21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Local Gaussian Directional Pattern for Face Recognition

Adin Ramirez Rivera, Jorge Rojas, and Oksam Chae
Kyung Hee University
South Korea, 1 Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do 446-701
{adin, jarc, oschae}@khu.ac.kr

Abstract

*We propose a novel local feature descriptor, Local Gaussian Directional Pattern (LGDP), for face recognition. LGDP encodes the directional information of the face's textures (*i.e.*, the texture's structure) in a compact way, producing a more discriminating code than other methods. The structure of each micro-pattern is computed by using a derivative-Gaussian compass mask, and encoded by using its prominent directions and sign—which allows it to distinguish among similar structural patterns that have different intensity transitions. Moreover, our descriptor extracts several facial characteristics by varying the size of its mask, to recover features that may be missed in just one resolution. We construct the face descriptor by concatenating the LGDP's distributions extracted from a uniform grid of the face. We perform several experiments in which our descriptor performs consistently under illumination, noise, expression and age variations.*

1 Introduction

In face analysis, a key issue is the descriptor of the face appearance [8]. The descriptor's efficiency depends on its representation and the ease of extracting it from the face. Ideally, a good descriptor should have a high variance among classes (between different persons), but little or no variation within classes (same person in different conditions).

We classify the descriptors into two classes according to their features: *global* and *local*. The global-feature descriptors are called holistic methods. These methods treat the face as a whole, and extract a descriptor from it in such way. Although these methods have been studied widely, local descriptors have gained attention because of their robustness to illumination and pose variations. Heisele *et al*. showed the validity of the component-based methods, and how

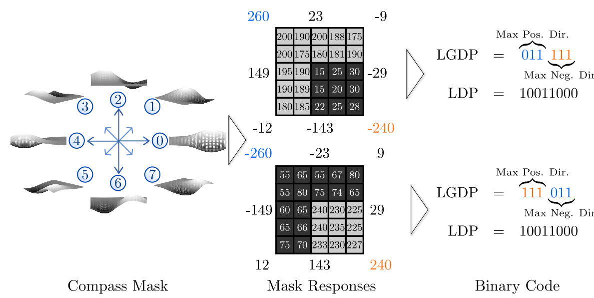


Figure 1. LGDP code computation.

they outperform holistic methods [3]. The local-feature methods compute the descriptor from parts of the face, and then gather the information into one descriptor. Among these methods are, Gabor features, Elastic Bunch Graph Matching (EBGM), and Local Binary Pattern (LBP) [1]. LBP achieves better performance than previous methods, thus it is widely used nowadays. Newer methods tried to overcome the shortcomings of LBP, one of them is the Local Directional Pattern (LDP) introduced by Jabid *et al*. [4], and extended by Kabir *et al*. [5]. This method encodes the directional information in the neighborhood, instead of the intensity. These methods use other information, instead of intensity, to overcome noise and illumination variation problems. However, these methods still suffer in non-monotonic illumination variation, random noise, and changes in pose, age, and expression conditions. Although some methods, like Gradientfaces [7], have a high discrimination power under illumination variation, they still have low recognition capabilities for expression and age variation conditions.

In this paper, we propose a novel face descriptor, Local Gaussian Directional Pattern (LGDP), for robust face recognition that encodes the structural information and the intensity variations of the face's texture. This approach allows us to distinguish intensity changes (*e.g.*, from bright to dark and vice versa) in the texture, that otherwise will be missed. Furthermore, our descriptor uses the information of the entire neighborhood, instead of using sparse points for its computation like LBP. Hence, our approach conveys more information into the code, yet it is more compact—as it is six bit long. Moreover, we use different resolutions for coding to acquire characteristics that may be neglected by just one resolution, and combine them to extend the encoded information. The introduction of multiple encoding levels produces an improvement in the detection process.

2 Local Gaussian Directional Pattern

LBP [1] encodes the local intensity by using the center pixel as a threshold for a sparse sample of the neighboring pixels. The few pixels used limit the accuracy of the method, because it discards most information in the neighborhood. To avoid these problems, all the neighborhood's pixels can be used, as LDP [4] does. Although using more information makes LDP more stable, it still encodes the information in a similar way as LBP: by marking the maximum absolute directions in a bit string. This encoding scheme, however, misses some directional information (the responses' sign) by treating all directions equally. Also, it is sensitive to illumination changes, and noise. To avoid these problems we propose to use a derivative-Gaussian compass mask to avoid the noise perturbation, and to make our method robust to illumination changes. Moreover, we implicitly include the sign of the directional responses to increase the structural information encoded.

The proposed Local Gaussian Directional Pattern (LGDP) is a six bit binary code assigned to each pixel of an input image that represents the structure of the texture and its intensity transitions. Hence, we create the pattern by computing the edge response of the neighborhood using a derivative-Gaussian compass mask, and by taking the most positive and negative directions of those edge responses. This codding scheme is illustrated in fig. 1. The positive and negative responses provide valuable information on the structure of the neighborhood, as they reveal the gradient direction of bright and dark areas in the neighborhood. Furthermore, the distinction between dark and bright responses, allows LGDP to differentiate between blocks with the positive and the negative direction swapped (which is equivalent to swap the bright and the dark areas of the neighborhood as shown in fig. 1), because it generates a different code for each of them, while other methods may mistake the swapped regions as one. Furthermore, these transitions occur often in the face, *e.g.*, the top and bottom edges of the eyebrows and mouth have different intensity transitions.

To describe the face, we combine different sizes of the mask to compute the LGDP code. The change in the mask's size allows our method to capture features in the face that otherwise may be overlooked. It is vital to provide descriptive features for long range pixel interaction. However, neither LBP nor LDP take this into account. We find that combining the local shape information, the relation between the edge responses, and relating the information from different resolutions can better characterize the face's appearance.

2.1 Compass mask

Inspired by the Kirsch mask, we use the derivative of a skewed Gaussian to create an asymmetric compass mask that we use to compute the edge response on the smoothed face. This mask is robust against noise and illumination changes, while producing strong edge responses. Hence, given a Gaussian mask defined by:

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (1)$$

where x, y are location positions, and σ is the width of the Gaussian bell; we define our mask as:

$$M_\sigma(x,y) = G'_\sigma(x+k,y) * G_\sigma(x,y), \quad (2)$$

where G'_σ is the derivative of G_σ with respect to x, $*$ is the convolution operation, and k is the offset of the Gaussian with respect to its center—in our experiments we use one fourth of the mask diameter for k. Then, we generate a set of compass masks (as shown in fig. 1), $\{M_\sigma^0\text{–}M_\sigma^7\}$, by rotating M_σ, $45°$ apart, in eight different directions.

2.2 Coding scheme

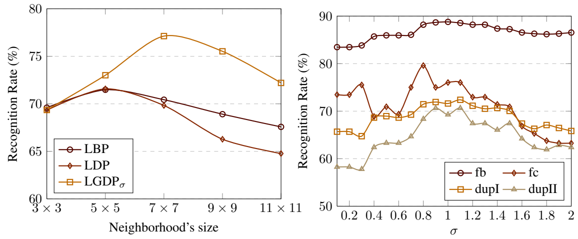
We generate the code, LGDP$_\sigma$, by analyzing the edge response of each mask, M_σ^0–M_σ^7, that represents the edge significance in its direction, and by combining the dominant information. Given that the responses are not equally important, the presence of a high negative or positive value signals a prominent dark or bright area. Hence, we used a fixed position for the positive (as the three most significant bits) and negative values (as the three least significant bits) to encode the sign information. Thus, we define the code as:

$$\text{LGDP}_\sigma(x_c,y_c) = 8i^\sigma_{x_c,y_c} + j^\sigma_{x_c,y_c}, \quad (3)$$

where (x_c,y_c) is the central pixel of the neighborhood being coded, $i^\sigma_{x_c,y_c}$ is the direction number of the maximum positive response, and $j^\sigma_{x_c,y_c}$ is the direction number of the minimum negative response defined by:

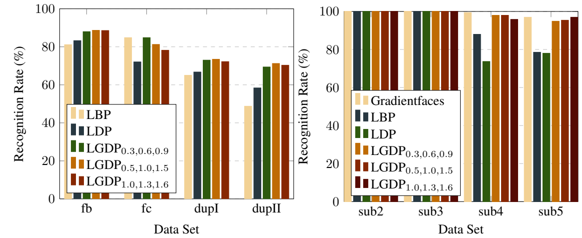
$$i^\sigma_{x_c,y_c} = \arg\max_i\{M^i_\sigma(x_c,y_c) \mid 0 \le i \le 7\}, \quad (4)$$

$$j^\sigma_{x_c,y_c} = \arg\min_j\{M^j_\sigma(x_c,y_c) \mid 0 \le j \le 7\}. \quad (5)$$



(a) Varying Neighborhood (b) Varying σ

Figure 2. Recognition rate on FERET [6].



(a) FERET [6] (b) Yale B [2]

Figure 3. Comparison of several methods.

2.3 Face descriptor

Each face is represented by a LGDP histogram (LH) that contains fine to coarse information of an image. It is achieved by computing the LGDP$_\sigma$ code at n different σ_n (LGDP$_{\sigma_1,\ldots,\sigma_n}$) and by concatenating them into one histogram, a multi-LGDP histogram (MLH), that is defined as: $\mathrm{MLH}_{\sigma_1,\ldots,\sigma_n} = \prod_{i=1}^{n} \mathrm{H}_{\sigma_i}$, where $\prod$ is the concatenation operation, H_{σ_i} is the histogram of the LGDP$_{\sigma_i}$ code, and n is the number of σ used—in our experiments we limit ourselves to three. Given that the histogram omits location information, to aggregate it to the descriptor, we divide the face image into a grid of N small regions, $\{R^1,\ldots,R^N\}$, and extract the $\mathrm{MLH}^i_{\sigma_1,\ldots,\sigma_n}$ from each region, R^i. Finally, the LH is computed by concatenating those histograms $\mathrm{LH} = \prod_{i=1}^{N} \mathrm{MLH}^i_{\sigma_1,\ldots,\sigma_n}$. The LH is used during the face recognition process, by comparing the encoded feature vector from one person with all other candidate's feature vector with the Chi-Square dissimilarity measure.

3 Experiments and Results

We evaluate the performance of the proposed algorithm under expression, age, pose, and illumination variation. We cropped and normalized all images to 100×100 pixels. In our experiments, every image is partitioned into 10×10 regions for all the methods. We compared the performance of LGDP against LBP [1] and LDP [4], and also test the illumination robustness against Gradientfaces [7].

3.1 Neighborhood's size variation

The increment in the neighborhood size increases the input data to the code. However, this data increment not necessarily will lead to incorporate more information into the code. Hence, we analyze the impact of different mask sizes for the face recognition problem. For the proposed method, we use different sizes of the derivative-Gaussian mask that depend on the given width value, σ. For LBP, we change the radius of the neighborhood. And for LDP, we increase the outer ring pattern. The average results of this variation, in the FERET [6] database, are shown in fig. 2a. This figure shows that the average accuracy of LDP and LBP quickly drops as the size of the neighborhood increases. However, LGDP maintains the accuracy throughout the size increment. Despite the close difference at the 3×3 neighborhood, in which LDP and LBP have 0.025% and 0.24% more recognition rate, the combined LGDP outperforms these methods. However, LGDP takes advantage of the sign information, which allows it to distinguish between regions with different intensity transitions and to compactly encode the sign into the structural information, providing a more reliable and stable code, regardless of the neighborhood's size. Therefore, in general the increment in the mask does not assure more discrimination power, if that information is not well used.

3.2 Results on FERET database

To evaluate the expression, pose, and age variation robustness of our method, we perform the evaluation in the FERET [6] database. First, in fig. 2b, we evaluate the LGDP$_\sigma$ code. These results also show that for medium neighborhood's sizes ($0.5 \le \sigma \le 1.5$) the recognition rate peaks. The results in fb, are as good as their combined counterpart—*cf*. fig. 3a. However, in the age variation data sets (dupI and dupII) and the illumination variation data set (fc) their accuracy is lower in comparison to the combined LGDP and the other two methods—*cf*. fig. 3a. Hence, we investigate the inclusion of different resolution masks into one code.

Figure 3a shows the results of three different combined LGDP$_{\sigma_1,\sigma_2,\sigma_3}$ codes. We choose to investigate the combination of the small (LGDP$_{0.3,0.6,0.9}$), medium (LGDP$_{0.5,1.0,1.5}$), and large neighborhoods (LGDP$_{1.0,1.3,1.6}$) sigmas. In general, the results of our codes outperform the results of LDP and LBP in the expression and age variation data sets. For the inten-

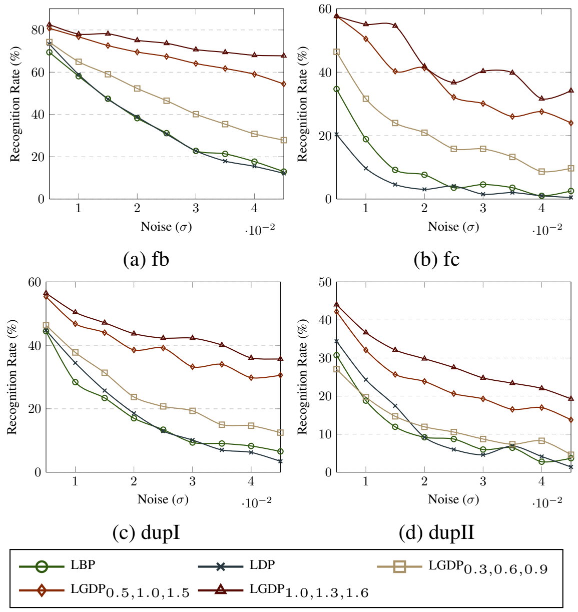


Figure 4. Accuracy on FERET with noise.

sity variation data set (fc), LBP has the same accuracy as the best LGDP code. However, for extreme illumination variation LBP's performance considerably drops in comparison to LGDP—*cf*. fig. 3b. As for the LGDP combined codes, the medium neighborhood code ($LGDP_{0.5,1.0,1.5}$) outperforms the other two. This high accuracy is due to the balanced σ combination that recovers small to large characteristics, instead of picking only small or large characteristics. And the noise robustness (created by perturbing the probe images with Gaussian noise) is shown in fig. 4, where LGDP outperforms other methods due to its derivative-Gaussian and multi-resolution mask. Furthermore, $LGDP_{1.0,1.3,1.6}$ has better performance in the presence of noise over the other two combined LGDP codes.

3.3 Results on Yale B database

We use the Yale B [2] database to evaluate the robustness of our method against illumination variation. The difficulty of this database increases for the subsets four and five, due to the illumination angles that covers half of the face with shadows. Nevertheless, our method is able to recover face features in the dark areas, as it does not rely on intensity like LBP.

We evaluate our method against Gradientfaces, LDP, and LBP; and we show the results in fig. 3b. For the last two sets, the recognition rate of LDP and LBP decreases significantly. The LDP's low accuracy is because the method cannot distinguish intensity changes. On the other hand, LGDP takes advantage of such intensity changes, as it differentiates between similar structures with different intensity changes.

The recognition rate difference, in average in the last two data sets, between gradientfaces and LGDP is of 1.5%. Although gradientfaces has a higher accuracy than LGDP under different illumination, it is not robust against expression and age variation. Gradientfaces has a non-acceptable recognition rate of 7% in fb, and 1% in dupI and dupII in the FERET database. However, LGDP showed to be more reliable in different variation conditions.

4 Conclusion

We introduced a novel encoding scheme, LGDP, that takes advantage of the structure of the face's textures and that encodes them efficiently. LGDP uses directional information that is more stable against noise than intensity, to code the different patterns. Additionally, we use a derivative-Gaussian compass mask to extract this directional information. This mask is stable against noise and illumination variation. The code scheme that we presented, inherently, uses the sign information of the directions which allows it to distinguish similar texture's structures with different intensity transitions. Moreover, we proposed a face descriptor that combines the information from several neighborhoods at different sizes to encode micro patterns at those levels. Hence, LGDP is able to recover more information and use it to increase its discriminating power.

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(12):2037–2041, Dec. 2006.

[2] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):643–660, 2001.

[3] B. Heisele, T. Serre, and T. Poggio. A component-based framework for face detection and identification. *Int. J. Comput. Vision*, 74:167–181, Aug. 2007.

[4] T. Jabid, M. H. Kabir, and O. Chae. Local directional pattern (LDP) for face recognition. In *ICCE*, 2010.

[5] M. Kabir, T. Jabid, and O. Chae. A local directional pattern variance (LDPv) based face descriptor for human facial expression recognition. In *AVSS*, 2010.

[6] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000.

[7] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu. Face recognition under varying illumination using gradientfaces. *IEEE Trans. Image Process.*, 2009.

[8] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, Dec. 2003.