Pattern Recognition 64 (2017) 118–129

Contents lists available at ScienceDirect

Pattern Recognition

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

Color texture image retrieval based on Gaussian copula models of Gabor wavelets

Chaorong Li[a,c,*,1], Yuanyuan Huang[b,c], Lihong Zhu[a]

[a] Department of Computer Science and Information Engineering, Yibin University, Yibin 644000, China
[b] Department of Network Engineering, Chengdu University of Information Technology, Chengdu 610225, China
[c] University of Electronic Science and Technology of China, Chengdu 611731, China

ARTICLE INFO

Keywords:
Color Texture retrieval
Image representation
Copula model
Gabor wavelets

ABSTRACT

Color texture retrieval is a hot research area in image analysis. In this paper, we propose an efficient color texture retrieval method by using copula model based on Gabor wavelets. When Gabor wavelets are used to decompose color image, three types of dependence exist in the decomposed subbands of Gabor wavelets: color dependence, scale dependence and direction dependence. We analyze these dependencies and then capture them by using Gaussian copula function. Four copula schemes are developed, and accordingly four KLDs (Kullback-Leibler distances) of the copula schemes are introduced for color texture image retrieval. The evaluations of the proposed method are performed on several color texture databases including two large texture databases ALOT and STex. Experimental results demonstrate the proposed method has better performance than the state-of-the-art retrieval methods.

1. Introduction

Texture image retrieval have been broadly used in diverse applications of image processing such as multimedia storage system, multimedia web search, video query of accident investigation and digital library [1]. Image representation and feature matching are the two necessary processes in image retrieval. Image representation refers to the way that representing the image with discriminative information (called features) extracted from image. Feature matching refers to the comparison between two images with their features based on a similarity measure. Usually, image representation is the critical step because it directly affects the performance of retrieval method. During last two decades, a plethora of image representation methods have been developed such as wavelet based method [2], texton based method [3], dictionary learning [4] and local descriptor [5].

Wavelet transform and its extensions [6,2,7,8] are well-known tools for image processing. Wavelet transform represents an image $f(x)$ with different-resolution wavelet functions φ_n: $f(x) = \sum_{n=0}^{+\infty} c_n \varphi_n$. Ordinarily, we will get the approximation of the image by using finite optimal terms of c_n in terms of wavelet theory. The output subbands of wavelet are sparse (large number or the coefficients are close to zero) and the energies of wavelet subbands can be treated as the feature of texture image [9,10]. Besides the energie, the parameters of General Gaussian Model (GGM) [2,11] or Gaussian Mixture Model (GMM) [12] of the wavelet subbands are often used as the features of image since GGM / GMM is capable of well-fitting the coefficients of wavelet subbands.

According to the visual perception mechanism of human beings, texture image can be reviewed as an arrangement of textons (texture primitives) [13,14]. With the texton concept, researchers have put forward various of texton based methods for image analysis, such as texton co-occurrence matrix [15], multi-texton histogram [16], and MR8 texton [3]. Texton in fact describes the local structure information in the image. Similarly, Olshausen and Field [4] presented the concept of over-complete dictionary based on texture atom which is analogical to texton, and they pointed out that an image can be represented as a linear combination of atoms (ϕ_n) in the dictionary: $f(x) = \sum_{n=0}^{N} a_n \phi_n$ [17]. Thenceforward a large number of dictionary learning algorithms have been developed for image representation [18–20]. Although, dictionary learning shows powerful ability of image representation by adopting learning approach, it has much higher computational cost for training the dictionary and finding the optimal linear combination of atoms.

Another concept related to texton and texture atom is local descriptor such as LBP [21] and its extensions(CLBP [22], LBPV [23] and LTP [24]), BSIF-TOP [25], SIFT [26] and its extensions(PCA-SIFT [27], SURF [28] and ASIFT [29]), etc. LBP is a simple and quite

* Corresponding author at: Department of Computer Science and Information Engineering, Yibin University, Yibin 644000, China.
E-mail address: lichaorong88@163.com (C. Li).
[1] This is the specimen author footnote.

http://dx.doi.org/10.1016/j.patcog.2016.10.030
Received 18 September 2015; Received in revised form 5 September 2016; Accepted 27 October 2016
Available online 05 November 2016
0031-3203/

efficient local descriptor. It labels the pixels of an image by thresholding the neighborhood of each pixel (center pixel), then converts the thresholding results in the neighborhood into a binary number. Different from LBP which creates a binary number from a neighborhood, SIFT calculates the orientation histogram from a local region located around the keypoints detected by difference-of-Gaussian filter. In fact, SIFT is keypoint descriptor and more suitable for objects retrieval and recognition since it extracts the features from the local region around the keypoints but not each pixel of image.

Aforementioned methods focus on extracting the local structure characteristics (called local features) such as circle, square patch, or local small region in image. Up to now the effective approach for organizing the local features is to utilize the statistical information of local features. For example, researchers tend to use probability distribution to fit wavelet coefficients, use histogram technique to capture the feature distribution of LBP output, use linear combination of atoms to represent image in dictionary learning method. It should be pointed out that these representation methods combined with statistical technique, lose the spatial information (which is no less important than the statistical information) of local features in a texture image, but fortunately, they work out fine in most cases. For this reason, in this work, we also use probability distribution (copula model) to describe the image features extracted by Gabor wavelets.

As the reflection of real world, the color image captured by sensor such as camera is the most common. Therefore, it is meaningful to consider the color components when one design an image representation method. Yu et al [30] presented a color texture moments method by using local Fourier transform based on HSV color space, and derived eight characteristic maps for describing co-occurrence relations of image pixels in each color channel. CIE Lab color space has been employed by researchers because of its perceptual uniformity: Liapis and Tziritas [31] extracted the texture features in CIE Lab color space; Guang-Hai and Jing-Yu [32] proposed color difference histograms which count the color difference between two pixels in CIE Lab space. Color feature combing other features such as texture feature and shape information can achieve higher retrieval efficiency. Hiremath and Pujari [33] served the color moments and the moments on the Gabor filter responses as local descriptors of color and texture respectively; the shape information is captured in terms of edge images. Similarly, by considering the texture feature and color feature in color space, Cong Bai et al [34] developed the histogram of feature vectors which are constructed from the subbands of wavelet transform.

In most of the cases, image representation methods extract the color features at each color channels independently. In fact there exist dependencies caused by linear transform in the color space. In recent years, copula has been used to capture the color dependence and achieved success [35,36]. In this work we propose a texture retrieval method with copula and the Gabor wavelets [37] which consists of Gabor filters. Gabor filter can well model the cells of visual cortex of human, therefore, it is useful to generate the textons of image [38,39]. Our method not only considers the RGB color dependence but also captures the dependencies in the subbands of Gabor wavelets by using copula which separates a dependence structure among variables from its marginal distributions. According to structure characteristic of Gabor wavelets subbands in the RGB color space, we develop four copula schemes for the representation of the RGB texture image.

Regarding the feature matching, in this work, Kullback-Leibler Distance (KLD) is used. For the sake of computational efficiency we derive the closed-form KLD between two copula models by summing the closed-form KLD of marginal models and the closed-form KLD of copula functions. Compared with [35], which employs the DTCWT to decompose image and ML (Maximum Likelihood) to retrieval images, our method uses Gabor wavelets to decompose images, and uses KLD as the similarity of copula models. Our previous work also uses Gabor wavelets to extract rotation-invariant feature of grayscale image [40] by taking into account the scale dependence of Gabor wavelets. In this work we further study the other dependencies (such as directional dependence) of Gabor wavelets and color dependence in the RGB color space.

2. Background

2.1. Gabor wavelets

Gabor filters are linear filters used for signal analysis. The frequency and orientation representation of Gabor filters are resemblant to human visual system, and they have been found to be particularly useful for texture image representation. The function of Gabor filter is defined with a Gaussian kernel function modulated by multiplying a sinusoidal plane wave:

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} e^{j\omega x}, \tag{1}$$

where $\frac{1}{\sqrt{2\pi}\sigma}e^{-x^2}2\sigma^2$ is the Gaussian function and $e^{j\omega x}$ is the sinusoidal plane wave. The 2D form of Gabor filter is [37]

$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\pi\left[\frac{(x-x_0)^2}{2\sigma_x^2}+\frac{(y-y_0)^2}{2\sigma_y^2}\right]} e^{j[u_0x+v_0y]}, \tag{2}$$

where (x_0, y_0) is the center of spatial domain and (u_0, v_0) is the optimal spatial frequency of the filter in the frequency domain. σ_x and σ_y are the standard deviations of Gaussian function along x and y. In order to achieve the multiresolution and directionality of standard wavelet, Eq. (2) can be expressed as

$$p_{m,n} = a_{\max} f^{-m} p(x', y'), \tag{3}$$

where $x' = a_{\max} f^{-m}(x\cos\theta + y\sin\theta)$, $y' = a_{\max} f^{-m}(-x\sin\theta + y\cos\theta)$. $\theta = n\pi/K$ indicates the direction parameter; $n = 0, 1, \ldots, K-1$ and $m = 0, 1, \ldots, S-1$ (K and S are the numbers of directions and scales respectively); $a_{\max}$ indicates the maximum frequency and f indicates the spacing factor between kernels in the frequency domain. $p_{m,n}$ represents a serial of Gabor filters, and it is often referred to as Gabor wavelets.

2.2. Copula model for image representation

2.2.1. Copula model

Copula C: $[0, 1]^d \to [0, 1]$ is a type of multivariate distribution function [41] which has the following properties: (1) $C(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_d) = 0$, the copula is zero if one of the arguments is zero; (2) $C(1, \ldots, 1, u, 1, \ldots, 1) = u$, the copula is equal to u if one argument is u and all others 1; (3) C is d-increasing, i.e., for each $B = \prod_{i=1}^{d} [x_i, y_i] \subseteq [0, 1]^d$ the C-volume of B is non-negative $\int_B dC(u) \geq 0$. Copula is often used to capture the dependence structure of variables in which the dependence exists.

Sklar theorem [42] states that a multivariate Cumulative Distribution Function (cdf) $H(x_1, \ldots, x_n)$ of a random vector $x = [x_1, \ldots, x_n]$ can be expressed in terms of its marginal cdfs $F_1(x_1), \ldots, F_n(x_n)$ and a copula C. That is

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{4}$$

Conversely, given a copula C and the marginal cdfs $F_i(x_i)$ of random variables x_i, we can defined a multivariate distribution by using Eq. (4). Moreover, if the copula C is continuous and differentiable, then the density of C is given by

$$\frac{c(u_1, \ldots, u_n) = \partial^d C(u_1, \ldots, u_n)}{\partial u_1 \cdots \partial u_n}, \tag{5}$$

where $u_i = F_i(x_i)$. Then the Probability Density Function (pdf) $h(x_1, \ldots, x_n)$ is given as follows