

# FACE RECOGNITION WITH SALIENT LOCAL GRADIENT ORIENTATION BINARY PATTERNS

*Shu Liao, Albert C. S. Chung*

Lo Kwee-Seong Medical Image Analysis Laboratory,  
Department of Computer Science and Engineering,  
The Hong Kong University of Science and Technology, Hong Kong.

## ABSTRACT

This paper proposes a new face recognition method. There are two novelties in the proposed method. First, a new saliency measure function is designed to detect the most salient regions in facial images and determine their corresponding best scales. Second, a new type of image feature, called local gradient orientation binary pattern (LGOBP) is proposed, which captures the neighborhood gradient orientation information which is not considered in the conventional local binary patterns (LBP) to give more discriminant power. LGOBPs are extracted from the most salient regions selected by the proposed saliency measure function. The proposed method is evaluated on the FRGC version 2 database by comparing it with several widely used methods. Experimental results show that the proposed method achieves the highest recognition rate among all the compared methods.

*Index Terms*— Face recognition, Machine vision

## 1. INTRODUCTION

Automatic face recognition (AFR) is an active research topic in computer vision. Its applications include biometrics, law enforcement, and surveillance development. It has also gained an increasing interest in building natural human-computer interaction (HCI) systems. Face recognition is a challenging task because factors such as pose, illumination, facial expression make it difficult to achieve high face recognition rate. As the demand of such applications increases, many novel face recognition methods have been proposed in the last several decades.

Face recognition can be factorized into two essential parts: (1) feature extraction from facial images; and (2) similarity measure and classifier design. Many feature extraction methods have been proposed regarding to the first part. For example, Wiskott *et al.* applied the Gabor wavelet features in the elastic bunch graph matching algorithm (EBGM) for face recognition [1]. Turk *et al.* used the principle component analysis (PCA) to compute facial image features [2]. This is commonly called the "eigenface". Manli and Martinez [3] selected the principal components in a two-stage LDA. In

recent years, a new feature extraction method, called local binary patterns (LBP) was proposed. LBP was originally used in texture classification [4]. It has also been applied in face recognition [5].

In this paper, we propose a new feature extraction approach for face recognition. The main contributions of this paper are: (1) A new saliency measure function is designed to detect the most salient regions in the facial images and determine their best scales with respect to the saliency. This saliency measure function is defined based on the generalized survival exponential entropy (GSEE); (2) Based on the most salient regions detected by the saliency measure function, a new region descriptor called local gradient orientation binary patterns (LGOBP) is proposed to extract anatomical features from those salient regions. LGOBP is developed based on the conventional local binary patterns (LBP) [4]. LGOBP encodes higher order pixel-wise information compare to the conventional LBP as it takes gradient orientations of the neighboring pixels into consideration which contain rich directional information. Moreover, LGOBP preserves the monotonic gray-level transformation invariant property of the conventional LBP.

The proposed method is evaluated on the FRGC version 2 database [6]. It is also compared with other widely used methods. Experimental results show that the proposed method achieves the highest recognition rate.

## 2. METHODOLOGY

In this section, we will first introduce the new saliency measure function defined based on the generalized survival exponential entropy (GSEE) [7] and show how to use it to detect the most salient regions with their best scales from the facial images. Then the LGOBP region descriptor is given to extract anatomical features from the most salient regions detected by the saliency measure function.

### 2.1. Saliency Measure Function Based on GSEE

Faces are topological objects, in order to preserve the spatial structural information in facial images, local based features such as LBP [5] need to be extracted from different regions of

facial images. A simple strategy used in LBP [5] is to divide the facial images uniformly into non-overlapping rectangular regions. However, as different structures of human face have different physical sizes, using rectangular regions with the same scale to characterize them is problematic. Also, this strategy results in information redundancy. Therefore, in order to increase the discriminant power of the local based features, salient regions with proper scales should be first located at the facial images, then local based features are extracted from those salient regions.

In this paper, we measure saliency based on the local image complexity. The generalized survival exponential entropy [7] (GSEE) is adopted to statistically measure the local image complexity. GSEE is defined by Equation 1:

$$GSEE_{\alpha,\beta}(X) = \left( \frac{\int_{R_+^m} \bar{F}_{|X|}^\alpha(x) dx}{\int_{R_+^m} \bar{F}_{|X|}^\beta(x) dx} \right)^{\frac{1}{\beta-\alpha}}, \quad (1)$$

where  $\alpha \geq 0$  and  $\beta \geq 0$  define the order of GSEE and  $\alpha \neq \beta$ ,  $X = (X_1, \dots, X_m)$  is a random vector in  $R^m$ .  $|X|$  denotes the random vector with components  $|X_1|, \dots, |X_m|$ . The notation  $|X| > x$  means that  $|X_i| > x_i$  for  $x_i \geq 0, i = 1, \dots, m$ . The multivariate survival function  $\bar{F}_{|X|}(x)$  of the random vector  $|X|$  is defined by:

$$\bar{F}_{|X|}(x) = P(|X| > x) = P(|X_1| > x_1, \dots, |X_m| > x_m) \quad (2)$$

for  $x \in R_+^m$  with  $R_+^m$  defined by Equation 3:

$$R_+^m = \{x \in R^m : x = (x_1, \dots, x_m), x_i \geq 0, i = 1, \dots, m\}. \quad (3)$$

The random vector  $X$  is the local intensity probability distribution histogram of a squared region with scale  $s$  (i.e. the side length of the square) centered at each referencing pixel  $v$ . Compare to the conventional Shannon's entropy, GSEE has several advantages: (1) GSEE is always nonnegative; (2) GSEE have consistent definitions in both the continuous and discrete domains; (3) The Shannon's entropy is calculated based on the density of the random variable  $p(X)$ . However,  $p(X)$  may not exist as pointed out in [8]. GSEE is calculated based on the survival function which always exists.

For each pixel  $v$ , the best scale  $S_v$  associated with it is defined as:

$$S_v = \arg \max_s GSEE_{\alpha,\beta}(H(s, v)), \quad (4)$$

where  $H(s, v)$  denotes the local intensity probability distribution histogram calculated from the squared region centered at  $v$  with side length  $s$ .

A region with complex image structures will give a large value of GSEE. However, if that region has high degrees of self-similarity over a large range of scales, it should also be

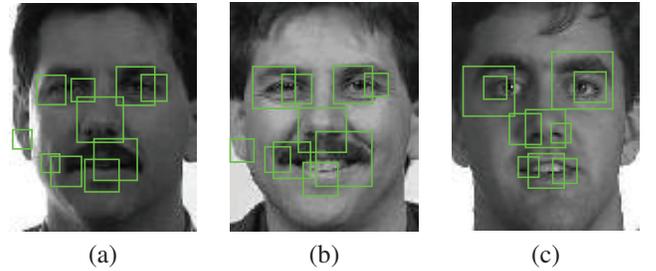
considered as non-salient. A typical example is a highly textured region. Therefore, a salient region should have complex image structures within a narrow range of scales. Thus, the saliency measure function of each pixel  $v$  is defined by the GSEE value of  $H(S_v, v)$  multiplied by the differentiation of  $H(S_v, v)$  over the scale space:

$$M(v) = GSEE_{\alpha,\beta}(H(S_v, v)) \cdot (S_v \cdot \sum_i \left\| \frac{\partial H_i(s, v)}{\partial s} \Big|_{s=S_v} \right\|), \quad (5)$$

where  $H_i(s, v)$  denotes the  $i$ th element of  $H(s, v)$ .

Based on the experimental results, it is found that  $\alpha = 3, \beta = 5$  is a good choice as the order parameters of GSEE. The scale space is defined from 4 to 20 pixels, that is, the best scale of each pixel is selected from the set of candidate scales  $\{s | s = 4, 5, \dots, 20\}$ .

Figure 1 shows an example of using the saliency measure function in Equation 5 for three facial images obtained from the FERET database [9]. Images in Figure 1(a) and Figure 1(b) are from the same person but with different facial expressions, while facial image in Figure 1(c) is from another person. The top 10 most salient regions are highlighted with the green squares, with the associated referencing pixel at the center of each region. The side length of each green square denotes the best scale of it. It is observed that the most salient regions detected in Figures 1(a) and 1(b) are located at similar positions of the most distinctive structures such as corners of eyes, mouths and noses with similar best scales. It matches with the fact that they are facial images of the same person. On the other hand, the most salient regions detected in Figure 1(c) correspondingly have large best scale differences between those in Figures 1(a) and 1(b), indeed, because Figure 1(c) is a facial image of another person.



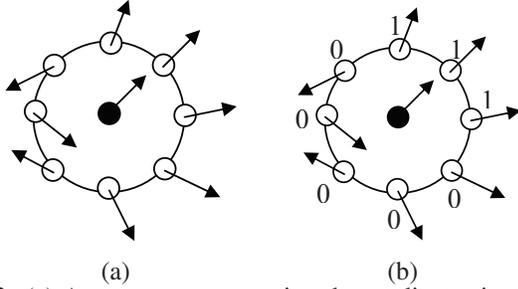
**Fig. 1.** The top 10 most salient regions detected by the proposed saliency measure function of three facial images are highlighted with green squares, with the corresponding referencing pixel at the center of each region. The side length of each green square denotes the best scale associated with it. (a) and (b) are from the same person with different facial expressions, (c) is from another person.

It is worth pointing out that determining the best scale of each salient region is important, because saliency is closely related to scales. For example, for the salient region detected surrounding the nose in Figure 1(a), it will be considered as non-salient if only a small scale is evaluated.

## 2.2. Local Gradient Orientation Binary Patterns

After extracting the most salient regions with their best scales, we can extract local based features from those regions as signatures to describe the facial images. In this section, a new region descriptor named local gradient orientation binary patterns (LGOBP) developed from LBP [4] is designed to extract features from the salient regions. LGOBP encodes higher order pixel-wise information than LBP and preserves the monotonic gray-level transformation invariant property of LBP.

For each pixel  $v_c$ , let  $\psi(\nabla_{v_c})$  denote its gradient orientation angle. Then a circularly symmetric neighborhood system with radius  $R$  is defined centered at pixel  $v_c$ , with  $k$  neighboring pixels uniformly located on the circle, neighboring pixels which are not exactly located at the image grid are interpolated by using the linear interpolation method. We denote the  $k$  neighboring pixels as  $v_i$  ( $i=1,\dots,k$ ) and let  $\psi(\nabla_{v_i})$  denote the gradient orientation angle of  $v_i$ . Figure 2(a) shows an example with  $R = 2$  and  $k = 8$ , with arrows denoting the gradient orientations of different pixels.



**Fig. 2.** (a) Arrows are representing the gradient orientations of different pixels; (b) Binary numbers are assigned to each neighboring pixel by applying Equation 7.

Then, we partition the gradient orientation space uniformly into four subspaces. Pixels with gradient orientations fall in the same subspace are considered as with the same gradient orientations. More precisely, each pixel is given one of the four labels according to its gradient orientation. For each pixel  $v$ , its label  $l(v)$  is determined by Equation 6:

$$l(v) = \begin{cases} 1, & \text{if } \psi(\nabla_v) \in [0, \frac{\pi}{2}) \\ 2, & \text{if } \psi(\nabla_v) \in [\frac{\pi}{2}, \pi) \\ 3, & \text{if } \psi(\nabla_v) \in [\pi, \frac{3\pi}{2}) \\ 4, & \text{if } \psi(\nabla_v) \in [\frac{3\pi}{2}, 2\pi) \end{cases} \quad (6)$$

The pixel-wise interaction information between the centering pixel  $v_c$  and the neighboring pixels  $v_i$  is then analyzed by comparing their corresponding labels obtained from Equation 6. Finally, each neighboring pixel is assigned with a binary number by Equation 7:

$$B(v_i) = \begin{cases} 1, & \text{if } l(v_i) = l(v_c) \\ 0, & \text{if } l(v_i) \neq l(v_c) \end{cases} \quad (7)$$

Figure 2(b) shows the resulting binary numbers assigned to each neighboring pixel by using Equation 7. The resulting binary patterns are called local gradient orientation binary patterns (LGOBP). LGOBP encodes the second order pixel-wise information as the gradient orientations of each pixel already contain the first order pixel-wise interaction properties, while LBP [4] only considers the first order information as it thresholds the neighboring pixels by only comparing their intensities to the intensity of the center pixel. LGOBP is also monotonic gray-level transformation invariant because although the gradient magnitude of each pixel changes according to monotonic gray-level transformation, their gradient orientations remain the same.

Similarly to the definition of uniform LBP [4], a LGOBP is a uniform LGOBP only if the bit-wise transactions between "0" and "1" of the neighboring pixels are less than or equal to two. For example, the resulting LGOBP shown in Figure 2 is a uniform LGOBP as it has only two bit-wise transactions between "0" and "1". Also, similar to LBP [5], we use the histogram of uniform LGOBP as the signature for each salient region.

The final feature to represent a facial image therefore is extracted as follows: First, we use the saliency measure function proposed in Section 2.1 to find out the top  $N$  most salient points associate with their corresponding salient regions of best scales. In this paper  $N = \lceil 0.15X \rceil$ , where  $X$  is the total number of pixels of a facial image. Then, we sort this  $N$  most salient regions with respect to their saliency values calculated by Equation 5. Finally, the LGOBP features are extracted from each salient region in order (i.e. from the salient region with the first largest saliency value to the salient region with the  $N$ th largest saliency value determined by Equation 5) and then concatenate the LGOBP features extracted from each salient region in the same order to form the final feature vector.

The support vector machine (SVM) [10] with the Gaussian Radius Basis Function (RBF) kernel was used as the classifier in this work.

## 3. EXPERIMENTAL RESULTS

In this section, the proposed method is evaluated on the FRGC version 2 database [6]. It is also compared with the BEE Baseline algorithm and the conventional LBP [5]. Before we perform the experiment, the facial images are normalized and cropped to the size of  $128 \times 128$  pixels to extract the facial region based on the eye positions provided by the FRGC. Figure 3 shows some example facial images from the FRGC database.

In the FRGC version 2 database, there are 12776 images taken from 222 subjects in the training set, 16028 target images taken under the controlled illumination condition. We conduct experiment 1 and experiment 4 protocols to evaluate the performance of the proposed method. In experiment 1, there are 16028 query images taken under the controlled il-



**Fig. 3.** Sample facial images in the FRGC Database.

illumination condition. In experiment 4, there are 8014 query images taken under the uncontrolled illumination condition. Experiment 4 is the most challenging protocol in FRGC as the uncontrolled large illumination variation brings significant difficulties to achieve high recognition rate.

The experimental results for the FRGC version 2 database in experiment 1 and experiment 4 are evaluated by using the Receiving Operator Characteristics (ROC), which is the face verification rate (FVR) versus the false accept rate (FAR). There are three ROC values, ROC 1 corresponding to images collected within semester, ROC 2 corresponding to images collected within year and ROC 3 corresponding to images collected between semesters. Tables 1 and 2 list the performance of different approaches on the face verification rate (FVR) at the false accept rate (FAR) of 0.1% in experiment 1 and experiment 4:

Methods	FVR at FAR = 0.1% (in %)		
	ROC 1	ROC 2	ROC 3
1. BEE Baseline	77.63	75.13	70.88
2. <i>LBP</i>	86.24	83.84	79.72
3. <i>GSEE + LGOBP</i>	97.63	95.15	91.03

**Table 1.** The FVR value of different approaches at FAR = 0.1% in experiment 1 of the FRGC version 2 database.

Methods	FVR at FAR = 0.1% (in %)		
	ROC 1	ROC 2	ROC 3
1. BEE Baseline	17.13	15.22	13.98
2. <i>LBP</i>	58.49	54.18	52.17
3. <i>GSEE + LGOBP</i>	76.28	74.36	75.09

**Table 2.** The FVR value of different approaches at FAR = 0.1% in experiment 4 of the FRGC version 2 database

It is observed from Tables 1 and 2 that the proposed method significantly outperforms the BEE baseline algorithm (PCA) and the conventional LBP method under both the controlled and uncontrolled conditions. Especially for experiment 4, which is the most challenging protocol in FRGC version 2 database, the proposed method achieves the FVR rate of 75.09% in ROC 3, as compared with 52.17% for the conventional LBP method and 13.98% for the BEE baseline algorithm.

#### 4. CONCLUSION

In this paper, a new face recognition method is proposed. The proposed approach consists of two steps: First, a saliency measure function based on the generalized survival exponential entropy (GSEE) is proposed to detect the most salient regions in the facial images. Then, the local gradient orientation binary pattern (LGOBP) is designed to extract features from those salient regions to represent the facial images. The proposed method is evaluated by using the FRGC version 2 database. It is shown that the proposed method outperforms the conventional LBP method and the BEE baseline algorithm, which implies the robustness of the proposed method against the variations of expression, lighting and aging.

#### 5. REFERENCES

- [1] L. Wiskott, J.M. Fellous, N. Kruger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *PAMI*, vol. 19, no. 7, pp. 775–779, 1997.
- [2] M. Turk and Pentland A., "Eigenfaces for recognition," *J. Cogn. Neurosci*, vol. 3, pp. 71–86, 1991.
- [3] Z. Manli and A.M. Martinez, "Selecting principal components in a two-stage lda algorithm," in *CVPR*, 2006, vol. 1, pp. 132–137.
- [4] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *PAMI*, vol. 24, no. 7, pp. 971–987, 2002.
- [5] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local binary patterns," in *ECCV*, 2004, vol. 1, pp. 469–481.
- [6] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek, "Overview of the face recognition grand challenge," in *CVPR*, 2005, pp. 947–954.
- [7] K. Zografos and S. Nadarajah, "Survival exponential entropies," *IEEE Trans. on Information Theory*, vol. 51, no. 3, pp. 1239–1246, 2004.
- [8] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.
- [9] P.J. Phillips, H. Wechsler, J. Huang, and P.J. Rauss, "The feret database and evaluation procedure for face recognition algorithms," *IVC*, vol. 16, no. 5, pp. 295–306, 1998.
- [10] V.N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.