Skew Detection Using the Radon Transform*

Bogdan Raducanu, Costin-Anton Boiangiu, Alexandra Olteanu, Alexandru Ștefănescu, Florin Pop, Ion Bucur

*Computer Science Department, Automatic Control and Computers Faculty, University POLITEHNICA of Bucharest (e-mail: BRaducanu@gmail.com, {costin.boiangiu, florin.pop, ion.bucur}@cs.pub.ro)

Abstract: In an automatic document conversion system, which builds digital documents from scanned articles, there is the need to perform various adjustments before the scanned image is fed to the OCR system. This is because the OCR system is prone to error when the text is not properly identified, aligned, de-noised, etc. Such an adjustment is the detection of page skew, an unintentional rotation of the page, which probably occurred during scanning. In this work we have used the Radon transform to detect the angle at which a page was skewed.

Key-Words: Skew, Radon, Digitization, OCR

1. INTRODUCTION

Document image processing has become an important technology for automating digitization processes.

A typical document digitization system consists of the following steps:
- Image pre-processing: noise reduction, skew correction, etc;
- Hierarchical segmentation: detect page layout, recognize text areas (also tables, figures, etc), then text paragraphs, individual lines, then words, characters (Boiangiu et al. 2008 b,c);
- Supervised recognition;
- OCR with spell-checking ability;
- Output to a popular format (PDF, html).

A digital image of a hard document is almost always skewed during the scanning process, most likely because the document was not correctly placed inside the scanner. Since the layout analysis engine and optical character recognition engine rely heavily on heuristics and thus are very sensitive to page skew, an important pre-processing step is required to determine the angle of the skew and then correct it.

This paper* analyzes a novel method for detecting the skew angle based on the Radon transform, which is also briefly discussed.

The methods presented here are appropriate for small angle rotations, in the interval $[-10, +10]$ degrees. Larger imperfections should be treated through other means. Also, non-linear distortions can appear if the page was not flat on the scanner. This kind of distortions is not discussed here, though the methods presented can be enhanced to deal with them.

In the following section, a survey of well established skew detection methods is presented (Hull et al. 1998). Afterwards, an introduction to the Radon transform is made and the new method is analyzed. The following methods, as well as our new method, assume that the input image contains some amount of text. The paragraph lines are useful as a reference to establish the skew angle.

2. EXISTING METHODS

A good simple method for determining the skew angle of a document image employs a vertical projection profile (Nakano et al, 1991). This is an array with the number of locations equal to the number of rows in the image. Each position in the array contains the number of black pixels for the corresponding row. The obtained histogram has a maximum amplitude and frequency when the image is not skewed, since that is where the most co-linear black pixels are found. One approach to implementing this method is to rotate the input image through the desired range of angles and determine the projection profile for every rotation. Every histogram is compared to each other and through various feature detections, the correct angle is determined. A heuristic is used to compare the histograms, and there are various approaches for this. One example is to sum the squared differences in adjacent cells of the histogram (Posl, 1986). A speed improvement can be obtained by using a range search for the angle. The angle is first searched in a coarse interval, after which the search is repeated for a finer resolution range. Another technique for skew detection works by extracting all features from the image (characters or connected entities). Each feature contributes with a point to a new image (Farrow et al., 1994). All computations are performed on the new image only. Thus, the number of memory accesses is reduced. For example, the bottom center point of each connected component is used to form a new image (Baird, 1987) (Baird, 1991). Other methods are
optimized for different image encoding and compression schemes (Smith, 1995).

Yet another class of techniques uses the Hough transform (Boiangiu et al. 2008a), a well known image processing method used to detect straight lines in the image. The lines are detected by carrying out a voting process in the image space and constructing a parameter space. Each pixel in the image is counted as a vote on all the lines it may be on. A line is represented by the equation:

\[ \rho = x\cos\theta + y\sin\theta \]  

(1)

An accumulator array is parameterized by a range of values for \( \rho \) and \( \theta \). The image pixels are iterated and for every pixel \((x,y)\) and angle \(\theta\) the \(\rho\) parameter is calculated and a vote is cast in the \((\rho,\theta)\) location of the accumulator. At the end, the parameter space is inspected and local maxima will correspond to actual lines. To detect the skew, the Hough method is applied and straight lines are extracted from the skewed image. An average of the slopes of those lines will most likely be the skew angle, as most co-linear black pixels are found on lines which are parallel to the image bottom line. Down-sampling the image can help improve the speed of the implementation.

Another type of methods for determining the skew of a document works by extracting information that represents skew from local characteristics of the image (Ejiri et al., 1993) (Ham et al., 1994).

Such local characteristic information is usually obtained by using masks. For each region that matches a mask a vote is cast. For example, if it is known that the image contains handwritten text, then a mask can be applied to the image and the average angle at which the text was written can be calculated (Rundle, 1974).

3. THE RADON TRANSFORM

The Radon transform has many definitions (Toft, 1996). The one presented here is perhaps the easiest to understand. The Radon transform \(g(\rho,\tau)\) is the integral of a continuous 2D function \(f(x,y)\) over a collection of slanted lines.

\[ g(\rho,\tau) = \int_{-\infty}^{\infty} f(x, \rho x + \tau) dx \]

The Radon transform is linear – the Radon transform of a weighted sum of functions is the same weighted sum of the individually Radon transformed functions. This is an important property which will be used to approximate the transform by means of a computer program. Also, the Radon transform of a rotated, scaled or translated image can be obtained knowing the Radon transform of the original image and the affine transformation parameters (Kunyansky, 1992) (Spitz, 1992).

From the definition, it is found that the Radon transform is actually a projection of the image space along specified directions. It is a fundamental property that is helpful for image processing. The inverse of this transform is very important in medical imagery, electron microscopy, etc. as it can reconstruct the original image from orientated projections (Deans, 1992) (Freeman et al., 2003) (Natterer, 2001).

3.1 The Normal Radon Transform

Notice that the integral in Eq. 2 suffers from the inability to deal with steep lines (lines with infinite slopes). To overcome this, a new form of the Radon transform may be used (this is not of significant interest here since most of the skew angles are supposed to appear in the vicinity of 0).

\[ g(\rho,\tau) = \frac{1}{|\sin\theta|} \int_{-\infty}^{\infty} f(x, \frac{\rho}{\sin\theta} - xcot\theta) dx \]

(3)

The line is now represented through polar coordinates (see Fig. 1).

\[ \rho = x\cos\theta + y\sin\theta \]  

(1)

The parameter \( \rho \) is the shortest distance from the origin of the coordinate system to the line, and \( \theta \) is an angle corresponding to the angular orientation of the line.

All lines can be described by choosing \( 0 \leq \theta \leq 2\pi \) and \( \rho \geq 0 \).

3.2 Computation

In digital image processing a discrete transform is used to approximate the integral (Brady, 1998). The numerical algorithm computes the integral along parallel beams that originate from a reference line and pass the entire image space (see Fig. 2). Each beam adds the black pixels it encounters in an accumulator array. This is possible because the Radon transform is considered linear.

To provide a good estimation, each pixel is divided into four subpixels and each subpixel's contribution is proportionally split into the two nearest bins, according to the distance between the projected location and the bin centers. This is an oversampling technique that is used in the same way as in the graphical anti-aliasing methods. Oversampling using factors
greater than two (on both axes) is possible but in practice the benefit would be hardly noticeable.

Consider the following example where the Radon transform was applied on a simple image containing a black square for a rotation angle of 45 degrees.

The peak in Fig.5 corresponds to the diagonal of the square, since that is where the most co-linear black pixels are found.

The following sections will show how the Radon transform can be used to approximate the skew angle of a document. This method is closely related to the projected profile method briefly presented in the previous section.

4. THE METHOD

Today, it is considered to be a classic of French literature, though it is overshadowed by its many subsequent adaptations. The novel was translated into English in 1911. It has since been adapted many times into film and stage productions, the most notable of which were the 1925 film depiction and Andrew Lloyd Webber’s 1986 musical, starring Michael Crawford as the Phantom, Steve Barton as Raoul, and Sarah Brightman as Christine. The Phantom of the Opera musical is now the longest running Broadway show in history and the most lucrative entertainment enterprise of all time, its worldwide box office over the past 20 years has outgrossed even the highest grossing film in history, 1997’s Titanic.

Fig. 4 – Input Image

Fig. 5 – Radon Transform for 45 degrees
The following images show an example of applying the Radon transform over a black and white image. The histograms are plotted and displayed for comparison. Also, the filtered entities and bottom-centre points can be observed.

Fig. 8 shows the Radon transform array of the image in Fig. 6. The horizontal axis represents the reference line used to calculate the integral, while the vertical axis represents amplitude. The peaks are due to the horizontal lines formed by the bottom-centre points of the characters. It can be seen that without skew the Radon transform is formed by uniformly spaced peaks with high values. Between two peaks the function is close to 0.

The skewed image exhibits lower magnitude peaks and the pixels are also distributed between adjacent peaks.

We have tested our collection of scanned newspaper articles and compiled the following table of results.

<table>
<thead>
<tr>
<th>Skew</th>
<th>Average Run-time</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1°</td>
<td>5s</td>
<td>90%</td>
</tr>
<tr>
<td>3°</td>
<td>8s</td>
<td>85%</td>
</tr>
<tr>
<td>5°</td>
<td>10s</td>
<td>85%</td>
</tr>
<tr>
<td>10°</td>
<td>15s</td>
<td>82%</td>
</tr>
</tbody>
</table>

Table 1 - Results

5. SKEW CORRECTION

Skew correction is not a trivial task either. After the skew angle has been determined, there needs to be a way to fix the skew. The simplest approach is to rotate the image in the opposite direction and correct the skew. However, this method is time and memory consuming for large images. Other, more efficient, methods exist which provide a good level of accuracy while having the advantage of a greater running speed and lesser memory used [22].

Consider the rows of the image movable along the horizontal axis. If each row going upward progressively moves further to the right the picture is “italicized”. This transformation is called a shear. The second kind of shear is at right angles to the first shear – a column preserving transformation. Together they form a near-rotation of the image which may appear as accurate for small angles.

6. IMPROVEMENTS

The method presented here is generic and can mostly be improved by plugging in other heuristics, other entity filters and feature point selectors.

The proposed skew detection method is largely dependent on the heuristic function, thus, a careful analysis is required to find such a function that correctly describes the nature of the transform array.

One simple and accurate approach is the array variance. Because of the nature of the transform array for the unskewed image, the variance of this vector is large, due to the high valued peaks and low valued pauses. For the tested images this variance was found to be as large as 5 times the variance of a 3 degree skewed image. It is a very powerful measure of the resulting histogram.

A very useful property of the array variance in case of the Radon Transform is that the values tend to be monotonous in the neighborhood of the expected solution. This is a very
important observation because it may lead to a significant speed optimization: instead of performing almost linear search of a solution in what the algorithm considers to be a small search interval, a binary-like search can be used instead and thus significantly reducing the computational demands.

Another good heuristic function is to sum the squared differences of the adjacent transform array cells. If peaks have low value neighbor cells, this sum will be quite large and the correct angle can be found. Simple methods like thresholding can also be advantageous for particular types of images.

A good optimization, if it doesn’t affect accuracy is to discard a small percentage of pixels from the input image. This would not affect the Radon transform histogram shape, but would increase performance and decrease memory usage. As a result, down-sampling the image can (in most of the cases up to 150 Dots-Per-Inch) boost running speed and maintain level of quality of results.

7. CONCLUSIONS

A good skew detection and correction algorithm is vital to any document digitization system. In this work we have explored the possibilities of the Radon transform in order to implement such a detection algorithm. The correction algorithm was not thoroughly discussed, but some ideas were referenced.

The Radon transform is a very important tool for image processing and feature detection, particularly because of its resistance to noise and imperfections. It can also be used to detect lines or other shapes (in a general form) and most of the applications will find its robustness very useful in noisy environments.

ACKNOWLEDGEMENTS

The research presented in this paper is supported by national project "SORMSYS – Resource Management Optimization in Self-Organizing Large Scale Distributed Systems", Project CNCSIS-PN-II-RU-PD ID: 201 (Contract No. 5/28.07.2010), and it was co-founded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labor, Family and Social Protection through the Financial Agreement POSDRU/89/1.5/S/62557.

REFERENCES


Baird H.S., Apparatus and method for skew control of document images, AT&T Bell Laboratories, 1991


Farrow G.S.D., M. A. Ireton, C. S. Xydeas, Detecting the skew angle in document images, Signal Processing: Image Communication 6, 1994


Postl W., Detection of linear oblique structures and skew scan in digitized documents, Proceedings of the 8th International Conference on Pattern Recognition, 1986


