Pattern Recognition 48 (2015) 2621–2632

Contents lists available at ScienceDirect

Pattern Recognition

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

Exploring space–frequency co-occurrences via local quantized patterns for texture representation

Tiecheng Song*, Hongliang Li, Fanman Meng, Qingbo Wu, Bing Luo

School of Electronic Engineering, University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, West Hi-Tech Zone, 611731 Chengdu, Sichuan, China

ARTICLE INFO

Article history:
Received 5 July 2013
Received in revised form
10 October 2014
Accepted 5 March 2015
Available online 13 March 2015

Keywords:
Texture classification
Image descriptor
Local Binary Pattern (LBP)
Textons
Local Fourier Transform (FT)

ABSTRACT

Local Binary Pattern (LBP) has shown its power in texture classification and face recognition. However, the LBP operator is performed in the original image space, and it lacks deeper pixel interactions to capture a richer description. In this paper, we propose to explore space–frequency co-occurrences via local quantized patterns for texture representation. The proposed method proceeds in two channels. In each channel, the multi-resolution spatial maps are first obtained by specific spatial filtering, and local frequency features (spectral maps) are subsequently extracted by applying the local Fourier transform to the spatial map. Two types of quantization via global thresholding are employed to quantize the spatial and spectral maps into three and two levels, respectively. The quantized patterns are then jointly encoded to construct a space–frequency co-occurrence histogram. Finally, the two-channel histograms are combined to characterize the texture. Without resort to the texton-based representation, our method directly encodes the joint information in the space and frequency domains while preserving the robustness to image rotation, illumination, scale and viewpoint changes. Extensive experiments have been conducted on three well-known texture databases, and our method achieves the best classification results compared with state-of-the-art approaches investigated.

1. Introduction

Image features play an important role in various computer vision tasks such as image matching [1], texture retrieval [2] and visual classification [3]. In the real world, images usually possess a variety of appearance changes such as rotation, scale and illumination variations due to different imaging conditions. Moreover, different classes of images often have very similar appearances. Therefore, designing an effective image representation that is robust for intra-class changes and discriminative for inter-class variations is a fundamental problem in many practical applications. In this paper, we focus on constructing an effective image representation for texture classification task.

Texture images have been widely studied in the past decades. Many approaches have been proposed for texture analysis including statistical, model-based and signal processing methods, such as co-occurrence matrices [4], Markov random fields [5], hidden Markov model [6] and image filtering [7]. Among these methods, Local Binary Pattern (LBP) [3] and texton-based methods [8–11] may be the most popular ones. The basic idea of LBP is to explore the local difference co-occurrences to extract pixel-wise feature. Due to its computational simplicity and invariance to illumination and rotation changes, LBP has received a lot of attention in texture classification [3] and face recognition [12]. The texton-based methods employ a texton dictionary learned from a set of filter responses (e.g., VZ_MR8 [9]) or from original image patches (e.g., VZ_Joint [10] and RP [11]) to summarize a texture. These methods are data-dependent owing to the texton dictionary learning, and they need intensive nearest-neighbor computation to encode each local image feature. Bypassing the texton-based representation, Weber Local Descriptor (WLD) [13] and Basic Image Features (BIFs) [14] were recently developed. The former is inspired by Weber's law, which constructs a 2-D histogram by encoding the differential excitations and orientations at certain pixel positions. The latter offers a natural quantization of filter responses into several distinct types of local image structure. It was reported that BIFs achieved the best classification performance on the KTH-TIPS database [15].

Motivated by LBP, a lot of variants have been recently proposed to encode local micro-structure information. Regarding local neighbor patterns, a diversity of geometric structures or topological patterns (e.g., the ellipse, disk, ring) were designed and evaluated in [16–19]. As for the encoded features, the high-order derivative direction information was explored both in Local Directional Pattern (LDP) [20] and Local Tetra Patterns (LTrPs) [21], while Gabor and monogenic features were exploited in [22–24] for face recognition. However, these methods either involve costly Gabor filtering, or were originally devised for face recognition in constrained environments. They are not suitable for texture classification under the uncontrolled imaging

* Corresponding author. Tel.: +86 13699426260; fax: +86 28 61830064.
E-mail address: tggwin@gmail.com (T. Song).

http://dx.doi.org/10.1016/j.patcog.2015.03.003
0031-3203/

conditions (such as rotation). Recently, the rotation invariant binary Gabor patterns (BGP) [25] were proposed for texture classification, which explore even- and odd-symmetric Gabor filters and the binarization operator to encode texture features. Following the line of patch-like features, average gray-values of blocks [26], self-similarity distances [27] and gradient magnitudes [28] were utilized for difference coding. To be robust to noise, Local Ternary Pattern (LTP) [29] quantizes the local difference into three levels and performs the split ternary coding on the quantized sequence. The self-adaptive quantization thresholds coupled with the N-nary coding were employed in Local Energy Pattern (LEP) [30] for material and dynamic texture classification. Completed LBP (CLBP) [31] was presented by decomposing the local difference into sign and magnitude components, and they are jointly encoded with the center pixel. In [32], Local Binary Count (LBC) was defined by only counting the 1's in the binary sequence. Dominant LBP (DLBP) [33] and Dominant Neighborhood Structure (DNS) [34] were, respectively, developed to combine local and global features to improve classification performance. Similar to LBP, circular neighborhoods centered at every pixel were considered in [35–39] to extract rotation invariant frequency features using the Fourier transform. Another recent trend is to extend LBP from 2-D plane to 3-D volume and from grayscale space to color space, such as Volume LBP (VLBP) [40], Uniform Spherical Region Descriptor (USRD) [41], Gabor Volume based LBP (GV-LBP) [42], Color LBP (CLBP) [43] and Local Color Vector Binary Patterns (LCVBP) [44]. A comprehensive study of LBP and its applications can be found in [45].

Despite its popularity, the basic LBP operator is performed in the original image space and it lacks deeper pixel interactions in different feature domains. In this paper, we propose to explore space–frequency co-occurrences via local quantized patterns for texture representation. The main contributions of this work include the following:

- Two-channel space–frequency feature spaces are designed to obtain locally invariant multi-resolution features.
- Two types of quantization via global thresholding are developed to partition the feature spaces. Our coarse quantization has three advantages: (i) it leads to the stable space partition; (ii) it enables the compact feature representation; (iii) it bypasses the texton-dictionary learning and involves no costly nearest-neighbor computation.
- The joint space–frequency coding is explored. Not only does this operator provide a richer and more discriminative description, but it also offsets the information loss caused by our coarse quantization.

With carefully designed space and frequency features, our method is robust to image rotation, illumination, scale and viewpoint changes. Combining the complementary space–frequency features in two channels, the discriminative power of our method is further enhanced. Therefore, our feature representation is expected to obtain good classification results.

The rest of the paper is structured as follows: Section 2 briefly reviews the principle of LBP and its recent variants. Section 3 elaborates our proposed method. Section 4 reports and analyzes the comparison results of our method against the state of the art, and Section 5 concludes this paper. A preliminary version of this work has appeared in [46].

2. LBP and its recent variants

2.1. Local Binary Pattern (LBP)

For a given pixel in an image, the LBP label [3] is defined as

$$\mathrm{LBP}_{P,R}=\sum_{p=0}^{P-1}s(g_p-g_c)2^p,\quad s(x)=\begin{cases}1, & x\geq 0\\ 0, & x<0\end{cases}\tag{1}$$

where g_c and g_p denote the intensity values of the center pixel and its sampling neighbors, respectively, P is the numbers of neighbors, and R is the sampling radius. Note that for sampling neighbors that fall out of the integer coordinates, their corresponding gray values can be obtained by bilinear interpolation.

The rotation invariant LBP [3] is obtained by

$$\mathrm{LBP}^{ri}_{P,R}=\min\{\mathrm{ROR}(\mathrm{LBP}_{P,R},i)|i=0,1,\ldots,P-1\}\tag{2}$$

where $\mathrm{ROR}(z,i)$ denotes the circular bit-wise shifting on the P-bit number z i times.

The uniform patterns [3] are further extensions of the original LBP operator. They were developed based on the empirical observation that uniform patterns can occupy over 90% among all the LBP patterns. A uniformity measure is defined by

$$U(\mathrm{LBP}_{P,R})=|s(g_{p-1}-g_c)-s(g_0-g_c)|+\sum_{p=1}^{P-1}|s(g_p-g_c)-s(g_{P-1}-g_c)|\tag{3}$$

The patterns that have U value of at most 2 are designated as "uniform", i.e., patterns with at most one 0–1 (or 1–0) transition in a circular bit string. The rotation invariant uniform LBP is defined as

$$\mathrm{LBP}^{riu2}_{P,R}=\begin{cases}\sum_{p=0}^{P-1}s(g_p-g_c), & U(\mathrm{LBP}_{P,R})\leq 2\\ P+1 & \text{otherwise}\end{cases}\tag{4}$$

The texture is finally represented as a histogram accumulated by the pattern labels of $\mathrm{LBP}^{riu2}_{P,R}$. That is, each rotation invariant uniform pattern has one separate histogram bin and all non-uniform patterns share a single bin. The resulting histogram has $K=P+2$ bins. Formally, for an image of size $W\times H$, the histogram is calculated as follows:

$$H(k)=\sum_{i=1}^{W}\sum_{j=1}^{H}\delta(\mathrm{LBP}^{riu2}_{P,R}(i,j)==k)\tag{5}$$

where $k\in\{0,\ldots,K-1\}$, K is the number of pattern labels, and

$$\delta(z)=\begin{cases}1 & \text{if } z \text{ is true}\\ 0 & \text{otherwise}\end{cases}\tag{6}$$

The multi-resolution LBP [3] representation can be achieved by combining multiple LBP operators via varying (P,R). In our work, we set $P=8$ when $R=1$; $P=16$ when $R=2$; and $P=24$ when $R=3$.

2.2. Completed LBP (CLBP) and Local Binary Count (LBC)

CLBP [31] was proposed to improve the discrimination power of LBP. As illustrated in Fig. 1, CLBP decomposes the local gray-level differences into two complementary components, i.e., the signs and the magnitudes. Accordingly, two operators, i.e., CLBP-Sign (CLBP_S) and CLBP-Magnitude (CLBP_M), are defined to encode the signs and magnitudes, respectively. The former is equivalent to the original LBP operator in (1), and the latter is defined as

$$\mathrm{CLBP\_M}_{P,R}=\sum_{p=0}^{P-1}t(m_p,c)2^p,\quad t(x,c)=\begin{cases}1, & x\geq c\\ 0, & x<c\end{cases}\tag{7}$$

where m_p is the magnitude of the difference between g_p and g_c, and c is set as the mean value of m_p from the whole image. Additionally, the center pixel is also encoded by the operator CLBP-Center (CLBP_C) as follows:

$$\mathrm{CLBP\_C}_{P,R}=t(g_c,c_I)\tag{8}$$

where c_I is the mean intensity values of pixels from the whole image.