



6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8
September 2016, Cochin, India

An Improved Method for Handwritten Document Analysis using Segmentation, Baseline Recognition and Writing Pressure Detection

Abhishek Bal^{a,*} and Rajib Saha^a

^aRCC Institute of Information Technology, Beliaghata, Kolkata, West Bengal 700015, India

Abstract

Handwritten document analysis is a scientific technique for identifying and understanding the personality of a writer through the strokes and patterns revealed by writer's handwriting. This research proposed an off-line handwritten document analysis through segmentation, skew recognition and writing pressure detection for cursive handwritten document. The proposed segmentation method is based on modified horizontal and vertical projection that can segment the text lines and words even if the presence of overlapped and multi-skewed text lines. Proposed work also present orthogonal projection based baseline recognition and normalization method as well as writing pressure detection method that can predict the personality of a writer from the baseline and writing pressure. The proposed method was tested on more than 550 text images of IAM database and sample handwriting image which are written by the different writer on the different background. The proposed method also provides a comparative study of the details analysis of the proposed method with other existing methods.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: Handwritten Document; Segmentation; Skew Recognition; Normalization; Projection; Pressure Detection; Personality Prediction.

1. Introduction

Handwritten document analysis is a demanding research area throughout the previous few years. Due to the cursive nature and high inconsistency of handwriting styles, handwriting analysis techniques should be more robust.

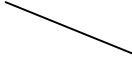
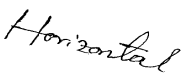

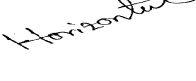

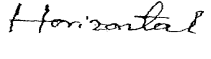
* Corresponding author. Tel.: +91-9038281735,+91-8296325464
E-mail address: abhisheknew1991@gmail.com

Generally, handwriting analysis for document image has four steps that are pre-processing^{2,7,8,9,31,32,33,34}, segmentation^{10,16,17,20,22,23,24,25,26,27,28}, feature extraction^{35,36} and classification^{35,36}. This paper proposed an off-line handwritten document analysis through segmentation, skew recognition and writing pressure detection for cursive handwritten document that can predict the personality of a writer from the baseline and writing pressure. The proposed segmentation method based on modified horizontal and vertical projection. To tolerate the text lines overlapping and multi-skewed text lines, proposed work modify present horizontal projection technique, which can segment the text lines and words even if text lines are overlapped with the help of vertical projection. In order to verify the stability of proposed word segmentation technique, the threshold which is used to differentiate between intra-word and inter-word gaps from each other is not fixed. If the threshold value is small then it may cause over segmentation whereas large threshold may cause under segmentation. Present work also proposed orthogonal projection¹⁴ based skew recognition, normalization and writing pressure detection that can predict the human personality.

Handwriting analysis is called Brain writing because a writer can control his or her mind about what he or she want to write but cannot control how to write. How to write fully operated from writer mind with the help of psychomotor. So, handwriting is a kind of mirror that can reflect the entire personality of a writer. Analysis of handwriting helps us understand the writer’s personality in a way which reveals his or her behaviour, motivations, desires, fears, emotional outlays and many other aspects. The most important handwriting features that can predict lots of personality trails are baseline, writing pressure, letter ‘t’, slant, the lower loop of letter ‘y’, speed of handwriting, size of letters, and others. This paper proposed a method that can predict the personality of a writer from the features that are extracted from his handwriting. The personality traits revealed by the baseline and the writing pressure of the handwriting as found in a writer handwriting are explored in this paper. The two parameters, the baseline and the writing pressure are the inputs to the proposed rule-based algorithm which outputs the personality traits of the writer based on these two parameters. The evaluation of the baseline using the orthogonal projection method and the evaluation of the writing pressure utilizes the grey-level value of actual handwriting portion are presented in this paper. These parameters from handwriting reveal a lot of accurate information about the writer.

Technically baseline or skew^{1,3,11,12,13,14} is defined as the alignment of the text lines and words with respect to the horizontal direction. The baseline in one’s handwriting reveals a lot of accurate information about the writer. Baseline in one’s handwriting is the line along which the writing flows. The three most common baselines found in any handwriting are ascending, descending and level as shown in Table 1. It is also true the baseline or skew tells specific personality trait of the writer such as pessimistic optimistic and level etc. Skew reveals the human’s or writer’s mentality^{35,36} about reaching his or her goals and the energy he or she applies to various situations.

Table 1. Personality trails represented by different baselines

Baseline Features	Baseline	Emotional Identification	Sample
Descending Baseline		Pessimistic	
Ascending Baseline		Optimistic	
Straight Baseline		Level	

Generally, machine printed document skew occurs at the time of scanning process due to the incorrect arrangement of the pages. Whereas skew^{3,11,12,13,14} in handwritten document can occur due to the human's behaviour as well as by the scanner during the scanning process. Generally, skew recognition in handwritten document is more difficult than the skew in printed document due to the variation of present mind condition of the writer and the difficulties at the time of scanning. In the case of the correct orientation of the pages, handwritten document still

Download English Version:

<https://daneshyari.com/en/article/571023>

Download Persian Version:

<https://daneshyari.com/article/571023>

[Daneshyari.com](https://daneshyari.com)