Automatic Denoising of 2D Color Face Images Using Recursive PCA Reconstruction

Hyun Park and Young Shik Moon

Department of Computer Science and Engineering, Hanyang University, 1271 Sa-Dong, Ansan, Kyunggi-Do 425-791, Korea
{hpark, yymoon}@cse.hanyang.ac.kr

Abstract. In this paper, we propose a denoising method based on PCA reconstruction for removing complex color noise components on human faces, which is not easy to remove by using vectorial color filters. The proposed method is composed of the following six steps: training of canonical eigenface space using PCA, automatic extraction of facial features using active appearance model and alignment of the input face to mean shape, reconstruction of an initial noise free face, relighting of reconstructed face using a bilateral filter, extraction of noise regions using the variances of skin color of training data, and reconstruction using partial information of input images (except the noise regions) and blending of the reconstructed image with the original image. Experimental results show that the proposed denoising method maintains the structural characteristics of input faces, while efficiently removing noise components with complex colors.

1 Introduction

Denoising and reconstruction of color images have been extensively studied in the field of computer vision and image processing. There have been some attempts to remove noises on color images. Early attempts removed noises on color images through independent smoothing of RGB channels. Generally, almost all approaches focus on a variety of filtering processes applied appropriately to the color vectorial data. The color filters such as vector median and directional filters are used for removing Gaussian white noise or impulse noise in the field of computer vision and image processing [1], [2], [3]. The nonlinear color filters such as WMF (weighted median filter) and CWMF (center weighted median filter) efficiently remove impulse and salt & pepper noises [4], [5]. The decomposition using PCA (principal component analysis), kernel PCA, ICA (independent component analysis) and wavelet transform is also applied to denoising [6], [7], [8]. In spite of these efforts, many of denoising methods have been performed on gray images and they removed mostly simple noises such as gaussian noises or impulse noises. Moreover, complex color noise components on human faces are difficult to remove by general color filtering processes.

Therefore, we propose a new denoising method based on recursive PCA reconstruction, which maintains the structural characteristics of input face and efficiently removes complex color noise components on input faces. The proposed method is...
composed of the following four steps. First, we construct a canonical eigenface space using PCA. Next, we automatically extract facial features of an input face using the multi-level active appearance model (MAAM). To minimize the reconstruction error by geometric misalignment, we align the input face to the reference shape using the extracted facial feature points. Next, we reconstruct an initial noise free face by projecting the input face onto constructed canonical eigenface space. We carry out the proposed relighting method so that both the reconstructed face and the input face have the same illumination condition. Then, we extract noise regions using the variances of vector magnitude and vector angle of the skin color at each pixel position of the training data. Finally, if the extracted noise regions are less than 40% of the total face region, we reconstruct the noise-free face once again using partial information of the input image (except the noise regions), and the reconstructed noise-free face is blended appropriately with the original image.

2 Training of Canonical Eigenface Space Using PCA

Generally, the original eigenface space is effective for recognition and reconstruction, but it is not robust against various illumination changes and geometric misalignments. In this paper, complex color noise components on training faces are manually removed, then training images are aligned to the mean shape of AAM and normalized to ‘zero mean and unit length’, as in equation (1).

We refer to this training face as the normalized canonical face. The training data consists of still images of 100 different frontal-view human faces, all without glasses and with a neutral expression. We construct a canonical eigenface space using PCA and normalized canonical training faces. Finally, the canonical eigenface space is constructed by removing the rest of complex color noises that have not been removed manually, which is performed by selecting only 95% of the principal components.

\[ x = \frac{X - E\{X\}}{\|X - E\{X\}\|} \quad E\{X\} = \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n \quad \Sigma = \frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x})(x_n - \bar{x})^T \]  

(1)

where \( x \) is the normalized face of original face \( X \), \( \bar{x} \) is the mean vector, \( \Sigma \) is the covariance matrix, \( N \) is the number of faces in the training set.

\[ \Lambda = \text{diag}(\sigma_i^2) = \Phi^T \Sigma \Phi \quad \alpha = \Phi^T_m (x - \bar{x}) \quad x^* = \sum_{i=1}^{m} \alpha_i \phi_i \]  

(2)

In equation (2), \( \Lambda \) is a diagonal matrix in which diagonal terms are eigenvalues of \( \Sigma \), and \( \sigma_i^2 \) is the variance of training faces in the direction of \( i \)th eigenvector. \( \Phi \) is an eigenvector matrix, \( \alpha \) is a principal component vector, \( m \) is the number of eigenvectors, and \( x^* \) is the reconstructed face by projection onto the canonical eigenface space [6], [9]. If we use partial information of the input face and the scaled eigenvectors \( \sigma_i \phi_i \) as a basis, reconstructed face \( x^* \) is defined by equation (3) [9].

\[ x^* = \sum_{i=1}^{m} \alpha_i \sigma_i \phi_i = \Phi \cdot \text{diag}(\sigma_i) \cdot \alpha \]  

(3)
3 Automatic Extraction of Facial Feature Points Using AAM

In order to improve the efficiency and robustness of the matching algorithm, a facial feature template is matched by using the MAAM (Multi-level active appearance model) based on color images. The training method for AAM uses the Jacobian learning scheme. As in the original AAM method, these AAMs are built at each level of a scale-pyramid for coarse-to-fine fitting based on multi-resolution [10], [11].

![Face textures and shapes (landmarks) for training the AAM](image)

**Fig. 1.** Face textures and shapes (landmarks) for training the AAM

![Search-based initialization of deformation templates using multi-level AAM](image)

**(a) Original image. (b) Search result in level 2 (50\% scaled down of level 1). (c) Search result in level 1 (50\% scaled down of level 0). (d) Final search result in level 0 (not scaled down).**

The training faces are acquired in 1187 x 1190 bitmap color format. As shown in Figure 1(b), the facial structures are manually annotated using 94 total landmarks of eyebrows, eyes, nose, mouth, and jaw. As shown in Figure 2, a multi-resolution pyramid is scaled into three levels. Figure 2(d) is the final search result in which parameters of the combined model for AAM are optimized (translation, scaling, rotation, texture model parameters, and shape model parameters).

4 Reconstruction of Noise Free Face Reflecting Various Facial Colors

In the PCA reconstruction in color domain, the difference between the color distribution of input face and that of training faces causes PCA reconstruction error. Especially, the reconstruction error becomes larger as the illumination condition changes. Therefore, the direct projection of color face onto the canonical eigenface space, using $x^*$ of equation (2), may not be robust. We solve this problem by using the polynomial regression approximation.
As shown in (1) of Figure 3, in order to minimize the difference of facial color between the input face and the training faces, we convert the facial color of input face to the facial color of the mean face using the Polynomial-Least Squares Fitting (PLSF) that is carried out by equation (5) and (6). Next, by projecting the input face onto the canonical eigenface space, the initial noise free face is reconstructed, as in (2), (3) of Figure 3. Finally, we convert the facial color of reconstructed face to the facial color of the input face.

\[
\begin{align*}
\mathbf{x}_{\text{facial color } A} &= \{x_1, x_2, x_3, \ldots, x_n\} \\
\mathbf{y}_{\text{facial color } B} &= \{y_1, y_2, y_3, \ldots, y_n\} \\
\mathbf{y}_{\text{facial color } A} &= \{y_1, y_2, y_3, \ldots, y_n\}
\end{align*}
\]  \hspace{1cm} (4)

\[
y_{\text{facial color } A} = \text{PLSF}(\mathbf{x}_{\text{facial color } A}, \mathbf{y}_{\text{facial color } B}) = ax^2 + bx + c = y \hspace{1cm} (5)
\]

\[
a = \frac{(n\sum x^2 y) - (\sum x^2 \sum x)}{(n\sum x^2) - (\sum x^2)^2}, \quad b = \frac{\sum xy}{\sum x^2}, \quad c = \frac{\sum x^2 y - \sum x^2 \sum x^2 y}{(n\sum x^2) - (\sum x^2)^2} \hspace{1cm} (6)
\]

where \(\mathbf{x}_{\text{facial color } A}\) is a reference face with facial color A, \(\mathbf{y}_{\text{facial color } B}\) is an input face with facial color B, \(\mathbf{y}_{\text{facial color } A}\) is a converted input face.

5 Relighting of Reconstructed Face

We can not always assume that both the reconstructed face and the input face are on the same illumination condition. To compensate the difference of illumination conditions, we propose a relighting method using the bilateral filter in HSV color domain [12]. Our relighting method is composed of the following two steps: initial relighting of reconstructed face using the bilateral filter, compensation of the initial relighting to exclude noise influences.

5.1 Relighting of Reconstructed Face Using Bilateral Filter

The bilateral filter combines a classic low-pass filter with an edge-stopping function that attenuates the filter kernel weights when the intensity difference between pixels is
large. As shown in Figure 4, the bilateral filter separates a color face image into the small scale and the large scale. The small scale represents the detailed shape of image and the large scale represents the illumination of image. We apply the bilateral filter to each RGB color channel separately with the same standard deviation parameters for all three channels. The output of bilateral filter using equation (7) and (8) is the large scale. The small scale is computed to divide color image by its large scale. By combining the saturation and value (brightness) of the large scale input face, the hue of large scale reconstructed face, and the small scale reconstructed face, we perform the relighting of the reconstructed face. The computation is done in the log domain to take the intensity ratios in account. Spatial variance $\sigma_i$ is equal to 2% of the image diagonal and variance $\sigma_g$ is equal to 0.4 for intensity influence [13], [14].

$$J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s)g(I_p - I_s)I_p$$  \hspace{1cm} (7)

$$k(s) = \sum_{p \in \Omega} f(p-s)g(I_p - I_s), \quad f(x) = g(x) = \frac{1}{2} e^{-\frac{x^2}{\sigma^2}}$$  \hspace{1cm} (8)

Here $p$ and $s$ are pixel positions, $I_p$ and $I_s$ are pixel values (or each color channel values) at $p$ and $s$ pixel positions respectively, $k(s)$ is normalization term and $\Omega$ is the size of filter mask.

5.2 Compensation of Relighting Using Extracted Noise Region Information

The large scale representing the illumination of color face is affected by complex color noise components that are difficult to remove by Gaussian filtering of the
bilateral filter. Therefore, the result of relighting might have a little noise effects. To prevent these noise effects occurred by the initial relighting, we propose the modified joint bilateral filter using noise region such as equation (9), (10) and (11).

\[
J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s)g(e(I_p)-e(I_s))e(I_p)
\]  

(9)

\[
k(s) = \sum_{p \in \Omega} f(p-s)g(e(I_p)-e(I_s))
\]  

(10)

\[
e(I_p) = \begin{cases} 
\text{if } p \in \text{noise region, the pixel of } p \text{ index on reconstructed face} \\
\text{if } p \notin \text{noise region, the pixel of } p \text{ index on input face}
\end{cases}
\]  

(11)

6 Extraction of Noise Regions Using the Variance of Skin Color

In order to extract complex color noise regions automatically, we propose a noise detection method based on the vector magnitude map (VMM) and vector direction map (VDM) of training data that is computed by equation (12) and (13) [2], [3]. The vector magnitude represents the distance and the vector direction represents the angle between two vectors at the same pixel position. As in equation (14), we use the standard deviation of VMM and VDM at each pixel position as the threshold value. The threshold value for noise extraction is determined by experiments for over detection of color noise regions. So the threshold value has been empirically selected to 2 times standard deviation of VMM and VDM, based on noise detection rate.

\[
\sigma_{VM,i}^2 = \frac{1}{N} \sum_{n=1}^{N} (D(x_{n,i}, \bar{x}_i) - D(x_{n,i'}, \bar{x}_i'))^2, \quad D(a,b) = \left( \sum_{k=1}^{d} (a^k - b^k)^2 \right)^{1/2}
\]  

(12)

\[
\sigma_{VD,i}^2 = \frac{1}{N} \sum_{n=1}^{N} \left( A(x_{n,i}, \bar{x}_i) - A(x_{n,i'}, \bar{x}_i') \right)^2, \quad A(a,b) = \cos^{-1} \left( \frac{a \cdot b}{|a||b|} \right)
\]  

(13)

\[
x_{\text{input},i} \text{ is noise pixel if } \begin{cases} 
2\sigma_{VM,i} < D(x_{\text{input},i}, x_{\text{recon},i}) \text{ and} \\
2\sigma_{VD,i} < A(x_{\text{input},i}, x_{\text{recon},i})
\end{cases}
\]  

(14)

Here \( \bar{x} \) is the mean vector of the training set, \( N \) is the number of data in the training set, \( D(a,b) \) is the distance between two vectors, \( A(a,b) \) is the angle between two vectors, \( a \) and \( b \) are pixel vectors, \( k \) is the dimension of a pixel vector, \( \sigma_{VM,i} \) is the standard deviation of VMM at \( i \) th position and \( \sigma_{VD,i} \) is the standard deviation of VDM at \( i \) th position.

7 Reconstruction Using Partial Information and Blending

Generally, the least squares minimization (LSM) method using orthogonal projection and the original PCA reconstruction are not robust when input images have
intra-sample outliers. We can regard complex color noise components on facial images as intra-sample outliers. Therefore, we reconstruct the optimal noise free face using the robust PCA based on the singular value decomposition (SVD). The robust PCA computes the optimal principal components by using the partial information of input images (except the noise regions), and then we construct the noise free face by using the optimal principal components [9], [15]. If the area of the extracted noise region is more than 40% of the total face region, we use the reconstructed noise free face by orthogonal projection.

We define an error function $E(\alpha)$ such as equation (16), as the sum of square errors which are the difference between pixel values in non-noise regions and its reconstructed ones. Our goal is to find the optimal $\alpha^*$ so as to minimize the error.

$$\alpha^* = \arg \min_{\alpha} E(\alpha)$$  \hspace{1cm} \text{(15)}$$

$$E(\alpha) = \sum_{j=1}^{p} \left( \tilde{x}(j) - \sum_{i=1}^{m} \alpha_i \sigma_i \phi_i(j) \right)^2$$  \hspace{1cm} \text{(16)}$$Here $\tilde{x}(j)$ are pixels of the input image except the noise regions, $p$ is the number of pixels in the non-noise regions. If $q_i = \sigma_i \phi_i$, the equation (16) is replaced by equation (17).

$$E(\alpha) = \sum_{j=1}^{p} \left( \tilde{x}(j) - \sum_{i=1}^{m} \alpha_i q_i \phi_i(j) \right)^2 = |\tilde{x} - Q \alpha|^2$$  \hspace{1cm} \text{(17)}$$

$$Q = U W V^T, \quad Q^+ = V W^+ U^T$$  \hspace{1cm} \text{(18)}$$

$$W^+ = \text{diag} \begin{pmatrix} w_i^{-1} & \text{if } w_i \neq 0 \\ 0 & \text{otherwise} \end{pmatrix}$$  \hspace{1cm} \text{(19)}$$

$$\alpha = Q^+ \tilde{x}, \quad \alpha^* = \sigma \alpha = \sigma Q^+ \tilde{x}$$  \hspace{1cm} \text{(20)}$$

$$x^{\text{recon}} = \sum_{i=1}^{m} \alpha_i^* \phi_i + \tilde{x}$$  \hspace{1cm} \text{(21)}$$

We can get the optimal $\alpha^*$ by using the pseudo-inverse of $Q$ that is computed by SVD. As the reconstructed face by orthogonal projection, we apply the facial color transfer and the relighting to the reconstructed face by equation (21). Finally, we blend the reconstructed noise free face with the original face by equation (22). In equation (22), b is a blending ratio. In this paper, we use 0.9 as the blending ratio. O is the original face, and R is the final reconstructed face.

$$\text{Final noise free face } = \begin{cases} b \cdot O + (1-b) \cdot R & \text{if noise free region} \\ (1-b) \cdot O + b \cdot R & \text{if noise region} \end{cases}$$  \hspace{1cm} \text{(22)}$$
8 Experimental Results

We evaluate the performance of the proposed denoising method by removing noise components on the frontal face. For this experiment, we manually insert complex color noise components such as pockmarks, pimples and blotches on a clear face. As shown in Figure 6, the proposed denoising method maintains the structural characteristics of input face, while efficiently removing complex color noise components. As we carry out the denoising process repeatedly, more detailed information on the face is blurred. However, this blurring is negligible. As shown in Figure 7, we also evaluate the performance of the proposed denoising method by comparing with multilevel inpainting method, TV inpainting method, 7x7 WMF and 7x7 CWMF. Experimental results show that the proposed denoising method is more efficient in terms of smoothness, visual impression and denoising effect than the other methods.
Fig. 6. Noise removal results by using recursive PCA reconstruction (a) noise free input face, (b) input face with arbitrary noise components, (c) reconstructed results when the removal is performed once, (d) reconstructed results when the removal is performed three times

9 Conclusions

In this paper, we propose a denoising method based on PCA reconstruction for removing complex color noises on human faces, which is difficult to remove by using general color filters. The proposed method maintains the structural characteristics of input faces, while efficiently removing complex color noises on input faces. Experimental results show that the proposed denosing method efficiently removes complex color noise components on input face images.
Fig. 7. Comparison of original face with noise removed results by proposed method, inpainting methods and spatial filter methods (a) noise free input face, (b) input face with arbitrary noise components, (c) result by the proposed method (executed once), (d) result by the proposed method (iterated three times), (e) result by multilevel inpainting method, (f) result by TV inpainting method, (g) result by 7x7 WMF, (h) result by 7x7 CWMF

Acknowledgement

This research was supported by the MIC (Ministry of Information and Communication) of Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment).
References