2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Similarity Michaelis-Menten Law Pre-processing Descriptor for Face Recognition

Suli Ji[1], Baochang Zhang[1,2]*, Dandan Du[1] ,Biao He[1]
1. Science and Technology on Aircraft Control Labority
School of Automation Science and Electrical Engineering
Beihang University
Beijing, China
2. Pattern Analysis and Computer Vision (PAVIS)
Istituto Italiano di Tecnologia (IIT)
Via Morego 30, 16163, Italy
Suliji@asee.buaa.edu.cn, baochang.zhang@iit.it

Jianzhuang Liu
Shenzhen Key Lab for CVPR, Shenzhen Institutes and Advanced Technology, Chinese Academic of Sciences, Chinese Department of Information Engineering, The Chinese University of Hong Kong, China Media LBP, Huawei Technologies Co. Ltd
jzliu@ie.cuhk.edu.hk

***Abstract*—This paper presents a non-linear pre-processing method based on Similarity Michaelis-Menten law (SMML) for face recognition. Similarity Michaelis-Menten law can be used to explain visual sensitivity in the vertebrate retina. We preprocess input images using SMML, and then employ Local Binary Pattern (LBP) for face feature extraction. Advantages of SMML include improvement of light adaption, noise effect, detection right rate, robustness and efficiency, which inspire us exploit it for face pre-processing descriptor for the first time in the field of face recognition. And the parameters of SMML are spatiotemporally and locally estimated by the input image itself employing Sobel , which shows its advantages for face recognition. Extensive experiments clearly demonstrate the superiority of our method over the ones which only use LBP on FERET database in many aspects including the robustness against different facial expressions, lighting and aging of the subjects.**

Keywords—retina; Michaelis-Menten law; LBP; face recognition

I. INTRODUCTION

The intensity of an image is the main element from a camera used for face recognition. Many extrinsic factors, such as illumination variations , various kinds of captured scenes and cameras, appear on face images, which degrades the performance of face recognition systems dramatically. In order to overcome these problems, many methods, such as various image preprocessing, feature extraction, and feature selection methods are investigated. Ojala et al. introduce LBP operator [1], which is a powerful means of texture description and labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. Later the operator was extended to use neigborhoods of different sizes [2], which uses circular neighborhoods and bilinearly interpolating the pixel values allow any radius and number of pixels in the neighborhood. Ahonen *et al.* presented a approach based on Local Binary Pattern (LBP) histograms for face recognition [3], which considered both shape and texture information in representing face images and achieved a good result for face recognition systems on the standard FERET database.

This paper gives much attention on the image preprocessing methods. There are many kinds of preprocessing methods for face recognition, including filtering, grayscale manipulation, sharpening and so on. Beaudot W.H.A discussed sensory coding in the vertebrate retina [4], which was about an adaptive control of visual sensitivity. Jeanny Hérault and Barthélémy Durette [5] presented a model of the retina, whose properties include sampling, spatiotemporal filtering, color-coding and non-linearity, and their consequences on the processing of visual information. Meylan et al proposed image contrast enhancement model based on the local retinal [8]. In retina modeling, simulation for the performance of the human retina is used in the method of illumination normalization. There are two layers of the retina in human eye which are outer plexiform layer and photoreceptors as shown in Fig. 1. In [9], two adaptive nonlinear functions and a Difference of Gaussians filter (DoG) are combined together for illumination normalization. The model achieved good performance in face recognition.

We propose a new preprocessing method, Similarity Michaelis-Menten Law (SMML), for face recognition based on Local Binary Pattern (LBP). SMML is successfully applied to various kinds of captured scenes of different dynamic ranges and different keys. Unlike other methods that work well only with certain kinds of images, the results of SMML show that the proposed preprocessing operator successfully and significantly improves the performance of the original LBP methods in all cases. SMML can be considered as an improvement of the retina filter. The motivation is that natural ability of human retina enables human to see the objects in a variable lighting condition.

The rest of this paper is organized as follows: In Section II, the retina architecture and model of signal processing are briefly introduced. In Section III, we describe and analyze the theory of SMML. In Section IV, the proposed method is described in detail. Finally, experimental results and discussion are given in Section V.

II. THE RETINA ARCHITECTURE AND MODEL OF SIGNAL PROCESSING REVIEW STAGE

A. Architecture of Our Human Eyes and Retina

Fig. 1 is the structure of human eye (left) with a thin piece of retina enlarged in a photomicrograph (right), revealing its architecture of different layers (drawing by Helge Kolb (4)). The retina consists 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 photoreceptor layer implements adaptive light intensity adjustment, and the inner plexiform layer aims to enhance local contrast. Researchers combine two adaptive nonlinear functions and a Difference of Gaussians filter [9], which is related to the performance of two layers of the retina: the photoreceptor and the outer plexiform layer.

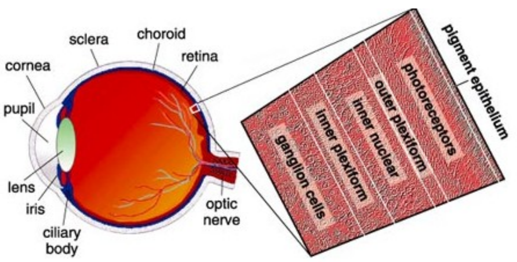

Fig. 1. Human eyes structures and the layers of a thin piece of retina

B. Process of signals transmission

The photoreceptor layer is made of two categories of sensory cells (Fig. 2): cones and rods. Cones are dedicated to photopic vision and colors. Rods are active in scotopic vision. So rods are generally used for low-light vision and cones for daylight, bright-colored vision. Although both rods and cones respond to light, they report quite different image properties. Rods, detecting dim light, usually respond to relatively slow changes. Cones, dealing with bright signals, can detect rapid light fluctuations. In both cases, photoreceptors begin the process of decomposing images into separate parts. Both rods and cones respond to light directly over them. However, their receptive fields are very narrow. An image continues to be broken into component elements at the first synapses of the visual pathway, these synapses are between photoreceptors and bipolar cells. Amacrine cells receive information from bipolar cells and transmit information to ganglion cells , other bipolar and amacrine cells.

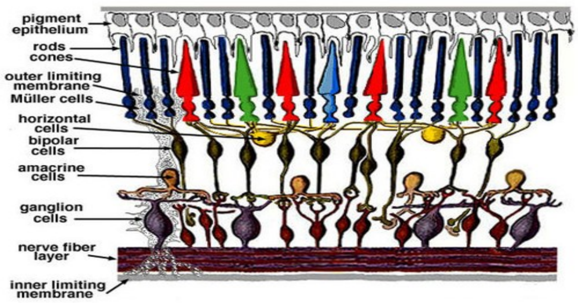

Fig. 2. The architecture of retina

III. THEORY OF SMML

A. An Electrical Model of Signal Processing in the Retina

Among neuron cells, only ganglion cells and a few amacrine cells can produce action potential. A biophysical first-order model of phototransduction is as follows. At a molecular level, they can be described and modeled by a cascade of enzymatic reactions. As we know previously in Fig.2, phototransduction is a complex process, and we only apply some aspects, and we will therefore limit our scope to the membrane stage of transduction, that is to say the action of light-dependent messenger is in membrane channels which account for the generation of presynaptic potential.

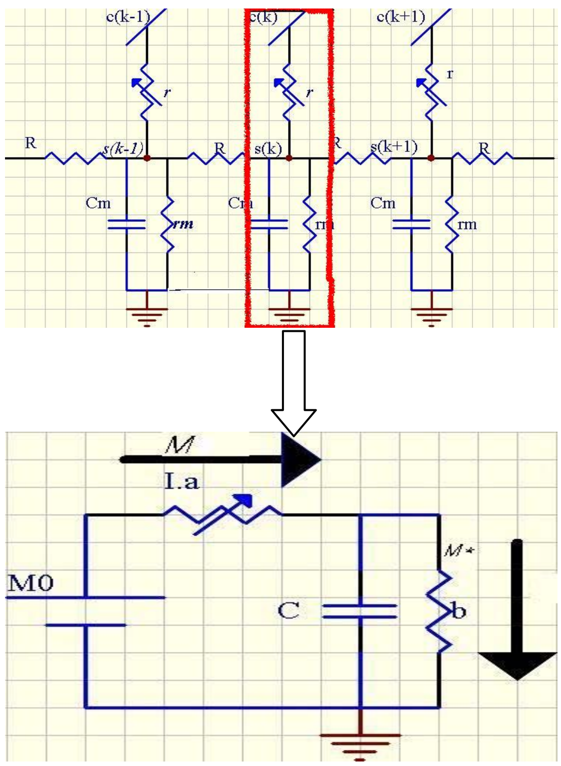

Fig. 3. The above is the electrical model between neighboring photoreceptors, and the below is a molecular simplified model (c(k) to M_0, s(k) to M^*, Cm to C, rm to b and r to I.a)

In the above electrical model, a membrane capacity and a leakage conductance in the red box together form tuned circuits and show the function of the photoreceptors, which is adaptive for light intensity. The mathematical expression of Fig.3 for modeling the binding of the internal messenger onto a membrane channel [2], is as followed:

$$C\frac{\partial M^*}{\partial t}+bM^*=Ia(M_0-M^*) \quad (1)$$

M^* and M_0 denote voltages. If β is related to the ratio b/C where b is a leakage conductance, C to a membrane capacity, M_0 to a potential generator, and α to the ratio a/C which is a conductance modulated by the input signal I, then we have:

$$\partial M^*/\partial t=\alpha I(M_0-M^*)-\beta M^* \quad (2)$$

Despite its simplicity, this model can account for several properties well suited for adaptation to a varying visual environment. The solution of (2) is:

$$M^*=Exp(-(\alpha I+\beta)*t)+\frac{IM_0}{I+\sigma} \quad (3)$$

Where $\sigma=\frac{\beta}{\alpha}$. As a temporal consequence: when the ambient light is high, σ increases and the overall gain dV/dI decreases; when it is low, σ decreases and the overall gain increases. If t is a constant, the solution of above equation (2) changes into:

$$M^*=\frac{IM_0}{I+\sigma} \quad (4)$$

Where $\sigma=\frac{\beta}{\alpha}$.

B. Michaelis-Menten Law and Its Parameters

The model mentioned (subsection III.A) expresses an overall property of photoreceptors: nonlinear and compressive. Particularly when stimulated with a constant light intensity I, equation (2) shows two basic properties: (i) its fast dynamics exhibits a temporal low-pass filtering of the input signal for a time constant (up to several tens of milliseconds) being inversely proportional to light intensity I, and (ii) the response, at steady state, follows a Michaelis-Menten law such as (4).

$$V(I)=\frac{IV_{\max}}{I+\sigma} \quad \sigma=\frac{\beta}{\alpha} \quad (5)$$

Where $V_{\max}$ is the saturation value of the potential equal to M_0, and σ is a dissociation constant which acts on the compression effect. This relation becomes quasi-linear with a slope $V_{\max}/\sigma$ when $I<<\sigma$, and converges towards $V_{\max}$ when $\sigma<<I$. Such an intensity response relation has been observed in direct photocurrent measurement as well as in voltage recording in cone segments of salamander and turtle. Fig. 4 illustrates the Michaelis-Mentren equation for different values

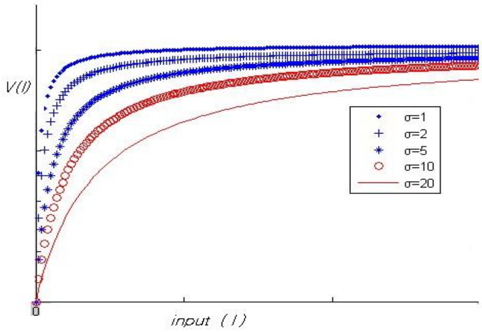

Fig. 4. Michaelis-Mentren equation for different values of σ

of σ when V_{max} equals 255, output $V(I)$ has higher sensitivity for small value of σ. However, for high values of σ, the sensitivity of the output shows a very small change. Hence, Michaelis-Menten equation enhances the local dynamic range in dark area of image whereas the bright area remains no effected.

C. Preprocessing Model based on Similarity Michaelis-Menten Law

The range of light intensities coded by a photoreceptor is incredibly wide: from 1 to 10^6, which do not comply with neurons, where a maximal range of 1.5-2 decades is allowable. Fortunately, in current life we never meet sudden variations of a 10^6 range of intensities. This makes possible the use of an adaptive process to the mean ambient light. The adaptation law resides at the level of photoreceptors, it is typically a Similarity Michaelis-Menten law, for which the response x to a stimulus intensity X is [2]:

$$x=\frac{X^n}{X^n+X_0^n} \quad (6)$$

We are inspired from (5) and (6), and get

$$V(I(p))=\frac{I(p)^n}{(I(p)+\sigma)^n}*I_{\max} \quad (7)$$

where p is the current pixel; σ is the local adaptation factor at pixel p; $I(p)$ is the intensity of the input image; I_{max} is the maximum value of intensity.

We preprocess the images by (7), and then apply the LBP for face recognition. $\frac{I(p)^n}{(I(p)+\sigma)^n}$ plays the part of compression for images. However, the key to tackling the problem lies in how to choose σ and n. We do extensive experiments to find σ, which is related to the local intensity of the input images and is changeable with the local light. Intensity images are treated by an enhancement process to elevate the intensity values of low-intensity (dark) pixels using a specifically designed nonlinear transfer function defined by (7). Fig. 4 motivates us to think about some contour operators. Among all the available operators such as Sobel operator, Laplacian operator, Roberts operator and Scharr operator, we find Sobel operator is best suitable. The exponent n usually takes values among 1, 2, 3, 4, 5, and experiments show that n has little influence on the FERET sets, but n can be changed on different conditions. Thus, we will consider n=1 in the sequel. The parameter σ is capable of adaptation, mainly under the molecular dynamics of light transduction. Its adaptation is respectively thought about from biology and graphics:

(1) In biology, Due to the Horizontal cells feedback and Calcium ions dynamics in photoreceptors, it also adapts to the neighborhood activity. The model is that of a spatial low-pass filtering.

(2) In graphics, the adaptaion factor σ is related to his neighborhood pixels, and chages with the intensity of neighborhood pixels. The model is an adaptive one.

IV. Face Description With SMML and LBP

A. Local Binary Patterns (LBP)

Ojala et al firstly introduces LBP operator [1], which is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3-neighborhood of each pixel with the center value and the values of the pixels in the thresholded neighborhood [Fig. 5(b)] are multiplied by the weights given to the corresponding pixels [Fig. 5(c)]. The result for this example is shown in [Fig. 5(d)]. Finally, the values of the eight pixels are summed to obtain the number (169) of this texture unit and considering the result as a binary number. Fig.5 is for an illustration of the basic LBP operator.

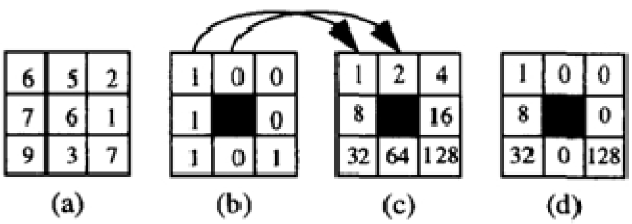

Fig. 5. Basic LBP operator

To be able to deal with textures at different scales, the LBP operator was later extended to use neighborhoods of different sizes, Defining the local neighborhood as a set of sampling points evenly spaced on a circle centered at the pixel to be labeled allows any radius and number of sampling points. Bilinear interpolation is used when a sampling point does not fall in the center of a pixel.

Another extension to the original operator is the definition of so called uniform patterns. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or from 1 to 0.

In our experiment, we choose the basic LBP, then use the histogram of the labels as a texture descriptor. A histogram of the labeled image $f_l(x,y)$ can be defined as

$$H_i = \sum_{x,y} I\{f_l(x,y) > f_m(x,y)\}, \quad i = 0,...,n-1, \tag{8}$$

In which n is the number of different labels produced by the LBP operator .

$$I\{A\} = \begin{cases} 1, A & is \;\; true \\ 0, A & is \;\; false \end{cases} \tag{9}$$

This histogram contains information about the distribution of the local micro-patterns, such as edges, spots and flat areas, over the whole image.

3x3-neighborhood has 256 kinds of different result. Efficient face representation includes much spatial information. The image is divided into regions $R_0, R_1 \dots R_{m-1}$ [2] and the spatially enhanced histogram is defined as

$$H_{x,y} = \sum_{x,y} I\{f_l(x,y) = i\} I\{(x,y) \in R_j\}, i = 0,...m-1. \tag{10}$$

In this histogram, we effectively have a description of the face on three different level of locality: the labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

B. Sobel Operator and SMML Pre-processing

The Sobel operator is used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. On the other hand, the gradient approximation that it produces is relatively crude, in particular for high frequency variations in the image.

Sobel operator has 3x3 convolution kernel and 5x5 convolution kernel. We use two 3x3 convolution kernel G (Gx and Gy), as shown in Fig. 6, which are convolved with the original input image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If we define I_{org} as the source image, then we can get output value I_0 of each pixel and replace σ with it before calculating LBP.

$$I_o = I_{org} * \mathrm{G} \tag{11}$$

* denotes the convolution operation.

1	**2**	**1**
0	**0**	**0**
-1	**-2**	**-1**

(a) Gx

1	**0**	**-1**
2	**0**	**-2**
1	**0**	**-1**

(b) Gy

Fig. 6. (a) and (b) are the Sobel convolution kernel: Horizontal direction operator Gx and Vertical direction operator Gy

We employ our proposed algorithm to process some images showing in Fig. 6. The following procedure is adopted:

(1) Two-direction Sobel operators are employed to find the edges of images, and we choose the one with big convolution value and the kernel size is set to 3x3.

(2) SSML simulates the performance of the adaptation of illumination by function (7). In this step, we preprocess the images on a per-pixel. Function (7), the notation of *I(p)* denotes the garyscale value of input images. σ is the value that step (1) gets. σ can change with convolution value, witch show illumination adaptation.

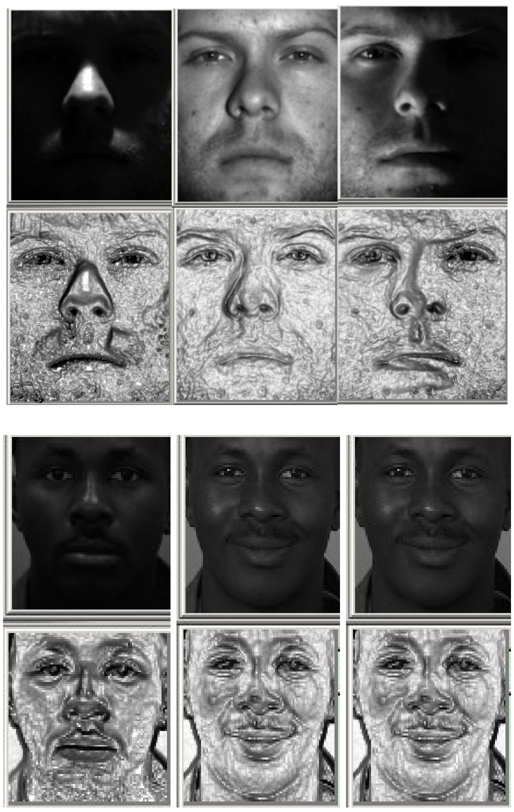

Fig. 7. Input images and processed images by the proposed algorithm

V. EXPERIMENTAL RESULTS AND DISCUSSION

The comparative experiments between LBP and our method are done on the FERET face database, which is widely exploited to evaluate face recognition algorithms. In the experiments, the cropped face is normalized to 88×88 pixels. We use the same gallery and probe sets as in the standard FERET evaluation protocol. Fa containing 1196 frontal images of 1196 subjects is used as the gallery set, while Fb, Fc, Dup 1 and Dup 2 are used as the probe sets. As shown in Fig. 8 and Fig. 9, the recognition accuracy of our method is higher than that of LBP on the FERET dataset. For example, we get 5% improvement than LBP in the Dup 1 set, 17% in Dup2, and 28% in Fc. We also employ different operators to find an light adaptive one, and extensive experiments show Sobel operator is better than others in Fig. 10.

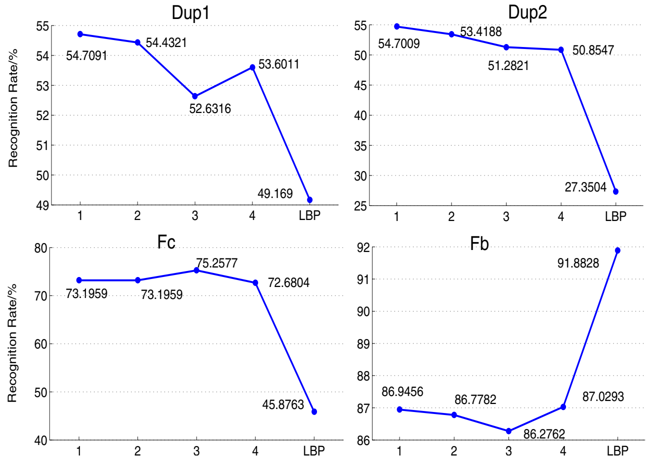

Fig. 8. Shows the result through preprocessing input images with function (6) (n=1 2 3 4) and LBP get in Dup1, Dup2, Fb, Fc. The vertical axis is face recognition rate

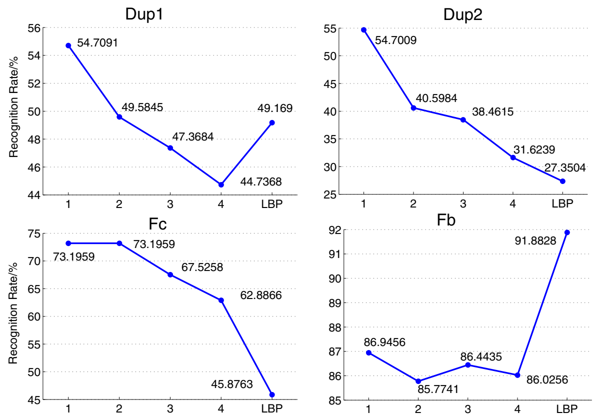

Fig. 9. Shows the result that equation (6) preprocesses the input images (n=1 2 3 4) and then extract LBP feature for face recognition in Dup1, Dup2, Fb, Fc. The vertical axis is face recognition rate

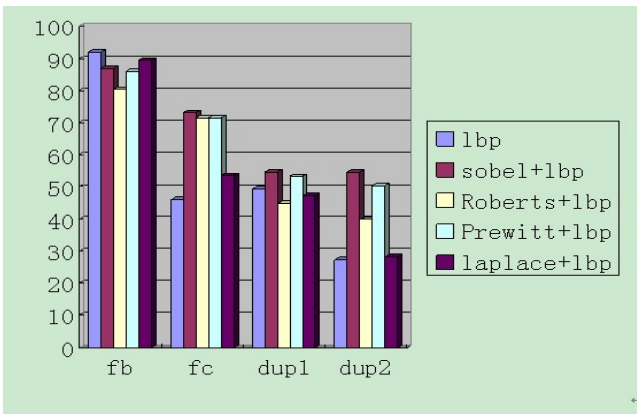

Fig. 10. σ gets different values, when different operators convolute input images. Thus we achieve different recognition rates in FERET database. The vertical axis is face recognition rate. (n=1)

Fig. 10. shows that Sobel preprocessing and LBP can get better recognition rate. So we choose Sobel. The proposed method has been compared with some methods and the results are shown in Table 1. From the recognition rate in Table 1, we can see that our proposed method has improved the recognition rate compared to other algorithms.

TABLE I
Recognition rates of different methods on FERET database

Method	Recognition rates			
	Fb	Fc	Dup1	Dup2
LBP	91.8%	45.8%	49.2%	27.3%
LBP + proposed preprocessing	86.9%	73.2%	54.7 %	54.7%
5-benchmark+5-pyramid LBP	87%	22%	54%	33%
5-benchmark+5-pyramidLBP+proposed preprocessing	85.6%	60.8%	57%	49.5%

A series of experiments were done, and the result shows: (i) Sobel Operator, Laplacian, Roberts and Scharr Operator respectively describe σ, then we found Sobel Operator is best for face detection based-LBP. Sobel Operator can compute an approximation of the gradient of the image intensity function and extract information of the images in edges and reduce other texture information. (ii) for different n, we also do extensive experiments, which show that the result is good when n is 1. Of course, we can choose different n for different data sets. At the same time, the result vertifies the SSML.

VI. CONCLUSION AND FUTURE WORK

In this paper, we present a non-linear pre-processing method based on Similarity Michaelis-Menten law (SMML). The efficiency of the method is estimated with respect to the performance of face recognition. In some other fields, the method may be provide a new direction to our train of thought. Similarity Michaelis-Menten law can be used to explain visual sensitivity in the vertebrate retina. We preprocess input images with SMML, and then employ Local Binary Pattern (LBP) for face feature extraction. And that parameters of SMML are locally estimated by the input image itself by employing *Sobel* to estimate its parameters, and show its advantage for face detection. In the experiments, the very high recognition rates are achieved for face recognition associated with the proposed preprocessing technique on FERET database. Advantages of SMML include improvement of light adaption, noise effect, detection right rate, robustness and efficiency, which inspire us exploit it for face pre-processing descriptor for the first time in the field of face recognition.

ACKNOWLEDGEMENT

The work was supported in part by the Natural Science Foundation of China, under Contracts 60903065, 61039003 and 61272052, in part by the Fundamental Research Funds for the Central Universities, and by the Program for New Century Excellent Talents University of Ministry of Education of China.

REFERENCES

[1] Ojala, T., Pietik¨ainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, vol. 29, 1996.

[2] Ojala, T., Pietik¨ainen, M., M¨aenp¨a¨a, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24 971–987, 2002.

[3] T. Ahonen, A. Hadid, and M. Pietikäinen, Face Description with Local Binary Patterns: Application to Face Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007.

[4] Beaudot W.H.A. Sensory coding in the vertebrate retina: Towards an adaptive control of visual sensitivity. Network: Computation in Neural Systems. 7(2):317-323, 1996.

[5] Jeanny Hérault, Barthélémy Durette: Modeling Visual Perception for Image Processing. In proc. of the 9th international work conference on Artificial neural networks. 662-675, 2007.

[6] Helga Kolb: How the Retina Works, American Scientist. 91, (1) , 2003.

[7] Li Tao and VijayanK.Asari: Adaptive and integrated neighborhood-dependent approach for nonlinear enhancement of color images. Journal of Electronic Imaging,14(4), 043006 , 2005.

[8] Meylan L, Alleysson D, Süsstrunk S. Model of retinal local adaptation for the tone mapping of color filter array images[J]. JOSA A, 2007, 24(9): 2807-2816.

[9] Vu N S, Caplier A. Illumination-robust face recognition using retina modeling[C]//Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, 2009: 3289-3292.