Digital Signal Processing 48 (2016) 269–275

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

A novel illumination-robust local descriptor based on sparse linear regression ☆

Zuodong Yang [a,b], Yong Wu [a,b], Wenteng Zhao [a,b], Yicong Zhou [c], Zongqing Lu [a,b], Weifeng Li [a,b,*], Qingmin Liao [a,b]

[a] Department of Electronic Engineering, Graduate School at Shenzhen, Tsinghua University, China
[b] Shenzhen Key Laboratory of Information Science and Technology, Guangdong, China
[c] Department of Computer and Information Science, University of Macau, Macau, China

ARTICLE INFO

Article history:
Available online 13 October 2015

Keywords:
Face recognition
Illumination-insensitive representation
Local descriptor
Sparse linear regression

ABSTRACT

Robust face recognition under uncontrolled illumination conditions is an important problem for real face recognition systems. In this paper, we introduce a novel illumination-robust local descriptor named Sparse Linear Regression Binary (SLRB) descriptor. The SLRB descriptor is a bit string by binarizing the sparse linear regression coefficients in a local block. It is an illumination-insensitive descriptor based on the locally linear consistency assumption under the Lambertian reflectance model. We use the cosine similarity and Hamming similarity as the similarity measure for the SLRB descriptor of two different images respectively. Experimental results on the Extended Yale-B and CMU-PIE face database show a promising performance compared to the existing representative approaches.

1. Introduction

Face recognition has recently received a lot of attention and has been applied to the fields of entertainment, smart cards, information security and law enforcement and surveillance [1]. Despite tremendous advance in face recognition has made, robust face recognition under uncontrolled illumination conditions is still challenging [2,3]. In recent years, there are a number of approaches for dealing with face image variations due to illumination changes. They can be roughly classified into four categories.

The first category attempts to handle the illumination normalization problem with traditional image processing methods. Histogram Equalization (HE) obtains an image with a high dynamic range and a great deal of details based on information in the histogram [4]. Gamma Intensity Correction (GIC) was introduced to normalize the overall image intensity at the given grey-level by Gamma transformation [5]. Logarithm Transform (LT) was proposed to perform illumination normalization in face images under uncontrolled illumination conditions [6]. These methods are mainly based on intensity transformation and used as pre-processing methods.

Using the illumination samples, the second category learns the model of face images under varying illumination. In [7] the author made the explanation that arbitrary illumination conditions could be modeled by an image basis and showed that five eigenfaces suffice to represent face images under a wide range of lighting conditions. In [8], it showed the fact that a set of images of an object, which has a fixed pose under varying illumination conditions, form a convex illumination cone in the space of images. In [9], it was proved that the intensity of the object surface obtained with arbitrary distance light sources spans a 9-dimension linear subspace based on a spherical harmonic representation. In [10], the authors propose a novel framework named Face Analysis for Commercial Entities (FACE) and adopt normalization ("correction") strategies to address illumination variations. These methods depend on a statistical model or a physical model, and can settle the illumination variations well. However, they require a large amount of training samples under varying illumination conditions in most cases, which makes them not practical for real face recognition systems.

The third category deals with illumination variations by removing the illumination component. Jobson et al. introduced the Retinex approach to obtain the reflectance component by estimating the illumination component [11,12]. In [13] the author enhanced the illumination removal phase and used double-density dual-tree complex wavelet transform (DD-DTCWT) filtering

☆ This work was supported in part by the National Natural Science Foundation of China (Grants No. 61271393 and No. 11002082).

* Corresponding author at: Department of Electronic Engineering, Graduate School at Shenzhen, Tsinghua University, China.

E-mail addresses: yangzd13@mails.tsinghua.edu.cn (Z. Yang), wuyong11@mails.tsinghua.edu.cn (Y. Wu), zhaowt13@mails.tsinghua.edu.cn (W. Zhao), yicongzhou@umac.mo (Y. Zhou), luzq@sz.tsinghua.edu.cn (Z. Lu), Li.Weifeng@sz.tsinghua.edu.cn (W. Li), liaoqm@tsinghua.edu.cn (Q. Liao).

http://dx.doi.org/10.1016/j.dsp.2015.09.010
1051-2004/

to extract the reflectance portion. Some methods were proposed to remove the illumination component in transformation domain as well, such as the Homomorphic filtering approach [14], discrete cosine transform in the logarithm domain [15], wavelet transform in the frequency domain [16], etc. These methods employ the fact that illumination is a low frequency component and can be applied to a single image without many training samples.

The forth category attempts to find a representation which is insensitive to illumination variations. Wang et al. defined the ratio of the albedo of one image as the Self-Quotient Image (SQI) which is independent of illumination [17]. In [18] the authors proposed the Relative Image Gradient (RIG) feature which is robust against to illumination variations. Gradientfaces takes the similar idea but utilizes the ratio between x-gradient and y-gradient [19]. Both RIG and Gradientfaces are extracted in the gradient domain and can be applied to a single sample. In addition, in [20] and [21], the authors introduced the Weberface and Generalized Weberface (GWF) which employ the relative intensity difference between the center pixel and its neighborhoods derived from the Weber's law. Most of the above methods make use of the Lambertian reflectance model and are based on the assumption that the illumination component is characterized by slow variations while the reflectance component varies drastically.

Apart from the above methods that dedicate to illumination normalization, some approaches for texture classification have been employed for illumination-robust face recognition as well. Local Binary Pattern (LBP) [22] is one of the most commonly used methods. LBP is a local descriptor of texture which processes the difference between the intensity of the center pixel and its neighborhoods with binary encoding. It has been widely applied to illumination-robust face recognition due to its tolerance of monotonic illumination variations and computational simplicity. Nevertheless, LBP is sensitive to noise when the image region is near-uniform. To improve the noise robustness, Local Ternary Pattern (LTP) was proposed in [23], which utilized ternary encoding instead of binary encoding in LBP.

In this paper, we propose a novel local binary descriptor named the Sparse Linear Regression Binary (SLRB) descriptor based on the sparse linear regression. The SLRB descriptor is a bit string obtained by binarizing the sparse linear regression coefficients in a local patch. We prove the SLRB descriptor to be robust to illumination based on the following assumptions.

1) The intensity of the center pixel, $f(x_0, y_0)$, can be linearly expressed by these of its neighborhoods and the linear combination coefficients α_i are consistent in a local block, which is similar to the assumption used in Locally Linear Embedding (LLE) [24]:

$$f(x_0, y_0) = \sum_{j=1}^{N} \alpha_j f(x_j, y_j) + \epsilon, \tag{1}$$

where (x_0, y_0) is the coordinate of the center pixel, (x_j, y_j) is the coordinate of the surrounding pixel and ϵ is the residual term.

2) The intensity of an image, $f(x, y)$, can be expressed as the product of its illumination component, $i(x, y)$, and reflectance component, $r(x, y)$, which is indicated in [14]:

$$f(x, y) = i(x, y) r(x, y). \tag{2}$$

In addition, the illumination component varies slowly while the reflectance component changes abruptly.

We summarize the characteristics of the proposed SLRB descriptor as follows.

1) Based on the Lambertian reflection model, we can prove that the SLRB descriptor is an illumination-insensitive feature and can be effectively applied to illumination-robust face recognition compared with the LBP and LTP features.

2) The SLRB descriptor is a bit string with a low dimension and we can simply employ the cosine distance and Hamming distance as similarity measures. Therefore, the SLRB descriptor is quite efficient and requires less computation complexity than other methods.

3) The lasso regression [25] exhibits the stability of the ridge regression method. Binary encoding scheme is a common method in face recognition and can reduce local noise. As a result, the SLRB descriptor is robust to noise on account of lasso regression and binary encoding.

The rest of the paper is organized as follows. We present our SLRB descriptor, the illumination-robust face recognition algorithm based on the SLRB descriptor and prove its illumination robustness in the next section. In Section 3, we illustrate some experiments by applying our face recognition algorithm on the Extended Yale Face Database B and CMU-PIE face database. Finally, we conclude our paper in Section 4.

2. Sparse linear regression binary descriptor

2.1. Local descriptor based on linear regression

As mentioned above, suppose that the intensity of the center pixel is a linear combination of the intensity of its neighborhoods as Eq. (1). To obtain the linear combination coefficients, we make a further assumption that the linear combination coefficients are the same in a local block. In an $N \times N$ block, we suppose that the intensity of the center pixel is linearly expressed by that of the surrounding eight pixels in a patch. Then there are $(N-2)^2$ "center pixels" $f^{(k)}, k = 1, \cdots, (N-2)^2$, that form a column vector, dependant variable $\mathbf{f} = [f^{(1)}, f^{(2)}, \cdots, f^{((N-2)^2)}]^T$.

For each patch, we have

$$\begin{aligned} f^{(k)} &= \sum_{j=1}^{8} \alpha_j f_j^{(k)} + \epsilon_k \\ &= \mathbf{x}_k^T \boldsymbol{\alpha} + \epsilon_k, \qquad k = 1, 2, \cdots, (N-2)^2, \end{aligned} \tag{3}$$

where $f_j^{(k)}$ is the intensity of the j-th surrounding pixel of the k-th center pixel, the surrounding pixel vector $\mathbf{x}_k = [f_1^{(k)}, f_2^{(k)}, \cdots, f_8^{(k)}]^T$, and the linear regression coefficient vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_8]^T$.

These $(N-2)^2$ equations can be stacked together and written as

$$\begin{bmatrix} f^{(1)} \\ f^{(2)} \\ \vdots \\ f^{((N-2)^2)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_{(N-2)^2}^T \end{bmatrix} \boldsymbol{\alpha} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{(N-2)^2} \end{bmatrix}. \tag{4}$$

Eq. (4) can be reformulated as

$$\mathbf{f} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \tag{5}$$

in which the surrounding pixel matrix $\mathbf{X} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_{(N-2)^2}^T]^T$ and the residual vector $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \cdots, \epsilon_{(N-2)^2}]^T$.

By solving Eq. (5), the linear regression coefficient vector $\boldsymbol{\alpha}$ would be determined as a descriptor of the current block. Fig. 1 illustrates the procedures of the linear regression coefficients in a block.

2.2. Sparse coefficients and binary encoding

Generally, we can determine the linear regression coefficient vector $\boldsymbol{\alpha}$ by the ordinary least square (OLS) estimation: