

OptiFuzz: a robust illumination invariant face recognition system and its implementation

Bima Sena Bayu Dewantara¹  · Jun Miura¹

Received: 6 May 2015 / Revised: 13 June 2016 / Accepted: 17 June 2016 / Published online: 15 July 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Vision-based human face detection and recognition are widely used and have been shown to be effective in normal illumination conditions. Under severe illumination conditions, however, it is very challenging. In this paper, we address the effect of illumination on the face detection and the face recognition problem by introducing a novel illumination invariant method, called OptiFuzz. It is an optimized fuzzy-based illumination invariant method to solve the effect of illumination for photometric-based human face recognition. The rule of the Fuzzy Inference System is optimized by using a genetic algorithm. The Fuzzy's output controls an illumination invariant model that is extended from Land's reflectance model. We test our method by using Yale B Extended and CAS-PEAL face databases to represent the offline experiments, and several videos are recorded at our campus to represent the online indoor and outdoor experiments. Viola–Jones face detector and mutual subspace method are employed to handle the online face detection and face recognition experiments. Based on the experimental results, we can show that our algorithm outperforms the existing and the state-of-the-art methods in recognizing a specific person under variable lighting conditions with a significantly improved computation time. Other than that, using illumination invariant images is also effective in improving the face detection performance.

Keywords Illumination invariant · Optimized fuzzy · Fuzzy Inference System · Illumination ratio · Illumination invariant model

1 Introduction

Vision-based human face recognition has been shown to be effective under normal illumination conditions. When it is used under severe illumination conditions, however, the recognition rate drops rapidly. We cannot always expect to have good illumination, and changes in illumination are hard to predict and control. When we move through various environments, including both indoor and outdoor, the appearance of facial features changes, thereby making it difficult to recognize faces (see Fig. 1).

In face recognition problems, the most important factor is a constancy; for each human face's part, color constancy in appearance-based approaches and feature constancy in feature-based approaches will be very helpful for distinguishing each person. A region of interest (ROI) of the face should also be constant. An ROI that is shifted from a correct face region would cause inaccurate recognition results.

Several illumination invariant methods have been proposed to solve these problems. Both appearance-based and feature-based methods are widely used. The appearance-based methods convert an input image into the desired output image and are categorized into three groups: histogram transformation, photometric-based methods and gradient-based methods. Feature-based methods transform an input image into a set of features or descriptors instead of using the intensity of each pixel directly.

Histogram transformation such as histogram equalization (HE) and block histogram equalization (BHE) [1] apply a contrast enhancement to the image which works very

Electronic supplementary material The online version of this article (doi:[10.1007/s00138-016-0790-6](https://doi.org/10.1007/s00138-016-0790-6)) contains supplementary material, which is available to authorized users.

✉ Bima Sena Bayu Dewantara
bima@aisl.cs.tut.ac.jp

Jun Miura
jun.miura@tut.jp

¹ Toyohashi University of Technology, Toyohashi, Japan



Fig. 1 Different visual appearances caused by different illumination conditions

fast. However, the equalized image frequently experiences washout and over-equalized effects in a narrow region with a very large distribution of intensity. The effects of light on a face is manifested by photometric-based methods. It uses human perception theory and illumination properties. Multi-scale retinex (MSR) [2], Fuzzy retinex (FR) [3], self-quotient image (SQI) [4] and classified appearance-based quotient image (CAQI) [5] are able to produce a fairly invariant image conversion. However, these methods are less able to control the final color constancy and sometimes fail to eliminate the boundary region between different illumination effects. These problems can be minimized by specifying some parameters either adaptively or with optimization techniques performed by adaptive scale retinex (ASR) [6] and optimized-SQI (Opt-SQI) [7], respectively. However, both of these methods require very long computation periods.

Logarithmic discrete cosine transform (LDCT) [8] discards several first DCT's coefficients which are related to shadow effects. This process runs very fast, but the output image becomes unclear. SVD-Face [9] uses the normalized coefficients of the singular value decomposition (SVD) which are insensitive to different illumination conditions for describing the underlying structures of faces. Mean estimation (ME) [10] removes the illumination component by subtracting the mean estimation from the original image. The mean estimation is locally obtained from the exponential form of the illumination–reflection model instead of using smoothing techniques. However, both of these methods are difficult to control for the final color constancy of the two areas affected by different lighting effects.

Approaches using a gradient can be found on both appearance-based and feature-based methods. GradientFace (GF) [11] and block difference of inverse probabilities and block variation of local correlation coefficient (BDIP-BVLC) [12] are examples of appearance-based gradient methods. The result of the GF is poor when used for facial images with low lighting, while the sketch-like converted image of the BDIP-BVLC often experiences strong edge artifacts at boundaries and frequently fails to maintain their final color constancy. Weber local descriptor (WLD) [13], local binary pattern histogram (LBPH) [14] and logarithm gradient histogram (LGH) [15] also use a gradient to build features. LGH

successfully combines the magnitude and the orientation of the logarithmic gradient to form a robust histogram feature. Its achievement outperforms WLD and LBPH.

We propose a novel appearance-based illumination invariant method that works very fast and can control the final color constancy while reducing the boundary effects to keep the high recognition rate. It controls the image contrast using an illumination invariant model that is supported by the Fuzzy Inference System (FIS) [16]. We optimize FIS rules using genetic algorithm (GA) [17], to find the most optimal illumination ratio which is a crucial component in the illumination invariant model. Our goal is the implementation of the method in real conditions. We consider that the appearance-based methods are easier to be used instead of using feature-based methods because of their flexibility when combined with other systems such as Viola–Jones face detector [18]. Here, we show that our illumination invariant method is effective in two face-related tasks, face detection and face recognition.

The remainder of this paper is organized as follows. Section 2 discusses the illumination normalization model and our modification for addressing the human face recognition problem. Section 3 presents the Fuzzy Inference System and its optimization. Section 4 describes the development of an online face recognition system. Section 5 presents the experimental results and discussions. Section 6 concludes our present work and possible future works.

2 Illumination invariant face recognition

The human face has distinctive components, texture and contour. When the face is exposed to light coming from various directions, its appearance from the front side will vary considerably. Most illumination invariant methods [2–10, 12] are capable of producing invariant facial appearances; however, they make the face look flat, unclear and create a boundary effect, differing final color constancy, etc.

Our method takes the advantage of the illumination normalization model proposed in [19]. This model was originally designed for normalizing a partially shadowed area of landscapes using a pair of appearance-based segmented

regions (shadowed and shadow-free regions). The model considers not only environment-based lighting effects but also a direct lighting effect. The model is derived from those two lighting effects and forms a kind of contrast ratio of an image.

2.1 Illumination normalization model

In general, the reflectance model [20] can be expressed as:

$$I_{x,y} = L_{x,y}R_{x,y}, \tag{1}$$

where $I_{x,y}$ is an image intensity at pixel (x, y) that is captured by the camera, $L_{x,y}$ is the illumination that comes to the surface of object and $R_{x,y}$ is the reflectance property of the surface.

Referring to the model by Guo et al. [19], an extended version of the reflectance model can be expressed as:

$$I_{x,y} = \left\{ d_{x,y}L_{x,y}^d \cos(\theta_{x,y}) + L_{x,y}^e \right\} R_{x,y}, \tag{2}$$

where $L_{x,y}^d$ and $L_{x,y}^e$ are representing the amount of the direct light and the environment light, respectively. $R_{x,y}$ is the surface reflectance. $\theta_{x,y}$ is an angle between the direct lighting direction and the surface normal. $d_{x,y}$ is a value between 0 and 1 indicating how much direct light gets to the surface (albedo). To simplify (2), they define $k_{x,y} = d_{x,y} \cos(\theta_{x,y})$ and get a simplified model :

$$I_{x,y} = \left\{ k_{x,y}L_{x,y}^d + L_{x,y}^e \right\} R_{x,y}. \tag{3}$$

where $k_{x,y}$ is henceforth referred to as the shadow coefficient. Since [19] only focused on shadow and shadow-free pixels, they used $0 \leq k_{x,y} \leq 1$. From (3), a shadow-free pixel can be represented by setting $k_{x,y} = 1$ (i.e., direct lighting):

$$I_{x,y} = \left\{ L_{x,y}^d + L_{x,y}^e \right\} R_{x,y}, \tag{4}$$

and a shadowed pixel when $k_{x,y} < 1$, as follows

$$I_{x,y}^* = \left\{ k_{x,y}L_{x,y}^d + L_{x,y}^e \right\} R_{x,y}. \tag{5}$$

Based on (4) and (5), a new relationship between shadowed and shadow-free pixels can be obtained as:

$$\begin{aligned} I_{x,y} &= \left\{ L_{x,y}^d + L_{x,y}^e \right\} R_{x,y} \frac{\left\{ k_{x,y}L_{x,y}^d + L_{x,y}^e \right\}}{\left\{ k_{x,y}L_{x,y}^d + L_{x,y}^e \right\}} \\ &= \frac{\left\{ L_{x,y}^d + L_{x,y}^e \right\}}{\left\{ k_{x,y}L_{x,y}^d + L_{x,y}^e \right\}} I_{x,y}^* \end{aligned} \tag{6}$$

We can see that a shadow-free pixel, $I_{x,y}$, can be generated from a shadowed pixel, $I_{x,y}^*$, by a specific ratio. If we use $r_{x,y} = L_{x,y}^d / L_{x,y}^e$ to simplify (6), then we get:

$$I_{x,y} = \frac{\{r_{x,y} + 1\}}{\{k_{x,y}r_{x,y} + 1\}} I_{x,y}^* \tag{7}$$

where $r_{x,y}$ is called as an illumination ratio of pixel (x, y) .

2.2 Illumination invariant model for human face

The model in [19] can also be used in our case since the illumination problem has similar effects, i.e., shadow and reflection. However, estimating $k_{x,y}$ by a physics-based approach is very hard, because measuring the actual values of $d_{x,y}$ and $\theta_{x,y}$ are very difficult. Instead, we utilize information provided by the image. Firstly we attempt to classify the appearance of face regions into normal and abnormal conditions. Normal condition ($k_{x,y} = 1$) is an actual appearance of a human face in normal lighting. Abnormal conditions ($k_{x,y} \neq 1$) correspond to an appearance which is largely affected by the condition of direct lighting, i.e., shadow ($k_{x,y} < 1$) and reflection ($k_{x,y} > 1$).

Equation (7) is valid only for certain conditions such as the presence of shadow ($k_{x,y} < 1$) and shadow-free ($k_{x,y} = 1$); the presence of reflection effect ($k_{x,y} > 1$) is not taken into account. Although the shadow effect is a more crucial problem than the reflection effect, these effects also potentially produce misrecognition. Therefore, we consider all potential effects by dividing the range of human face intensities and assigning their properties as shown in Fig. 2.

Considering that the most of human face is composed of skin, we segment human skin intensity based on [21–23] where the human skin intensities are located between 30 and 200 (when maximum intensity is 255). To obtain the conversion of image that is invariant and to keep the final color constancy, we use the maximum intensity of human skin and make a rough approximation as follows:

$$\hat{k}_{x,y} \approx \frac{I_{x,y}^*}{200}, \tag{8}$$

where $\hat{k}_{x,y}$ is an approximated shadow coefficient at position (x, y) that shows a better normalized intensity of the human face. We then make a minor modification to Eq. (7), that is,

$$I_{x,y} = \frac{\{r_{x,y} + 1\}}{\{\hat{k}_{x,y}r_{x,y} + 1\}} I_{x,y}^* \tag{9}$$

where we use a new shadow coefficient $\hat{k}_{x,y}$ instead of $k_{x,y}$.

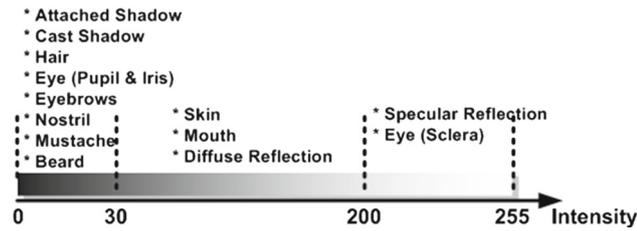


Fig. 2 Human face color/intensity segmentation in a normal illumination

3 Fuzzy Inference System and its optimization

We now have an illumination invariant model for human faces. The method of obtaining an illumination ratio, $r_{x,y}$, used in [19] cannot be used in our case because their approach is used exclusively for a partially shadowed area. In our case, the lighting may cause a full shadow effect on the face. It is difficult to always get a pair of shadowed and shadow-free regions from a given image.

Usually, we can estimate and reconstruct an appearance by comparing the current appearance with the intrinsic appearance which is commonly known in advance. For example, in the case of the pupil (see Fig. 3a), the intensity of a pupil (marked by yellow cross) is usually very low (VL) in the image regardless of shadow effects. We can thus develop reasoning as follows: If the intrinsic appearance is very low

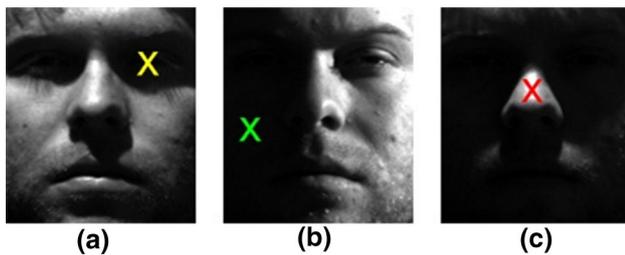


Fig. 3 Differences of facial appearance due to direct lighting: **a** the pupil's appearance which is always *dark* even it is covered by shadow or not, **b** the cheek's appearance which is *darker* because of shadow effect and **c** the nose's appearance which is *brighter* because of specular effect

Fig. 4 Illustration of block histogram equalization

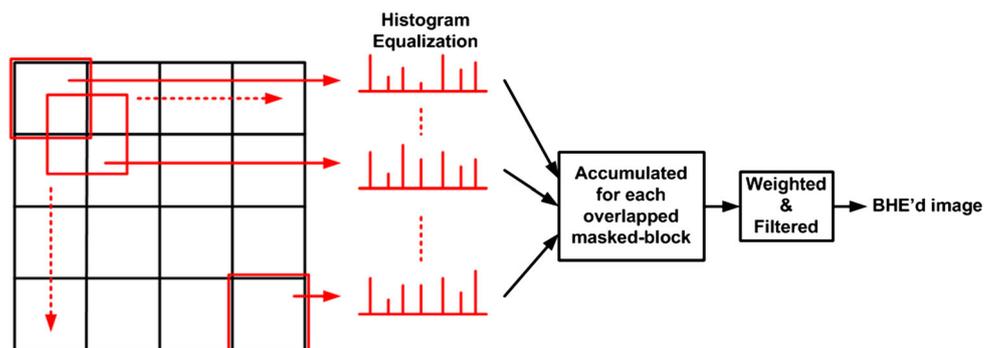


Fig. 5 Block histogram equalization; original image (left), BHE'd image (right)

(VL) and the current appearance is very low (VL), then $r_{x,y}$ should be very low (VL). Another example, in the case of the cheek (see Fig. 3b), its intensity (marked by green cross) is usually medium (M, in the range of normal skin intensity in Fig. 2). We thus make the following reasoning; if the intrinsic appearance is medium (M) and the current appearance is very low (VL), $r_{x,y}$ should be high (H).

To get the original intrinsic appearance is, however, difficult. Therefore, we apply a local contrast adjustment to the input image and make a rough approximation of the intrinsic appearance image. We use the block histogram equalization (BHE) [1] to see if a pixel is inherently dark or is inherently bright but appears dark due to a shadow effect. An input image is divided into blocks. A masking block with size 12x12 pixels is applied to the image and histogram equalization is performed. The block is moved every 6 pixels to overlap with an adjacent block to avoid a blocking effect on the adjacent blocks. Then, we apply a weighted sum of neighboring adjacent blocks to smooth the boundaries. Figure 4 shows the illustration of BHE process. The BHE'd result (I^{int}) is shown in Fig. 5. At the same time, we use the input image as the current appearance image (I^{cur}).

3.1 Fuzzy Inference System

For now, we have two pieces of information from the intrinsic appearance image, I^{int} , and the current appearance image, I^{cur} . We can make a simple reasoning from both images. We realize our idea by using a Fuzzy Inference System (FIS) [16]

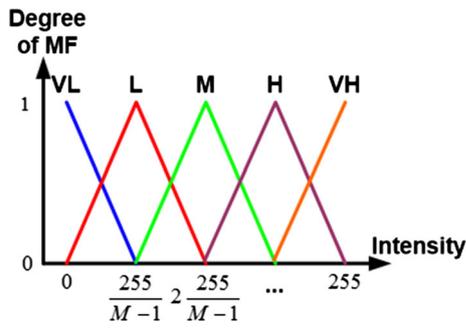


Fig. 6 Fuzzy membership functions using triangular shape

as proposed in [24] because FIS has the capability to combine two types of information based on a determined rule. We argue that the human face has a similar shape and contour; the same lighting direction may produce the same effect for every person. Therefore, it is reasonable to use a simple rule to enhance the affected face region. FIS can be decomposed into three main parts: fuzzification and membership function, the inference engine and rule base, and defuzzification.

3.1.1 Fuzzification and membership function

The fuzzification converts a crisp input into a linguistic variable using the membership function. Linguistic variables are labels or concepts corresponding to partitions of a state space, such as low, medium and high. During reasoning, the variables are referred to by the linguistic terms, and the Fuzzy membership function determines the correspondence between linguistic variables with the numerical values. In the case of using five members, for example, we set five linguistic variables, very low (VL), low (L), medium (M), high (H) and very high (VH), for both intrinsic and current appearances, as shown in Fig. 6.

We set our Fuzzy membership using a triangular function. The reason for using the triangular function is it outperformed other functions (please refer to Sect. 5.1 for more detail). Based on the triangular shape, we can derive a degree of the membership function, that is,

$$\mu = \begin{cases} \frac{x-a}{b-a} & \text{if } a < x \leq b \\ \frac{c-x}{c-b} & \text{if } b < x \leq c \\ 0 & \text{if others} \end{cases} \quad (10)$$

where μ is a degree of membership function, x is the crisp input, a is the left boundary of membership function, b is the center of membership function and c is the right boundary of membership function. The parameters of the membership function of each input can be expressed as

$$b_m = (m - 1) \frac{255}{M - 1}, a_m = b_{m-1}, c_m = b_{m+1} \quad (11)$$

where $m = 1, 2, \dots, M$ is the index of membership. M is the number of Fuzzy membership functions used in our system. M can be 3, 5, 7 or 9. Figure 6 shows the shape of the membership function.

3.1.2 Inference engine and rule base

We have two kinds of memberships, $\mu_{I_m^{cur}}$ and $\mu_{I_m^{int}}$ to represent the degree of the membership function of the current appearance and the intrinsic appearance images, respectively. We used the product of $\mu_{I_m^{cur}}$ and $\mu_{I_m^{int}}$ instead of the min operator to form the rule as follows

$$\mathbf{B} = [\mu_{I_m^{cur}} \mu_{I_m^{int}}]_{m=1,2,\dots,M}, \quad (12)$$

where \mathbf{B} is the T-norms between $\mu_{I_m^{cur}}$ and $\mu_{I_m^{int}}$. There are M memberships for each of the current and the intrinsic appearances, and therefore, we have $F = M^2$ dimensions in total, for which a Fuzzy rule \mathbf{R} needs to be defined.

3.1.3 Defuzzification

Finally, defuzzification is performed by converting the Fuzzy output of the inference engine to crisp by adopting center of area (CoA), that is,

$$r_{x,y} = \frac{\sum_{f=1}^F B_f R_f}{\sum_{f=1}^F B_f}, \quad (13)$$

where $r_{x,y}$ is the illumination ratio and $f = 1, 2, \dots, F$ is the number of inference \mathbf{B} as well as the number of \mathbf{R} .

3.2 Optimizing rules of Fuzzy Inference System

The most important part of the Fuzzy Inference System is a rule base. Manual selection of the rule is very difficult due to the vast variation of inputs, consideration of the image size and range of intensities. Optimizing the FIS rule base is, therefore, crucial. We use the genetic algorithm (GA) to produce an optimal rule that satisfies all illumination conditions. The Fuzzy optimization using GA is shown in Algorithm 1.

We randomize H -sets of the F -dimensional FIS rule, \mathbf{R}_h ($h = 1, 2, \dots, H$) to initialize the population. Each vector \mathbf{R}_h represents a set of the gene to represent a solution to the optimization problem being addressed and called as an individual. A fitness function associated with each individual is used to evaluate the appropriateness of the solution. We use an averaged normalized cross-correlation-based distance to compute the fitness value, $f(\mathbf{R}_h)$, as follows

$$f(\mathbf{R}_h) = \frac{1}{N - 1} \sum_{n=2}^N \sum_{x,y} \frac{(I_{x,y}^1 - \mu_{I^1}) (I_{x,y}^n - \mu_{I^n})}{\sigma_{I^1} \sigma_{I^n}}, \quad (14)$$

Algorithm 1: Genetic algorithm for optimizing Fuzzy’s rule

Input: A set of training image, $I_n, n = \{1, 2, \dots, N\}$
 Number of iteration, $Iter$
 Number of population, H
 Number of Fuzzy’s rule element, F

Output: Optimized Fuzzy’s rule

Initialization: Randomize H sets of Fuzzy’s rule vector,
 $\mathbf{R}_h = \{R_1, R_2, \dots, R_F\}, h = \{1, 2, \dots, H\}$

Begin

1. for $i := 1$ to $Iter$ step 1 do
2. for $h := 1$ to H step 1 do
3. for $n := 1$ to N step 1 do
4. Convert I_n using Eq. (9) for \mathbf{R}_h
5. end for
6. Choose the conversion result of I_1 as a target, and calculate the distances between I_1 and $I_{2:N}$ using Eq. (14) to obtain the fitness value of $\mathbf{R}_h, f(\mathbf{R}_h)$
7. end for
8. Evaluation: sorting the fitness value, $f(\mathbf{R}_h)$ from the highest to the lowest
9. Reproduction: $\mathbf{R}_1 \leftarrow \mathbf{R}$ with the highest fitness value from the sorting
 $\mathbf{R}_2 \leftarrow \mathbf{R}$ which is randomly chosen from the sorting
 \mathbf{R}_1 and \mathbf{R}_2 are set as a new pair of parent
10. Crossover between \mathbf{R}_1 and \mathbf{R}_2
11. Mutation between \mathbf{R}_1 and \mathbf{R}_2
12. New composition of population \mathbf{R}_h is formed. Two new individuals as the result of crossover and mutation operations replace the worst two individuals in the current population.
13. end for
14. The most optimized \mathbf{R} is obtained after $Iter$ step

where $h = 1, 2, \dots, H$ is the index of individual, $n = 2, 3, \dots, N$ is the index of training image excluding the first image as the normal reference; (x,y) is a pixel position; $I_{x,y}^1$ and $I_{x,y}^n$ are the first and the n th normalized image, respectively. μ_{I^1} and μ_{I^n} are the mean of the first and the n th normalized image, respectively. σ_{I^1} and σ_{I^n} is the standard

deviation of the first and the n th normalized image, respectively.

The aim of applying the genetic operators is to transform an individual into a new individual with a higher fitness value. The reproduction operator performs a natural selection function to choose the individuals with the highest fitness values that have a greater probability to be selected as a pair of parents to produce better offspring. The crossover operator chooses a pair of genes by randomly selecting a point and exchanging their tails. The mutation operator randomly mutates the values of the same gene position to optimize each individual. The optimization process is iterated until required conditions are met (e.g., determined number of generation or fitness value threshold).

Figure 7 shows the block diagram of our proposed method. The complete diagram shows the optimization process, where an input image is converted to an invariant image using our illumination invariant model. The GA-based optimized rule ensures that the Fuzzy Inference System produces an appropriate illumination ratio for a better invariant image. When the optimization process has finished, the genetic algorithm (red dashed) is released. Our system is then ready to directly enhance the input image.

4 Online face recognition system

To deal with an online system, we designed our system as shown in Fig. 8. First, we convert the input image using our illumination invariant method. The converted image is then fed into the face detector process to search for and detect human faces in the image. The detected face is then aligned using the face alignment process to get a better face position and orientation. Finally, the aligned face is recognized using the mutual subspace method to determine the person’s identity.

Fig. 7 Our proposed system for the illumination invariant face recognition

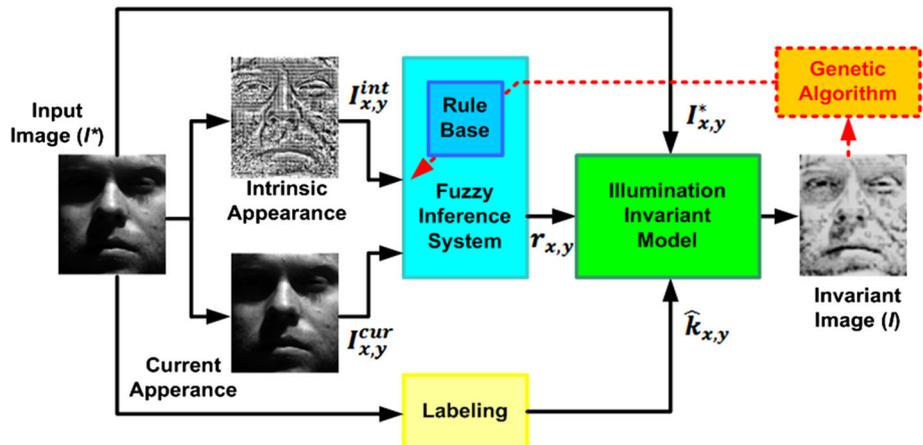


Fig. 8 Our robust illumination invariant face recognition system

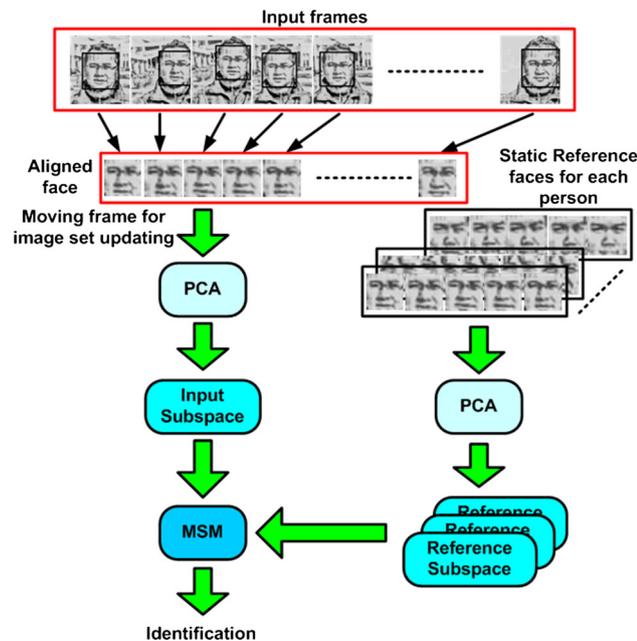
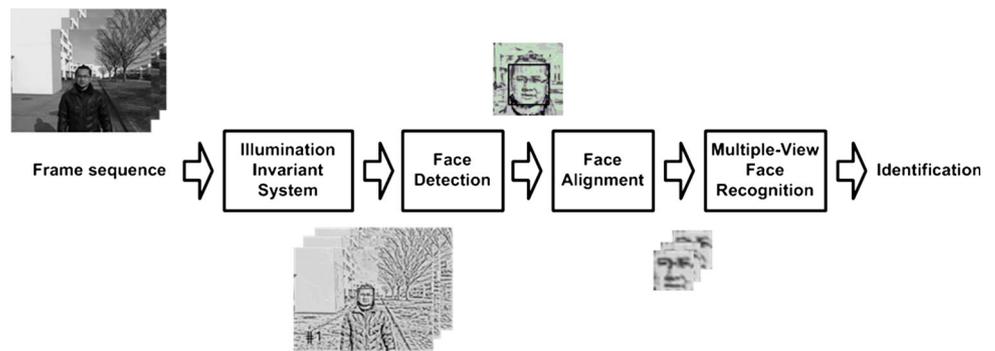


Fig. 9 Flow of face recognition based on MSM

4.1 Face detection

The Viola–Jones face detector [18] is applied to the illumination invariant images in order to detect human faces online in real scenes. We trained the detector using 4000 positive images and 7000 negative images; where all images were converted to invariant images using the illumination invariant method before training.

4.2 Face alignment

A detected face from the face detector needs to be further processed. Face recognition is feeble against rotation and scaling in the image. Aligning the cropped face is, therefore, very important to minimize the misrecognition. We adopt a multiple-scale and multiple-angle template matching-based face feature detection [26] to detect eye position, measure the rotation angle and correct the scale.

We first capture several templates of eyes. The angle of rotation of each template is limited up to $\pm 45^\circ$. Each template is then enhanced using the illumination invariant method. From the detected eye positions, an angle of rotation can be calculated. The face image is then rotated using the calculated angle. Unnecessary parts such as hair, neck and background are then discarded. After the alignment process has completed, the aligned image is resized to 32×32 pixels.

4.3 Face recognition

Pattern matching using a single view of the face frequently causes misrecognition. We employ the mutual subspace method (MSM) [27], a powerful matching technique that tolerates variations in face patterns. It utilizes multiple canonical angles between the input subspace and the reference subspace. The MSM recognizes a temporal aligned face image sequence as shown in Fig. 9.

5 Experimental results

5.1 Selecting the best function for Fuzzy membership

There are two functions that are commonly used in Fuzzy systems, triangular function and Gaussian/Bell function. To ascertain which function is the best, we compared the performance of our method using both functions. We use a set of images of one subject in the Yale B Extended face database [25] for optimization. The database is fairly representative because it provides a sufficient variety of illumination conditions. We have tested several sizes of image, namely 168×192 pixels, 84×96 pixels, 42×48 pixels and 21×24 pixels. We chose a resized image with size of 42×48 pixels because it is the smallest size that can maintain recognition accuracy. Similar appearance images are the objective of the optimization.

We performed the GA training for a 5×5 Fuzzy rule. The reason of using the 5×5 Fuzzy rule is explained in Sect. 5.2. We iterate and record the fitness value up to 10,000 itera-

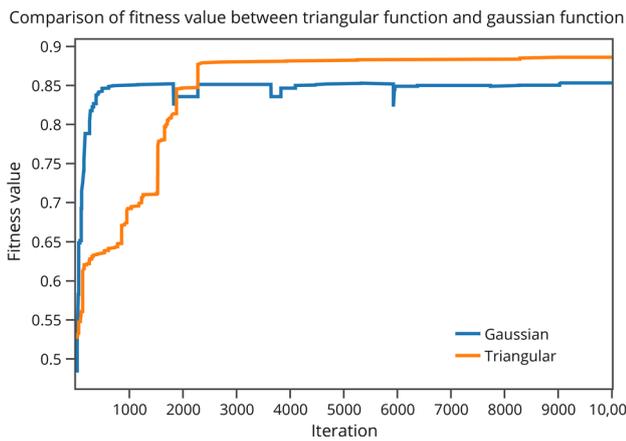


Fig. 10 Comparison of fitness value evolution between the triangular function and the gaussian function

tions to guarantee its convergence as shown in Fig. 10. Even though the gaussian function can reach higher fitness value faster than the triangular function, but its fitness stops around 0.85, while the triangular function can achieve higher value (around 0.88). These achievements affect to the recognition results (using 38 subjects of Yale B Extended face database) as shown in Table 1. From this reason, we decided to use the triangular function.

5.2 Fuzzy rule optimization

After getting the best function for our Fuzzy membership, we perform the optimization of the Fuzzy rule. We utilize the same set and size of images, iteration number and objective as used in Sect. 5.1 for training using different sizes of the Fuzzy rule as shown in Table 2. The evolution of fitness value of each size of Fuzzy rule is shown in Fig. 11. Based on the training outcomes and referring to the test results using the 10 subjects from Yale B Extended face database, we used the 5 × 5 rule because it has a rapid transition without experiencing the local optimum problem. It also achieves a high fitness value and produces the best invariant image as shown in Fig. 12. The best achievements using NN and eigenfaces (PCA) show the stability of the 5 × 5 rule. The 5 × 5 rule is shown in Fig. 13.

Figure 13 shows a mapping of inputs (current appearance and intrinsic appearance) and an output of the Fuzzy Inference System-based reasoning. For example, if the intensity of a particular pixel has current appearance = VL and intrinsic appearance = VL, then output = 10. It means that our

Table 2 Comparison results of different size of the Fuzzy rule

Rule	Max fitness	Evaluation method		Time (ms)
		NN (%)	PCA (%)	
3 × 3	0.7956	97.66	84.38	8.02
5 × 5	0.8863	99.84	98.75	8.41
7 × 7	0.8912	93.28	74.69	10.04
9 × 9	0.8933	99.69	89.06	13.56

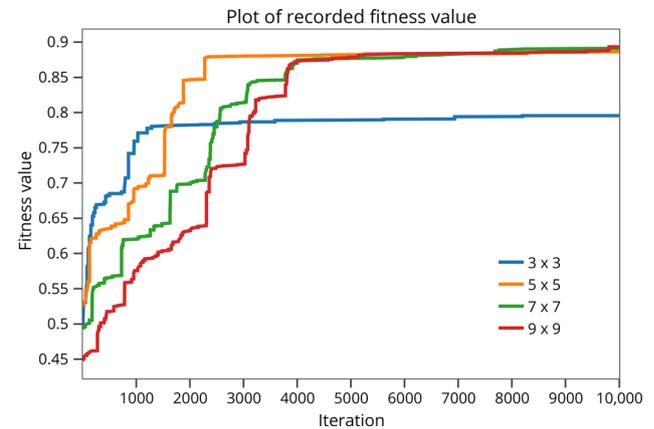


Fig. 11 Evolution of GA’s fitness value for different size of the Fuzzy rule for 10,000 iterations

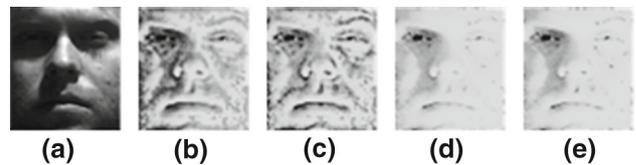


Fig. 12 Comparison results of different size of the FIS rule; **a** the original image, **b** the converted image using a 3 × 3 rule, **c** the converted image using a 5 × 5 rule, **d** the converted image using a 7 × 7 rule, and **e** the converted image using a 9 × 9 rule

illumination invariant model will receive a small illumination ratio, which gives only a small adjustment to the intensity of that pixel.

5.3 Experiment using face database

Our experiments use two public face databases to measure and compare the quality of all methods using our best-defined parameters as shown in Table 3. Yale B Extended [25] and CAS-PEAL [28] face databases are the representative face databases that deal with the illumination problem. We use

Table 1 Comparison of face recognition performance using the triangular function and the gaussian function

Function	Nearest neighbor (%)	Principal component analysis (%)	Time (ms)
Triangular	94.24	92.85	8.41
Gaussian	91.49	85.61	78.99

5 x 5		Current Appearance				
		VL	L	M	H	VH
Intrinsic Appearance	VL	10	1	0	0	20
	L	2	0	0	0	2
	M	57	7	2	0	14
	H	142	137	97	57	0
	VH	145	118	148	145	31

Fig. 13 A 5 × 5 Fuzzy rule generated by GA

nearest neighbor (NN), eigenfaces (PCA) [29] and support vector machines (SVMs) [30] for evaluation. The classifier settings are summarized in Table 4.

5.3.1 Experimental results using Yale B Extended face database

The Yale B Extended face database consists of 38 subjects with 64 illumination variations. Each image is already cropped and aligned at size of 168 × 192 pixels. Thirty-eight frontally illuminated face images are chosen as a training set, while the remaining are used as a testing set.

Table 3 Parameters setting of proposed method and other methods in comparison

Method	Parameters	Value	
None	–	–	
HE	–	–	
BHE [1]	Block size	12	
MSR [2]	Filter size (Gaussian)	3, 5, 7	
FR [3]	Filter size (Gaussian)	3	
ASR [6]	Fuzzy membership function	Triangular, 5 × 5	
	Smoothing iteration, T	50	
	SQI [4]	Filter size (Gaussian)	3, 5, 7
CAQI [5]	Filter size (Gaussian)	3, 5, 7	
	α	0.05	
	Opt-SQI [7]	Filter size (Gaussian)	3, 5, 7, 9, 11, 13, 15, 17
		β_k	1, 1/15, 2/15, 7/15, 1, 1, 1, 3/15
m_k		4/15, 1/15, 1/3, 0, 4/15, 14/15, 14/15, 3/5	
DCT [8]	α	0.98	
	Ddis (discarded DCT coefficient)	15	
BDIP-BVLC [12]	Block size	3	
GF [11]	–	–	
OptiFuzz (Ours)	Block size (Appearance model)	12	
	Fuzzy membership function	Triangular, 5 × 5	
	Fuzzy rule	See Fig. 13	

Table 4 Parameters setting of the classifier methods

Classifier	Parameters	Method/value
NN	Classification type	Pearson product moment correlation coefficient
PCA [29]	Number of principal components	38
SVMs [30]	SVM type	nu-SVC multi-class classification
	Kernel type	Linear

We reimplement the existing appearance-based illumination invariant methods as listed in Table 3. The illumination invariant results of each method are shown in Fig. 14. We can see that our method and ASR are able to maintain the final color constancy and minimize the boundary effect. The quantitative results are shown in Table 5. We also include the results of the other new methods such as the SVD-Face and the mean estimation. The (*) mark indicates that the result are directly taken from the paper.

The existing appearance-based illumination invariant methods such as HE, MSR [2], FR [3], CAQI [5] and GF [11] are not robust enough when tested using very harsh illumination conditions. They fail to control the final color constancy of the converted images. The remaining methods such as BHE [1], ASR [6], SQI [4], Opt-SQI [7], DCT [8] and BDIP-BVLC [12] can maintain the final color constancy of

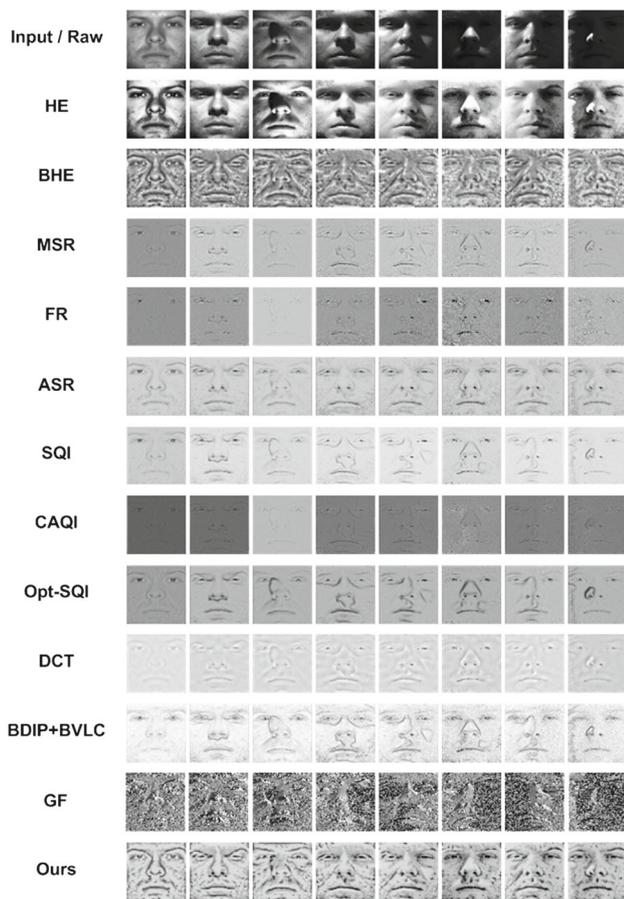


Fig. 14 Comparison of illumination normalization methods for Yale B Extended face database

most of the converted images; however, some of them fail to reduce boundary effects. Considerable differences in the performance of each classifier show that these methods do not have consistent stability.

SVD-Face [9] and ME [10] seem to perform well regarding the performance. However, we do not know the consistency of their stability since we do not have data for the other classifiers. However, one can note that our proposed method has the best performance for face recognition and the best stability in the experiments. It successfully maintains the final color constancy, reduce the boundary effects and provide useful facial textures that are very helpful to maintain a high recognition rate.

We also compare the performance of our proposed method with the feature-based methods as shown in Table 6. Local binary pattern histogram (LBPH) [14], Weber local descriptor (WLD) [13] and logarithm gradient histogram (LGH) [15] are the examples of feature-based illumination invariant method. The (*) mark indicates that the result are directly taken from the paper. Our method also still shows better results for all the classifiers.

Table 5 Recognition rate of appearance-based illumination invariant method for Yale B Extended face database (%)

No	Methods	Recognition rate (%)		
		NN	PCA	SVMs
1	Raw	41.98	40.46	32.81
2	HE	46.83	53.95	47.12
3	BHE [1]	90.25	86.76	87.34
4	MSR [2]	83.80	66.82	12.21
5	FR [3]	65.17	59.42	12.91
6	ASR [6]	92.97	90.38	81.58
7	SQI [4]	87.54	86.72	88.90
8	CAQI [5]	53.50	60.57	7.77
9	Opt-SQI [7]	91.90	89.93	92.02
10	DCT [8]	91.94	86.76	30.88
11	BDIP + BVLC [12]	93.96	85.07	82.81
12	GF [11]	56.78	54.15	52.38
13	SVD-Face [9]*	93.50	N/A	N/A
14	ME [10]*	N/A	92.00	N/A
15	OptiFuzz(Ours)	94.24	92.85	94.78

The significant results are in bold

Table 6 Recognition rate of feature-based illumination invariant method for Yale B Extended face database (%)

No	Methods	Recognition rate (%)		
		NN	PCA	SVMs
1	LBPH [14]	48.91	49.06	48.44
2	WLD [13]	50.63	60.16	31.41
3	LGH [15]*	93.30	N/A	N/A

5.3.2 Experimental results using CAS-PEAL face database

The CAS-PEAL face database is a huge Chinese face database that consists of pose, expression, accessories and lighting datasets in one package. We use only the lighting dataset which consists of 1042 subjects with 20 illumination variations. We chose the first 20 subjects and captured the whole area of face using the eye position ground truth information that is included in the distributed dataset. We rotate, align, crop and resize each detected face at 168×192 pixels.

We used four images (one image of frontally illuminated face, one image under fluorescent lighting, one image under incandescent lighting, and one image is randomly chosen under an elevation angle of $\pm 45^\circ$) of each subject as the training set. The remaining images are used as the testing set.

We use the same reimplementation as we have done for Yale B Extended face database when evaluating the CAS-PEAL face database. The illumination invariant results are shown in Fig. 15. The quantitative results are shown in

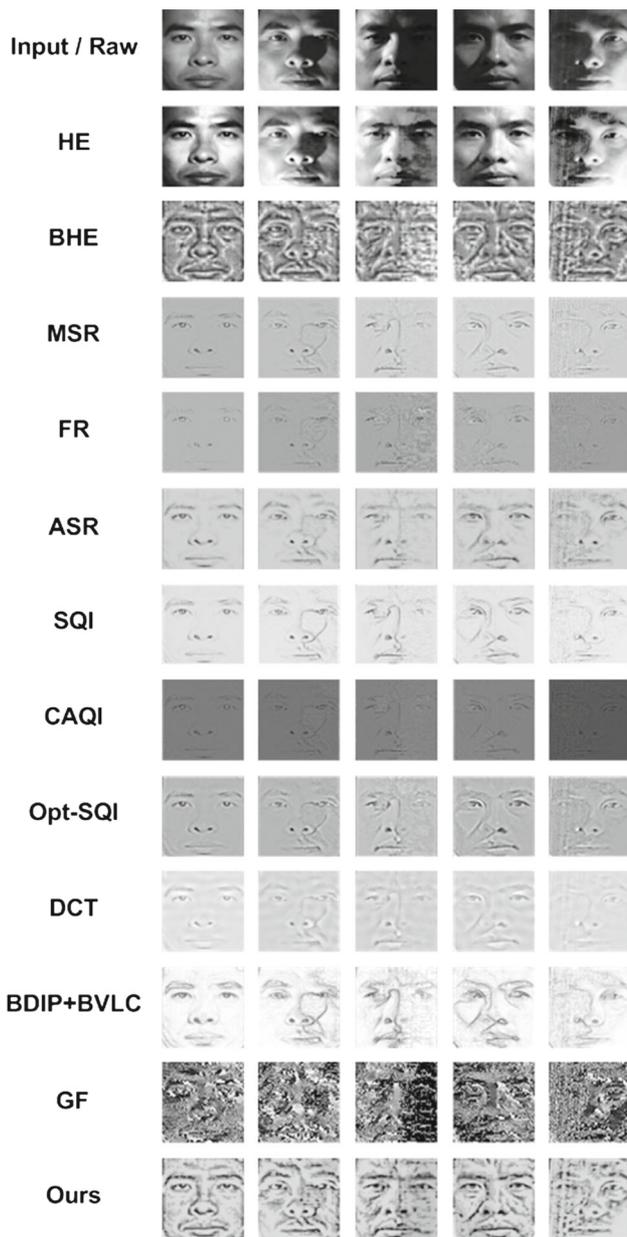


Fig. 15 Comparison of illumination normalization methods for CAS-PEAL face database

Table 7. We also include the result of the Mean Estimation method. The (*) mark indicates that the result is directly taken from the paper and the (**) mark indicates that each method is evaluated using only one frontally illuminated face image of each subject as the training image and the remaining used for testing. Our proposed method successfully outperforms the other existing and the state-of-the-art appearance-based illumination invariant methods.

We also compare the performance of our proposed method with the feature-based methods as shown in Table 8. Local binary pattern histogram (LBPH) [14] and Weber local descriptor (WLD) [13] represent the feature-based illumina-

Table 7 Recognition rate of appearance-based illumination invariant method for CAS-PEAL face database (%)

No	Methods	Recognition rate (%)		
		NN	PCA	SVMs
1	Raw	30.00	44.38	50.63
2	HE	33.13	53.44	54.06
3	BHE [1]	79.06	76.25	75.63
4	MSR [2]	60.00	52.19	60.31
5	FR [3]	38.44	28.44	34.38
6	ASR [6]	71.25	65.94	76.88
7	SQI [4]	70.94	71.25	77.81
8	CAQI [5]	27.50	30.94	39.38
9	Opt-SQI [7]	73.44	75.63	72.50
10	DCT [8]	64.69	56.25	60.00
11	BDIP + BVLC [12]	70.63	62.81	74.38
12	GF [11]	53.75	46.25	47.81
13	ME [10]*	N/A	23.21**	N/A
14	OptiFuzz(Ours)	79.06	79.06	80.94
15	OptiFuzz(Ours)	N/A	51.25**	N/A

The significant results are in bold

Table 8 Recognition rate of feature-based illumination invariant method for CAS-PEAL face database (%)

No	Methods	Recognition rate (%)		
		NN	PCA	SVMs
1	LBPH [14]	38.13	37.81	40.31
2	WLD [13]	33.13	34.69	28.13

tion invariant method. Our method still shows better results for all classifiers.

5.3.3 Computation time

We evaluate the computation time of all illumination invariant methods for still image databases. It is obtained by averaging the computation time of 2394 images in Yale B Extended face database and 320 images in CAS-PEAL face database. Evaluation is conducted using Microsoft Visual C++ running on a personal computer equipped with 2.66 GHz Intel processor supported by 2 GB of RAM. The result is summarized in Table 9.

5.4 Experiment for face detection

In most online face recognition problems, the first step is finding the face itself. Finding the face in a normally illuminated image is not as difficult as finding it in a harshly illuminated image. Using of illumination invariant images improves the detection in harshly illuminated images. We conducted three

Table 9 Computation time of each method

No	Category	Methods	Time (ms)
1	Raw	Raw	0.08
2	Photometric	HE	0.16
3		BHE [1]	2.64
4		MSR [2]	7.73
5		FR [3]	4.80
6		ASR [6]	69.21
7		SQI [4]	34.16
8		CAQI [5]	33.88
9		Opt-SQI [7]	214.81
10		DCT [8]	5.94
11	ME [10]	9.00	
12	SVD-Face [9]	50.00	
13		OptiFuzz(Ours)	8.41
14	Gradient	BDIP + BVLC [12]	38.84
15		GF [11]	1.52
16	Features	LBPH [14]	3.12
17		WLD [13]	11.63
18		LGH [15]	N/A

The significant results are in bold

different experiments to prove the effectiveness of our illumination invariant images.

We apply Haar-based detector (Viola–Jones face detector) [18] for detecting face in both original and converted images. For detecting original images, the Haar-based detector was trained using Viola–Jones original database. To be used in our system, we trained the Haar-based detector using 4000 positive images and 7000 negative images; all images were converted to invariant images using the illumination invariant method before training. In the first experiment, we used the Yale B Extended face dataset (5 persons, 65 images each). In the second experiment, we used Real1 dataset (day, indoor, 3 persons, 80 images in sequential each). In the third experiment, we used Real2 dataset (night, outdoor, 3 persons in one frame, 250 images of sequential video capture). We set the detection setting to output only one biggest face region among the ones found in the image. The results of each experiment are shown in Fig. 16 and Table 10.

True positive (TP) detection means that the detector can precisely detect face areas, while false positive (FP) detection means the detector fails to precisely detect face areas, i.e., shifted, detects other object and the detected face area is too small or too large. We validated the results by manually



Fig. 16 Comparison of the performance of face detection without and with the illumination invariant method. Upper images show the input images. Middle images show the face detection results without illumination invariant method. Lower images show the face detection results

with illumination invariant method. The first and the second column, the third and the fourth column and the fifth and the sixth column show the results using “Yale B Extended” face database, “Real1” video dataset and “Real2” video dataset, respectively

Table 10 Comparison of the performance of face detection without and with illumination invariant method

Dataset	No of person	No of image	Original		Applying illumination invariant					
			TP + FP	% (TP + FP)	TP	% (TP)	TP + FP	% (TP + FP)	TP	% (TP)
Yale B Ext.	5	325	273	84.0	228	70.0	324	99.7	322	99.0
Real1	3	240	228	95.0	176	73.0	240	100.0	197	82.0
Real2	3	250	216	86.4	181	72.0	230	92.0	205	82.0

The significant results are in bold

checking each image produced in each experiment. From the experimental results, we show that using illumination invariant images is effective in increasing the number of correct

detected faces which are very helpful for supporting face recognition.

5.5 Experiment for online face recognition

5.5.1 Experimental setup

We have examined each illumination invariant method performances to the offline face recognition problems. In this experiment, we evaluate the performance of the best six of the appearance-based illumination invariant methods which were successfully reimplemented by us, i.e., BHE [1], DCT [8], BDIP-BVLC [12], ASR [6], Opt-SQI [7] and ours.

We captured 75 illumination invariant face images of three people to create the reference set in this experiment. Twenty-five images were associated with each person. Each image was resized to 32×32 pixels. We then recorded three 640×480 pixels videos with each person for testing. Each person

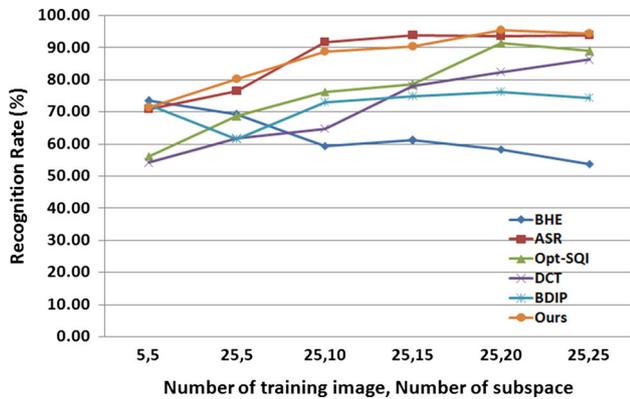


Fig. 17 Recognition rate correspond to the number of training image and number of subspace

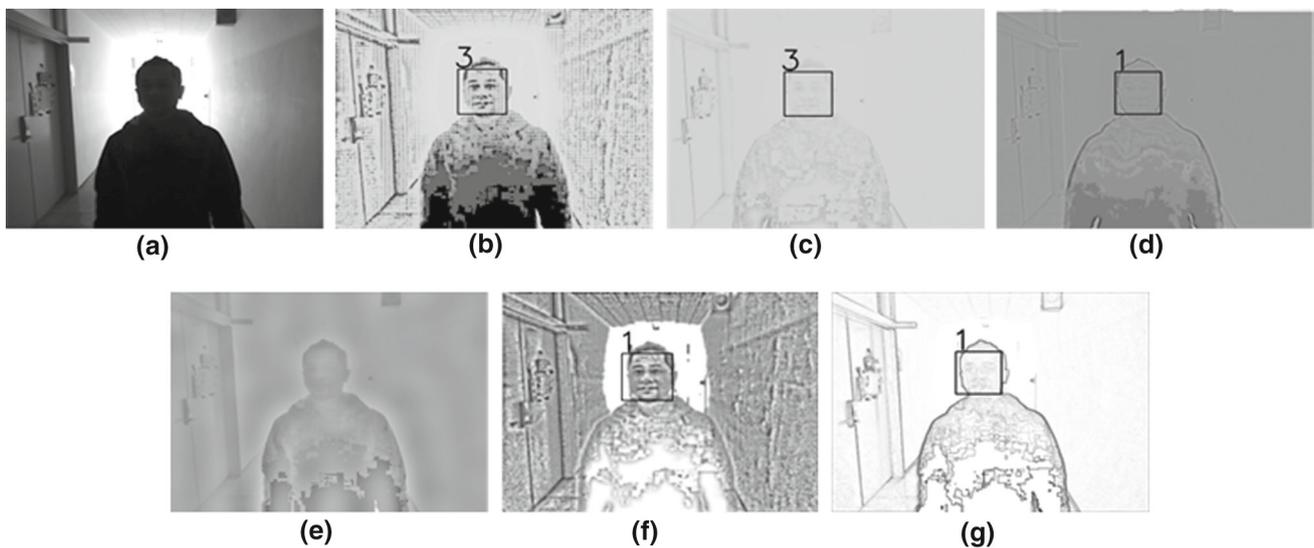


Fig. 18 A comparison of online illumination invariant methods: **a** the original input image of person #3, **b** ours, **c** ASR, **d** Opt-SQI, **e** DCT, **f** BHE and **g** BDIP-BVLC. A number on the top left of the bounding box shows the recognition identity. **b** is the converted image which look

quite clear and can be recognized correctly. The face region in **c**, **d**, **e** and **g** seem to be vague where the face properties cannot be seen clearly. Viola–Jones face detector fails in **e**, while MSM fails in **d**, **f** and **g**



Fig. 19 Experimental results for online system using our illumination invariant method; **a** the original input frames and **b** the converted frames and its recognition results that are informed on the top left of the bounding box

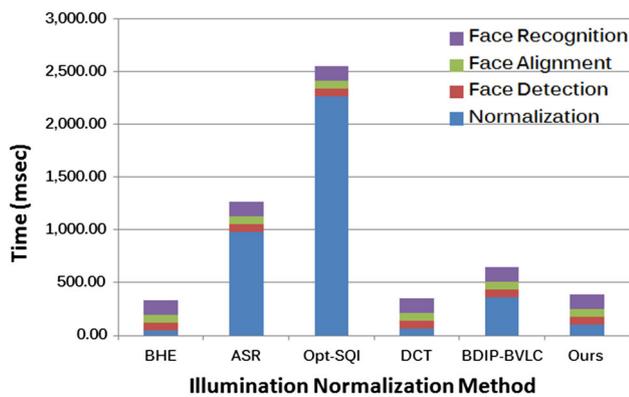


Fig. 20 Comparison of computation time of each illumination invariant method using 25 training data and 10 subspaces

walked following the camera at a distance of about two meters and were conditioned to always look ahead. Each video was captured at our campus in both indoor and outdoor scenes. The indoor video was taken in a corridor with various lighting conditions. The outdoor video was taken during the daytime. The duration of each video recording was one minute. Each video was captured at 30 frames per second by a web camera.

Since the performance of the MSM depends on the number of the reference images and the subspaces used, the experiment is divided into two parts: first, using only five reference images and five subspaces and second, using 25 reference images and increasing the number of subspace steps by five. The performance of each illumination invariant method is measured by counting the number of correctly recognized people.

5.5.2 Experimental results

The comparison result of the selected illumination invariant methods is shown in Fig. 17. Our method and ASR equally outperform the others for all combinations in the number of training images and subspace. The results of both methods increase from around 70–80% by increasing the number of training images, even while maintaining the same number of subspace. Overall, from the methods in comparison, Opt-SQI and DCT recorded the highest increases in performance. The recognition rate increases when we increase the number of subspace. In contrast, the recognition rate of BDIP-BVLC and BHE shows constant and negative trends, respectively. Nevertheless, Opt-SQI and DCT frequently fail to control final appearances in some harsh illumination conditions, whereas BHE and BDIP-BVLC are frequently too rough. As a consequence, important parts of the face are not clear and tend to be vague as shown in Fig. 18.

Figure 19 shows the experimental results of three people using our illumination invariant method. Our method can produce a stable invariant image while retaining the clarity

of the parts of the face. With these capabilities, the classifier will be easier to recognize the identity of the person.

5.5.3 Computation time

The average computation time of the illumination invariant method, the face detection, the face alignment and the multiple-view face recognition processes of each illumination invariant method are presented in Fig. 20. Although our illumination invariant method is slightly slower than BHE and DCT, it is much faster than BDIP-BVLC, ASR and Opt-SQI. From Fig. 17, even though the performance of ASR is comparable to our method, based on Fig. 20, our method is 3.57 times faster than ASR. This means that our method is more applicable for real implementation (≈ 3 fps).

6 Conclusion

We have proposed a novel illumination invariant method based on an extended reflectance model combined with a Fuzzy Inference System whose rules are optimized using a genetic algorithm. Our method is very effective in helping face recognition task since it only requires very low computation time with high performance. Our method can produce better illumination invariant images. Using these images is proved to be effective especially under very environmental illumination conditions because our method guarantees that face features can still be seen clearly. The texture of each object can be enhanced and is very helpful for a face detector to distinguish between human face and others. By those advantages, our method directly contributes to two face-related tasks, face detection and face recognition.

Based on the experimental results, our method successfully identifies people in very poor lighting conditions. However, all of the above results can only be achieved in a single view. For our next step, we want our method to identify people from multi-views, where people do not have always to look ahead or at the camera.

Acknowledgments We would like to thank to the Directorate General of Higher Education, Ministry of Research, Technology, and Higher Education of Indonesia for financially supporting the first author under Grant No. 224/E4.4/K/2012. This work is also in part supported by JSPS Kakenhi No. 25280093.

References

- Xie, X., Lam, K.M.: Face recognition under varying illumination based on a 2-D face shape model. *Pattern Recogn.* **38**(2), 221–230 (2005)
- Gross, R., Brajovic, V.: An image pre-processing algorithm for illumination invariant face recognition. In: 4th International Confer-

- ence on Audio and Video Based Biometric Person Authentication, pp. 10–18 (2003)
3. Nam, G.P., Park, K.R.: New fuzzy-based retinex method for the illumination normalization of face recognition. *Int. J. Adv. Robot. Syst.* **9**, 103 (2012)
 4. Wang, H., Li, S.Z., Wang, Y., Zhang, J.: Self quotient image for face recognition. *Int. Conf. Image Process.* **2**, 1397–1400 (2004)
 5. Nishiyama, M., Kozakaya, T., Yamaguchi, O.: Illumination Normalization using Quotient Image-based Techniques, Recent Advances in Face Recognition, pp. 97–108 (2008)
 6. Park, Y.K., Park, S.L., Kim, J.K.: Retinex method based on adaptive smoothing for illumination invariant face recognition. *Sig. Process.* **88**, 1929–1945 (2008)
 7. Perez, C.A., Castillo, L.E.: Illumination compensation for face recognition based on genetic optimization of the self-quotient image method. In: *International Symposium on Optomechatronic Technologies*, pp. 322–327 (2009)
 8. Chen, W., Er, M.J., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans. Syst. Man Cybern.* **36**(2), 458–466 (2006)
 9. Kim, W., Suh, S., Hwang, W., Han, J.J.: SVD face: illumination-invariant face representation. *IEEE Signal Process. Lett.* **21**(11), 1336–1341 (2014)
 10. Luo, Y., Guan, Y.P., Zhang, C.Q.: A Robust Illumination Normalization Method Based on Mean Estimation for Face Recognition. *ISRN Machine Vision*, pp. 1–10 (2013)
 11. Zhang, T., Tang, Y.Y., Fang, B., Shang, Z., Liu, X.: Face recognition under varying illumination using gradientfaces. *IEEE Trans. Image Process.* **18**(11), 2599–2606 (2009)
 12. Ju, Y.A., Soo, H.J., Kim, N.C., Kim, M.H.: Face Recognition Using Local Statistic of Gradients and Correlations. In: *18th EUSIPCO*, pp. 1169–1173 (2010)
 13. Rui, T., Yang, Z., Liu, F., Jiang, S., Li, H.: Block-based face recognition using WLD. *PCM* **2013**, 811–822 (2013)
 14. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. *ECCV* **2004**, 469–481 (2004)
 15. Zhu, J.Y., Zheng, W.S., Lai, J.H.: Logarithm gradient histogram: a general illumination invariant descriptor for face recognition. In: *The 10th IEEE International Conference and Workshop on Automatic Face and Gesture Recognition*, pp. 1–8 (2013)
 16. Mamdani, E.H.: Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Trans. Comput. C* **26**(12), 1182–1191 (1977)
 17. Holland, J.H.: *Genetic Algorithms*, pp. 66–72. Scientific American, New York (1992)
 18. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Conference on Computer Vision and Pattern Recognition* (2001)
 19. Guo, R., Dai, Q., Hoeim, D.: Single-image shadow detection and removal using paired regions. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2033–2040 (2011)
 20. Land, E.: Recent advances in retinex theory. *Vision. Res.* **26**(1), 7–21 (1986)
 21. Peer, P., Solina, F.: An automatic human face detection method. In: *Proceedings of the 4th Computer Vision Winter Workshop*, pp. 122–130 (1999)
 22. Elgammal, A., Muang, C., Hu, D.: *Skin Detection - A Short Tutorial*. Encyclopedia of Biometrics. Springer, Berlin (2009)
 23. Grove, G., Zerweck, C., Damia, J.: Human skin coloration using the RGB color space model. In: *The 4th International Symposium of L'Oreal Institute for Ethnic Hair and Skin Research* (2007)
 24. Dewantara, B.S.B., Miura, J.: Fuzzy-based illumination normalization for face recognition. In: *IEEE Workshop on Advanced Robotics and Its Social Impacts*, pp. 131–136 (2013)
 25. Georghiadis, A.S., Belhumeur, P.N., Kriegman, D.J.: From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 643–660 (2001)
 26. Nanaa, K., Rizon, M., Rahman, M.N.A., Almejrads, A., Aziz, A.Z.A., Mohamed, S.B.: Eye detection using composite cross-correlation. *Am. J. Appl. Sci.* **10**(11), 1448–1456 (2013)
 27. Fukui, K., Yamaguchi, O.: Face recognition using multi-viewpoint patterns for robot vision. In: *11th International Symposium of Robotics Research*, pp. 192–201 (2003)
 28. Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D.: The CAS-PEAL large-scale chinese face database and baseline evaluations. *IEEE Trans. Syst. Man Cybern. Part A Syst. Humans* **38**(1), 149–161 (2008)
 29. Turk, M.A., Pentland, A.P.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**, 71–86 (1991)
 30. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**(27), 1–27 (2011)



Bima Sena Bayu Dewantara received the B.Eng. degree in information technology from Electronic Engineering Polytechnic Institute of Surabaya, Indonesia, and M.Eng. degree in electrical engineering from Sepuluh Nopember Institute of Technology, Indonesia, in 2004 and 2010, respectively. He joined the Department of Electronic Engineering at Electronic Engineering Polytechnic Institute of Surabaya, Indonesia, as lecturer in 2005. Then, he moved to the

Department of Informatics and Computer Science in 2007. He is currently a PhD candidate in Graduate School of Computer Science and Engineering at Toyohashi University of Technology, Japan. His research interests include pattern recognition, computer vision, machine learning and robotics system.



Jun Miura received the B.Eng. degree in mechanical engineering in 1984, the M.Eng. and the Dr.Eng. degrees in information engineering in 1986 and 1989, respectively, all from the University of Tokyo, Tokyo, Japan. In 1989, he joined Department of Computer-Controlled Mechanical Systems, Osaka University, Suita, Japan. Since April 2007, he has been a Professor at Department of Computer science and Engineering, Toyohashi University of Technology, Toyohashi, Japan. From March 1994 to February 1995, he was a Visiting Scientist at Computer Science Department, Carnegie Mellon University, Pittsburgh, PA. He received several awards including Best Paper Award from the Robotics Society of Japan in 1997, Best Paper Award Finalist at ICRA-1995 and Best Service Robotics Paper Award Finalist at ICRA-2013. Prof. Miura published over 180 papers in international journals and conferences in the areas of intelligent robotics, mobile service robots, robot vision and artificial intelligence.