International Journal of Computer Vision and Image Processing
Volume 6 • Issue 2 • July-December 2016

Scale Space Co-Occurrence HOG Features for Word Spotting in Handwritten Document Images

C. Thontadari, Department of Computer Science, Kuvempu University, Shimogga, India
C. J. Prabhakar, Department of Computer Science, Kuvempu University, Shimogga, India

ABSTRACT

In this paper, the authors proposed a Scale Space Co-occurrence Histograms of Oriented Gradients method (SS Co-HOG) for retrieving words from digitized handwritten documents. The poor performance of HOG based word spotting in handwritten documents is due to that HOG ignores spatial information of neighboring pixels whereas Co-HOG captures the spatial information of neighboring pixels through counting the occurrence of the gradient orientations of two or more neighboring pixels. The authors employed three scale parameter representation of an image and at each scale, they divide the word image into blocks and Co-HOG features are extracted from each block and finally concatenate them into form a feature descriptor. The proposed method is evaluated using precision and recall metrics through experimentation conducted on popular datasets such as IAM and GW and confirmed that their method outperforms for both the datasets.

KEYWORDS

Co-Occurrence, Dynamic Time Warping, Feature Descriptor, Feature Extraction, Histogram of Orientation Gradients, Optical Character Recognition, Scale Space

1. INTRODUCTION

Document Image Analysis (DIA) is one of the active research areas, which attracts research community due to its complexity and the increasing requirement for accessing the content of digitized contents. Optical Character Recognition (OCR) has been investigated for a few decades with huge achievement, which helps to automate the human process. OCR techniques usually recognize words by processing characters independently, works well with machine printed fonts against clean backgrounds. Generally, the large amount of document images is stored in digital libraries, and processing of these documents through the OCR system requires high computation cost because of complexity involved in understanding the page layout of documents, irregular writing styles, faded ink, stained paper and other undesirable factors. In order to overcome these problems, researchers have proposed a technique called word spotting. Word spotting technique is a moderately new alternative for character recognition and retrieval in both printed and handwritten documents.

Handwritten word spotting is the pattern classification task which retrieves words from the document images that are similar to the query word. Generally, a typical word spotting system consists of three main modules: preprocessing, features extraction and feature matching. Among them, features extraction is one of the most important factors for achieving high retrieval performance, because features with strong discriminative information can be well classified even using simplest classifier.

DOI: 10.4018/IJCVIP.2016070105

Copyright © 2016, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

The literature survey reveals that Histogram of Oriented Gradients (HOG) descriptor is widely used in several recognition applications because of its discriminating ability compared to other existing feature descriptors. The HOG feature descriptor is developed by Dalal, et al. (2005) for human detection using Support Vector Machine (SVM) classifier. The HOG has been successfully applied in many research fields such as word spotting task (Rodriguez et al., 2008; Terasawa et al., 2009), body parts detection (Corvee et al., 2010), face recognition (Deniz et al., 2011; Shu et al., 2011), character recognition (Newell et al., 2011), text/non-text classification problem (Minetto et al., 2013) and vehicle detection in traffic video (Arrospide et al., 2013).

Rodriguez et al. (2008) have proposed local gradient histogram features for word spotting in unconstrained handwritten documents. A sliding window moves from left to right over a word image. At each position, the window is subdivided into cells, and in each cell, histogram of orientations is accumulated. Slit style HOG features for handwritten document image word spotting is proposed by Terasawa et al. (2009). Newell et al. (2011) have extended the HOG descriptor to include features at multiple scales for character recognition. Saidani et al. (2015) have proposed a novel approach for Arabic and Latin script identification based on Histogram of Oriented Gradients feature descriptors. HOG is first applied at word level based on writing orientation analysis. Then, they are extended to word image partitions to capture fine and discriminating details. The unsupervised segmentation-free HOG based word spotting method was proposed by Almazan, et al. (2014). Documents are represented by a grid of HOG descriptors, and a sliding-window approach is used to locate the document regions that are most similar to the query.

The fundamental inspiration of Histogram of Oriented Gradients descriptor is that local object appearance and shape can often be characterized rather well by the distribution of the local intensity gradients or edge directions, even without particular information of the equivalent gradient or edge positions. The orientation analysis is robust to lighting changes since the histogram provides translational invariance. The HOG feature descriptor summarizes the distribution of measurements within the image regions. Compared with the SIFT feature descriptor, the HOG feature descriptor is evaluated over the whole image and does not require locations of suitable key-points. When extracting HOG features, the orientations of gradients are usually quantized into histogram bins and each bin has an orientation range. A histogram of oriented gradients falling into each bin is computed and then normalized to overcome the illumination variation. The orientation of gradients from all blocks are then concatenated together to form a feature descriptor of the whole image.

HOG feature descriptor captures orientation of only isolated pixels, whereas spatial information of neighboring pixels is ignored. Co-occurrence Histogram of Oriented Gradients (Co-HOG) (Watanabe et al., 2009) descriptor is an extension of the original HOG descriptor that captures the spatial information of neighboring pixels. Instead of counting the occurrence of the gradient orientation of a single pixel, gradient orientations of two or more neighboring pixels are considered. For each pixel in an image block, the gradient orientations of the pixel pair formed by its neighbor and itself are examined. Co-occurrence Histogram of Oriented Gradients is dominant descriptor widely used in object detection because Co-HOG features accurately represent significant characteristics of the object structure. At the same time, it is more efficient compared with HOG and therefore more suitable for real-time applications. The pedestrian detection method is proposed by Watanabe et al. (2009) based on extraction of the gradient orientation of neighboring pixel pairs. Ren et al. (2010) have proposed object detection method using Co-HOG features with variable location and variable size blocks which captures the characteristics of object structure. Face recognition using weighted Co-HOG features is proposed by Do (2012). The weighted voting Co-HOG features for scene text recognition are proposed by Tian et al. (2013). The character recognition in natural scenes using ConvCo-HOG features are proposed by Su et al. (2014) and multilingual scene character recognition using Co-HOG and ConvCo-HOG features are proposed by Tian et al. (2015).

In this paper, we proposed to extract Co-HOG features of the word image at multiple scales and these features are called as Scale Space Co-HOG (SS Co-HOG) features for word spotting in

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/scale-space-co-occurrence-hog-features-for-word-spotting-in-handwritten-document-images/171132

Related Content

Insect Recognition Using Sparse Coding and Decision Fusion
Hanqing Lu, Xinwen Hou, Cheng-Lin Liuand Xiaolin Chen (2018). *Computer Vision: Concepts, Methodologies, Tools, and Applications (pp. 1746-1767).*
www.irma-international.org/chapter/insect-recognition-using-sparse-coding-and-decision-fusion/197023

Breast Cancer Diagnosis System Based on Wavelet Analysis and Neural Networks
K. Taifi, S. Safi, M. Fakirand A. Elbalaoui (2014). *International Journal of Computer Vision and Image Processing (pp. 1-16).*
www.irma-international.org/article/breast-cancer-diagnosis-system-based-on-wavelet-analysis-and-neural-networks/111472

iVAS: An Interactive Visual Analytic System for Frequent Set Mining
Carson Kai-Sang Leungand Christopher Carmichael (2011). *Visual Analytics and Interactive Technologies: Data, Text and Web Mining Applications (pp. 213-231).*
www.irma-international.org/chapter/ivas-interactive-visual-analytic-system/48399

A Texture Features-Based Robust Facial Expression Recognition
Jayati Krishna Goswami, Sunita Jalal, Chetan Singh Negiand Anand Singh Jalal (2022). *International Journal of Computer Vision and Image Processing (pp. 1-15).*
www.irma-international.org/article/a-texture-features-based-robust-facial-expression-recognition/283963

Learning Aided Digital Image Compression Technique for Medical Application
Kandarpa Kumar Sarma (2016). *Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing (pp. 400-422).*
www.irma-international.org/chapter/learning-aided-digital-image-compression-technique-for-medical-application/141645