LOG-DOMAIN POLYNOMIAL FILTERS FOR ILLUMINATION-ROBUST FACE RECOGNITION

Yinyan Jiang[†], *Yong Wu*[†], *Weifeng Li*[†], *Longbiao Wang*[♯], *and Qingmin Liao*[†]

[†]Shenzhen Key Lab. of Information Sci&Tech,
Department of EE/Graduate School at Shenzhen, Tsinghua University, China
[♯]Nagaoka University of Technology, Japan

ABSTRACT

This paper proposes a novel face image descriptor local surface pattern (LSP) for illumination-robust face recognition. It is assumed that the discrete array of pixel values comes about by sampling an underlying smooth surface on the domain of the image. The proposed method efficiently estimates the underlying local surface information, which is approximately represented as linear projection coefficients of the pixels in a local patch. Thus, by filtering local image patches using the polynomial filters and binarizing the filter responses via thresholding, the method can compute a binary code for each pixel in the face image. Then the distribution of the code over suitable image regions is used for face representation. Furthermore, we prove that applying zero-mean filters in log-domain may enable the responses to be more robust to illumination variations. The experimental results on Extended Yale-B and FERET *fc* databases illustrate the effectiveness of our proposed method in illumination-robust face recognition.

Index Terms— face recognition, illumination insensitive, local patterns, surface fitting

1. INTRODUCTION

Face recognition has remained an active research area, with much attention being directed towards extracting robust and discriminant features. Currently, local pattern features, such as Local Binary Patterns (LBP) [1] and Local Quantized Patterns (LQP) [2], have proved effective face descriptors in face recognition due to their extreme simplicity and micro-patterns. Local patterns exploit local information about higher-level image content via patterns of qualitative local gray-level relationship.

Local patterns usually share a same three-step framework. First, certain relationship of pixels in local neighborhood is captured. Second, an encoded image is derived from the relationship according to encoding rules. Third, the number of occurrences of each possible pattern over suitable encoded image regions is counted and then concatenated into a joint histogram.

There are many local patterns focusing on improvements on the first step. LBP [1][3] represents the local information via the difference between the neighbors and central pixel. Three-Patch LBP (TPLBP) and Four-Patch LBP (FPLBP) [4] are proposed to capture the local patch similarities instead of pixel similarities. Local derivative pattern (LDerP) [5] extracts local high-order derivative information instead of the first-order circular derivative pattern in LBP. Local directional pattern (LDirP) [6] adopts kirsch masks to represent the neighboring directional information with edge responses in 8 directions .

Many local patterns attempt to explore the encoding rules. Local ternary pattern (LTP) [7] adopts a ternary instead of binary encoding rule to alleviate the sensitivity to the near-uniform image regions. Local quantized patterns (LQP) [2] incorporates vector quantization into encoding phase to reduce the length of the histogram dimensions when increasing the sampled neighbors.

Inspired by local surface fitting in [8], we propose Local Surface Patterns (LSP) to improve LBP by exploring local surface information of face images. LSP assumes that there is an underlying smooth surface behind a face image, and the surface conveys the inherent structure information of the face. LSP efficiently captures the underlying local surface information via surface fitting coefficients, which are derived by filtering local patches with a set of linear polynomial filters (PF). Polynomial filters also include more high-order edge information, compared with the simple first-order filters in LBP. Our polynomial filters are zero-mean and orthogonal to each other, which may result in the compactness of patterns. Moreover, we prove that the fusion of the logarithmic transform and zero-mean filters may provide more robustness in the presence of illumination variations.

2. ALGORITHMS

In this section, we will firstly provide a brief review of the polynomial filters. Then the local surface patterns (LSP) is introduced. Finally we will present that the fusion of logarithmic transform and zero-mean filter is capable of removing

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

the illumination component.

2.1. Polynomial Filters

In [8], Robert M. Haralick proposed the facet model which exploits surface fit concept [9] to accomplish step edge detection. The facet model assumes that there exists an underlying gray tone intensity surface and the digital image should be regarded as an observed noisy sampling of the surface. The surface is a real-valued function s defined on the domain of the image , which is a bounded and connected subset of the real plane $\mathbb{R}^2$. In some sense, numeric digital image operations should be explained in terms of their actions on the underlying surface. Therefore, there must then involve fitting a function s to the sampled data before operating.

Of course, it is impossible to recover the true function s from the observed noisy sample. However, we can assume some parametric form for the underlying function s, and then use the sampled data to estimate the the parameters. Finally these parameters are kind of representations of the underlying function s, i.e. smooth surface.

Assuming that the function s takes the parametric form of a polynomial in the row and column coordinates, [8] exploits the discrete Chebyshev polynomials as the basis set. Let an index set Ω be given, and the number of elements in Ω is N. A discrete orthogonal polynomials over Ω are constructed, denoted as $\{P_0(x,y),\ldots,P_{N-1}(x,y)\}$. With regard to a 3×3 patch, the index set R is $\{-1,0,1\}\times\{-1,0,1\}$, where the operator $\times$ is tensor product, and the discrete polynomial set $P_n(x,y)$ is

$$\begin{aligned}
&P_0 = 1, P_1 = x, P_2 = y,\\
&P_3 = x^2 - 2/3, P_4 = xy, P_5 = y^2 - 2/3,\\
&P_6 = \left(x^2 - 2/3\right)y, P_7 = \left(y^2 - 2/3\right)x,\\
&P_8 = \left(x^2 - 2/3\right)\left(y^2 - 2/3\right).
\end{aligned} \tag{1}$$

Now the underlying surface $s(x,y),(x,y)\in\Omega$ takes the form

$$s(x,y) = \sum_{n=0}^{N-1} a_n P_n(x,y), \tag{2}$$

where the coefficients $a_0,\ldots,a_{N-1}$ represent the surface information to recover. Let the observed data be $f(x,y)$, for each $(x,y)\in\Omega$. To estimate the coefficients, is to minimize the reconstruction error

$$e^2 = \sum_{(x,y)\in\Omega} [f(x,y) - \sum_{n=0}^{K} a_n P_n(x,y)]^2, K \leq N-1. \tag{3}$$

Setting the derivatives of the above objective with respect to a_m to zero, we obtain for each index $(x,y)\in\Omega$

$$a_m = \sum_{(x,y)\in\Omega} w_m(x,y) f(x,y) = \mathbf{w}_m^T \mathbf{f}, \tag{4}$$

(a)

$$\begin{bmatrix}1&0&0\\0&-1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&1&0\\0&-1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&1\\0&-1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&-1&1\\0&0&0\end{bmatrix}$$

$$\begin{bmatrix}0&0&0\\0&-1&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&0&0\\0&-1&0\\0&1&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&-1&0\\1&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\1&-1&0\\0&0&0\end{bmatrix}$$

(b)

$$\begin{bmatrix}-1&-1&-1\\0&0&0\\1&1&1\end{bmatrix}\begin{bmatrix}-1&0&1\\-1&0&1\\-1&0&1\end{bmatrix}\begin{bmatrix}1&1&1\\-2&-2&-2\\1&1&1\end{bmatrix}\begin{bmatrix}1&0&-1\\0&0&0\\-1&0&1\end{bmatrix}$$

$$\begin{bmatrix}1&-2&1\\1&-2&1\\1&-2&1\end{bmatrix}\begin{bmatrix}-1&0&1\\2&0&-2\\-1&0&1\end{bmatrix}\begin{bmatrix}-1&2&-1\\0&0&0\\1&-2&1\end{bmatrix}\begin{bmatrix}1&-2&1\\-2&4&-2\\1&-2&1\end{bmatrix}$$

Fig. 1: The comparison of filters used in LBP and LSP. (a) Eight filters used in LBP. (b) Eight polynomial filters (PF) used in Local Surface Patterns (LSP) for surface fitting.

where

$$w_m(x,y) = \frac{P_m(x,y)}{\sum_{(x,y)\in\Omega} P_m^2(x,y)}, \tag{5}$$

and $\mathbf{w}_m, \mathbf{f} \in \mathbb{R}^{N\times1}$ are the vector notation of weights $\{w_m(x,y),(x,y)\in\Omega\}$ and patch $\{f(x,y),(x,y)\in\Omega\}$ respectively.

Eq. (4) implies each fitting coefficients a_m can be computed as linear combination of the data values. The weight is just an approximate normalization of an evaluation of the the polynomial P_m at the index (x,y). Regarding to a patch of N pixels, we can obtain N weights vector $\mathbf{w}_m$ for each polynomial basis $P_m, m=0,\ldots,N-1$. We name the weights vector $\mathbf{w}_m$ as polynomial filter (PF), from which we furthermore develop a Local Surface Patterns (LSP). The polynomial filters $(P_m, m=1,\ldots,8)$ for 3×3 window are compared with the simple filters used in LBP in Fig. 1. It can be observed that filters in LBP compute simple first-order derivative, whereas our polynomial filters compute higher-order derivative.

2.2. Local Surface Pattern

With polynomial filters, the proposed Local surface pattern (LSP) attempts to recover the underlying surface information as the pattern and compute a binary code assigned to each pixel of a given image. The surface information is approximately represented via the local fitting coefficients $a_m, m=0,\ldots,8$, which also can be regarded as the responses of polynomial filters.

Note that the first polynomial filter ($m=0$)is indeed the mean filter, and the filter response value of a image will always be non-negative, which is meaningless for the LBP-like code scheme. Thus, we only make use of the other eight zero-mean filters. Given a pixel at (x,y), LSP calculates 8 fitting coefficients $a_m, m=1,\ldots,8$ by weighting the local patch centering at (x,y) with the polynomial filter according to Eq. (4). It's clear that bit strings for all image pixels, can be computed conveniently by 8 convolutions. These coefficients not only contain the underlying surface information, but also can be

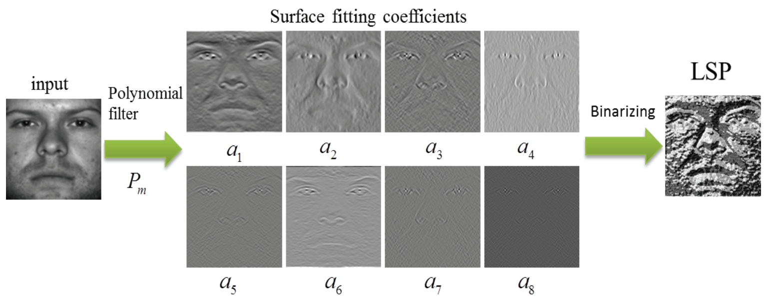

Fig. 2: The computation of LSP. The input image is convoluted with the last 8 polynomial filters $P_m, m = 1, \ldots, 8$ to obtain the local surface fitting coefficients $a_m, m = 1, \ldots, 8$, which are an approximate representation of the underlying surface.

regarded as directional derivative of the local patch. The resulting LSP code can be expressed as follows:

$$LSP(x,y) = \sum_{m=1}^{8} q(a_m) \times 2^{(m-1)}, q(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{6}$$

The LSP computation is shown in Fig. 2.

Spatial histogram is exploited in LSP to model the distribution of the patterns as in original LBP [3]. Our method firstly divides the encoded image into 8×8 non-overlapping regions, and computes a histogram within each of these regions. Finally these spatial histograms are concatenated to form a global face descriptor.

2.3. Illumination Removing

In order to handling illumination variations, we propose to perform filtering in log-domain with polynomial filters rather than intensity domain. According to the Lambertian reflectance model, a face image f could be expressed by

$$f(x,y) = r(r,c)\, i(x,y), \tag{7}$$

where $f(x,y)$ is the image pixel value, $r(x,y)$ denotes the reflectance and $i(x,y)$ denotes the illumination at each pixel (x,y). i depends mainly on the lighting source and the position between the source and the object, while r can be considered as face feature, which includes the characteristics in both surface texture and $3D$ shape. Many previous methods attempt to obtain r by removing i from f.

Now we consider a 3×3 patch and convert the matrix form in Lambertian model to vector form, we have

$$\mathbf{f} = \mathbf{r}^T \mathbf{i}, \tag{8}$$

where $\mathbf{f}, \mathbf{r}, \mathbf{i} \in \mathbb{R}^{9 \times 1}$, correspond respectively to f, i, i. If we filter the f using $\mathbf{w}$ in vector form in the logarithm domain instead of pixel intensity domain, thus we obtain

$$\mathbf{w}^T \log \mathbf{f} = \mathbf{w}^T \log \mathbf{r} + \mathbf{w}^T \log \mathbf{i}, \tag{9}$$

where logarithm transform is performed on each element of the vectors.

It's commonly assumed that the illumination component $i(x,y)$ varies quite slowly except for the shadow boundaries. Then in a local patch, the illumination value can be treated as a constant. Therefore when the filter $\mathbf{w}$ is zero-mean, we have

$$\mathbf{w}^T \log \mathbf{i} = \alpha \mathbf{w}^T \mathbf{1} = 0, \tag{10}$$

which could be regarded as a way to remove the illumination component. By substituting Eq.(10) into Eq.(9), the output

$$\mathbf{w}^T \log \mathbf{f} = \mathbf{w}^T \log \mathbf{r}, \tag{11}$$

is indeed an illumination insensitive representation of the central pixel of the local patch.

As just proved, applying a zero-mean filter to the logarithm transformed image, can theoretically eliminate the effect of the illumination variations. Therefore, we extract the local surface patterns of the image in the logarithm domain instead of directly pixel intensity.

2.4. Related to Other Work

LBP has proven a discriminant and effective feature for face recognition. However, few number of neighbor pixels used in generating a bit results in limited accuracy, because most of the information in the neighborhood is ignored and it is very sensitive to intensity changes in the presence of noise, especially when the size of the neighborhood increases. Moreover, the simple first-order derivative in 8 directions is also a limitation to accuracy.

To overcome these drawbacks, LderP [5] and LDirP [6] respectively exploit high-order derivative and kirsch masks for exploring more useful information conveyed by more pixels in a neighborhood. Differ from them, our proposed method LSP assumes that there is an underlying surface behind the face image, and attempts to encode local surface information via surface fitting by polynomial filters. The use of structure information improves the modeling capacity of LSP.

3. FACE RECOGNITION EXPERIMENTS

Our experiments are carried out on two face databases with large illumination variations, namely, Extended Yale-B [10] and *fc* subset of FERET database[11] to illustrate the effectiveness of our proposed LSP. There are many distance and similarity measures developed for histogram matching. Here we use histogram intersection to measure the similarity between two histograms. The nearest neighborhood rule is used as the classifier.

Extended Yale-B. In the extended Yale-B database, there includes 38 subjects under 9 poses and 64 illumination conditions. All images are divided into 5 subsets according to the angle between the light source direction and the central camera axis. In our experiments, images with the most neutral light conditions ('A+000E+00') were used as the gallery, and only frontal images in subsets were used as probes. All images are cropped and resized to 120×120.

FERET-fc. In the standard FERET database, the basic gallery *fa* contains $1,196$ images of $1,196$ subjects. There are four probe sets with different environment variations. One of the probe sets, the *fc* set is designed for evaluating the illumination variation. The *fc* set includes 194 images of 194 subjects taken in the same time under significantly different lighting conditions. For our experiment, all images are aligned, cropped and resized to 128×128 based on the location of the eyes.

3.1. Results and Discussions

The comparative experiments between LSP and LBP were first conducted on Extended Yale-B database. We take into consideration two kinds of LBP operators, $\text{LBP}_{8,2}(u)$ [1] and LBP(8) [3]. $\text{LBP}_{8,2}(u)$ denotes the LBP operator based on a symmetric neighborhood of 8 members on a circle of radius 2 with only uniform patterns used. LBP(8) and our proposed LSP(8) are both based on the 3×3 neighborhood with all 8 bits used.

Table 1 and Table 2 respectively list the comparison results of recognition rates in intensity domain and log-domain on Extended Yale-B. We can see that LSP(8) outperforms LBP(8) and $\text{LBP}_{8,2}(u)$ on all five subsets in both intensity and logarithm domain. From the comparison between Table 1 and Table 2, it's clear that the fusion of logarithmic transform and zero-mean filters improves the accuracy, as all filters in LBP and LSP are zero-mean. The fusion of logarithmic transform and zero-mean filters improves LBP no more than 0.8%, but improves LSP 2.5%.

Table 1: The comparison results of recognition rates (%) in pixel intensity domain on Extended Yale-B.

Method	S1	S2	S3	S4	S5	Avg.
$LBP_{8,2}(u)$	100	100	96.92	61.03	34.87	78.56
$LBP(8)$	100	100	98.68	73.76	43.70	83.23
$LSP(8)$	100	100	99.34	79.66	48.88	85.58

We also validate LSP on the *fc* set of FERET database to further illustrate its robustness to illumination variations.

Table 2: The comparison results of recognition rates (%) in log-domain on Extended Yale-B.

Method	S1	S2	S3	S4	S5	Avg.
$LBP_{8,2}(u)$	100	100	97.58	61.79	34.87	79.04
$LBP(8)$	100	100	98.24	75.48	46.22	83.99
$LSP(8)$	100	100	99.12	83.46	57.84	88.08

Table 3: The comparison results of recognition rates (%) in log-domain on *fc* subset of FERET.

Method	Original	Log-domain
$LBP_{8,2}(u)$	80.93	81.96
$LBP(8)$	68.04	72.16
$LSP(8)$	83.51	89.69

Table 3 lists the comparison results of recognition rates in intensity domain and log-domain on *fc* of FERET. The results demonstrate that LSP performs best in both original intensity domain and log-domain. The fusion of logarithmic transform and our polynomial filters gains more than 6% in accuracy.

From the above results we may have following observations:

- Local neighborhoods contain lots of useful information, and LSP successfully encodes local surface structure and high-order derivative information, which is otherwise ignored by LBP. The property that polynomial filters are orthogonal to each other also make LSP more compact and reduce the redundancy of patterns.

- LSP and LBP share the important property of being invariant to monotonic gray-level changes, because their patterns are linear combinations of the pixel intensity in a local patch. However, illumination changes may not always be monotonic in the presence of noise. Furthermore, the effect of noise will be enlarged in log-domain as logarithmic transform compresses the intensity range. The better performance of LSP in both intensity domain and log-domain, demonstrates that LSP is more robust to noise, for more than a single neighbor are taken into consideration and polynomial filters capture higher-order information than LBP.

4. CONCLUSIONS

This paper has presented a novel image feature based on LSP codes for illumination-robust face recognition. The local surface information is exploited by measuring the relationship between the neighboring pixels and the central pixel with local fitting coefficients. LSP turns out to be more robust to illumination variations than LBP both in intensity and log-domain. However, it remains an interesting issue how to capture the information in the local patch.

5. REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietik*ä*inen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. on Patt. Analysis and Machine Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.

[2] S. Hussain, T. Napoléon, and F. Jurie, "Face recognition using local quantized patterns," in *British Machive Vision Conference*, 2012.

[3] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[4] L. Wolf, T. Hassner, and T. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Trans. on Patt. Analysis and Machine Intell.*, vol. 33, no. 10, pp. 1978–1990, 2011.

[5] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. on Image Process.*, vol. 19, no. 2, pp. 533–544, 2010.

[6] T. Jabid, Md. Hasanul Kabir, and Oksam Chae, "Facial expression recognition using local directional pattern (ldp)," in *IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 1605–1608.

[7] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proc. IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures*, 2007, pp. 235–249.

[8] Robert M. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Trans. on Patt. Analysis and Machine Intell.*, , no. 1, pp. 58–68, 1984.

[9] J. Prewitt, *Object enhancement and extraction*, vol. 75-149, Academic Press, New York, 1970.

[10] K. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. on Patt. Analysis and Machine Intell.*, vol. 27, no. 5, pp. 684–698, 2005.

[11] P. J. Phillips, H. Moon, P. Rizvi, and P. Rauss, "The feret evaluation method for face recognition algorithms," *IEEE Trans. on Patt. Analysis and Machine Intell.*, vol. 22, no. 10, pp. 1090–1104, 2000.