ADAPTIVE RADIAL PROJECTION ON FOURIER MAGNITUDE SPECTRUM FOR DOCUMENT IMAGE SKEW ESTIMATION

Luan Pham[*]   *Phu Hao Hoang*[*]   *Xuan Toan Mai*[†]   *Tuan Anh Tran*[†]

[*] Cinnamon AI - Department of Research

Floor 10th, No. 36 Hoang Cau street, O Cho Dua Ward, Dong Da District, Hanoi City, Viet Nam.

[†] Faculty of Computer Science & Engineering, Ho Chi Minh City-University of Technology (HCMUT),

268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam

[†] Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam.

phamquiluan@gmail.com, phuhao1998@gmail.com, mxtoan@hcmut.edu.vn, trtanh@hcmut.edu.vn [†]

ABSTRACT

Skew estimation is one of the vital tasks in document processing systems, especially for scanned document images, because its performance impacts subsequent steps directly. Over the years, an enormous number of researches focus on this challenging problem in the rise of digitization age. In this research, we first propose a novel skew estimation method that extracts the dominant skew angle of the given document image by applying an Adaptive Radial Projection on the 2D Discrete Fourier Magnitude spectrum. Second, we introduce a high quality skew estimation dataset DISE-2021 to assess the performance of different estimators. Finally, we provide comprehensive analyses that focus on multiple improvement aspects of Fourier-based methods. Our results show that the proposed method is robust, reliable, and outperforms all compared methods. The source code is available at github.com/phamquiluan/jdeskew.

Index Terms— Document Image Skew Estimation, Fourier Transform-based method, Adaptive Radial Projection

1. INTRODUCTION

In a document processing system, the input image should be ensured to be in a straight position since a graceless skew angle might seriously affect the performance of subsequent steps [1]. Spatial information is still a crucial feature for most deep learning architectures to achieve good performance in various document processing tasks, such as document layout analysis, information extraction, and others. A skew estimation algorithm receives a digital document image and outputs the main skew angle. It is usually coupled with an affine transformation to correct the image.

By 2010, a lot of research had been conducted to tackle this problem, using different schools of thought such as Fourier transform [2, 3], projection profiles analysis [4, 5], Hough transforms [6, 7, 8], and others [9, 10, 11]. Since the dawn of DISEC2013 competition [12], many studies [13, 14, 15] have been continuously carried out to improve the performance of the skew estimation task. The angle range from $-15°$ to $+15°$ and the threshold of $0.1°$ to count the number of correct estimations have been used to compare different estimators.

In this work, we propose a skew estimation method that works robustly over a larger angle range from $-44.9°$ to $44.9°$ degrees. We take advantage of the Fourier-based approach [3, 2, 13] because it can deal with a variety of document images without many assumptions about its kind. The only assumption is that the main skew angle
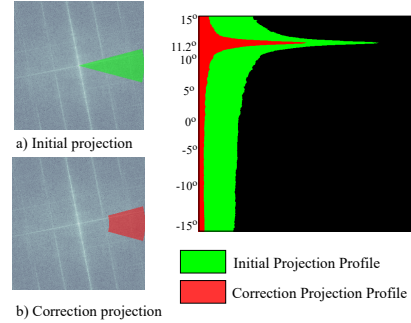

Fig. 1: The overview of our adaptive projection.

is represented on the dominant line of the Fourier Magnitude Spectrum, see Fig. 1. To extract that angle precisely, we introduce a novel Adaptive Radial Projection step. This step is inspired by radial integration in [2, 16, 17], and Radon Transform in [3]. The difference is that we perform radial projection twice, then aggregate their results. The first projection is applied directly on the magnitude from the origin as in [16, 17]. The second projection aims to produce more accurate results by discarding the DC component and low frequencies. The empirical results show our proposed method can extract the skew angle efficiently and its performance surpasses all the compared methods [18, 19, 20, 21, 13].

We also introduce a new dataset called DISE 2021 and propose a way to validate the straightness via the verification mask. Our dataset is aggregated from three different document datasets to ensure its variety: DISEC 2013 [12], RDCL 2017 [22], and RVL-CDIP [14] . We release the DISE 2021 dataset with its straight images, verification masks, and two skew versions of 15 degrees and 44.9 degrees. Previously, we lacked a standard skew dataset in the range of 44.9 degrees and the straightness verification process is still ambiguous in the literature. Our dataset has high quality and is reviewed rigorously by our human annotators.

Additionally, we provide thorough experiments to compare the effectiveness of different factors on the Fourier-based approach under a standard dataset and criteria. These variants are inspired by [3, 2, 13, 23]. To the best of our knowledge, there are no comprehensive analyses on different improvement aspects of this approach in the literature currently.
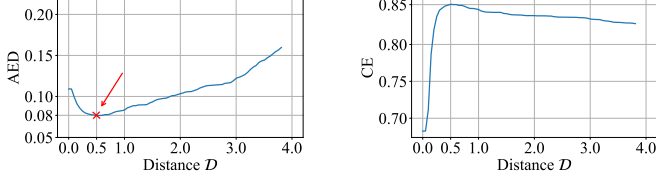
Fig. 2: The change of AED and CE levels when tuning the distance $\mathcal{D}$. The chosen threshold (**x**) should bring the lowest AED value.

2. PROPOSED METHOD

2.1. Overview

The skew estimation method we present in this paper consists of three main steps: preprocessing, 2D Discrete Fourier Transform (2D-DFT), and Adaptive Radial Projection. First, the input image $I \in \mathbb{R}^{H \times W \times C}$, where H, W, C denote its height, width and channel respectively, is subject to essential preprocessing steps that produce a binary image $B \in \{0, 1\}^{H \times W}$. Second, the image B is passed through a 2D-DFT and normalization steps to get its magnitude spectrum $M \in \mathbb{R}^{H \times W}$. Now the oblique angle is clearly shown on the magnitude spectrum. Finally, we apply the Adaptive Radial Projection on the spectrum M to extract the output angle θ precisely.

2.2. Adaptive Radial Projection

This step is composed of two separate radial projections: the initial projection and the correction projection. The initial projection is applied directly on the magnitude from the magnitude center as in [16, 17]. The correction projection is inspired by [2, 3]. However, instead of applying a square-shaped mask [2] or a donut-shaped mask [3], we simply move the left-most element of the integration line away from the magnitude center a distance $\mathcal{W}$, see Fig. 1b. Finally, we aggregate these two outputs in a way that leverages the advantages and mitigates the disadvantages of the second projection as we analyse in Sec. 4.3.3 via ablation experiments on discarding the DC component and low spatial frequencies.

Given the spectrum $M \in \mathbb{R}^{H \times W}$, the projection radius $\mathcal{R} = \min(H, W)$, and the angle list $[\theta_{\min}, \ldots, \theta_{\max}]$, the initial projection value $\mathcal{A}(\theta_i)$ at the angle θ_i is calculated as follows, see Fig. 1a

$$\mathcal{A}(\theta_i) = \sum_{s=0}^{\mathcal{R}} M[c_y + s \cdot \cos(\theta_i), c_x - s \cdot \sin(\theta_i)], \quad (1)$$

where (c_x, c_y) is the center coordinate, and $M[y, x]$ is the value of magnitude spectrum at coordinate (x, y) The correction projection value $\mathcal{B}(\theta_j)$ at angle θ_j is computed as follows, see Fig. 1b.
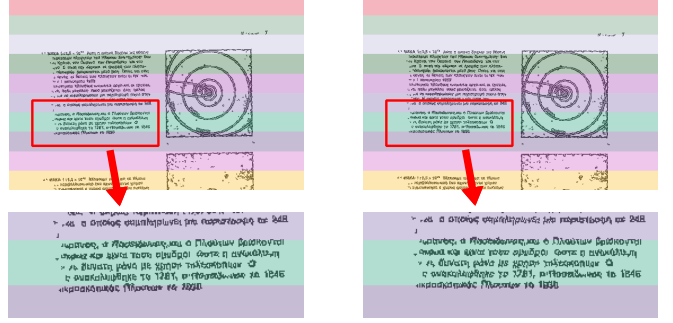
$$\mathcal{B}(\theta_j) = \sum_{s=\mathcal{W}}^{\mathcal{R}} M[c_y + s \cdot \cos(\theta_j), c_x - s \cdot \sin(\theta_j)]. \quad (2)$$

The candidate angles θ_a and θ_b is determined by

$$\theta_a = \underset{m}{\arg\max} \, \mathcal{A}(m), \ \theta_b = \underset{n}{\arg\max} \, \mathcal{B}(n) \quad (3)$$

The final output angle $\theta_{\mathcal{F}}$ is aggregated by the following rule:

$$\theta_{\mathcal{F}} = \begin{cases} \theta_a, \text{if } |\theta_a - \theta_b| > \mathcal{D} \\ \theta_b, \text{otherwise} \end{cases} . \quad (4)$$

Fig. 3: An example of low-quality annotation in DISEC2013, deskewed by DISEC2013 ground-truth angle (left) and our system (right).

The search of parameters. The window size $\mathcal{W}$ and distance $\mathcal{D}$ are chosen as follows. Firstly, we use $\theta_{\mathcal{B}}$ as the default output and search for $\mathcal{W}$ which brings the highest correct estimation rate [12], on the development set. Due to the computational expensive, the coarse search space has 20 elements from 0 to $\frac{1}{3} \times H$. The fine space includes another 20 elements around the coarse result with a deviation of $\frac{2}{30} \times H$. The search space is configured empirically, which could be improved. Secondly, we find $\mathcal{D}$, which brings the optimal Average Error Deviation score on our system. In Fig. 2, at the left-most, $\mathcal{D} = 1$, which means $\theta_{\mathcal{A}}$ is the output angle by default. At the right corner, $\theta_{\mathcal{B}}$ is the output angle without consideration. Within that range, the output angle is chosen as in Eq. (4). The suitable $\mathcal{D}$ minimizes the AED score while possibly bringing the highest CE. The detailed configuration is presented in Sec. 4.2.

3. DATA

3.1. Verification Mask

Although [12] and [24] argued that the skew angle of $0.1°$ might be detectable to a human observer, we see that such a small angle is hard to verify normally and it might lead to unexpected biases [15]. We propose the verification mask to ensure the image is in a straight position during the data review process. Firstly, the images are overlaid with a color mask horizontally or vertically. Secondly, the annotators will draw three to five red boxes that clearly show the alignment of different document components such as text lines, ruling lines, tables, and figures. In this way, we can detect noisy samples easily, see Fig. 3.

3.2. DISE 2021

In this research, we build a new DISE2021 dataset from 95 images from DISEC2013 dataset [12], 70 images from RDCL dataset [22], and 324 images from RVL-CDIP dataset [14]. The composed dataset contains various types of documents, multiple languages, and typography features. Firstly, all the images are ensured and verified to be in a straight position, as described in Sec. 3.1. Secondly, we use the same generating algorithm as in [12] to generate skew images in the range $-15°$ to $+15°$. The dataset is split into two development/test sets by a ratio of $0.7/0.3$ that results in 3399 development images and 1491 testing images. When generating the skew dataset in the range from $-44.9°$ to $44.9°$, we double the augmented image that results in 6980 development images and 2800 testing images. From now on, without clearly specifying DISE 2021 ($44.9°$), the dataset

of $15°$ degrees is used to compare the performance in experiments by default.

At the time of conducting this research, we are unable to obtain the official version of DISEC 2013 dataset [1]. The version we use in this research is obtained by contacting [25]. From manual inspection, we found that there are some samples, which do not fit to our context, where the text orientation is more considerate, see Fig. 3. In 2015, [15] also figured out a similar problem when they got a higher Correct Estimation score by simply subtracting their output with a small offset. From these points, we see that tuning parameters for higher CE in this dataset might lead to wrong conclusions.

4. EXPERIMENT

4.1. Criteria

To evaluate the robustness of skew estimators, we use three standard criteria AED, TOP80, and CE [12]. Nevertheless, we also inspect the sorted absolute error curve to supply a meaningful insight of the error distribution and to observe the worst estimation.

4.2. Threshold Configurations

Image height. Because the size of the document image varies remarkably from a small letter to a big poster or drawing document, it affects the system speed as well as the threshold design. To address this problem, we resize the input image to a pre-defined image height while preserving its aspect ratio. We use five different heights in this research 1024, 1500, 2048, 3072, 4096 for performance comparison in terms of both accuracy and speed.

The range of θ. Since the projection range of our Adaptive Radial Projection is configurable and the angle range is known beforehand, we set the couple $\{\theta_{min}, \theta_{max}\}$ equal to $\{-15, 15\}$ and $\{-44.9, 44.9\}$ when dealing with the corresponding range.

The window size $\mathcal{W}$ and the distance $\mathcal{D}$. After searching, we have 5 sets of parameters $\{H, \mathcal{W}, \mathcal{D}\}$ for each image height config. $\{1024, 247, 0.7\}$, $\{1500, 328, 0.55\}$, $\{2048, 304, 0.55\}$, $\{3072, 307, 0.45\}$, $\{4096, 250, 0.5\}$.

4.3. Analysis of the impact of different factors

In this section, we aim to verify the performance gain potential of the following ideas: (*a*) Split the image by $N \times N$ parts. (*b*) Which is better, between Power spectrum and Magnitude spectrum? (*c*) Discarding the DC component and low spatial frequencies.

For simplicity, we use the baseline estimator as described in Sec. 2 except that it uses $\theta_{\mathcal{A}}$ as default without applying the correction projection result $\theta_{\mathcal{B}}$ to realize experiments in this section.

4.3.1. Image division

This experiment aims to confirm the impact of image division. It was proposed by [2] and [3]. The general idea is instead of directly examining the Magnitude Spectrum of the whole image; we do it on image blocks, then find a way to aggregate their output. In this experiment, we test this idea with multiple block sizes, which are proportional to the image height H. Firstly, we tile the input image to multiple parts of block size $N \times N$ and convert each part to the frequencies domain. Secondly, we normalize the magnitude spectrum of each block and average them. Finally, we use radial

---

Table 1: Effects of image division with different block size.

Block Size	1024			2048		
	AED	TOP80	CE	AED	TOP80	CE
$N = h * 0.1$	0.95	0.75	0.01	0.91	0.55	0.03
$N = h * 0.2$	0.56	0.43	0.03	0.51	0.33	0.02
…	…	…	…	…	…	…
$N = h * 0.9$	0.16	0.11	0.39	0.15	0.08	0.57
$N = h$	0.14	0.09	**0.47**	0.12	0.07	**0.62**

Table 2: Comparison of Magnitude and Power Spectrum

	1024			2048		
	AED	TOP80	CE	AED	TOP80	CE
Magnitude Spectrum	0.14	0.09	**0.47**	0.12	0.07	**0.62**
Power Spectrum	0.2	0.14	0.28	0.19	0.15	0.35

projection analysis to extract the angle from the combined magnitude spectrum. Table 1 indicates that the level of efficiency steadily decreases as we split the image into smaller parts. For instance, the CE scores drop from 0.47 to 0.01 and from 0.62 to 0.03 in the image heights of 1024 and 2048 respectively. We see that there are two main reasons for the deficiency: the outlier blocks and the smallness of magnitude spectrum affect the overall result.

Obviously, the CE only counts if and only if $E(i) \leq 0.1$ [12]. Geometrically, $\frac{x}{y} = \tan(\theta)$, so to show 1 pixel in x, we need $\frac{1}{\tan(0.1)} \approx 573$ pixels in y. It means the square image must have the size of $573 \times 2 = 1146$ pixels. This not only explains the plummet but also the climb of performance when we increase the image height.

4.3.2. Power spectrum and magnitude spectrum

While [13] leverages Magnitude spectrum M, [16, 17] use Power Spectrum $\mathcal{P}$, where $\mathcal{P}(u, v) = M(u, v)^2$. In this work, we also conduct the quantitative comparison of these twos when using the radial projection to extract the skew angle from the spectrum. In Tab. 2, it can be clearly seen the magnitude spectrum can reach remarkably higher performance in terms of evaluation score. While the AED of magnitude spectrum trials around 0.12 to 0.14, the power spectrum results fluctuate from 0.19 to 0.2. The Correct Estimation rate between these two experiments also has a large margin. In experiments of image height 1024 and 2048, the magnitude spectrum gets 47% and 62% correct rate while the opponent gets only 28% and 35% accordingly.

#### 4.3.3. Discarding the DC component and small frequencies

The effectiveness of removing the DC component and small frequencies in the Fourier spectrum is investigated in this experiment. Previously, [2] uses a squared zero window, while [3] uses a round-shaped zero mask to override the spectrum center. In this work, we simply move the start point of the radial projection far from the spectrum center a distance $\mathcal{W}$, see Fig. 1b. The results of the experiment are shown in Tab. 3. On the one hand, significant improvement in the correct estimation rate of all image heights is clearly shown. There is an increase from 47% to 65% percent in the image height of 1024, 58 to 74, and 62 to 78 in the image heights of 1500 and 2048 respectively. We conclude that it is possible to reach a surprisingly high correct estimation rate by having a right tuned $\mathcal{W}$.

On the other hand, it is noticeable that while we are expecting the AED to maintain a negative correlation with CE, which means higher CE should couple with lower AED, but it turns out that this

Table 3: Efficiency comparison when discarding the DC component and small frequencies by different window size on our baseline.

Window Size	1024			3072		
	AED	TOP80	CE	AED	TOP80	CE
$W = 15$	0.14	0.09	0.47	0.1	0.06	0.68
$W = 35$	0.13	0.09	0.49	0.1	0.06	0.69
$W = 55$	0.13	0.09	0.51	0.11	0.06	0.7
$W = 75$	0.13	0.09	0.53	0.11	0.06	0.72
$W = h/10$	0.13	0.08	0.55	0.74	0.04	**0.81**
$W = h/8$	0.14	0.08	0.57	1.23	0.04	0.77
$W = h/6$	0.14	0.07	0.6	2.19	0.36	0.65
$W = h/4$	0.18	0.07	**0.65**	5.03	2.9	0.32
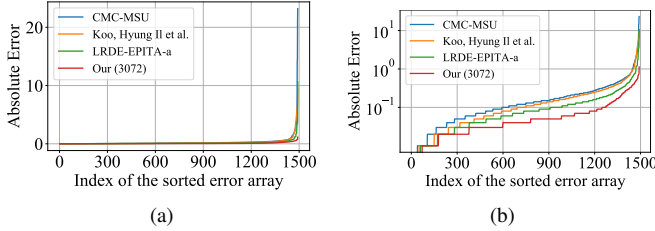
Fig. 4: (a) The sorted absolute error curve and (b) its log-scale version of four different methods on DISE 2021 ($15°$) dataset.

is not true in this experiment. The average error goes up with the correct estimation rate. This phenomenon shows over $10°$ worst estimation at the tail of the sorted error curve range from $5°$ to $15°$, causing the AED to go up dramatically. In case of image height 3072, even though the best 80 percent of estimation results are able to achieve 0.04 in TOP80 score, the total average error is up to a surprisingly high number of 0.74. Inspecting the spectrum magnitude, the zero mask in our expectation could help to remove the noisy signals, but in bad cases, it might also remove the dominant "line" in the magnitude spectrum. Hence, this causes difficulties for the later radial projection to get the right results. From the industrial point of view, this system might ruin the user experience when it gives completely wrong estimation.

To increase in CE without bringing the negative effect to AED, we perform radial projection twice as described in Sec. 2.2, the second angle $\theta_\mathcal{B}$ is chosen if it does not differ from $\theta_\mathcal{A}$ greater than a distance $\mathcal{D}$. By this way, we gain improvements not only in the significant reduction in AED but also in the increase in CE rate. For example, the best AED goes from 0.74 down to one-tenth and the CE increases from 0.81 to 0.86 in the image height of 3072 pixels, see Table Tab. 4.

4.4. Experimental Results

Table 4 illustrates the evaluation scores of different methods on all three datasets: DISEC 2013, DISE 2021 ($15°$), and DISE 2021 ($49.9°$). The proposed method outperforms all other methods on two DISE 2021 datasets. The configuration of $\{H, \mathcal{W}, \mathcal{D}\} = \{3072, 307, 0.45\}$ brings the best scores of $0.07, 0.04, 0.86$ on the DISE 2021 ($15°$) dataset and $0.06, 0.02, 0.88$ on the DISE 2021 ($44.9°$) dataset in the AED, TOP80, and CE criteria respectively. While the current solutions of CMC-MSU [21, 12] and LRDE-EPITA-a [13, 12] do not support the range of $44.9°$, [20] shows competitive performance on this range. [18] and [19] show a reasonable performance on the DISE 2021 ($15°$) dataset when they reach 0.24 and 0.09 respectively in the TOP80 criterion, but their AED scores are too high compared to others.

Table 4: Quantitative comparison of different methods on the DISE2021 $15°$ and $44.9°$ datasets.

Method	DISE 2021				DISE 2021 (45)			
	AED	TOP80	CE	WE	AED	TOP80	CE	WE
Our (2048)	0.08	0.04	0.84	1.13	0.06	0.03	0.87	1.06
Our (3072)	0.07	0.04	**0.86**	1.13	0.06	0.02	**0.88**	1.06
Our (4096)	0.08	0.04	0.83	1.18	0.06	0.03	0.86	1.06
FredsDeskew	10.82	0.09	0.54	109	12.44	0.1	0.51	110
PypiDeskew	16.59	0.24	0.2	141	21.79	2.51	0.14	179
Koo, Hyung Il \etal	0.22	0.09	0.48	9.43	0.24	0.09	0.48	82.37
CMC-MSU	0.27	0.11	0.43	23.2	-	-	-	-
LRDE-EPITA-a	0.14	0.06	0.66	10.61	-	-	-	-

In the matter of the Worst Error (WE), our method maintains a low level around $1°$ over all configurations. While the other approaches go from around $10°$ to over $20°$, the implementation of PypiDeskew and FredsDeskew have the Worst Error over 100 degrees. In addition, figure Fig. 4 presents the sorted absolute error curve of four different methods. By examining this curve, we see that all four methods perform excellently over 95% of cases. The difference reveals in the last 5% of worst cases. The log-scale version of the error curve, in Figure Fig. 4b, shows that our method consistently achieve better compared to the other three solutions. This insight is worthy to verify the reliability when comparing different methods on a skew estimation dataset, especially when handling hard cases.

The speed of our method. The running time of our method is around 1 second per image in the image height of $1024, 1500, 2048$ on the single-threaded implementation. The multi-threaded implementation in our machine [1] brings the throughput up to nearly 37 images per second in image height 1024. Stahlberg et al. [15] system has got the best throughput of 5.3 documents per minute. The LRDE-EPITA-a solution [13] can process an image in around 7 seconds, our method beats this solution in both terms of accuracy and speed.

5. SUMMARY

We have presented a novel method, which involves the Fourier transform and an Adaptive Radial Projection, to robustly solve the skew estimation problem in document image. Our method can extract the dominant oblique angle precisely, and independently from different languages, document types and structures. We have also created a new high quality dataset to benchmark the skew estimators. Finally, various experiments and detailed analyses have been contributed to the literature. Future works might include speed enhancements toward a real-time estimator and the evaluation of other system performance such as table extraction, OCR system, and information extraction when integrating this skew estimation method into their document processing pipeline.

---

[1] https://aws.amazon.com/ec2/instance-types/c5/

6. REFERENCES

[1] Tuan Anh Tran, Kanghan Oh, In-Seop Na, Guee-Sang Lee, Hyung-Jeong Yang, and Soo-Hyung Kim, "A robust system for document layout analysis using multilevel homogeneity structure," *Expert Systems With Applications*, vol. 85, no. 1, pp. 99–113, 2017. 1

[2] GS Peake and TN Tan, "A general algorithm for document skew angle estimation," in *International Conference on Image Processing*. IEEE, 1997, vol. 2, pp. 230–233. 1, 2, 3

[3] Scott Lowther, Vinod Chandran, and Subramanian Sridharan, "An accurate method for skew determination in document images," in *Digital Image Computing Techniques and Applications*, 2002, vol. 1, pp. 25–29. 1, 2, 3

[4] G Ciardiello, G Scafuro, MT Degrandi, MR Spada, and MP Roccotelli, "An experimental system for office document handling and text recognition," in *International Conference on Pattern Recognition*, 1988, pp. 739–743. 1

[5] A Bagdanov and Junichi Kanai, "Projection profile based skew estimation algorithm for jbig compressed images," in *International Conference on Document Analysis and Recognition*. IEEE, 1997, vol. 1, pp. 401–405. 1

[6] Sargur N Srihari and Venugopal Govindaraju, "Analysis of textual images using the hough transform," *Machine vision and Applications*, vol. 2, no. 3, pp. 141–153, 1989. 1

[7] Stuart C Hinds, James L Fisher, and Donald P D'Amato, "A document skew detection method using run-length encoding and the hough transform," in *International Conference on Pattern Recognition*. IEEE, 1990, vol. 1, pp. 464–468. 1

[8] Daniel S Le, George R Thoma, and Harry Wechsler, "Automated page orientation and skew angle detection for binary document images," *Pattern Recognition*, vol. 27, no. 10, pp. 1325–1344, 1994. 1

[9] BV Dhandra, VS Malemath, H Mallikarjun, and Ravindra Hegadi, "Skew detection in binary image documents based on image dilation and region labeling approach," in *International Conference on Pattern Recognition*. IEEE, 2006, vol. 2, pp. 954–957. 1

[10] Ray Smith, "A simple and efficient skew detection algorithm via text row accumulation," in *International Conference on Document Analysis and Recognition*. IEEE, 1995, vol. 2, pp. 1145–1148. 1

[11] Su Chen, Robert M Haralick, and Ihsin T Phillips, "Automatic text skew estimation in document images," in *International Conference on Document Analysis and Recognition*. IEEE, 1995, vol. 2, pp. 1153–1156. 1

[12] Alexandros Papandreou, Basilios Gatos, Georgios Louloudis, and Nikolaos Stamatopoulos, "Icdar 2013 document image skew estimation contest (disec 2013)," in *International Conference on Document Analysis and Recognition*. IEEE, 2013, pp. 1444–1448. 1, 2, 3, 4

[13] Jonathan Fabrizio, "A precise skew estimation algorithm for document images using knn clustering and fourier transform," in *International Conference on Image Processing*. IEEE, 2014, pp. 2585–2588. 1, 3, 4

[14] Adam W Harley, Alex Ufkes, and Konstantinos G Derpanis, "Evaluation of deep convolutional nets for document image classification and retrieval," in *International Conference on Document Analysis and Recognition*. IEEE, 2015. 1, 2

[15] Felix Stahlberg and Stephan Vogel, "Document skew detection based on hough space derivatives," in *International Conference on Document Analysis and Recognition*. IEEE, 2015, pp. 366–370. 1, 2, 3, 4

[16] Wolfgang Postl, "Detection of linear oblique structures and skew scan in digitized documents," IEEE, 1986, pp. 687–689. 1, 2, 3

[17] Wolfgang Postl, "Method for automatic correction of character skew in the acquisition of a text original in the form of digital scan results," Feb. 2 1988, US Patent 4,723,297. 1, 2, 3

[18] Stéphane Brunner, "A python library used to deskew a scanned document.," https://github.com/sbrunner/deskew, 2017. 1, 4

[19] Fred Weinhaus, "Fred's imagemagick textdeskew script.," http://www.fmwconcepts.com/imagemagick/textdeskew/index.php, 2017. 1, 4

[20] Hyung Il Koo and Nam Ik Cho, "Robust skew estimation using straight lines in document images," *Journal of Electronic Imaging*, vol. 25, no. 3, pp. 033014, 2016. 1, 4

[21] Oleg Naydin, "The solution of cmc-msu team in disec 2013 competition.," https://github.com/OLEGator30/DISEC-13, 2013. 1, 4

[22] Christian Clausner, Apostolos Antonacopoulos, and Stefan Pletschacher, "Icdar2017 competition on recognition of documents with complex layouts-rdcl2017," in *International Conference on Document Analysis and Recognition*. IEEE, 2017, vol. 1, pp. 1404–1410. 1, 2

[23] Tuan Anh Tran, Hong Tai Tran, In Seop Na, Guee Sang Lee, Hyung Jeong Yang, and Soo Hyung Kim, "A mixture model using random rotation bounding box to detect table region in document image," *Journal of Visual Communication and Image Representation*, vol. 39, no. 1, pp. 196–208, 2016. 1

[24] Jonathan J Hull, "Document image skew detection: Survey and annotated bibliography," in *Document Analysis Systems II*, pp. 40–64. World Scientific, 1998. 2

[25] PV Bezmaternykh and Dmitry P Nikolaev, "A document skew detection method using fast hough transform," in *International Conference on Machine Vision*. SPIE, 2020, vol. 11433, p. 114330J. 3