PAPER

Illumination Normalization-Based Face Detection under Varying Illumination

Min YAO^{†a)}, Nonmember, Hiroshi NAGAHASHI^{††}, and Kota AOKI^{††}, Members

SUMMARY A number of well-known learning-based face detectors can achieve extraordinary performance in controlled environments. But face detection under varying illumination is still challenging. Possible solutions to this illumination problem could be creating illumination invariant features or utilizing skin color information. However, the features and skin colors are not sufficiently reliable under difficult lighting conditions. Another possible solution is to do illumination normalization (e.g., Histogram Equalization (HE)) prior to executing face detectors. However, applications of normalization to face detection have not been widely studied in the literature. This paper applies and evaluates various existing normalization methods under the framework of combining the illumination normalization and two learning-based face detectors (Haar-like face detector and LBP face detector). These methods were initially proposed for different purposes (face recognition or image quality enhancement), but some of them significantly improve the original face detectors and lead to better performance than HE according to the results of the comparative experiments on two databases. Meanwhile, we propose a new normalization method called segmentation-based half histogram stretching and truncation (SH) for face detection under varying illumination. It first employs Otsu method to segment the histogram (intensities) of the input image into several spans and then does the redistribution on the segmented spans. In this way, the non-uniform illumination can be efficiently compensated and local facial structures can be appropriately enhanced. Our method obtains good performance according to the experiments.

key words: illumination normalization, Otsu method, segmentation-based, Haar-like face detector, LBP face detector

1. Introduction

Face detection is a popular research topic and many efforts have been made to improve its effectiveness and efficiency [1]–[4]. Face detection methods can be mainly clustered into four categories [1]: knowledge-based methods, feature invariant methods, template matching methods, and appearance-based methods. In general, appearance-based methods, which are also known as learning-based methods [3], demonstrate superior performance and thus have dominated the recent advances in face detection.

The most famous learning-based method is Haar-like face detector [5]. Inspired by this seminal work, many extended or improved versions have been proposed to tackle various specific issues [6]–[13]. Among these issues, the illumination problem is one of the most important factors

decreasing the performance of learning-based methods.

One solution to the illumination problem confronting learning-based methods is to use illumination invariant features. Since Haar-like facial features are considered sensitive to illumination, many alternatives have been created, such as modified census transform (MCT), improved local binary patterns (ILBP) [11], multi-block local binary patterns MLBP [12] and the recently proposed Haar local binary patterns (HLBP) [14]. However, it has been pointed out that actually there are no illumination invariant features [15]. These feature-focused modifications based on learning algorithms are still vulnerable to varying illumination. There is also a need of “re-learning” when the goal is to improve the detection performance and the learning process is a heavy task.

Another solution tries to utilize skin color information to balance and compensate the illumination [16]–[18]. However, skin colors vary largely in different illumination. It is difficult to model the extremely illuminated skin colors without prior knowledge. On the other hand, the illumination of the test images used in these researches are still even enough and well controlled.

A third way is to do illumination normalization as a preprocessing step. The famous Histogram Equalization (HE) has been applied to face detection [5], [19]. However, to our best knowledge, other normalization methods besides HE, have been rarely tried to solve the illumination problem in face detection. Applications of normalization to face detection have not been widely studied in the literature.

In this paper, we focus on the manipulation of combining the illumination normalization and two famous learning-based face detectors (Haar-like and LBP face detectors). It is aimed at maintaining the advantages of pre-trained learning-based face detectors while improving the resistance to varying illumination. The preliminary study can be found in [20]. Figure 1 shows the flowchart of the general framework. We summarize the contributions of this paper as follows:

- 1) We widely apply a variety of existing normalization methods (besides HE) to normalize the illumination in face detection and evaluate them on large databases. This work is inspiring since the influence of prepended normalization on the face detection performance has not received the attention that it deserves. The results demonstrate that several methods show potential for further improvements of the two famous face detectors (better than HE). We also reach an interesting conclusion from the comparative study.

Manuscript received September 27, 2013.

Manuscript revised January 20, 2014.

[†]The author is with the Department of Information Processing, Tokyo Institute of Technology, Yokohama-shi, 226–8503 Japan.

^{††}The authors are with the Imaging Science and Engineering Laboratory, Tokyo Institute of Technology, Yokohama-shi, 226–8503 Japan.

a) E-mail: yao.m.aa@m.titech.ac.jp

DOI: 10.1587/transinf.E97.D.1590

Fig. 1 General framework.

2) We propose a new illumination normalization method called segmentation-based half histogram truncation and stretching (SH) to tackle the illumination problem in face detection. Otsu method is firstly used to segment the histogram (intensities) of the input image into several spans. Then, the truncation and stretching is operated to redistribute the segmented spans. The proposed method aims to expand the intensity range containing faces or facial organs to a large extent in order to enhance the facial structures. According to the experimental results, our method can significantly improve the original face detectors and yield better performance than most of the other tested methods.

The rest of this paper is organized as follows. Section 2 introduces the learning-based face detectors used in this work. In Sect. 3, descriptions of the related normalization methods are provided. In Sect. 4, the proposed method is explained in detail. Then the experiments are presented in Sect. 5, followed by the final conclusions in Sect. 6.

2. Face Detectors Used in This Work

A learning-based face detector refers to the face detection technique which constructs a classifier based on a specific learning algorithm using sophisticatedly selected facial features. Widely used learning algorithms and facial features are boosting-based algorithms [3], [5], [21]–[24] and neighborhood comparison (intensity contrast) illustrators [5], [10], [25]. The basic concept of these learning-based methods is to calculate corresponding facial features in multi-scaled analysis windows and to use a specific boosting algorithm for training out of an effective classifier. But being subject to the training samples and efficiency requirement, the learning-based face detector tends to be weak in detecting faces under varying illumination.

In this work, we utilize two learning-based (gentle-boosting-based) face detectors [26], namely, Haar-like face detector and LBP face detector[†]. They are advantageous for the richness of feature set, efficient feature selection and fast evaluation of used features. Figure 2 illustrates some Haar-like features (see (a)) and LBP features (see (b)). As seen from these features, they largely depend on the discriminability of neighbourhood comparison values. However, harsh illumination makes these patterns useless and makes these two detectors incapable under varying illumination. Correspondingly, we apply various normalization

Fig. 2 (a) are some Haar-like features shown in stick figures (left) and of real faces (right); (b) are some LBP features.

methods, trying to gain proper contrasts and preserve these patterns in the given image.

3. Related Normalization Methods

The normalization methods applied in this work can be generally grouped into three categories: histogram equalization variants, histogram remapping methods, and face illumination normalization techniques. In this section we will give an overview of them.

3.1 Histogram Equalization Variants

Histogram equalization (HE) is a commonly used method for image contrast enhancement. It obtains a uniform histogram for the output image. However, HE often results in lacks of local enhancement and yields undesirable artifacts and noise [27], [28]. Later, adaptive histogram equalization (AHE) [29] was introduced to improve the local contrast enhancement. But it also suffers from the noise amplification in “flat” regions and “ring” artifacts at strong edges [30]. Contrast-limited AHE (CHE) was proposed to reduce these drawbacks. On the other hand, many bi-histogram equalization methods have been proposed to overcome the problems of HE such as brightness preserving bi-histogram equalization (BBHE) [31], equal area dualistic sub-image histogram equalization (DSIHE), minimum mean brightness error bi-histogram equalization (MMBEBHE). They divide the histogram of an input image into two parts based on the mean value, median value, and minimum mean difference threshold, respectively and then equalize each part independently. In [27], it was considered that the histogram division approaches in these methods may cause more annoying side effects and thus a new normalization method called range limited bi-histogram equalization (RHE) was proposed. RHE uses Otsu method to choose a proper threshold for histogram division and then does the range limitation and equalization to achieve minimum absolute mean brightness error. Among the above mentioned methods, we will examine HE, AHE, CHE, and RHE.

3.2 Histogram Remapping Methods

Since many normalization methods are based on the image histogram, in [32], a more general concept of histogram remapping was discussed, which remaps the input histogram to a target distribution of different shapes. In this work, we will test three histogram remapping methods with target distribution shape of normal distribution (ND), lognormal distribution (LN), and exponential distribution (EX) [32].

[†]Note that the implementation of LBP face detector in [26] uses Multi-block LBP (MLBP) features. In these features, the neighborhood mean value is used instead of one pixel value. The mean value can be changed by the normalization and thus LBP face detector is likely to benefit from the normalization. For more details about the MLBP features, refer to [12]. In this paper, we call MLBP features as LBP features for simplicity.

3.3 Face Illumination Normalization Techniques

Face detection is highly related to face recognition and numerous face illumination normalization methods were developed to tackle illumination problem in face recognition [33]–[35]. In this paper, we will investigate four face illumination normalization techniques for face detection. They are ASR (adaptive single-scale retinex) [36], DCT (discrete cosine transform-based method) [37], TT (Tan and Triggs' method) [38], and WF (Weber face) [39]. All of them are classical or state-of-the-art methods. They are particularly selected in our tests due to their abilities in 1) maintaining most of original visual patterns (unlike GRF [40]), and 2) removing many illumination effects (unlike the early work WA [41]).

4. Segmentation-Based Half Histogram Truncation and Stretching (SH)

As illumination has different effects in different regions, a possible way to enhance local facial structures is to firstly partition the input image into several regions, and then execute normalization within each region separately. Hence SH is formally defined by two procedures: image segmentation and histogram redistribution.

In [42], face images are roughly partitioned into four geometrically regular parts for illumination invariant face recognition. Yet regular regions are somewhat arbitrary and difficult to conquer the non-uniform illumination problem. AHE, CHE and some bi-histogram equalization methods also adopt the idea of local normalization, but they will lead to undesired artifacts, saturated brightness or lacks of local contrast enhancement. SH uses Otsu method [43] for image segmentation, similar to RHE [27], but the difference is that SH applies multi-segmentation and redistribution of the histogram to a large range. Furthermore, during the redistribution, histogram truncation is used to enlarge the range rather than limiting the range in RHE. Figure 3 gives a simple illustration of SH.

4.1 Otsu Image Segmentation

Given an input image $I(x, y)$ with n pixels and a total num-

Fig. 3 Illustration of the proposed method. Segmentation of the input image histogram (top) and output histogram after redistribution (bottom).

ber of L grey levels, the probability of pixel occurrence is denoted by $p_r(r_q) = n_q/n$, where n_q ($q = 0, 1, \dots, L-1$) is the number of pixels with r_q grey level.

If intensities of this image are segmented into N spans ($S_1, S_2, \dots, S_N$), a series of thresholds $\mathbf{k} = \{k_1, k_2, \dots, k_{N-1}\}$ are required. Then the cumulative probability ω_m , mean level μ_m for each span and the mean intensity of the whole input image μ_T are given by $\omega_m = \sum_{q \in S_m} p_r(r_q)$, $\mu_m = \sum_{q \in S_m} r_q p_r(r_q) / \omega_m$, and $\mu_T = \sum_{q=0}^{L-1} r_q p_r(r_q)$. Otsu method selects the optimal $\mathbf{k}^* = \{k_1^*, k_2^*, \dots, k_{N-1}^*\}$ that maximizes the inter-class variance σ_B^2 :

$$\mathbf{k}^* = \arg \max_{1 \leq k_1 < \dots < k_{N-1} < L} \{\sigma_B^2(\mathbf{k})\}, \quad (1)$$

where $\sigma_B^2 = \sum_{m=1}^N \omega_m (\mu_m - \mu_T)^2$. After this segmentation, intensity values within each span are more uniform than those in the global level, and so is the illumination.

4.2 Histogram Redistribution

Previously, Otsu method segments the input image into N intensity spans; there are $N-1$ thresholds $k_1^*, k_2^*, \dots, k_{N-1}^*$ in ascending order. These spans are separated into two groups: the lowest intensity span and the rest intensity spans. The redistribution step contains the histogram truncation and stretching [44], which is operated in the two groups of spans independently. Truncation means to cut off a specific percentage of the lower and upper ends of one span group so as to enlarge the intensity range and remove spurious effects caused by few very bright or dark pixels. For the purpose of enlarging the range in the dark regions, the lowest intensity span is stretched to the low half band of the target distribution and the rest intensity spans to the high half band. In this way, the target distribution can cover the entire band. There are two assumptions in this redistribution.

1) Faces are probably with low intensity values (i.e., in the dark region) under varying illumination.

2) Even if faces are not wholly in the dark region, the facial organs are probably with low intensity values (i.e., in the dark region) under varying illumination.

In this sense, the stretched redistribution of the lowest intensity span to the low half band is supposed to be able to emphasize the facial structures.

As mentioned above, the region belonging to the lowest intensity span can be labelled corresponding to grey levels $\mathbf{r}_{\text{low}} = \{r_0, r_1, \dots, r_{k_1^*-1}\}$ and the rest regions to grey levels $\mathbf{r}_{\text{high}} = \{r_{k_1^*}, r_{k_1^*+1}, \dots, r_{L-1}\}$. L denotes the total number of grey levels of the input image. The number of pixels to be cut off at both ends of the histogram within the lowest intensity span can be calculated by

$$C_{\text{low}} = \sum_{q=0}^{k_1^*-1} n_q \times \gamma_1, \quad (2)$$

where, γ_1 indicates the cut-off percentage. With histogram redistribution, the goal is to get a new intensity value $r'_q(x, y)$ of target pixel (x, y) with initial intensity r_q :

$$r'_q = \begin{cases} 0, & r_q \leq r_s \\ \left(\frac{r_q - r_s}{r_t - r_s} \right) \times 127, & r_s < r_q < r_t \\ 127, & r_q \geq r_t \end{cases}, \quad (3)$$

where r_s and r_t are the cut-off intensity values of the lower and upper end of the histogram within the lowest intensity span, respectively. They satisfy the following equations.

$$\sum_{q=0}^s n_q \leq C_{\text{low}} < \sum_{q=0}^{s+1} n_q \quad (4)$$

$$\sum_{q=t}^{k_1^*-1} n_q \leq C_{\text{low}} < \sum_{q=t-1}^{k_1^*-1} n_q \quad (5)$$

It can go through a similar process within the rest intensity spans. Firstly calculate the number of pixels to be cut off at the end by

$$C_{\text{high}} = \sum_{q=k_1^*}^{L-1} n_q \times \gamma_2, \quad (6)$$

where γ_2 indicates the cut-off percentage within the intensity spans under consideration. Then find the cut-off intensity values r_u and r_v , letting

$$\sum_{q=k_1^*}^u n_q \leq C_{\text{high}} < \sum_{q=k_1^*}^{u+1} n_q, \quad (7)$$

$$\sum_{q=v}^{L-1} n_q \leq C_{\text{high}} < \sum_{q=v-1}^{L-1} n_q. \quad (8)$$

Finally the new pixel intensity value $r_q(x, y)'$ to the high half band can be got by

$$r'_q = \begin{cases} 128, & r_q \leq r_u \\ \left(\frac{r_q - r_u}{r_v - r_u} \right) \times 127 + 128, & r_u < r_q < r_v \\ 255, & r_q \geq r_v \end{cases}. \quad (9)$$

Figure 4 gives some visual samples using different normalization methods. It can be seen that some of the existing methods may change the original facial patterns, some may introduce intensity saturation, and some are likely to result in lacks of contrast enhancement. Our method can efficiently compensate the non-uniform illumination and appropriately enhance the local facial structures.

5. Experiments

In the experiments, we used the pre-trained implementation of Haar-like face detector and LBP face detector [26]. This would allow us to understand how illumination normalization can affect the original learning-based face detectors without re-learning. Two databases are used for assessment: extended Yale B database [45] and natural database.

5.1 Parameter Selection

The parameters in the applied normalization methods were

Fig. 4 Sample images and their corresponding processed images using different normalization methods.

Fig. 5 Images processed by SH using different parameters.

set as recommended or used in the reference papers. In our method, N is the number of spans segmented by Otsu method; γ_1 and γ_2 denote the percentage of pixels to be cut off within the lowest intensity span and within the rest intensity spans of the input image, respectively. If N is too small, the dark region which probably contains faces or facial organs can not be enhanced enough. If it is too large, the faces or facial organs may not be included in the low intensity span and thus will not be appropriately enhanced as well. γ_1 and γ_2 are used to further enlarge the local enhancement but there will be losses of useful details if they are too large. Figure 5 shows some output images processed by our method using different parameters where $\gamma_1 = \gamma_2 = \gamma$. Figure 6 shows the relation between the detection rate of Haar-like face detector and the parameters based on the natural database. Since $N = 4$ with $\gamma = 5$ yields the highest rate, we

Fig. 6 Detection rate versus N and γ .

Fig. 7 Sample images of subject 1, 3, 10, 11 and 16 (from left to right).

used these values through our experiments. There are also other parameter selection methods such as cross validation, which can be used in the future work.

5.2 Results on Extended Yale B Database

Extended Yale B database contains 16128 gray images of 28 human subjects with 9 poses and 64 illumination conditions (576 images for each subject). It is a popular database and is commonly used in the testing task of face recognition and detection. Based on this database, we evaluate how these normalization methods perform with different subjects and different light sources.

Test with Different Subjects. Since the illumination conditions for each subject are similar, we did test on 5 selected subjects (subject 1, 3, 10, 11, and 16) with totally 2880 faces, which vary in gender, region (western/eastern), skin color, and facial structure. Sample images of the 5 selected subjects are shown in Fig. 7. Tables 1 and 2 show the comparative results (in the form of Recall/Precision/F-measure) with Haar-like face detector and LBP face detector, respectively. The results of the original face detectors (ORI) are also given as the baseline. Recall, Precision and F-measure were calculated by

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (10)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (11)$$

$$\text{F-measure} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (12)$$

where TP, FN, and FP are the number of true positives, false negatives, and false positives, respectively. Therefore, Recall reflects the ability to detect true faces, Precision reflects the resistance to face-like non-face patterns, and F-measure demonstrates the performance accounting for both precision and recall.

From Table 1, we can see that the performance of

Fig. 8 Sample images from subset 1 to 5 (from left to right).

each method varies largely for different subjects. Generally speaking, subject 3, 10, and 16 get similar performance; subject 1 has the best recall; the recall of subject 11 is much lower than the other subjects. ASR reaches very high results for subject 11 who is with the black skin. But its precision is not stable. By contrast, histogram equalization variants, histogram remapping methods, and our method have relatively stable performance. The face illumination normalization techniques DCT, TT, and WF have rather poor performance, even worse than the original Haar-like face detector for some subjects. In Table 2, the rank of the performance for different subjects is similar to that with Haar-like face detector. This demonstrates that subjects have a specific influence on detection results. This time, ASR get unsatisfactory results for the entire subjects; DCT, TT, and WF can not improve the original LBP face detector at all. Still, histogram equalization variants, histogram remapping methods, and our method obtain stable and significant improvements of the original face detector for all the subjects (except AHE).

Test with Different Light Sources. All face images of the five subjects were divided into 5 subsets according to different lighting angles, resulting in subset 1 (0° to 12° , 315 images), subset 2 (13° to 25° , 540 images), subset 3 (26° to 50° , 540 images), subset 4 (51° to 77° , 630 images), and subset 5 (above 78° , 855 images). Sample images of one subject from subset 1 to 5 are shown in Fig. 8. As can be seen from the results shown in Tables 3 and 4, the more extreme the illumination condition is, the worse the general performance of all the methods becomes and, however, the more improvements of the original face detectors can be achieved. The recall of ASR is high and stable with Haar-like face detector under different light sources but low with LBP face detector, compared with histogram equalization variants, histogram remapping methods, and our method. The other face illumination normalization techniques except ASR are unreliable for all of the lighting conditions.

Tables 5 and 6 give the average performance using Haar-like face detector and LBP face detector, respectively. Since the detection rates of TT with both detectors are extremely bad, its precisions are meaningless and listed just for reference. The case of WF with LBP face detector is similar. It can be noticed that most of the face illumination normalization techniques are not effective in improving the original detectors except for ASR. ASR accomplishes the highest recall with Haar-like face detector. But the precision and its performance with LBP face detector is not as comparable as the histogram-based methods and our method. When looking at the F-measure, face illumination normalization techniques have the lowest performance generally. Histogram equalization variants perform better. Histogram remapping

Table 1 Results (Recall(%)/Precision(%)/F-measure(%)) using Haar-like face detector on different subjects.

Subject	1	3	10	11	16
ORI	87.15 / 99.60 / 92.96	74.48 / 96.62 / 84.12	72.22 / 99.76 / 83.79	59.20 / 98.27 / 73.89	80.73 / 97.28 / 88.24
ASR	92.36 / 63.26 / 75.09	98.78 / 62.60 / 76.63	98.26 / 97.42 / 97.84	99.83 / 99.65 / 99.74	99.13 / 98.96 / 99.05
DCT	97.57 / 98.60 / 98.08	59.72 / 93.99 / 73.04	78.99 / 100.00 / 88.26	44.79 / 98.47 / 61.58	63.72 / 99.73 / 77.75
TT	63.54 / 100.00 / 77.71	10.59 / 100.00 / 19.15	30.38 / 100.00 / 46.60	19.44 / 100.00 / 32.56	24.65 / 100.00 / 39.55
WF	95.49 / 98.21 / 96.83	68.40 / 97.77 / 80.49	67.19 / 99.49 / 80.21	69.44 / 100.00 / 81.97	73.44 / 99.76 / 84.60
AHE	99.13 / 93.00 / 95.97	80.21 / 82.49 / 81.55	84.72 / 93.67 / 88.97	77.43 / 88.14 / 82.44	83.68 / 93.05 / 88.12
RHE	99.65 / 99.31 / 99.48	93.58 / 98.36 / 95.91	90.63 / 93.72 / 92.14	84.03 / 97.38 / 90.21	93.23 / 98.17 / 95.64
CHE	97.92 / 99.65 / 98.77	83.51 / 98.36 / 90.33	88.54 / 99.61 / 93.75	75.35 / 99.09 / 85.60	88.19 / 97.32 / 92.53
HE	99.83 / 98.97 / 99.39	92.36 / 96.90 / 94.58	92.36 / 90.94 / 91.65	85.24 / 81.83 / 83.50	94.79 / 98.56 / 96.64
EX	100.00 / 90.57 / 95.05	96.18 / 80.76 / 87.80	91.49 / 94.28 / 92.86	89.76 / 95.04 / 92.32	94.62 / 99.63 / 97.06
LN	100.00 / 99.65 / 99.83	93.06 / 98.17 / 95.54	96.01 / 98.40 / 97.19	83.68 / 96.59 / 89.67	96.01 / 98.22 / 97.10
ND	100.00 / 100.00 / 100.00	94.44 / 96.28 / 95.35	96.53 / 97.37 / 96.95	86.11 / 87.94 / 87.02	96.35 / 97.54 / 96.94
SH	100.00 / 99.83 / 99.91	94.62 / 97.85 / 96.20	96.70 / 98.76 / 97.72	85.59 / 99.20 / 91.89	96.70 / 98.93 / 97.81

Table 2 Results (Recall(%)/Precision(%)/F-measure(%)) using LBP face detector on different subjects.

Subject	1	3	10	11	16
ORI	92.19 / 94.32 / 93.24	64.41 / 83.56 / 72.75	67.88 / 90.09 / 77.43	51.04 / 62.42 / 56.16	74.83 / 83.85 / 79.08
ASR	80.21 / 83.70 / 81.91	53.13 / 91.62 / 67.25	61.98 / 84.60 / 71.54	43.40 / 92.94 / 59.17	79.86 / 91.27 / 85.19
DCT	94.62 / 85.42 / 89.79	15.97 / 31.08 / 21.10	52.60 / 77.30 / 62.60	10.24 / 48.36 / 16.91	30.38 / 68.36 / 42.07
TT	0.17 / 25.00 / 0.34	0.00 / 0.00 / -	0.00 / 0.00 / -	0.00 / 0.00 / -	0.00 / 0.00 / -
WF	5.73 / 100.00 / 10.84	0.17 / 100.00 / 0.35	0.87 / 100.00 / 1.72	0.00 / - / -	18.40 / 96.36 / 30.90
AHE	94.79 / 65.23 / 77.28	56.94 / 55.97 / 56.45	65.28 / 73.29 / 69.05	47.05 / 46.64 / 46.85	75.69 / 57.75 / 65.51
RHE	96.53 / 93.45 / 94.96	74.65 / 87.93 / 80.75	72.22 / 81.09 / 76.40	59.20 / 70.60 / 64.40	81.60 / 91.98 / 86.48
CHE	95.14 / 85.09 / 89.84	64.93 / 75.86 / 69.97	76.91 / 85.19 / 80.84	59.20 / 57.80 / 58.49	80.90 / 80.76 / 80.83
HE	96.35 / 94.55 / 95.44	71.35 / 89.15 / 79.27	75.00 / 91.53 / 82.44	62.15 / 73.06 / 67.17	82.99 / 90.87 / 86.75
EX	97.22 / 95.73 / 96.47	69.79 / 92.41 / 79.53	72.57 / 90.48 / 80.54	60.76 / 79.37 / 68.83	83.33 / 92.49 / 87.67
LN	96.18 / 94.86 / 95.52	69.44 / 90.50 / 78.59	75.52 / 85.29 / 80.11	61.98 / 72.86 / 66.98	81.77 / 92.53 / 86.82
ND	97.05 / 94.75 / 95.88	71.53 / 91.35 / 80.23	76.39 / 81.78 / 78.99	65.45 / 75.86 / 70.27	83.16 / 91.06 / 86.93
SH	97.22 / 95.40 / 96.30	71.35 / 92.57 / 80.59	75.69 / 90.27 / 82.34	63.19 / 76.79 / 69.33	83.33 / 93.75 / 88.24

Table 3 Results (Recall(%)/Precision(%)/F-measure(%)) using Haar-like face detector on different light sources.

Subset	1	2	3	4	5
ORI	100.00 / 97.52 / 98.75	99.07 / 99.07 / 99.07	92.96 / 98.82 / 95.80	76.35 / 98.36 / 85.97	37.43 / 96.97 / 54.01
ASR	98.73 / 81.63 / 89.37	99.81 / 81.30 / 89.61	99.81 / 79.62 / 88.58	97.62 / 80.92 / 88.49	94.62 / 80.58 / 87.04
DCT	66.03 / 96.30 / 78.34	74.63 / 98.53 / 84.93	67.59 / 99.18 / 80.40	69.84 / 99.10 / 81.94	66.67 / 97.60 / 79.22
TT	50.48 / 100.00 / 67.09	37.22 / 100.00 / 54.25	34.44 / 100.00 / 51.24	33.17 / 100.00 / 49.82	11.81 / 100.00 / 21.13
WF	99.05 / 99.36 / 99.21	93.52 / 99.02 / 96.19	84.07 / 98.91 / 90.89	81.11 / 99.22 / 89.26	43.51 / 98.41 / 60.34
AHE	96.83 / 88.15 / 92.28	92.96 / 90.78 / 91.86	76.11 / 86.89 / 81.15	80.16 / 91.65 / 85.52	84.91 / 91.55 / 88.11
RHE	99.68 / 97.82 / 98.74	99.44 / 97.99 / 98.71	92.22 / 99.01 / 95.49	89.05 / 96.56 / 92.65	87.25 / 96.38 / 91.59
CHE	100.00 / 98.75 / 99.37	99.81 / 99.45 / 99.63	94.44 / 99.42 / 96.87	88.41 / 98.24 / 93.07	67.37 / 98.29 / 79.94
HE	100.00 / 94.59 / 97.22	99.63 / 94.72 / 97.11	94.44 / 98.46 / 96.41	90.00 / 92.95 / 91.45	87.25 / 88.81 / 88.02
EX	99.68 / 92.08 / 95.73	99.26 / 93.06 / 96.06	88.89 / 90.91 / 89.89	91.59 / 91.59 / 91.59	94.97 / 90.52 / 92.69
LN	100.00 / 99.37 / 99.68	99.81 / 99.26 / 99.54	94.07 / 98.07 / 96.03	92.54 / 97.82 / 95.11	88.30 / 97.55 / 92.69
ND	100.00 / 97.22 / 98.59	100.00 / 97.12 / 98.54	95.74 / 98.29 / 97.00	93.33 / 95.77 / 94.53	89.71 / 92.97 / 91.31
SH	100.00 / 99.37 / 99.68	99.63 / 98.90 / 99.26	95.56 / 98.85 / 97.18	93.81 / 98.50 / 96.10	89.82 / 99.10 / 94.23

methods and our method obtain equally good results, ranking the top, where our method yields the best score (slightly better than the histogram remapping methods) with both detectors. On the whole, the improvements of LBP face detector are not as large as Haar-like face detector as LBP features are more robust to illumination than Haar-like features.

5.3 Results on Natural Database

The images in Extended Yale B database were taken under controlled light sources. In order to evaluate the illumination insensitive face detection under natural environments, a natural database was created. It contains 840 images with 932 faces. These images were collected from the internet, Bao database [46], and the real-world photographs. Most of

the faces are illuminated in extreme or non-uniform way and meanwhile embedded in complicated natural backgrounds. Figure 9 shows several detection results of our method and the results of the original face detectors. The performance statistics using the two face detectors are given in Tables 7 and 8, respectively. From Table 7, it can be seen that most of the normalization methods (except DCT, TT, and WF) with Haar-like face detector generate much better results than the original detector. Considering the F-measure, SH and EX are the best two methods and HE comes the third. In Table 8, the original LBP face detector demonstrates much higher recall than Haar-like face detector and the improvements achieved by the normalization methods are not so significant. But SH and EX still get the highest values of F-measure. These results demonstrate that normalization as a

Table 4 Results (Recall(%)/Precision(%)/F-measure(%)) using LBP face detector on different light sources.

Subset	1	2	3	4	5
ORI	97.14 / 84.30 / 90.27	94.63 / 87.20 / 90.76	80.19 / 85.74 / 82.87	57.46 / 79.56 / 66.73	47.49 / 78.53 / 59.18
ASR	76.51 / 90.94 / 83.10	72.78 / 88.91 / 80.04	80.93 / 91.23 / 85.77	48.10 / 86.57 / 61.84	53.92 / 84.59 / 65.86
DCT	32.70 / 63.19 / 43.10	35.37 / 69.96 / 46.99	40.37 / 67.91 / 50.64	46.67 / 71.88 / 56.59	43.04 / 68.40 / 52.84
TT	0.00 / 0.00 / -	0.00 / 0.00 / -	0.00 / 0.00 / -	0.16 / 16.67 / 0.31	0.00 / 0.00 / -
WF	0.63 / 100.00 / 1.26	1.30 / 77.78 / 2.55	21.30 / 98.29 / 35.01	1.75 / 100.00 / 3.43	1.17 / 100.00 / 2.31
AHE	90.16 / 64.25 / 75.03	85.74 / 65.67 / 74.38	67.59 / 62.82 / 65.12	55.24 / 51.71 / 53.42	58.13 / 57.06 / 57.59
RHE	97.46 / 87.22 / 92.05	94.44 / 89.01 / 91.64	82.22 / 86.05 / 84.09	61.59 / 81.51 / 70.16	65.96 / 83.68 / 73.77
CHE	97.46 / 78.32 / 86.85	94.81 / 80.00 / 86.78	86.30 / 79.66 / 82.84	67.62 / 75.94 / 71.54	53.92 / 71.36 / 61.43
HE	97.78 / 89.53 / 93.47	95.93 / 90.56 / 93.17	82.41 / 87.77 / 85.00	64.76 / 87.74 / 74.52	64.91 / 85.65 / 73.85
EX	97.78 / 94.48 / 96.10	94.44 / 91.56 / 92.98	80.56 / 90.06 / 85.04	62.22 / 89.09 / 73.27	66.08 / 88.84 / 75.79
LN	97.46 / 90.03 / 93.60	95.56 / 90.69 / 93.06	83.15 / 88.04 / 85.52	64.13 / 86.32 / 73.59	63.27 / 83.62 / 72.04
ND	97.78 / 91.12 / 94.33	95.93 / 89.00 / 92.34	83.33 / 86.87 / 85.07	66.03 / 85.42 / 74.49	67.25 / 84.93 / 75.07
SH	98.10 / 91.42 / 94.64	95.56 / 90.53 / 92.97	83.33 / 89.29 / 86.21	65.71 / 89.22 / 75.69	65.73 / 90.06 / 76.00

Table 5 Average performance using Haar-like face detector.

Method	ORI	ASR	DCT	TT	WF	AHE	RHE	CHE	HE	EX	LN	ND	SH
Recall(%)	74.76	97.67	68.96	29.72	74.79	85.03	92.22	86.70	92.92	94.41	93.75	94.69	94.72
Precision(%)	98.31	80.72	98.27	100.00	98.99	90.17	97.40	98.81	93.27	91.49	98.25	95.85	98.91
F-measure(%)	84.93	88.39	81.04	45.82	85.21	87.53	94.74	92.36	93.09	92.93	95.95	95.27	96.77

Table 6 Average performance using LBP face detector.

Method	ORI	ASR	DCT	TT	WF	AHE	RHE	CHE	HE	EX	LN	ND	SH
Recall(%)	70.07	63.72	40.76	0.03	5.03	67.95	76.84	75.42	77.57	76.74	76.98	78.72	78.16
Precision(%)	83.18	88.18	68.90	2.86	97.32	59.81	85.41	76.91	88.09	90.50	87.46	87.13	90.04
F-measure(%)	76.06	73.98	51.22	0.07	9.57	63.62	80.90	76.16	82.50	83.05	81.88	82.71	83.68

Table 7 Performance statistics using Haar-like face detector.

Method	ORI	ASR	DCT	TT	WF	AHE	RHE	CHE	HE	EX	LN	ND	SH
recall(%)	72.86	82.89	55.62	41.08	36.80	85.33	77.51	82.89	84.60	86.55	79.83	82.40	87.16
precision(%)	96.28	96.44	94.40	92.05	92.90	94.45	95.20	95.76	96.78	96.59	96.03	95.87	97.41
F-measure(%)	82.95	89.15	70.00	56.80	52.71	89.66	85.44	88.86	90.28	91.30	87.18	88.63	92.00

Table 8 Performance statistics using LBP face detector.

Method	ORI	ASR	DCT	TT	WF	AHE	RHE	CHE	HE	EX	LN	ND	SH
recall(%)	89.36	73.35	78.73	12.22	20.90	90.46	89.61	90.10	91.56	91.81	89.73	90.10	92.30
precision(%)	95.06	94.94	95.69	87.72	88.14	93.32	94.34	94.97	95.05	96.41	94.59	94.49	96.79
F-measure(%)	92.12	82.76	86.38	21.46	33.79	91.87	91.91	92.47	93.28	94.05	92.10	92.24	94.49

Fig. 9 Sample results by using the original Haar-like face detector, our method with Haar-like face detector, the original LBP face detector, and our method with LBP face detector (from top row to bottom row).

preprocessing is very potential for improving the pre-trained face detectors. With regard to the general performance, our method is desirable for illumination normalization in face detection.

5.4 Discussions

It is noteworthy that EX and ASR with Haar-like face detector get good recalls when the faces are with black skin or under harsh illumination as the conditions in subset 5. It is because that they largely suppress the influence from the harsh illumination or dark skin colors and maintain the Haar-like features.

ASR, DCT, TT, and WF were all initially proposed for illumination normalization in face recognition. However, in our experiments, DCT, TT, and WF perform poorly in face detection and ASR with LBP face detector also yields bad result. Although these four techniques can result in similar

face images of one subject after normalizing the illumination, which facilitate face recognition, they are not suitable for face detection with the face detector pre-trained by using normal training samples. If one wants to make use of the illumination suppression capability of these techniques, a possible way is to apply them to the training samples as well. Nevertheless, the computational cost will be very high and extensive implementations will be very difficult.

On the other hand, the relatively high improvements achieved by histogram equalization variants, histogram remapping methods, and our method indicate that these methods can balance the image illumination and get proper contrasts. Especially, SH enhances the local facial structures more than purely removing the illumination; it is advantageous in terms of stability. It is also supposed that the truncation operation in SH helps decrease the false positives and the results seem to prove so.

Another notable thing is that TT yields extremely bad results in face detection in our experiments while DCT has relatively “good” results (better than TT and WF). However, in face recognition, TT has the best performance among the four face normalization techniques while DCT performs the worst (according to another study of ours [47]). What’s more, AHE, CHE, and RHE were all developed to improve HE in the field of image enhancement. However, HE outperforms them in face detection under varying illumination. Then, we can draw an interesting conclusion that effective normalization methods in face recognition or image enhancement are not necessarily useful in improving the normally pre-trained face detectors. In summary, illumination processing in face detection is different from that in face recognition and image enhancement. Face recognition is more focusing on eliminating the illumination to make faces of the same subject more similar to each other. Image enhancement aims to improve the image visual quality without distortions (detail losses). For face detection, emphasizing faces to distinguish them from the background (including illumination) is the core work.

Regarding our method SH, it achieves good performance in average. But the parameters were experimentally selected and could not guarantee good results with all images. SH sacrifices some image details for better local enhancement and this leads to false negatives in some cases. If the background of the image is too dark (darker than the facial organs) in a large range, the facial features are likely to be included in the light regions. Then the facial features can be lost to some extent because the intensity range of the light regions may be compressed during the computation of SH.

6. Conclusions

In this paper, we applied and evaluated a variety of existing normalization methods for face detection under varying illumination. Some of them significantly improved the performance of the pre-trained face detectors (Haar-like face detector and LBP face detector). An interesting conclusion

was also made that effective normalization methods in face recognition or image enhancement are not necessarily useful in face detection using the normally pre-trained face detectors. Meanwhile, we proposed a new normalization method to compensate non-uniform illumination and enhance the local facial structures. This method proved to be effective in raising the performance of the original face detectors under varying illumination. In the future, we will study ways of achieving a better trade-off between the local contrast enhancement/illumination removal and the loss of useful information.

References

- [1] M. Yang, D. Kriegman, and N. Ahuja, “Detecting faces in images: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.1, pp.34–58, 2002.
- [2] E. Hjeltnas and B. Low, “Face detection: A survey,” *Comput. Vis. Image Und.*, vol.83, no.3, pp.236–274, 2001.
- [3] C. Zhang and Z. Zhang, “A survey of recent advances in face detection,” Technical Report, Microsoft Research, 2010.
- [4] M. Yao, K. Aoki, and H. Nagahashi, “Photograph based pair-matching recognition of human faces,” *Proc. IEICE Gen. Conf.* 2011, p.146, 2011.
- [5] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.1, pp.34–58, 2002.
- [6] T. Mita, T. Kaneko, and O. Hori, “Joint Haar-like features for face detection,” *Proc. ICCV*, vol.2, pp.1619–1626, 2005.
- [7] S. Li, L. Zhu, Z. Zhang, A. Blake, H. Zhang, and H. Shum, “Statistical learning of multi-view face detection,” *Proc. ECCV*, pp.67–81, 2002.
- [8] M. Jones and P. Viola, “Fast multi-view face detection,” Technical Report, Mitsubishi Electric Research Laboratories, TR2003-96, 2003.
- [9] M. Jones, P. Viola, and D. Snow, “Detecting pedestrians using patterns of motion and appearance,” Technical Report, Mitsubishi Electric Research Laboratories, TR2003-90, 2003.
- [10] B. Froba and A. Ernst, “Face detection with the modified census transform,” *Proc. IEEE Int’l. Conf. Automatic Face and Gesture Recognition*, pp.91–96, 2004.
- [11] H. Jin, Q. Liu, H. Lu, and X. Tong, “Face detection using improved LBP under Bayesian framework,” *Proc. Int’l. Conf. Image and Graphics*, pp.306–309, 2004.
- [12] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Li, “Face detection based on multi-block LBP representation,” *Proc. Int’l. Conf. Biometrics*, 2007.
- [13] M. Yao, K. Aoki, and H. Nagahashi, “Segmentation-based illumination normalization for face detection,” *Proc. IEEE Int’l. Workshop Computer Intelligence and Applications*, 2013.
- [14] A. Roy and S. Marcel, “Haar local binary pattern feature for fast illumination invariant face detection,” *British Machine Vision Conference*, 2009.
- [15] H. Chen, P. Belhumeur, and D. Jacobs, “In research of illumination invariants,” *Proc. CVPR*, pp.254–261, 2000.
- [16] R.L. Hsu, M. Abdel-Mottaleb, and A. Jain, “Face detection in color images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.5, pp.696–706, 2002.
- [17] C. Erdem, S. Ulukaya, A. Karaali, and A. Erdem, “Combining Haar feature and skin color based classifiers for face detection,” *Proc. ICASSP*, pp.1497–1500, 2011.
- [18] Z. Tabatabaie, R. Rahmat, N. Udzir, and E. Kheirkhah, “A hybrid face detection system using combination of appearance-based and feature-based methods,” *Int’l. J. Computer Science and Network Security*, vol.9, no.5, pp.181–185, 2009.

- [19] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognit. Lett.*, vol.21, pp.463–471, 2000.
- [20] M. Yao and H. Nagahashi, "Study on illumination insensitive face detection based on normalization techniques," *IEICE Technical Report*, vol.113, no.196, pp.1–6, 2013.
- [21] H.T. Luo, "Optimization design of cascaded classifiers," *Proc. CVPR*, vol.1, pp.480–485, 2005.
- [22] S.W. Lyu, "Infomax boosting," *Proc. CVPR*, vol.1, pp.533–538, 2005.
- [23] J. Sochman and J. Matas, "Waldboost-learning for time constrained sequential detection," *Proc. CVPR*, vol.2, pp.150–156, 2005.
- [24] J. Sun, J.M. Rehg, and A.F. Bobick, "Automatic cascade training with perturbation bias," *Proc. CVPR*, vol.2, pp.276–283, 2004.
- [25] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.7, pp.971–987, 2002.
- [26] OpenCV documentation, <http://docs.opencv.org/>
- [27] C. Chao, Q. Chen, and X. Sui, "Range limited bi-histogram equalization for image contrast enhancement," *Optik*, vol.124, pp.425–431, 2013.
- [28] R. Gonzalez and R. Woods, *Digital image processing*, 2nd ed., Prentice-Hall, Upper Saddle River, New Jersey, 2002.
- [29] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B.T.H. Romeny, and J.B. Zimmerman, "Adaptive histogram equalization and its variations," *Comput. Vis. Graph. Image Process.*, vol.39, no.3, pp.355–368, 1987.
- [30] J.B. Zimmerman, S.B. Cousins, K.M. Hartzell, M.E. Frisse, M.G. Kahn, and H. Nagahashi, "A psychophysical comparison of two methods for adaptive histogram equalization," *J. Digital Imaging*, vol.2, pp.82–91, 1989.
- [31] Y.T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consum. Electron.*, vol.43, pp.1–8, 1997.
- [32] V. Struc, J. Zibert, and N. Pavesic, "Histogram remapping as a pre-processing step for robust face recognition," *WSEAS Trans. Information Science and Applications*, vol.6, pp.520–529, 2009.
- [33] X. Zou, J. Kittler, and K. Messer, "Illumination invariant face recognition: A survey," *IEEE Int'l. Conf. Biometrics: Theory, Applications, and Systems*, pp.1–8, 2007.
- [34] Y. Li, C. Wang, and X. Ao, "Illumination processing in face recognition," *Face Recognition*, M. Oravec (Ed.), ISBN: 978-953-307-060-5, InTech, 2010.
- [35] H. Han, S. Shan, X. Chen, and W. Gao, "A comparative study on illumination preprocessing in face recognition," *Pattern Recognit.*, vol.46, pp.1691–1699, 2013.
- [36] Y.K. Park, S.L. Park, and J.K. Kim, "Retinex method based on adaptive smoothing for illumination invariant face recognition," *Signal Process.*, vol.88, no.8, pp.1929–1945, 2008.
- [37] W. Chen, M.J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Syst., Man, Cybern., B*, vol.36, no.2, pp.458–466, 2006.
- [38] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol.19, no.6, pp.1635–1650, 2010.
- [39] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on Weber's law with application to face recognition," *IEEE Signal Process. Lett.*, vol.18, no.8, pp.462–465, 2011.
- [40] T. Zhang, Y.Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Trans. Image Process.*, vol.18, no.11, pp.2599–2606, 2009.
- [41] S. Du and R. Ward, "Wavelet-based illumination normalization for face recognition," *IEEE Int'l. Conf. Image Process.*, vol.2, 2005.
- [42] G. An and J.W.Q. Ruan, "An illumination normalization model for face recognition under varied lighting conditions," *Pattern Recognit. Lett.*, vol.31, no.9, pp.1056–1067, 2010.
- [43] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol.9, no.1, pp.62–66, 1979.
- [44] M. Leszczynski, "Image preprocessing for illumination invariant face verification," *J. Telecommunications and Information Technology*, vol.nr4, pp.19–25, 2010.
- [45] A.S. Georgiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.23, no.6, pp.643–660, 2001.
- [46] Face detection homepage, <http://www.facedetection.com>
- [47] M. Yao and H. Nagahashi, "Analysis of noteworthy issues in illumination processing for face recognition," Submitted.

Min Yao received B.E. from Southeast University, China in 2009 and M.E. from Tokyo Institute of Technology, Japan in 2011, respectively. At present, she is a Ph.D. candidate in Department of Information Processing, Tokyo Institute of Technology. Her research interests include pattern recognition, image processing, and face detection & recognition in difficult conditions.

Hiroshi Nagahashi received his B.S. and Dr.Eng. degrees from Tokyo Institute of Technology in 1975 and 1980, respectively. Since 1990, he has been with Imaging Science and Engineering Laboratory, Tokyo Institute of Technology, where he is currently a professor. His research interests include pattern recognition, computer graphics, image processing, and computer vision.

Kota Aoki received B.E. and Ph.D. degrees from Tokyo Institute of Technology in 2001 and in 2006, respectively. He is currently an assistant professor in Imaging Science and Engineering Laboratory at Tokyo Institute of Technology. His research interests include computer vision, image processing, and statistical modeling.