Skewed Angle Detection in Text Images Using Orthogonal Angle View

Seongah CHIN and Moonwon CHOO
Department of Multimedia
Sungkyul University, KOREA
#147-2, Anyang 8-dong, Manan-gu, Anyang-city, Kyunggi-do, KOREA
schin@soback.kornet.net,mchoo@hana.sungkyul.ac.kr

Abstract: In this paper we propose skewed angle detection methods for images that contain text that is not aligned horizontally. In most images text areas are aligned along the horizontal axis, however there are many occasions when the text may be at a skewed angle (denoted by $0 < \theta \leq \pi$). In the work described, we adapt the Hough transform, Shadow and Threshold Projection methods to detect the skewed angle of text in an input image using the orthogonal angle view property. The results of this method are a primary text skewed angle, which allows us to rotate the original input image into an image with horizontally aligned text. This utilizes document image processing prior to the recognition stage.

1. Introduction

Images convey vast amounts of information and many images contain embedded or overlaid text. The text in this image may be as meaningful as the image itself. For coding or understanding document images it is essential to identify and possibly separate text, images and graphical regions into physically distinct segments of the page in order to be able to process them properly[1][2]. One of complexity of document processing is thought of as the skewed angle detection on embedded or overlaid text images since graphical or picture objects in images may affect to the skewed angle without regarding text skew.

Most existing page segmentation algorithms do not handle text images with skew [4],[5]. Yu and Jain designed a skewed detection algorithm using hierarchical Hough transform to connected components[6]. Reference[7], however, computes the orientation of text strings without using Hough transform whereas they calculate the angle by arctan computation on the centers of a pair of connected components.

Now, we proposed a skewed angle detection scheme focusing on embedded text images into other objects as well as document images. We note that text can be aligned any orientation with skewed angles. Highly accumulated angles might have angle distance 90°, orthogonal angle view made by vertical and horizontal stroke in text images. Orthogonal angle view sufficiently provides peak detection after mapping the image space into Hough space. We are able to rotate and translate input images in terms of new sizes and mainly skewed angle as shown in **Fig. 1**. In section2, we describe whole methodology. Orthogonal angle view property is introduced as feature extraction to make it possible the following technical procedures. As classification steps various one-dimensional functions are formulated and defined. Section 4 and 5 provide us experimental results and conclusion.

2. Methodology

The method is primarily based on recognition of orthogonal angle view in edge points from text images. As preprocessing stages, we adapt edge detection method to produce edge images, which provides mechanisms to find out orthogonal angle view. In order to explore candidate skewed angles, we investigate Hough transform mapping edge image space into Hough space. Shadow vector and Threshold Projection are introduced as significant procedures.

2.1 Orthogonal Angle View

Now we define skewed angle view as shown **Fig.1** where θ is the angle made with respect to the vertical x axis. Text in images may be aligned horizontally therefore we introduced the concept of angle view. This illustrates how text angles relate to angles detected as shown in **Fig. 2** The skewed angle θ frequently hinders us from extracting accurate text region in images as preprocessing steps to recognition.

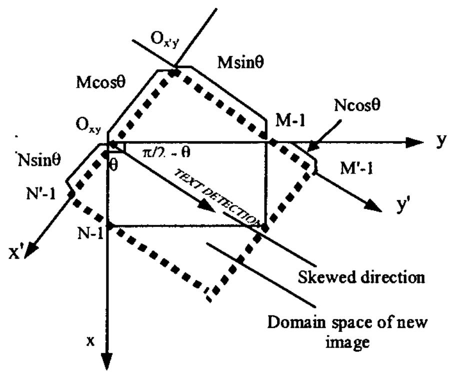


Fig. 1.
Relationship of the new image axes(x',y') to the original axes(x,y) and the position of the new origin $O_{x'y'}$ with the skewed angle θ

The main idea of the method is based on the principles of orthogonal angle view that text images almost carry. On the edge images, we can see high appearance of edge points on the strokes as shown in **Fig. 2.** Those cross stroke edge lines has orthogonal angle distance called Orthogonal Angle View. These features are mainly dealt as fundamental starting point to the next mechanisms. Even if we can see many straight lines at various angles in text images, we focus on identifying the angle views θ_a and θ_b

into in only two directions as shown in **Fig 2**. These angles allow us to transform our edge image properly.

Since the Hough transform not only extracts line segments but also detects non-continuous collinear points, which is not sufficient way to detect an accurate skewed angle, we should derive a way to detect only angles θ_a and θ_b which are orthogonal.

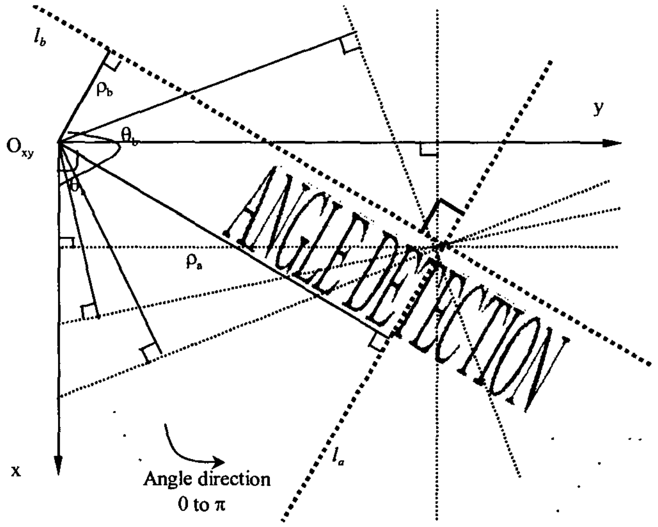


Fig. 2. Relation between Orthogonal Angle View and Text images

2.2 Various One dimensional function of the angle

Generally speaking, edges are defined as the boundary between two regions with relatively distinct gray levels. We obtain edge images using Prewitt edge mask and calculate the Hough matrix on the edge points.

The Hough transform is well known for detecting straight lines in various low level image processing applications. This method regards transforming each of the x- y image space points into a collinear points in a θ-ρ parameter space which is defined by the parametric representation, where θ is the angle made by x axis and ρ is the normal distance from the line to the origin of the image. All points at a straight line in image space can be represented by one point in which the angle θ and the distance ρ are called normal representation of a line. The Hough Transform maps edge points in the image space into an accumulator array representing $H(\theta, \rho)$ using the following mapping function [3],[8],[9]:

$$\rho = x\cos\theta + y\sin\theta \qquad (1)$$

As we realize edge points are highly accumulated along the strokes in text, these features can be seen as maximum number of occurrences on the Hough matrix along angle direction called Shadow Vector, $P_s(\theta)$ in (2). In order to identify the angle and distance, we note that the more frequencies at the angle, the more points at the straight line in image space. However we should be careful when there exist multiple peaks detected in an accumulator array since the main angle may not necessarily to be the maximum peak. Hence, we decide on a threshold for frequencies of angles, which allows us to retain the angles greater than the threshold in $H(\theta, \rho)$. The summation of the frequency over angles after thresholding called Threshold Projection Vector $P_\tau(\theta)$ (4) offers the solution of proper main text angle.

$$P_s(\theta) = \max_{\rho} H(\theta,\rho) \qquad (2)$$

$$P_p(\theta) = \sum_{\rho=1}^{\max\rho} H(\theta,\rho) \qquad (3)$$

$$P_\tau(\theta) = \sum_{\rho=1}^{\max\rho} u(H(\theta,\rho) - \tau_\theta)H(\theta,\rho) \qquad (4)$$

Where $1 \le \rho \le \max\rho = \sqrt{N^2 + M^2}$ for a fixed θ with image size N×M, τ_θ: threshold value and $H(\theta, \rho)$: Hough matrix of the image.
And

$$u(x) = \begin{cases} 1 & \text{if } (x \ge 0) \\ 0 & \text{otherwise} \end{cases}$$

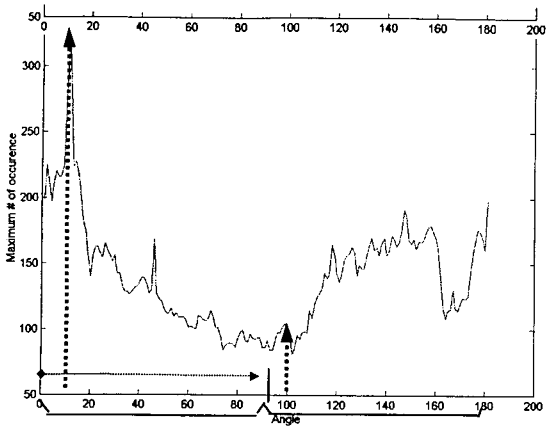
The final goal is to find the angle associated with a straight line which should be easily detected. We need to compute the summations along angles in the Hough matrix after thresholding on the number of occurrence in the Hough matrix. This implies that the angle density is high at the specific angles. In other words, we need to know which straight line contains a great number of image on text images. This line offers many intersection points at the specified angle. In order to extract useful angle from the Hough matrix, we need various (one dimensional) function of the angle(θ) derived from the Hough matrix. Some of the functions we use are (2), (3) and (4).

2.3 Peak Detection

We use an orthogonal angle view in text, which is useful for detecting peaks since orthogonal angles have a 90° difference between the two angles(θ_a and θ_b) as shown **Fig. 1**. To find the angles, we implement a peak detection mask as follows : The mask extracts peaks that have a distance of 90° by investigating the shadow $P_s(\theta)$ which displays only the maximum number of occurrences at angles. Sequentially, PeakMASK computes at angles (0+n-1) to (90+n-1) at the n[th] shift where n=1[st] up to 91[th]. In **Fig. 3(a)**, we see the mechanism to find the peak candidates by applying PeakMASK from 0 to π/2 on the maximum number of occurrence of each angle.

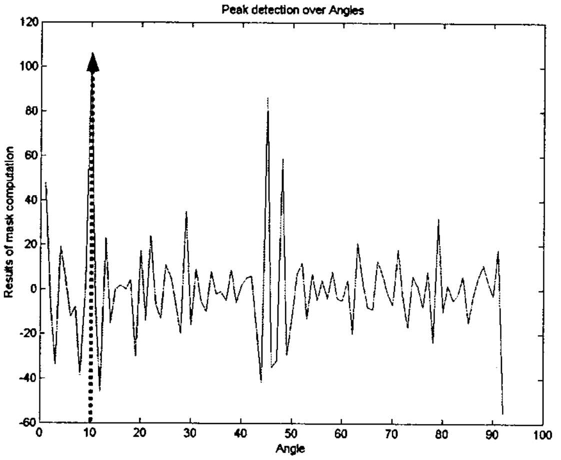
PeakMASK=[-1 2 −1 0 0… 0…0 -1 2 -1]^T (90)

Peak candidates are visible in **Fig. 3 (b)** at θ_p=10 which automatically implies another peak at $\theta_{p'}$ =100 due to the



1[st] time PeakMASK.................91[st] time PeakMask $P_s(\theta)$

(a)



Peak detected θ_p =10 (implies $\theta_{p'}$ =100)

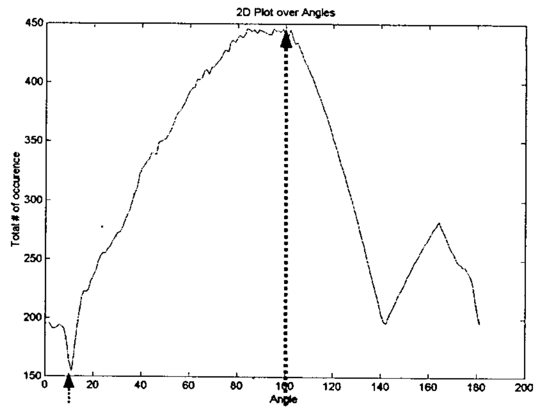
(b)

Fig. 3. Peak Detection
(a) Angles and maximum number of occurrence at each angle and use of PeakMask for input **Fig. 5**
(b) Peak detection

90° angle distance at PeakMASK. The next section introduces how do we determine the main skewed angle between the two peak candidates θ_p and $\theta_{p'}$.?

2.4 Angle Projection and the Skewed Angle

In this section we suggest the way to determine the main skewed angle between the two peaks obtained in the previous section. This can be done by counting the number of occurrence at both angles. The results of the counts are determined by threshold projection, $P_\tau(\theta)$ along angles at Hough matrix $H(\theta,\rho)$. Since we note that the maximum peak is not always necessarily the main angle, we have turned our attention to threshold projection . In other words, there is possibility that Hough matrix $H(\theta, \rho)$ contains many high values at the angle θ, but none of the angles is the maximum value. We then choose a threshold value τ_θ from $H(\theta, \rho)$ then compute the summation of only the number of occurrence of those values which is greater than the threshold along the angle θ, called the projection. We can correctly pick the peak over threshold projection $P_\tau(\theta)$, which gives us a proper mainly skewed angle. The threshold projection vector was defined in (4). By choosing



$\theta_p = 10$ $\theta_{p'} = 100$ $P_\tau(\theta)$

Fig. 4. Threshold Projection of Fig. 2
Peak candidates : $\theta_p = 10$, $\theta_{p'} = 100$

the threshold value, by the mean of the number of occurrences, this will detect angle θ_a since text lines hold many vertical strokes l_a as shown in **Fig. 2.** Higher values at the Hough matrix implies that more intersection points at sinusoidal curves indicating a line existence in the image space The higher peak between θ_p and $\theta_{p'}$ in the angle projection matrix $P_\tau(\theta)$ directly represents the main skewed angle in the image as shown **Fig. 4.** In other words, as we can compare $P_\tau(\theta_p)$ with $P_\tau(\theta_{p'})$, the angle which holds the higher number of occurrence becomes the mainly skewed angle. In **Fig. 4** we see $P_\tau(\theta_{p'} = 100)$ is higher than, $P_\tau(\theta_p = 10)$. Thus, our $\theta_s = 100$. We found the main skewed angle θ_s ,which is composed of the x axis and θ_a since text lines hold many vertical strokes l_a as shown in **Fig. 2.** Thus, we can directly transform edges images with respect to angle θ_s, without any changes, in order to perform rotation and translation properly.

3. Experimental Results

For evaluating the proposed method, we have carried out experiments on more than one hundred image files which are classified by three groups book covers; commercial bill boards, document images and road signs. Image files were captured from a scanner and most of the experiments were conducted in a MATLAB environment. In our experiments we found that 10 to 12 point size of text should be captured with a minimum of 400dpi (dot per inch) resolution since lower resolution cause significant data loss especially with small type fonts. This proposed methods successfully detect accurate skewed angle in any arbitrary angle. In **Fig. 6**, (a), (c) and (e) are skewed input images and (b), (d) and (f) rotated and extracted images applying transformations. **Table 1.** contains information of each image in terms of some parameters such size, resolution (DPI), skewed angle error denoted by e_s which can be computed by difference between the real skewed angle θ_r and the detected skewed angle θ_s.

4. Conclusion

The worked here presented is a novel skewed angle detection algorithm using orthogonal angle view property. This work is focusing on figuring out the skewed angle especially on embedded text images into any objects. The approach proposed here seems to be promising since we follow the principles to identify orthogonal angle view and have shown acceptable experimental results. This can utilize text segmentation, license plate detection, road sign detection and text image retrieval

Trials	*Size*	*DPI*	θ_s	θ_r	e_s
Fig.5 (a)	600X450	96	100	100	0
Fig. 6(a)	580X434	300	59	60	1
Fig. 6(b)	532X262	96	119	120	1
Fig. 6(c)	231X 233	300	45	47	2
Fig.6 (d)	443X444	96	70	70	0

Table1. data sets and its information

References

[1] Rasshid. Malik and Seongah CHIN, "Extraction of Text in images," IEEE International Conference On Information Intelligence, ICIIS99, pp. 534-537, Oct. 1999

[2] Kamran Etemad, David Doermann, and Rama Chellappa, "Multiscale Segmentation of unstructured Document Pages Using Soft Decision Integration," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 1, 19, No. 1, pp. 92-96, Jan. 1997

[3] Robert M. Haralick, Linda G. Shapiro, Computer and Robot Vision, Volume1, Addison Wesley,1992

[4] Anil K. Jain and B. Yu, "Document Representation and Its Application to Page Decomposition," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, pp. 294-308

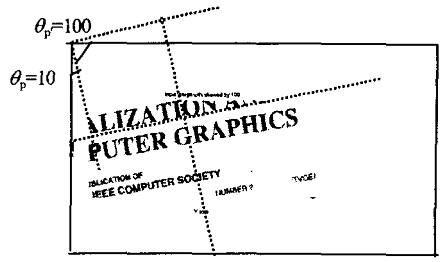
[5] H. Baird, "Anatomy of a Versatile Page Reader, " Proc. IEEE, Vol. 80, pp. 1,059-1,065, 1992

[6] B. Yu and A. Jain, "A Robust and Fast Skew Detection Algorithm for Generic Documents," Pattern Recognition, Vol. 29, pp. 1,599-1,629, 1996

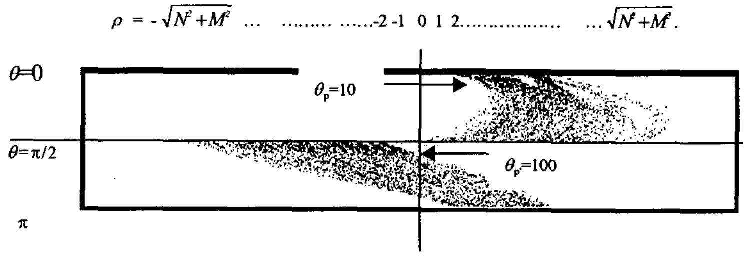
[7] C. L. TAN and P. O. NG, "Text Extraction Using Pyramid," Pattern Recognition, Vol. 31, No.1, pp. 63-72, 1998

[8] R. Duda and P. Hart "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Graphics and Image Processing, Vol. 15, No. 1, pp. 11-15, Jan. 1972

[9] Rafael C., Gonzalez Richard E. Woods, Digital Image Processing, Addison Wesley, 1993



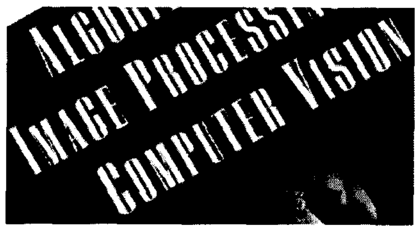
(a)



(b)

Fig. 5. Hough Matrix
(a) Input image N=450, M=600
(b) Hough Space Matrix generated by O'Gorman and Clowe's

(a)



(b)

(c)

(d)

Fig. 6. Sample experimental images