Lighting-and Personal Characteristic-Aware Markov Random Field Model for Facial Image Relighting System

Ching-Ting Tu¹, Hwei-Jen Lin², and Chin-Yu Chang³
¹Dept. of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan
²Dept. of Computer Science and Information Engineering, Tamkang University, New Taipei City, Taiwan
³Dept. of Computer Science and Information Engineering, Tamkang University, New Taipei City, Taiwan

Corresponding authors: Ching-Ting Tu (e-mail: ctu@nchu.edu.tw) and Hwei-Jen Lin (e-mail: 086204@mail.tku.edu.tw)

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant MOST 109-2221-E-005-056-MY2.

ABSTRACT In this study, we proposed an example-based facial image relighting framework to relight an unseen input 2D facial image from a specified source lighting condition to a target lighting. Such a relighting framework is highly challenging, since the highlight or shadow areas in the relighted image usually follow the specific facial feature characteristics of the input image. When facial images have to be lighted in a specified lighting condition style followed a few provided pairwise lighting examples, the problem becomes even more complex. In contrast to existing learning-based relighting framework that directly predicts the intensity of the target image, we use a personal and lighting-specific transformation to formulate the appearance correlation between source and target lighting conditions. We propose a lighting-and personal characteristic-aware Markov Random Field Model to estimate the transformation parameter of the input subject, which is integrated with additional classifiers to determine input facial regions should be enhanced or maintained. Experimental results show that the proposed kernel facial relighting model can avoid overfitting problem and generates vivid and recognizable results despite the scarcity of training samples. Furthermore, the relighted results successfully simulate the individual lighting effects produced by the specific personal characteristics of the input image, such as the nose and cheek shadows. Finally, the effectiveness of the proposed framework is demonstrated using a robust face verification test for images taken under side-light conditions.

INDEX TERMS Face image synthesis, Markov random field, Adaboost

I. INTRODUCTION

Frontal lighting conditions are an essential requirement for most face-related applications, such as face identification, recognition, and animation. However, in many real-world situations, frontal lighting cannot be guaranteed. Accordingly, the problem of developing systems for the automatic synthesis of facial images under different lighting conditions has attracted growing interest in the field of computer vision and machine learning in recent years. Furthermore, in some entertainment applications, there is a need for headshot portraits taken under particular lighting conditions to present a certain lighting style. However, developing an automatic facial image relighting system capable of reproducing the unique light and shadow style of a particular lighting condition is particularly challenging.

Histogram matching approaches (such as methods proposed in [1, 17, 25]) are typical image relighting methods, where the objective is to convert the statistical intensity distribution of the input image to that of the (target) reference image. Such methods work well in natural photography, but do not produce image with a light effect depending on specific image features or structure. Some well-known image lighting frameworks have been proposed to analyze specific image content or features with the help of machine learning models to determine the relationship of the image appearance between the source and the target lighting conditions. For example, the tone mapping method proposed in [14] integrates with a principal component analysis (PCA) clustering algorithm to consider the semantic lighting factors of image transformation tasks; the high dynamic range image generation method in [12] applies a convolutional neural network (CNN) to map input images to target outputs. However, lighting factors tend to be more complex, for example, the image intensity distribution of different images lighting varies under the same light source.

In this study, we propose a facial relighting framework for relighting a 2D image of a given source lighting condition in a target light whose style is defined by some pairwise light examples from different subjects. In order to generate vivid
and recognizable results, the proposed framework meets the following constraints: (i) the personal characteristics of relighted images generated by the proposed framework should be close to the original input; (ii) The result must be visually similar to the target light conditions, such as following common shapes and positions of shadows and highlights, as indicated in the provided examples; (iii) The results of different subjects' lighting and shadow effects should have different characteristics, depending on their personal characteristics.

Unlike existing relighting frameworks, which predict target intensity directly, we propose a lighting transformation based on personal and light-specific parameters to map source images to the target results. Here, the transformation parameters of each input subject are estimated using a novel proposed lighting- and personal characteristic-aware MRF model that introduces two classifiers to determine whether the input facial region that must be enhanced or maintained and to select its on-demand references for parameters estimation.

The remainder of this paper is organized as follows. Section 1.A briefly reviews previous example-based relighting frameworks. Section 2 presents the formulation of the relighting transformation, and the proposed example-based face-relighting framework. Section 3 introduces the lighting and personal characteristic-aware Markov random field models to predict the personalized transformation parameters. Section 4 presents and discusses the experimental results. Section 5 provides some brief concluding remarks and indicates the major contributions of the present research.

A. RELATED WORK

There are several well-known learning-based models that have been successfully applied for image transformation tasks (the facial image relighting task is one of image transformation applications). In this section, we will discuss how different kernel models apply to face lighting tasks and their individual abilities to avoid overfitting problems.

1) KERNEL MODELS FOR RELIGHTING TRANSFORMATION

Broadly speaking, existing example-based facial image relighting frameworks can be broadly classified as either single-reference image-based approaches or multiple-reference image-based approaches. Examples of single reference approaches are 3D face model-based frameworks in [5, 26] and 2D image-based framework in [2, 4, 18]. These frameworks calculate the ratio of intensity between the input image and the reference image, and apply it to match the light and shadow distribution of the input to the reference image. The critical problem with this method is that it follows the unique characteristics of a reference image of a different subject, creating unrealistic and unexpected light effects.

On the other hand, some relighting systems proposed in [6, 21, 28] are leaned on based on multiple-references, namely training facial image datasets, where image appearance correlation between the source and target lighting conditions of these provided references is constructed by different kernel models. For example, kernel principal component analysis (kernel PCA) and Tensor PCA model are used in face relighting frameworks in [28] and [6], respectively, where the coordinates of these parametric-based model capture the significant facial and lighting features of provided reference samples, which are then used the feature for leaning the image appearance mapping between different lighting conditions.

On the other hand, authors in [21] apply the Markov Random field (MRF) model to construct the mapping of image appearance between different lighting conditions in non-parametric manner, where in general, MRF-based model can produce clearer but less variational relighting results.

Recently, neural network-based generative models, such as the generative adversarial network and the autoencoder network, have been widely used for image transformation tasks. Authors in [34] and [27] applied the conditional hourglass network and the conditional U-net model, respectively, to relight facial images from the source lighting-condition to the target condition. Asymmetric joint GAN model and RelightGAN are proposed in facial relighting frameworks [8] and [29], respectively, where with the help of the discriminator model, the generator model can generate relighting results with higher quality. Generally speaking, the NN-based approaches can generate relighting results with more diversity appearances and more realistic image quality, but it required more training examples to learn the model parameters.

2) SEMANTIC CONSTRAINT FOR IDENTITY PRESERVATION AND LIGHTING CONDITION

In order to make the results preserve the personal characteristics as the input instead of the same as the provided reference(s), some existing leaning-based relighting frameworks include the identity consistency constraint for avoiding the overfitting problem. For example, in the relighting frameworks [2] and [18], the input facial image is decomposed as multiple image scale components, then only parts of the image components, which are manually determined, are enhanced while others are retained to preserve the personal characteristics of the input subject. Facial relighting systems in [15, 16] proposed to calculate facial appearance maps or task-aware masks first, and then use to build the relighting process, i.e. determine different facial regions should be enhanced or preserved according to lighting and identity semantics.

For the NN-based frameworks, GAN-based relighting frameworks in [8, 29], similar to the well-known cycleGAN [35] and PIX2PIX [11] approaches, utilize the concept of the bi-direction transformation to include the additional identity and lighting semantic loss function to improve the relighting quality. In contrast, autoencoder-based relighting systems in [30, 34] included additional lighting condition labels to improve model stability; the facial relighting framework [9], similar to frameworks [2, 18], adopted multiple image scales
as features for training NN-model, where each image (scale) component concentrate specialized facial features.

3) PROPOSED LIGHTING-AND PERSONAL CHARACTERISTIC-AWARE MRF MODEL KERNEL APPROACH

Generally speaking, the performances of existing example-based kernel models for image transformation/synthesis tasks are highly dependent on the collected training reference samples. However, it is difficult to collect pairwise facial images that are properly lit in both the source lighting condition and the target lighting condition to train the transformation model in a supervised way. In this study, we proposed a lighting- and personal characteristic-aware MRF model for the lighting transformation task based on a limited size of training dataset. To avoid the overfitting problem, we do not estimate the intensity of target result directly, instead we included a simple linear relighting transformation for the relighting purpose, where the personalized parameters of relighting transformation are dynamically estimated for each input individual. This preserves the facial structure of the input because each input grayvalue is enhanced instead of being replaced by a pattern from another subject.

The proposed MRF models the correlation between input appearance and its corresponding transformation parameter in a patch-based manner, such correlation being constructed with multiple references in a non-parametric way. Especially the transformation parameters of the different regions of the face are not completely derived from the training reference; on the contrary, the parameters of some regions of the face are calculated directly from the appearance of the input. Such on-demand reference selection process for different facial regions are guided by their individual identity and lighting semantic meaning. Therefore, the relighting image of the face is visually similar to the appearance of the face in the target lighting condition, while maintaining the personal characteristics of the input subject. Accordingly, the relighting results provided a useful foundation for downstream face verification/recognition processes.

II. EXAMPLE-BASED FACE RELIGHTING FRAMEWORK

A. BAYESIAN FORMULATION FOR FACIAL IMAGE RELIGHTING

This study develops an example-based facial image relighting system capable of relighting the appearance of an input 2D facial image, \( I_{x_{\text{test}}} \) from a specified source lighting condition \( X \) to a target lighting condition \( Y \). \( I_{y_{\text{test}}} \). We model the mapping function form the source lighting condition \( X \) to the target condition \( Y \) at the pixel position \((u, v)\) by using a simple linear transformation by:

\[
\hat{I}_y(u, v) = W(u, v) \ast (I_x(u, v) - \hat{I}_x(u, v)) + \hat{I}_y(u, v),
\]

where \( W(u, v) = \frac{(I_x(u, v) - I_x(u, v)) \otimes G(\sigma)}{I_x(u, v) - I_x(u, v) \otimes G(\sigma)} \) \( (1) \)

In (1), the bias values, \( \hat{I}_x(u, v) \) and \( \hat{I}_y(u, v) \), are taken as the average component values within patch regions centered at coordinates \((u, v)\) in image \( I_x \) and image \( I_y \), respectively; the gain \( W(u, v) \) is defined as the ratio between the unbiased source component value and its corresponding target component value followed by processing a Gaussian filter.

Instead of predicting the target pixel value \( I_y(u, v) \) directly, our framework predicts the personalized transformation parameters for different input, i.e. the enhancement matric \( W(u, v) \) and \( \hat{I}_{y_{\text{test}}}(u, v) \) of all pixels, in this way, the facial structure of input source image \( I_{x_{\text{test}}} \) can be better preserved. We seek the optimal solution to maximize the posterior probability \( p(I_y|I_x) \) by personalized enhancement matric \( W^* \) and \( \hat{I}_y^* \), i.e.,

\[
\hat{I}_y^*(W^*, \hat{I}_y^*) = \arg \max_w (I_y | W, \hat{I}_y) p(I_y | W, \hat{I}_y) \quad (2)
\]

To carefully model (2), we follow the multiscale relighting framework in [18], and simply decompose a facial image \( l \) as its global component \( l^G \) and the local detailed component \( l^D \) facial model by:

\[
l^G = l \otimes G(\sigma) \quad \text{and} \quad l^D = l - l^G \quad (3)
\]

where \( G(\sigma) \) is the Gaussian blurring function with kernel size \( \sigma \). \( l^G \) is a smooth version of the original input capturing the overall global face shading effect, e.g., lighting or shading regions of an image distribution; the residual between \( l \) and the \( l^G \) is then taken as the local detailed component, \( l^D \), which indicates individual local detailed features, such as clear shadow edges or facial feature contours. We let the posterior probability \( p(I_y|I_x) \) in (2) as the energy function \( E(I_y | I_x, M) \), and the solution strategy is naturally divided into these two components:

\[
\hat{I}_y^*(W^*, \hat{I}_y^*) = \arg \max_{\hat{l}^G, \hat{l}^D} E(I_y | \hat{l}^G, \hat{l}^D) \mid I_x, M^G \quad + E(I_y | \hat{l}^D, \hat{l}^D) \mid I_x, M^D \quad (4)
\]

where two maps, a global map \( M^G \) and a local detailed map \( M^D \), are included to indicate the identity and lighting condition of different facial regions.

B. PROPOSED RELIGHTING FRAMEWORK

As shown in Figure 1, to solve (4), the proposed relighting
framework decomposes the unseen input facial image, \( I_{x,\text{test}} \), under a particular source lighting condition, \( X \), as its global lighting distribution component \( I_{L,\text{test}}^G \) and the local detailed feature component \( I_{D,\text{test}}^D \). Consequently, two adaptive maps, \( M_{G,\text{test}} \) and \( M_{D,\text{test}} \), of the input \( I_{x,\text{test}} \) are detected by two classifiers, \( H_L^c(\cdot) \) and \( H_{ID}^c(\cdot) \), to expose his/her lighting-aware or personal characteristics-aware regions when changing lighting condition from \( X \) to \( Y \). We proposed lighting-and-personal characteristic-aware Markov Random Field models, \( M^G \) - MRF\textsuperscript{G/D} and \( M^D \) - MRF\textsuperscript{D/D}, for predicting the transformation parameter matrix in (1), \( (W_{G,\text{test}}^G, I_{G,\text{test}}^G) \) and \( (W_{D,\text{test}}^D, I_{D,\text{test}}^D) \), for the global and local components, respectively. In particular, in the proposed \( M^{G/D} - \text{MRF}^{G/D} \) model, personalized \( M_{G,\text{test}}^c \) maps capture the facial semantic properties of different facial regions and guide the process of selecting on-demeane references to predict parameters for each region. The final structural map \( W_{G,\text{test}}^G \) and \( I_{G,\text{test}}^G \) are applied to adjust the input source components \( I_{G,\text{test}}^G \) and \( I_{D,\text{test}}^D \) to generate corresponding target results \( I_{G,\text{test}}^T \) and \( I_{D,\text{test}}^T \), and then merged into the final output \( I_{y,\text{test}} \).

![FIGURE 1. Proposed system flowchart consisting of three major procedures; global and local facial component decomposition, extraction of lighting and personal characteristic-aware region \( N^a \) and \( N^b \) by using classifiers \( H_L^c(\cdot) \) and \( H_{ID}^c(\cdot) \) and lighting- and personal characteristic-aware MRF modeling for lighting enhancement matrix, \( (W_{test}^G, I_{test}^G) \), inference](image)

III. LIGHTING AND PERSONAL CHARACTERISTIC-AWARE MARKOV RANDOM FIELD MODELS

The present study uses the YaleB [7] database for the purposes of training classifiers, \( H_L^c(\cdot) \) and \( H_{ID}^c(\cdot) \), and also for construct the correlation between the input appearance and its corresponding personalized parameters in \( M^{G/D} - \text{MRF}^{G/D} \). In particular, \( N \) pairwise samples from \( N \) different subjects are collected in the dataset, i.e., \( DB = \{(x_i, y_i)\}_{i=1}^N \) where \( x_i \) is the \( i \)'th training image sample under the target lighting condition \( X \), while \( y_i \) is the corresponding facial image of the same individual under a particular target lighting condition, \( Y \). The global and local component datasets, \( DB^G = \{(I_{x,i}, I_{y,i}^G(W_{i}^G, I_{G,i}^G))\}_{i=1}^N \) and \( DB^D = \{(I_{x,i}^D, I_{y,i}^D(W_{i}^D, I_{D,i}^D))\}_{i=1}^N \), are generated for modeling (4), and each subject’s \( I_{y,i}^G \) can be transformed by its personalized parameter matrix \( (W_{i}^G, I_{G,i}^G) \) (formulated as (1)).

A. LIGHTING- AND PERSONAL CHARACTERISTIC-AWARE REGION DETECTORS FOR MRF GRAPH ADAPTION

1) IDENTITY- AND LIGHTING-AWARE REGION DETECTORS

We train the identification and lighting condition classifiers \( H_L^c(\cdot) \) and \( H_{ID}^c(\cdot) \) by using Adaboost algorithm. There are three main training processes, namely the generation of the feature pool where possible facial appearance patterns in target conditions are collected, the weak classification construction that selects important feature from the pool for classification tasks, and the weak classification ensemble processes where the strong \( H_L^c(\cdot) \) and \( H_{ID}^c(\cdot) \) classifications are formed by the selected weak classifiers.

In contrast to the lighting condition classification framework proposed in [3], in which several lighting-sensitive facial regions and lighting contrast patterns are manually specified, we define the feature pattern pool \( F^G \) in a data-driven manner for \( H_L^c(\cdot) \). The feature pool \( F^G \) is collected from \( N \) training images in target lighting conditions \( (I_{y,i}^G) \), each training image provides \( N^a \) facial patch patterns from different facial positions. From \( F^G \), the boost algorithm selects \( M \) significant features for lighting condition classification and builds \( M \) corresponding lighting weak classifiers. The weak classifier \( h_m^i(I^G; f_m^p, \theta_m^p) \) of the \( m \)'th feature \( f_m^p \), classifies the sample \( I^G \) by:

\[
h_m^i(I^G; f_m^p, \theta_m^p) = \begin{cases} 1 & |Val_m^G(I^G) \geq \theta_m^p \\ -1 & |Val_m^G(I^G) < 0 \end{cases}
\]

where \( Val_m^G(I^G) = \max_{p \in \text{loc}(f_m^p)} \frac{f_m^p}{\|f_m^p\|} \|I^G\| \|f_m^p\| \)

where \( Val_m^G(I^G) \) is the feature representation of \( I^G \) based on
the feature $f_m^G$; loc($f_m^G$) is the patch location, where $f_m^G$ is extracted from, i.e. loc($f_m^G$) is located at one of grids in $V_{x,x,test}^G$ in Fig. 2(a); $V_{x,x,test}^G(I^G)$ is calculated as the maximal correlation feature response of $I^G$ by applying $f_m^G$ to $I^G$ at neighborhood regions of N(loc($f_m^G$)); the threshold $\theta_m$ is leaned from the training dataset $DB^G$. These $M$ weak classifiers are the ensemble of the lighting classifier $H_L$ ( ), and the facial regions of these $M$ selected features, loc($H_L$ ( )), are located at lighting-aware facial regions.

The learning process of the boosting algorithm of $H_{ID}( )$, is similar to that of $H_L$( ). The feature pool $F^D$ contains patterns of local detailed component facial patches, which are from $N$ training samples $\{I^D_{x,y}\}$ at different facial positions. $M'$ patch patterns are selected from $F^D$ to classify the identity of the pairwise sample $(I^D_{x,y}, I^D_{x,y}')$ under two different lighting conditions. The weak identity classifier $h_m((I^D_{x,y}, I^D_{x,y}'); f_m, \theta_{m'})$ of the selected feature $f_m^D \in F^D$ is formulated as:

$$h_m((I^D_{x,y}, I^D_{x,y}'); f_m, \theta_{m'}) = \begin{cases} 1 & |Va(I^D_{x,y}, I^D_{x,y}')| \geq \theta_{m'} \\ -1 & |Va(I^D_{x,y}, I^D_{x,y}')| < 0 \end{cases},$$

where

$$Va(I^D_{x,y}, I^D_{x,y}') = \max_{p_i \in N(loc(I^D_{x,y}))} \left( \frac{f_m^D(I^D_{x,y}) f_m^D(I^D_{x,y}')}{\|f_m^D\| \|f_m^D\|'} \right),$$

Here the feature response value, $Va(I^D_{x,y}, I^D_{x,y}')$ is measured in two terms: the outer-max term evaluates the correlation between the feature pattern $f_m^D$ and the facial patch of $I^D_{x,y}$ cropped from the loc($f_m^D$) neighborhood regions with a maximum value of correlation, and the inner-max term evaluates the image patch similarity of this pairwise sample $(I^D_{x,y}, I^D_{x,y}')$. Note that, similar to (5), the similarity scores in (6) is evaluated within a search window $N(\cdot)$ to solve the problem of image alignment. The identity classifier $H_{ID}( )$ is the ensemble of $M'$ weak classifiers, and the $M'$ facial regions, loc($H_{ID}(\cdot)$), are discriminated for identity classification under two lighting conditions.

These two patch-based classifiers, $H_L( \cdot )$ and $H_{ID}( \cdot )$, are used to determine the facial regions of the $I_{x,test}^G$ input in different facial position, which should be improved to change the appearance of the image, or be preserved to maintain its individual characteristics. As shown in Figure 2(a), the lighting conditions of the $N_K$ facial region in a given test $I_{x,test}^G$ are determined by $H_L(\cdot)$ in a personalized map $M^G = \{M^G_{p1}, M^G_{p2}, ..., M^G_{pK}\}$. The $pi$'th map entry, $M_{p_i}^G$, of $I_{x,test}^G$, is determined as:

$$M_{p_i}^G = \begin{cases} 0 & p_i \in \text{loc}(H_L) \text{ and } h_{pi}(I_{p_i}^G) = 1 \\ 1 & \text{otherwise} \end{cases}$$

where if the $pi$'th facial region is located at one of $M$ selected lighting discrimination positions, $p_i \in \text{loc}(H_L(\cdot))$, the corresponding weak classifier $h_{pi}(\cdot)$ is applied to determine the lighting condition label of $I_{p_i}^G$. On the other hand, the map $M^D = \{M^{Dp1}, M^{Dp2}, ..., M^{DpK}\}$ of a given pairwise local detailed component, $(I_{x}^D, I_{x}'^D)$ is determined by $H_{ID}(\cdot)$. The $pi$'th map entry $M^{Dp_i}$ is denoted as:

$$M^{Dp_i} = \begin{cases} 0 & p_i \in \text{loc}(H_{ID}) \text{ and } h_{pi}(I_{p_i}^D) = 1 \\ 1 & \text{otherwise} \end{cases}$$

where, if the $pi$'th facial patch is located at one of selected identity discrimination positions, $p_i \in \text{loc}(H_{ID}(\cdot))$, the corresponding weak classifier $h_{pi}(\cdot)$ is applied to determine the identity consistency of the pairwise input at this region position.

According to (7) and (8), two types of facial regions are defined: (i) the fixed-type regions (corresponding to $M^{G/D}$ entries equal to 0): It means the corresponding facial region in the input image (or the initial result) is already classified as the appearance taken under lighting condition $Y$, thus the result patch can be calculated directly from the input for preserving the facial feature characteristics of the input subject, and (ii) the inference-type regions ($M^{G/D}$ entries equal to 1): as their name indicates, the appearance of these regions should be enhanced, and their personal-specific parameter matrices for relighting transformation are to be determined from the training refences.

$$\begin{bmatrix} I_x & I_y \\ I_x' & I_y' \end{bmatrix}$$

$\text{FIGURE 3.}$ Global and local detailed adaptive maps $M^G$ and $M^D$ of three testing sample subjects. The white and red dots are located at regions of loc($H_L(\cdot)$) and loc($H_{ID}(\cdot)$). (b) Retrieved subjects of an example $I_x$, using $M^G$ and $M^D$ as feature; (q1) and (q2) query image by using $M^G$ and $M^D$ as the feature representation, respectively; DB-1 and DB-2 correspond to the most similar and different subjects retrieved from the database. Note that the white dots in $M^G$ indicate the corresponding facial regions of input $I_x$ are similar to the appearance in the target lighting condition $Y$. On the other hand, white dots in $M^D$ indicate that the pairwise examples from different lighting condition at these regions are identity-consistent.
2) DISCUSSION

Fig. 3(a) shows three adaptive maps $M^G$ and $M^D$ of three different examples of images, white and red dots are located in the discriminatory feature regions, i.e. loc($H_x$) and loc($H_D$).

- Analysis of $M^G$: the entries corresponding to $M^G_{pi}=0$ (in (7)) are marked as the white dots, which means that the facial region of $I^G_{x,\text{test}}$ at this position is already similar to the facial appearances in the target lighting condition $Y$, even though $I^G_{x,\text{test}}$ is form the lighting condition $X$. Since the effects of lighting on such facial features are observable, only some regions of $I^G_{x,\text{test}}$ are miss-classified as from the lighting condition $Y$ (white dots).

- Analysis of $M^D$: the white dots of the adaptive maps $M^D$ indicate the locations of entries $M^D_{pi}=0$ (in (8)), it means the facial characteristics of $I^D_{x,\text{test}}$ and $I^D_{y,\text{test}}$ at the corresponding regions are consistent, although these two images are obtained under different lighting conditions.

- $M^G\!/$ $M^D$ are personalized maps: In Fig. 3(b), the maps $M^G_{x,\text{test}}$ and $M^D_{y,\text{test}}$ have been used as a feature representation to search the most similar and most different subjects from the database shown in (DB-1) and (DB-2), respectively. Generally, the $M^G$ focus on the facial brightness feature, so that the brightness of the query image (q1 in (b) and its closest image (DB-1 of q1 in (b)) are evenly dispersed, and the brightness of the most different subject (DB-2 of q1) is located on the lower and upper halves of the face. For the case of q2 depended on the $M^D$, the brow, nose, and chin characteristics of its most similar subject (DB-1 of q2) are more similar than those of the most different subject (DB-2 of q2).

B STRUCTURAL LIGHTING ENHANCEMENT MATRIX PREDICTION

We integrate recognized properties of different facial regions of the input image $I^G_{x,\text{test}}$ and guide the transformation parameter $W^{G/D}_{x,\text{test}}$ and $I^{G/D}_{y,\text{test}}$ estimation process predicted by a novel lighting-and personal characteristic-aware Markov Random Field Model, $M^G\!/$ $M^D$ – $MRF^{G/D}$. Different from most existing MRF-based image synthesis (transformation) systems [13, 22, 31, 33], where the MRF graph is usually static for all testing subjects, the recognized maps $M^G_{x,\text{test}}$ and $M^D_{y,\text{test}}$ change the static MRF graph into a personalized version, as shown in Fig. 2(b). Under the patch-based MRF formulation, the energy function $E(I^{G/D}_{y,\text{test}}|I^G_{x,\text{test}}, M)$ in (4) becomes:

\[
E(I^G_{y,\text{test}}|I^G_{x,\text{test}}, M) = \sum_{pi} \Phi(I^G_{y,\text{test}}(W^{pi}_{y,\text{test}}, I^G_{pi})|M^{pi}) + \sum_{(pi, pj)\in N} \psi(I^G_{y,\text{test}}(W^{pi}_{y,\text{test}}, I^G_{pi}), I^G_{y,\text{test}}(W^{pj}_{y,\text{test}}, I^G_{pj}))
\]

(9)

where the given detected $M^{pi}$ is used to determine if the $pi$’th facial region should be enhanced or not. The data term, $\phi(\ )$, corresponds to the likelihood probability of Bayesian formulation (2), which evaluates the correlation between the observation input patch $I^y_{pi}$ and its corresponding hidden parameter matrix ($W^{G/D,pi}$ and $I^{G/D,pi}$) to be estimated; the smoothness term $\psi(\ )$ corresponds to the prior probability in (2), which is used to evaluate the fitness score of the hidden parameter matrix of two adjacent patches ($W^{G/D,pi}$ and $I^{G/D,pi}$ and ($W^{G/D,pi}$ and $I^{G/D,pi}$), where the $pi$’th patch is located in the neighborhood of $pi$’th patch.

Specifically, the data term of the $pi$’th patch, $\phi(I^G_{pi}, I^G_{pi}(W^{pi}_{y,\text{test}}, I^G_{pi})|M^{pi})$ is formulated as:

\[
\phi(I^G_{pi}, I^G_{pi}(W^{pi}_{y,\text{test}}, I^G_{pi})|M^{pi}) = \delta(M^{pi} = \exp \left( - \frac{||I^G_{pi}(W^{pi}_{y,\text{test}}, I^G_{pi})||^2}{(\eta\sigma)^2} \right) + \delta(M^{pi} = 0)c, (10)
\]

where $\delta(M^{pi} = 1)$ is an indicator function; we intend to evaluate the candidate’s fitness score of the inference-type node in the first term, and for the second term for the fixed-type node; we let the constant $c$ equal to 1; $\eta$ is determined by cross-verification for normalization purposes; $\Phi_{x}(\ )$ is a feature mapping function that map a patch candidate $I^{G}_{pi}(W^{pi}_{y,\text{test}}, I^G_{pi})$, for example selected from $k$’th training sample, to its the corresponding patch in the source lighting condition $X$, $I^{G}_{x,k}$. The smoothness term $\psi(\ )$ uses to evaluate whether the parameter metrics of two adjacent nodes $pi$ and $pj$, are well overlapped:

\[
\psi(I^{pi}_{y,\text{test}}, I^{pi}_{y,\text{test}}(W^{pi}_{y,\text{test}}, I^{pi}_{y,\text{test}}), I^{pj}_{y,\text{test}}(W^{pj}_{y,\text{test}}, I^{pj}_{y,\text{test}})) = \exp \left( - \frac{\|od_{ij}(I^{pi}_{y,\text{test}}, I^{pj}_{y,\text{test}})) - od_{ij}(I^{pi}_{y,\text{test}}, I^{pj}_{y,\text{test}})\|}{(\eta\sigma)^2} \right) (11)
\]

where $od_{ij}(\ )$ is used to define the overlapped region of these two adjacent patches. Note that, the normalization terms $\eta$ and $\eta'$ in both (10) and (11) play important roles in the prediction result; please refer to APPENDIX.

1) NON-PARAMETRIC FACE MODELING AND INFERENCE PROCESS

Fig. 4 illustrates the process of proposed $M^G\!/$ $MRF^G$ graph model to predict the parameter matrix of the global component, $W^G$ and $I^G_y$, and the same processes can be applied to the local detailed component. The $N_{x}$ overlapped facial (observed) patches of $I_{x,\text{test}}=[I_{x,\text{test}}, I_{x,\text{test}}, \ldots, I_{x,\text{test}}]$ are corresponding to $N_{x}$ observation nodes of $M^G\!/$ $MRF^G$. Each observation patch, $I^{G,pi}_{x,\text{test}}$, has a corresponding hidden node as its hidden parameter matrix $W^{G,pi}_{x,\text{test}}$ and $I^{G,pi}_{y,\text{test}}$, to be determined. The MRF-inference stage to solve (9) contains
two strategies, the on-demand candidate selection for hidden nodes for face modeling, and the optimization process to select optimal result for the final enhancement matrix prediction.

For each inference-type region (see two orange illustrated type patches (five green nodes), the results are classified as the target lighting condition Y. Therefore, the hidden pattern $W_{G,pi}^{g,pi}$ is calculated directly as the corresponding patch $I_{k,pi}^{g,pi}$, so that the result intensity does not change.

Once each node has its on-demand candidate(s), the optimal result of structural matrix ($W_{D,D}^{G,pi}$, $W_{D,D}^{g,pi}$) are determined by using the belief propagation (BP) algorithm [22] to maximize (9) with constraints (10) and (11). The overall algorithm for the prediction processes of the lighting enhancement matrix, ($W_{D,D}^{G,pi}$, $W_{D,D}^{g,pi}$) by $M_{D,D}^{G,pi}$ model is provided in Table 1.

2) DISCUSSION
**M* \( \text{test}^D \) and \( M_{\text{test}}^G \):** For all lighting conditions, the discriminative areas (i.e., white and red dots) of \( M_{\text{test}}^G \) are mainly concentrated in the shadow areas of the face, where inference-type nodes (red areas) represent a greater percentage of the total number of nodes. That is, most relighted regions of \( I_{x,\text{test}}^G \) must be derived incorporated into training samples. On the contrary, the discriminating areas of \( M_{\text{test}}^D \) are concentrated mainly in the facial features (e.g., eyes or mouth), with about 50% of fixed-type (white) and inference-type (red) nodes. In other words, certain facial regions of result \( I_{y,\text{test}}^D \) are calculated directly from the input \( I_{x,\text{test}}^D \) to retain the personal characteristics of the input subject.

- **Create personal and lighting-specific effect:** Generally speaking, human shadows and highlights are very sensitive to individual characteristics. In Fig. 5, \( M^G \) and \( M^D \) differ depending on the specific lighting condition and the unique characteristics of the input subject, so that the results generated by the proposed framework not only reasonable but also more diverse. Consequently, the relighting results reflect a personalized shadow effect, while the removal results successfully maintain the facial appearance of the input facial features (see the second, fourth, and sixth rows). Note that, for each prediction result \( I_{y,\text{test}} \) shown in the last row, its corresponding ground-truth can refer to the image \( I_{x,\text{test}} \) shown in the first row of the opposite lighting transformation case. However, some artifacts still exist in the final results in the case, if the input image with large shadow areas (e.g., cases c5 to c0).

- **Visual comparison of results of MRF \( G^D / D \) and MRF \( G^G / D \):** Fig. 6 compares the relighting performances of the proposed MRF \( G^G / D \) – MRF \( G^D / D \) graph (M1) with the traditional MRF graph (M2, without the additional adaptive map), where for a fair comparison, each of these two methods combines the global and local components processes individually and merges them. For visualization purposes, the first and second columns for each subject show the relighting result and its particular region result obtained after increasing the contrast, respectively. Fig. 6(a) shows \( I_{x,\text{test}}^G \) of three testing subjects, estimated results produced by the proposed approach (M1) show a more significant lighting contrast around facial features than those produced by the traditional MRF model (M2). In Fig. 6(b), the corresponding estimated \( I_{x,\text{test}}^D \) components of these testing subjects are shown, where the estimated facial features produced by the proposed M1 approach are closer to the original input image characteristic than those produced by M2. The final results \( I_{x,\text{test}}^D \), which integrate the global and local components, are shown in Fig. 6(c). The proposed method generates better and clearer mouth shape in the first two subject than the typical M2 method. Furthermore, there are still some artifacts in the synthesized result of the right eye region produced by M2 method.

**FIGURE 6.** Comparison of the relighted results produced by the MRF model without adaptive map (the typical MRF model, M2) and with adaptive map (the proposed model, M1). Note that the first row shows the corresponding reference ground-truth (GT) images. (a) Global component results; (b) results of local detailed component; and (c) final synthesized results consisting of both global and local detailed texture components. Note that for each subject, the second columns present particular local regions of the facial images for visual comparison purposes.

**IV. EXPERIMENTAL RESULTS**

The present study uses the YaleB database for training and testing purposes. The database contains 38 individuals, 32 of whom are randomly selected to collect the training sample images, while the remaining 6 subjects are used for testing purposes. For each image, an affine transformation process was performed to geometrically align the image with a 192x168 image size using three reference points. This present study considers 9 different lighting conditions, namely the normal (natural) lighting condition (referred to as c0) and eight different lighting conditions (referred to as c1–c8), as shown in Fig. 7. In the configurations of the MRF model and the classifier, each image is decomposed into \( N_k \) overlapping patch grids, with image patch sizes of 20 x 20 and the overlapping distance between adjacent patches set to
10 pixels. In particular, \( \eta \) and \( \eta' \) in (10) and (11) play an important role in MRF relighting modeling, and are determined by cross-verification (see APPENDIX); In this study, \( \eta \) and \( \eta' \) of the \( W^s_{\text{test}} \) estimation process are set to 1.6 and 1.3, \( I_{\text{test}}^s \) to 1.6 and 1.2, \( W^d_{\text{test}} \) to 2.7 and 2.3, and \( I_{\text{test}}^d \) to 2.7 and 1.4. In addition, the search window for the MRF modeling and the classifier learning processes are defined as the border area of 5 pixels expanded from the center of the image patch. Finally, for the inference-type MRF nodes, the number of state candidates required for MRF optimization was set to 5.

1) VISUALIZATION ANALYSIS OF THE RELIGHTING RESULT UNDER DIFFERENT LIGHTING CONDITIONS

![Figure 8](image1.png)

**FIGURE 8.** Relighting results for four testing example subjects under the eight specified lighting conditions. (a) Relighted results from source natural lighting condition \( c_0 \) to target lighting conditions \( c_1 \sim c_8 \); and (b) corresponding relighting removal cases.

Figure 8(a) shows the proposed algorithm to illuminate the images of four testing subjects form source lighting condition \( c_0 \), to 8 different target lighting conditions, \( c_1 \sim c_8 \). Figure 8(b) shows their corresponding lighting removal result, where the input images taken under the source lighting conditions, \( c_1 \sim c_8 \), are restored to the \( c_0 \) source condition. Overall, these images support the following observations:

- Results on different subjects have personalized, not monotonous lighting effects: Each row of Fig. 8(a) shows that shadows of each synthesis result vary greatly depending on facial characteristics, contours, and positions of a particular light source. For example, in \( c_3 \), the shapes of shadow in the nose regions differs significantly from that of another. To be specified, the relighted images under the same lighting condition in Fig. 9 are different, as are the distribution of each symbol in Fig. 10. At the same time, the distributions of each symbols (results of a particular lighting

![Figure 9](image2.png)

**FIGURE 9.** Distribution of real and relighted image results under eight lighting conditions. Note that the numbers 0, 1, 2,..., 8 ' denote the real images, while symbols ‘&’, ‘*’, ‘+’, ‘@’, ‘!’ denote the relighted results produced by the proposed framework.

- Results for different lighting conditions have different lighting effects, and successfully follow the training samples: Each column of Fig. 8(a) indicates that the result of the synthesis of each individual have different shading distributions in different facial areas, depending on the lighting conditions. For example, facial shadows are distributed in different parts of the face, such as cheeks or nose edges, depending on the location of the light source in relation to the face.

- Results of the same subject’s lighting removal under different lighting conditions are consistent with the identity: On the hand, in Fig. 8(b), each column shows that each subject’s restored images from different lighting source condition successfully preserve the input subject’s particular facial characteristics. However, under source lighting condition \( c_8 \), since the input image has large areas of facial shadows, the light removal results contains some artificial ghosting effects, e.g., the eyes of the third subject lost the geometric symmetry.

- Distribution of real and relighted images under different lighting conditions is consistent: Fig. 9 analyzes the image relighting results obtained using the proposed framework by projecting the relighted images on the manifold space constructed by the real training images from all 9 different light conditions. Fig. 9 shows different symbols for the results of the lighting and the removal of testing subjects displayed in Fig. 8(a) and (b), which are spatially close to the real image distribution, which is marked in different numbers.

Overall, the results presented in Fig. 9 and Fig. 10 confirm that the proposed framework ensures the individual’s lighting or shadow areas are visually different from each another. To be specified, the relighted images under the same lighting condition in Fig. 9 are different, as are the distribution of each symbol in Fig. 10. At the same time, the distributions of each symbols (results of a particular lighting
condition) are closed to the distributions of its corresponding ground-truth lighting condition to other lighting conditions.

2) COMPARISON WITH OTHER RELIGHTING MODELS

Figures 10, 11,12 compare the relighting performance of the proposed method with that of 5 existing example-based relighting models, induding PCA+KCCA (the kernel approach in [10]), PCA+DDCM (the kernel approach in [23]), 2DDCM (the kernel approach in [24]), CVAE [20] (the deep leaning-based kernel approach in relighting framework [8]), and MRF approach (simplified kernel approach in [21]). All of these 4 methods are trained using the same training data set as the proposed approach. We also include two relighting models, the histogram matching (the kernel approach in [25]) and headshot [18], which require a user-specified reference rather than a training dataset. We manually select the most suitable image for each testing subject as the reference.

- PSNR/SSIM quality assessment for relighting results: As shown in Fig. 10, the proposed MG/D-MRF, MRF, 2DDCM, and CVAE models have higher performance than the other models in terms of SSIM/PSNR value, and their visual comparisons under the lighting condition c3 are shown in Fig. 11. (a visual comparison of MRF model is shown in Fig 6). Although the results of the 2DDCM model are highest in SSIM/PSNR score of Fig.10, but our results are more detailed and clear contours than those 2DDCM/CVAE model.

- Light removal results for the robust facial recognition: The performance of the various relighting methods was further evaluated by the top 10 ranking performance of the face retrieval process in Fig. 12. Different relighting models are used to transfer images from side lighting conditions (c1~c8) to a natural light condition (c0), which are used as queries for the face identification database. Note that the retrieval process was performed in the PCA subspace, and the similarity of the query images to the database images was evaluated by the Education distance. The face recognition environment was conducted using two databases, namely: Face database: the database containing all 38 subjects from the training and testing databases form the normal light condition c0, each image has its own annotated identity. Query database: the database contains 48 query images of 6 testing subjects under eight lighting conditions, and different lighting models are then applied to generate their removal images for comparison with the database image.

- Discussion for model performance: The results presented in Fig. 12 show that the proposed relighting method yields a better retrieval performance than the other methods. In particular, the retrieval results are consistent with the PSNR results presented in Fig. 10(b). Fig. 12 confirms that the proposed approach can retain the individual characteristics of input images since the semantic notable regions of the testing subjects are automatically and adaptively detected and integrated into the proposed lighting model. Results produced by 2DDCM/CVAE methods also achieve good retrieval.
performance since they preserve the complete 2D facial geometry of the testing image. In contrast, the headshot and histogram matching methods only consider a single reference image, and hence the relighting results tend to be similar to the reference image. As a result, their retrieval performances are degraded. Finally, the PCA-based CCA and DCM methods both consider multiple learning samples for the relighting process. However, they suffer from an overfitting problem, which destroys the geometric characteristics of the facial image and degrades the retrieval performance accordingly.

To further evaluate the overfitting issue, two additional testing procedures were performed to test the proposed approach.

- **Cross-dataset testing:** The training data set was taken from the YableB database, while the testing samples were taken from the CMU database [19], where the same geometry normalization operation was applied to the images in both datasets. The corresponding relighting results are presented in Fig. 13. It is seen that even when the testing images are taken from a different database than that used for training purposes, the proposed framework still produces relighting results with realistic shadow patterns and facial features that resemble those of the input image. However, some obvious artificial noise is still evident in some of the predicted results (marked by red circles) due to inconsistencies between the images in the two databases.

- **Facial appearance-inconsistency testing:** In Fig. 15, training and testing datasets were both compiled from the YableB database. However, the testing images were rendered geometrically inconsistent with the training images by introducing resolution, geometry, out-of-plane, and facial expression variations. It is seen that the proposed method achieves a reliable prediction performance even when geometric inconsistencies exist between the testing images and the training images. We infer that our proposed MRF graph model is adaptively changed, and hence an acceptable relighting result is achieved. A detailed inspection of the global and local adaptive maps, \( M^G_{\text{test}} \) and \( M^D_{\text{test}} \), confirms that the extracted maps are changed following these image inconsistencies.

\[ \text{FIGURE 13. (a) Three test subjects taken from CMU database, (b) corresponding relighting results for eight target lighting conditions learned from the YableB training dataset. Note that the red circles indicate obvious noise artifacts in the re-lighted image.} \]

\[ \text{FIGURE 14. Application of the proposed approach to testing samples with variations in facial appearances. (a) Geometry correction, (b) low resolution, (c) out-of-plane variation, and (d) facial expression variation. The relighting result is provided in the second row. Note that, for comparison purposes, in each column, the left testing images have appearance inconsistencies with the training samples, while the right testing images are appearance consistent with the training samples.} \]

\[ \text{V. CONCLUSION} \]

This study has proposed a novel lighting- and personal characteristic-aware Markov random field (MRF) framework for relighting facial images. The experimental results have shown that the proposed framework has three main advantages compared to existing example-based face relighting schemes. (i) The adaptive MRF graph model ensures that the results are not only similar to the training samples under the target lighting conditions, but also preserves the individual personal characteristics of the system input. (ii) The proposed framework can produce a relighting image based on input the facial characteristics, so the results are not only reasonable but also more diverse, i.e. the distributions of the produced shadow and highlight is not monotonous. (iii) The proposed model provides acceptable results for unseen test subjects, despite the limited number of training samples. This is because the proposed framework detects areas that need to be enhanced based on system input facial characteristic and then use these areas to dynamically change the MRF graph to select the on-demand reference for each patch. Overall, the proposed model can be applied to the image transformation tasks, especially for the supervised learning situations where only limited training pairwise samples are available. In a future study, the capabilities of the
proposed framework will be extended to a wide variety of face synthesis/transformation applications, including non-
photorealistic rendering, super-resolution, and facial expression synthesis.

ACKNOWLEDGMENT
This work was supported in part by the Ministry of Science and
Technology of Taiwan under Grant MOST 109-2221-E-
005-056-MY2.

APPENDIX:
For most MRF models [22, 31, 32], the data term and the
smoothness term play important roles in determining the
optimal hidden node state. To define suitable and
discriminative values, $\eta$ and $\eta'$ in (10) and (11) for selecting
discriminative candidates, a cross-validation process is
applied. In practice, for example, to determine the appropriate
value of $\eta$, 30 different values of $\eta$ (ranging from 0–3) are
considered. For each value, we construct a covariance matrix
with size $N \times N$ ($N$: the number of training samples), and its
entry $(i, j)$ is calculated by the average similarity values of $N_k$
patches for the $i$th and $j$th training subjects, where the patch
similarity value is defined using the formulation in (10) and
(11) based on this specified $\eta/\eta'$ value. Finally, the correlation
matrix of a given $\eta/\eta'$ with highest entropy value is the
optimum.

Fig. 15 shows the entropy curves of the data (likelihood)$\phi(\cdot)$ and smoothness (prior)$\psi(\cdot)$ terms for different $\eta/\eta'$
values, where the gain and base matrices ($W^{G/B}_i$ or $I^{G/B}_i$) are
evaluated separately. Note that, in each case, the horizontal
axis represents the different $\eta$ or $\eta'$ value, while the vertical
axis shows the corresponding response entropy value. In
general, the higher the value of entropy, the more
discriminatory it is to select training sample for modeling
construction (refer to the three covariance matrices in the first
subfigure, $\phi(\cdot;\eta)$, where the optimum is marked by red
circles).

FIGURE 15. Entropy curves for selecting $\eta$ or $\eta'$ in normalization terms
of the MRF data term $\phi(\eta)$ and the smoothness term $\psi(\eta')$. For each
figure, there are 30 different values of $\eta$ or $\eta'$; each curve have an optimal
value indicated in red circle, and the correlation matrix of the similarity
scores of all the training patches based on that value is the maximum
entropy compared to the other values.

REFERENCES
[1] Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, “Automatic Contrast-
Limited Adaptive Histogram Equalization with Dual Gamma Correction,”
IEEE Access, pp. 11782-11792, 2018.
[2] X. Chen, H. Wu, X. Jin and Q. Zhao, “Face Illumination Manipulation
Using A Single Reference Image by Adaptive Layer Decomposition,”
IEEE Transactions on Image Processing, pp. 4249-4259, 2013.
[3] X. Chen, X. Jin, H. Wu and Q. Zhao, “Learning Templates for Artistic
Portrait Lighting Analysis,” IEEE Transactions on Image Processing,
pp. 608-618, 2015.
[4] J. Chen, G. Su, J. He, and S. Ben, “Face Image Relighting using Locally
Constrained Global Optimization,” ECCV, pp. 44-57, 2010.
[5] P. Choudhary, S. Manwar, A. Kumar, and M. Son, “Real-Time Face
Relighting via Adaptive Normal Mapping,” SIGGRAPH ASIA Mobile
Graphics and Interactive Applications, 2017.
[6] X. Geng, K. Smith-Miles, Z. H. Zhou, and L. Wang, “Face Image
Modelling by Multilinear Subspace Analysis with Missing Values,”
IEEE Transactions on System, Man, and Cybernetics, Part B (TSMC-
B), pp.881-892, 2011.
[7] 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. Intelligence,
p. 643-660, 2001.
[8] X. Han, H. Yang, G. Xing, and Y. Liu, “Asymmetric Joint GANs for
Normalizing Face Illumination from a Single Image,” IEEE Transactions
on Multimedia. PP. 1-1, 2019.
[9] X. Han, Y. Liu, H. Yang, G. Xing, and Y. Zhang, “Normalization of
Face Illumination with Photorealistic Texture via Deep Image Prior
Synthesis,” Neurocomputing 386, pp. 305-316, 2020.
[10] H. Huang and H. He, “Super-Resolution Method for Face Recognition
Using Nonlinear Mappings on Coherent Features,” IEEE Trans.
Neural Networks, pp. 121-130, 2011.
[11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image
Translation with Conditional Adversarial Networks, “ CVPR, pp. 5967-
5976, 2017.
[12] H. Jiang, K. Bang, J. Jang, and D. Hwang, “Dynamic Range Expansion
Using Cumulative Histogram Learning for High Dynamic Range
Image Generation,” IEEE Access, pp. 38554-38567, 2020.
[13] C. Li and M. Wand, “Combining Markov Random Fields and
Convolutional Neural Networks for Image Synthesis,” CVPR, pp.
2479-2486, 2016.
[14] H. Li, X. Jia, and L. Zhang, “Clustering Based Content and Color
Adaptive Tone Mapping,” Comput. Vis. Image Underst., pp. 37-49,
2018.
[15] L. Liang and X. Zhang, “Adaptive Label Propagation for Facial
Appearance Transfer,” IEEE Trans. Multimedia, pp. 3068-3082, 2019.
[16] L. Liang, L. Jin, and Y. Xu, “PDE Learning of Filtering and
Propagation for Task-Aware Facial Intrinsic Image Analysis,” IEEE
Trans cybernetics, 2020.
[17] T. V. Luan, H.-Y. Lin, “Extending and Matching a High Dynamic
Range Image from a Single Image,” Sensors, 2020.
[18] Y. Shih, S. Paris, C. Barnes, W. T. Freeman and F. Durand, “Style
Transfer for Headshot Portraits,” ACM Trans. Graph. 33(4), pp. 1-148,
2014.
[19] T. Sim, S. Baker, and M. Bsat, “The CMU Pose, Illumination, and
Expression Database,” IEEE Trans. Pattern Anal. Mach. Intell., pp.
1615-1618, 2003.
[20] K. Sohn, H. Lee, and X. Yan, “Learning Structured Output
Representation using Deep Conditional Generative Models,” NIPS,
pp. 3483-3491, 2015.
[21] Y. Song, L. Bao, and S. He, Q. Yang, M.-H. Yang, “Stylizing Face
Images via Multiple Exemplars,” Comput. Vis. Image Underst. 162, pp.
135-145, 2017.
[22] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman, “Learning
Gaussian Conditional Random Fields for Low-Level Vision,” CVPR,
pp. 1-8, 2007.
[23] C.-T. Tu, J.-J. J. Lien, “Automatic Location of Facial Feature Points and Synthesis of Facial Sketches Using Direct Combined Model,” IEEE Trans. Systems, Man, and Cybernetics, Part B, pp. 1158-1169, 2010.

[24] C.-T. Tu, M.-C. Ho, and M.-Y. Lin, “A New Approach for Face Hallucination Based on A Two-Dimensional Direct Combined Model,” Pattern Recognit., pp. 1-20, 2017.

[25] T. F. Y. Vicente and D. Samaras, “Single Image Shadow Removal via Neighbor-Based Region Relighting” ECCV Workshops, pp. 309-320, 2014.

[26] Y. Wang, L. Zhang, Z. Liu, G. Hua, Z. Wen, Z. Zhang and D. Samaras, “Face Relighting from a Single Image under Arbitrary Unknown Lighting Conditions,” IEEE Trans. Pattern Anal. Mach. Intell. 31(11), pp. 1968-1984, 2009.

[27] Z. Wang, X. Yu, M. Lu, Q. Wang, C. Qian, and F. Xu, “Single Image Portrait Relighting via Explicit Multiple Reflectance Channel Modeling,” ACM Trans. Graph, pp.220:1-220:13, 2020.

[28] X. Xie, W. S. Zheng, J. Lai, P. C. Yuen and C. Y. Suen, “Normalization of Face Illumination Based on Large-and Small-Scale Features,” IEEE Transactions on Image Processing 20(7), pp. 1807-1821, 2011.

[29] W. Xu, X. Xie, and J. Lai, "RelightGAN: Instance-level Generative Adversarial Network for Face Illumination Transfer", IEEE Transactions on Image Processing, pp. 3450-3460, 2021.

[30] Z. Xu, K. Sunkavalli, S. Hadap, and R. Ramamoorti, “Deep Image-Based Relighting from Optimal Sparse Samples,” ACM Trans. Graph, pp. 126:1-126:13, 2018.

[31] M. Zhang, N. Wang, X. Gao, and Y. Li, “Markov Random Neural Fields for Face Sketch Synthesis,” Int. Joint Conf. Artif. Intell., pp. 1142–1148, 2018.

[32] M. Zhang, N. Wang, Y. Li, and X. Gao, “Neural Probabilistic Graphical Model for Face Sketch Synthesis,” IEEE Trans. Neural Networks Learn. Syst, pp. 2623-2637, 2020.

[33] H. Zhou, Z. Kuang, and K.-Y. K. Wong, “Markov Weight Fields for Face Sketch Synthesis,” In IEEE Conference on Computer Vision and Pattern Recognition, pp. 1091–1097, 2012.

[34] H. Zhou, S. Hadap, K. Sunkavalli, and D. Jacobs, “Deep Single-Image Portrait Relighting,” ICCV, pp. 7193-7201, 2019.

[35] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks", IEEE International Conference on Computer Vision (ICCV), pp. 2242-2251, 2017.

Ching-Ting Tu is an Associate Professor at the Department of applied mathematics, National Chung Hsing University, Taiwan. She joined the department of computer science and information engineering at Tamkang University, Taiwan, in 2011 as an Assistant Professor, and was later promoted as an Associate Professor in 2017. Her research interests include computer vision and pattern recognition, and multimedia information analysis.

Hwei-Jen Lin received the B.S. degree in applied mathematics from National Chiio Tung University, Hsinchu, Taiwan, in 1981, and the M.S. and Ph.D. degrees in mathematics from Northeastern University, Boston, MA, in 1983 and 1989, respectively. She is currently a Professor with the Department of Information Engineering, Tamkang University, Taipei, Taiwan. Her current research interests include pattern recognition, image processing, computer algorithms, and formal language.

Chin-Yu Chang received the M.S. degree from department of computer science and information engineering at Tamkang University, Taiwan, in 2017. Her research interests include face recognition and computer vision.