Pattern Recognition Letters 31 (2010) 1403–1411

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Fast and robust skew estimation of scanned documents through background area information

Angélica A. Mascaro, George D.C. Cavalcanti *, Carlos A.B. Mello **

Center of Informatics, Federal University of Pernambuco, Brazil

ARTICLE INFO

Article history:
Received 24 March 2009
Received in revised form 23 December 2009
Available online 23 March 2010
Communicated by A.M. Alimi

Keywords:
Skew angle estimation
Document analysis
Noisy documents

ABSTRACT

Skew correction of scanned documents is a crucial step for document recognition systems. Due to the problem of high computational costs of the state-of-the-art methods, we present herein a variation of a parallelograms covering algorithm. This variation strongly reduces the computational time and works over noisy documents and documents containing non-textual elements, like: stamps, handwritten components and vertical bars. Experimental studies with different databases show that this variation overcomes well-known techniques, achieving better results over synthetic rotated documents and real scanned documents.

1. Introduction

The correction of the skew angle of a scanned document is a very important step towards automatic document recognition systems. A skew in a document image can interfere in the whole posterior process, like layout identification and segmentation, compromising recognition. To solve this problem, several algorithms were developed to estimate the skew angle of a document image (Cattoni et al., 1998; Hull, 1998).

Methods based on the Hough transform (Le et al., 1994; Min et al., 1996; Amin and Fisher, 2000; Singh et al., 2008; Hinds et al., 1990; Pal and Chaudhuri, 1996; Vailaya et al., 2002) convert the Cartesian coordinates in Polar coordinates and the number of black pixels is accumulated in a vector. Peaks values in this vector correspond to straight lines in the image which is probably related to text lines or lines of tables and forms. These methods are popular because of its robustness, but generally demands high computational time and space to find the peak on the Hough Plane. Also, most techniques were developed to work over printed documents relying on the fact that a document contains a minimum space of printed text area. Amin and Fisher (2000) applied Hough transform to the last line of segmented text blocks and grouped pixels into connected components to reduce computational cost.

Another category includes the methods based on projection profile (Postl, 1986; Li et al., 2007; Nicchiotti and Scagliola, 1999; Baird, 1987; Ciardiello et al., 1988; Kanai and Bagdanov, 1998) in which the amount of black pixels in each line is calculated. Skew detection algorithms based on projection profile assume that most of the document is composed by text lines and its accuracy generally decays in the presence of other elements, such as graphics or noise. Several works were developed aiming to overcome these problems. Baird's projection profile based algorithm (Baird, 1987) uses connected component analysis and creates a projection profile using a single point to represent each connected component. Ishitani (1993) proposed a skew detection method based on maximum variance of transition-counts (which can be considered a variation of a projection analysis approach) to deal with images containing a mixture of text areas, photographs, figures, charts or tables.

Skew estimation techniques based on cross-correlations (Avanindra, 1997; Gatos et al., 1997; Akiyama and Hagita, 1990; Yan, 1993) estimate the document skew by measuring vertical deviations along the image. Avanindra (1997) selected random small regions of the image to compute the interline cross-correlation of a document to reduce the time consuming and the influence of graphics which affects the cross-correlation methods accuracy.

There are also methods based on nearest neighbor clustering (Lu and Tan, 2003; Smith, 1995; Shivakumara and Kumar, 2006) which are based on the assumption that characters in a line are aligned and close to each other. Methods in this category start by labeling connected black pixels to group them in blocks and, after that, in larger blocks with similar features. Finally, they try to estimate the skew based on the mutual distance and spatial relationship between the text lines. The effectiveness of this kind of

* Correspondence to: G.D.C. Cavalcanti, Federal University of Pernambuco, Center of Informatics, Av. Prof. Luis Freire, Cidade Universitária 50740-540, Recife, PE, Brazil. Tel.: +55 81 2126 8430x4346; fax: +55 81 2126 8438.
** Corresponding author.
E-mail addresses: aam3@cin.ufpe.br (A.A. Mascaro), gdcc@cin.ufpe.br (G.D.C. Cavalcanti), cabm@ieee.org (C.A.B. Mello).
URL: http://www.cin.ufpe.br/~viisar (G.D.C. Cavalcanti).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.03.016

analysis is dependent on the quality of the binarization process and on the degree of noise in the image. Lu et al. (2007) proposed a technique to estimate the document skew based on the observation of the white runs that span the interline spacing of the text. Ávila–Lins algorithm for skew estimation (Ávila and Lins, 2005) groups neighbors of connected components and draws imaginary lines following the text lines which are used to detect both skew angle and landscape/portrait orientation. Saragiotis and Papamarkos (2008) filtered the characters in the image based on features such as density and width to height ratio to avoid including other components in the analysis and used linear regression to estimate an imaginary base line for the text lines. The novelty of the work was the capability to deal with more than one skew angle in a document.

Most part of the traditional skew estimation techniques has high computational cost and generally make some restriction about the document input type, such as need for a minimum text area. A common problem for the traditional skew estimation methods is to deal with documents containing complex layouts, multiple font styles and sizes, noise or high amount of non-text regions such as figures or tables. Variants of traditional methods commonly aims at reducing the time complexity of the algorithms or at selecting the data involved in the computation to avoid interference of non-textual components.

It is presented herein a variation of the parallelograms covering technique for skew estimation proposed by Chou et al. (2007). This method follows the idea that a document is composed by a combination of rectangular objects, such as text lines, forms, figures, tables, etc.; and make no assumptions about the type of input documents, with no need for a minimum text region area. Chou et al.'s algorithm constructs parallelograms at various angles to decide the best skew angle. Our proposal improves Chou et al.'s algorithm changing the criterion to evaluate the best angle – now based on the background area. With this variation it is possible to effectively reduce the search for the correct skew angle. We also propose a variation in the method which aims to deal with noisy images and documents with complex layouts, such as those containing forms, tables and handwritten components.

This paper is organized as follows. Section 2 presents the main idea of the Chou et al.'s original method; the modifications that we propose in this method is presented in Section 3. Section 4 shows an experimental study with synthetic rotated and real scanned document images, while Section 5 presents some final remarks.

2. Skew angle estimation based on parallelograms

Chou et al. (2007) proposed an algorithm to estimate the skew angle of documents by constructing parallelograms at various angles and then deciding which one best fits the objects in the image. This algorithm considers that a document is a combination of rectangular objects, such as text lines, forms, figures, tables, etc. The process starts by drawing parallel lines (called scan lines) at a certain angle θ. A scan line is a row with one pixel width that crosses the image at an angle from left to right. These scan lines are vertically divided into fixed size regions (slabs) – see Fig. 1(a).

Dividing the image into slabs also divides the scan lines. These subdivisions of the scan lines are called sections. Chou et al. used a slab with 450 pixels width. As expected, in a left-to-right reading, the rightmost slab can be smaller than the others according to the width of the image.

In the parallelograms construction phase, each section of the scan lines is examined by a certain angle. If this section contains at least one black pixel, it is turned to gray; otherwise, it stays white. Adjacent gray sections form parallelograms. The scan lines are skewed at different angles and the size of the region not covered by parallelograms is evaluated. As exposed in Fig. 1(a)–(c), when the lines are drawn in the same angle of the document inclination, the white region is larger.

By doing so, the angle which produces the largest white region is considered as the skew angle of the document. Fig. 1(b) shows the parallelograms that were constructed with scan lines drawn at $-6°$, which is the same angle of the document skew (Fig. 1(a)). Fig. 1(c) shows the parallelograms with scan lines at $1°$. As we can see, the white region at $-6°$ is larger than at $1°$.

Chou et al.'s algorithm performs a full search for skew angles between $-15°$ and $+15°$. In general, we expected that digitized documents have a skew angle around $0°$. Also, this approach is vulnerable to noise and other components, such as large amounts of vertical bars and handwritten components.

3. Proposed algorithm

We propose modifications in Chou et al.'s original algorithm. These modifications are related to: how to measure the size of the region not covered by parallelograms (Section 3.1); how to perform an efficient search for the best angle (Section 3.2) and how to avoid the undesired interference caused by components such as noise and vertical separators (Section 3.3).

3.1. Analysis of the background area

In Chou et al.'s algorithm, the skew angle of the document is estimated by calculating the largest white region achieved within the tested angles. To measure the size of the white region at an angle θ, Chou et al. proposed to count the number of white sections at that angle. However, when the value of the angle increases (in positive or negative directions), the total number of sections covering the whole image is also changed; and this is a problem. As illustrated in Fig. 2(a), at $0°$ we have n scan lines, which is also the number of rows of the image. In Fig. 2(b), the scan lines are rotated and the total number of scan lines is now equal to $n + a$, where a is the number of extra small lines.

Aiming to make a fair evaluation, it is expected that the total number of sections stays the same for every angle. In other words, if you have 1000 sections at angle θ, for example, you also should have 1000 sections at all other angles. As shown in Fig. 2(b), an im-

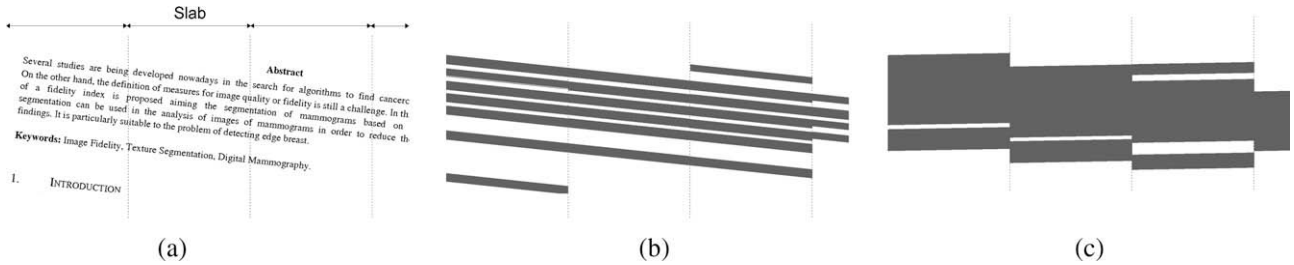

Fig. 1. (a) An image with $-6°$ skew angle. (b) Parallelograms constructed at $-6°$. (c) Parallelograms constructed at $1°$.

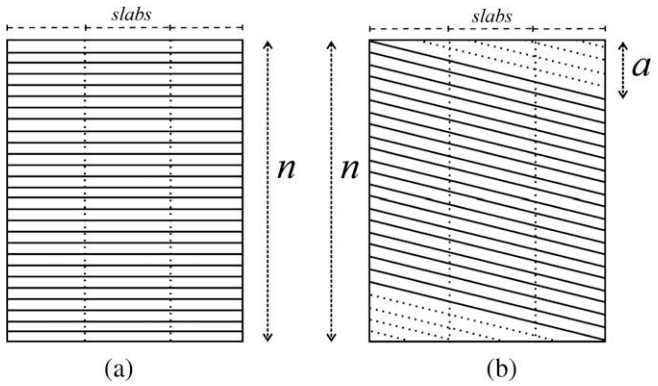

Fig. 2. (a) A document with scan lines drawn at 0°. (b) The scan lines drawn at an angle different from 0°. The total number of scan lines is increased with extra small lines.

age portion must be ignored to assure that the total number of scan lines stays constant. Thus only the lower or the upper part of the smaller lines (represented by the dashed lines in Fig. 2) should be considered. This is a problem because the ignored part can contain valuable information to estimate the inclination of the document.

On the other hand, because of slab division, even setting up the number of scan lines to a constant (by ignoring the upper or the lower part), a different number of sections per angle is produced. This happens because, in the upper and the lower parts of the image, the lines are smaller than the ones in the middle of the document. For example, in Fig. 2(a), there are three slabs then the total number of sections is $3 \times n$. In Fig. 2(b) the total number is changed because of the small lines (dashed lines); some sections disappear and other ones have a small size.

Another problem with Chou et al.'s approach is due to the fact that the procedure of counting the number of white sections gives the same weight for sections with different number of pixels. Thus, small sections which contain only few pixels have the same "importance" in the analysis as sections with the full slab width. This happens in the lower (or upper) part of the image and with the sections in the last slab at the right. Rather than using the number of white sections, we propose a more efficient alternative to measure the white region based on its area. We count the number of white pixels in background (i.e., the number of pixels that is not covered by parallelograms), instead of counting the number of white sections. The background in the document is the white paper, not painted with ink.

An efficient way to measure this area avoiding the parallelogram image construction is given by:

$$\textbf{if } nb \textbf{ then } wc = wc + ss, \tag{1}$$

where *nb* means that no black pixel is detected in the current section, *wc* is a counter which stores the number of white pixels and *ss* is the size of the section.

This strategy naturally gives a weight to a section proportionally to its size. A benefit of using the background area is that it is not necessary to care about fixing the total number of sections. This information is preserved by the total area (number of pixels) of the document at all angles. Thus, the entire image can be used to estimate the skew angle and no portion is ignored.

3.2. Reducing the search space

Chou et al.'s algorithm limited their skew angle search within the interval $[-15°,+15°]$ with no decrease in the accuracy rate as real scanned documents usually have small skew angles. To decrease time processing, instead of search through all angles within $-15°$ and $+15°$, Chou et al. proposed the following search strategy:

1. Search for the best skew angle β within $[-15°,+15°]$ with a step size of 2°;
2. Select the best skew angle γ, within the three angles $\beta - 1$, β and $\beta + 1$;
3. Finally, search for the best skew angle δ within $[\gamma - 1, \gamma + 1]$ with a step size of 0.1°.

The best skew angle at each step is the one that achieves the greatest number of white sections (in our approach, the largest white area) and δ is the final estimated skew angle of the document.

One alternative to speed up the first one of these three steps is to observe the variation of the white region size at each angle. This size should increase as the angle of the scan lines becomes closer to the real skew of the document and should decrease as the angle of the scan lines gets more distant from the real skew angle. However, this desired behavior is not observed when the number of white sections is used as a measure of the white region. On the other hand, the area of the white region is an option that has the expected behavior.

In Fig. 3, it is shown the behavior of the white region measured in the first step of the algorithm – when looking for the β angle from $-15°$ to $+15°$ with step size of 2°. Fig. 3(a) shows a scanned document with a skew angle close to 0° (some parts of the document are covered for privacy purposes). Fig. 3(b) shows the behavior of the white counter using the number of sections as proposed by Chou et al. It is possible to see that Chou et al.'s original method failed because the angles close to 15° and $-15°$ have greater values. This is caused by the size and by the variation in the number of sections, as discussed before. Fig. 3(c) shows the behavior of measuring the white region through the area, as it is proposed here. We can see that, close to 0°, the white counter assumes its greater value and decreases when the scan lines becomes more distant from it. Fig. 3(d) shows the curve of measuring the white area of the same document now skewed with an angle of 7°. We can see that the peak of the curve occurs exactly at 7°. Thus, the behavior of the curve can be helpful in early stopping the search for the best skew angle β. In other words, if we detect that the value of the white counter is decreasing, we can stop the search for the β angle and move to the next step. Another point to consider in real cases is that most of the scanned images have around 0° of skew angle. So, we propose to start our search evaluating 0° and then watching the behavior of the curve to continue in the direction that it grows.

$B(\theta)$ is defined as a measure of the white region (in our case: the background area). So, we propose a new alternative to speed up the search for the skew angle as follows:

1. Search for the best skew angle within 0°, 2°, $-2°$.
   if $B(0°) > B(-2°)$ and $B(0°) > B(+2°)$
   then assign 0° to β and move to the next step.
   else continue searching for the best skew angle β in the direction of the largest $B(\theta)(\theta = 2°$ or $\theta = -2°)$ with a step size of 2°.
   Stop when $B(\theta)$ begins to decrease;
2. Select the best skew angle γ, within the three angles $\beta - 1$, β and $\beta + 1$;
3. Finally, search for the best skew angle δ within $[\gamma - 0.6°, \gamma + 0.6°]$ with a step size of 0.1°.

Using this strategy, it is possible to note a reduction in the search space for the skew angle β. The best case occurs when $B(0°)$ is larger than $B(-2°)$ and larger than $B(2°)$, so β is set to 0°. Based on the fact that most of the document images in real problems has low skew

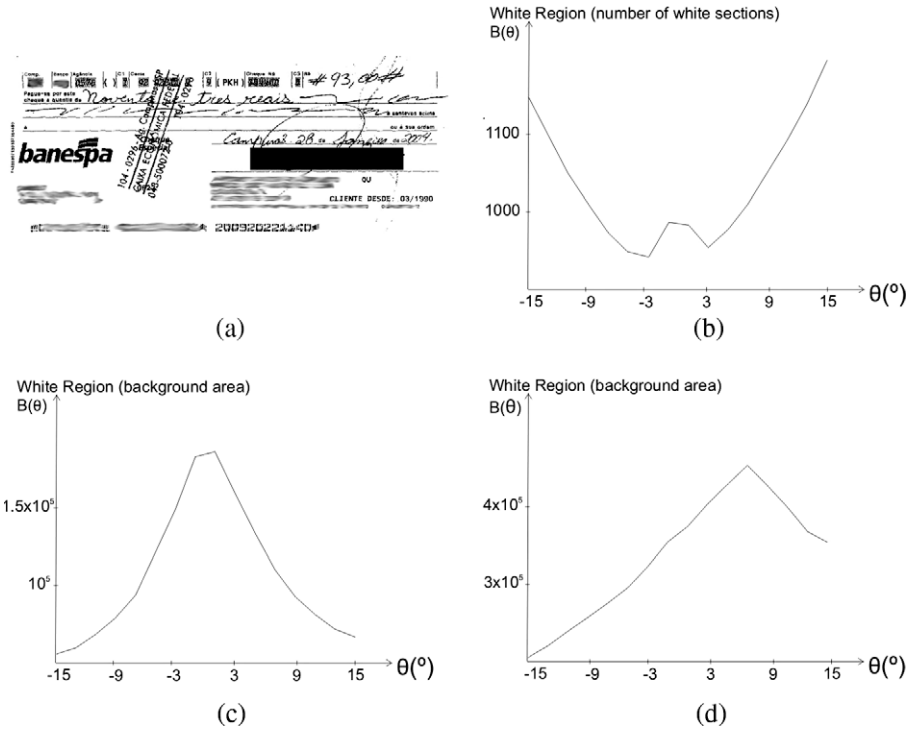

Fig. 3. (a) A scanned document with an angle close to 0°. Measuring the white region of the document with different skew angles using: (b) the number of white sections as proposed by Chou et al.; (c) the white area as we propose; and (d) the white area again but after a 7° rotation of the original image.

angles, the computation time decreases when compared to the original approach. The worst case occurs when β is +15° or −15°. Even in these cases, the search space for β should be reduced by half in comparison with the original search described by Chou et al., because only one path (positive or negative) is explored. The third step also reduces the search space: we only search for the skew angle from $\gamma - 0.6°$ to $\gamma + 0.6°$ instead of $\gamma - 1°$ to $\gamma + 1°$. This is done to avoid redundancy with the previous steps.

Another point to consider is that, in the first and second steps, the values of $B(\beta)$ and $B(\gamma)$ should be stored to be reused in the following steps; these values do not need to be recalculated. To avoid local minimum, we suggest to stop the search for the best skew angle β only when $B(\theta)$ decreases after two consecutive iterations. It is also important to mention that, when limiting the first step (search for β) within [−15°,+15°], the maximum range achieved is [−17°,+17°]. This is obtained because the second step searches for $\beta + 1$ and $\beta - 1$, and the third step searches for $\gamma - 1$ and $\gamma + 1$. If, for example, $\beta = -15°$ and $\gamma = -16°$, then δ can reach −17°.

3.3. Avoiding the interference of noise and vertical separators

The presented method is vulnerable to images with noise in the background. This happens because every section that has at least one black pixel is turned to gray in the construction of the parallelograms.

This vulnerability can also happen in documents with a large number of vertical separators (vertical bars), such as tables or forms. It is important to mention that the original idea of Chou et al. of using slabs was designed to work just over images with vertical bars. But it is easy to find a combination of vertical separators which may lead to failure (such as an image containing a vertical separator on each slab, that would lead to turn to gray all sections for all angles).

Based on that, instead of turning to gray the sections with more than one black pixel, we propose a smoother rule: to use a threshold T to decide whether the section should be turned to gray or not. The section should only be turned to gray if it contains a percentage of black pixels above this threshold. The percentage is calculated as:

$$T = nbp/ss, \tag{2}$$

where nbp is the number of black pixels in the section and ss is the size of the section. So, rewriting the Eq. (1) in Section 3.1:

$$\textbf{if } p > T \textbf{ then } wc = wc + ss, \tag{3}$$

where p is the percentage of black pixels in the section, T is the threshold, wc is a counter which stores the number of white pixels and ss is the size of the section.

Even in documents which have components that cannot be perfectly covered by parallelograms, the use of a threshold T brings benefits. This occurs in documents with non-textual elements, like stamps or handwritten components. After experimental study (shown further), T was set to 0.018.

Fig. 4 presents examples in which Chou et al.'s original method performed unsatisfactorily. Fig. 4(a) shows a bank check with noise, handwritten components and a stamp. Fig. 4(b) shows a document containing a large amount of vertical separators. The new proposal using the white area and the threshold estimated correctly the skew angle over these images.

4. Experimental studies

We evaluated our proposal over a collection of images. The original Chou et al.'s approach was implemented and tested for comparison. We computed the error of the skew angle estimation as the difference between the estimated angle and the target angle. This is called *estimation error*.

Chou et al. compared their technique with the following ones: a projection-based method (Postl, 1986), a maximum variance of transitions counts method (Ishitani, 1993) and a cross-correlations

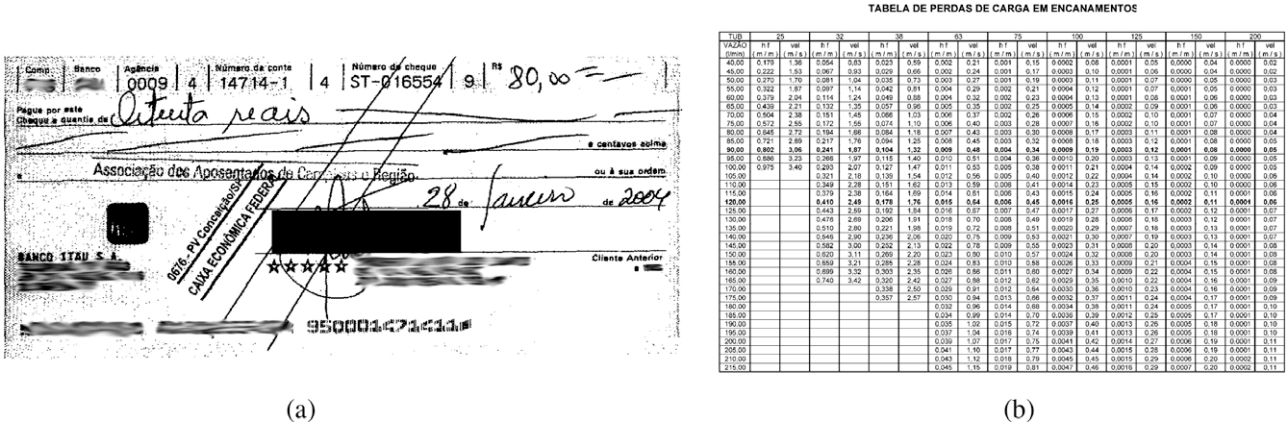

Fig. 4. Examples of documents with non-textual elements that can mislead the skew angle estimation proposed by Chou et al. (a) A Brazilian bank check with noise, handwritten components and a stamp (some information are hidden for identity preserving purposes). (b) A document containing vertical separators.

method (Avanindra, 1997). As conclusion, Chou et al. showed that their method achieved the best results in the estimation of the skew angle and also in the computational time. This section shows the comparison between our algorithm, Chou et al.'s original approach, Baird projection profile method (Baird, 1987), and Ávila and Lins' nearest neighbor clustering approach capable to detect both orientation and skew angle (Ávila and Lins, 2005). For all the experiments the slab width was set to 450 pixels for Chou et al.'s and our approaches. Previous experiments were presented in Mascaro and Cavalcanti (2008).

4.1. Databases

To examine the performance of the proposed method we constructed two disjoint databases with document images: one synthetic and one real. The first one was used to compute the estimation error of the skew angles and vary the threshold T to decide its best value. To ensure that each document in this database had 0° of skew angle, we collected images from electronic documents. After that, the images were artificially rotated with an angle within [−15°,+15°]. For this database, it was used a variety of documents composed by a mixture of textual and non-textual components (like forms, tables, pictures, etc.) with 200 dpi. The number of images in this database is 644.

In order to evaluate the difference in performance using the threshold described in Section 3.3, we simulate noisy images by adding salt and pepper noise in the first database. Fig. 5 illustrates a zooming into an image from this database. The error over images with the salt and pepper noise was computed apart.

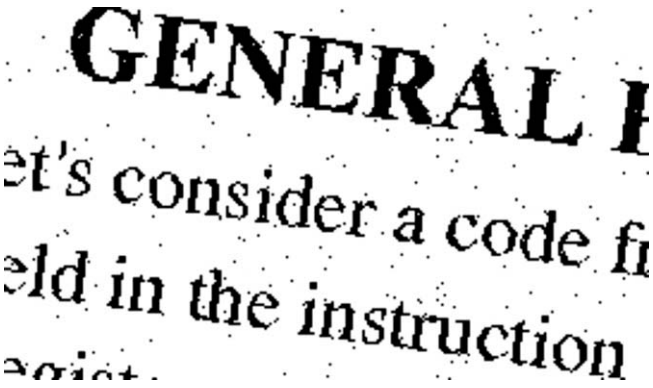

Fig. 5. Piece of a document from salt and pepper database with density 0.01.

The second database was used to evaluate the performance of the method over real scanned images. We collected 3,268 images including 2,744 Brazilian bank checks, 303 forms, 167 payrolls and 54 bank payment slips. Fig. 6 shows sample images that compound the second database. We can see that the images have very different layouts.

As it can be seen, the documents from the bank payment slip and forms datasets contain vertical separators and multiple font styles and sizes. And the images from the bank checks dataset contain handwritten components, stamps and much noise.

A third dataset was introduced: the database provided by Chou et al. (2007). It contains 500 images, which are generated by scanning a collection of different documents, with a resolution of 300 dpi. The database is divided into five categories and each document was artificially skewed at five different angles within [−15°,+15°], totalizing 100 images per category, as follows: (1) English documents with no figures, forms, or tables; (2) documents in Chinese or Japanese; (3) documents composed of horizontal text lines and large-scale figures; (4) documents composed of text lines and tabular regions; and (5) documents in several languages.

4.2. Results

Table 1 presents the results of estimating the skew angle of the first database which is formed by synthetic rotated documents. We varied the value of the threshold T and evaluated the error over this database. The column labeled as "Proposed" shows the results of our approach with $T = 0.018$ which generated the minimum average error. The other methods in comparison are presented in columns labeled as "Chou et al.", "Baird" and "Ávila–Lins".

Baird's projection profile method achieved the best result due to its high accuracy in dealing with documents containing clear separation between the text lines. The proposed method achieved second best result. The improvement over Chou et al.'s results occurred especially in images with a large amount of vertical separators. Although Ávila–Lins' algorithm also worked well with plain clean text, the reason for main errors is the orientation correction. For example, a document with 0° skew angle being evaluated as up-sided-down, result in an estimation error of 180°.

To choose the better value for the threshold T, we varied its value from $T = 0.010$ to $T = 0.050$ and measured the average error over the synthetic database. For $T = 0.010$ to $T = 0.018$ the average error over the database decreased, achieving a minimum of 0.0165° at $T = 0.018$ with a standard deviation of 0.0384°. For values above $T = 0.018$, the average error started to increase. For all following presented results, T was set to 0.018.

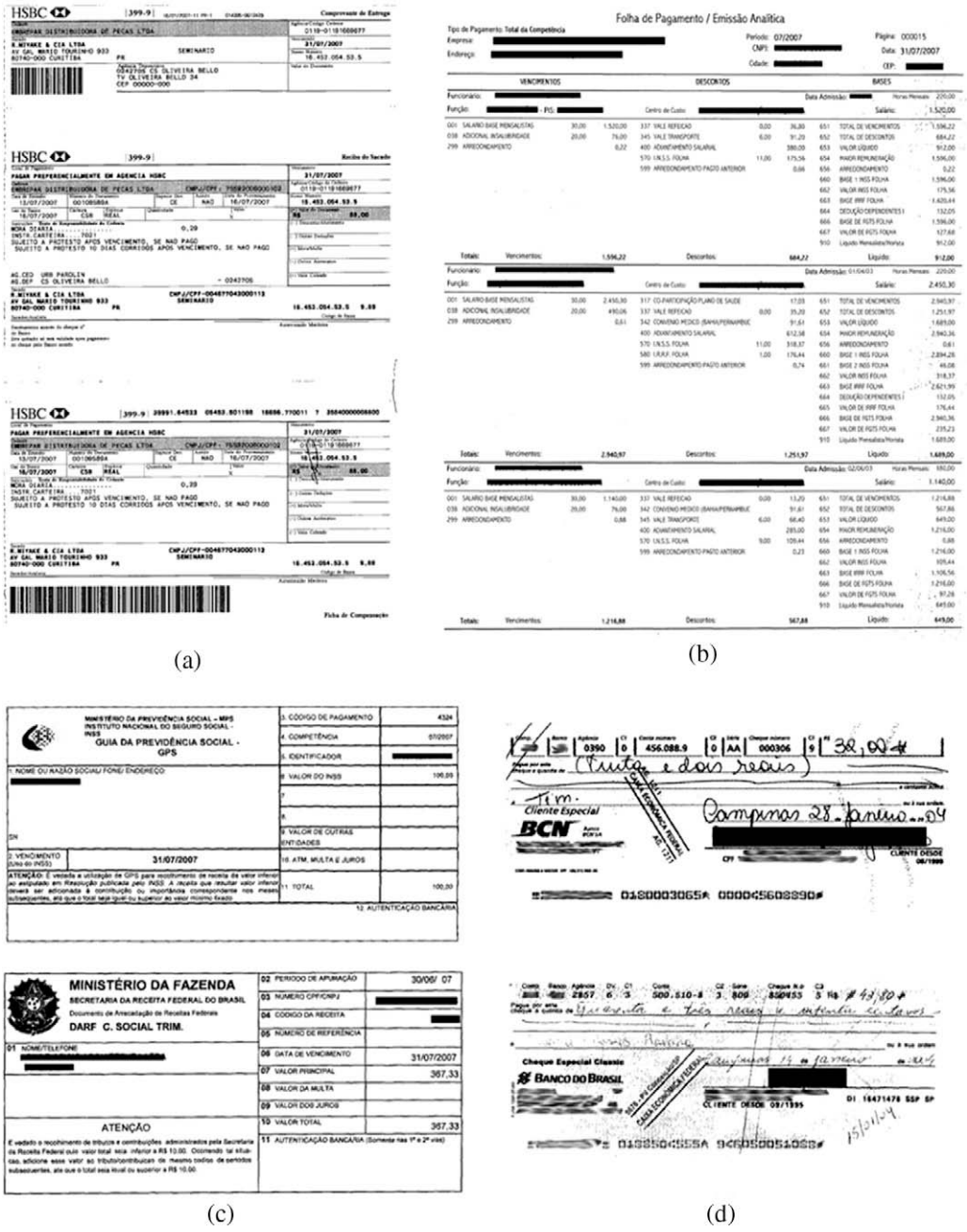

Fig. 6. Examples of database composed with real scanned images of documents. (a) A bank payment slip, (b) a payroll, (c) two examples of forms and (d) two Brazilian bank checks (data was hidden for identity preserving purposes).

Table 1
Error rates in degrees (°) from images of the synthetic database. $\bar{x}$ is the average error and s is the standard deviation.

	Estimation error ($\bar{x} \pm s$)			
	Proposed	Chou et al.	Baird	Ávila–Lins
Database 1	0.0165 ± 0.0384	1.381 ± 5.196	*0.0002 ± 0.004*	16.768 ± 52.350

Table 2
Error rates in degrees (°). Images from synthetic database (Database 1) with salt and pepper noise.

Density	Estimation error ($\bar{x} \pm s$)			
	Proposed	Chou et al.	Baird	Ávila–Lins
0.01	*0.0261 ± 0.0457*	5.2292 ± 6.7752	7.4738 ± 4.0803	92.5650 ± 27.3920
0.02	*0.1080 ± 0.7982*	12.3573 ± 7.4525	7.4550 ± 4.0220	15.9090 ± 7.4550
0.03	*0.2896 ± 1.6229*	15.9641 ± 8.5011	7.4620 ± 4.0560	7.4620 ± 4.0560

Aiming to verify the gain in performance using the threshold T we added salt and pepper noise with different densities in images of the synthetic database. As shown in Table 2, the error over Chou et al.'s approach increased, especially with larger density (d) noise. The noise causes Chou et al.'s algorithm to turn the sections to gray, misleading the analysis of the white sections. Baird's approach also greatly increased its average error. This happens because the analysis of such methods

Table 3
Difference to 0°. Images from the real database.

Category	Estimation error ($\bar{x} \pm s$)			
	Proposed	Chou et al.	Baird	Ávila–Lins
Bank payment slip	0.3291 ± 0.6716	0.3056 ± 0.7362	*0.3056 ± 0.7175*	70.1519 ± 88.2205
Form	*0.1079 ± 0.2340*	9.4488 ± 8.3749	0.1792 ± 0.4774	24.5109 ± 61.5573
Bank check	*0.1438 ± 0.2598*	15.7570 ± 4.3809	0.1681 ± 0.1914	106.6098 ± 88.2100
Payroll	0.1743 ± 0.1864	*0.1611 ± 0.1839*	0.1836 ± 0.1453	36.3384 ± 72.2529

rely in evaluating the text components, which are sensible to noise.

Conversely, the proposed approach continued to work fine. The use of the threshold T added robustness to the parallelogram method and the average error suffered a small increase. For all the densities, the proposed approach achieved the best results.

To analyze the skew correction in a practical environment, the new approach was tested over real scanned images. The images in this dataset present a skew close to 0°. Table 3 presents a quantitatively analysis of the results in this dataset: the values represent the difference to 0°. However, these values do not represent error rates, just an estimation of it, assuming that images are close to 0°.

In this manner, it is possible to compare the performance of each approach. Both proposed and Baird's approaches achieved very satisfactory results for all images in this database. However, the original Chou et al.'s algorithm got confused in most part of the skew angles. For the payroll and bank payment slip dataset, Chou et al.'s algorithm achieved satisfactory results. However, for bank checks and forms datasets, Chou et al.'s approach had a low performance in almost all images; Fig. 7 shows examples. For Ávila–Lins' algorithm, the orientation correction persisted as a problem.

The success of the proposed approach can be attributed to: (i) evaluation of the white region through background area; (ii) ability to deal not only with plain text, but also with: handwritten compo-

Fig. 7. Examples of the skew correction of real scanned images. (a) Scanned document of the category "Forms". (b) Result of applying the correction to the form based on the Chou et al.'s estimation – Chou et al.'s method detected −17° of skew. Our proposed approach detected 0° skew angle. (c) A scanned Brazilian bank check. (d) Result of applying the correction to bank check based on the Chou et al.'s estimation – Chou et al.'s method detected −17° of skew. Our proposed approach detected 0.3° skew angle.

Table 4
Error rates in degrees (°). Images from database provided by Chou et al.

Category	Estimation error ($\bar{x} \pm s$)			
	Proposed	Chou et al.	Baird	Ávila–Lins
#1	0.1700 ± 0.1314	*0.149 ± 0.129*	0.2420 ± 0.1689	0.2130 ± 0.1668
#2	*0.1070 ± 0.1365*	0.1390 ± 0.1430	0.1730 ± 0.1704	50.1693 ± 48.8462
#3	0.4890 ± 0.1456	*0.2310 ± 0.1350*	0.5480 ± 0.1337	0.5690 ± 0.1454
#4	0.1190 ± 0.1361	*0.1110 ± 0.1270*	0.1380 ± 0.1421	41.3349 ± 75.9839
#5	*0.0730 ± 0.0737*	0.0770 ± 0.0750	0.0831 ± 0.0830	64.8600 ± 86.7820

Table 5
Gain in computation time (in %) using the reduced search space when compared with the Chou et al.'s approach.

Database	Gain (%)
Database 1 (images synthetically rotated)	12.35
Database 2 (real scanned images)	51.30

nents, stamps, noise and vertical separators – due to the threshold. This happens because our new approach makes a better evaluation about the white space (the background) using the area to measure it. As the images contain components that are not plain text (handwritten components, stamps, noise, vertical separators, etc.), the proposal of using a threshold brought the ability to deal with this kind of features.

The last dataset to be evaluated is the database provided by Chou et al. The average errors over this database are presented in Table 4. It can be noticed that our new proposal and Chou et al.'s approach had very similar error rates. Baird's algorithm also showed satisfactory results. For Ávila–Lins' approach, the high error rate in categories #2, #4 and #5 is due to orientation failure. Various Chinese documents in category #2 were evaluated as being in landscape mode, i.e., for a 0° skew angle, the estimation error was 90°. For category #4 and #5 the errors were due to images resulting in up-side-down orientation.

4.3. Computational time

With the reduction in the search space proposed here, the computational time is also reduced. Table 5 shows the average difference (in %) in the computational time between our and Chou et al. approaches. The second row in Table 5 presents the results for the real database (images with skew angles close to 0°), composed with real scanned images. We can see that reducing the search space gave approximately 50% reduction in the computational time. It is important to mention that the Chou et al.'s method was faster than a projection-based method (Postl, 1986), a maximum variance of transitions counts method (Ishitani, 1993) and a cross-correlations method (Avanindra, 1997). Based on that, we can say that the approach presented here is faster than these techniques too. Both algorithms were implemented in MatLab by the same programmer with no optimization code. They were also tested in the same computer in equal conditions. The source code of our approach can be downloaded in the following site: http://www.cin.ufpe.br/~viisar.

5. Final remarks

Using parallelograms to fit the objects of a document is useful to estimate the skew angle of document images. In this paper we proposed a variation of Chou et al.'s method based on parallelogram covering. Comparing with Chou et al.'s approach, a projection profile method and a nearest neighbor method, through experimental tests over synthetic rotated images we showed that the proposal achieved better results over noisy images and documents containing vertical separators, like tables and forms. Through the real scanned images database we noted that our new approach can make a better evaluation of the background in the document and also works better with printed images containing handwritten components, noise and vertical separators.

We also showed an efficient procedure to reduce the search space, reducing the computational time of the algorithm. The proposed approach saves more time when the images have skew angles close to 0°, as it starts searching from this angle; the time consumption is proportional to the skew angle of the image.

The objective of this work was to deal with images containing a single angle along the image. It was not prepared to deal with images containing multiple skew along the document. In handwritten documents, it is possible to have multiple skew angles for each text line, especially when the person does not have a baseline. We leave this as future work to evolve the present approach to deal with this kind of skew.

References

Akiyama, T., Hagita, N., 1990. Automated entry system for printed documents. Pattern Recognition 23 (11), 1141–1154.

Amin, A., Fisher, S., 2000. A document skew detection method using the Hough transform. Pattern Anal. Appl. 3 (3), 243–253.

Avanindra, S., 1997. Robust detection of skew in document images. IEEE Trans. Image Process. 6 (2), 344–349.

Ávila, B., Lins, R., 2005. A fast orientation and skew detection algorithm for monochromatic document images. In: Proceedings of the ACM Symposium on Document Engineering, pp. 118–126.

Baird, H., 1987. The skew angle of printed documents. In: Proceedings of Society of Photographic Scientists and Engineers, pp. 21–24.

Cattoni, R., Coianiz, T., Messelodi, S., Modena, C.M., 1998. Geometric Layout Analysis Techniques for Document Image Understanding: A Review. ITC-IRST Technical Report #9703-09.

Chou, C., Chu, S., Chang, F., 2007. Estimation of skew angles for scanned documents based on piecewise covering by parallelograms. Pattern Recognition 40 (2), 443–455.

Ciardiello, G., Scafuro, G., Degrandi, M., Spada, M., Roccotelli, M., 1988. An experimental system for office document handling and text recognition. In: International Conference on Pattern Recognition, pp. 739–743.

Gatos, B., Papamarkos, N., Chamzas, C., 1997. Skew detection and text line position determination in digitized documents. Pattern Recognition 30 (9), 1505–1519.

Hinds, S., Fisher, J., D'Amato, D., 1990. A document skew detection method using run-length encoding and the Hough transform. In: International Conference on Pattern Recognition, pp. 464–468.

Hull, J., 1998. Document image skew detection: Survey and annotated bibliography. Document Analysis Systems II. World Scientific Pub. Co. Inc. pp. 40–64.

Ishitani, Y., 1993. Document skew detection based on local region complexity. In: International Conference on Document Analysis and Recognition, pp. 49–52.

Kanai, J., Bagdanov, A., 1998. Projection profile based skew estimation algorithm for JBIG compressed images. Internat. J. Doc. Anal. Recognition, 43–51.

Le, D.S., Thoma, G.R., Wechsler, H., 1994. Automated page orientation and skew angle detection for binary document images. Pattern Recognition 27 (10), 1325–1344.

Li, S., Shen, Q., Sun, J., 2007. Skew detection using wavelet decomposition and projection profile analysis. Pattern Recognition Lett. 28 (5), 555–562.

Lu, Y., Tan, C.L., 2003. Chamzas. Improved nearest neighbor based approach to accurate document skew estimation. In: International Conference on Document Analysis and Recognition, pp. 503–507.

Lu, S., Wang, J., Tan, C.L., 2007. Fast and accurate detection of document skew and orientation. In: International Conference on Document Analysis and Recognition, pp. 684–688.

Mascaro, A.A., Cavalcanti, G.D.C., 2008. Estimating the skew angle of document through background area information. In: Brazilian Symposium on Computer Graphics and Image Processing, pp. 87–94.

Min, Y., Cho, S.-B., Lee, T., 1996. A data reduction method for efficient document skew estimation based on Hough transformation. In: International Conference on Pattern Recognition, pp. 732–736.

Nicchiotti, G., Scagliola, C., 1999. Generalized projections: A tool for cursive handwriting normalization. In: International Conference on Document Analysis and Recognition, pp. 729–732.

Pal, U., Chaudhuri, B., 1996. An improved document skew angle estimation technique. Pattern Recognition Lett. 17 (8), 899–904.

Postl, W., 1986. Detection of linear oblique structures and skew scans in digitized documents. In: International Conference on Pattern Recognition, pp. 687–689.

Saragiotis, P., Papamarkos, N., 2008. Local skew correction in documents. Internat. J. Pattern Recognition Artif. Intell. 22 (4), 691–710.
Shivakumara, P., Kumar, G., 2006. A novel boundary growing approach for accurate skew estimation of binary document images. Pattern Recognition Lett. 27 (7), 791–801.
Singh, C., Bhatia, N., Kaur, A., 2008. Hough transform based fast skew detection and accurate skew correction methods. Pattern Recognition 41 (12), 3528–3546.
Smith, R., 1995. A simple and efficient skew detection algorithm via text row accumulation. In: International Conference on Document Analysis and Recognition, pp. 1145–1148.
Vailaya, A., Zhang, H., Yang, C., Liu, F., Jain, A., 2002. Automatic image orientation detection. IEEE Trans. Image Process. 11 (7), 746–755.
Yan, H., 1993. Skew correction of document images using interline cross-correlation. CVGIP. Graphical Models Image Process. 55 (6), 538–543.