Neurocomputing 199 (2016) 77–89

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Principal curvatures based rotation invariant algorithms for efficient texture classification

Kazım Hanbay [a,*], Nuh Alpaslan [b], Muhammed Fatih Talu [b], Davut Hanbay [b]

[a] Bingol University, Department of Informatics, Turkey
[b] Inonu University, Department of Computer Engineering, Turkey

ARTICLE INFO

Article history:
Received 17 October 2015
Received in revised form
3 March 2016
Accepted 11 March 2016
Available online 25 March 2016

Keywords:
Feature extraction
Principal curvatures
Rotation invariance
Texture classification

ABSTRACT

The histograms of oriented gradients (HOG) and co-occurrence HOG (CoHOG) algorithms are simple and intuitive descriptors. However, the HOG and CoHOG algorithms based on gradient computation still have some shortcomings: they ignore meaningful textural properties and are unstable to noise. In this paper, two new efficient HOG and CoHOG methods are proposed. The proposed algorithms are based on the Gaussian derivative filters, and the feature vectors are obtained by means of principal curvatures. The feature vectors are rotation invariant by means of the rotation invariance characteristic of principal curvatures (i.e. eigenvalues). The experimental results on the CUReT, KTH-TIPS, KTH-TIPS2-a, UIUC, Brodatz album, Kylberg and Xu datasets confirm that the developed algorithms have higher classification rates than state-of-the-art texture classification methods. The classification results also demonstrate that the developed algorithms are more stable to noise and rotation than the original HOG and CoHOG algorithms.

1. Introduction

Texture is a considerable visual descriptor, used in analyzing and recognizing object surfaces [1]. Texture examination has a major role in many pattern recognition and image analysis applications such as remote sensing, object recognition, biometrics, plant image classification and content-based image retrieval [2,3]. Considering decades of research efforts on texture analysis, it remains a difficult problem and needs intensive research. Due to complexity of real world, analyzing texture images effectively is very difficult. For instance, a real world texture will have varied image variations such as translation, rotation, illumination change, scale change, viewpoint change and affine transformation. Thus, a robust texture classification method should not only capture highly discriminative information but also be robust to environmental conditions [4,5].

The demand of invariant texture classification has stimulated the research on this task. In order to adapt the undesirable environmental factors, numerous rotation invariant classification methods have been proposed [6–8]. Discrete wavelet transform (DWT) [9], complex wavelet transform (CWT) [10], discrete Fourier transform (DFT) [11] and cellular automaton (CA) [12] are extensively used in texture examination. To tackle the problem of classifying natural and synthetic texture images, a robust approach that combines concepts from CA and corrosion modeling to obtain a descriptor vector has been suggested [12]. Li et al. [13] recommended a rotation invariant feature descriptor which is robust to illumination and noise by utilizing Rapid-transform. Recently, the statistical-based methods have caught remarkable attention [14,15]. For instance, Ojala et al. [14] developed local binary pattern (LBP) method to extract the textural information. LBP is a powerful and effective method in describing local image patterns. Due to their good texture characterization, the LBPs were extensively used to other pattern recognition applications [16,17]. Multifractal models have proposed to improve robustness to local scale and orientation changes [18]. These methods successfully obtain texture information and are mathematically invariant to various complex transformations. Patch-based models provide an intuitive representation for texture classification [19]. They constitute a dictionary of active image patches. Textural features are calculated using this dictionary. This methodology can be used successfully in texture classification problems [19].

In general, there are two methods for rotation invariant texture classification: extracting rotation invariant features [20] and obtaining local rotation invariant features [14,15]. The first group methods are simple and easy to compute. They do not include complex algorithmic steps. Texton dictionary-based algorithms are also rotation and scale invariant [21,22]. But these algorithms are time-consuming and impractical methods, and they need enough training images to obtain texton dictionary.

In recent years, deep neural networks (DNNs) have attracted considerable attention on object detection and texture classification.

* Corresponding author.
E-mail address: kazimhanbay@gmail.com (K. Hanbay).

http://dx.doi.org/10.1016/j.neucom.2016.03.032
0925-2312/

DNNs include multiple levels of non-linear operations. They produce compositional models where the texture image is defined as layered composition of image features. Many methods were proposed for training deep models, and have been implemented successfully in the field of computer vision [23–26]. Ren et al. [23] developed a novel deep learning configuration to learn local visual features. A method that combines convolution neural network and mid-level features extracted from video sequences is developed in [24]. To construct the object appearance model, this method uses a hierarchical feature extraction schema. To detect the appearance variations, the likelihood function was used. Qi et al. [25] suggested a Convolutional Neural Network (ConvNet) as a feature extractor to calculate mid-level features from video sequences. Then first and second order statistics were calculated on by using the mid-level features. However, the computational cost of ConvNet was expensive due to size of feature vector and huge amounts of training data. Wang and Hu [27] proposed a new procedure for texture classification by obtaining a high level feature using DNN. They construct a semantic feature model to obtain a high level textural feature. For image recognition, Hee et al. [28] developed a deep residual learning framework. They reformulate the network layers as learning residual functions with reference to the layer inputs. Recently, Shuai et al. [29] have applied Convolutional Neural Networks (CNNs) to learn discriminative features and classifiers for local patch classification. They achieved superior labeling results by learning distinguishing features and classifiers to discriminate visually dissimilar image pixels. In the last years, many supervised learning methods have been proposed including shareable methods [30] and generative methods [31]. In these methods, layer features are calculated by a deep learning framework from a large-scale data. However, these methods adopt unsupervised learning methods to learn filters for feature extraction. There are several problems about the use of DNNs in texture classification, including the choice of training model for complex data like natural images, computation time and overfitting.

In order to describe objects effectively, gradient-based methods were proposed, which are edge orientation histogram HOG [32], co-occurrence HOG (CoHOG) [33]. Texture classification with HOG is considered more robust against illumination variations of objects and invariance to the affine transformations such as rotation, scale and position. CoHOG algorithm computes gradient orientations of neighboring pixel pairs and consequently obtains more spatial and contextual information, making it more capable to express the texture accurately and effectively. Thus, CoHOG can describe more varied shapes. Its effectiveness for object detection, human tracking and face detection has been presented in [33].

The detection of important pattern in textures is based on two basic concepts [34]: differential geometry and scale-space theory. The geometrical approach models utilize the assumption that basic patterns can be extracted from the texture pattern based on anisotropic variations of gray level pixel intensities. Recently, the differential methods have more concerns [35,36]. Gradient-based algorithms were developed in the last years. These algorithms are robust in obtaining inclusive textural context. However, the gradient computation cannot deal with different types of images, and it has low classification accuracy for noisy images. Principal curvatures are more robust differential method. They have sign and magnitude information. The principal curvatures are more effective in obtaining the discriminative points in pixel data than conventional gradient computation [21]. We have already developed principal curvatures-based methods in [5,37]. These methods are rotation invariant versions of HOG and CoHOG using Gaussian derivatives and principal curvatures from eigenvalues of Hessian matrix. Because of using principal curvatures information, the suggested HOG and CoHOG methods can cope with rotated textures. Our developed methods in [5,37] have only used magnitude of principal curvatures. However, these methods are more sensitive to noise. Instead of using the local magnitude of principal curvatures as a feature extractor, the information of two principal curvatures can be evaluated by using shape describer index described in [38]. This makes the developed HOG and CoHOG methods less sensitive to the rotation changes and robust to noise.

In this paper, we propose two new descriptors for the classification of texture images using a principal curvatures and surface classification approaches. We use the principal curvatures computation as the basis for extraction of textural features. The concavity and convexity of the image surface itself is used to classify different texture regions. A new formulation is used to calculate regions of concave and convex image surface. These two local descriptors are the counterparts of two state-of-the-art principal curvature based methods: Eig(Hess)-HOG and Eig(Hess)-CoHOG [5]. For the novel proposed methods, comprehensive noise experiments have been carried out on different datasets. Specifically, classification accuracies have been tested with datasets including both rotation invariance and noise invariance.

The rest of this paper is presented as follows: In Section 2, the original HOG and CoHOG algorithms were presented. The detailed expression of the developed algorithms was demonstrated in Section 3. In Section 4, accuracy of proposed algorithms was tested on seven texture datasets. In addition, a series of noisy and rotation analysis experiments were executed. Furthermore, the distinctive properties of proposed algorithms are demonstrated in a detailed way. The classification results with the well-known texture classification methods performed on Brodatz, CUReT, KTH-TIPS, KTH-TIPS2-a, Kylberg and XU datasets are given in Section 4. Conclusions are given in Section 5.

2. Related works

2.1. Histograms of oriented gradients (HOG)

The principal idea behind the HOG algorithm is to express an image as a group of local histograms. The HOG algorithm contains two computation units. One is the cell and the other is the block [4]. HOG descriptor vectors are calculated from all the overlapping blocks in a detection window. Fig. 1 demonstrates the relationship between cell and block components.

The size of the cell is $M \times N$ pixels and the size of the block is $2M \times 2N$ pixels. One block consists of four different cells. For each cell, histograms are calculated by gradient magnitude and orientation. During this process, grouping of magnitudes and orientations are carried out for better classification results [32]. Finally, the histogram of all cells and larger spatial blocks $n \times m$ are normalized by combining all histograms belonging to one block. The feature vector for the image is calculated by combining histograms from each block and the cell.

The number of orientation bins is called as N_o. Each histogram demonstrates a special region around the keypoint. Each of these regions corresponds to the cells of a $N_c \times N_c$ grid. We set the parameters of the descriptor to $N_c = 4$ cells for each spatial direction and $N_o = 8$ bins for each histogram. As a result, totally $N_o \times N_c^2 = 128$ elements are obtained in a HOG feature vector.

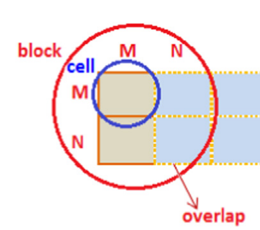

Fig. 1. Cells and blocks used in HOG algorithm.