]	The author used the vector connecting the connected component to its first nearest neighbor only and used the vector angle histogram as the criterion function.
	O'Gorman	1993	[18]	The author used k nearest neighbors of each connected component.
	Jiang	1999	[19]	The author proposed Focused Nearest-Neighbor Clustering (FNNC).
	Liolios	2001	[20]	The connected components of each text line placed in a cluster.
	Lu and Tan	2003	[21] [22]	K-Nearest-Neighbor Chain (K-NNC) is proposed for accuracy improving.
	Konya et al.	2010	[23]	Euclidean minimum spanning tree (MST) is used.
	Kleber et al.	2014	[24]	FNNC method is enhanced by the analysis of paragraphs/lines.
	Fabrizio	2014	[25]	The author used KNN clustering and Fourier transform.
cross-correlation	Yan	1993	[26]	The cross-correlation between the two vertical lines with a certain distance is calculated and the sum of them as the criterion function used.
	Chaudhuri	1997	[27]	Instead of finding the correlation for the entire image, the author calculated it for small randomly chosen areas. He considered the median maximum cross-correlation as the criterion to obtain the skew.
	Chen	1999	[28]	The author added the verification stage to Chaudhuri's method for determining the suitability of the area. In order to improve accuracy, he calculated horizontal and vertical cross-correlations and used the one with more specific peaks.
	Brodic and Milivojevic	2013	[29]	The cross-correlation between log-polar transformation of text and ellipse image used.
piece-wise painting algorithm	Alaei	2011	[30]	See [30].
piece-wise covering by parallelogram	Chou et al.	2007	[31]	The scan lines are drawn at different angles and for each angle, a parallelogram is drawn. The area of complementary parallelograms is then calculated as a criterion function at each angle.
	Mascaro et al.	2010	[32]	Authors improved the above method in three contexts: getting the size of parallelogram complementary regions, efficient searching for the best angle, preventing undesirable interference caused by components such as noise, vertical separator.
transition counts	Ishitani	1993	[33]	The variance of the transition counts for skew detection used.
	Chen and Wang	2000	[34]	The feature selection sub-step is added before calculating the variance and selected the region of the image containing enough text.
morphology	Chen and Haralick	1994	[35]	Authors extracted the image text lines using morphological operators. In this method, fill the space between the characters by using morphological operators.
	Guru et al.	2013	[36]	Authors segmented each word using morphological operations, estimated skew of each word by fitting a minimum circumscribing ellipse, clustered the words using adaptive k-means clustering to identify the multiple blocks and estimated the average orientation of each block.
	Wagdy et al.	2014	[37]	Authors divide the document into connected component using Erosion Morphological operation, then detect the skew using the extreme points of the largest connected component.

A. Variables Initialization

This step decides the range of skew angles. This range specifies the minimum and maximum skew angle that will be estimated, and can be extended arbitrarily to include all the possible angles. However, a much wider range results in more training time. Meanwhile, this step specifies the step size which is the smallest skew that will be estimated.

B. Data Collection

This step chooses a suitable number of document images and removes their skew (manually) and normalizes their size. Then, it groups the document images according to their type, orientation or other specifications (optional). Finally, it syntactically rotates each image with the angles of the determined range with the desired step size. The results are the training images. We perform the next steps for each group. More precise grouping improves the accuracy.

C. Feature Extraction

This step calculates a set of features for each training image. The known image skew determines the class of features set. This paper examines fourteen features as shown in Table II. These features depend on the document skew variation. Appendix I presents the values of these features for a specific data set.

TABLE II. The employment features

Feature No.		Feature Description	Ref.
1	horizontal projection profile (HPP)	The sum of squared differences between adjacent elements of HPP	[2]
2		The variance of the HPP elements	[3]
3		The difference between the sum of local maximum heights and the sum of local minimum heights	[7]
4		The minimum of local minimum heights	-
5		The sum of local minimum heights	-
6	vertical projection profile (VPP)	The sum of the squares of the VPP elements	[8]
7		The variance of the VPP elements	-
8		The difference between the sum of local maximum heights and the sum of local minimum heights	-
9		The minimum of local minimum heights	-
10		The sum of local minimum heights	-
11	The area of complementary parallelograms obtained from the piece-wise covering by the parallelogram method for scan lines at an angle of zero degrees		[31]
12	The variance of the transition counts for the scan lines at an angle of zero degrees		[33]
13	The variance of the horizontal projection profile when the image is rotated by 1° in a counterclockwise direction		-
14	The variance of the horizontal projection profile when the image is rotated by 1° in a clockwise direction		-

D. System Training

Step C prepares the training data with specified class. This step uses a supervised learning algorithm to train a skew estimation system and utilizes it to estimate the skew of new images. Several supervised learning algorithms exist, including Discriminant Analysis, Support Vector Machines, Nearest Neighbour Classification, Radial Basis Function Networks and Bayesian Networks. According to the type of the training data and their class, this paper employs Discriminant Analysis.

E. Skew Estimation

This step calculates the feature set for each new image and determines the unknown image class (image skew) by the trained system. If we group the document images in the second step, we have to determine the image group at first, and then use the corresponding trained system for skew estimation.

F. Direction Detection

The features discussed here produce approximately equal values for two opposite skew directions. Therefore, we perform the skew estimation in two steps. The first step estimates the absolute value of unknown skew. The second step finds the skew direction (skew sign). In the second step, a 1° image rotation in two opposite directions and measuring the variance of the horizontal projection profile of each direction finds the skew sign. If the variance of the clockwise direction is greater than or equal to the other variance, the sign is negative. Appendix II presents the validity of this approach.

When there is no skew, we will get the maximum mean and minimum variance for the horizontal projection profile. Any skew results in the mean reduction and the variance increment. As a result, the mean and variance are suitable features in direction detection. If we measure these features before and after skewing, their differences will show the direction. The positive mean difference and negative variance difference show the correct skew direction.

IV. Performance Evaluation

This section explains the elements of performance evaluation, including the data set and performance criteria. This paper employs the old documents of Kharazmi University as data set, and accuracy and speed as the performance criteria. Upcoming sections present these two in detail.

A. Data set

The skew estimation methods can be evaluated using a data set of images with a known skew angle. This section explains how to create such a data set and details of the employed data set. We can build a data set of images with a known skew angle in two ways:

1. Collecting a data set of document images with zero degree angles (no skew), then rotating each image into a set of angles manually (synthetic data set).
2. Collecting a document image data set, then manually finding the skew of each image (a real data set, which means that the rotation of the images is created during the scanning process).

This paper evaluates the proposed method and compares it with other methods using an image data set of Kharazmi University's old documents containing 152 images. We divide the images into three groups of similar documents containing 77, 28 and 47 images. We employed 70% of images as training, and the rest for test. We synthetically generate images by rotating the images of each group in the range of ±15° with 1° and 0.1° step sizes. It means that we rotate an image step by step with angles in the ranges [-15, -14, -13, …, 13, 14, 15] and [-15, -14.9, -14.8, …, 14.8, 14.9, 15]. We further show the efficiency of the proposed method for a large dataset by using another data set from old documents of Kharazmi University that contains 3000 images of students' certificates. The employed system for skew estimation was trained with 50 images.

B. Performance Evaluation Criteria

We evaluate the skew estimation methods by speed in terms of computational time (T) in seconds, and the accuracy as the closeness of the estimated and the actual skew in terms of error. For this purpose, suppose we have a data set of N images with known skew, and we implement a skew estimation method. Suppose θ_i and $\hat{\theta}_i$ show the skew and its estimated one of the image number i respectively, then, equation (1) shows the absolute error of each image.

$$e_i = \left|\theta_i - \hat{\theta}_i\right| \qquad i = 1, \ldots, N \tag{1}$$

Three statistical factors of error, including maximum, mean and variance show the algorithm accuracy. Equations (2)-(4) show these three factors.

1. The maximum of absolute error:
$$e_{max} = max(e_i) \qquad i = 1, \ldots, N \tag{2}$$
2. The mean of absolute error:
$$\mu = \frac{1}{N}\sum_{i=1}^{N} e_i \tag{3}$$
3. The standard deviation of absolute error:
$$SD = \left[\frac{1}{N}\sum_{i=1}^{N}(e_i - \mu)^2\right]^{\frac{1}{2}} \tag{4}$$

In addition to the above criteria, the performance evaluation employs the percent of correct skew estimation and the percent of estimated skew with the absolute error in an interval.

V. Experimental Results

This section presents the experimental results by employing each feature, subset of features, the impact of employing different number of training images and noise, and non-textual components in Sections A-D respectively.

A. The Effect of Features

This section shows the results when employing each feature and all features mentioned in Table II. Table III shows the results for the first image group. Where all features and feature 11 are used we get almost the best result based on error parameters.

TABLE III. Results for images rotated to angles in the range of ±15° with 0.1° step size when using each feature. The first three data are the error parameters in degree and the last one is the computational time in second.

Feature No.	e_{max}	μ	SD	T
1	10.5	0.87	0.99	111
2	7.6	0.81	0.72	112
3	10.5	1.93	1.79	151
4	11.7	2.45	1.88	143
5	11.7	3.26	2.42	135
6	6	1.64	1.46	118
7	4.7	0.55	0.45	108
8	15	3.27	2.23	141
9	13.4	1.98	2.24	142
10	10.9	2.28	1.84	140
11	0.5	0.04	0.06	576
12	9.8	2.27	1.75	152
13	1	0.55	0.21	103
14	1	0.55	0.21	111
All	**0.2**	**0.01**	**0.02**	**681**

B. Subsets of Features

This section studies the employing of a subset of features. Table IV shows the results with and without feature selection. Methods 1, 2 and 3 first employ a feature selection method, feature selection without cross-validation, feature selection with 2-fold cross-validation, feature selection with 10-fold cross-validation respectively and, method 4 directly use all 14 features without employing any feature selection method.

TABLE IV. Results of use or non-use of different feature selection methods for images rotated to angles in the range of ±15° with 0.1° step size

Group	Method	Training Time	Selected Features	e_{max}	μ	SD	Time
1	1	6130	11, 14, 13, 2, 1, 7, 5, 10, 12, 3	0.3	0.01	0.02	689
	2	5072	11, 12	0.5	0.04	0.06	622
	3	24715	11, 14, 13	0.2	0.01	0.02	562
	4	1397	All Features	0.2	0.01	0.02	681
2	1	4595	11, 14, 13, 2, 1, 7	0.3	0.02	0.04	286
	2	5404	11, 14, 13	0.5	0.02	0.06	279
	3	27764	11, 14, 13, 2	0.3	0.02	0.05	278
	4	956	All Features	0.3	0.02	0.05	326
3	1	6366	13, 14, 11, 4, 7, 10, 5, 12, 9, 8, 3, 2	0.6	0.08	0.08	738
	2	5963	13, 14, 11	0.8	0.07	0.09	598
	3	24691	13, 14, 11	0.8	0.07	0.09	598
	4	1481	All Features	0.8	0.08	0.09	730

The employed feature selection methods choose a subset of features. They start from an empty set and sequentially add any non-selected features to create the candidate subsets. They train and test using any candidate subset, and calculate the mean absolute error, and then choose the subset with the lowest mean absolute error. This process continues until there is no change to the mean absolute error. In the case of not using cross validation, they use the total data to train and test for each candidate subset. When using k-fold cross validation, the data is partitioned into k parts, and, on each occasion, one part is used for testing and others for training. The average results of k times train and test are considered the mean absolute error of the candidate subset. Table IV shows that different feature selection method results in different subsets of features and produces different errors and increases the training time. Also, because of the different procedure of the selection in each feature selection method, the training time can be very different even though the same final selection of features are used. Therefore, based on a case, either a feature selection method or all the features can be used. We use all the features for the presented results.

C. Effect of Training with Different Number of Images

This section investigates the effect of using different numbers of training images. Table V gives the results for the first group of images, when training the system with different numbers of images. As seen in Table V, however, when the number of training images is greater, the errors will be lower.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE V. Results for images rotated to angles in the range of ±15° with 0.1° step size when the system was trained with a different number of images

Training Images Number	e_{max}	μ	SD
10	3.4	0.06	0.31
20	0.3	0.01	0.02
30	0.3	0.01	0.02
40	0.2	0.01	0.02
50	**0.2**	**0.01**	**0.02**

D. *Effect of Noise and Non-Textual Components*

The presence of noise in the scanned document images is inevitable. Types of noise are ruled line, marginal, clutter, stroke like pattern, salt and pepper, background noise [38], Gaussian and Speckle noise [39]. If the skew detection system trained based on images with some of these noises or some non-textual components, the system can estimate other images with that noise or non-textual components. For example, we show a sample image of the first group of images in Fig. 2 containing noise and non-textual components. The image skew is 0.6° and it is estimated by the system correctly.

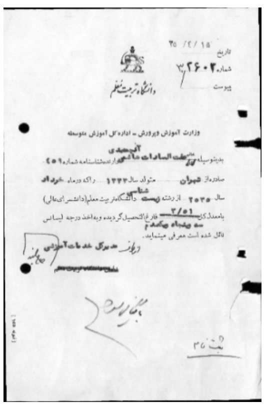


Fig. 2. A sample image of the first group of images

To show the effect of noise on the proposed method, we add some usual types of noise to the train and test images of the fourth group. Table VI shows the results. The results reveal that if the system is trained with noisy image, it can estimate the skew of new image with similar noise almost correctly.

TABLE VI. Results for images rotated to angles in the range of ±15° with 0.1° step size when we add some noise to images

Type of added noise	Parameters of the noise	e_{max}	μ	SD
salt & pepper	noise density=0.05	0.7	0.04	0.06
salt & pepper	noise density=0.1	0.6	0.05	0.06
gaussian	mean=0.1 and variance=0.01	0.2	0.03	0.05
gaussian	mean=0.1 and variance=0.05	0.9	0.06	0.07
speckle	variance=0.04	0.6	0.04	0.06
speckle	variance=0.1	0.9	0.04	0.07

VI. The Comparison Results

This section compares the proposed method with existing methods based on the evaluation criteria in different situations. Section A presents a general comparison. Sections B and C provide a comparison based on different numbers of images and different skew ranges, a fixed image number and different step size or angle ranges, respectively. Finally, Section D provides a comparison without grouping the images, and by grouping images based on their orientation. We implement all the algorithms by MATLAB and run them on Intel ® Core ™ i5-2410M CPU@2.30GHz processor.

A. *General Comparison*

Table VII gives the results for the introduced image groups when using the proposed method, and some of the best previous existing methods. According to this table, the proposed method is more accurate and faster than the other methods.

TABLE VII. The results for the image groups with angles in the range of ±15° with 0.1° step size

Group	Method	e_{max}	μ	SD	Percent of Estimation with e=0°	Percent of Estimation with 0° < e ≤ 0.2°	Time
1	Proposed method	**0.2**	**0.01**	**0.02**	**95.8**	**4.1**	**681**
	[7]	1.3	0.29	0.27	13.0	29.8	22644
	[8]	7.7	3.64	2.18	0.4	3.0	12069
	[16]	0.7	0.05	0.07	11.3	87.0	1423
	[32]	17.5	0.69	1.52	7.6	21.2	34538
2	Proposed method	**0.3**	**0.02**	**0.05**	**80.5**	**18.3**	**326**
	[7]	0.6	0.11	0.11	20.3	58.6	13137
	[8]	7.6	3.46	2.10	0.4	4.1	9503
	[16]	0.8	0.08	0.10	11.3	75.5	825
	[32]	2	0.03	0.07	68.6	30.9	29840
3	Proposed method	**0.8**	**0.08**	**0.09**	**42.4**	**48.6**	**729**
	[7]	5.9	0.50	0.93	7.9	27.5	28095
	[8]	7.8	3.47	2.12	0.2	2.9	20258
	[16]	1	0.28	0.30	8.7	43.3	1797
	[32]	1.4	0.15	0.24	53.2	24.8	66530
4	Proposed method	**0.8**	**0.04**	**0.06**	**63.2**	**34.6**	**498**
	[7]	1.1	0.20	0.19	13	45.2	14451
	[8]	7.8	1.98	2.32	12.8	16.3	11434
	[16]	1.7	0.97	0.18	0.4	55.3	1155
	[32]	16.9	0.20	0.67	25.7	39.9	37532

B. *Comparisons in Different Numbers of Images and Different Angle Ranges*

This section compares the proposed method with other methods using different numbers of images and different angle ranges. Table VIII shows the results of skew estimation of 50 and 500 randomly selected images from the first group with angles in the range of ±15° and all images in the range of ±15°, ±10° and ±5° with 0.1° step size.

TABLE VIII. The results of skew estimation using different number of images of first group with different angle ranges

	50 image		500 image		All images (8127 image)					
	-15:0.1:15		-15:0.1:15		-15:0.1:15		-10:0.1:10		-5:0.1: 5	
Method	μ	T	μ	T	μ	T	μ	T	μ	T
Proposed method	**0.01**	**5**	**0.01**	**43**	**0.01**	**681**	**0.00**	**472**	**0.00**	**237**
[7]	0.30	138	0.28	1385	0.29	22644	0.29	10073	0.29	2563
[8]	3.65	75	3.57	740	3.64	12069	2.38	5402	1.14	1362
[16]	0.05	9	0.95	88	0.05	1423	0.05	938	0.05	469
[32]	0.55	213	0.67	2136	0.69	34539	0.53	22465	0.65	10087

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

As seen in Table VIII, for the same range, the time taken by all methods increased with greater numbers of images, but in most cases, the time taken by the proposed method is approximately linearly increased. When the skew estimation of greater numbers of images is considered, however, the proposed method achieves a remarkable result. According to this table, for the same number of images, in the various angle ranges the proposed method is more accurate and faster than the other methods. Therefore, the proposed method is more suitable for a much larger range of angles and lower step sizes.

C. Comparison Using a Fixed Number of Images and Different Step Sizes or Ranges of Angles

This section compares the skew estimation time using the proposed method and other methods for a fixed number of images and different step sizes or ranges of angles. For this purpose, we randomly select 500 images from the first group of images that are rotated to angles in the range of ±15° with 1° step size (set of images 1), in the range of ±15° with 0.1° step size (set of images 2), and in the range of ±10° with 1° step size (set of images 3). Table IX shows skew estimation time for the three image sets.

TABLE IX. Skew estimation time of the three sets

Method	Set of Image 1	Set of Image 2	Set of Image 3
Proposed method	43	43	43
[7]	144	1385	96
[8]	75	740	51
[16]	33	88	30
[32]	934	2136	868

In the sets 1 and 2, the number of images and range of angles are equal, but the step size of set 2 are lower than set 1. As seen in Table IX, in this case, the processing time of our proposed method does not change much, while it increases for other methods. In addition, in the sets 1 and 3, the number of images and step size are equal, but the range of the third set is smaller than the first set. As it can be seen, in the larger range of angles, the processing time of the proposed method does not change much, while it increases for other methods. Thus, for a fixed number of images, a larger range of angles and a smaller step size, the processing time of other methods increases. In the proposed method, the range and step size only affect the training time initially for a set of similar images, and do not change the skew estimation time very much. The number of images only affects the time of skew estimation using the trained system.

D. Comparison on Non-Classified Images and Classified based on Orientation

This section provides a comparison when the data set is not classified or classified based on image orientation. The first group of images has a vertical orientation, and the images of the second and third categories are oriented horizontally. The results of the first group are reported in Table VII; only the skew estimation results of the second and third groups together are shown here. In the proposed method, the system is trained using 20 training images from each category. Table X gives the skew estimation results for three introduced groups together (non-classified images) and the second and third groups together (classified images based on their orientation). As seen in this table, the proposed method, even when not using classified images or classified images based on their orientation, works well, but not remarkably better than the other methods. We do not utilize the fourth group of images in this comparison, because this group is very much larger and the results may be biased.

TABLE X. The results for non-classified and classified images based on their orientation, rotated to angles in the range of ±15° with 0.1° step size

groups	Method	e_{max}	μ	SD	Percent of e=0°	Percent of 0 < e ≤ 0.2°	Skew Detection Time
1 & 2 & 3	Proposed	**1.6**	**0.20**	**0.22**	**21.02**	**42.82**	**1781**
	[7]	5.9	0.33	0.58	12.46	33.46	63876
	[8]	7.8	3.55	2.15	0.33	3.14	41831
	[16]	1	0.12	0.21	10.45	70.94	4047
	[32]	17.5	0.41	1.15	31.91	23.90	130910
2 & 3	Proposed	**1.1**	**0.10**	**0.11**	**38.27**	**44.90**	**1062**
	[7]	5.9	0.37	0.79	11.85	37.42	41232
	[8]	7.8	3.46	2.11	0.23	3.28	29762
	[16]	1	0.21	0.27	9.55	53.62	2624
	[32]	2	0.11	0.21	58.15	26.78	96372

VII. Comparative Analysis

This section compares the proposed method with the existing ones and shows a diagram of mean absolute errors for easier comparison.

Sadri and Cheriet [7] and Papandreou and Gatos [8], respectively, use horizontal and vertical projection profiles. They rotate each image to different angles and compute criterion functions of any angle. The angle corresponding to the maximum criterion function is considered as the skew angle. However, scanned document images have large files, and their rotation takes a lot of time. The proposed method computes a set of features for each image, and the trained system determines the image skew. So, in the proposed method, image rotation to different angles is not necessary. Therefore, the time of image rotation is removed and the processing time is greatly reduced.

As seen in the results, method [8] did not work well with our images. The method works well with written languages where most of the letters include at least one vertical line, such as those with Latin alphabets. The Persian language does not have this feature. However, we have used that feature in our method and achieved a good result, because our method uses that feature on more separate individual images with different skews.

Kumar and Singh [16] employed Hough Transform. They update the cells of accumulator array for each black pixel on the image and use a desirable range of angles with a step size equal to one-tenth size of the range. Then, they select an interval of angles and repeat the process. This process is repeated until the desired precision is reached. So, in this method, the calculations are also repeated. Since the proposed method does not require repetition, the processing time is less.

Mascaro, Cavalcanti and Mello [32] perform the calculations for angles 0°, 2° and -2°. Then, they repeat them for new intervals, which are selected based on the comparison of the previous results. They repeat this process until they achieve the desired precision. Because of the elimination of calculation repetition in the proposed method, the processing time is less.

The proposed method calculates a feature set for each image with a known skew. Each set of features is used as training data and the skew specifies its class. There is a number of training data for each skew. Due to the similarity of training images, the values of training data in each skew (class) are close to each other. When the training is completed, for every skew and every feature in that skew, a range is determined. The test images are similar to the training images, so their values are similar to the training data. Therefore, the trained system can estimate the skew of test images very well.

For an easier comparison, the mean absolute error for the first image group that are rotated to angles in the range of ±15° with 0.1° step size is shown in Fig. 3. As it can be seen, in the proposed method, the mean absolute error for all angles is almost identical and is less than the other methods.

VIII. Conclusion

This paper presented an adaptive skew estimation algorithm suitable in situations where we have a large number of scanned document images with varying skew. This method collects several same sized document images without skew and rotates them to angles in a pre-determined range with a desired step size, and calculates a set of features, and then it employs a supervised learning algorithm to train the system for skew estimation. The employed of learning method significantly improves the skew detection time. The accuracy improvement comes from the appropriate selection of skew detection features, the learning algorithm and its dependency on images.

The results show that the proposed method provides better speed and accuracy in comparison to similar algorithms. This method works well in all skew ranges up to 0.1° step size, as long as the system has been trained for that range and step size. The only limitation of the proposed method is its image dependency. The proposed method works better for the more similar training and test images. This method is especially suitable for specific OCR systems in organizations with similar documents, such as for cheque recognition in a bank, and letter recognition in an organization.

Appendix I. Feature Extraction

To show the possibility of using introduced features, Fig. 4 shows the values of these features. The horizontal axis plots the values of a feature for the first image group that are rotated to angles in the range of ±15° with 1° step size. The vertical axis is the class (angle) of each image. The symmetry about zero of vertical axis shows that the values of features at any absolute value of angle are in a specific interval. Clearly, more distinguishable intervals indicate the feature that more accurately detects the absolute value of image skew. If we want to specify the absolute value of angle, the intervals are different from each other, so we can use these features for skew detection in the proposed method.

Appendix II. Direction Detection

According to the symmetry of curves in Fig. 4 based on zero skew, the feature values for positive and negative angles have approximately the same interval. Therefore, training using the values of these features is unable to determine the skew sign, and we employ features 13 and 14 for this purpose. Fig.5. (a) and (b) shows the values of these features for the introduced data set.

To show the validity of the direction detection feature, Fig.5.(c) shows the values of features 13 and 14 diagrammatically. The diagonal line shows equal values for these two features. Thus, the positive values of angles (feature 13) are under the diagonal line, and the negative values (feature 14) are above the diagonal line. Therefore, the skew sign can be specified by comparing these two features.

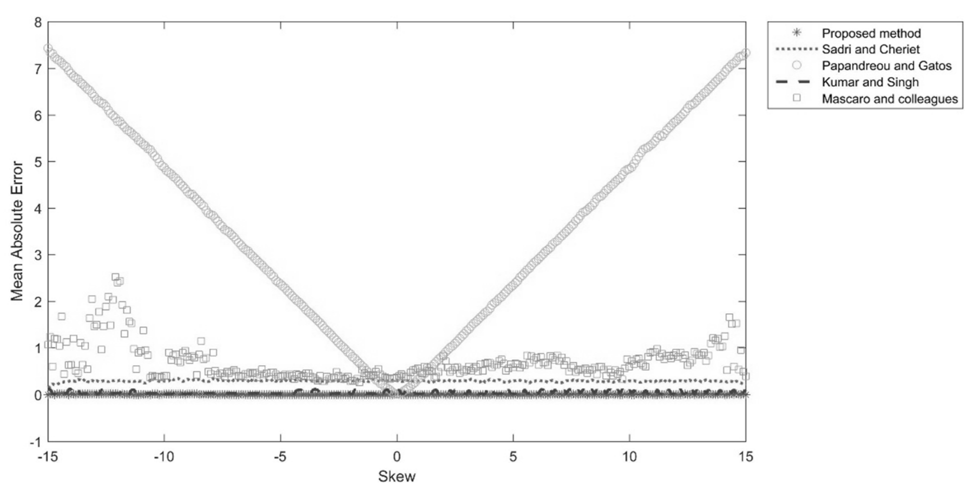


Fig. 3. Mean absolute error for the first group of images that are rotated to angles in the range of ±15° with 0.1° step size

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

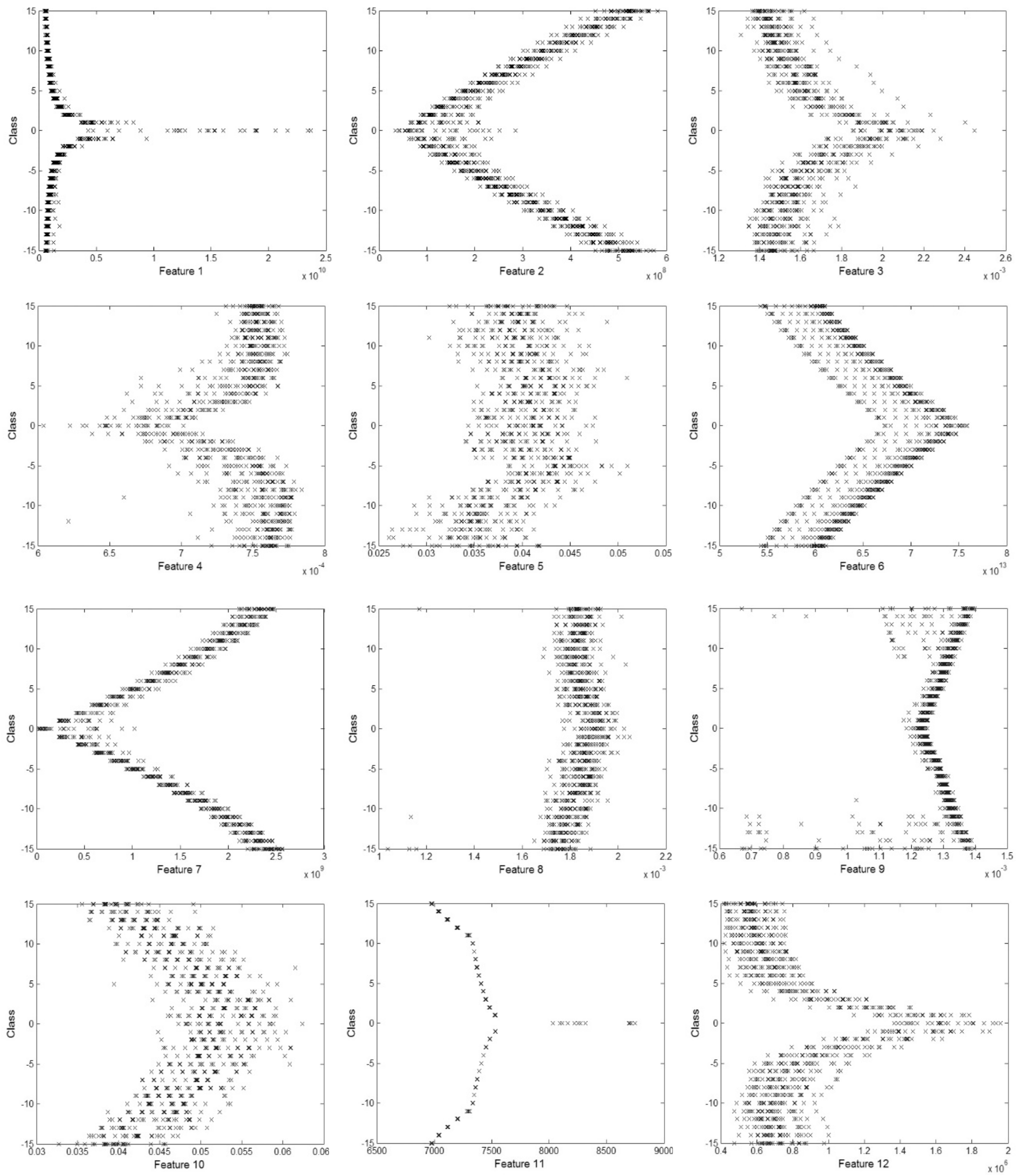
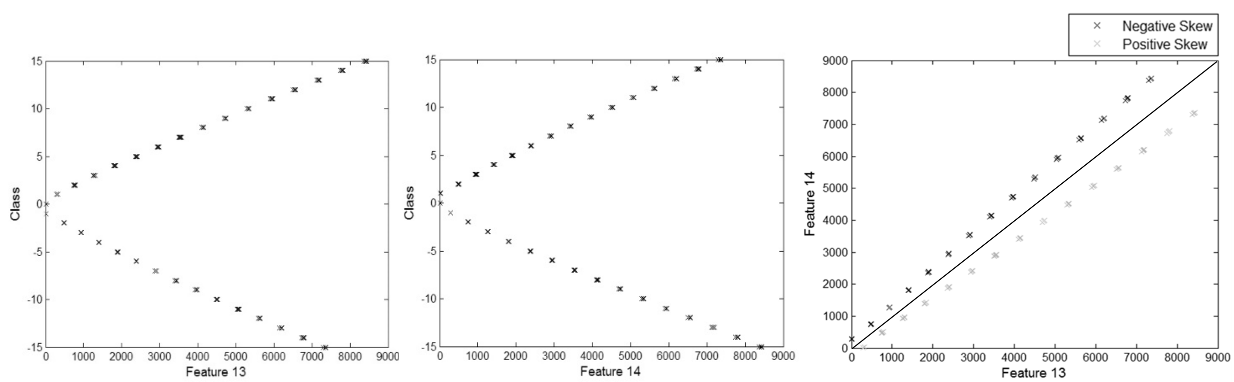


Fig. 4. The values of features 1 to 12



(a) (b) (c)

Fig. 5. (a) The values of features 13, (b) The values of features 14, and (c) The value of features 13 and 14 for images rotated to angles in the range of ±15° with 1° step size

REFERENCES

[1] S. BarekatRezaei, A. Sarrafzadeh, and J. Shanbehzadeh, "Skew Detection of Scanned Document Images," presented at the International MultiConference of Engineers and Computer Scientists, Hong Kong, 2013.

[2] W. Postl, "Method for automatic correction of character skew in the acquisition of a text original in the form of digital scan results," United States Patent, 1988.

[3] D. S. Bloomberg and G. E. Kopec, "Method and apparatus for identification and correction of document skew," United States Patent, 1993.

[4] D. S. Bloomberg, G. E. Kopec, L. Dasari, and H. S. Baird, "Measuring document image skew and orientation," presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1995.

[5] E. Kavallieratou, N. Fakotakis, and G. Kokkinakis, "Skew angle estimation in document processing using Cohen's class distributions," Pattern Recognition Letters, vol. 20, pp. 1305-1311, 1999.

[6] S. Li, Q. Shen, and J. Sun, "Skew detection using wavelet decomposition and projection profile analysis," Pattern Recognition Letters, vol. 28, pp. 555-562, 2007.

[7] J. Sadri and M. Cheriet, "A New Approach for Skew Correction of Documents Based on Particle Swarm Optimization," in Document Analysis and Recognition, 10th International Conference on, 2009, pp. 1066-1070.

[8] A. Papandreou and B. Gatos, "A Novel Skew Detection Technique Based on Vertical Projections," in Document Analysis and Recognition (ICDAR), International Conference on, 2011, pp. 384-388.

[9] A. Papandreou, B. Gatos, S. J. Perantonis, and I. Gerardis, "Efficient skew detection of printed document images based on novel combination of enhanced profiles," International Journal on Document Analysis and Recognition (IJDAR), vol. 17, pp. 433-454, 2014.

[10] D. Brodic, M. Jevtic, Z. N. Milivojevic, and V. Tasic, "Text Skew Estimation Based on the Horizontal Entropy Calculation," MIPRO/CIS, 2015.

[11] S. N. Srihari and V. Govindaraju, "Analysis of textual images using the Hough transform," Machine Vision and Applications, vol. 2, pp. 141-153, 1989.

[12] S. C. Hinds, J. L. Fisher, and D. P. D'Amato, "A document skew detection method using run-length encoding and the Hough transform," in Pattern Recognition, Proceedings., 10th International Conference on, 1990, pp. 464-468 vol.1.

[13] B. Yu and A. K. Jain, "A robust and fast skew detection algorithm for generic documents," Pattern Recognition, vol. 29, pp. 1599-1629, 1996.

[14] A. Amin and S. Fischer, "A Document Skew Detection Method Using the Hough Transform," Pattern Analysis & Applications, vol. 3, pp. 243-253, 2000.

[15] B. Epshtein, "Determining Document Skew Using Inter-line Spaces," in Document Analysis and Recognition (ICDAR), International Conference on, 2011, pp. 27-31.

[16] D. Kumar and D. Singh, "Modified Approach of Hough Transform for Skew Detection and Correction in Documented Images," International Journal of Research in Computer Science, vol. 2, pp. 37-40, 2012.

[17] A. Hashizume, P.-S. Yeh, and A. Rosenfeld, "A method of detecting the orientation of aligned components," Pattern Recognition Letters, vol. 4, pp. 125-132, 1986.

[18] L. O'Gorman, "The document spectrum for page layout analysis," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 15, pp. 1162-1173, 1993.

[19] X. Jiang, H. Bunke, and D. Widmer-Kljajo, "Skew detection of document images by focused nearest-neighbor clustering," in Document Analysis and Recognition, Proceedings of the Fifth International Conference on, 1999, pp. 629-632.

[20] N. Liolios, N. Fakotakis, and G. Kokkinakis, "Improved document skew detection based on text line connected-component clustering," in Image Processing, Proceedings. International Conference on, 2001, pp. 1098-1101 vol.1.

[21] Y. Lu and C. L. Tan, "A nearest-neighbor chain based approach to skew estimation in document images," Pattern Recognition Letters, vol. 24, pp. 2315-2323, 2003.

[22] Y. Lu and C. L. Tan, "Improved Nearest Neighbor Based Approach to Accurate Document Skew Estimation," presented at the Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 1, 2003.

[23] I. Konya, S. Eickeler, and C. Seibert, "Fast Seamless Skew and Orientation Detection in Document Images," in Pattern Recognition (ICPR), 20th International Conference on, 2010, pp. 1924-1928.

[24] F. Kleber, M. Diem, and R. Sablatnig, "Robust Skew Estimation of Handwritten and Printed Documents Based on Grayvalue Images," presented at the Proceedings of the 22nd International Conference on Pattern Recognition, 2014.

[25] J. Fabrizio, "A precise skew estimation algorithm for document images using KNN clustering and fourier transform," in Image Processing (ICIP), IEEE International Conference on, 2014, pp. 2585-2588.

[26] H. Yan, "Skew Correction of Document Images Using Interline Cross-Correlation," CVGIP: Graphical Models and Image Processing, vol. 55, pp. 538-543, 1993.

[27] A. Chaudhuri and S. Chaudhuri, "Robust detection of skew in document images," Image Processing, IEEE Transactions on, vol. 6, pp. 344-349, 1997.

[28] M. Chen and X. Ding, "A robust skew detection algorithm for grayscale document image," in Document Analysis and Recognition, Proceedings of the Fifth International Conference on, 1999, pp. 617-620.

[29] D. Brodic and Z. N. Milivojevic, "Log-polar Transformation as a Tool for Text Skew Estimation," ELEKTRONIKA IR ELEKTROTECHNIKA, vol. 19, 2013.

[30] A. Alaei, U. Pal, P. Nagabhushan, and F. Kimura, "A Painting Based Technique for Skew Estimation of Scanned Documents," in Document Analysis and Recognition (ICDAR), International Conference on, 2011, pp. 299-303.

[31] C.-H. Chou, S.-Y. Chu, and F. Chang, "Estimation of skew angles for scanned documents based on piecewise covering by parallelograms," Pattern Recognition, vol. 40, pp. 443-455, 2007.

[32] A. A. Mascaro, G. D. C. Cavalcanti, and C. A. B. Mello, "Fast and robust skew estimation of scanned documents through background area information," Pattern Recognition Letters, vol. 31, pp. 1403-1411, 2010.

[33] Y. Ishitani, "Document skew detection based on local region complexity," in Document Analysis and Recognition, Proceedings of the Second International Conference on, 1993, pp. 49-52.

[34] Y.-K. Chen and J.-F. Wang, "Skew detection and reconstruction based on maximization of variance of transition-counts," Pattern Recognition, vol. 33, pp. 195-208, 2000.

[35] S. Chen and R. M. Haralick, "An automatic algorithm for text skew estimation in document images using recursive morphological transforms," in Image Processing, IEEE International Conference, 1994, pp. 139-143 vol.1.

[36] D. S. Guru, M. Ravikumar, and S. Manjunath, "Multiple Skew Estimation in Multilingual Handwritten Documents," International Journal of Computer Science Issues, vol. 10, September 2013.

[37] M. Wagdy, I. Faye, and D. Rohaya, "Document image skew detection and correction method based on extreme points," in Computer and Information Sciences (ICCOINS), International Conference on, 2014, pp. 1-5.

[38] A. Farahmand, A. Sarrafzadeh, and J. Shanbehzadeh, "Document Image Noises and Removal Methods," in Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, 2013, pp. 436-440.

[39] A. S. Reddy, "An Overview on Different Noise Models and Filtering Techniques," International Journal of Computer and Communication System Engineering (IJCCSE), vol. 2, pp. 330-335, 2015.