1-19-2010

2x1D Image Registration and Comparison

Geng Zheng
Boise State University

Elisa H. Barney Smith
Boise State University

Nader Rafla
Boise State University

Tim Andersen
Boise State University

 DOI: 10.1117/12.839931

2x1D Image Registration and Comparison

Geng Zheng, Elisa H. Barney Smith,* Nader Rafla and Tim Andersen

Boise State University, Boise, Idaho, USA

ABSTRACT

This paper presents a novel *2×1D phase correlation* based image registration method for verification of printer emulator output. The method combines the basic phase correlation technique and a modified 2x1D version of it to achieve both high speed and high accuracy. The proposed method has been implemented and tested using images generated by printer emulators. Over 97% of the image pairs were registered correctly, accurately dealing with diverse images with large translations and image cropping.

Keywords: Image registration, image comparison, processing accelerations

1. INTRODUCTION

The traditional method used to develop and verify printer design is to construct a physical printer (prototype) and use it to print thousands of test pages, which are then manually analyzed for defects. To improve quality, speed time to market, and reduce cost, printer manufacturers have begun using print engine emulators for much of the design validation. Emulators work by digitally capturing the output of the printer, bypassing the physical printing stage and capturing the laser trace intended for the print drum to create a digital image. We propose to automate validation of printer emulator output using a phase correlation based method that can detect two types of image misalignments: (1) orientational and (2) translational, under high speed and high accuracy constraints.

Section 2 introduces the method of image registration we use and discusses its relation to previous work. Section 3 describes the experiments. Further analysis and possible improvements are discussed in Section 4. Section 5 concludes the paper.

2. IMAGE REGISTRATION METHODS

The two broad classes of image registration methods are:[1] (1) Feature-based methods which identify edges, corners, contours, or other features common to the images; (2) Area-based methods which match overall images "blindly" by measuring their similarity (usually the overall intensity change of the image pairs).[2] Feature-based methods are suitable when common features are present in both images. But printer validation requires testing a variety of images, making it hard to design a good set of features, therefore we focused on area-based image registration.

2.1 Phase Correlation

One approach for image registration maximizes the correlation between image pairs using normalized cross-correlation, which can be calculated directly in the spatial domain. To make the similarity measure independent of uniform changes of image intensities, normalized cross-correlation is usually calculated.

Given two images f_1 and f_2, the 2D normalized cross-correlation of them is defined as

$$c = \frac{\sum_x \sum_y (f_1(x,y) - \mu_1)(f_2(x,y) - \mu_2)}{\sqrt{\sum_x \sum_y (f_1(x,y) - \mu_1)^2 \sum_x \sum_y (f_2(x,y) - \mu_2)^2}}, \tag{1}$$

where μ_1 and μ_2 are the means of f_1 and f_2, respectively. If f_1 and f_2 only differ by a displacement $(\Delta x, \Delta y)$, i.e.,

$$f_2(x,y) = f_1(x + \Delta x, y + \Delta y), \tag{2}$$

* Send correspondence to Elisa Barney Smith. E-mail: EBarneySmith@BoiseState.edu

Image Processing: Machine Vision Applications III, edited by David Fofi, Kurt S. Niel,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7538, 75380V · © 2010 SPIE-IS&T
CCC code: 0277-786X/10/$18 · doi: 10.1117/12.839931

the cross-correlation of f_1 and f_2 is

$$c(\Delta x, \Delta y) = \frac{\sum_x \sum_y (f_1(x,y) - \mu_1)(f_2(x - \Delta x, y - \Delta y) - \mu_2)}{\sqrt{\sum_x \sum_y (f_1(x,y) - \mu_1)^2 \sum_x \sum_y (f_2(x - \Delta x, y - \Delta y) - \mu_2)^2}}. \tag{3}$$

To determine the relative translation the normalized cross-correlation (Equation 3) must typically be calculated for all possible displacements $(\Delta x, \Delta x)$. However, the Fourier transform of the correlation between two images is the product of the Fourier transform of one image and the complex conjugate of the Fourier transform of the other image. Thus, computational cost can be reduced by computing the correlation in the frequency domain. This phase correlation method was first proposed by Kuglin and Hines in 1975.[3] Since then, several related methods have been proposed.[2,4–6] The basic phase correlation estimates translational misalignment. If f_1 and f_2 are two images, then F_1 and F_2 are their Fourier transforms, respectively. If $f_1(x,y)$ and $f_2(x,y)$ only differ by a displacement $(\Delta x, \Delta y)$, their Fourier transforms are related by

$$F_2(\omega_x, \omega_y) = F_1(\omega_x, \omega_y) \cdot e^{-j(\omega_x \Delta x + \omega_y \Delta y)}. \tag{4}$$

So, in the frequency domain, the images have the same magnitude, differing by a phase shift related to their relative translation in the spatial domain. The normalized cross-power spectrum of the images is defined as

$$C(\omega_x, \omega_y) = \frac{F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)}{|F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)|} = e^{j(\omega_x \Delta x + \omega_y \Delta y)}, \tag{5}$$

where F_2^* is the complex conjugate of F_2. The correlation of the images is obtained by taking the inverse Fourier transform of their cross-power spectrum:

$$c(x,y) = \mathcal{F}^{-1}\{C(\omega_x, \omega_y)\} = \delta(x + \Delta x, y + \Delta y). \tag{6}$$

The function $c(x,y)$ is an impulse function which has a peak at $(-\Delta x, -\Delta y)$. The magnitude of the peak is the correlation of the images. Thus the relative translation can be found by locating the peak.

2.2 Orientation Identification with Phase Correlation

Orientational misalignment is the special case of rotational misalignment when the relative rotation of two images is 90°, 180°, or 270°. While the basic phase correlation technique can only identify translational misalignment, the Fourier-Mellin transform can be used to detect rotation and scaling misalignment.[6] However, this is computationally expensive. Other methods[7–11] have been developed to estimate rotational misalignments, but they require the images to be predominantly text. In this research because the orientation misalignment is not from manual misplacement of a physical piece of paper, but rather from possible algorithmic misdesign in the printer firmware, a simple but effective method is used instead. For a given image pair, the target image is rotated by 0°, 90°, 180°, and 270°. The correlation at each possible orientation is measured and the highest correlation is selected.

2.3 2D Image Registration Using 2×1D Phase Correlation

While computing the phase correlation is faster than computing cross-correlation in the spatial domain, it is still slow. Computing the phase correlation for four orientations takes more than 7 seconds for 1000×900 pixel images. To reduce this we developed a modified phase correlation technique *2×1D Phase Correlation.* This creates two 1D images from the projection of the reference image along the x-direction and y-direction, called R_X and R_Y respectively. Similarly, two 1D images, T_X and T_Y, are created from the target image. 2×1D phase correlation uses standard phase correlation to compute the correlation between the 1D image pairs. The correlation between R_X and T_X determines translation in the x-direction, while the R_Y and T_Y correlation determines translation in y.

A disadvantage of 2×1D phase correlation is that due to smoothing during creation of the projection, the intensity variation of the 1D image projection can be very small, increasing the likelihood of error. To compensate for this, instead of only locating the highest peak, the highest N correlation peaks in each x and y dimension are recorded. Then, at the corresponding N^2 translation combinations the 2D spatial correlation is calculated to find the most likely translation.

3. EXPERIMENTAL SETUP AND RESULTS

In this paper, two printer emulators, a reference and target, were used to generate a variety of grey-scale images for testing. Each emulator was used to generate 1136 images. 49% of the 1136 images were predominantly text; 13% contained natural scenes or human faces; 10% were line drawings; and 28% were a mixture of image types. The reference images were 1011×850 pixels, and the target images were originally 817×1100 pixels, but were rotated and scaled to have the same orientation and size as the reference images, so the only misalignment was that added for testing purposes. All of the tests were run on an Intel Pentium D 3 GHz CPU with 2 GB (3 GHz) RAM.

3.1 Translation Estimation

Varying degrees of translation were tested. These are grouped into six categories which are defined as:

- T_1: small translations of 1, 2, 3, 4, 5 pixels in only one (x or y) dimension.
- T_2: medium translations of 5, 10, 15, 20 pixels in only one dimension.
- T_3: large translations of 30, 40, 50, 100, 200, 450 pixels in only one dimension.
- T_4: small translations of 1, 2, 3, 4, 5 pixels in both x and y dimensions (all permutations).
- T_5: medium translations of 5, 10, 15, 20 pixels in both x and y dimensions (all permutations).
- T_6: large translations of 30, 40, 50, 100, 200, 450 pixels in both x and y dimensions (all permutations).

The target image in each image pair was subjected to each translation in each category, yielding 1136 x {10, 8, 11, 25, 16, 30} pairs tested. Translation estimation was evaluated for 2×1D phase correlation and the traditional 2D phase correlation. Since speed was a factor being evaluated, 2D phase correlation with downsampling was also tested. Table 1 shows the means and standard deviations of the estimation errors. All methods have high accuracy with a standard deviation of only a few pixels if the translation amount is small or medium (i.e., T_1, T_2, T_4, T_5). In cases where translation is large (T_3, T_6) the error still averages just a few pixels, but the standard deviation is higher. For large translations, the error is higher for 2×1D phase correlation than 2D phase correlation. This is because the pixels that are translated off the page affect the sum of those that remain for the 1D case. So, the more expensive (computationally) 2D phase correlation should be used to register images with large translational misalignment. But for small or medium translations, 2×1D phase correlation performs comparably well and is much faster. Thus, determining which approach to use for a given image pair is needed, and a method to do this is proposed in Section 4.

3.2 Orientation Identification

To identify image orientation, the algorithm first computes the correlation between the reference image and target image with 0° rotation. The target images are then rotated by 90°, 180°, 270° and the correlations computed for each orientation. Experiments indicate that the larger the difference between the correlation scores at the 4 rotations, the more likely the algorithm can identify the correct orientation. To test the robustness of the algorithms at identifying the correct orientation, various versions of the target images were created by translating them in the magnitude groupings described in Section 3.1. Both the introduced orientation and translation modifications could cause the target images to be cropped.

The correlation results for correct versus different orientation are shown graphically in Figure 1. Figure 1a shows the distribution of the correlation computed by the 2D phase correlation. Figures 1b and c show the distribution of the correlation computed by the 2×1D phase correlation in the x- and y- direction. The red lines in these figures represent the correlation when the reference and target images have the same orientation. This correlation was usually high, but had a wide range of values. The shaded lines represent the largest correlation of the other three orientations. The separation is greater when using 2D correlation, but the separation for 2×1D phase correlation is still significant.

Table 1. Mean and Standard Deviation of Translation Estimation Error. μ = Mean σ = Standard Deviation

		Phase Correlation			
		2D	2D with downsampling	2×1D (N=5)	2×1D (N=3)
T_1	μ	0.111	0.797	0.107	0.107
	σ	2.245	2.186	2.085	2.085
T_2	μ	0.149	0.911	0.168	0.371
	σ	2.425	3.007	2.649	6.594
T_3	μ	0.564	1.579	1.161	1.500
	σ	10.922	11.051	19.200	20.732
T_4	μ	0.123	0.511	0.119	0.119
	σ	2.265	2.158	2.106	2.106
T_5	μ	0.198	0.701	0.236	0.540
	σ	2.609	2.894	2.990	8.303
T_6	μ	1.015	1.990	2.211	2.404
	σ	14.980	14.821	25.559	26.089

Table 2 shows the statistics for the correlations computed by the 2×1D phase correlation and 2D phase correlation algorithms. The column 'Orientation: Same' gives the correlations of the image pairs having the same orientation. The column 'Orientation: Others' gives the highest correlations from the other three orientations. The column labeled Δ Correlation shows the difference between the correlation of image pairs with the same orientation, and the maximum of the correlations with the alternate orientations. The larger the differences, the more robust the algorithm is. As expected, both algorithms are better at identifying rotation if translation misalignment is small (T_1 and T_4). 2×1D phase correlation performs well on images that have large intensity variations, such as high contrast picture images and text only images. Low contrast images, however, may not be registered correctly due to relatively small intensity variations.

Table 3 shows the percentage of incorrect orientation identifications made by the phase correlation algorithms. All had less than 1% error on small T and for large T all were less than 4%. The computationally expensive 2D phase correlation algorithm made the fewest incorrect identifications. 2D phase correlation with downsampling also made fewer incorrect identifications compared to the 2×1D phase correlation algorithm, but since the x and y correlation will be used in conjunction, their combined error rate is lower than the individual rates. For all cases the percentage of incorrect identifications made by both 2D phase correlation algorithms are below 1%. The percentage of incorrect identifications of the 2×1D phase correlation algorithm is high. The percentage of incorrect identification of registering the 1D image projected to the x axis is larger than the one of the 1D image projected to the y axis, because the size of the image in the y axis is smaller than the one in the x axis and thus the 1D image along y axis contains more information.

4. ANALYSIS AND FURTHER STUDIES

The faster 2×1D phase correlation method is not as accurate as 2D phase correlation for large translation misalignments. However, it is possible to approach the speed of 2×1D phase correlation and the accuracy of 2D phase correlation by using 2D phase correlation when 2×1D phase correlation appears to be failing. With this approach, the decision of 2×1D phase correlation is first checked, and if a quality criteria is not met then 2D phase correlation is used. Three possible criteria are explored in our experiments:

- Criterion #1: always use 2×1D phase correlation (baseline).
- Criterion #2: use 2×1D phase correlation unless both the 1D phase correlations along the x-direction and the y-direction indicate different orientations.
- Criterion #3: use 2×1D phase correlation unless both the 1D phase correlations along the x-direction and the y-direction indicate different orientations or when the measured correlations indicate the same orientation but the correlations between the orientations differ by less than 0.1.

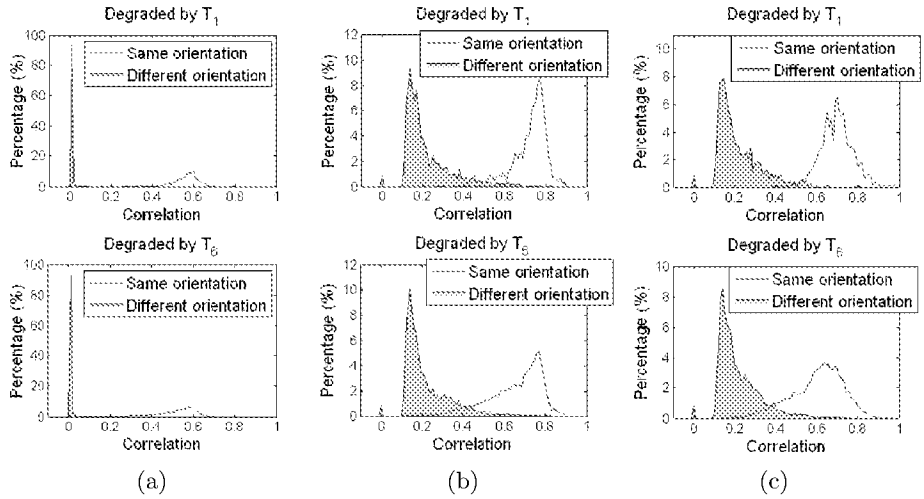

Figure 1. Histograms of correlations for same and different orientations for cases T_1 and T_6. (a) 2D phase correlation (b) 2×1D phase correlation -x, and (c) 2×1D phase correlation -y.

The test results for each criteria are shown in Table 4. Here $P_{2\times1D}$ is the percentage of image pairs that satisfy the criteria and do not pass on testing to 2D phase correlation. P_{err} stands for the percentage of incorrect identifications caused by using 2×1D phase correlation. For criteria #2 & #3, over 90% of the image pairs are registered using the faster 2×1D phase correlation and the percentage of error is below 1%.

Criterion #3 caused the fewest incorrect identifications, and for small translation, T_1 and T_4, over 95% of the image pairs were registered using 2×1D phase correlation. This results in 1.5 hours savings on 1000 images. Even with larger translation less than 10% of images are registered using 2D phase correlation, still saving nearly 1.5 hours on 1000 images. If higher accuracy (more than 99.5%) is required on images with large translations, this can be accomodated with criteria #3 by increasing the correlation difference threshold.

The 2D phase correlation accurately detects orientation alignment (0°) for all four image types as shown in Figure 2. The 2×1D phase correlation algorithm identifies the orientation of text images in the x-direction (the longer side) very well. This is because the text lines of these images are perpendicular to the x-direction and thus the 1D signals in the x-direction exhibit larger intensity variation. When dealing with picture or graphic images, however, registering the 1D signals in the y-direction is more likely to return a correct identification since each term of the 1D signal in the y direction contains more image information than in the x-direction.

The results shown in Figure 2 indicate that images with predominantly text content are more suitable for automated testing of printer emulators for orientation errors. If the 2×1D phase correlation algorithm is used, the results corresponding to the 1D signals perpendicular to text lines are preferred.

5. CONCLUSION

A new phase correlation based image registration method that is computationally fast as well as accurate was presented. The standard 2D phase correlation is very accurate, but considerably slow. By projecting the image data along the axes, the 2D FFT can be reduced to two 1D FFTs. The summation can cause some information to be averaged away, and the effects of image cropping may corrupt the projection signal. This can be compensated for by evaluating the 2×1D phase correlation first, and if the accuracy in the two projections is not adequate, then falling back to 2D phase correlation.

ACKNOWLEDGMENTS

The work presented in this paper was supported by a grant from Hewlett Packard Corporation. The authors wish to thank them for this support.

Table 2. Correlations computed by 2D and 2×1D Phase Correlation algorithms for orientation and translation
μ = Mean, σ = Standard Deviation

		2D Phase Correlation						2×1D Phase Correlation					
		without Downsampling			with Downsampling			X			Y		
		Orientation		Δ	Orientation		Δ	Orientation		Δ	Orientation		Δ
		Same	Others	Corr.	Same	Others	Corr.	Same	Others	Corr.	Same	Others	Corr.
T_1	min	0	0	0	0	0	-0.011	0.347	0	-0.136	0	0	-0.087
	max	1	0.290	1	1	0.635	1	0.987	0.807	0.737	0.892	0.817	0.715
	μ	0.562	0.010	0.553	0.522	0.017	0.505	0.697	0.224	0.471	0.721	0.223	0.498
	σ	0.078	0.011	0.078	0.106	0.019	0.108	0.081	0.108	0.101	0.079	0.119	0.156
T_2	min	0	0	0	0	0	-0.067	0.193	0	-0.316	0	0	-0.118
	max	1	0.290	1	1	0.635	1	0.987	0.807	0.737	0.892	0.817	0.716
	μ	0.558	0.010	0.548	0.500	0.017	0.483	0.663	0.221	0.440	0.719	0.222	0.497
	σ	0.081	0.010	0.080	0.107	0.017	0.108	0.100	0.108	0.146	0.103	0.116	0.157
T_3	min	0	0	0	0	0	-0.082	0.129	0	-0.323	0	0	-0.285
	max	1	0.290	1	1	0.434	1	0.987	0.808	0.737	0.892	0.816	0.716
	μ	0.534	0.010	0.525	0.407	0.017	0.389	0.622	0.219	0.401	0.685	0.220	0.465
	σ	0.099	0.010	0.099	0.065	0.016	0.067	0.125	0.105	0.161	0.124	0.113	0.168
T_4	min	0	0	0	0	0	-0.010	0	0	-0.136	0	0	-0.089
	max	1	0.290	1	1	0.635	1	0.987	0.807	0.737	0.892	0.816	0.715
	μ	0.561	0.010	0.552	0.633	0.017	0.616	0.692	0.224	0.467	0.721	0.223	0.498
	σ	0.079	0.011	0.079	0.142	0.019	0.144	0.102	0.108	0.140	0.101	0.118	0.156
T_5	min	0	0	0	0	0	-0.068	0	0	-0.315	0	0	-0.118
	max	0.791	0.289	0.903	0.628	0.494	0.864	0.983	0.804	0.737	0.892	0.817	0.716
	μ	0.553	0.010	0.543	0.586	0.017	0.569	0.630	0.219	0.411	0.717	0.221	0.496
	σ	0.083	0.010	0.083	0.146	0.016	0.148	0.116	0.109	0.156	0.105	0.114	0.158
T_6	min	0	0	-0.001	0	0	-0.049	0	0	-0.370	0	0	-0.448
	max	0.741	0.273	0.682	0.717	0.432	0.633	0.974	0.804	0.743	0.893	0.816	0.716
	μ	0.512	0.010	0.502	0.398	0.017	0.381	0.597	0.221	0.377	0.652	0.218	0.434
	σ	0.110	0.010	0.109	0.066	0.015	0.067	0.138	0.105	0.172	0.138	0.109	0.179

Table 3. Percentage of Incorrect Orientation Identification

	Phase Correlation			
	2D	2D with downsampling	2×1D X	2×1D Y
T_1	0.079	0.114	0.484	0.968
T_2	0.396	0.451	1.496	1.133
T_3	0.432	0.520	2.225	1.512
T_4	0.159	0.173	0.511	0.968
T_5	0.792	0.825	2.465	1.287
T_6	0.795	0.860	3.539	2.409

REFERENCES

[1] Zitova, B. and Flusser, J., "Image registration methods: A survey," *Image and Vision Computing* **21**, 977–1000 (October 2003).

[2] Stone, H. and Wolpov., R., "Blind cross-spectral image registration using prefiltering and fourier-based translation detection," *IEEE Transactions on Geoscience and Remote Sensing* **40**, 637–650 (March 2002).

[3] Brown, L., "A survey of image registration techniques," *ACM Computing Surveys* **24**, 325–376 (December 1992).

[4] Averbuch, A. and Keller, Y., "Fft based image registration," in [*Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on*], 3608–3611, IEEE (2002).

Table 4. Criteria Evaluation
$P_{2\times1D}$ = Percentage of the Registrations Using 2×1D Phase Correlation
P_{err} = Percentage of the Incorrect Identification Caused by Using the 2×1D Phase Correlation

		Criterion #1	Criterion #2	Criterion #3
T_1	$P_{2\times1D}$	100	98.55	95.90
	P_{err}	0.66	0	0
T_2	$P_{2\times1D}$	100	97.59	94.81
	P_{err}	1.73	0.11	0
T_3	$P_{2\times1D}$	100	96.60	92.40
	P_{err}	2.78	0.17	0
T_4	$P_{2\times1D}$	100	98.52	95.93
	P_{err}	0.69	0	0
T_5	$P_{2\times1D}$	100	96.68	93.68
	P_{err}	2.75	0.22	0
T_6	$P_{2\times1D}$	100	95.24	90.12
	P_{err}	4.56	0.60	0.05

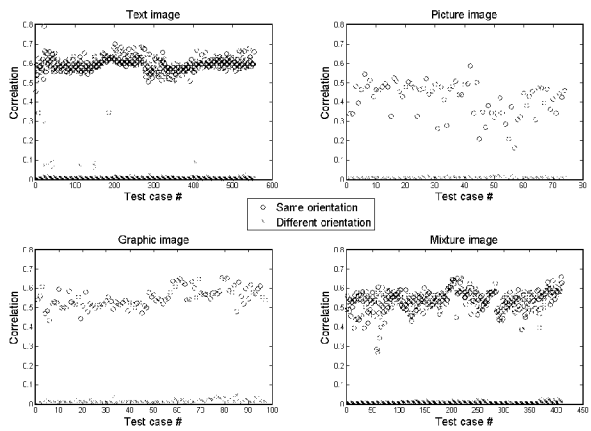

Figure 2. Correlation computed by 2D phase correlation algorithm.

[5] Keller, Y., Averbuch, A., and Israeli, M., "Pseudo-polar based estimation of large translations rotations and scalings in images," *IEEE Transactions on Image Processing* **14**, 12–22 (January 2005).

[6] Reddy, B. and Chatterji, B., "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transactions on Image Processing* **3**, 1266–1270 (August 1996).

[7] Amin, A. and Wu., S., "Robust skew detection in mixed text/graphics documents," in [*Proceedings of the 8th International Conference on Document Analysis and Recognition*], 247–251, IAPR (2005).

[8] Jiang, H., Han, C., and Fan, K., "A fast approach to detect and correct skew documents," in [*Proceedings of the 9th International Conference on Pattern Recognition*], 7276–7276 (check these!), IAPR (1996).

[9] Lu, Y. and Tian, C., "Improved nearest neighbor based approach to accurate document skew estimation," in [*Proceedings of the 9th International Conference on Document Analysis and Recognition*], 503–507, IAPR (2003).

[10] Okun, O., Pietiekain, M., and Sauvalo, J., "Robust document skew detection based on line extraction," in [*Proceedings of the 11th Scandinavian Conference on Image Analysis*], 457–464 (1999).

[11] Yu, C., Tang, Y., and Suen, C., "Document skew detection based on the fractal and least squares method," in [*Proceedings of the Third International Conference on Document Analysis and Recognition*], 1149–1152, IAPR (1995).