Optik 127 (2016) 6583–6589

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Local spiking pattern and its application to rotation- and illumination-invariant texture classification

Songlin Du, Yaping Yan, Yide Ma *

School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 11 January 2016
Accepted 6 April 2016

Keywords:
Local image description
Local spiking pattern (LSP)
Texture classification
Rotation-invariant
Illumination-invariant

ABSTRACT

Automatic classification of texture images is an important and challenging task in the applications of image analysis and scene understanding. In this paper, we focus on the problem of the classification of texture images acquired under various rotation and illumination conditions and propose a new local image descriptor which is named local spiking pattern (LSP). Specifically, the proposed LSP uses a 2-dimensional neural network, which is made up of a series of interconnected spiking neurons, to generate binary images by iteration. The binary images are then encoded to generate discriminative feature vectors. In classification phase, we use a nearest neighborhood classifier to achieve supervised classification. Finally, LSP is evaluated by comparison with some state-of-the-art local image descriptors. Experimental results on Outex texture database show that LSP outperforms most of the other local image descriptors in the noiseless case and shows high robustness when texture images are distorted by salt & pepper noise.

1. Introduction

With the rapid development of multimedia technologies, accurate description and automatic representation of the content of digital image have become essential problems in image analysis and scene understanding. It is widely accepted that human visual system (HVS) is highly adaptive to extract structural information for scene perception [1–3]. Inspired by this biological law, image processing and computer vision researchers have discovered that local image structure description is very important in many applications, such as texture classification [4], image fusion [5], image retrieval [6], image denoising [7,8], image quality assessment [9], face recognition [10,11], human action recognition [12], and moving object detection [13]. Although structural information is extremely important in most of the image processing applications, accurate description of local image structure is still an open problem [14–16]. In recent years, many local image structure description methods have been proposed, such as spatial statistics [17], transformed-domain statistics [18–20], edge preservation [21], sparse representation [22] and neural coding [23].

Of all existing image structure description methods, local binary pattern (LBP) and its variants [24–30], which are simple yet effective, are the most popular ones. In particular, the LBP operator assigns a label to every pixel of an image by thresholding the 3×3 neighborhoods of each pixel with the gray-level value of the center pixel and considers the result as a binary number [24,25]. The histogram of the labels are usually used for image structure description purposes. Besides LBP, a variety of LBP-based variants have also been constructed. Zhang et al. [26] proposed a local derivative pattern (LDP) by encoding the local derivative variations based directional pattern features, and applied the LDP to face recognition. Tan and Triggs [27]

* Corresponding author. Tel.: +86 09318912786.
E-mail address: ydma@lzu.edu.cn (Y. Ma).

http://dx.doi.org/10.1016/j.ijleo.2016.04.002

extended LBP to a local ternary pattern (LTP) by introducing a threshold, and found that the LTP is more discriminant and less sensitive to noise in uniform regions. Murala et al. [28] proposed to encode images with the direction of pixels that are calculated by horizontal and vertical derivatives, and constructed what is now known as local tetra pattern (LTrP). Fan and Hung [29] represented each pixel in an image as a vector by computing the gray-level values between the referenced pixel and the adjacent pixels with diverse distances from different directions, and proposed a local vector pattern (LVP). They experimentally demonstrated that the LVP can outperform LBP in face recognition. Nguyen et al. [30] introduced a support LBP model to establish the relationship among all of the pixels in local region, and also proved that the support LBP model improves the performance of the original LBP method. Undoubtedly, LBP and its variants have won tremendous success in the field of image structure description. Nevertheless, all of the descriptors mentioned above use either geometry or mathematical methods to improve description performance. In this work, we take a different path to construct a bio-inspired local image structure descriptor. In particular, our method is achieved by artificially modeling the neuronal mechanism of the mammal visual cortex, because we use an artificial neural network, which is made up of some interconnected spiking neurons, to generate binary images by iteration. The binary images are then encoded to generate discriminative feature vectors for rotation- and illumination-invariant texture classification.

The rest of this paper is organized as follows. In Section 2, the neuronal mechanism of the primary visual cortex, the details of the proposed LSP, and the feature extraction and classification methods, are described successively. The proposed approach is evaluated by experiments in Section 3. Finally, some conclusions are drawn in Section 4.

2. Methodology

2.1. Neuron model

Researches in neuroscience have discovered that: on one hand, if a neuron receives an external stimulus during its resting period, its membrane potential is mainly charged by the stimulus directly; on the other hand, its membrane potential is also modulated by the postsynaptic action potential from its neighborhood neurons [31]. The spiking cortical neuron model, which is evolved from the neuron model of pulse-coupled neural network (PCNN) [32], was designed by Zhan et al. [31] and is shown in Fig. 1, contains three main elements: internal activity $U_{i,j}$, dynamic threshold $E_{i,j}$, and output $Y_{i,j}$. The membrane potential of the spiking cortical neuron is calculated by combining of the external stimulus and the synaptic modulation, which is similar to the actual neuronal electrical activity.

By combining the external stimulus, i.e., $S_{i,j}$, and the outputs of the neighborhood neurons in last iteration, i.e., $Y_{k,l}[n-1]$, the internal activity of the neuron located at (i, j) in the current iteration can be described as

$$U_{i,j}[n] = S_{i,j} + \beta S_{i,j} V_U \sum_{k,l} w_{i,j,k,l} Y_{k,l}[n-1], \tag{1}$$

where (i, j) is the coordinate of the neuron, and n denotes the number of iterations. $w_{i,j,k,l}$ is a constant synaptic weight matrix which is inversely related to the distances between the central neuron (i, j) and its neighborhood neurons located at (k, l). β is a constant coefficient which adjusts linking strength. V_U is the initial amplitude of $U_{i,j}$. As mentioned above, $S_{i,j}$ is the external stimulus which can be regarded as the gray-level value of the pixel which is corresponded to the central neuron, and $Y_{k,l}[n-1]$ denotes the outputs of the neighborhood neurons in last iteration.

By combining the dynamic threshold of the neuron in last iteration, i.e., $E_{i,j}[n-1]$, and the output of the neuron in last iteration, i.e., $Y_{i,j}[n-1]$, the dynamic threshold of the neuron located at (i, j) in the current iteration can be described as

$$E_{i,j}[n] = e^{-\alpha_E} E_{i,j}[n-1] + V_E Y_{i,j}[n-1], \tag{2}$$

where n still denotes the number of iterations, V_E is the initial amplitude of $E_{i,j}$, and α_E is a decaying coefficient.

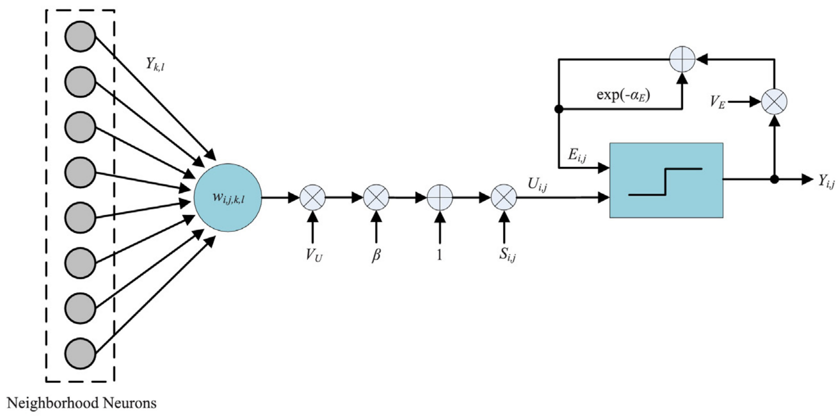

Fig. 1. The spiking cortical neuron model.