J. Vis. Commun. Image R. 26 (2015) 305–316

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Texture classification and discrimination for region-based image retrieval

Mohsen Zand *, Shyamala Doraisamy, Alfian Abdul Halin, Mas Rina Mustaffa

Department of Multimedia, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 1 April 2014
Accepted 4 October 2014
Available online 20 October 2014

Keywords:
Region-based image retrieval
Texture feature extraction
Texture classification
Gabor wavelet
Curvelet filters
Polynomials
ImageCLEF
Outex

ABSTRACT

In RBIR, texture features are crucial in determining the class a region belongs to since they can overcome the limitations of color and shape features. Two robust approaches to model texture features are Gabor and curvelet features. Although both features are close to human visual perception, sufficient information needs to be extracted from their sub-bands for effective texture classification. Moreover, shape irregularity can be a problem since Gabor and curvelet transforms can only be applied on the regular shapes. In this paper, we propose an approach that uses both the Gabor wavelet and the curvelet transforms on the transferred regular shapes of the image regions. We also apply a fitting method to encode the sub-bands' information in the polynomial coefficients to create a texture feature vector with the maximum power of discrimination. Experiments on texture classification task with ImageCLEF and Outex databases demonstrate the effectiveness of the proposed approach.

1. Introduction

The rapid growth of image data on the internet has spurred the demand for methods and tools for efficient search and retrieval. Although many researches have been done in the field of image search and retrieval, there are still many challenging problems to be solved. As the semantic gap is considered to be the main issue, recent works have focused on semantic-based image retrieval. Most of the proposed approaches learn image semantics by extracting low-level features from entire image. However, such approaches fail to take into consideration the semantic concepts that occur in the images. In this paper, we focus on the high-level semantic identification at the region level. This is because analyzing the visual features included in the images gives more intuition about images. By learning these features at the region level, high-level semantics of images can be built. The approaches in which region information is employed to extract semantic concepts of images are known as region-based image retrieval or RBIR [1–3].

One issue in semantic understanding of image regions in RBIR is the extraction of effective and discriminatory features. Many researches have been done in global image features extraction and representation, but not much attention has been paid to region-based features [2–6]. Ideally, the extracted features must match the human perceptions of images. Most image retrieval systems apply three well-known color, shape and texture features. Color is the most common feature since it is invariant to distortion and scale. Although color feature is well-defined and widely used in image retrieval systems, it is unable to distinguish between different objects with the same color. Shape feature is not as important as other features in RBIR as regions' shapes are more vulnerable than regions' color and texture features [2,7].

Texture is an important determinant of region class in RBIR due to its capability to distinguish regions with similar colors and shapes. Although texture feature is very useful in RBIR, they are difficult to model. Ideally, a texture feature of an object should be consistent with human perceptual intuitions of the object, like directional/chaotic and smooth/rough [8].

Basically, texture analysis has four main categories: (1) texture feature, (2) texture discrimination, (3) texture classification, and (4) shape from texture. In this paper, we consider only texture discrimination and classification. Many different methods have been proposed for texture feature extraction. In general, these methods are categorized into spatial and spectral approaches [9,10].

The spatial approaches are further classified into structural, statistical and model-based approaches. In structural approaches such as Voronoi tessellation [11], texture feature is described using a set of texture primitives and their placement rules. Statistical texture

* Corresponding author.
E-mail addresses: zand.mohsen@gmail.com (M. Zand), shyamala@upm.edu.my (S. Doraisamy), alfian@upm.edu.my (A.A. Halin), masrina@upm.edu.my (M.R. Mustaffa).

http://dx.doi.org/10.1016/j.jvcir.2014.10.005
1047-3203/

features are usually based on low level statistics of grey level co-occurrence matrices (GLCM) [12,13]. Although these features are compact and robust, they are insufficient to describe a large variety of textures. Model-based approaches such as Markov random fields (MRF) [14] and fractal dimensions (FD) [15] apply stochastic (random) or generative models to describe texture features. As these models fall into an optimization problem, they usually need complex computations [16].

In spectral methods such as discrete cosine transform [18], Fourier transform [19], wavelet filters [20], Gabor [21] and curvelet features [22], texture images are transformed into the frequency domain using a set of spatial filters. Then, the statistics of the spectral information at different scales and orientations form the texture descriptor. Due to the large neighborhood support of the filters, spectral methods can generate sufficient number of features to classify variety of texture images.

However, the varying rotations of real-world textures suggest the need for rotation–invariant methods. Among the many spatial methods that can be considered, LBP [17] is the most widely used. It combines structural and statistical approaches by computing the occurrence histogram for rotation–invariant texture classification. In LBP, the values of neighboring pixels are turned into binary values using the central pixel as the threshold. This local binary gray-scale information is encoded to characterize a structural pattern. Although rotation invariance is achieved by only selecting rotation–invariant uniform local binary patterns, it is not scale-invariant. The LBP-based approaches often also fail in detecting large-scale textural structures. Many Gabor- and wavelet-based algorithms were also proposed for rotation–invariant texture classification [23–25]. Han and Ma [25] proposed to create texture features from a rotation–invariant and a scale-invariant Gabor representation by summations of the conventional Gabor filters. Recently, [2,26] presented a circular shifting of the curvelet texture features to generate rotation–invariant texture representations.

However, both the Gabor wavelet and the curvelet filters capture a large volume of unnecessary information which reduces their distinguishing power in texture classification. To overcome this issue, sub-band coefficients are produced in multiple orientations and scales and analyzed in the pre-processing step. In some earlier works, generalized Gaussian density was used to model wavelet coefficients [20,27,28]. In most of the texture extraction methods, texture feature vectors consist of statistical information, which are calculated from all sub-bands generated by applying either Gabor or curvelet transforms on a given image [23,29,30]. Zhang et al. [2] used mean and standard deviation to create a texture feature vector from curvelet sub-bands for each image region. Although both Gabor and curvelet transforms represent image texture features sufficiently by sub-band coefficients, using mean and standard deviation and other statistical information can lead to misclassifications as the discrimination power reduces. For instance, each image pair in Fig. 1 shows two different textures being classified into the same group as they have similar or slightly different means and standard deviations.

Mohamadzadeh and Farsi [7] used down-sampling to create a reduced size feature vector. However, random down-sampling increases the risk of losing the key information in the respective sub-band.

To overcome these problems, we propose the application of polynomial coefficients that are unique in representing data points. It is a combination of sub-band coefficients computed by the Gabor

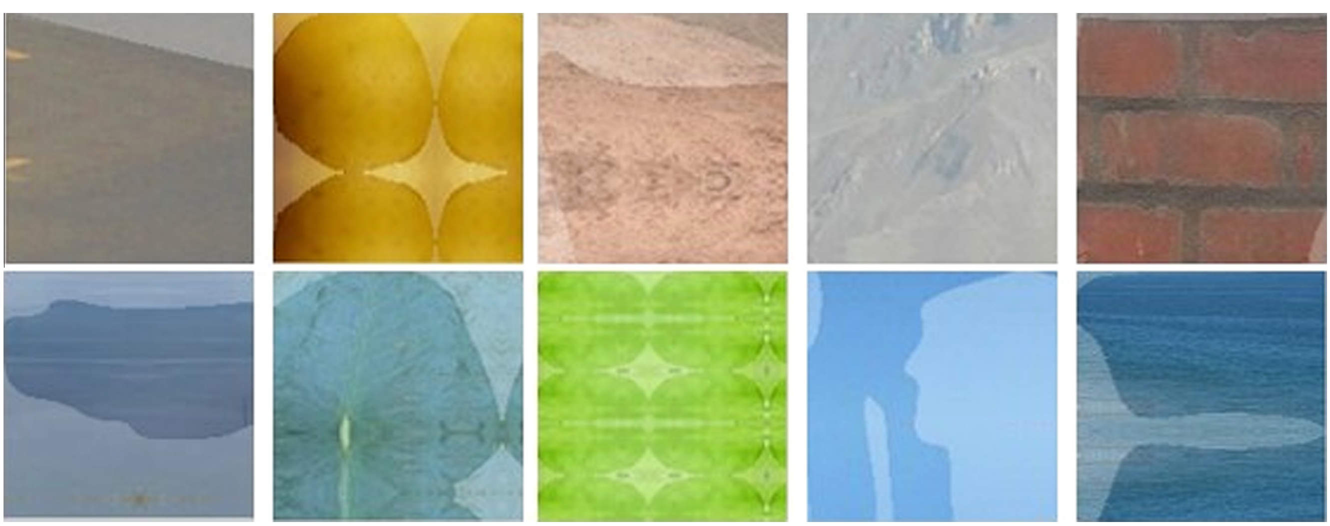

Fig. 1. Pairs of different textures with similar statistical values obtained from curvelet sub-bands, which lead to the misclassification.

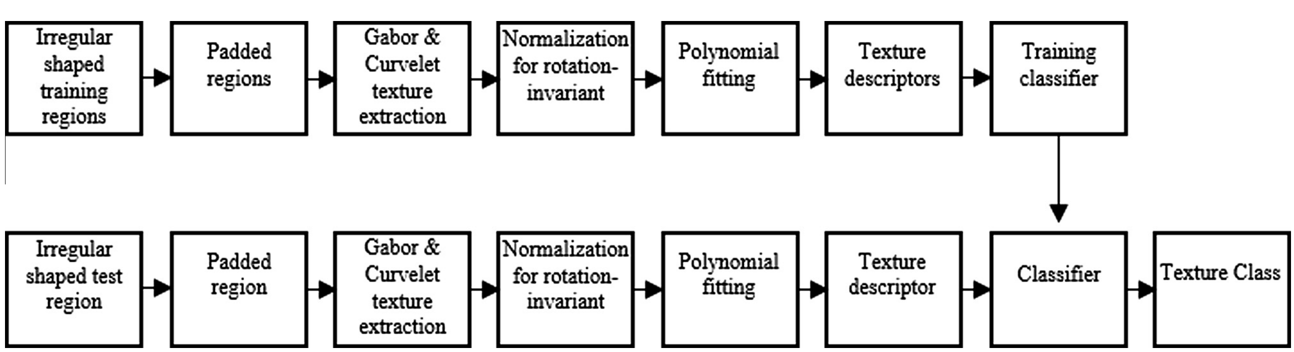

Fig. 2. The block diagram of the proposed method.