Neurocomputing 182 (2016) 111–117

Contents lists available at ScienceDirect

Neurocomputing

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

Brief Papers

Completed local similarity pattern for color image recognition

Jing Li [a,b], Nong Sang [a,b,*], Changxin Gao [a,b]

[a] National Key Laboratory of Science and Technology on Multi-Spectral Information Processing, Wuhan, China
[b] School of Automation, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 7 June 2015
Received in revised form
19 October 2015
Accepted 2 December 2015
Communicated by Jinhui Tang
Available online 15 December 2015

Keywords:
Color image recognition
Local binary pattern
Color word
Completed local similarity pattern

ABSTRACT

In this paper, we propose a simple yet effective local color image descriptor, completed local similarity pattern (CLSP), for face recognition. We represent the color image as the co-occurrence of its image pixel color quantization information and the local color image textural information. Specifically, CLSP consists of two complementary components: color label and local similarity pattern. Expressed by soft color label for every image pixel, we adopt the k-means clustering method and the soft-assignment coding method to summarize and encode the image pixel color information globally ignoring local neighboring pixels. While based on the similarity of color information between central pixel and its neighbors, local similarity pattern (LSP) is used to encode the local spatial textural feature of the color image ignoring their color value. Therefore, LSP has the merits of robustness and compactness of coding based feature extraction method. The joint distribution (2-D histogram), which unifies the color label and LSP, is used for a given image representation. Moreover, seven different color spaces are selected and fused to compensate each other for color image feature extraction. Experimental results on GeorgiaTech, AR, NUST-RWFR and LFW face database show that CLSP outperforms state-of-the-art color image feature extraction methods.

1. Introduction

As one of the most focused research topics in image processing, pattern recognition and computer vision, face recognition has been widely applied in many fields, such as information security, smart cards, entertainment, law enforcement, video surveillance and human–computer interaction. Image feature extraction serves as one of the most critical steps for face recognition. Although numerous approaches have been proposed and tremendous progress has been made, during the past decades, it is still could not perform as well as desired under uncontrolled conditions. Therefore, how to extract discriminative and robust feature is of vital importance to face recognition.

Generally, the grayscale image feature extraction methods could be broadly summarized into two categories based on their properties, i.e., holistic methods and local methods. The holistic methods generally extract feature from a facial image by treating the image as a whole. Principal component analysis (PCA) [1,2], linear discrimination analysis (LDA) [3,4], independent component analysis (ICA) [5], locality preserving projection (LPP) [6,7], local linear embedding (LLE) [8], local discriminant embedding (LDE) [9], marginal fisher analysis (MFA) [10], discriminant simplex analysis (DSA) [11], nonnegative graph embedding (NGE) [12,13], clustering-guided sparse structural learning (CGSSL) [14] and robust structured subspace learning (RSSL) [15] etc. are the typical ones of this kind. These methods are liable to be influenced by face image pose, illumination, scale and so on, and variations in these factors can largely degrade their recognition performance. The local methods usually consider several regions or sets of isolated points, from which features for classification are extracted. Classical methods such as local binary pattern (LBP) [16,17], scale-invariant feature transform (SIFT) [18,19], speeded-up robust features (SURF) [20], Weber local descriptor (WLD) [21], Weber local binary pattern (WLBP) [22], monogenic binary coding (MBC) [23], histograms of local dominant orientation (HLDO) [24] and enhanced local directional pattern (ELDP) [25] and their variants have been widely examined. Compared with holistic methods, local methods are distinctive and invariant to many kinds of geometric and photometric transformations, and they have been gaining more and more attention because of their promising performance.

As one of the representative local image descriptors, local binary pattern (LBP) was first introduced by Ojala et al. [16], and it has shown a high discriminative ability for texture classification due to its invariance to monotonic gray level changes. After that, many variants of LBP have been introduced to further improve its

* Corresponding author at: National Key Laboratory of Science and Technology on Multi-Spectral Information Processing, Wuhan, China and School of Automation, Huazhong University of Science and Technology, Wuhan, China.
E-mail addresses: hustlijing@hust.edu.cn (J. Li), nsang@hust.edu.cn (N. Sang), cgao@hust.edu.cn (C. Gao).

http://dx.doi.org/10.1016/j.neucom.2015.12.005
0925-2312/

performance. Heikkilä et al. [26] proposed center-symmetric LBP (CS-LBP), which is different to LBP in that it compares center symmetric pairs of pixels against a center pixel, rather than comparing each pixel with the center. This halves the number of bits of binary patterns for the same number of neighbors. Guo et al. [27] proposed the completed LBP (CLBP), which includes the information contained in the center pixel and magnitudes of local differences as complementary to the signs used by the original LBP. Considering that high frequency texture regions have high higher variances and do more contribution to discriminate image, Guo et al. [28] proposed variance weighted LBP (VLBP) histogram image representation method rather than quantized the variance into N bins. Tan and Triggs [29] proposed local ternary pattern (LTP), which quantizes the image pixel gray-level differences between center and neighbor pixel into a ternary value instead of the binary one. Huang et al. [30] proposed local circular pattern (LCP), which improves the LBP and its variants by replacing the binary quantization with a clustering method, resulting in higher discriminative ability as well as better robustness to noise.

Existing works have demonstrated that color information can provide more discriminant information for image representation. Much work has been attempted to extend the original LBP descriptor to the color image representation. The most adopted approach, termed CaLBP, first applies LBP operator independently on each of the different color spectra, then LBP histogram of each spectrum is concatenated to form final image feature vector [31]. As such, it is limited to encoding the texture patterns of only color pixel variations derived from each individual spectral-band image. Mäenpää et al. [32] proposed multispectral LBP (MSLBP), which uses monochrome LBP computed from each spectrum independently and opponent LBP that captures the spatial correlation between spectra, to describe the color image. Proposed by Lee et al., local color vector binary pattern (LCVBP) [33,34] consists of color norm binary pattern and color angular binary patterns, which are derived by applying LBP to the color image norm and angular features respectively.

Most of existing color face recognition methods [35–40] first select an appropriate color space, e.g., RIQ, YIQ, RCrQ, RGBr, RRGb; then extract the feature of each spectrum by introducing conventional method, e.g., PCA, DFT, Gabor, DCT, LBP, separately; and finally concatenate feature vector of each spectrum together into one for color image representation. However, these methods simply concatenate spectrum-wise color features ignoring the complex relationship between different spectra of the color image.

In this paper, we present a simple yet effective color image descriptor, completed local similarity pattern (CLSP), for color image recognition. We represent the color image as the occurrence of its center pixel color quantization information and the local color image textural information. Our proposed approach can be summarized as follows. First, we assign soft color labels to each pixel of the color image using k-means clustering and soft-assignment coding methods. Meanwhile, according to the similarity of color information between central pixel and its neighbors, local similarity pattern (LSP) is used to encode the local spatial textural information of the color image. Next, the joint distribution (2-D histogram) of color label and LSP within an image or sub-image is used for color image representation, where each cell corresponds to a pair of color label and LSP. Since we use localized soft-assignment coding method, each pixel of the image contributes to multiple color labels. Thus, the histogram aggregated from many pixels in the image serves as a representation of the spatial distribution of color information for the image. Finally, we treat this histogram as a feature vector in a discriminative classifier to recognize the color image. Moreover, since there is no single color space which has the characteristic of invariance against varying imaging and illumination conditions, color spaces including original RGB, HSL, HSV, YIQ, YCbCr, $L^*a^*b^*$ and $L^*u^*v^*$ are selected and fused to compensate each other for color image feature extraction. It is another difference between our CLSP and abovementioned binary color image descriptors. Experimental results on four databases show the effectiveness of the proposed method.

The rest of this paper is organized as follows. In Section 2, we briefly review the LBP and existing binary color image descriptors. In Section 3, we describe our proposed CLSP based image representation method in detail. Section 4 presents the experimental results on four databases. Finally, we conclude the paper in Section 5.

2. Related work

2.1. Local binary pattern

LBP defines an image texture descriptor by comparing a center pixel which is used as a threshold with those pixels in its local neighborhood. Given a central pixel I_c of P neighborhood ones I_p, $p=0,1,\ldots,P-1$ in the image, the LBP coding number is computed by comparing its value with those of its neighborhoods as follows:

$$LBP(I_c) = \sum_{p=0}^{P-1} B(g_p - g_c) \times 2^p, \tag{1}$$

$$B(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \tag{2}$$

where g_c is the gray value of the central pixel I_c, g_p is the value of its neighbor I_p. Because only the signs of the differences between the center pixel and its neighbors rather than their exact values are used to define the encoded pattern number, LBP can achieve gray-scale invariance.

2.2. Multispectral LBP

Mäenpää et al. [32] proposed multispectral LBP (MSLBP), which uses monochrome features computed from each spectrum independently and opponent features that capture the spatial correlation between spectra, to describe the color image. That is to say, the center pixel for a neighborhood and the neighborhood itself can be taken from any spectrum.

$$MSLBP_{i,j}(I_c) = \sum_{p=0}^{P-1} B(v_{i,p} - v_{j,c}) \times 2^p, \quad i,j \in [1,\ldots,N] \tag{3}$$

$B(\cdot)$ is the binary function as defined in (2), i and j are the indexes of the spectrum and N is the total number of the spectra. In general, N is set to 3 for the three color models. i and j are denoted as two of the N spectra. If i is equal to j, it is called monochrome LBP operator and it will be same as (1), otherwise, it is regarded as an opponent LBP operator. For a three-channel color model, there are three monochrome and six opponent LBP operators. The main drawback of the descriptor is that it is not invariant to illumination change.

2.3. Local color vector binary pattern

Lee et al. [33,34] proposed local color vector binary pattern (LCVBP), which consists of color norm binary pattern and color angular binary patterns, for color image representation.

Let I be the original RGB or other color spaces color image, $I_i, i=1,\ldots,N$ denotes the ith spectral-band image, the pixel at z can be represented as a N-component vector, denoted by $v=[v_1,\ldots,v_N]^T$, where the ith element v_i is the pixel value at z of