

# Correcting for Variable Skew

A. Lawrence Spitz

Document Recognition Technologies, Inc.  
616 Ramona Street, Suite 20  
Palo Alto, California 94301, USA

**spitz@docrec.com**

**Abstract.** The proliferation of inexpensive sheet-feed scanners, particularly in fax machines, has led to a need to correct for the uneven paper feed rates during digitization if the images produced by these scanners are to be further analyzed. We develop a technique for detecting and compensating for this type of image distortion.

## 1 Introduction

Many techniques have been developed for detecting the dominant skew angle of printed documents [2]. However sometimes more than one skew angle is represented in the document image. Baird [3] points out that the paste-up of text columns may result in slightly (but significantly) different skew angles and notes the importance of segmenting the columns first in order that individual skew corrections can be applied to the individual columns. Spitz [6] and Antonacopoulos [1] have described techniques for detecting the presence of multiple discrete skew angles in a single document image. Spitz showed the angular distribution of the “power” (as defined by Baird) for both instances of multiple discrete skews. This is reproduced in Fig. 1.

This paper presents a solution to a different problem, that of documents which contain a continuously variable skew as a function of vertical distance along the page.

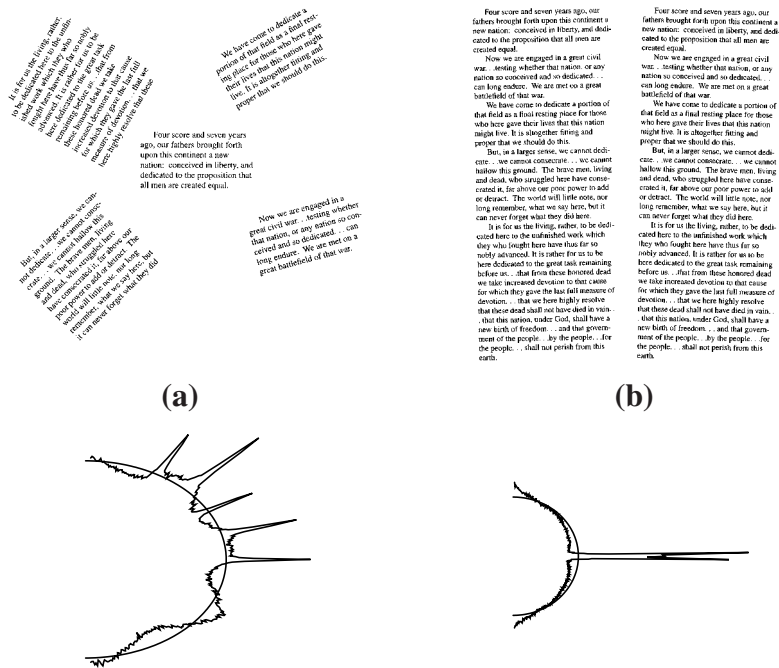
In Section 2, we describe the type and range of skew angle variability seen. Section 3 describes the segmentation of the image and determination of the dominant skew angle for the individual segments. Section 4 describes the correction of the skew for each of these segments. We discuss the re-assembly of the image in Section 5.

Throughout Sections 2 through 5 we will describe the algorithms using a much simplified page image specifically synthesized for this study because doing so makes the processes much easier to understand. In Section 6 we will provide a real-world example.

## 2 Image Characteristics

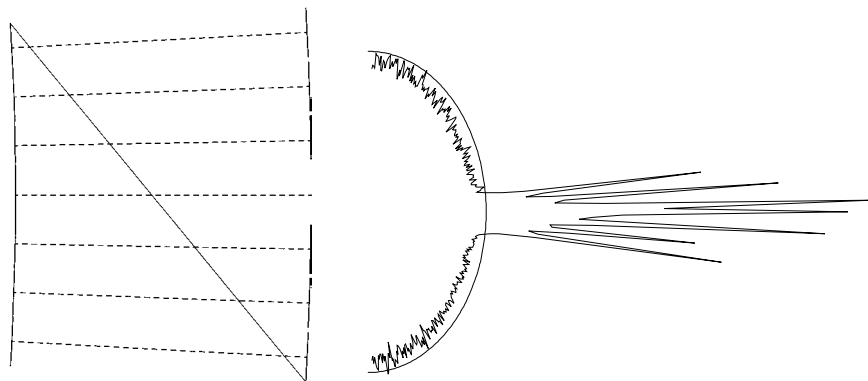
Inexpensive sheet-feed scanners often feed paper unevenly, that is as the page passes the CCD array from top to bottom, the left and right edges do not move at exactly the same rates. The effect of this is to produce a distorted image with continuously variable skew from top to bottom.

In order to simplify development and illustration of the techniques described here, we synthesized an number of images demonstrating variable skew, one of which is shown in Fig. 2 This image has been printed and scanned to introduce noise to the basic structure. Note that though this simulation is useful for didactic purposes it is an insufficiently accurate representation of “real-world” distortion because continuous variation of skew angle is not apparent due to the small number of nominally horizon-



**Fig. 1.** Two document images and their respective angular power distribution. (a) shows five text blocks and five distinct peaks in the distribution. (b) illustrates the ability to detect small differences in skew. Note the double peak in the power distribution.

tal lines. Note that the lines to be made horizontal are dashed in order to simulate individual characters in the way that they generate pass codes and that the “diagonal” line will be able to demonstrate, by its continuity or lack of continuity, the degree to which distortion of the input image (in the digital domain) introduces artifacts.



**Fig. 2.** Synthesized image with continuously variable

Note that most images that demonstrate variable skew do not exhibit the monotonic, and nearly linear, skew variation shown in Fig. 4 for the test image.

### 3 Skew Detection

Skew angle detection is fundamentally a two stage process. In the first stage one determines the locations of the features on which alignment quality will be assessed. In the second, various tests are applied to determine if a proposed alignment is a good one relative to a prior sample alignment or other standard. Depending on the choice of alignment features, the computational expense of calculating tentative alignments may be variously divided between the calculation of the alignment measure itself or in the testing of the quality of the particular alignment.

We adopt the method of testing rotational alignment developed and described by Baird [3], but we use a different set of fiducial points.

#### 3.1 Fiducial Points

As we have described in earlier papers, we detect skew on the basis of the positions of white CCITT pass codes [6][7]. Pass codes are described by Hunter and Robinson [4] and are a part of the Group IV facsimile coding scheme. CCITT pass codes themselves do not carry color information; however it is easy to add. If the process of decompression is adding white pixels at the time a pass code is encountered in the compressed data stream, that pass code is labeled white. Hence the presence of a white pass code indicates the presence of the bottom of a black structure.

Text documents contain images of characters. Each character has at least one bottom point. A, H, K, M, N, Q, W, X, h, k, m, n and w have more than one bottom point. The bottom of most characters (except g, j, p, q, y) lie on a baseline. Thus character bottoms are a reasonable structure on which to determine an alignment.

Of course facsimile images can contain graphical elements other than character images such as rulings and logotypes. These structures too generate pass codes.

The white pass positions for the test document shown in Fig. 1 are shown in Fig. 3.

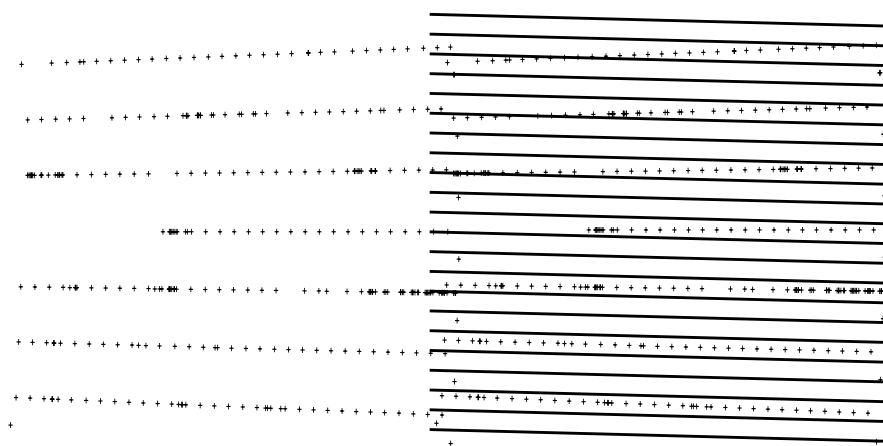
The calculation of alignment is made efficient by processing this relatively small number of fiducial points as representative of the entire image.

#### 3.2 Alignment

Alignment of fiducial points is determined by overlaying the spatial distribution of fiducial points with parallel bins. This system of bins is rotated through the desired range of angles and the populations of each of the bins is calculated at each angle. For each angle the “power” of the alignment is calculated by computing the sum of squares of the bin populations. This power measurement is therefor maximized when the variance of the bin population is greatest since the total number of fiducial points is constant. The power is greatest when the largest proportion of fiducial points falls within the smallest number of bins.

#### 3.3 Multiple Skew Angles

We have now detected the dominant skew angle, that is the angle at which a particular bin has the maximum population of passes. We then iteratively remove the passes



**Fig. 3.** White pass positions in the test documents and a set of rotationally aligned bins overlaying the fiducial points. Note that the actual bins would be much smaller than those.

found in that bin, leaving behind passes aligned at the same angle but at different positions on the page, and re-compute the dominant skew angle.

In the example shown in Fig. 2, we find seven dominant skew angles, one for each nominally horizontal line. Along with the skew angle we calculate the vertical position along the vertical centerline of the page at which this skew angle exists. If there is significant skew angle difference between adjacent points on the vertical centerline we interpolate to reduce this angular distance between samples. Currently we re-sample to reduce the inter-sample skew angle difference to less than 0.5 degrees.

We can demonstrate the distribution of skew with a plot of skew angle vs. vertical distance down the page measured on the vertical centerline. Such a plot is shown in Fig. 4. Note the filled squares denoting the original 7 detected skew angles, the skew profile, shown by the solid line, determined by fitting a curve to the detected skew angles. This skew profile is subsequently sampled at regular intervals in order to provide the basis for rotating, slicing and re-assembling the image.

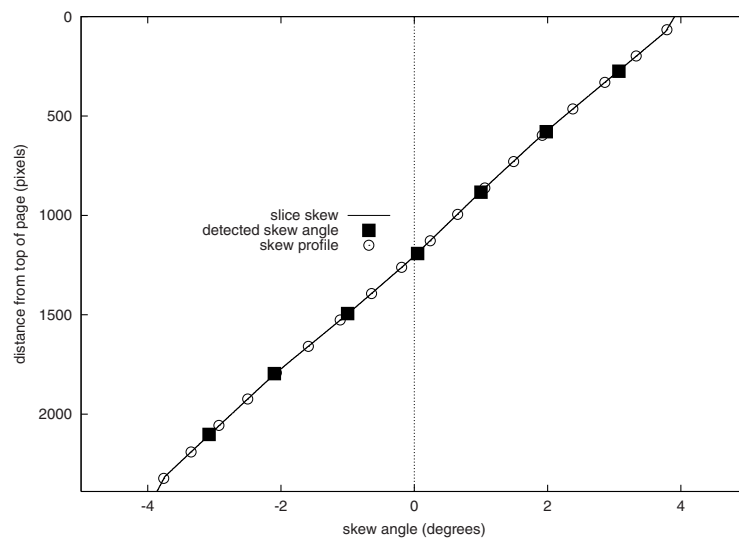
For each skew angle thus determined we rotate the source image by the appropriate angle to correct for the skew. For the test image the original 7 skew angles are increased to 18. Ten of these 18 are shown in Fig. 5

Rotation is accomplished using the three-shear method due to Paeth [5]. Each shear is implemented by looping over the source pixels and distributing fractions to each of the destination pixels. This has an “anti-aliasing” effect.

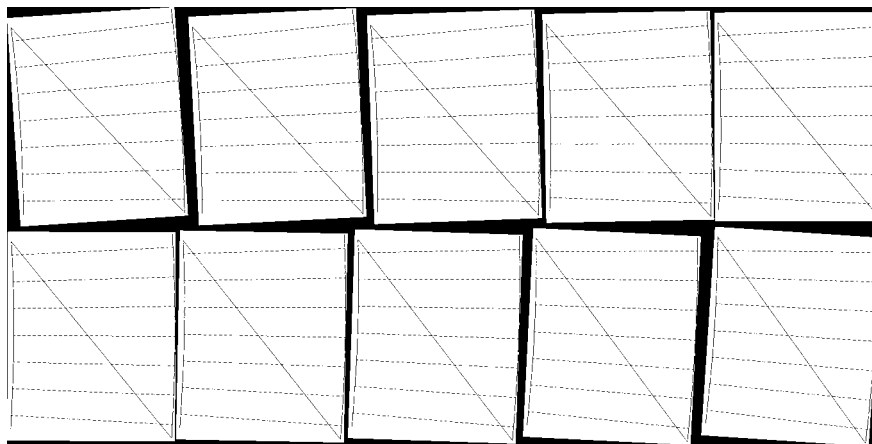
#### 4 Slicing the Image

For reasons of efficiency we wish to process the image in as few slices as possible consistent with maintaining image quality sufficient for further recognition processes such as OCR.

To this end we determine the minimum number of slices of constant height necessary to ensure that the maximum difference of skew angle from slice center to slice



**Fig. 4.** Plot of detected skew angles, the skew profile that they define and the samples along the profile used for rotating the slices of the image.



**Fig. 5.** Ten of the 18 rotated images resulting from finding the dominant skew angle in a test

center is less than a pre-determined threshold, in this instance 0.5 degrees. We constrain this number to never be less than the number of discrete skew angles detected earlier or greater than 100. Therefore at the maximum number of slices, necessary for the most severely warped images, a 200 spi facsimile of a letter or A4 page would result in slices that are approximately 22 pixels in height.

Once the number of horizontal slices is determined, the skew profile is sampled at equal intervals of vertical distance. Each sampling point is then associated with a verti-

cal position and an angular rotation. For each such point the appropriately rotated image is cropped according to the slice center position.

### 5 Re-assembling the Image

The resulting set of images slices is then stacked from top to bottom. The results of slicing and splicing the image are shown in Fig. 6.

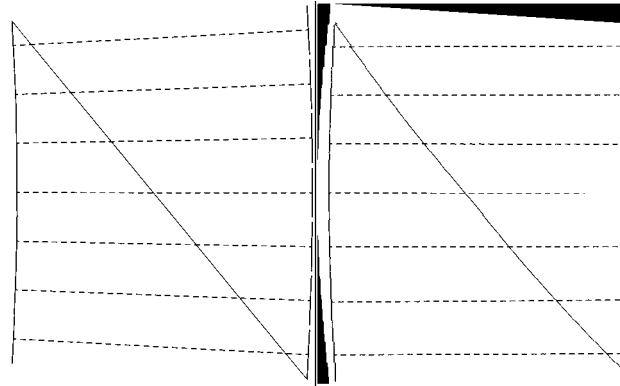


Fig. 6. Distorted and corrected test image.

### 6 Real World

Fig. 7 shows an example of a “real world” facsimile image which demonstrates variable skew. The pass code positions on which that skew determination is based is shown in Fig. 8.

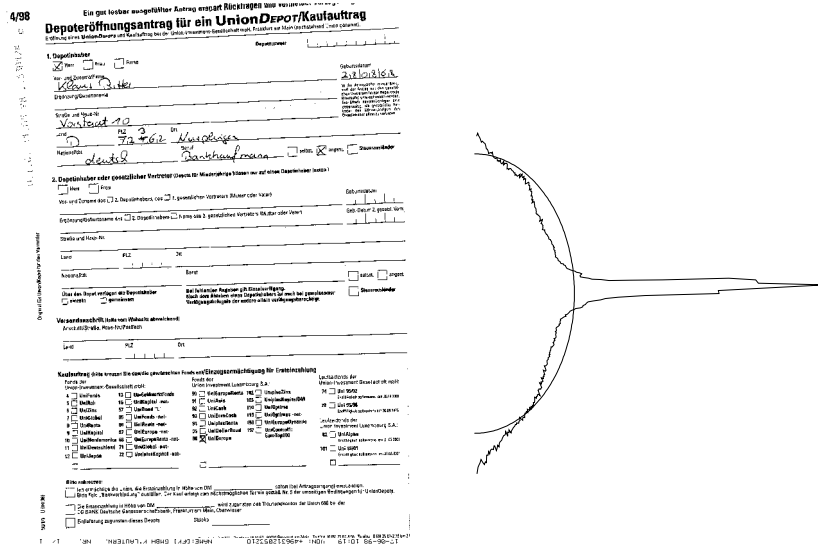


Fig. 7. An actual facsimile image demonstrating variable skew and the plot of the angular distribution of “power”. Note the broad peak relative to those shown in Fig. 1.

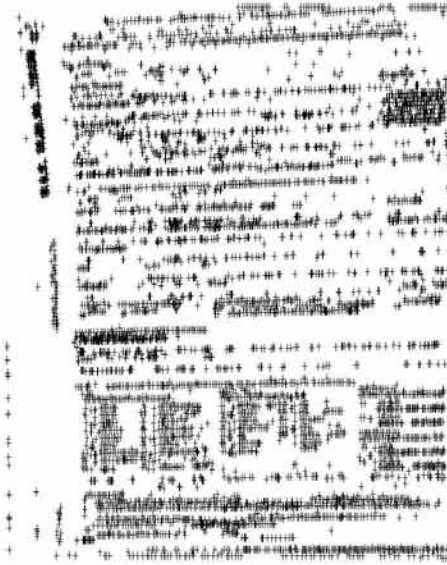


Fig. 8. White pass positions in the “real world” documents

The skew angles detected in the document as a function of vertical position on the page, the skew profile and the selected sampling points are shown in Fig. 9.

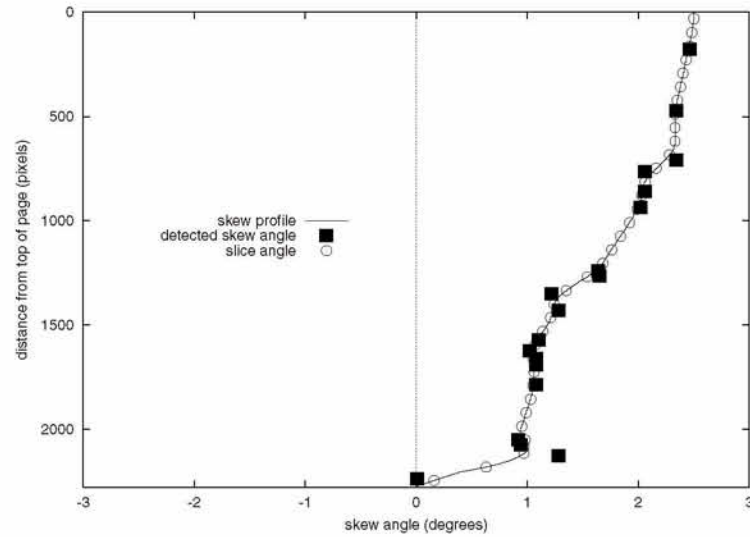


Fig. 9. Plot of detected skew angles, the skew profile that they define and the samples along the profile used for rotating the slices of the image.

Fig. 10 shows the facsimile document both before and after variable skew detection and correction.

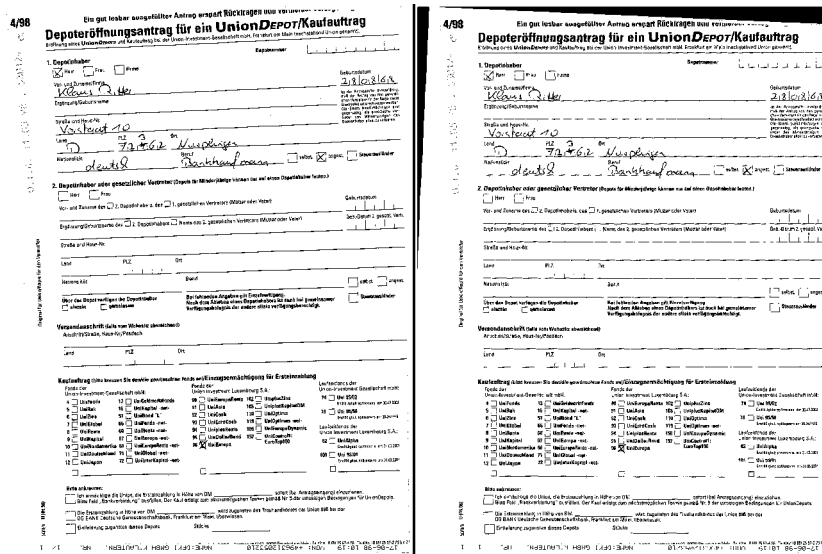


Fig. 10. Distorted and corrected “real world” fax images.

## 7 Conclusions

It might be possible to achieve somewhat better results both in terms of the quality of the corrected image and in reduction of computational expense, if the slices taken were not constrained to be of constant height. That is, if there were a large range of the vertical span of the document where the detected skew angle were relatively constant it would be more efficient to make this slice larger and only deal with small slices in areas of rapidly varying skew angle.

To do this would reduce the number of rotation operations and crops and the resultant image would have fewer artifacts where the slices are re-joined. These may be very positive attributes deemed worthwhile for particular applications, but in our experience the described technique is good enough and has the virtue of simplicity.

It might also be advantageous to ensure that the points on the skew profile selected for slicing the image have minimum mean-square distance from the detected skew angles, thus ensuring that the structures that generated the peak in the power distribution become more nearly horizontal in the corrected image

## References

1. Apostolos Antonacopoulos, “Local Skew Angle Estimation From Background Space in Text Regions”, *International Conference on Document Analysis and Recognition*, pp 684-688. 1997.
2. Andrew D. Bagdanov, Junichi Kanai, “Evaluation of Document Image Skew Estimation Techniques”, *SPIE*, pp 343-353, 1996.
3. Henry S. Baird, “The Skew Angle of Printed Documents”, *SPSE Symposium on Hybrid Imaging Systems*, pp 21-24, 1987.



4. Roy Hunter & A. Harry Robinson, "International Digital Facsimile Coding Standards". *Proceedings of the IEEE*, Vol. 68, No. 7, pp 854-867, July 1980.
5. Alan W. Paeth, "A Fast Algorithm for General Raster Rotation", *Graphics Interface '86*, pp 77-81, 1986.
6. A. Lawrence Spitz, "Analysis of Compressed Document Images for Dominant Skew, Multiple Skew and Logotype Detection", *Computer Vision and Image Understanding*, 70, 3, pp 321-334, 1998.
7. A. Lawrence Spitz." Skew Determination in CCITT Group 4 Compressed Images", *Symposium on Document Analysis and Information Retrieval*, Las Vegas, Nevada, 1992.