

# Performance Analysis of Retina and DoG Filtering Applied to Face Images for Training Correlation Filters

Everardo Santiago Ramírez<sup>1</sup>, José Ángel González Fraga<sup>1</sup>, Omar Álvarez Xochihua<sup>1</sup>, Everardo Gutierrez López<sup>1</sup>, and Sergio Omar Infante Prieto<sup>2</sup>

<sup>1</sup> Facultad de Ciencias, Universidad Autónoma de Baja California,  
Carretera Transpeninsular Tijuana-Ensenada, Núm. 3917, Colonia Playitas,  
Ensenada, Baja California, C.P. 22860  
{[everardo.santiagoramirez](mailto:everardo.santiagoramirez@uabc.edu.mx), [angel\\_fraga\\_aomar](mailto:angel_fraga_aomar@uabc.edu.mx), [everardo.gutierrez](mailto:everardo.gutierrez@uabc.edu.mx)}@uabc.edu.mx

<sup>2</sup> Facultad de Ingeniería, Arquitectura y Diseño, Universidad Autónoma de Baja California, Carretera Transpeninsular Tijuana-Ensenada, Núm. 3917, Colonia Playitas, Ensenada, Baja California, C.P. 22860  
[sinfante@uabc.edu.mx](mailto:sinfante@uabc.edu.mx)

**Abstract.** Recognizing facial images with nonhomogeneous illumination is a challenging task. Retina and difference of Gaussians filtering have been applied to facial images in order to remove variations in illumination. This paper presents a study on how these preprocessing operations affect or improve the performance of the correlation filters in a face recognition task. These preprocessing operations were applied to CMU and YaleB facial databases, which containing images with homogeneous and nonhomogeneous illumination, respectively. Results show that these operations improve the performance of correlation filters for recognizing facial images with variable illumination.

**Keywords:** Retina Filtering, Difference of Gaussians, Correlation Filters, Facial Recognition.

## 1 Introduction

A facial image captured in an unconstrained environment is exposed to different light sources. These factors cause nonhomogeneous illumination conditions, which degrade the performance of face recognition algorithms that are based in correlation filters. The main problem is related to local shadows that change the appearance of facial features (eyes, nose, mouth and cheeks) and distort their limits (edges). This affects the performance of correlation filters, which use both the form and the content of the object.

Several approaches to the problem of variable illumination in face images have been proposed in the literature. In [1] the authors proposed an algorithm based on DCT (Discrete Cosine Transform) supported by the reduction of brightness gradient, reduction of spectrum components of low-spatial frequency and

spectral characteristics fusion conditioned on the average intensities. Another approach based on DCT was proposed in [2], where the DCT was employed to compensate for illumination variations in the logarithm domain. Because illumination variations lie mainly in the low-frequency band, an appropriate number of DCT coefficients were truncated to minimize variations of the different illumination conditions. In [3], a hybrid approach of PCA (Principal Component Analysis) and correlation filters was presented, while in [4] a low-pass Gaussian filter was used to estimate the illumination on the face image. Although these preprocessing operations improve the global illumination of facial images, the edge of the facial features is not properly recovered.

An approach to address the problem of nonhomogeneous illumination is the retina and difference of Gaussians (DoG) filtering. However, there is no evidence that these preprocessing operations improve the performance of correlation filters. In this paper, we present a study on how the retina and the DoG filtering influence the performance of correlation filters in the facial recognition task. These preprocessing operations were applied to two sets of facial images, one set with homogeneous illumination and the other with nonhomogeneous illumination. The results obtained in this work show that preprocessing improve the performance of the correlation filter for recognizing facial images with nonhomogeneous illumination.

The remainder of the paper is organized as follows. In Section 2, basic concepts of retina filtering, DoG filtering and correlation filters are presented. In Section 3, the employed evaluation methodology and the obtained results are shown. Finally, Section 4 presents the main conclusions of this work.

## 2 Theoretical Basis

### 2.1 Retina Filtering

Retina modeling [5] mimics the performance of the human retina, removing the variations in illumination by combining two nonlinear adaptive functions. Let  $f(x, y)$  be a normalized facial image. The first nonlinear function is a low-pass filter given by:

$$F_1(x, y) = f(x, y) \geq G_1(x, y) + \frac{\overline{f_{in}}}{E_v}, \quad (1)$$

where  $F_1(x, y)$  is the light adaptation factor, the symbol  $\geq$  denotes the convolution operation,  $\overline{f_{in}}$  is the mean of the input image and  $G_1(x, y)$  is a low-pass Gaussian filter, given by:

$$G_1(x, y) = \frac{1}{2\pi\sigma_1^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_1^2}\right). \quad (2)$$

$E_v$  in Eq. 1 influences directly the amount of edge information to be recovered. Small values for  $E_v$  are better than large values for retrieving

edge information of facial features, but they only remove a few variations in illumination. Then,  $f(x, y)$  is processed according to  $F_1(x, y)$ :

$$f_{la_1}(x, y) = (\max\{f(x, y) | + F_1(x, y)\}) \frac{f(x, y)}{f(x, y) + F_1(x, y)}. \quad (3)$$

The term  $(\max\{f(x, y) | + F_1(x, y)\})$  is a normalization factor, where  $\max\{f(x, y) |$  is the maximum intensity value in the input image. The second nonlinear function works similarly. The improved image is given by:

$$f_{la_2}(x, y) = (\max\{f_{la_1}(x, y) | + F_2(x, y)\}) \frac{f_{la_1}(x, y)}{f_{la_1}(x, y) + F_2(x, y)}, \quad (4)$$

where:

$$F_2(x, y) = f_{la_1}(x, y) \geq G_2(x, y) + \frac{f_{la_1}}{E_v}, \quad (5)$$

and

$$G_2(x, y) = \frac{1}{2\pi\sigma_2^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_2^2}\right). \quad (6)$$

## 2.2 Difference of Gaussians

The image  $f_{la_2}(x, y)$  has a uniform texture. In order to provide the filters with a better discrimination capability, an edge enhancement is performed by applying a DoG filter as follows:

$$f_{bip}(x, y) = DoG \geq f_{la_2}(x, y), \quad (7)$$

where DoG is given by:

$$DoG = \frac{1}{2\pi\sigma_{ph}^2} e^{-\frac{x^2+y^2}{2\sigma_{ph}^2}} - \frac{1}{2\pi\sigma_H^2} e^{-\frac{x^2+y^2}{2\sigma_H^2}}, \quad (8)$$

where the terms  $\sigma_{ph}^2$  and  $\sigma_H^2$  correspond to the standard deviations of the low-pass filters.

## 2.3 Composite Correlation Filters

Correlation pattern recognition is based on the selection or creation of a reference signal  $\mathbf{h}$ , called correlation filter, and then determining the degree to which the analyzed image  $f(x, y)$  resembles the reference signal [6]. Applying a correlation filter to a test image yields a correlation plane  $g(x, y)$ :

$$g(x, y) = f(x, y) \geq h(x, y) = \cup^{-1}\{F(u, v) \times H(u, v)\}, \quad (9)$$

where  $F(u, v)$  and  $H(u, v)$  are the Fourier Transforms (FT) of  $f(x, y)$  and  $\mathbf{h}$ , respectively. Since  $\mathbf{h}$  is a vector, it is reshaped to a bidimensional signal denoted

by  $H(u, v)$ . The symbols  $\otimes$  and  $\cup^{-1}$  are the element-wise multiplication, the conjugate complex and the inverse of the FT, respectively.

Let  $f_i(x, y) / T$  be the  $i$ th training image and  $F_i(u, v)$  its FT. The nonlinear filtering [7] applied to  $F_i(u, v)$  is given by:

$$F_i^k(u, v) = \sqrt[k]{F_i(u, v)} \exp(i\varphi_{F_i(u, v)}). \quad (10)$$

where  $0 < k < 1$  is the nonlinearity factor. Let  $\mathbf{x}_j$  be a column-vector obtained by scanning lexicographically  $F_i^k(u, v)$ , and the vectors  $\mathbf{x}_j$  the columns of matrix  $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ , the nonlinear Synthetic Discriminant Function (SDF) filter is given by [8,7]:

$$h = X^k ((X^k)^+ X^k)^{-1} \mathbf{u}, \quad (11)$$

where  $^{-1}$  denotes the inverse of matrix,  $+$  indicates the transpose operation and  $\mathbf{u} = [u_1, u_2, \dots, u_N]^+$  is a vector of size  $T$  containing the desired values at the origin of the correlation output for each training image.

The Minimum Average Correlation Energy (MACE) filter is based on the principles of the SDF, but is focused to produce a sharp, high peak for authentic images. Let  $D = \frac{1}{N^2d} \sum_{i=1}^N (X_i X_i^+)$  be a matrix where  $X_i$  is a diagonal matrix whose elements correspond to  $F_i(u, v)$ . Consider also the matrix  $X$  defined previously, MACE filter is given by [9]:

$$h = D^{-1} X (X^+ D^{-1} X)^{-1} \mathbf{u}. \quad (12)$$

In a similar fashion, the Unconstrained Optimal Trade-off SDF (UOTSDF) filter produces sharp, high peaks, and it is also tolerant to noise. The UOTSDF filter is given by [10]:

$$h = (\alpha D + \overline{1 - \alpha C})^{-1}, \quad (13)$$

where  $C$  is a diagonal matrix whose diagonal elements  $C(k, k)$  represent the noise power spectral. This filter is optimal to be used when the input images are susceptible to noise at low-light conditions. Thus the UOTSDF filter is an optimal choice as well due to its reduced complexity that alleviates the need to invert a Gram matrix present in the MACE filter.

### 3 Retina and DoG Filtering for Enhancing Images for Training Correlation Filters

#### 3.1 Setting Up the Experiment

Figure 1 depicts the process for recognizing a facial image. Given an input image, with either homogeneous or nonhomogeneous illumination, it is preprocessed by the retina and DoG filtering described in subsections 2.1 and 2.2. After the original image is preprocessed, the enhanced image is correlated with stored filters. Each correlation output  $g_i(x, y)$  is analyzed in order to find a correlation

peak and measure its sharpness by the Peak-to-SideLobe Ratio (PSR) metric. A person is correctly recognized if  $PSR \in \tau$ , where  $\tau$  is a recognition threshold defined experimentally. The  $PSR$  metric is given by [9]:

$$PSR = \frac{(peak\ value)\ \mu_{area}}{\sigma_{area}}. \quad (14)$$

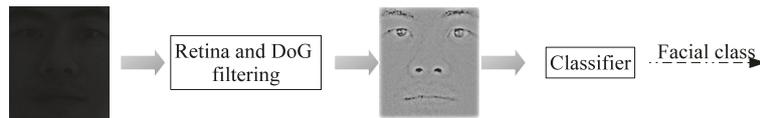


Fig. 1. Process for recognizing a facial image.

The CMU [11] and YaleB[12] facial databases were used to calculate the performance of correlation filters trained with face images preprocessed by retina and DoG filtering. The CMU database contains 13 facial classes. Each facial class consists of a total of 75 different facial images of a same person with homogeneous illumination and different facial expressions. Each image has a resolution of  $64 * 64$  pixels. The YaleB database contains 38 face classes. Each class contains between 60 and 64 different facial images of a same person. Each image has a resolution of  $192 * 168$  pixels captured under nonhomogeneous illumination.



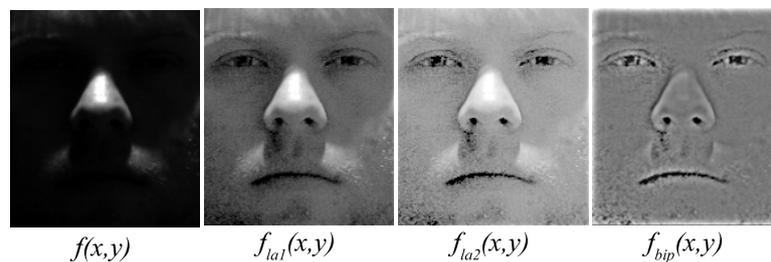
Fig. 2. Sample of training set for the  $i$ th facial class in YaleB (row 1) and CMU (row 2).

Let  $A_j$  be the set of images of the  $j$ th facial class. Each set was split into two subsets  $T_j \cup P_j = A_j$ , where  $T_j$  and  $P_j$  are the sets of training images and the

set of test images. Set  $T_j$  for the CMU database contains four manually selected face images with different facial expressions, while the set  $T_j$  for YaleB database contains four manually selected face images with nonhomogeneous illumination. The test sets  $P_j$  have the following characteristics. For the YaleB test set,  $P_j$  contains between 56 – 60 different facial images, while the CMU test set  $P_j$  contains 71 different facial images. For evaluation purposes, it is assumed that the images contained in  $P_j$  are unknown, so they can be used in order to compute the performance of correlation filters trained with images contained in  $T_j$ .

### 3.2 Numerical Results

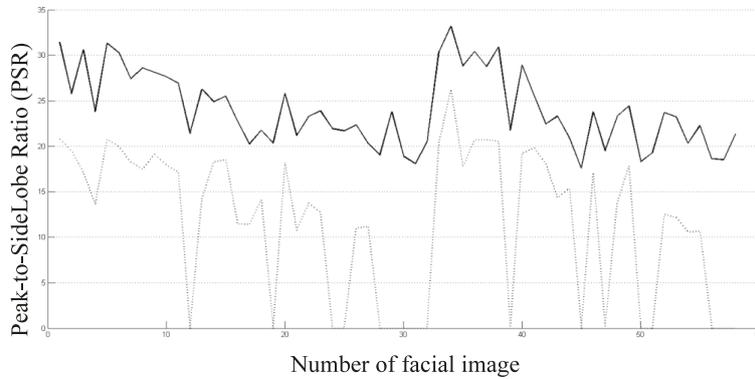
Figure 3 shows the process for improving an image with nonhomogeneous illumination. Firstly, the input image  $f(x, y)$  is filtered by the low-pass filter given in Eq. 4 in order to produce  $f_{la_1}(x, y)$ . Secondly, the light adaptation process is performed by applying the low-pass filter given in Eq. 5 to image  $f_{la_1}(x, y)$  for producing the image  $f_{la_2}(x, y)$ . As it can be observed, the local shadows were removed in the image  $f_{la_2}(x, y)$ . However, the edges of facial features are not easily distinguishable. This problem is solved by applying a DoG filter as is shown in Eq. 7 for obtaining the final image  $f_{bip}(x, y)$ . This filter produces strong edges for the distinct facial features and illumination variations.



**Fig. 3.** Enhancing a facial image by applying retina and DoG filtering. From left to right: input image  $f(x, y)$ ,  $f_{la_1}(x, y)$  is the image  $f(x, y)$  processed by the Gaussian filter given in 4,  $f_{la_2}(x, y)$  is the image  $f_{la_1}(x, y)$  adapted to light according to Eq. 5, and finally the image  $f_{bip}(x, y)$  produced by applying the DoG filter in Eq. 7 to  $f_{la_2}(x, y)$ .

Figure 4 shows the *PSR* performance of a nonlinear SDF filter correlated with 64 different test images of the same class. The *PSR* values for a filter trained with original images (dotted line) are lower than the *PSR* values obtained by a nonlinear SDF filter trained with preprocessed face images (solid line). As it can be observed, enhanced images with retina and DoG filtering are easily recognized with a recognition threshold  $\tau = 10$ .

The FERET protocol states how an evaluation is conducted and how the results are computed in order to measure the performance of face recognition



**Fig. 4.** PSR performance for a nonlinear SDF filter trained with preprocessed face images (solid line) and a nonlinear SDF filter trained with no preprocessed face images (dotted line).

algorithms [13]. Computing performance requires three sets of data. The first is a gallery  $G$  which contains correlation filters trained with sets  $T_j$ . The other two are probe sets. A probe is a face image  $f_j(x, y)$  that is correlated with a filter  $h_j / G$  for recognition, where recognition can be verification or identification. The first probe set is  $P_G$  which contains different face images of people registered in  $G$  (these face images are different from those used to build  $G$ ). The other probe set is  $P_N$ , which contains face images of people who are not in  $G$ .

In the identification task, an algorithm determines if a probe  $f_j(x, y)$  corresponds to a known person by comparing it against filters  $h_j$  stored in  $G$ . If a match is found, then the algorithm identifies the person in the probe  $f_j(x, y)$ . In the verification task, a person presents his biometric sample (probe  $f_j(x, y)$ ) to a system and claims to be a person registered in  $G$ . The algorithm then compares the probe  $f_j(x, y)$  with the stored filter  $h_j$  of the person in the gallery. The claimed identity is accepted if there is a match between the image  $f_j(x, y)$  and the filter  $h_j$ , otherwise, the claimed identity is rejected.

Four FERET metrics were used in this work to compute the performance of the correlation filters. Detection and identification rate (DIR), given in Eq. 15, provides the fraction of probes in  $P_G$  that are correctly identified. False alarm rate (FAR), given in Eq. 16, provides the performance when a probe is not of someone in the gallery (i.e.,  $f_j(x, y) / P_N$ ). This type of probe is also referred to as an imposter. The verification rate (VR), given in Eq. 17, provides the performance when a claimed identity is correctly verified. Finally, Eq. 18 shows the false accept rate (FACR) which provides the times that a claimed identity is declared as authentic, however the identities of the probe  $f_j(x, y)$  and the filter  $h_j$  are different.

$$DIR(\tau, 1) = \frac{\sqrt{\sum \{f_j(x, y) : rank(f_j(x, y)) = 1\}}}{\sqrt{P_B}} \sqrt{\sum \{s_{i,j} \in \tau\}} \quad (15)$$

$$FAR(\tau) = \frac{\sqrt{\sum \{f_j(x, y) : max_i \{s_{i,j}\} \in \tau\}}}{\sqrt{P_N}} \sqrt{\sum \{s_{i,j} \in \tau\}} \quad (16)$$

$$VR(\tau) = \frac{\sqrt{\sum \{f_j(x, y) : s_{i,j} \in \tau \{id(\mathbf{h}_j = id(f_j(x, y)))\}}}{\sqrt{P}} \sqrt{\sum \{s_{i,j} \in \tau\}} \quad (17)$$

$$FACR(\tau) = \frac{\sqrt{\sum \{s_{i,j} : s_{i,j} \in \tau \{id(\mathbf{h}_j \neq id(p_j(x, y)))\}}}{\sqrt{(P-1)B}} \sqrt{\sum \{s_{i,j} \in \tau\}} \quad (18)$$

Function  $rank()$  in Eq. 15 sorts in descending order those results where  $psr \in \tau$ . The identity of a probe  $f_j(x, y)$  is associated to the filter  $\mathbf{h}_j$  with  $rank() = 1$ .  $s_{i,j}$  is the similarity score obtained by correlating  $f_j(x, y)$  with  $\mathbf{h}_j$  and both are the same facial class. On the other hand,  $s_{i,j}$  is the similarity score obtained when  $f_j(x, y)$  is correlated with a filter  $\mathbf{h}_j$ . In this work,  $s_{i,j}$  and  $s_{i,j}$  are computed in terms of the  $PSR$  metric.

**Table 1.** Performance of nonlinear SDF filter using FERET testing protocol.

|                          | CMU   |      |       |      | YaleB |      |       |      |
|--------------------------|-------|------|-------|------|-------|------|-------|------|
|                          | DIR   | FAR  | VR    | FACR | DIR   | FAR  | VR    | FACR |
| Retina and DoG filtering | 98.69 | 0.55 | 98.59 | 3.15 | 95.53 | 5.95 | 95.65 | 0.44 |
| Original images          | 98.68 | 8.18 | 98.78 | 0.08 | 77.76 | 7.47 | 77.90 | 0.03 |

Table 1 shows the results obtained by the nonlinear SDF filter. Retina filtering improves the performance of the filter in terms of DIR and VR metrics on nonhomogeneous illumination conditions. The values for the FAR metric in both databases has decremented highly using the preprocessed images, while the values for FACR has increased slightly.

**Table 2.** Performance of MACE filter using FERET testing protocol.

|                          | CMU   |      |       |      | YaleB |      |       |      |
|--------------------------|-------|------|-------|------|-------|------|-------|------|
|                          | DIR   | FAR  | VR    | FACR | DIR   | FAR  | VR    | FACR |
| Retina and DoG filtering | 98.59 | 0.00 | 98.59 | 0.00 | 83.53 | 0.00 | 83.62 | 0.11 |
| Original images          | 94.15 | 0.00 | 94.15 | 0.00 | 64.50 | 7.58 | 64.82 | 0.20 |

The performance of MACE filter trained with preprocessed images is improved in all FERET metrics, such as can be observed in Table 2. The UOTSDF filter obtained an improved performance in the YaleB database using images enhanced with retina and DoG filters for recognizing face images with nonhomogeneous illumination. However, the performance remained the same for

the CMU database, which indicates that the retina and DoG filtering does not affect the quality of of face images with homogeneous illumination.

**Table 3.** Performance of UOTSDF filter using FERET testing protocol.

|                          | CMU   |      |       |      | YaleB |      |       |      |
|--------------------------|-------|------|-------|------|-------|------|-------|------|
|                          | DIR   | FAR  | VR    | FacR | DIR   | FAR  | VR    | FacR |
| Retina and DoG filtering | 98.37 | 0.29 | 98.33 | 0    | 87.74 | 0    | 87.74 | 0.01 |
| Original images          | 98.37 | 0.29 | 98.33 | 0    | 70.27 | 4.33 | 70.27 | 0.06 |

The results presented in this section showed that the images enhanced with retina and DoG filtering improve the performance of correlation filters. However, the nonlinear SDF filter obtained the best performance reaching values above 95% in DIR and VR metrics.

## 4 Conclusions

It has been presented a study of the effect of retina and DoG filtering applied to face images for training correlation filters for face recognition. Applying these preprocessing operations on facial images removes illumination variations significantly, improving the regions affected by local shadows. The edges of the facial features become visible as well as the small details in the facial skin. These changes improve the overall appearance of the facial image, and improve the performance of the correlation filters. The evaluated filters achieved a recognition rate above 94% for images with homogeneous illumination, while for images with nonhomogeneous illumination the filters reached performances greater than or equal to 70%.

**Acknowledgments.** This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT), with the scholarship number 344833/239152 for author whose name is given first above.

## References

1. Forczmanski, P., Kukharev, G., Shchegoleva, N.L.: An algorithm of face recognition under difficult lighting conditions. *Przeład Electrotechniczny* (2012) 201–204
2. Chen, W., Joo-Er, M., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Transactions on Systems, Man and Cybernetics* **36** (2006) 458–465
3. Savvides, M., Vijaya-Kumar, B., Khosla, P.: corefaces- robust shift invariant pca based correlation filter for illumination tolerant face recognition. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2004)

4. Qu, F., Ren, D., Liu, X., Jing, Z., Yan, L.: A face image illumination quality evaluation method based on gaussian low-pass filter. Proceedings of IEEE CCIS2012 (2012)
5. Vu, N.S., Caplier, A.: Illumination-robust face recognition using retina modeling. In: Proceedings of the 16th IEEE international conference on Image processing. ICIP'09, Piscataway, NJ, USA, IEEE Press (2009) 3253–3256
6. Vijaya-Kumar, B.V.K., Mahalanobis, A., Juday, R.: Correlation pattern recognition. Cambridge University Press (2005)
7. Javidi, B., Wang, W., Zhang, G.: Composite fourier-plane nonlinear filter for distortion-invariant pattern recognition. Society of Photo-Optical Instrumentation Engineers **36** (1997) 2690–2696
8. Casasent, D., Chang, W.T.: Correlation synthetic discriminant functions. Appl. Opt. **25** (1986) 2343–2350
9. Mahalanobis, A., Kumar, B.V.K.V., Casasent, D.: Minimum average correlation energy filters. Appl. Opt. **26** (1987) 3633–3640
10. Kiat-Ng, C.: Pda face recognition system using advanced correlation filters. Master's thesis, Carnegie Mellon University (2005)
11. Lab, A.M.P.: Cmu-face expression database (2001)
12. Georgiades, A., Belhumeur, P., Kriegman, D.: From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Mach. Intelligence **23** (2001) 643–660
13. Phillips, P.J., Grother, P., Micheals, R.: 14. In: Evaluation methods in face recognition. Springer (2005) 329–348