# 6

# Biologically Inspired Processing for Lighting Robust Face Recognition

Ngoc-Son Vu and Alice Caplier
*Grenoble Institute of Technology*
*France*

## 1. Introduction

Due to its wide variety of real-life applications, ranging from user-authentication (access control, ATM) to video surveillance and law enforcements, face recognition has been one of the most active research topics in computer vision and pattern recognition. Also, it has obvious advantages over other biometric techniques, since it is natural, socially well accepted, and notably *non-intrusive*. In reality, several reliable biometrics authentication techniques are available and widely used nowadays (such as iris or fingerprint), but they mostly rely on an active participation of the user. On the contrary, facial biometric demands very little cooperation from the user; thanks to this user-friendly capability, face recognition is said to be non-intrusive.

Over the last decades, significant progress has been achieved in face recognition area. Since the seminal work of Turk and Pentland (Turk & Pentland, 1991), where the Principal Component Analysis (PCA) is proposed to apply to face images (Eigenfaces), more sophisticated techniques for face recognition appear, such as Fisherfaces (Belhumeur et al., 1997), based on linear discriminant analysis (LDA), Elastic Bunch Graph Matching (EBGM) (Wiskott et al., 1997), as well as approaches based upon Support Vector Machines (SVM) (Phillips, 1999), or Hidden Markov Models (HMM) (Nefian & III, 1998; Vu & Caplier, 2010b), etc.

Nevertheless, face recognition, notably under *uncontrolled scenarios*, remains active and unsolved. Among many factors affecting the performance of face recognition systems, illumination is known to be one of the most significant. Indeed, it was proven, both theoretically (Moses et al., 1994) and experimentally (Adini et al., 1997) that image variation due to lighting changes is more significant than that due to different personal identities. In other words, the difference between two face images of the same individual taken under varying lighting conditions is larger than the difference between any two face images taken under the same lighting conditions, as illustrated in Fig. 1.

Inspired by the great ability of human retina that enables the eyes to see objects in different lighting conditions, we present in this chapter a novel method of illumination normalization by simulating the performance of its two layers: the photoreceptors and the outer plexiform layer. Thus, we say the algorithm biologically inspired.

The rest of the chapter is structured as follows: Section 2 briefly discusses the related work; Section 3 presents the model of retinal processing and its advantage. In Section 4, the

Fig. 1. Face appearance varies significantly due to different lighting conditions: (left) face images of two people taken under the same lighting conditions; (right) two face images of the same individual taken under varying lighting conditions.

proposed method is described in detail. Experimental results are presented in Section 5, and conclusion is finally given in Section 6.

## 2. Related work

It is possible to deal with problems of illumination at three different stages in the pipeline of face recognition: during the preprocessing, the feature extraction and the classification. Therefore, existing methods are usually divided into the three following categories:

### 2.1 Illumination invariant feature extraction

The methods of this category try to extract image features which are invariant to illumination changes. It was shown theoretically in (Moses et al., 1994) that in the general case there are no functions of images that are illumination invariant. In (Adini et al., 1997), the authors empirically showed that classical image representations such as edge maps, derivatives of the gray level as well as the image filtered with 2D Gabor-like functions are not sufficient for recognition task under a wide variety of lighting conditions. This observation was later formally proved in (Chen et al., 2000), where the authors showed that for any two images, there is always a family of surfaces, albedos and light sources that could have produced them. Although more recent work, such as Local Binary Patterns (LBP) (Ahonen et al., 2004), Patterns of Oriented Edge Magnitudes (POEM) (Vu & Caplier, 2010a), reveals that certain features are less sensitive to lighting conditions, face recognition based on feature extraction only performs not reliably enough under extreme lighting variations.

### 2.2 Illumination modeling

These approaches require a training set containing several images of the same individual under varying illumination conditions. A training phase is then performed so as to derive a *model* for every identy, which will be used for recognition task. Examples are Illumination Cone (Belhumeur & Kriegman, 1998), Spherical Harmonics (Basri & Jacobs, 2003). Although providing the high quality results in general, these algorithms are costly and in particular they require several images obtained under different lighting conditions for each individual to be recognized. They are therefore impractical for many applications, such as surveillance where there is strict constraint upon the computational time or face recognition in one sample circumstances.

### 2.3 Suppression of illumination variation

The most suitable choice is to deal with lighting variation during the preprocessing step, prior to other stages. Such algorithms transform the image to a canonical form where the illumination variation is erased. Classical algorithms such as histogram equalization, gamma

correction are simple examples whereas the more elaborated techniques are mostly based on properties of the human visual system, evidenced in (Land & McCann, 1971), known as the Retinex theory.

The Retinex theory aims to describe how the human visual system perceives the color/lightness of a natural scene. Our vision ensures that the perceived color/lightness of objects remains relatively constant under varying illumination conditions. This feature helps us identifying objects. Physics says applying red light on a green apple is not the same as applying white light on the same green apple, but our vision attempts to see the same color, regardless of the applied light. The goal of Land's Retinex theory was thus to understand and to find a computational model of how our vision system process the physical stimuli in such a way that color/lightness consistancy is preserved (in the remainder, only the gray images are considered since face recognition techniques perform well on gray images). Assuming that the intensity signal $I(x,y)$ is the product of the illumination $L(x,y)$ and the surface reflectance $R(x,y)$, i.e. $I(x,y) = L(x,y)R(x,y)$, the authors supposed that the reflectance value of a pixel $R(x,y)$ can be computed by taking the ratio of the pixel intensity with the illumination. The problem of obtaining $R$ from an input image $I$ can be solved therefore by estimating $L$. Using this observation, several methods have been presented, such as Single Scale Retinex (SSR), Multi Scale Retinex (MSR) (Jobson et al., 1997) as well as Self-Quotient Image (SQI) (Wang et al., 2004).

Actually, these algorithms are widely used for illumination normalization and also reach the state-of-the-art results. However, they still can not exactly estimate $L$, so large illumination variations are not completely removed. Another disadvantage of those algorithms is that the computational time is still relatively high: both MSR and SQI are "multi-scale" methods which require to estimate the illumination at various "scales". It is also worth noting that the term Retinex coming from the words "Retina" and "Cortex", meaning that both the eyes and the brain are involved in the process. However, to the best of our knowledge, the rule of brain is rather to build a visual representation with vivid details, whereas the natural properties of retina allow our eyes to see and to identify objects in different lighting conditions. That is the motivation for our retina based illumination normalization method.

Before going into details of our retina filter, we need to distinguish the difference between the method proposed in (Tan & Triggs, 2007), referred as PS in the follows, and ours. Although both consist of three steps (see (Tan & Triggs, 2007) for details of the PS method and Section 4 for ours), algorithms used in each step (except the second stage) are different. In their work, the authors do not point out that their algorithm is basically based on the performance of retina. Moreover, we will show that our algorithm is both more efficient and of lower complexity. It is also worth noting that our Gipsa-lab is one of pioneer laboratories on modeling the behavior of the retina, such as (Beaudot, 1994).

## 3. Retina: properties and modeling

The retina lies at the back of the eye (Fig. 2). Basically, it is made of three layers: the photoreceptor layer with cones and rods; the outer plexiform layer (OPL) with horizontal, bipolar and amacrine cells; and the inner plexiform layer (IPL) with ganglion cells. The goal here is not to precisely model the dynamics of retinal processing, such as is done, for example, in (Benoit, 2007). We aim at identifying which processing acts on the retinal signal for illumination normalization. This section demonstrates that bipolar cells not only remove illumination variations and noise but also enhance the image edges. It is worth noting that the

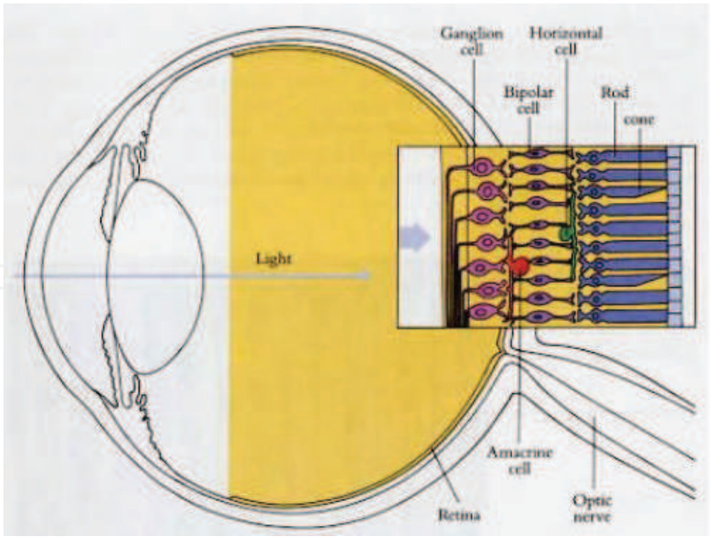



Fig. 2. The retina lies at the back of the eye. Light passes through the bipolar and amacrine cells and reaches the photoreceptors layers where it returns
[http://hubel.med.harvard.edu/bio.htm].

retina is capable to process both spatial and temporal signals but working on static images, we consider only the spatial processing in the retina.

### 3.1 Photoreceptors: light adaptation filter

Rods and cones have quite different properties: rods have the ability to see at night, under conditions of very low illumination (night vision) whereas cones have the ability to deal with bright signals (day vision). In other words, the photoreceptors are is able to adjust the dynamics of light intensity they receive: it plays a crucial rule as light adaptation filter. This property is also called the *adaptive* or *logarithmic compression*.
To exploit and mimic this property, an adaptive nonlinear function is usually applied on the input signal (Benoit, 2007):

$$y = \frac{x}{x + x_0}, \tag{1}$$

where $x$ represents the input light intensity, $x_0$ is the adaptation factor, and $y$ is the adapted signal.
Fig. 3 illustrates the adaptive nonlinear function for different values of $x_0$. If $x_0$ is small, the output has increased sensitively, otherwise when $x_0$ is large, there is not much change in sensitivity.
For an automatic operator, several methods are proposed to determine the adaptation factor $x_0$. One solution is to take $x_0$ equal to the average image intensity. This works if the image intensity is roughly balanced, meaning that the histogram is relatively flat around the mean value. However, when image regions are not lighted similarly, such a function will equalize those image regions identically (see Fig. 4(b)).
Therefore, $x_0$ should vary for each pixel. It can be obtained by applying a low pass filtering on the input image (lighting adapted image using these factors are shown in Fig. 4(c)). Another solution is to combine these two approaches: a low-pass filtering is applied on the input image and for each pixel, the adaptation factor is the sum of the image average intensity and the

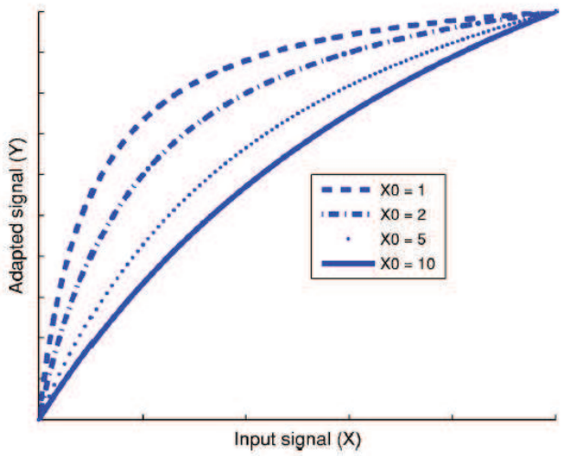



Fig. 3. Performance of nonlinear operations with different adaptation factors $x_0$.

(a) (b) (c) (d)

Fig. 4. (a): original images; images obtained with adaptation factor equal to: (b) the average of image intensity; (c): intensity of low-pass filtered image; (d) the sum of both (b) and (c).

intensity of the low-pass filtered image. Resulting images in Fig. 4 show that after applying the adaptive operators, the local dynamic range in dark regions are enhanced whilst bright regions remain almost unchanged. Among the images (b),(c) and (d), the images (d) are the best lighting adapted. Consequently, this combinational approach will be used in our model.

### 3.2 Outer Plexiform Layer (OPL)

Photoreceptors perform not only as a light adaptation filter but also as a low pass filter. This leads to an image in which the high frequency noise is strongly attenuated and the low frequency visual information is preserved. The signal is then transmitted and processed by horizontal cells which acts as a second low pass filter. Bipolar cells calculate the difference between photoreceptor and horizontal cell responses, meaning that bipolar cells act as a band pass filter: the high frequency noise and low frequency illumination are removed.

To model the behavior of bipolar cells, two low pass filters with different cutoff frequencies corresponding to performance of photoreceptors and horizontal cells (the cutoff frequency of horizontal cells is lower than that of photoreceptors) are often used, and then the difference of these responses is calculated. In our algorithm, two Gaussian low pass filters with different standard deviations corresponding to the effects of photoreceptors and horizontal cells are

used, and bipolar cells act like a Difference of Gaussians filter (*DoG*). As can be seen from Fig. 5, the very high and very low frequencies are eliminated whilst the middle ones are preserved. Note that, another advantage of the DoG filter is the enhancement of the image edges, which is believed useful for recognition task.

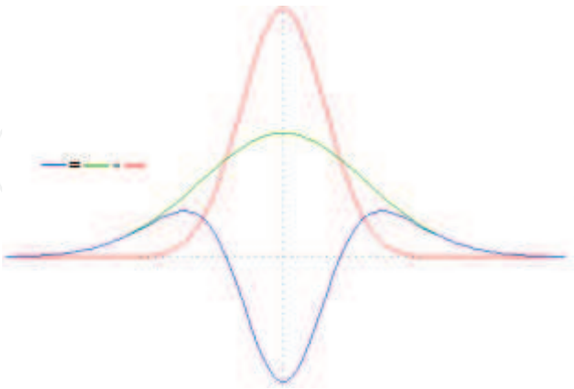


Fig. 5. Difference of Gaussians filter.

### 3.3 Inter Plexiform Layer (IPL)

In this last processing stage of the retina before the optic nerve, the information obtained through the processing of OPL is processed by the ganglion and amacrine cells. However, this layer rather deals with temporal information or movement and therefore is not related to this work.

## 4. Proposed method in detail

As pointed out above, a model with a nonlinear operator and a band pass filter can be used for illumination variation removal. In our model, *multiple* consecutive nonlinear operations are used for a more efficient light adaptation filter. Also, a truncation is used after the band pass filter to enhance the global image contrast.

### 4.1 Multiple logarithmic compressions

In (Meylan et al., 2007), being interested in the property of light adaptation of the retina, the authors modeled the behavior of the *entire* retina by two adaptive compressions, which correspond to the effects of OPL and IPL, respectively. By experiments, they showed that these duplex operations lead to a very good light adaptation filter with a good visual discrimination. Inspired by this observation, we propose to apply several adaptive operations in the first step of our model. In reality, (Vu & Caplier, 2009) already pointed out that using two consecutive adaptive functions leads to the optimal performance on the Yale B database (when images with the most neutral light sources are used ad reference). For an algorithm with generality, this work will automatically determines the optimal number of compressions.
The adaptation factor ($x_0$ in Equation 1) of the first nonlinear function is computed as the sum of the average intensity of the input image and the intensity of the low pass filtered image:

$$F_1(p) = I_{in}(p) * G_1 + \frac{\overline{I_{in}}}{2} \tag{2}$$

where $p = \{x, y\}$ is the current pixel; $F_1(p)$ is the adaptation factor at pixel $p$; $I_{in}$ is the intensity of the input image; $*$ denotes the convolution operation; $\overline{I_{in}}$ is the mean value of the input; and $G_1$ is a 2D Gaussian low pass filter with standard deviation $\sigma_1$:

$$G_1(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} \tag{3}$$

The input image is then processed according to Equation 1 using the adaptation factor $F_1$, leading to $I_{la_1}$ image:

$$I_{la_1}(p) = (I_{in}(max) + F_1(p)) \frac{I_{in}(p)}{I_{in}(p) + F_1(p)} \tag{4}$$

The term $I_{in}(max) + F_1(p)$ is a normalization factor where $I_{in}(max)$ is the maximal value of the image intensity.
The second nonlinear function works similarly, the light adaptation image $I_{la_2}$ is obtained by:

$$I_{la_2}(p) = (I_{la_1}(max) + F_2(p)) \frac{I_{la_1}(p)}{I_{la_1}(p) + F_2(p)} \tag{5}$$

with

$$F_2(p) = I_{la_1}(p) * G_2 + \frac{\overline{I_{la_1}}}{2} \tag{6}$$

and

$$G_2(x, y) = \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}} \tag{7}$$

When using more compressions, the next operations work in the same way and we finally get the image $I_{la_n}$ where $n$ is the number of nonlinear operators used.

(a) $I_{in}$ (b) $I_{la_1}$ (c) $I_{la_2}$ (d) $I_{la_2}$

Fig. 6. Performance of photoreceptors with different parameters. (a): original images; (b): images after one adaptive operator; (c) & (d): images after two operators with different parameters.

Fig. 6 shows the effect of the adaptive nonlinear operator on two images with different parameters. Visually, we observe that the images after two operations (Fig. 6(c),(d)) are

better adapted to lighting than those after a single operation (Fig. 6(b)). Another advantage of two consecutive logarithmic compressions, as argued in (Meylan et al., 2007), is that the resulting image does not depend on low pass filter parameters: we see any difference between the images in Fig. 6(c) ($\sigma_1 = 1, \sigma_2 = 1$) and those in Fig. 6(d) ($\sigma_1 = 1, \sigma_2 = 3$).

### 4.2 Difference of Gaussians filter and truncation

The image $I_{la_n}$ is then transmitted to bipolar cells and processed by using a Difference of Gaussians ($DoG$) filter:

$$I_{bip} = DoG * I_{la_n} \tag{8}$$

where $DoG$ is given by:

$$DoG = \frac{1}{2\pi\sigma_{Ph}^2} e^{-\frac{x^2+y^2}{2\sigma_{Ph}^2}} - \frac{1}{2\pi\sigma_H^2} e^{-\frac{x^2+y^2}{2\sigma_H^2}} \tag{9}$$

The terms $\sigma_{Ph}$ and $\sigma_H$ correspond to the standard deviations of the low pass filters modeling the effects of photoreceptors and horizontal cells.
In fact, the output image at bipolar cells $I_{bip}$ is the difference between the output image at the photoreceptors $I_{Ph}$ and that at horizontal cells $I_H$: $I_{bip} = I_{Ph} - I_H$, where $I_{Ph}$ and $I_H$ are obtained by applying low pass Gaussian filters on the image $I_{la_2}$: $I_{Ph} = G_{Ph} * I_{la_n}, I_H = G_H * I_{la_n}$.

(a) $I_{la_2}$ (b) $I_{Ph}$ (c) $I_H$ (d) $I_{bip}$

Fig. 7. Effect of the Difference of Gaussians filter: $I_{bip} = I_{Ph} - I_H$. (a) image after two non-linear operators; (b) image after by the $1^{st}$ low pass filter at photoreceptors; (c) image after by the $2^{nd}$ low pass filter at horizontal cells; (d) output image at bipolar cells.

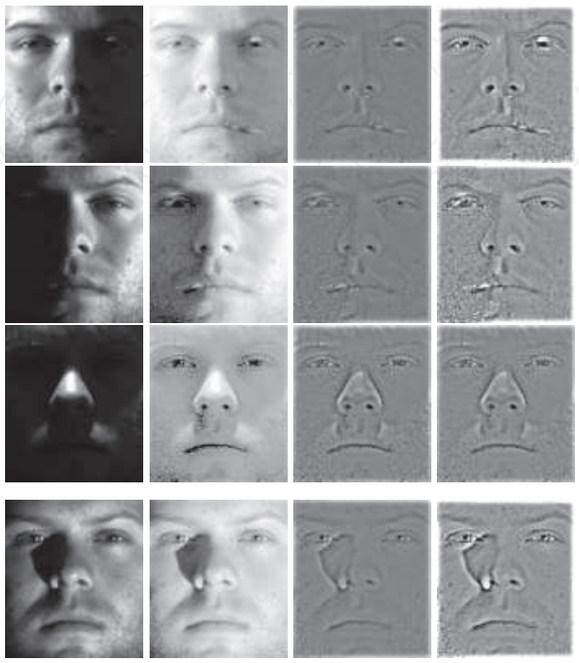
Fig. 7 shows the effect of Difference of Gaussians filter on the output image of two adaptive operators (Fig. 7(a)). We observe that the illumination variations and noise are well suppressed (Fig. 7(d)).
A drawback of the DoG filter is to reduce the inherent global contrast of the image. To improve image contrast, we further propose to remove several extreme values by truncation. To facilitate the truncation, we first use a zero-mean normalization to rescale the dynamic range of the image. The substraction of the mean $\mu_{I_{bip}}$ is not necessary because it is near to 0.

$$I_{nor}(p) = \frac{I_{bip}(p) - \mu_{I_{bip}}}{\sigma_{I_{bip}}} = \frac{I_{bip}(p)}{\sqrt{E(I_{bip}^2)}} \tag{10}$$

After normalization, the image values are well spread and are mainly lied around 0, some extreme values are removed by a truncation with a threshold $Th$ according to:

$$I_{pp}(p) = \begin{cases} max(Th, |I_{nor}(p)|) & \text{if } I_{nor}(p) \geq 0 \\ -max(Th, |I_{nor}(p)|) & \text{otherwise} \end{cases} \tag{11}$$

The threshold $Th$ is selected in such a way that the truncation can remove approximately 2-4% extreme values of image. Fig. 8 shows the main steps of the algorithm. We observe that after the truncation, the overall contrast of the image is improved (Fig. 8(d)). As properties, the proposed algorithm not only removes the illumination variations and noise, but also reinforces the image contours.



(a) $I_{in}$ (b) $I_{la_2}$ (c) $I_{bip}$ (d) $I_{pp}$

Fig. 8. Effects of different stages of the proposed algorithm. (a) input images; (b) images after the light adaptation filter; (c) output images at the bipolar cells; (d) final processed images.

## 5. Experimental results

### 5.1 Experiment setting

The performance of the proposed preprocessing algorithm is evaluated regarding face recognition application. Three recognition methods are considered, including Eigenface (Turk & Pentland, 1991), the Local Binary Patterns (LBP) (Ahonen et al., 2004) based and the Gabor filter based (Liu & Wechsler, 2002) methods:

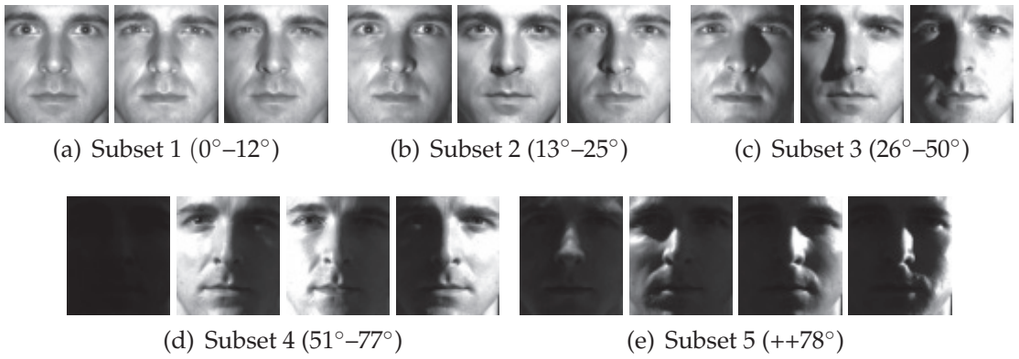
- Although the Eigenface method is very simple (many other methods lead to better recognition rate), the recognition results obtained by blending our preprocessing method with this recognition technique is very interesting because they show the effectiveness of our algorithm.
- Both LBP and Gabor features are considered to be robust to lighting changes, we report the recognition results when combining our preprocessing algorithm with these illumination robust feature based techniques in order to show that such features are not sufficient for wide variety of lighting changes and that our illumination normalization method improves significantly the performance of these methods.

The simple nearest neighbor classification is used to calculate the identification rates in all tests. Experiments are conducted on three databases: Extended Yale B, FERET (frontal faces) and AR, which are described in the rest of this section.

#### 5.1.1 Extended Yale B database

The Yale B face dataset (Georghiades & Belhumeur, 2001) containing 10 people under 64 different illumination conditions has been a *de facto* standard for studying face recognition under variable lighting over the past decade. It was recently updated to the Extended Yale B database (Lee et al., 2005), containing 38 subjects under 64 different lighting conditions. In both cases the images are divided into five subsets according to the angle between the light source direction and the central camera axis (0°–12°; 13°–25°; 26°–50°; 51°–77°; ++78°) (c.f. Fig. 9). Although containing few subjects and little variability of expression, aging, the extreme lighting conditions of the Extended Yale B database still make it challenging for most face recognition methods.

Normally, the images with the most neutral light sources (named "A+000E+00" and an example is shown in the first image of Fig. 9) are used as reference, and all other images are used as probes. But in this work, we conduct more difficult experiments where reference database also contains images of *non-neutral* light sources. We use face images already aligned by the authors and then resize them which are originally of 192 × 168 pixels to 96 × 84 pixels.



(a) Subset 1 (0°–12°) (b) Subset 2 (13°–25°) (c) Subset 3 (26°–50°)

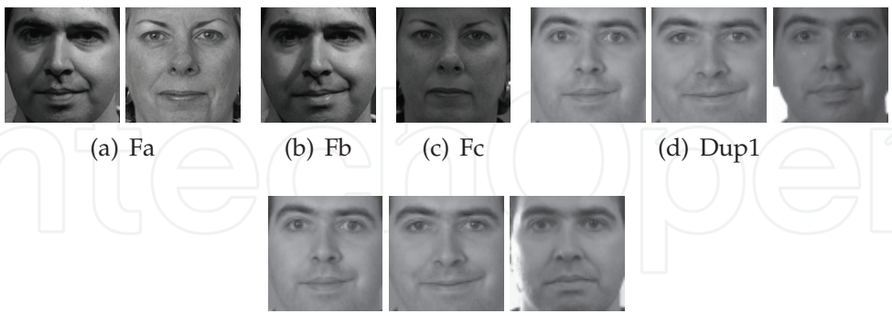
(d) Subset 4 (51°–77°) (e) Subset 5 (++78°)

Fig. 9. Example images from the Extended Yale B database. Images from this challenging database is divided into 5 subsets according to the angle between the light source direction and the central camera axis. (a) Subset 1 with angle between 0° and 12°; (b) Subset 2 with angle between 13° and 25°; (c) Subset 3 with angle between 26° and 50°; (d) Subset 4 with angle between 51° and 77°; (e) Subset 5 with angle larger than 78°. Example of filtered images can be seen in Fig. 8.

#### 5.1.2 FERET database

In the FERET database (Phillips et al., 2000), the most widely adopted benchmark for the evaluation of face recognition algorithms, all frontal face pictures are divided into five categories: Fa, Fb, Fc, Dup1, and Dup2 (see example images in Fig. 10). Fb pictures were taken at the same day as Fa pictures and with the same camera and illumination condition. Fc pictures were taken at the same day as Fa pictures but with different cameras and illumination. Dup1 pictures were taken on different days than Fa pictures but within a year. Dup2 pictures were taken at least one year later than Fa pictures. We follow the standard FERET tests, meaning that 1196 Fa pictures are gallery samples whilst 1195 Fb, 194 Fc, 722 Dup1, and

234 Dup2 pictures are named as Fb, Fc, Dup1, and Dup2 probes, respectively. As alignment, thanks to the available coordinates of eyes, facial images are geometrically aligned in such a way that centers of the two eyes are at fixed positions and images are resized to $96 \times 96$ pixels.



(a) Fa (b) Fb (c) Fc (d) Dup1

(e) Dup2

Fig. 10. Example images of the FERET database. (a): Fa pictures with neutral expressions; (b): Fb pictures taken at the same day as Fa pictures and with the same camera and illumination condition; (c): Fc pictures taken at the same day as Fa pictures but with different cameras and illumination; (d): Dup1 pictures were taken on different days than Fa pictures but within a year; (e): Dup2 pictures were taken at least one year later than Fa pictures.

#### 5.1.3 AR database

The AR database (Martinez & Benavente, 1998) contains over 4000 mug shots of 126 individuals (70 men and 56 women) with different facial expressions, illumination conditions and occlusions. Each subject has up to 26 pictures in two sessions. The first session, containing 13 pictures, named from "AR–01" to "AR–13", includes neutral expression (01), smile (02), anger (03), screaming (04), different lighting (05 - 07), and different occlusions under different lighting (08 - 13). The second session exactly duplicates the first session two weeks later. We used 126 "01" images, one from each subject, as reference and the other images in the first section as probes. In total, we have 12 probe sets, named from "AR–02" to "AR–13". The AR images are cropped and aligned in a similar way as images in the FERET database. Fig. 11 shows example images from this dataset whereas the images shown in Fig. 12 are preprocessed images, from the probe set "AR–07", containing images of two side lights on.

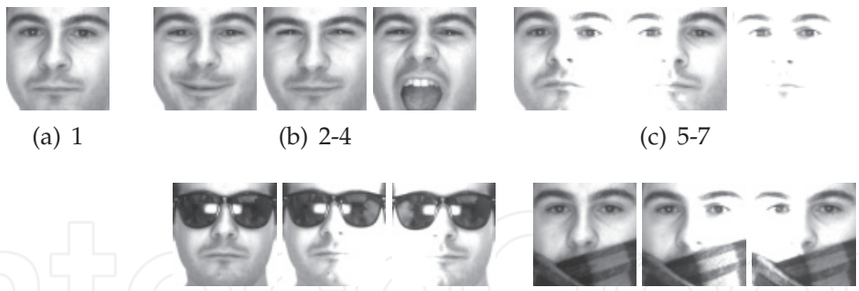
### 5.2 Parameter selection

This section considers how the parameters effect to the filter performance. Parameters varied include the number of compressions and the standard deviations associated ($\sigma_1, \sigma_2$ in the case of two compressions), the standard deviations $\sigma_P, \sigma_H$ and the threshold $Th$. As recognition algorithm, in this section, we use the simple Eigenface method associated and the cosine distance. The Yale B database containing images of 10 different people is used.

#### 5.2.1 Number of compressions et parameters

The compression number $n$ used in the first step is turned variable whilst other parameters are fixed ($\sigma_P = 0.5, \sigma_H = 3.5$, et $Th = 4$). Nearly eight hundred tests were carried out:

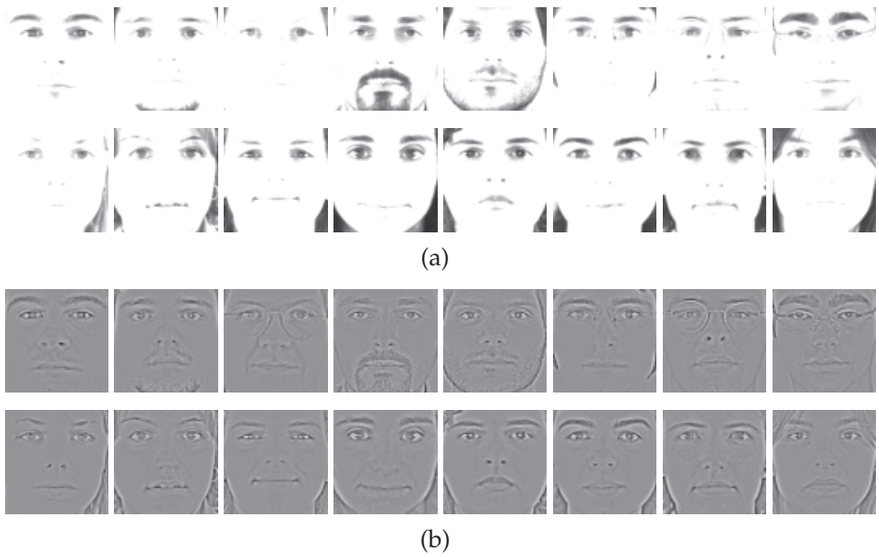
1. $n = 1, 2, 3, 4$. In the follows, the corresponding filters are denoted $\mathbf{F}^1, \mathbf{F}^2, \mathbf{F}^3, \mathbf{F}^4$.



(a) 1 (b) 2-4 (c) 5-7

(d) 8-10 (e) 11-13

Fig. 11. Example images from the AR database. (a) image of neutral expression; (b) images of different expressions: smile (02), anger (03), screaming (04); (c) images of different lighting conditions (05 - 07); (d) & (e) images of different occlusions under different lighting (08 - 13).



(a)

(b)

Fig. 12. Examples of the AR-07 subset: (a) original images; (b) processed images.

2. For each $n$, we vary the standard deviations $\sigma = 1, 2, 3$ (those are used to calculate the adaptation factors). The corresponding filters are referred as $\mathbf{F}^n_{(\sigma_1,...,\sigma_n)}$. In total, 12 filters (3 for each n) are considered.
3. For each filter, 64 different experiments are carried out: for a given illumination angle, 10 images of 10 people in this angle are used as reference and the rest of the database is used as test.
4. We then calculate the average of results obtained across the subset to which reference images belong.

Fig. 13 shows the average recognition rates obtained when the reference images belong to different subsets with different filters (for clarity, we depict only the results of 7 filters). On the horizontal axis of this figure, $(i_1, i_2, ..., i_n)$ means $\mathbf{F}^n_{(\sigma_1,...,\sigma_n)}$. For example, (1) corresponds to $\mathbf{F}^1$ with the standard deviation $\sigma_1 = 1$. It is clear that:

1. Multiple adaptive operations always lead to better compression rates than only one.

| $\sigma_{Ph}, \sigma_H$ | Sous-ensembles | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 0.5 & 3 | 100 | 98.3 | 99.8 | 98.6 | 100 |
| 0.5 & 3.5 | 100 | 98.3 | 99.4 | 97.9 | 100 |
| 1 & 3.5 | 100 | 98.3 | 99.0 | 96.8 | 99.7 |
| 1 & 4 | 99.8 | 97.9 | 96.4 | 95.7 | 99.3 |

Table 1. Recognition rates on the Yale B database for different $\sigma_{Ph}$ & $\sigma_H$.

2. Regarding the performance of multiple compressions, all $\mathbf{F}^2$, $\mathbf{F}^3$, and $\mathbf{F}^4$ perform very well.
3. The performance of $\mathbf{F}^n$ ($n = 2, 3, 4$) is similar: *the values $\sigma_i$ are therefore not important.* In reality, the final results of $\mathbf{F}^3$ are slightly better than those of $\mathbf{F}^2$ and $\mathbf{F}^4$ with the differences of 0.2 and 0.3% respectively. But for complexity constraints, we use $\mathbf{F}^2$ with $\sigma_1 = \sigma_2 = 1$.

The proposed method produces very good results in all cases. When reference images belong to the first four subsets, the subset 1 is the easiest query; when the reference images belong to subset 5, the subset 5 is the easiest test. This is easy to understand. However, surprisingly, the subset 5 is often easier to process than subsets 2, 3, 4 (see Fig. 13 (b), (c) and (d)).

#### 5.2.2 Parameters of DoG and truncation

*DoG* filter parameters are the two standard deviations $\sigma_{Ph}$ and $\sigma_H$ which define the low and high cutoff frequencies of the band pass filter, respectively (refer to Fig. 5). A critical constraint is $\sigma_{Ph} < \sigma_H$. In this work, we choose the values $\sigma_{Ph} \in \{0.5; 1\}$, $\sigma_H \in \{3; 4\}$ [1]. By varying these values in corresponding intervals, we find that $\sigma_{Ph} = 0.5$ and $\sigma_H = 3$ give better results than others. Table 1 shows the average rates obtained when the reference images belong to subset 3.

Regarding the truncation threshold $Th$, we first analyze the distribution of image values. Remind that these values are mainly lied around 0 and their average is 0. By choosing randomly 20 images $I_{nor}$, we find that with a threshold $Th \in \{3, 4\}$, we can remove in average 3-4% extreme values per image. We then evaluate the effect of the retina filter by varying $Th$ in this range and we observe that the obtained results are almost similar. However, if the truncation is not applied, the recognition rate is degraded about 1-2%, depending on the subsets. This shows the effectiveness of the truncation.

**The optimal parameters are:** $\sigma_1 = \sigma_2 = 1, \sigma_{Ph} = 0.5, \sigma_H = 3$, and $Th = 3.5$.

### 5.3 Results on the Extended Yale B database

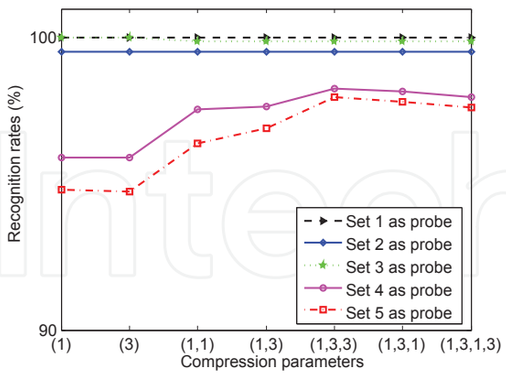
In the follows, we compare the performance of our method with the state of the art methods. Thanks to the "INface tool" software [2], we have codes in Matlab of several illumination normalization methods. Among the available methods, we consider the most representative methods, such as MSR, SQI, and PS [3]. Parameters are used as recommended by the authors.

Table 2 presents results obtained on the Extended Yale B dataset when the reference set contains images acquired under ideal lighting condition (frontal lighting with angle 0°) and the test contains the rest of database. The reported results are divided into two groups: ones

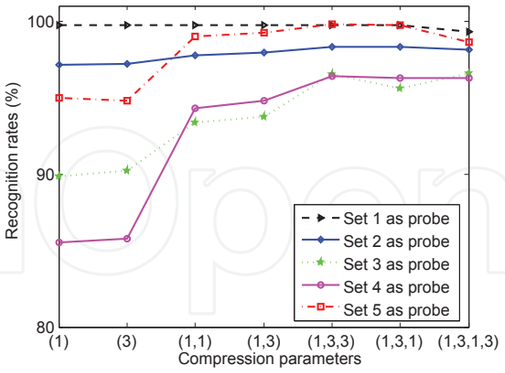
[1] The cutoff frequencies should depend on the quality of images. With a blurred image whose information lies mainly in low frequency, applying a filter with $\sigma_{Ph}$ "too high" will cause a lost of information. For a fully automatic parameter choice, a quality metric images should be used: $\sigma_{Ph}$ and $\sigma_H$ should be chosen in such a way that not two much information of image is removed.

[2] http://uni-lj.academia.edu/VitomirStruc

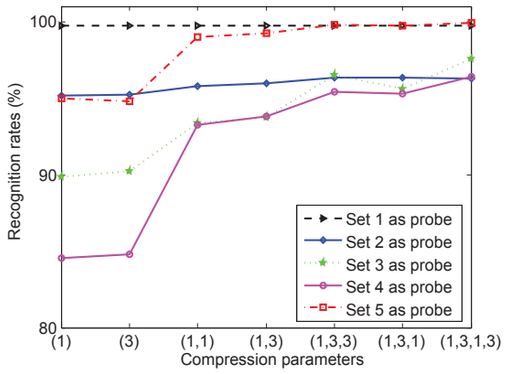
[3] The code for this method is available from http://parnec.nuaa.edu.cn/xtan/



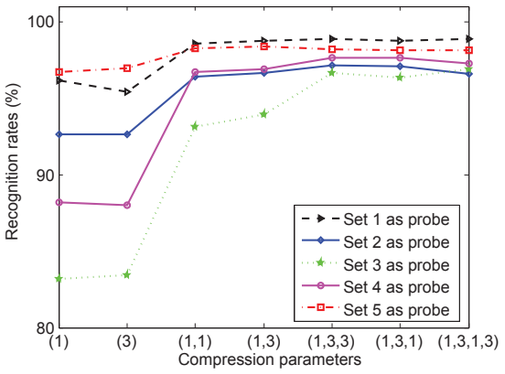

(a) Images in set 1 as reference



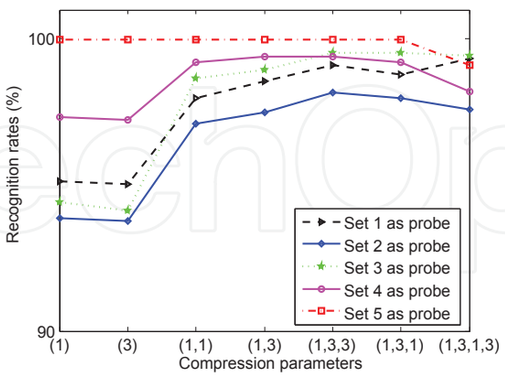

(b) Images in set 2 as reference




(c) Images in set 3 as reference




(d) Images in set 4 as reference




(e) Images in set 5 as reference

Fig. 13. Recognition rates on the Yale B database for different adaptive operations with different parameters.

| | Methods | Subsets | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 10 individuals | Without preprocessing | 100 | 98.3 | 64.2 | 32.9 | 13.7 |
| | Histogram equalization (HE) | 100 | 98.3 | 65.8 | 35 | 32.6 |
| | MSR | 100 | 100 | 96.7 | 85 | 72.1 |
| | SQI (Wang et al., 2004) | 100 | 100 | 98.3 | 88.5 | 79.5 |
| | PS (Tan & Triggs, 2007) | 100 | 100 | 98.4 | 97.9 | 96.7 |
| | LTV (Chen et al., 2006)$^{+}$ | 100 | 100 | 100 | 100 | 100 |
| | **Retina filter** | **100** | **100** | **100** | **100** | **100** |
| | *Cone-cast (Georghiades & Belhumeur, 2001)** | *100* | *100* | *100* | *100* | - |
| | *Harmonic image (Basri & Jacobs, 2003)** | *100* | *100* | *99.7* | *96.9* | - |
| 38 individuals | HE | 98.9 | 97.6 | 56.5 | 23.6 | 21.4 |
| | MSR | 100 | 100 | 96.7 | 79.5 | 65.7 |
| | SQI | 100 | 99.8 | 94.0 | 85.5 | 77.0 |
| | PS | 100 | 99.8 | 99.3 | 99.0 | 96.6 |
| | **Retina filter** | **100** | **100** | **99.7** | **99.3** | **98.8** |

$^{+}$ *This method is about 500 times lower than ours.*
* *These methods belong to the second category which aims at modeling the illumination. These methods require a training set of several images of the same individual under different lighting conditions (e.g. 9 images per person for the Cone cast method). The authors do not report on the subset 5.*

Table 2. Recognition rate obtained on the Extended Yale B when using the Eigenface recognition technique associated with different preprocessing methods and using frontal lighting images as reference.

obtained on the Yale B database containing only 10 individuals and the others obtained on the Extended Yale B database containing 38 individuals (in fact, researchers mostly conduct the experiments on the Yale B database). It can be seen from Table 2 that:

1. The recognition performance drops dramatically very significantly on subsets 4 and 5 when any preprocessing is used (the first row of the table).
2. The proposed method reaches the very strong results and outperforms all competing algorithms on both datasets. On the Yale B database, we obtain the perfect rates, even on the most challenging subset. The LTV method also performs very well but it is about 500 times slower than our algorithm.

We now consider more difficult tests when there are illumination variations on both reference and test images. 30 different tests are carried out on the Extended Yale B set. For each experiment, the reference database contains 38 images of 38 persons for a given angle of illumination and the test database contains all the rest (in fact, there are in total 64 different tests but we randomly choose 30 different lighting conditions: 5 conditions for each of the first four subsets and 10 for the subset 5). Presented in Table 3 are the average recognition rates which clearly show that the proposed method works very well even in challenging test where there are illumination variations on both the reference and probe images.

### 5.4 Results on the FERET database

The aim of this section is to prove the following advantages of our method:

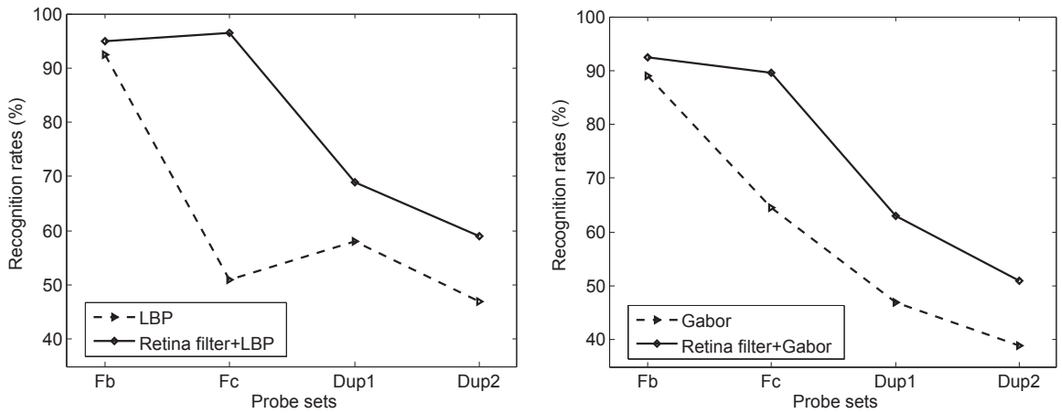
1. It enhances the methods based on features which are considered to be robust to illumination variations.

| Method | MSR | SQI | PS | Proposed |
|---|---|---|---|---|
| Rate | 75.5 | 81.6 | 97.8 | 99.1 |

Table 3. Average recognition rates obtained on the Extended Yale B database when combining the simple Eigenface recognition technique with different preprocessing methods.

2. It improves the face recognition performance in all cases whether there are or not illumination variations on images.

To this end, two features being considered illumination invariant are used for representing face, including LBP (Ahonen et al., 2004) and Gabor wavelets (Liu & Wechsler, 2002). We follow the standard FERET evaluation protocol for reporting the performance: the subset Fa is reference whilst Fb, Fc, Dup1 & Dup2 subsets are probes.



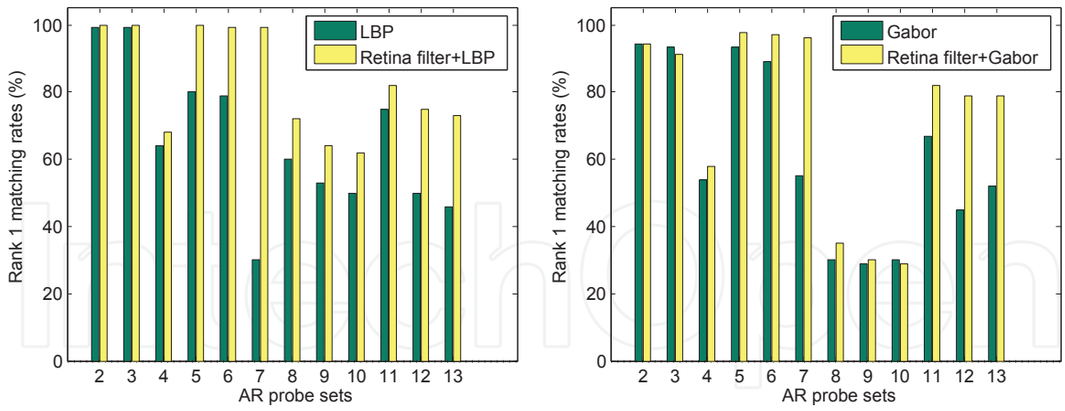

(a) (b)

Fig. 14. Performance of the retina filter on the FERET database when combining with different recognition methods: (a) LBP; (b): Gabor

Fig. 14 clearly shows that the retina preprocessing improves significantly the performance of the two considered methods on all four probes. The considerable improvements obtained on the Fc set confirm that when a good illumination normalization method is used in prior, the robustness of facial features is increased even if these features are considered to be robust to lighting changes. Regarding the results obtained on Fb, Dup1 & Dup2 sets, we can see that the retina filtering is useful for face recognition in all cases whether there are or not lighting variations on images. The reason is that our filter not only removes illumination variations but also enhances the image contours which are important cues for distinguishing individuals.

### 5.5 Results on the AR database

We repeat the similar experiments as in the previous section on the AR database. All 126 images "AR-01" (one for each individual) are used as reference. The recognition results are assessed on 12 probe sets, from "AR02" to "AR13", and are shown in Fig. 15. As can be seen from this figure, our method always leads to very good results. This again proves high efficiency of retinal filter.




(a) (b)

Fig. 15. Performance of the retina filter on the AR database when combining with different recognition methods:: (a) LBP; (b): Gabor

### 5.6 Computational time

This section compares the complexity of different illumination normalization methods to show the advantage of our algorithm. We consider the time required for processing 2000 images of 192 $\times$ 168 pixels and show the average time for one image in Table 4.

| Method | LTV$^+$ | SQI | MSR | PS | Proposed |
|---|---|---|---|---|---|
| Time (s) | 7.3$^+$ | 1.703 | 0.126 | 0.0245 | 0.0156 |

$^+$ *In (Tan & Triggs, 2007), the authors showed that the LTV method is about 300 times slower than the PS method. We estimate therefore that this algorithm is about 500 times slower than ours.*

Table 4. Time required to process an image of 192x168 pixels.

As can be seen from Table 4 or more visually from Figure 16, our method is of very low complexity. Using the code implemented in Matlab (on a desktop of Dual core 2.4 GHz, 2Gb Ram), we can process about 65 images of 192x168 pixels per second; our algorithm is a real-time one. Our method is about 1.57 times faster than the PS method and significantly faster than the others. It maybe worth noting that the most consuming stages in our method are convolutions. Let $mn$ be the image size, $w^2$ the mask size. To reduce the complexity, instead of directly using a 2D mask, which leads to the complexity of $O(mn \times w^2)$, we can use two successive 1D convolutions, leading to a *linear complexity* $O(mn \times 2w)$. In the model, $w = 3\sigma$ where $\sigma$ is the standard deviation of the Gaussian filter. As we use small deviations ($\sigma_1 = \sigma_2 = 1$), the computational time of convolutions is not important. On the contrary, the standard deviations in MSR and SQI are much bigger (e.g. $\sigma_1 = 7, \sigma_2 = 13, \sigma_3 = 20$). That is the reason why our method is very fast.

### 5.7 Illumination normalization for face detection

This section shows another application of our retina filter, i.e. lighting normalization for face detection. Regarding the effects of each step of the proposed model, we observe that using the output image at photoreceptor layer (light adaptation filter) may improve well the face detection performance. For validation, we use the face detector proposed in (Garcia & Delakis, 2004) and calculate the face detection rates of 640 original images in the Yale

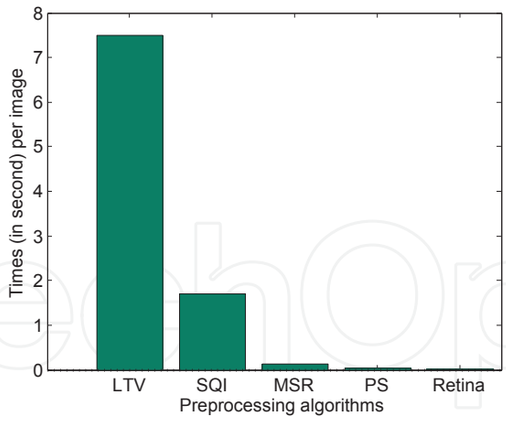
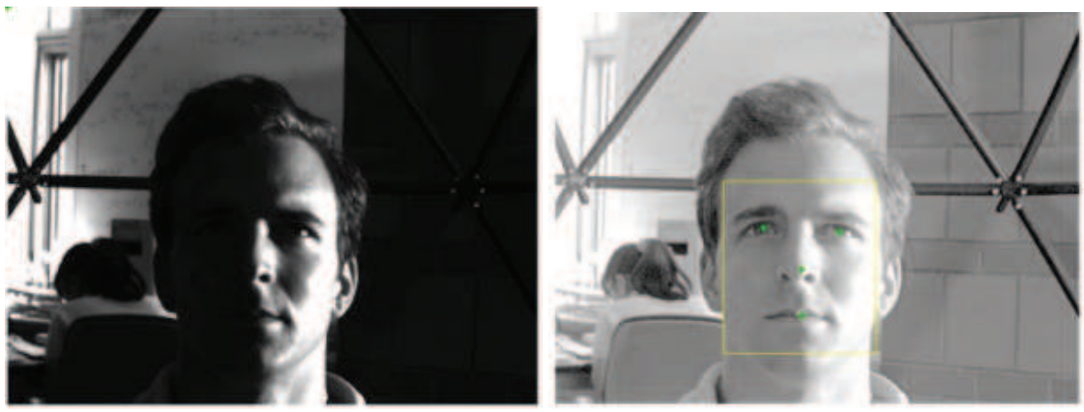



Fig. 16. Average computational time of different algorithms on a 192 $\times$ 168 image.

B database being preprocessed by different methods. Fig. 17(b) shows an example of correct detection and Table 5 compares the performance of different normalization methods. It is clear that our method improves significantly the face detection performance and also outperforms other preprocessing algorithms.



(a) Without preprocessing, face detector does not work

(b) With preprocessing, face detector works well

Fig. 17. Illustration of performance of the proposed algorithm for face detection.

| Method | Without preprocessing | HE | MSR | Proposed |
|---|---|---|---|---|
| Rate (%) | 12 | 98 | 99.0 | 99.5 |

Table 5. Face detection rate on the Yale B database with different preprocessing methods.

## 6. Conclusion

Face recognition has obvious advantages over other biometric techniques, since it is natural, socially well accepted, and non-intrusive. It has attracted substantial attention from various disciplines and contributed to a skyrocketing growth in the literature. Although these attempts, unconstrained face recognition remains active and unsolved. One of the remaining challenges is face recognition across illumination, which is addressed in this chapter. Inspired

by the natural ability of human retina that enables the eyes to see objects in varying illumination conditions, we propose a novel illumination normalization method simulating the performance of retina by combining two adaptive nonlinear functions, a Difference of Gaussian filter and a truncation. The proposed algorithm not only removes the illumination variations and noise, but also reinforces the image contours. Experiments are conducted on three databases (Extended Yale B, FERET and AR) using different face recognition techniques (PCA, LBP, Gabor filters). The very high recognition rates obtained in all tests prove the strength of our algorithm. Considering the computational complexity, ours is a real time algorithm and is faster than many competing methods. The proposed algorithm is also useful for face detection.

## 7. References


Adini, Y., Moses, Y. & Ullman, S. (1997). Face recognition: The problem of compensating for changes in illumination directions, *IEEE Trans. PAMI* 19: 721–732.

Ahonen, T., Hadid, A. & Pietikainen, M. (2004). Face recognition with local binary patterns, *European Conference on Computer Vision*, pp. 469–481.

Basri, R. & Jacobs, D. W. (2003). Lambertian reflectance and linear subspaces, *IEEE Trans. PAMI* 25(2): 218–233.
URL: *http://dx.doi.org/10.1109/TPAMI.2003.1177153*

Beaudot, W. (1994). *The neural information processing in the vertebrate retina: A melting pot of ideas for artificial vision*, PhD thesis, Grenoble Institute of Technology, Grenoble, France.

Belhumeur, P., Hespanha, J. & Kriegman, D. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, *IEEE Trans. PAMI* .
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.3247*

Belhumeur, P. & Kriegman, D. (1998). What is the set of images of an object under all possible illumination conditions?, *Int. J. Comput. Vision* 28(3): 245–260.
URL: *http://dx.doi.org/10.1023/A:1008005721484*

Benoit, A. (2007). *The human visual system as a complete solution for image processing*, PhD thesis, Grenoble Institute of Technology, Grenoble, France.

Chen, H., Belhumeur, P. & Jacobs, D. (2000). In search of illumination invariants, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*.

Chen, T., Yin, W., Zhou, X., Comaniciu, D. & Huang, T. (2006). Total variation models for variable lighting face recognition, *IEEE Trans. PAMI* 28(9): 1519–1524.
URL: *http://dx.doi.org/10.1109/TPAMI.2006.195*

Garcia, C. & Delakis, M. (2004). Convolutional face finder: A neural architecture for fast and robust face detection, *IEEE Trans. PAMI* 26(11): 1408–1423.
URL: *http://dx.doi.org/10.1109/TPAMI.2004.97*

Georghiades, A. & Belhumeur, P. (2001). From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. PAMI* 23: 643–660.

Jobson, D., Rahman, Z. & Woodell, G. (1997). A multiscale retinex for ridging the gap between color images and the human observation of scenes, *IEEE Trans. On Image Processing* 6: 965–976.

Land, E. & McCann, J. (1971). Lightness and retinex theory, *J. Opt. Soc. Am.* 61(1): 1–11.
URL: *http://dx.doi.org/10.1364/JOSA.61.000001*

Lee, K., Ho, J. & Kriegman, D. J. (2005). Acquiring linear subspaces for face recognition under variable lighting, *IEEE Trans. PAMI* 27(5): 684–698.
URL: *http://dx.doi.org/10.1109/TPAMI.2005.92*

Liu, C. & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Trans. Image Processing* 11: 467–476.

Martinez, A. M. & Benavente, R. (1998). The ar face database, *Technical report*.

Meylan, L., Alleysson, D. & Susstrunk, S. (2007). Model of retinal local adaptation for the tone mapping of color filter array images, *Journal of the Optical Society of America A* 24: 2807–2816.
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.109.2728*

Moses, Y., Adini, Y. & Ullman, S. (1994). Face recognition: The problem of compensating for changes in illumination direction, *European Conference on Computer Vision*.

Nefian, A. & III, M. H. (1998). Hidden markov models for face recognition, *ICASSP*, pp. 2721–2724.

Phillips, J. (1999). Support vector machines applied to face recognition, *NIPS*, Vol. 11, pp. 803–809.
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.2690*

Phillips, J., H.Moon & et al., S. R. (2000). The feret evaluation methodology for face-recognition algorithms, *IEEE Trans. PAMI* 22: 1090–1104.

Tan, X. & Triggs, B. (2007). Enhanced local texture feature sets for face recognition under difficult lighting conditions, *AMFG*, pp. 168–182.
URL: *http://portal.acm.org/citation.cfm?id=1775256.1775272*

Turk, M. & Pentland, A. (1991). Eigenfaces for recognition, *J. Cognitive Neuroscience 3* pp. 71–86.

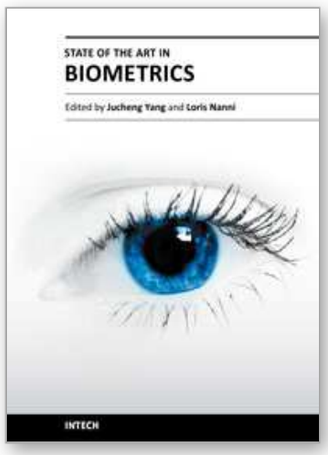
Vu, N.-S. & Caplier, A. (2009). Illumination-robust face recognition using the retina modelling, *International Conference on Image Processing*, IEEE.

Vu, N.-S. & Caplier, A. (2010a). Face recognition with patterns of oriented edge magnitudes, *European Conference on Computer Vision*.

Vu, N.-S. & Caplier, A. (2010b). Patch-based similarity hmms for face recognition with a single reference image, *International Conference on Pattern Recognition*.

Wang, H., Li, S. & Wang, Y. (2004). Generalized quotient image, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (2)*, pp. 498–505.

Wiskott, L., Fellous, J. M., Kuiger, N. & von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching, *IEEE Trans. PAMI* 19(7): 775–779.
URL: *http://dx.doi.org/10.1109/34.598235*




**State of the art in Biometrics**

Edited by Dr. Jucheng Yang

ISBN 978-953-307-489-4
Hard cover, 314 pages
**Publisher** InTech
**Published online** 27, July, 2011
**Published in print edition** July, 2011

Biometric recognition is one of the most widely studied problems in computer science. The use of biometrics techniques, such as face, fingerprints, iris and ears is a solution for obtaining a secure personal identification. However, the â€œoldâ€ biometrics identification techniques are out of date. This goal of this book is to provide the reader with the most up to date research performed in biometric recognition and descript some novel methods of biometrics, emphasis on the state of the art skills. The book consists of 15 chapters, each focusing on a most up to date issue. The chapters are divided into five sections- fingerprint recognition, face recognition, iris recognition, other biometrics and biometrics security. The book was reviewed by editors Dr. Jucheng Yang and Dr. Loris Nanni. We deeply appreciate the efforts of our guest editors: Dr. Girija Chetty, Dr. Norman Poh, Dr. Jianjiang Feng, Dr. Dongsun Park and Dr. Sook Yoon, as well as a number of anonymous reviewers



**InTech Europe**
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

**InTech China**
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen