Pattern Recognition 46 (2013) 2103–2116

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

Noise robust rotation invariant features for texture classification

Rouzbeh Maani [a,*], Sanjay Kalra [b,c], Yee-Hong Yang [a]

[a] Department of Computing Science, University of Alberta, Canada AB T6G 2E8
[b] Department of Medicine, University of Alberta, Canada
[c] Department of Biomedical Engineering, University of Alberta, Canada

ARTICLE INFO

Article history:
Received 29 June 2012
Received in revised form
20 November 2012
Accepted 9 January 2013
Available online 18 January 2013

Keywords:
Texture classification
Noise robust
Rotation invariant
Local frequency descriptors
Local binary patterns
FFT

ABSTRACT

This paper presents a novel, simple, yet powerful and robust method for rotation invariant texture classification. Like the Local Binary Patterns (LBP), the proposed method considers at each pixel a neighboring function defined on a circle of radius *R*. We define local frequency components as the magnitude of the coefficients of the 1D Fourier transform of the neighboring function. By applying different bandpass filters on the 2D Fourier transform of the local frequency components, we define our Local Frequency Descriptors (LFD). The LFD features are added dynamically from low frequencies to high. The features defined in this paper are invariant to rotation. As well, they are robust to noise. The experimental results on the Outex, CUReT, and KTH-TIPS datasets show that the proposed method outperforms state-of-the-art texture analysis methods. The results also show that the proposed method is very robust to noise.

1. Introduction

Texture classification is an important topic in image processing and has been used in many applications including automated inspection, image retrieval and medical image analysis. Image textures are defined as visual patterns appearing in images. Texture classification methods use chromatic and structural characteristics of images to characterize textures. The methods usually consist of four steps:

1. *Pre-processing:* The images are normalized in this step. The purpose of normalization is standardizing the intensity range, such that the extracted properties from the images are comparable.
2. *Feature extraction:* Textural features of images are extracted in this step. Different methods are used to find the textural features (e.g., statistical information, frequency analysis, etc.).
3. *Feature selection:* In this step, useful features are selected. Sometimes the number of features is huge. As well, some features may not be informative. The goal of this step is to reduce the number of extracted features by selecting those giving important textural information.
4. *Classification:* In this step, each image is assigned to one of the known texture classes. Basically, there are two sets: a training set and a test set. Classification is performed to assign images in the test set to one of the texture classes learned from the training set. Different classification methods can be used in this step. Some examples are support vector machines (SVMs), and nearest neighbor (NN) classifiers.

All methods referred to as texture analysis methods are used in the second step to extract textural features. The contribution of this paper is to present a new method for texture feature extraction based on the local frequencies in images. The proposed method provides features that are (1) invariant to rotation, (2) robust to noise, and (3) few in number.

In the next sections, this paper reviews relevant previous works (Section 2), presents our method (Section 3), explains the experimental results (Section 4), and concludes in Section 5.

2. Related works

There are many different methods for texture analysis; however, they can be categorized into four general groups. The first group uses statistical features. The main motivation behind these methods is based on the fact that the human visual system uses statistical features to distinguish textures. The co-occurrence matrix proposed by Haralick and Shanmugam [1] is one of the first known methods using this approach. The co-occurrence matrix represents the relationship between intensity levels for a given direction and distance in the image. More recently, the co-occurrence matrices have been extended to basic gray level

* Corresponding author. Tel.: +1 780 200 8812.
E-mail address: rmaani@ualberta.ca (R. Maani).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.01.014

Aura matrices by Qin and Yang [2,3]. The Run Length Matrices (RLM) [4,5] method defines a gray level run as consecutive pixels of the same gray level in a given direction, and the length of the runs is used to describe textures. There are other statistical methods such as using higher order statistics [6,7] and invariant moments of the images [8,9]. Recently, Local Binary Patterns (LBP) proposed by Ojala et al. [10] has been recognized as one of the most successful statistical methods and has been extended by different research groups [11–17]. The method represents the relationship of each pixel and its neighbors (located on a circle around the pixel) by a binary pattern and uses the histogram of these patterns for texture classification. LBP suffers from two major issues. First, the number of patterns increases exponentially with respect to the number of neighbors. Second, LBP and its variants are not robust to noise.

The second group of texture analysis methods uses structural features of images. These methods decompose textures into elements known as primitives or texels. The primitives and their spatial arrangements are used to characterize textures. For example, morphological operations are used to characterize textures [18]. Song's method [19] decomposes textures into a set of scale images, finds square texels of the same size at each scale, and uses the histogram of the texels as texture features. The method proposed by Gui et al. [20] extracts the size, position, periodicity, and spatial organization of texels to analyze textures. Khellah's method [21] uses the similarity between pixels and their surrounding neighbors within a predefined window and generates a global map called the dominant neighborhood structure. The features extracted from this map along with the features obtained from the LBP are used for texture classification. The key problem of the structural based methods is how to define texels that represent different texture structures. In general, the structural-based methods are better suited for textures with large structures (macrostructure) and do not work well on stochastic textures and microtextures [22].

The third class of texture methods defines textures as probability models. Some well-known models are Markov random field (MRF) [23,24], auto-regressive (AR) model [25,26], and Gibbs random field [27]. The key issue in these models is how to choose the correct model for a given texture and how to effectively map a texture into the selected probability model [22]. In addition, each model imposes some assumptions that may not be true for all textures. For instance, MRF assumes that the probability of each pixel depends only on its neighbors which may not be correct for all textures.

The fourth and last approach to analyze textures applies filters on images in either spatial or frequency domain. For instance, windowed Fourier filters are proposed by Azencott et al. [28] for texture classification. Gabor wavelet features are proposed by Manjunath and Ma [29]. The work of Chang and Kuo [30] employs tree-structured wavelet transform on textures to extract features. Rotation invariant textural features are extracted using Radon and translation-invariant wavelet transforms in the method proposed by Jafari-Khouzani and Soltanian-Zadeh [31]. The work of Chu and Chan [32] applies tunable Gabor filter banks to define rotation and scale invariant features. The method proposed by Haley and Manjunath [33] makes use of circular Gabor filters for texture classification. The main advantage of these methods that use frequency components is the capability of handling noise. However, the mentioned methods cannot capture local changes in textures. As a result some research studies use spatial domain to define textural features. The methods of Leung and Malik [34], Cula and Dana [35], and Varma and Zisserman [36] apply spatial filters to the textures, and use the histogram of cluster centers of the filter responses as features. Later, Varma and Zisserman [37] replace the filter responses with the local patches of the original image. Nonetheless, the spatial filter methods are not able to handle noise like their counterparts that use frequency information. Some methods use local frequencies of samples around pixels to capture the local changes. A popular approach is by taking the 1D Fourier transform on samples on a circle (or multiple circles) around pixels [38–40]. Any rotation makes a circular shift on the circular samples, keeping the magnitude of the frequency unchanged. The method of Arof and Deravi [38] uses two concentric circles around a pixel. The magnitude of the 1D Fourier transform of the samples and the difference of the samples with the center pixel are computed as features. A similar approach is used by Deng and Clausi [39] to construct anisotropic circular Gaussian MRF (ACGMRF) model. The method considers N concentric circular neighbors around a pixel and finds the parameters of the MRF using least squares estimation (LSE). Finally, it computes the magnitude of the 1D Fourier transform of the parameters to achieve rotation invariance. Recently, Liao and Chung [40] have proposed the composite Fourier domain (CFD) method. Considering samples located on three concentric circles around each pixel, the method computes the magnitude of the 1D Fourier transform on each circle. Then a global multidimensional Fourier transform is applied to form the composite Fourier domain. The null-space based linear discriminant analysis (nLDA) is used to construct the final features.

Our proposed method is inspired by the above discussed local Fourier analysis methods. We observed that none of these methods takes advantage of the local frequencies to make noise robust features. The development of such features is the main motivation and contribution of this paper. Among the mentioned methods, our method is similar to the CFD method of Liao and Chung [40]. We first take the 1D Fourier transform of the samples located on a circle around pixels. However, unlike CFD, we only use the low frequency components (not all). We observed that the low frequency components carry the major energy of the 1D signals. On the other hand, the high frequency components carry information that is very sensitive to noise. Therefore, by using the low frequency components, we not only reduce the number of frequency channels but also avoid the noise sensitive information. Similar to CFD, we use a 2D Fourier transform on the local frequency components. However, instead of using nLDA, we apply rotation invariant bandpass filters which provide noise robust features. In addition to the mentioned differences, a small image needs considerable space when transformed into the composite Fourier domain. As a result, the memory complexity of CFD makes the approach impractical for many applications. However, our proposed method does not require an excessive amount of memory which makes it practical for many applications.

3. The proposed method

Since the sampling method in our approach (as well as in some other works [38–40]) is similar to that in the LBP, we explain the relationship between our method and the LBP and compare the proposed method with the LBP and its variants. A brief overview of the LBP method and its variants is given in Section 3.1. The proposed method is introduced in Section 3.2. The implementation details are given in Section 3.3.

3.1. LBP and its variants

Traditionally, LBP considers N points on a circle with radius R at center pixel, t_c. These N points $(t_0, t_1, \ldots, t_{N-1})$ are the neighbors of the center pixel and their gray level values are determined by interpolation if they are not located at the center of pixels. Fig. 1 shows three popular configurations with radius of one, two,