Face Image Lighting Enhancement Using a 3D Model

Qiulin Chen
Purdue University
chen2114@purdue.edu

Jan P. Allebach
Purdue University
allebach@purdue.edu

Abstract

Image enhancement helps to generate balanced lighting distributions over faces. Our goal is to get an illuminance-balanced enhanced face image from a single view. Traditionally, image enhancement methods ignore the 3D geometry of the face or require a complicated multi-view geometry. Other methods cause color tone shifting or over saturation. Inspired by the new research achievements in face alignment and face 3D modeling, we propose an improved face image enhancement method by leveraging 3D face models. Given a face image as input, our method will first estimate its lighting distribution. Then we build an optimization process to refine the distribution. Finally, we generate an illuminance-balanced face image from a single view. Experiments on the FiveK dataset [5] demonstrate that our method performs well and compares favorably with other methods.

1. Introduction

Illumination imbalance causes a lot of problems. Performance of models working on different tasks, such as face recognition, skin tone detection, and emotion recognition are greatly affected by imbalanced illumination. A lot of effort has been devoted to processing the photos taken under different illumination environments. Once a model can precisely describe the geometric relationship between the human face and the illumination environment, powerful portrait enhancement becomes applicable by manipulating the light distribution. Image enhancement can help to balance the illumination distribution over face images.

Current enhancement methods can be divided into these five categories: 1) Histogram-based, 2) S-curve based, 3) Ratio-imaging-based, 4) Fusion-based, and 5) Learning-based methods. Histogram-based methods manipulate the histogram of the image directly to increase the image contrast, but fail when dealing with an underexposed image. S-curve methods apply pixel-level tone mapping to the original image. This increases the contrast in low lightness areas while compressing the value in high lightness areas. Bai and

Reibman [3] applied logarithmic mappings to the underexposure regions. They also built a model for underexposure and overexposure mapping detection. Ratio-imaging-based methods, to which our method belongs, separate the lighting distribution (shading) from the reflectance. These methods depend greatly on the accuracy of the lighting distribution estimation. Inaccurate estimation will result in ghosting artifacts, blurry restored images or uneven lightness. Fusion-based methods need images in multiple exposures and multiple s-curve mappings to fuse together. The final result is achieved by adjusting the illumination map and merging it with the reflectance. This will result in lower contrast and unnaturalness in the image. Learning-based methods perform well when dealing with small sets of exposure problems included in the training dataset, but they fail when dealing with images in the wild.

In this paper, we propose a new method to enhance face images. Inspired by the research on 3D face model and face alignment [4] [10] [34] [33], we enhance the face image by utilizing a 3D face model estimated from a single portrait. Our approach first estimates an initial lighting distribution of an input face image. Then we design an optimization process to further improve the lighting distribution. Based on our prior knowledge, two regularizers are chosen. They focus on local illumination and global illumination, respectively. To increase the computational efficiency, we convert our convex optimization process into a linear system.

In Section 2, we introduce related work on image decomposition, spherical harmonics, face alignment, and a 3D face morphable model. Then, we introduce our method in Section 3. In this section, we describe our method in three parts. In Section 4, we compare our model with others, both qualitatively and quantitatively. The experimental results demonstrate the effectiveness of our model for face image enhancement.

The major contributions of this work are:

- We apply a 3D face model to face image enhancement, which is rarely used in traditional face image enhancement methods.

- We design a loss function with two regularizers based

on prior knowledge about the face lighting distribution. These two regularizers greatly improve the face enhancement results.

- To simplify the optimization process, we convert our convex optimization process into a linear system.

- Our model performs well and the experimental results show that our model compares favorable with other methods.

2. Related Work

A lot of research has been done on face image enhancement. Image enhancement can help to improve the performance of models for different tasks such as face detection, face recognition, color constancy, and skin-tone detection. The method we use involves a lot of research from different areas including image enhancement, computer graphics, face alignment, and a 3D morphable model. The related work is introduced in the following order: 1) image decomposition, 2) spherical harmonics, 3) underexposed and overexposed image enhancement, 4) face alignment and a 3D face morphable model.

2.1. Image Decomposition

In computer graphics, researchers often treat an image as the multiplication between reflectance and irradiance [27]. The irradiance is also known as the lighting distribution or shading. It describes how the lightness strength is distributed over the surface of the objects in the image.

$$I_p = S_p R_p + C_p \tag{1}$$

In the above equation, Gross et al. [14] decomposed an image into three parts: shading (lighting distribution) S_p, reflectance R_p, and a specular term C_p. $S_p R_p$ represents the diffuse reflection, while C_p represents specular reflection. p represents the spatial coordinates. We decompose the images in the same way, except that we don't consider specular reflection independently.

2.2. Spherical Harmonics

Spherical harmonic functions provide a basis in frequency space to approximate functions distributed over spherical surfaces. For an arbitrary direction vector $\vec{\omega}$, which is represented in a spherical coordinate system as in (2), the spherical harmonic basis functions [17] are defined in (3)-(8).

$$\vec{\omega} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{2}$$

$$y_l^m(\theta, \phi) = \begin{cases} \sqrt{2}K_l^m \cos(m\phi)P_l^m(\cos\phi) & \text{if } m > 0, \\ K_l^0 P_l^0(\cos\theta) & \text{if } m = 0, \\ \sqrt{2}K_l^m \sin(-m\phi)P_l^{-m}(\cos\theta) & \text{if } m < 0. \end{cases} \tag{3}$$

$$K_l^m = \sqrt{\frac{(2l+1)}{(4\pi)}\frac{(l-|m|!)}{(l+|m|!)}} \tag{4}$$

$$(P_0^0)_p = 1, \tag{5}$$

$$(P_m^m)_p = (2m-1)!!\,(1-p^2)^{m/2}, \tag{6}$$

$$(P_{m+1}^m)_p = (p(2m+1)P_m^m)_p, \tag{7}$$

$$(P_l^m)_p = \left(\frac{p(2l-1)}{l-m}P_{l-1}^m - \frac{(l+m-1)}{l-m}P_{l-2}^m\right)_p \tag{8}$$

In (3), l is the band index and m represents the index in each band. m satisfies the constraint that $-l < m < l$. K_l^m are the normalization parameters while P_l^m are the Associated Legendre Polynomials (ALP) [9]. $!!$ represents a double factorial operation. A recurrence formula for ALP is provided from (5) to (8).

According to Ramamoorthi and Hanrahan [22], using the first nine basis terms is sufficient to approximate the irradiance map. The projection and combinations are calculated as in (9) and (10).

$$h_l^m = \int_{\Omega_{4\pi}} y_l^m(\vec{\omega})S_{sphere}(\vec{\omega})d\vec{\omega} \tag{9}$$

$$S'_{sphere}(\vec{\omega}) = \sum_{l=0}^{8}\sum_{m=-l}^{l} y_l^m(\vec{\omega})h_l^m \tag{10}$$

In (9), h_l^m represents the coefficient of the y_l^m basis when projecting the light distribution $S_{sphere}(\vec{\omega})$ to the basis functions. l represents the band index, while m represents the index in each band of spherical harmonics. ω represents the direction vector over the object surface. After getting all nine coefficients, we get the approximated light distribution $S'_{sphere}(\vec{\omega})$ over the object surface. We use a low-dimensional semidefinite programming (SDP) method to calculate the nine coefficients.

2.3. Face alignment and 3D face morphable model

In (10), we introduced a method to compute the illumination over a spherical surface. The basis functions are functions of the direction vector $\vec{\omega}$. Usually, multi-view images of a single face are needed to precisely deduce the face normal. The 3D data collection process is pretty challenging. A lot of datasets were built in this way [16] [26] [8] [18] [13]. However, in real applications, multi-view photographs of a single person with fixed poses are usually not available. The

face normal needs to be estimated from a single view. Face alignment and a 3D morphable model help to work around this problem [4] [10]. Based on 3DMM, Zhu et al. [34] designed their 3D dense face alignment (3DDFA) model to tackle the challenges in face alignment. We leverage their idea to extract face normals. Then, we use the face normals to calculate the lighting distributions as is shown in (10).

3. Face Image Enhancement Using Face Normal

Previously, we discussed the superiority of the spherical harmonics (SH) method in estimating the lighting distribution. We want to take advantage of this method to solve our problem focusing on face image enhancement. Our method can be divided into two parts. First, we use the spherical harmonics method to estimate an initial lighting distribution of the human face. This distribution is coarse and needs further refinement. Then we build our own quadratic loss function with two extra regularizers to optimize the lighting distribution.

3.1. Initialization of the lighting distribution

There are many different ways to choose an initialization of the lighting distribution. As is shown in [32], the maximum value across the red, green and blue (RGB) channels can be used as an initialization of the shading. However, directly using this to restore the image will result in a low contrast and unnatural image. Using the maximum value across the RGB channels will still contain a lot of high frequency components that are not only from lighting unevenness but also from reflectance. The enhanced image will have a low contrast problem. Instead, we use the estimation from the spherical harmonics (SH) as an initialization. The lighting distribution estimated from our method will be very smooth, and only reflect the variations in illumination.

3.2. Refine the lighting distribution

In order to refine the initial lighting distribution, we leverage the ideas from Xu et al. [29]. Our loss function is quadratic, and we can solve the optimization process by converting it into a linear system. Then, we can easily acquire a closed-form solution for it. This greatly reduces the computational complexity. Equation (11) shows that our loss function is composed of three parts. The first part comes from the original estimated lighting distribution. We want the global lightness order for both overexposure and underexposure regions to be similar to the original image. The second part controls the texture of the lighting distribution. We want the small texture components in the lighting distribution to be consistent with the original input image. This will help to reduce the artifacts in lighting unevenness and maintain the contrast level in the results. The third

part controls the global illumination uniformity. We want to make the lightness order between overexposure and underexposure regions closer to each other. By doing this, we will get an enhanced face image with a more uniform lighting distribution. To achieve this, we utilize the face masks from the previous 3D morphable model and the overexposure and underexposure masks achieved by using the model from Bai et al. [3]. This will help us to identify the overexposure and underexposure regions.

$$s' = \arg \min_s L(s) = L_{base}(s) + \lambda_g L_g(s) + \lambda_u L_u(s) \quad (11)$$

In (11), s and s' are the original and optimized lighting distributions. $L_{base}(s)$ represents the constraint on the lighting distribution s for lightness order similarity. $L_g(s)$ and $L_u(s)$ represent the constraints for local and global illumination uniformity in lighting distributions, respectively.

Lightness order similarity. The original images have overexposure and underexposure regions. During image enhancement, the lightness of underexposure regions may be mapped to be close to that of the overexposure regions to achieve a more uniformly illuminated face image. But the lightness order between the overexposure and underexposure regions should not be flipped. Violating this constraint may result in unnaturalness in the enhanced image. Equation (12) shows the details of the first term. Constraining the refined lighting distribution map s' to be close to the original estimation will help retain the original lightness order. Furthermore, the values of the original estimated lighting distribution are truncated to ensure that when we restore the enhanced image from the lighting distribution according to $I_{enhanced} = I_{origin}/s'$, the values will not be out of range. This constraint is similar to the color consistency constraint in [32]. But they use the maximum value across the RGB channels while we use the estimated lighting distribution from SH to be the initialization s'. The advantage of our method is described previously.

$$L_{base}(s) = \sum_p \left(s_p - s'_p \right)^2 \quad (12)$$

Local illumination uniformity. After refining the lighting distribution according to the first and third terms in (11), the lighting estimation is still very coarse. The gradients of the illumination channel of the original image are included inside the second term in the loss function (11). This is to ensure that the relative texture strength in the enhanced image is consistent with the input image. Equations (13) - (15) show details of the second term. There is a trade off in the estimated lighting distribution. If it contains too many texture components, the enhanced images will suffer from low contrast. If the distribution is too coarse with limited texture included, then the results may not achieve a uniform lighting distribution. Fine textures in the lighting distribution

can bring more flexibility to the local lightness adjustment. In practice, a certain amount of texture information can help to remove the shadows cast by distant lighting sources on the side of the captured scene.

$$L_g(s) = \sum_p \left(a_{x,p} \left(\frac{\partial s}{\partial x} \right)^2 + a_{y,p} \left(\frac{\partial s}{\partial y} \right)^2 \right) \quad (13)$$

$$a_{x,p} = \left(\left| \frac{\partial Y}{\partial x} \right|^\alpha + \epsilon \right)^{-1}, \quad (14)$$

$$a_{y,p} = \left(\left| \frac{\partial Y}{\partial y} \right|^\alpha + \epsilon \right)^{-1}, \quad (15)$$
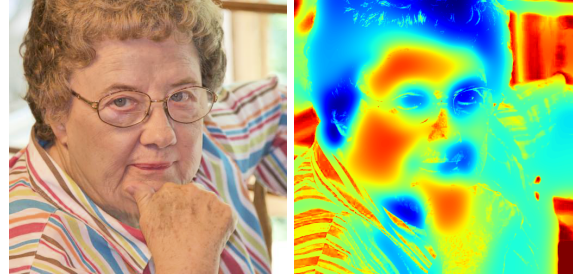
In (13), the gradients of the lighting distribution s are adjusted according to the gradients $\frac{\partial Y}{\partial x}$ and $\frac{\partial Y}{\partial y}$ of the luminance channels of the original image. Previously, a lot of research for skin detection worked successfully by using the orthogonal color space YC_bC_r [6, 7, 20, 24, 28]. The Y channel is an additive combination of the RGB components. So it preserves the high frequency image content. Thus, we choose the Y channel from the YC_bC_r color space as the illumination. λ_g in (11) is a parameter to control the weight of this term. p represents pixel locations. The hyperparameter α is used to control the sensitivity to the gradients. Large gradients in the original illumination channel will result in small reciprocal values in the weights $a_{x,p}$ and $a_{y,p}$. This will slack the constraints on the gradients in the lighting distribution. So it will generate relatively large gradients at some spatial locations in the lighting distribution. Then, due to the relationship $I_{enhanced} = I_{origin}/s'$, the locations in s' where the gradients are large can reduce the lightness nonuniformity in the original image. Thus, the enhanced image $I_{enhanced}$ will have an appearance much closer to that of photos taken under a uniform lighting environment. In Figure 1, we give an example.

Figure 1 shows the different enhanced images when the lighting distribution s contains different amounts of the texture information. Figure 1a is an enhanced image from a relatively smooth lighting distribution while Figure 1c is enhanced by retaining more texture in the lighting distribution. This is more obvious if we focus on the right eye regions of the subject in Figures 1b and 1d. Accordingly, the illumination around the right eye in the enhanced image in Figure 1c is more uniform than that in Figure 1a.

Global illumination uniformity. In some cases, when a subject's face is lighted by a distant lighting source on one side of the scene, one side of the face will have very low illumination values while the other side has very large ones. To balance the two regions, we need an extra constraint. This is a more coarse balancing compared to the previous local uniformity constraint. Equation (16) shows the details for this constraint.
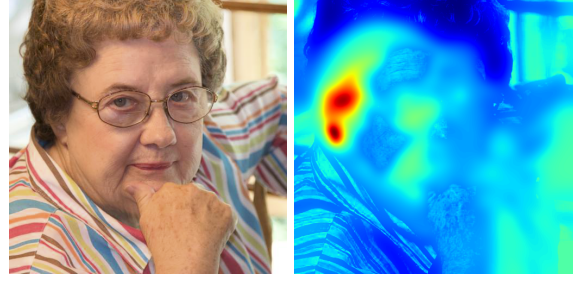
(a) Enhanced image from with less texture.

(b) Lighting distribution with less texture.

(c) Enhanced image from with more texture.

(d) Lighting distribution with more texture.

Figure 1. Comparisons between two enhanced images with different amounts of texture in the lighting distribution.

$$L_u(s) = \sum_p \left| \frac{M_u \odot I_p}{M_u \odot s_p} - \frac{M_o \odot I_p}{M_o \odot s_p} \right|^2 \quad (16)$$

In (16), L_u represents the global uniformity constraint in the loss function. The parameter λ_u controls its weight. $\odot$ represents element-wise multiplication. s represents the lighting distribution. I is the original input image. p represents the pixel coordinates in input image. M_u and M_o are masks for the underexposure and overexposure regions, respectively [3], [15]. This global illumination uniformity is to ensure that in the enhanced images, overexposure and underexposure regions will have exposure levels close to each other. The previous local illumination uniformity constraint will help to erase the shadows; but the differences between high exposure and low exposure regions will still exist. Our target is to make the lighting distributions within the overexposure masks be close to those within the underexposure masks. This is indicated by (17). So it is effectively equivalent if we take the reciprocal of the summand on both sides of the equation to get (18). This is equivalent to minimizing the loss function shown in (19). This approximation will simplify the optimization process.

$$\frac{1}{N} \sum_p \left(\frac{I \odot M_u}{s \odot M_u} \right) \approx \frac{1}{N} \sum_p \left(\frac{I \odot M_o}{s \odot M_o} \right) \quad (17)$$

$$\frac{1}{N} \sum_p \left(\frac{s \odot M_u}{I \odot M_u} \right) \approx \frac{1}{N} \sum_p \left(\frac{s \odot M_o}{I \odot M_o} \right) \quad (18)$$

$$L_u = \arg\min_s \frac{1}{N} \sum_p \left| \left(\frac{s \odot M_u}{I \odot M_u} \right) - \left(\frac{s \odot M_o}{I \odot M_o} \right) \right|^2 \quad (19)$$

For the overexposure and underexposure masks, we leverage the ideas from Bai and Reibman [3] and Guo et al. [15]. Usually, we can deploy a hard-thresholding method to extract the likely overexposure and underexposure regions. So the pixel values larger than 254 will be assigned to the overexposure region and those less than 6 to 10 will be assigned to the underexposure regions. But this method does not deal well with the transition cases. And this will cause artifacts around the region boundaries. So a soft classification model is invented. Instead of doing hard thresholding, Guo et al. [15] built a model to calculate the possibility for each pixel to be in overexposure class. They claimed that image pixels in the overexposure regions are desaturated and their lightness increases greatly. This results in an increase in L^* channel and a decrease in the a^* and b^* channels. Inspired by this idea, Bai and Reibman created a similar model by comparing the relative values in the L^*, a^* and b^* channels to assign likelihood to pixels. Based on their idea, we designed our overexposure and underexposure mask region detector. Equation (20) shows the details of our model.

$$\mathcal{L}(p) = \left(\frac{\tanh(\delta(127 + \alpha_T - ((G_\sigma * L^*)^2 + \|C\|_2))) + 1}{\tanh(\delta(127 - \alpha_T + ((G_\sigma * L^*)^2 - \|C\|_2))) + 1 + \epsilon} \right)_p \quad (20)$$

$$\|C\|_2 = (a^*)^2 + (b^*)^2 \quad (21)$$

$$M_u = \mathcal{L} >= 1 \quad (22)$$

$$M_o = \mathcal{L} < 1 \quad (23)$$

In (20), σ is a scale parameter to control the spread of the hyperbolic tangent function. G_σ is the Gaussian filter kernel. $*$ represents the convolution operation. α_T is the parameter that controls the weight assigned to the overexposure and underexposure regions. When α_T is larger, our model is more aggressive in assigning pixels to the underexposure class. When it is smaller, the overexposure class assignment will be more aggressive. ϵ is a small coefficient. Finally, we achieve our binary mask for overexposure and underexposure regions by comparing the relative likelihood strength with 1. If larger than 1, we recognize the current pixel as one from the underexposure region, otherwise it is

(a) Enhanced image from (b). (b) Lighting distribution with weak constraint.

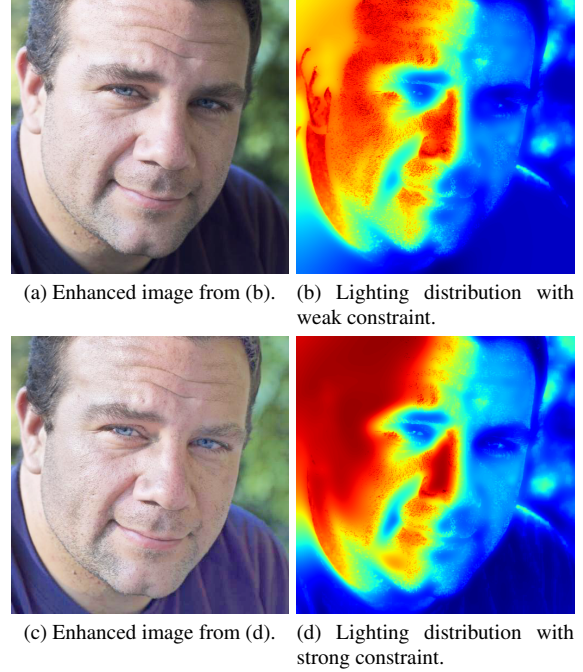(c) Enhanced image from (d). (d) Lighting distribution with strong constraint.

Figure 2. Comparisons between two enhanced images with different degrees of global illumination uniformity constraint in the lighting distribution, as controlled by the paramter λ_u in (11).

from the overexposure region. Empirically, we set the value of δ to be 1/60. For α_T, we set it to be 128.

To further clarify the functionality of the global illumination uniformity, we show two more examples in Figure 2. For the two enhanced images, by comparing the left and right half of the subject's face, we can see the lighting distribution on the subject's face is more uniform under the stronger constraint.

3.3. Optimization solver

Previously, we introduced the details in designing our loss function. We use approximations to both simplify the loss function and to achieve our goal to constrain the lighting distribution. Inspired by the Weighted Least Squares Method [11] [29] [32], our loss function is a convex function and we can easily acquire a closed form solution for it. Due to our simplification, we can convert the quadratic loss function into a linear system. The process to achieve the linear system equations is shown below.

$$\arg\min_s L(s) = \sum_p \left(s_p - s'_p\right)^2$$

$$+ \lambda_g \sum_p \left(a_{x,p}\left(\frac{\partial s}{\partial x}\right)_p^2 + a_{y,p}\left(\frac{\partial s}{\partial y}\right)_p^2\right)$$

$$+ \lambda_u \frac{1}{N}\sum_p \left|\left(\frac{s \odot M_u}{I \odot M_u}\right)_p - \left(\frac{s \odot M_o}{I \odot M_o}\right)_p\right|^2$$

$$(24)$$

$$a_{x,p} = \left(\left|\frac{\partial Y}{\partial x}(p)\right|^\alpha + \epsilon\right)^{-1}, \qquad (25)$$

$$a_{y,p} = \left(\left|\frac{\partial Y}{\partial y}(p)\right|^\alpha + \epsilon\right)^{-1}, \qquad (26)$$

$$(27)$$

$$\mathcal{L}(p) = \left(\frac{\tanh\left(\delta(127 + \alpha_T - ((G_\sigma * L^*)^2 + \|C\|_2)))\right) + 1}{\tanh\left(\delta(127 - \alpha_T + ((G_\sigma * L^*)^2 - \|C\|_2)))\right) + 1 + \epsilon}\right)$$

$$(28)$$

We have previously explained the variables in the above equations. Here, we just show how we achieve a simplified linear solver for this optimization problem. To simplify the expression, we use a high dimensional vector to represent the original image as $S \in R^{M \times N}$. M and N are the height and width, respectively, of the original input image. The loss function in this new format is shown in (29) and (30).

$$\arg\min_S L(S) = |S - S'|^2$$

$$+ \lambda_g\left(S^T D_x^T A_x D_x S + S^T D_y^T A_y D_y S\right) \quad (29)$$

$$+ \lambda_u S^T K^T K S$$

$$K = \left([I + \epsilon]_D^{-1}[M_u]_D - [I + \epsilon]_D^{-1}[M_o]_D\right) \qquad (30)$$

In (29), we show the loss function expressed in vector format. S' is the vector of the estimated lighting distribution. It is also used as the initialization value for S. The loss function is composed of a forward model and a two regularization terms weighted by λ_g and λ_u. We use a grid search method to find the values of these two weights. D_x and D_y are the Toeplitz matrices [25] for the forward difference. A_x and A_y are the inverse of the gradients of the original Y channel. They are diagonal matrices. Elements on the diagonal are shown in (25). $[\cdot]_D$ represents the diagonal form of the original matrix. ϵ is a small coefficient. I is the original input image.

The loss function (29) is a quadratic function in the variable S. So we can calculate the gradients of it in terms of

the variable S. We can convert the quadratic convex optimization problem to a linear system. So the problem is reformatted as in (31) - (33).

$$\frac{\partial L(S)}{\partial S} = (S - S')^T + \lambda_g\left(S^T D_x^T A_x D_x + S^T D_y^T A_y D_y\right)$$

$$+ \lambda_u S^T K^T K$$

$$= 0$$

$$(31)$$

The double coefficients are merged into the weight parameters λ_g and λ_u. A_x and A_y are diagonal matrices and they are symmetric.

$$\left(I_e + \lambda_g H^T + \lambda_u K^T K\right) S = S' \qquad (32)$$

$$H^T = D_x^T A_x D_x + D_y^T A_y D_y \qquad (33)$$

I_e is an identity matrix. $\left(I_e + \lambda_g H^T + \lambda_u K^T K\right)$ is the symmetric positive definite Laplacian matrix. The forward difference matrices are used to calculate the gradients along both x-axis and y-axis directions. And, by now, our loss function is in the format of $Ax = b$, a linear system, which is easy to solve. Also, we can use the sparsity of the matrices to accelerate the calculation process.

4. Compare with Other Methods

To measure the performance of our method, underexposed portrait images from the MIT-Adobe FiveK Dataset [5] are chosen to do a quantitative comparison between our method and other previous works. The thumbnails of the nine testing images are shown in Figure 5. We also choose one of these images for a qualitative comparison of the enhanced images.

4.1. Qualitative comparisons

First, we would like to compare our results with others qualitatively. The other methods that we choose for this comparison are contrast limited adaptive histogram equalization (CLAHE) [21], perceptually bidirectional similarity (PBS) [31], controllable image illumination enhancement (CIIE) [3], structure-revealing low-light (SRLL) enhancement [19], simultaneous reflectance and illumination estimation (SRIE) [12], low-light image enhancement using camera reponse model (LLCRM) [23], weighted adjustable histogram enhancement (WAHE) [2], and deep portrait relighting (DPR) [33]. For all these methods, we use their original code or pretrained model. All these methods are designed for image enhancement except that DPR is designed for face relighting. For the DPR method, we use $pyshtools$ to rotate the original illumination environment to be incident from the top-frontal direction of the subject. This ensures that the pretrained DPR network will receive the best
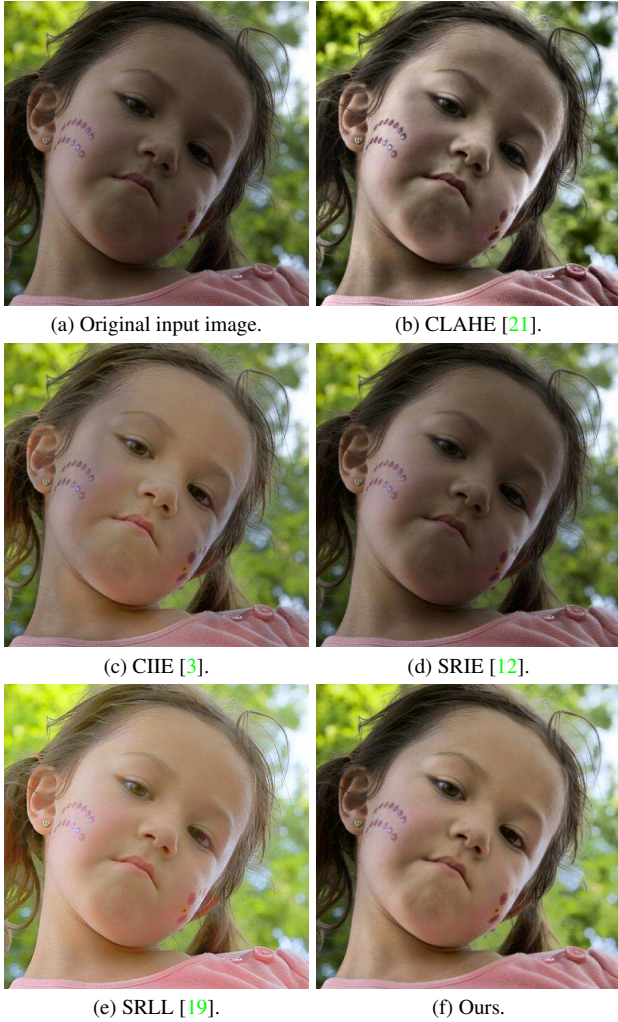
(a) Original input image.　　(b) CLAHE [21].

(c) CIIE [3].　　(d) SRIE [12].

(e) SRLL [19].　　(f) Ours.

Figure 3. Qualitative comparisons for different methods (CLAHE, CIIE, SRIE, SRLL, and ours).

(a) Original input image.　　(b) LLCRM [23].

(c) WAHE [2].　　(d) DPR [33].
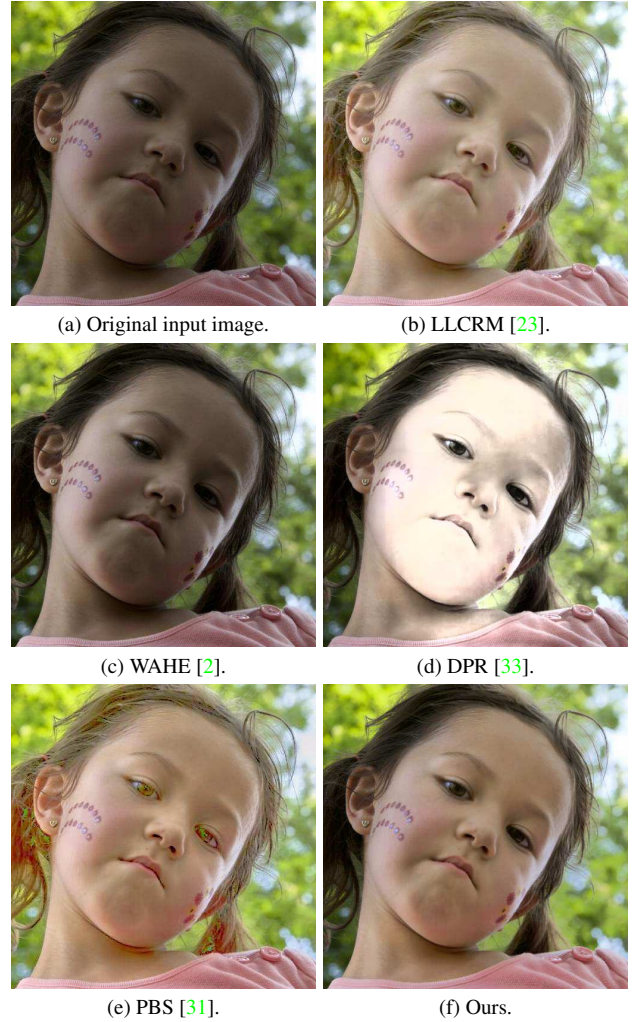
(e) PBS [31].　　(f) Ours.

Figure 4. Qualitative comparisons for different methods (LLCRM, WAHE, DPR, PBS, and ours).

illumination instead of an unbalanced illumination from a side direction. The qualitative results from these methods are shown in Figures 3 and 4. Our method compares favorably with all eight of the comparsion methods, some of which produce images that still look quite dark, or which have a very unnatural appearance.

4.2. Quantitative comparisons

In order to quantitatively compare our results with others, we choose several objective metrics to evaluate the results. Our model is good at balancing the lighting distribution and removing shadows. However, adjusting the relative lightness strength among different areas may risk reversing the lightness order. This may cause unnaturalness and artifacts. So one metric we choose is called the Lightness Order Metric (LOM) [3]. It measures the degree of unnaturalness of results by checking if the original lightness order

is retained. According to the original paper, smaller scores in the LOM indicate that the lightness order in original images are better retained. And the lighting distribution will be more natural. For measuring the contrast of images, we use a measure of enhancement (EME) [1]. For this metric, higher score values mean that the images are better enhanced with higher contrast. To quantify the enhancement of details and their visibility, we use the discrete entropy (DE) [30]. Higher DE values represent better enhancement in image details or visibility of images.

Based on our qualitative comparisons, it turns out that results from WAHE and SRIE are poorly enhanced images. The lighting distributions of the enhanced images are almost the same as those of the original inputs. And the underexposed regions are adjusted to be darker. Their enhancement targets are not in accordance with ours. And for DPR, the lighting distributions are always uneven and lots

Figure 5. Thumbnails of the nine testing images used to generate the results in Tables 1 - 3.

of artifacts are introduced. Based on this, we decided not to include them in the quantitative comparisons. Instead, we will include the remaining methods, which share similar enhancement targets and which yield relatively visually appealing results. So we choose to quantitatively compare our results with LLCRM, SRLL, CIIE, and PBS. The metrics we use are the three metrics we introduced previously. Comparison results using these three metrics are shown in Tables 1, 2, and 3. We choose nine images from the MIT-Adobe FiveK Dataset as examples. For each example, we use bold font to emphasize the method with best performance. For the Lightness Order Metric (LOM), our method gets better lightness naturalness compared with the other methods across all nine images. For the measure of enhancement (EME), our method also gets the highest score for each image. This implies that our enhanced images have good contrast. And for discrete entropy (DE), our method gets the highest average entropy. This shows that the our method enhances details well and increases the visibility of the poorly exposed images.

5. Conclusion

We proposed a pipeline to enhance poorly exposed face images. Our method can be separated into two steps: 1) estimating the original lighting distribution, and 2) optimizing that lighting distribution. To acquire an accurate lighting distribution, we leveraged ideas from 3D morphable models, face alignment, and face relighting to accurately deduce the face geometry from a single view. To further refine the lighting distribution from spherical harmonics, we design our own loss function. We also simplified the function to a convex quadratic form. To validate our method, we com-

Methods	LLCRM	SRLL	CIIE	PBS	Ours
Image 1	0.2141	0.2822	0.3577	0.1494	**0.1294**
Image 2	0.2500	0.3304	0.4022	0.2237	**0.1901**
Image 3	0.4888	0.3896	0.4078	0.2772	**0.0824**
Image 4	0.1867	0.5332	0.2405	0.1536	**0.1133**
Image 5	0.2338	0.2870	0.2652	0.1872	**0.1242**
Image 6	0.1420	0.2713	0.2825	0.1186	**0.0844**
Image 7	0.4319	0.3207	0.3884	0.2171	**0.0990**
Image 8	0.2149	0.2604	0.2821	0.1564	**0.0807**
Image 9	0.1415	0.4010	0.1936	0.1446	**0.1280**
Average	0.2560	0.3418	0.3133	0.1809	**0.1146**

Table 1. Comparison results using LOM metric [3] (Lower values are better). The images are shown in Fig. 5, and are numbered from left to right and top to bottom.

Methods	LLCRM	SRLL	CIIE	PBS	Ours
Image 1	6.3621	5.0363	12.7059	13.1127	**16.2816**
Image 2	10.5765	8.9197	13.4239	15.5792	**17.1697**
Image 3	42.1028	23.2063	60.6323	37.7957	**66.1502**
Image 4	9.0566	10.6031	10.2225	12.2546	**21.8791**
Image 5	7.5298	5.875	9.5969	10.5239	**11.0064**
Image 6	7.0793	5.3842	7.669	8.7756	**8.9927**
Image 7	13.243	9.0524	16.7294	18.6075	**20.6313**
Image 8	13.299	5.8689	25.8228	22.8722	**30.6223**
Image 9	9.6474	10.1326	11.1961	12.8121	**14.0283**
Average	13.2107	9.3421	18.6665	16.9259	**22.9735**

Table 2. Comparison results using EME metric [1] (Higher values are better). The images are shown in Fig. 5, and are numbered from left to right and top to bottom.

Methods	LLCRM	SRLL	CIIE	PBS	Ours
Image 1	5.1362	5.1586	4.9258	5.2606	**5.2641**
Image 2	5.0741	5.0806	4.8626	5.1994	**5.2537**
Image 3	5.0857	**5.1163**	4.8893	4.9514	4.8512
Image 4	4.9867	4.9619	5.134	5.1718	**5.3191**
Image 5	5.086	5.1171	5.0722	5.2042	**5.257**
Image 6	5.0839	5.0098	4.8257	5.1258	**5.2468**
Image 7	**5.0872**	5.0599	4.9484	5.0771	5.0088
Image 8	5.045	5.0529	4.8603	5.1529	**5.2047**
Image 9	4.8471	4.8901	4.9088	5.0302	**5.2268**
Average	5.0480	5.0497	4.9363	5.1304	**5.1814**

Table 3. Comparison results using DE metric [30] (Higher values are better). The images are shown in Fig. 5, and are numbered from left to right and top to bottom.

pared it both qualitatively and quantitatively with others. The comparison results show that our method performs well and compares favorably with other methods.

References

[1] S.S. Agaian, K. Panetta, and A.M. Grigoryan. Transform-based image enhancement algorithms with performance measure. *IEEE Transactions on Image Processing*, 10(3):367–382, 2001. 7, 8

[2] T Arici, S Dikbas, and Y Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE Transactions on Image Processing*, 18(9):1921–1935, 2009. 6, 7

[3] C. Bai and A. R. Reibman. Controllable image illumination enhancement with an over-enhancement measure. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pages 385–389, 2018. 1, 3, 4, 5, 6, 7, 8

[4] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3D faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, pages 187–194. ACM Press/Addison-Wesley Publishing Co, 1999. 1, 3

[5] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input / output image pairs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011. 1, 6

[6] Abbas Cheddad, Joan Condell, Kevin Curran, and Paul Mc Kevitt. A skin tone detection algorithm for an adaptive approach to steganography. *Signal Processing*, 89(12):2465–2478, 2009. 4

[7] Abhishek Choudhury, Marcus Rogers, Blair Gillam, and Keith Watson. A novel skin tone detection algorithm for contraband image analysis. In *2008 3rd International Workshop on Systematic Approaches to Digital Forensic Engineering*, pages 3–9, 2008. 4

[8] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pages 145 – 56, 2000. 2

[9] Shi-Hai Dong and R. Lemus. The overlap integral of three associated Legendre polynomials. *Applied Mathematics Letters*, 15(5):541–546, 2002. 2

[10] Bernhard Egger, William Smith, Ayush Tewari, Stefanie Wuhrer, Michael Zollhoefer, Thabo Beeler, Florian Bernard, Timo Bolkart, Adam Kortylewski, Sami Romdhani, Christian Theobalt, Volker Blanz, and Thomas Vetter. 3D morphable face models-past, present, and future. *ACM Transactions on Graphics*, 39(5):1–38, 2020. 1, 3

[11] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics*, 27(3):1–10, 2008. 5

[12] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2782–2790, 2016. 6, 7

[13] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-PIE. In *2008 IEEE International Conference on Automatic Face Gesture Recognition*, pages 1–8, 2008. 2

[14] R. Grosse, M.K. Johnson, E.H. Adelson, and W.T. Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2335 – 42, 2009. 2

[15] Dong Guo, Yuan Cheng, Shaojie Zhuo, and Terence Sim. Correcting over-exposure in photographs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 515–521, 2010. 4, 5

[16] Xiaofei He, Shuicheng Yan, Yuxiao Hu, P. Niyogi, and Hong-Jiang Zhang. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):328–340, 2005. 2

[17] Wojciech Jarosz. *Efficient Monte Carlo Methods for Light Transport in Scattering Media*. PhD dissertation, UC San Diego, 2008. 2

[18] H. A. Le and I. A. Kakadiaris. UHDB31: A dataset for better understanding face recognition across pose and illumination variation. In *IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 2555–2563, 2017. 2

[19] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018. 6, 7

[20] S.L. Phung, A. Bouzerdoum, and D. Chai. Skin segmentation using color pixel classification: analysis and comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):148–154, 2005. 4

[21] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3):355–368, 1987. 6, 7

[22] Ravi Ramamoorthi and Pat Hanrahan. An efficient representation for irradiance environment maps. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pages 497 – 500, 2001. 2

[23] Yurui Ren, Zhenqiang Ying, Thomas H. Li, and Ge Li. LECARM: Low-light image enhancement using the camera response model. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(4):968–981, 2019. 6, 7

[24] Khamar Basha Shaik, P. Ganesan, V. Kalist, B.S. Sathish, and J. Merlin Mary Jenitha. Comparative study of skin color detection and segmentation in HSV and YCrCb color space. *Procedia Computer Science*, 57:41–48, 2015. 4

[25] Gilbert Strang and Shev MacNamara. Functions of difference matrices are Toeplitz plus Hankel. *SIAM Review*, 56(3):525–546, 2014. 6

[26] Tiancheng Sun, J.T. Barron, Yun-Ta Tsai, Zexiang Xu, Xueming Yu, G. Fyffe, C. Rhemann, J. Busch, P. Debevec, and R. Ramamoorthi. Single image portrait relighting. *ACM Transactions on Graphics*, 38(4):79 (12 pp.), 2019. 2

[27] Zhen Wen, Zicheng Liu, and T.S. Huang. Face relighting with radiance environment maps. In *Proceedings of the*

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 2, pages 158 – 65, 2003. 2

[28] Kwok-Wai Wong, Kin-Man Lam, and Wan-Chi Siu. A robust scheme for live detection of human faces in color images. *Signal Processing: Image Communication*, 18(2):103–114, 2003. 4

[29] Li Xu, Qiong Yan, Yang Xia, and Jiaya Jia. Structure extraction from texture via relative total variation. *ACM Transactions on Graphics*, 31(6):139 (10 pp.), 2012. 3, 5

[30] Zhengmao Ye, Habib Mohamadian, and Yongmao Ye. Discrete entropy and relative entropy study on nonlinear clustering of underwater and aerial images. In *2007 IEEE International Conference on Control Applications*, pages 313–318, 2007. 7, 8

[31] Qing Zhang, Yongwei Nie, Lei Zhu, Chunxia Xiao, and Wei-Shi Zheng. Enhancing underexposed photos using perceptually bidirectional similarity. *IEEE Transactions on Multimedia*, 23:189 – 202, 2021. 6, 7

[32] Qing Zhang, Ganzhao Yuan, Chunxia Xiao, Lei Zhu, and Wei-Shi Zheng. High-quality exposure correction of underexposed photos. In *Proceedings of the 2018 ACM Multimedia Conference*, pages 582 – 590, 2018. 3, 5

[33] Hao Zhou, S. Hadap, K. Sunkavalli, and D. Jacobs. Deep single-image portrait relighting. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 7193 – 201, 2019. 1, 6, 7

[34] X. Zhu, X. Liu, Z. Lei, and S. Z. Li. Face alignment in full pose range: A 3D total solution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):78–92, 2019. 1, 3