Available online at www.sciencedirect.com

ScienceDirect

Pattern Recognition Letters 28 (2007) 555–562

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Skew detection using wavelet decomposition and projection profile analysis

Shutao Li [a,*], Qinghua Shen [a], Jun Sun [b]

[a] College of Electrical and Information Engineering, Hunan University, Changsha 410082, China
[b] Fujitsu R&D Center Co., Ltd., Eagle Run Plaza B1003, Xiaoyun Road No. 26, Chaoyang District, Beijing 100084, China

Received 22 February 2006; received in revised form 22 September 2006
Available online 28 November 2006

Communicated by A.M. Alimi

Abstract

In this paper, a novel document skew detection algorithm based on wavelet decompositions and projection profile analysis is proposed. First, the skewed document images are decomposed by the wavelet transform. The matrix containing the absolute values of the horizontal sub-band coefficients, which preserves the text's horizontal structure, is then rotated through a range of angles. A projection profile is computed at each angle, and the angle that maximizes a criterion function is regarded as the skew angle. Experimental results show that this algorithm performs well on document images of various layouts and is also robust to different languages. The effects of various wavelet basis, number of decomposition levels, and parameters of the criterion function are investigated too.

Keywords: Skew detection; Document analysis; Projection profile analysis; Wavelet transform

1. Introduction

Document skew detection is necessary for most document analysis system and many methods have been developed. Existing methods typically use: (1) projection profiles analysis (Bloomberg and Kopec, 1993; Bloomberg et al., 1995; Ishitani, 1993; Liolios et al., 2002; Postl, 1986); (2) nearest neighbors (Jiang et al., 1999; Liolios et al., 2001; Lu and Tan, 2003); (3) Hough transform (Amin and Fischer, 2000; Yu and Jain, 1996; Ham et al., 1994); (4) mathematical morphology (Das and Chanda, 2001; Najman, 2004); (5) cross-correlations (Akiyama and Hagita, 1990; Yan, 1993; Chaudhuri and Chaudhuri, 1997; Chen and Ding, 1999; Gatos et al., 1997).

The traditional projection profile (PJ) based approach for skew detection was proposed by Postl (1986). First, the input document is rotated through a range of angles and a projection profile is calculated at each angle. Features are then extracted from each projection profile to determine the skew angle. This is computationally expensive as it is performed directly on the original document image. Moreover, it is sensitive to the layout of the document image.

An improved projection profile based approach was proposed by Bloomberg and Kopec (1993). The original document image is down-sampled before the projection profile is computed. The following operations are based on the sampled image. Therefore, the image data to be processed is reduced and the computational cost is reduced significantly. However, a major weakness is that its detection accuracy is influenced by the document image layout. It often fails on document images with multiple font styles/sizes or those that contain a large amount of non-text regions (such as pictures, tables or graphics).

The second class of the skew detection methods is based on the nearest neighbors (Jiang et al., 1999; Liolios et al., 2001; Lu and Tan, 2003). Here, the angle between each

* Corresponding author. Tel.: +86 731 8672916; fax: +86 731 8822224.
E-mail addresses: shutao_li@hnu.cn, shutao_li@yahoo.com.cn (S. Li).

0167-8655/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2006.10.002

connected component and its nearest neighbor is computed, and a histogram of the angles is formed to determine the skew angle.

The Hough transform has also been used for document image skew detection (Amin and Fischer, 2000; Yu and Jain, 1996; Ham et al., 1994). The idea is that the highest number of co-linear pixels is on lines that are co-incident with the baseline of the text. Its disadvantage is that one has to first extract text regions in the document image before the Hough transform is applied. This is not trivial for document images whose layouts are complex. Moreover, both the nearest neighborhood and Hough transform-based methods are computationally expensive.

On the other hand, mathematical morphology based methods represent the skewed document images with a more compact representation of its components, particularly text objects, such as words, lines, and paragraphs, where they can be represented by their rectilinear bounding boxes (Das and Chanda, 2001; Najman, 2004). This kind of algorithm is fast as well as independent of script forms. However, it is difficult to choose the morphological operators.

Cross-correlations (CC) (Akiyama and Hagita, 1990; Yan, 1993; Chaudhuri and Chaudhuri, 1997; Chen and Ding, 1999; Gatos et al., 1997) and the maximum variance of transition counts (TC) (Ishitani, 1993; Chen and Wang, 2000) are also used for document image skew detection.

Chou et al. proposed a skew estimation method based on piecewise covering of objects (PCP) such as textlines, figures, forms, and tables (Chou et al., 2007). First, the document image is divided into a number of non-overlapping slabs in which each object is covered by parallelograms. Then, the skew angle is estimated based on these parallelograms or, equivalently, their complementary regions. Chou et al. demonstrate that their method is faster and more robust than PJ, TC and CC.

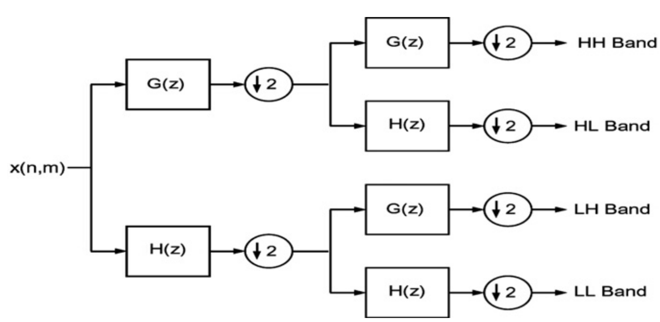

Fig. 1. Illustration of two-dimensional discrete wavelet transform.

Wavelet transforms have been widely used in image processing because of its inherent advantages such as sensitivity to abrupt variations, multi-resolution (Mallat, 1989; Daubechies, 1992). After one level of decomposition, the original image is decomposed into four sub-images which are approximations of the source image and details in the vertical, horizontal and diagonal directions.

In this paper, we present a document skew detection algorithm using wavelet decompositions and projection profile analysis. Using the wavelet decomposition, the horizontal structure of the document image can be extracted and emphasized, which will improve the accuracy of profile projection method. On the other hand, the computational cost is less than one tenth of that of the original after 2-level wavelet decomposition. Experimental results show that the proposed algorithm can deal with document images with complex layouts and multiple font size. It is also language independent. Moreover, there is no need to select text lines from the input document image before the projection profile is computed.

This paper is organized as follows. Section 2 introduces wavelet transform. Section 3 briefly describes the traditional projection profile based approach to skew estimation. In Section 4 the proposed algorithm is described in detail. Experimental results of the proposed algorithm are presented in Section 5. Discussions are given in Section 6. Finally, conclusion of this paper is presented in Section 7.

2. Two-dimensional wavelet decomposition

In our algorithm, a two-dimensional discrete wavelet transform (2D-DWT) is used. The roles of the highpass filters, low-pass filters and downsamplers in 2D-DWT are shown in Fig. 1. Here, $G(z)$ is a highpass filter, $H(z)$ is a lowpass filter, and "↓2" for downsampling. The lowpass

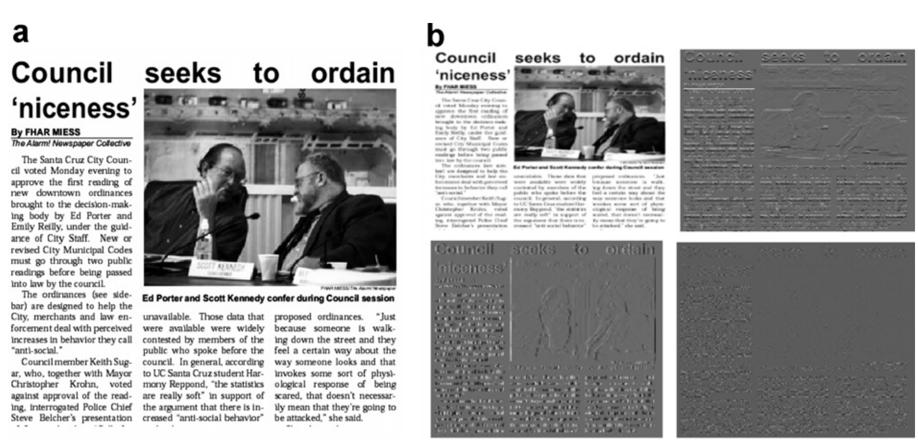

Fig. 2. An example of wavelet decomposition. (a) Original document image, (b) one-level decomposition using symlets wavelets.