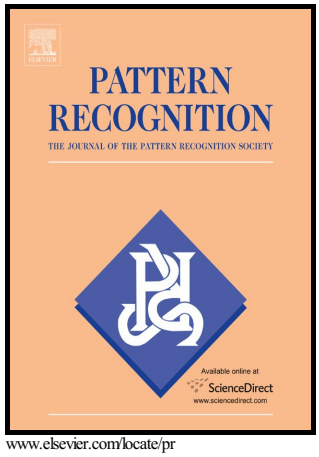
Writer identification using curvature-free features

Sheng He, Lambert Schomaker



www.elsevier.com/locate/pr

PII: S0031-3203(16)30305-3
DOI: http://dx.doi.org/10.1016/j.patcog.2016.09.044
Reference: PR5906

To appear in: *Pattern Recognition*

Received date: 12 February 2016
Revised date: 22 September 2016
Accepted date: 25 September 2016

Cite this article as: Sheng He and Lambert Schomaker, Writer identificatio using curvature-free features, *Pattern Recognitio* http://dx.doi.org/10.1016/j.patcog.2016.09.044

Writer identification using curvature-free features

Sheng He*, Lambert Schomaker

Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen, PO Box 407, 9700 AK, Groningen, The Netherlands

Abstract

Feature engineering takes a very important role in writer identification which has been widely studied in the literature. Previous works shown that the joint feature distribution of two properties can improve the performance. The joint feature distribution makes feature relationships explicit instead of roping that a trained classifier picks up a non-linear relation present in the data. In this paper, we propose two novel and curvature-free features: run-lengths of Local Binary Pattern (LBPruns) and Cloud Of Line Distribution (COLD) features for writer identification. The LBPruns is the joint distribution of the traditional run-length and local binary pattern (LBP) methods, which computes the run-lengths of local binary patterns on both binarized images and gray scale images. The COLD feature is the joint distribution of the relation between orientations and lengths of line segments obtained from writing contours in handwritten documents. Our proposed LBPruns and COLD are textural-based curvature-free features and capture the line information of handwritten texts instead of the curvature information. The combination of the LBPruns and COLD features provides a significant improvement on the CERUG data set, handwritten documents on which contain a large number of irregular-curvature strokes. The proposed features evaluated on other two widely used data sets (Firemaker and IAM) demonstrate promising results. Experimental results show that our proposed methods provide very good performance on irregular-curvature handwriting.

Keywords: writer identification, curvature-free, run-lengths of local binary pattern,

*Corresponding author

Email addresses: heshengxgd@gmail.com (Sheng He), L.Schomaker@ai.rug.nl (Lambert Schomaker)