# Face Image Illumination Processing Based on GAN with Dual Triplet Loss

Wei Ma[1], Xiaohua Xie[1,2,3(✉)], Jianhuang Lai[1,2,3], and Junyong Zhu[1,2,3]

[1] Sun Yat-sen University, Guangzhou, China
mawei23@mail2.sysu.edu.cn
{stsljh,xiexiaoh6,zhujuny5}@mail.sysu.edu.cn
[2] Guangdong Key Laboratory of Information Security Technology, Guangzhou, China
[3] Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, Beijing, China

**Abstract.** It is generally known that the illumination could seriously affect the performance of face analysis algorithms. Moreover, in most practical applications, the illumination is usually uncontrolled. A number of methods have been put forward to tackle the problem of illumination variations in face images, but they always only work on facial region and need to segment faces in advance. Furthermore, many illumination processing methods only demonstrate on grayscale images and require strict alignment of face images, resulting in limited applications in the real world. In this paper, we propose a face image illumination processing method based on the Generative Adversarial Network (GAN) with dual triplet loss. Through considering the inter-domain similarity and intra-domain difference between the generated images and the real images, we put forward the dual triplet loss. At the same time, we introduce the self-similarity constraint of the images in the target illumination field. Experiments on the CMU Multi-PIE face datasets demonstrate that the proposed method preserve the facial details well when relighting. The experiment of 3D face reconstruction also verifies the effectiveness of the proposed method.

**Keywords:** Face image · Illumination processing
Generative adversarial nets · Dual triplet loss

## 1 Introduction

Because of the great development of biometric recognition and machine learning, face analysis technologies, such as face detection, face recognition and 3D face reconstruction, have received great attention. Nowadays, in a highly constrained environment, many classical algorithms have been able to achieve nearly perfect performance. However, in the real world, the imaging environment in most applications is uncontrolled. For example, the user's posture or expression are not a neutral state, the illumination condition changes and so on. Compared


J.-H. Lai et al. (Eds.): PRCV 2018, LNCS 11258, pp. 150–161, 2018.
https://doi.org/10.1007/978-3-030-03338-5_13

with other interference factors, illumination has a greater impact on many face analysis algorithms. Therefore, the normalization of illumination is crucial for exploring the method of illumination invariant.

Over the years, a large number of methods on illumination invariance have been put forward. The invariant feature method is proposed to get the illumination invariant feature of images. Among them, Xie et al. [3] divided face images into large scale and small scale, and processed them separately. Recently, Wang et al. [4] proposed robust principal component analysis to eliminate the shadow produced by high-frequency features based on Xie's work. All these methods have achieved impressive results in the removal of soft shadows, but they are not effective in dealing with problems such as hard edge shadow caused by self occlusion. At the same time, these technologies can not be extended to color space, resulting in limited application in the real world.

With the development of 3D technology and deep learning, many researchers turn to use them to solve the illumination problems. Zhao et al. [5] propose a method for minimizing illumination difference by unlighting a 3D face texture via albedo estimation using lighting maps. Hold-Geoffroy et al. [6] trained a convolutional neural network to deduce the illumination parameters and reconstruct the illumination environment map. These methods are powerful and accurate. However, they are easily limited by data collection and unavoidable highly computing cost. In addition, most of the existing methods only focus on dealing with the carefully segmented face regions, which are not robust to the whole face images.

Inspired by the successful application of the Generative Adversarial Network in transfer learning [8] and domain adaptation [9], we propose to reformulate the face image illumination processing problem as a style translation task with a Generative Adversarial Network (GAN) in [10]. By using the circle reversible iterative scheme and via the multi-scale adversarial learning, we build the mapping from any complex illumination field to a target illumination field and its inverse mapping to effectively achieve the normalization of illumination without affecting any other non-illumination features of the image. In this paper, by analyzing the distance relationship between the generated image and the real image, an improved illumination processing method based on the dual triplet loss is proposed in order to better retain the details of the image and improve the quality of the generated image.

Overall, our contributions are as follows:

- We propose an improved illumination processing method based on Generative Adversarial Nets with dual triplet loss.
- We put forward the dual triplet loss through considering the inter-domain similarity and intra-domain difference between the generated images and the real images.
- We introduce the self-similarity constraint of the images in the target illumination field and add two image similarity indexes, SSIM and PSNR, to supplement the measure of similarity.

– We demonstrate that the proposed method can outperforms the state-of-the-arts realistic visualization results on non-strictly aligned color face images and eliminate the ill effects caused by illumination.

## 2 The Proposed Approach

### 2.1 Overall Network Framework

The overall network framework of our generative adversarial nets is shown in Fig. 1. The same as [10], our network consists of one generator and a pair of multi-scale discriminators with the same network structure but different classification constraint. We train $G$ to translate an input image $x$ under any lighting conditions into an expected lighting image $\tilde{x}'$ conditioned on the target illumination label $c'$, $G(x, c') \rightarrow \tilde{x}'$. And then reconstruct $\tilde{x}'$ to the input image conditioned on the original illumination label $c$ using the same $G$, $G(\tilde{x}', c) \rightarrow \tilde{x}$. The discriminator $D1$ distinguishes between the synthesized output images $\tilde{x}'$ and the real ones $x$, and classify the illumination category $\tilde{c}'$. The classification loss of real images used to optimize $D1$, and the fake images' used to optimize $G$. Similar but different, $D2$ distinguishes between $\tilde{x}'$ and a randomly selected picture $y'$ of maybe anybody's under target illumination condition and recognizes the identity $\tilde{l}'$ to optimize $G$ and $D2$.

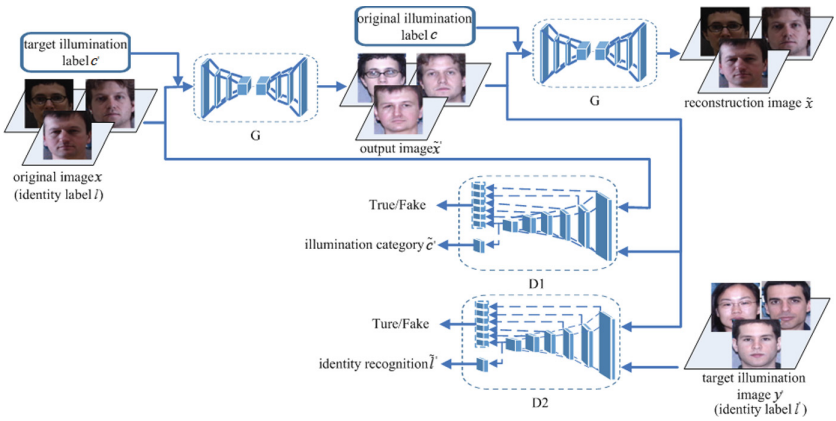



**Fig. 1.** Basic network architecture for face image illumination processing based on GAN with dual triplet loss.

### 2.2 Inter-domain Similarity and Intra-domain Difference

According to our research idea, face images under the same illumination conditions are divided into the same domain and our goal is to learn the mapping from any other illumination domain to the target illumination domain, which

refers the positive standard illumination in this paper. As shown in Fig. 2(a), the images before and after illumination normalization belong to different illumination domains, but their non-illumination information are same, which we call "inter-domain similarity". At the same time, the different images after normalization belong to the same illumination domain, but their non-illumination information are different, which we call "intra-domain difference".

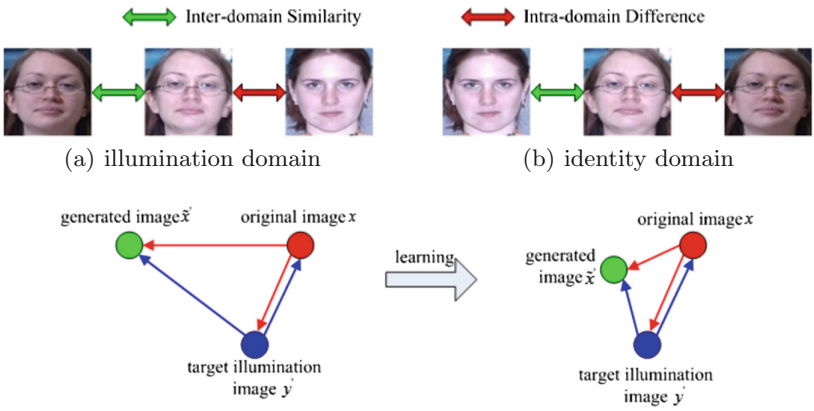



**Fig. 2.** Sketch of inter-domain similarity, intra-domain difference and the dual triplet loss.

Besides, as shown in Fig. 2(b). If we treat the non-illumination information as a symbol of the domain division, the two images before and after the normalization belong to the same identity domain, but their illumination information are different. That is, the two images have intra-domain difference now. Similarly, for any two different images after illumination normalization, they belong to different identity domains, but their illumination information are consistent. That is, the two images have inter-domain similarity now.

### 2.3 Dual Triplet Loss

Inspired by the thought of the triplet loss [11], we propose to construct a dual triplet loss based on the intra-domain difference and inter-domain similarity between the generated image and the real image. As is shown in Fig. 2(c).

The dual triplet loss include two triplet loss, each is composed of the original image $x$, the generated image $\tilde{x}'$ after illumination normalization and the real image $y'$ captured randomly from the target illumination domain. The first triplet loss takes $y'$ as anchor and takes $\tilde{x}'$ and $x$ as positive and negative sample respectively. The second triplet loss takes $x$ as anchor and takes $\tilde{x}'$ and $y'$ as positive and negative sample respectively.

Define $f(x)$, $f(\tilde{x}')$ and $f(y')$ are the features of $x$, $\tilde{x}'$ and $y'$ extracted from our multi-scale discriminant network. In the illumination domain, $x$ and $\tilde{x}'$ have inter-domain similarity. So the distance between them should be as small as possible and must be shorter than the distance between $y'$ and $x$. That is:

$$\|f(x) - f(\tilde{x}')\|_2^2 - \|f(x) - f(y')\|_2^2 < 0 \tag{1}$$

Similarly, in the identity domain, $\tilde{x}'$ and $y'$ have inter-domain similarity. So the distance between them should be as small as possible and must be shorter than the distance between $y'$ and $x$. That is:

$$\|f(y') - f(\tilde{x}')\|_2^2 - \|f(y') - f(x)\|_2^2 < 0 \tag{2}$$

In addition, $\tilde{x}'$ and $y'$ belong to the same illumination domain, but their non-illumination information are different. So, the distance between them should be larger than a minimum distance interval $\Delta_1$. That is:

$$\Delta_1 - \|f(y') - f(\tilde{x}')\|_2^2 < 0 \tag{3}$$

Similarly, in the identity domain, the distance between $\tilde{x}'$ and $x$ should be larger than a minimum distance interval $\Delta_2$. That is:

$$\Delta_2 - \|f(x) - f(\tilde{x}')\|_2^2 < 0 \tag{4}$$

In summary, the formula for calculating the loss function of dual triplet constraints is:

$$\begin{aligned} L_{dual-tri} &= \mathbb{E}[\|f(x) - f(\tilde{x}')\|_2^2 - \|f(x) - f(y')\|_2^2]_+ \\ &+ \mathbb{E}[\|f(y') - f(\tilde{x}')\|_2^2 - \|f(y') - f(x)\|_2^2]_+ \\ &+ \mathbb{E}[\Delta_1 - \|f(y') - f(\tilde{x}')\|_2^2]_+ + \mathbb{E}[\Delta_2 - \|f(x) - f(\tilde{x}')\|_2^2]_+ \end{aligned} \tag{5}$$

where $[\bullet]_+$ is a brief description of $max[\bullet, 0]$, which indicates that the loss is valid only when the result value of [] is greater than 0, otherwise it is recorded as 0. The threshold distance $\Delta_1$ is set as the minimum value of the feature distance between any two face images in the target illumination domain of the current training batch. Similarity, $\Delta_2$ is set to the minimum value of the distance between any two face images in the original identity domain.

### 2.4 Self-similarity Constraint and Reconstruction Loss

The ideal function of the generate network is transferring the input image to the target illumination and keeping the non-illumination information unchanged. Therefore, if we use any real image of target illumination domain as input, the generated image should be the same as the original, namely "self-similarity. Because the illumination scene of them are already the target illumination and don't need to be transferred.

Similar to the definition of the reconstruction loss in the previous article, we use the L1 distance to measure the error between the input and output image at first. The self-similarity constraint can be defined as

$$L_{rec-y'} = \mathbb{E}\|y' - G(y', c)\|_1 \tag{6}$$

L1 distance calculation is the sum of the absolute values of the corresponding pixel difference of all pixels between two images. The advantage is that it is convenient to calculate and can ignore the influence of the abnormal value in the image data, which is relatively stable and robust. But its disadvantage is also obvious, that is, the space between the pixels and their neighborhood is omitted, which may lead to the loss of high frequency information such as texture and detail. Based on the confirmation in [10], we use SSIM [12] and PSNR [13] to supplement the L1 distance in the image reconstruction constraint. Define:

$$\begin{aligned} L_{SSIM}(x_1, x_2) &= 1 - SSIM(x_1, x_2) \\ &= 1 - \frac{(2\mu_{x_1}\mu_{x_2} + c_1)(2\sigma_{x_1x_2} + c_2)}{(\mu_{x_1}^2 + \mu_{x_2}^2 + c_1)(\sigma_{x_1}^2 + \sigma_{x_2}^2 + c_2)} \end{aligned} \tag{7}$$

$$\begin{aligned} L_{PSNR}(x_1, x_2) &= 1 - \frac{PSNR(x_1, x_2)}{30} \\ &= 1 - \frac{1}{3}\log\frac{{MAX_x}^2}{MSE(x_1, x_2)} \end{aligned} \tag{8}$$

where $MAX_x$ is the maximum possible pixel value of the image. $MSE(x_1, x_2)$ is the mean squared error of $x_1$ and $x_2$. $\mu_{x_1}$, $\mu_{x_2}$, and $\sigma_{x_1}$, $\sigma_{x_2}$ are the average and variance of $x_1$ and $x_2$ respectively. $\sigma_{x_1x_2}$ is the covariance of $x_1$ and $x_2$. $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$ are two variables to stabilize the division with weak denominator, in which $L$ is the dynamic range of the pixel-values (1 in this paper). Special to note is that we use an empirical value of 30 to normalize the PSNR value.

Then the final cycle consistency loss of the generator can be written as

$$\begin{aligned} &L_{rec-all} = L_{rec-new} + \alpha_1 L_{rec-y'-new} \\ &= \mathbb{E}\|x - x_{rec}\|_1 + \alpha_2(L_{SSIM}(x, x_{rec}) + L_{PSNR}(x, x_{rec})) \\ &+ \alpha_1(L_{rec-y'} + \alpha_3(L_{SSIM}(y', G(y', c)) + L_{PSNR}(y', G(y', c))) \end{aligned} \tag{9}$$

We use $\alpha_2 = 0.5$, $\alpha_3 = 0.5$ and $\alpha_1 = 2$ in all of our experiments.

### 2.5 Loss Function

**Base Loss.** To stabilize the training process and generate higher quality images, we use Wasserstein GAN objective with gradient penalty as [8,10,14,15]. Define $\breve{x}_1$ and $\breve{x}_2$ are sampled uniformly along a straight line between a pair of real image and generated image, as well as a pair of target illumination image and

generated image. The discriminator network $D1$ and $D2$ update their parameters by minimizing the following loss:

$$L_{adv1} = \mathbb{E}[D1_{src}(x)] - \mathbb{E}[D1_{src}(G(x, c'))] - \lambda_{gp}\mathbb{E}[(\|\nabla_{\breve{x}_2} D1_{src}(\breve{x}_1)\|_2 - 1)^2] \tag{10}$$

$$L_{adv2} = \mathbb{E}[D2_{src}(y')] - \mathbb{E}[D2_{src}(G(x, c'))] - \lambda_{gp}\mathbb{E}[(\|\nabla_{\breve{x}_2} D2_{src}(\breve{x}_2)\|_2 - 1)^2] \tag{11}$$

where we use $\lambda_{gp} = 10$ for all experiments.

For an input image $x$ whose identity label is $l$ and a target illumination label $c'$, our goal is to translate $x$ into an output image $\tilde{x}'$, which is properly classified by $D1$ to $c'$ and recognized by $D2$ to $l$. The classification loss for illumination and identity classification task can be defined uniformly as

$$L_{cls1} = \mathbb{E}[logD1_{cls}(\hat{c}|\hat{x})] \tag{12}$$

$$L_{cls2} = \mathbb{E}[logD2_{cls}(\hat{c}|\hat{x})] \tag{13}$$

where $\hat{x}$ represents the image to be classified and the item $\hat{c}$ represents the proper label $\hat{x}$ should be in this classification task.

**Loss Function for Generator.** Define the illumination label and identity label of the synthesized output image as $\tilde{c'}$ and $\tilde{l'}$. So, the base objective functions to optimize $G$ can be written as

$$\begin{aligned} L_{G-base} &= L_{adv1}(x, G(x, c')) + L_{adv2}(y', G(x, c')) \\ &\quad + \alpha_4 L_{cls1}(\tilde{c'}, c) + \alpha_5 L_{cls2}(\tilde{l'}, l) \end{aligned} \tag{14}$$

where $\alpha_4$ and $\alpha_5$ are hyper-parameters that control the relative importance of illumination classification and identity recognition losses respectively, compared to the adversarial loss. We set $\alpha_4 = 1$ and $\alpha_5 = 1$. According to Eqs. (14, 9, 5), the overall objective functions to optimize $G$ can be written as

$$L_G = L_{G-base} + \alpha_6 L_{rec-all} + \alpha_7 L_{dual-tri} \tag{15}$$

The detailed description of all the individual loss functions was postpone above. We use $a_6 = 10$ and $a_7 = 10$ in all of our experiments.

**Loss Function for Discriminator.** The networks parameters of $D1$ and $D2$ can be optimized by minimizing a specifically designed adversarial loss $L_{adv1}$, $L_{adv2}$ and the aforementioned classification loss $L_{cls1}$, $L_{cls2}$ of the real one's respectively:

$$L_{D1} = -L_{adv1}(x, G(x, c')) + \alpha_8 L_{cls1}(\tilde{c'}, c) \tag{16}$$

$$L_{D2} = -L_{adv2}(y', G(x, c')) + \alpha_9 L_{cls2}(\tilde{l'}, l') \tag{17}$$

we set $a_8$ and $a_9$ as 1 in our experiments.

### 2.6 Model Training

We summarize the details of our algorithm training procedure in Algorithm 1. And we use the same history updating strategy as [10]. Moreover, we set $K_d = 5$, $K_g = 1$, $T = 1000$ and $lr_G = lr_D = 0.0001$ in the first 500 iterations, which both decay to 0 linearly in the following iterations.

**Algorithm 1.** Face Image Illumination Processing Based on the Dual Triplet Loss

**Input**: Real images $x$, identity label $l$, illumination label $c$ and target illumination label $c'$. Images with target illumination $y'$, identity label $l'$. Max number of steps $T$, number of the two discriminator network update per step $k_d$, number of generative network updates per step $K_g$, the learning rate of $lr_G$ and $lr_D$.

**Output**: The network parameters

```
for i = 1 : T do
    for k = 1 : k_d do
        Sample a batch of real images x and target illumination images y';
        Get G(x, c') with current network;
        If the history buffer is not null, update the batch content with half a
        batch images sampling from the buffer;
        Update network parameters of D1 by taking a Adam step on batch loss
        L_D1 in Eq. (16);
        Update network parameters of D2 by taking a Adam step on batch loss
        L_D2 in Eq. (17);
        Sample half a batch images from the original G(x, c') and add to the
        history buffer.
    end
    for k = 1 : k_g do
        Sample a batch of real images x and target illumination images y';
        Get G(x, c') and G(y', c) with current network;
        Reconstruct G(G(x, c'), c) and update network parameters of G by
        taking a Adam step on batch loss L_G in Eq. (15)
    end
end
```

## 3 Experimental Results and Analysis

Experiments were conducted on the CMU Multi-PIE Face Database [1] to verify the effectiveness of the proposed methods. Notably, all the images in this dataset are color images, which is always a challenge on illumination normalization for traditional methods. In our experiments, we restrict our attention merely to the frontal face images with neutral expression. All images are simply aligned and resized to $128 \times 128$ pixels, among which the first 2000 pictures were used for test and the others used for training.

### 3.1 Comparisons of the Visual Quality with Other Methods

For convenience, we denote our previous base method in [10] as GAN-base and denote this paper's method as GAN-DTL. In Fig. 3, we compare the visual results of normalized images between the proposed GAN-DTL method, GAN-base method and two baseline algorithms: NPL-QI [17] and ITI [18]. Same as other traditional methods, these two baseline algorithms can only process gray images and require strict alignment of face images. However, even on gray images,

they don't work well. For example, the NPL-QI method can't handle the extreme illumination conditions such as the first group and the third group. There is a general loss of detail in face after processing of the ITI method. And these two methods are not effective in dealing with the self occlusion of nose in the second groups. In contrast, our GAN-DTL method and GAN-base method achieve the best normalization performance and preserve more facial details and almost all appearance information, such as the hairstyle and hair color. At the same time, our GAN-DTL method provides a higher visual quality of normalization results on all kinds of test images. Different skin colors were preserved closer to the original ones, especially obvious on the first group image. And the details of eyeglass frame and whiskers in the third group are preserved more perfect. The result indicated that the proposed GAN-DTL method can preserve the details of generated images better and improve the quality of generated images.

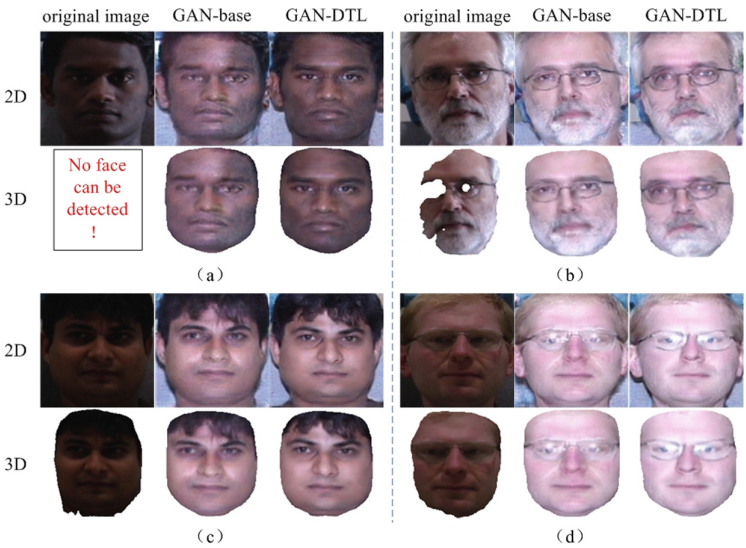
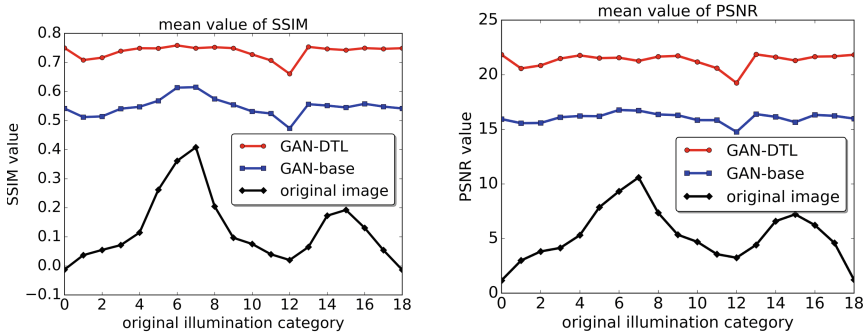



**Fig. 3.** Quantitative evaluation results comparison between the proposed GAN-DTL method, GAN-base method and two baseline algorithms.

### 3.2 Comparisons of the Ablation Study

We conduct ablation studies to show the superiority of our GAN-DTL method. We carry out the experiment on our 2000 test images. Take the face image of the same face under the target illumination as benchmark, we calculate the SSIM value and PSNR value of the original image, the generated image of GAN-base and the generated image of GAN-DTL respectively. And take the mean value according to the original illumination category then, which are drawn in black, blue and red curves in Fig. 4 respectively. As we can see, our GAN-DTL method

improves the evaluation results to a new height. The total average value of the SSIM is raised from 0.550 of the GAN-base method to 0.736 and the total average of the PSNR is raised from 16.048 to 21.324, which is consistent with the evaluation of the visual effect.




(a) (b)

**Fig. 4.** Comparisons of the ablation study SSIM and PSNR. (Color figure online)

### 3.3 Test of Face Algorithm Application

We use the online 3D face reconstruction from a single image algorithm [19] which is put forward by the team of nottingham university in 2017. As is shown in Fig. 5. As the initial 3D reconstructed image is not a positive angle of view, the angle and size of the pictures are slightly deviated when they are manually rotated to the front view. But it obviously does not affect the experimental comparison. In group (a), as the original image is in the dark light condition and the skin color of the face is black, the face can not be detected in the 3D reconstruction. In group (b), due to the uneven illumination of original images, the location of facial landmarking is not allowed, resulting in partial deletion of reconstructed 3D models. Similarly, in group (c) and group (d), the face region segmentation of the original image is inaccurate due to the influence of illumination on the location of facial landmarking, and the rough edge produced by the shadow in the chin area. However, in the four sets of images, the 3D model can be built very well and smoothly for the generated images after our GAN-DTL and GAN-base method illumination normalization. And our GAN-DTL method achieve the best results and illustrate the effectiveness of the proposed method in real-world applications.

## 4 Conclusion

In this paper, we propose a face image illumination processing method based on Generative Adversarial Nets with dual triplet loss. Through considering

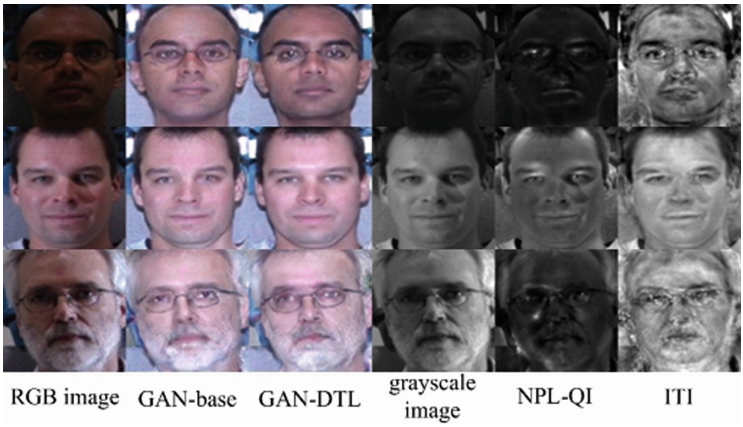



**Fig. 5.** 3D face reconstruction from a single image.

the inter-domain similarity and intra-domain difference between the generated images and the real images, we put forward the dual triplet loss. At the same time, we introduce the self-similarity constraint of the target illumination images and add two image similarity indexes, SSIM and PSNR, to supplement the measure of similarity. Experiments on the CMU Multi-PIE face datasets demonstrate that the proposed method preserve the details of generated images and improve the quality of generated images. The 3D face reconstruction experiment shows that the face images after our methods processing can eliminate the ill effects caused by illumination, and illustrates the effectiveness of the proposed methods in real-world applications.

**Acknowledgment.** This project is supported by the Natural Science Foundation of China (61672544, 61702566), Fundamental Research Funds for the Central Universities (No. 161gpy41), and the Tip-top Scientific and Technical Innovative Youth Talents of Guangdong special support program (No. 2016TQ03X263).

## References

1. Gross, R., et al.: Multi-pie. Image Vis. Comput. **28**(5), 807–813 (2010)
2. Adini, Y., Moses, Y., Ullman, S.: Face recognition: the problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 721–732 (1997)
3. Xie, X., et al.: Normalization of face illumination based on large-and small-scale features. IEEE Trans. Image Process. **20**(7), 1807–1821 (2011)
4. Wang, H., Ye, M., Yang, S.: Shadow compensation and illumination normalization of face image. Mach. Vis. Appl. **24**(6), 1121–1131 (2013)
5. Zhao, X., et al.: Minimizing illumination differences for 3D to 2D face recognition using lighting maps. IEEE Trans. Cybern. **44**(5), 725–736 (2014)
6. Hold-Geoffroy, Y., et al.: Deep outdoor illumination estimation. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, no. 2 (2017)

7. Wu, Z., Deng, W.: One-shot deep neural network for pose and illumination normalization face recognition. In: 2016 IEEE International Conference on Multimedia and Expo (ICME). IEEE (2016)
8. Choi, Y., Choi, M., Kim, M., et al.: StarGAN: unified generative adversarial networks for multi-domain image-to-image translation (2017)
9. Patel, V.M., et al.: Visual domain adaptation: a survey of recent advances. IEEE Sig. Process. Mag. **32**(3), 53–69 (2015)
10. Anonymous: Face image illumination processing based on generative adversarial nets. In: 24th International Conference on Pattern Recognition (ICPR) (2018)
11. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: a unified embedding for face recognition and clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
12. Wang, Z., et al.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
13. Hore, A., Ziou, D.: Image quality metrics: PSNR vs. SSIM. In: 2010 20th International Conference on Pattern recognition (ICPR). IEEE (2010)
14. Martin, A., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: International Conference on Machine Learning (2017)
15. Gulrajani, I., et al.: Improved training of Wasserstein GANs. Advances in Neural Information Processing Systems (2017)
16. Phillips, P.J., et al.: Overview of the face recognition grand challenge. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005. CVPR 2005, vol. 1. IEEE (2005)
17. Xie, X., et al.: Non-ideal class non-point light source quotient image for face relighting. Signal Process. **91**(4), 1048–1053 (2011)
18. Liu, J., et al.: Illumination transition image: parameter-based illumination estimation and re-rendering. In: 19th International Conference on Pattern Recognition, 2008. ICPR 2008. IEEE (2008)
19. Jackson, A.S., et al.: Large pose 3D face reconstruction from a single image via direct volumetric CNN regression. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE (2017)