Word Spotting Based on Bispace Similarity for Visual Information Retrieval in Handwritten Document Images

Ryma Ben Abdelaziz, Computer Science, Modeling, Optimization, and Electronic Systems Laboratory (LIMOSE), UMBB, Boumerdes, Algeria
Djamel Gaceb, Computer Science, Modeling, Optimization, and Electronic Systems Laboratory (LIMOSE), UMBB, Boumerdes, Algeria
Mohammed Haddad, Lab LIRIS, UMR CNRS 5205, University of Claude Bernard Lyon 1, F-69622, Villeurbanne, France

ABSTRACT

Retrieving information from a huge collection of ancient handwritten documents is important for indexing, interpreting, browsing, and searching documents in various domains. Word spotting approaches are widely used in this context but have several limitations related to the complex properties of handwriting. These can appear at several steps: interest point detection, description, and matching. This article proposes a new word spotting approach for the word retrieval in handwritten document, which mainly leverages the properties of image gradients for visual features detection and description. The proposed approach is based on the combination of spatial relationships with textural information to design a more accurate matching. The experimental results of the proposed approach demonstrate a higher performance over the Jeremy Bentham dataset, evaluated following the recent benchmarks of ICDAR 2015 Competition on Keyword Spotting for Handwritten Documents.

KEYWORDS

Features extraction, Gradient vectors, Handwritten documents, Image processing and analysis, Word spotting

DOI: 10.4018/IJCVIP.2019070103
1. INTRODUCTION

The ever-growing evolution in document digitization technologies serves several challenges one of which is handwriting recognition and word spotting. Word spotting consists of retrieving information from huge scanned document databases. The volume and complexity of such databases are making this task impossible manually (time-efficiency and cost-effectiveness). Many researchers around the world have tackled this challenge, and some of them are working on automated word-spotting systems. These systems are capable of extracting relevant and precise information from large and complex document databases. Content-Based Image Retrieval (CBIR) is an image retrieval technique (not specific to documents) used to retrieve information from large image databases (Ciciani, 1995). This technique relies on computing similarities between visual features (textures, colors, and shapes). CBIR is handy when it comes to performing a global search in a natural image, but becomes unfortunately very limited in case of a partial search (within the documents themselves). Therefore, other types of systems such as Document Image Retrieval Systems (DIRS) are more suitable for this type of images. They are based on two categories of techniques: techniques based on recognition (manual or automatic) and techniques based on word spotting. The recognition-based techniques are very accurate for recently written (modern handwriting), and good quality printed or handwritten documents, but lack effectiveness for old and degraded documents. Moreover, these techniques are often limited to a single language. Consequently, a handwriting recognition system for ancient documents or documents written in a rare language is very complicated to design and validate.

Afterwards, techniques of word spotting destined for documents have been proposed to overcome the problems resulting from recognition-based techniques. Our interest is for the word spotting techniques related to the interest point detection, description and matching on old and degraded handwritten document images.

The keypoints detectors and descriptors like SIFT (Lowe, 2004) and SURF (Bay, 2006) are widely used in various CBIR and DIRS applications due to their invariance to image scale and rotation and their robustness relatively efficient compared to older algorithms (corner, blob or region detectors). Several variants have been proposed in the literature (see section 2) to improve or adapt these methods to the specificities of certain application problems encountered. For each application, the performance of keypoint detection and description is compared for scale change, rotation change, blur change, illumination change, and/or affine transformations. The major drawback of these approaches (for detection or description phase) lies in the fact that the invariance to the change of illumination and colors is not taken into account (or taken into account in a partial and insufficient way in some variants). This disadvantage makes it difficult to apply these approaches to old and degraded handwritten documents with a very variable writing ink density (due to the use of the pen). For example, the SURF detector uses a fixed empirical threshold for keypoint filtering that is unrelated to the present local variations. This is a significant drawback on degraded handwritten,
Analysis and Comparison of Clustering Techniques for Chronic Kidney Disease With Genetic Algorithm

Computer Vision for Learning to Interact Socially with Humans
[www.igi-global.com/chapter/computer-vision-learning-interact-socially/73194?camid=4v1a](www.igi-global.com/chapter/computer-vision-learning-interact-socially/73194?camid=4v1a)

Online Handwritten Gujarati Word Recognition
[www.igi-global.com/article/online-handwritten-gujarati-word-recognition/224378?camid=4v1a](www.igi-global.com/article/online-handwritten-gujarati-word-recognition/224378?camid=4v1a)