Facial Ethnic Appearance Synthesis

Felix Juefei-Xu(✉) and Marios Savvides

Carnegie Mellon University, Pittsburgh, USA
felixu@cmu.edu, msavvid@ri.cmu.edu

Abstract. In this work, we have explored several subspace reconstruction methods for facial ethnic appearance synthesis (FEAS). In our experiments, our proposed dual subspace modeling using the Fukunaga Koontz transform (FKT) yields much better facial ethnic synthesis results than the ℓ_1 minimization, the ℓ_2 minimization and the principal component analysis (PCA) reconstruction method. With that, we are able to automatically and efficiently synthesize different facial ethnic appearance and alter the facial ethnic appearance of the query image to any other ethnic appearance as desired. Our technique well preserves the facial structure of the query image and simultaneously synthesize the skin tone and ethnic features that best matches target ethnicity group. Facial ethnic appearance synthesis can be applied to synthesizing facial images of a particular ethnicity group for unbalanced database, and can be used to train ethnicity invariant classifiers by generating multiple ethnic appearances of the same subject in the training stage.

Keywords: Soft biometrics · Ethnicity · Face synthesis · Fukunaga Koontz transform

1 Introduction

Within this decade, soft biometrics identification has gained more and more attention as an aid for the traditional face recognition in the biometrics world. Different from traditional hard biometrics such as iris, fingerprints, palmprints, and face [10–19, 31] that are difficult to change with the time and living behaviors and have high confidence in identifying subjects, the soft biometrics [8, 20, 25, 28–30, 34], on the other hand, focuses more on the physical and behavioral traits that are more prone to change with time and life style, and is less confident in subject identification if used alone. For example, the shape of the eyebrows, the presence of the beard and moustache, skin color, skin texture, color of the pupil, facial marks, gait patterns [9], and so forth, can all be considered as traits of soft biometrics. With the correct identification of these soft biometrics traits, we can infer the age, gender and the ethnicity of the subject. By doing so, we can dramatically narrow down the search space in the scenario of identifying or verifying the subject against a huge gallery database.

L. Agapito et al. (Eds.): ECCV 2014 Workshops, Part II, LNCS 8926, pp. 825–840, 2015.
DOI: 10.1007/978-3-319-16181-5_62

For ethnicity classification [1,5,7,26,33,35], it is crucial that the researchers obtain a balanced database with subjects from all ethnic groups[1] equally presented for both genders. This is one of the priorities before any learning algorithms are applied for ethnicity classification. Moreover, subjects in the database should be uniquely presented. In this way, the learning machine learns an ethnicity classifier instead of subject-dependent classifier. But unfortunately, database collection with high quality images covering all the ethnic group for both genders is pretty hard to accomplish[2]. That is why in this paper, we will be focusing on solving one of the biggest problems concerning the database creation for ethnicity classification by the synthesis of facial ethnic appearance. In this way, we can automatically and efficiently achieve the balance in the database while also keeping the subject uniqueness in the synthesized database.

The rest of this paper is organized as follows: in Section 2, we will describe our database with which the ethnicity-specific subspaces are built. Section 3 details the facial ethnic appearance synthesis using single subspace modeling methods such as the ℓ_2 minimization, principal components analysis reconstruction and the ℓ_1 minimization. Section 4 details the facial ethnic appearance synthesis using dual subspace modeling with Fukunaga Koontz transform. Experimental setup and results are discussed and analyzed in Section 5. Finally we present some conclusions of our work in Section 6.

2 Database

2.1 Database Collection

We have collected a database with a total of 6849 frontal mugshot-like images from 4 different ethnic groups: east asian, south asian, white and black. The statistics of our database is shown in Table 1.

Table 1. Statistics of our FEAS database

	Female	**Male**	**Total**
East Asian	559	477	1036
South Asian	86	138	224
White	2284	2256	4540
Black	482	567	1049
Total	3411	3438	6849

As can be seen from Table 1, the database is not balanced, the majority of the subjects are white people and there are very few south asians. Because of

[1] As is commonly adopted in the literature, the classification of ethnicity boils down to the 3-class case (asian, black and white), or the 4-class case (east asian, south asian, black and white). In this work, we consider the 4-class case, where we specifically separate south asians from east asians.

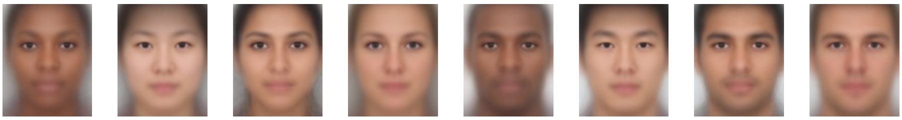
[2] White people dominates most of the ethnicity database publicly available, followed by black people. There are fewer east asian people and south asians are the rarest.

this limitation, the reconstruction performs worse if the target ethnic group is set to be south asian since there are not sufficient images to learn from. Our database also has a bias in age distribution. The majority of the subjects in the database are young adults from 20 to 30 years old. This bias tends to jeopardize the synthesis of the facial ethnic appearance when the query image is an aged subject. Figure 1 shows the mean faces from each of the ethnic group in our database for both genders.

2.2 Preprocessing

We localize and center the eye of each facial image using the modified active shape model (MASM) [32] and crop the rectified full image to be size of 84×68. The original size of the face in the image varies, and the reason we crop the face using this dimension is two-fold. First, this is a reasonable size to compute with using the reconstruction methods to be discussed. Second, some images in the database have low resolution and our cropping dimension of choice suppresses the artifacts and errors caused by up-sampling.

If high resolution synthesis is indeed desired in some applications, we can easily port the algorithms to GPUs using CUDA.



(a) (b) (c) (d) (e) (f) (g)

Fig. 1. Mean faces from our database. (a) Female black, (b) female east asian, (c) female south asian, (d) female white, (e) male black, (f) male east asian, (g) male south asian, and (h) male white.

3 Single Subspace Modeling for Facial Ethnic Appearance Synthesis

In this section, we show the use of single subspace for the synthesis of facial ethnic appearance. We construct a subspace using images from the target ethnic group. This ensures that any reconstruction obtained using components of this subspace has rich ethnic features of this particular ethnic group. We then reconstruct a given query face from the source ethnic group using this target subspace in order to synthesize the ethnic appearance.

Following this procedure, the synthesized facial image is supposed to preserve the subject identity as well as appear closer to the target ethnic group.

We first detail 3 well-established subspace methods using single subspace, namely the ℓ_2 minimization, principal component analysis reconstruction, and the ℓ_1 minimization and then show some results obtained by each of the methods, followed by analysis and discussion.

3.1 ℓ_2 Minimization

Here, we discuss the reconstruction of the given query face in the target subspace using the ℓ_2 minimization. We first build a subspace of faces with target ethnicity and find the linear combination of basis vectors from this subspace that matches the query image the closest. The weight vector $\mathbf{w}$ for the images spanning the subspace is found by minimizing the ℓ_2-norm.

Let $\mathbf{R}$ be a matrix of dimensions $d \times n$ where d is the number of pixels in each face image and n is the number of images spanning the subspace. In other words, each column of $\mathbf{R}$ is a vectorized image of a face from the target ethnic group. Let $\mathbf{x}$ be an incoming query image which is resized to the same size as the face images in the subspace and vectorized. Let $\mathbf{x}^*$ be the reconstructed image using the subspace $\mathbf{R}$. Let $\mathbf{w}^*$ be the $n \times 1$ array of optimal weights for each image in the subspace. The equations used in reconstruction are shown below:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \|\mathbf{x} - \mathbf{R}\mathbf{w}\|_2^2 = (\mathbf{R}^\top \mathbf{R})^{-1}\mathbf{R}^\top \mathbf{x} \tag{1}$$

$$\mathbf{x}^* = \mathbf{R}\mathbf{w}^* \tag{2}$$

3.2 Principal Component Analysis

In this part, we outline a PCA based synthesis of a facial image from the target ethnic group. The subspace is built using the same samples as in the ℓ_2 minimization method. PCA is applied to the data and the eigenvectors obtained from the subspace are used for projection and reconstruction. The matrix $\mathbf{R}$, with which the PCA subspace is built is of dimension $d \times n$ where d is the number of pixels in each image and n is the number of images used to construct the subspace. As before, each column of $\mathbf{R}$ is a vectorized image of the images from target set. Let $\mathbf{V}$ be the matrix of eigenvectors generated after performing PCA on $\mathbf{R}$ and $\boldsymbol{\mu}$ be the mean of the images in $\mathbf{R}$. The following equations show the projection and reconstruction of a query facial image $\mathbf{x}$ using this PCA subspace:

$$\mathbf{w} = (\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{V} \tag{3}$$

$$\mathbf{x}^* = \mathbf{V}\mathbf{w}^\top + \boldsymbol{\mu} \tag{4}$$

3.3 ℓ_1 Minimization

In this subsection, we discuss the application of the ℓ_1 minimization to the problem posed above to obtain a better reconstruction. It is done by using the basis pursuit or the basis pursuit de-noising (BPDN) [2,4,6,22–24,27] and provides a sparse set of weights. This indicates that we use a relatively sparse set of images from the given training set while trying to reconstruct the image with target ethnicity that is closest in an ℓ_2 sense to the query image. In other words, we use the same optimization function as for the ℓ_2 minimization case but add a regularization term which regularizes the ℓ_1-norm of the weights. Let $\mathbf{w}$ be the optimal set of sparse weights for images in the subspace and let $\mathbf{R}$ be the matrix containing images in the subspace. The modified optimization function is shown below:

$$\text{minimize} \quad \|\mathbf{w}\|_1 \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{R}\mathbf{w}\|_2^2 \leq \epsilon \tag{5}$$

The above optimization can be rewritten using Lagrange multiplier λ as shown below:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \|\mathbf{x} - \mathbf{R}\mathbf{w}\|_2^2 + \lambda\|\mathbf{w}\|_1 \tag{6}$$

For this approach to perform well, we need a large number of examples spanning our subspace since sparse reconstruction typically generalizes well when the dictionary is over-complete. In both the previous approaches, the lack of a constraint on $\mathbf{x}$ distributes the energy of $\mathbf{w}$ across a large number of its elements which results in a lot of artifacts left behind primarily from the boundaries of the face where there is a sharp change in intensity. The reconstruction is also more blurred. The ℓ_1 minimization is expected to perform better than the ℓ_2 minimization and the PCA reconstruction since it avoids these artifacts left behind and also produces a smoother reconstruction than the ℓ_2 minimization and PCA reconstruction.

4 Dual Subspace Modeling for Facial Ethnic Appearance Synthesis Using Fukunaga Koontz Transform

Ever since Fukunaga Koontz transform (FKT) came along in 1970 [3], it has been widely used for feature selection especially for general pattern recognition and image processing problems. Unlike traditional principal component analysis (Karhunen-Loève transform), the FKT incorporates data form both positive and negative classes and using eigen decomposition on the joint covariance matrix, in order to find the optimal basis vectors that very well represent one class while have least representation power on the other class. The intrinsic nature of the FKT formulation makes it a very good feature selection tool for two-class problems. More recently, Li *et al.* [21] managed to generalize the FKT to be applied to multi-class problems. We start with the basics of the FKT and We will take a further step in the FKT analysis to explore some very nice properties of the dual subspace modeling which may not be found in other literatures.

Let $\mathbf{X} \in \mathbb{R}^{d \times m}$ be the data set containing the source ethnic facial images, with each column a vectorized image with dimension d. Let $\mathbf{Y} \in \mathbb{R}^{d \times n}$ be the data set containing all the target ethnic facial images. Both $\mathbf{X}$ and $\mathbf{Y}$ are mean removed. The covariance $\mathbf{\Sigma}$ of both the source and target images are the summation of the covariance for each set $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_Y}$. The total covariance matrix $\mathbf{\Sigma}$ is symmetric and can be diagonalized using eigen-decomposition as:

$$\mathbf{\Sigma} = \mathbf{\Sigma_X} + \mathbf{\Sigma_Y} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^\top \tag{7}$$

where $\mathbf{\Phi}$ contains the entire span of eigenvectors of $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$ houses the corresponding eigenvalues of $\mathbf{\Sigma}$ on its diagonal.

Next, a pre-whitening step is applied in the FKT. Both the source and target data are transformed by a pre-whitening matrix $\mathbf{P} = \mathbf{\Phi}\mathbf{\Lambda}^{-\frac{1}{2}}$. So the transformed data $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ becomes:

$$\widehat{\mathbf{X}} = \mathbf{P}^\top \mathbf{X} \quad \text{and} \quad \widehat{\mathbf{Y}} = \mathbf{P}^\top \mathbf{Y} \tag{8}$$

Therefore, the covariance matrices of the transformed source data $\widehat{\mathbf{X}}$ and target data $\widehat{\mathbf{Y}}$ become:

$$\mathbf{\Sigma}_{\widehat{\mathbf{X}}} = \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top = \mathbf{P}^\top \mathbf{X}\mathbf{X}^\top \mathbf{P} = \mathbf{P}^\top \mathbf{\Sigma}_{\mathbf{X}} \mathbf{P} \tag{9}$$

$$\mathbf{\Sigma}_{\widehat{\mathbf{Y}}} = \widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^\top = \mathbf{P}^\top \mathbf{Y}\mathbf{Y}^\top \mathbf{P} = \mathbf{P}^\top \mathbf{\Sigma}_{\mathbf{Y}} \mathbf{P} \tag{10}$$

The transformed covariance matrix for both source and target data becomes:

$$\widehat{\mathbf{\Sigma}} = \mathbf{\Sigma}_{\widehat{\mathbf{X}}} + \mathbf{\Sigma}_{\widehat{\mathbf{Y}}} = \mathbf{P}^\top \mathbf{\Sigma}_{\mathbf{X}} \mathbf{P} + \mathbf{P}^\top \mathbf{\Sigma}_{\mathbf{Y}} \mathbf{P} \tag{11}$$

$$= \mathbf{P}^\top (\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{Y}}) \mathbf{P} = \mathbf{P}^\top \mathbf{\Sigma} \mathbf{P} = \mathbf{I} \tag{12}$$

So, the new covariance matrix is actually an identify matrix. This is because we have performed a global pre-whitening transformation instead of a class-specific pre-whitening transformation to de-correlate the data.

Here, we again perform an eigen-decomposition on the source covariance $\mathbf{\Sigma}_{\widehat{\mathbf{X}}}$, which yields:

$$\mathbf{\Sigma}_{\widehat{\mathbf{X}}} \mathbf{w} = \lambda \mathbf{w} \tag{13}$$

From Equation 12 we can obtain the following by multiplying $\mathbf{w}$ on both sides:

$$\mathbf{\Sigma}_{\widehat{\mathbf{X}}} \mathbf{w} + \mathbf{\Sigma}_{\widehat{\mathbf{Y}}} \mathbf{w} = \mathbf{w} \tag{14}$$

With Equation 12 and 14, we have:

$$\mathbf{\Sigma}_{\widehat{\mathbf{Y}}} \mathbf{w} = \mathbf{w} - \mathbf{\Sigma}_{\widehat{\mathbf{X}}} \mathbf{w} = \mathbf{w} - \lambda \mathbf{w} = (1 - \lambda)\mathbf{w} \tag{15}$$

This effectively means that the covariance matrix from two classes share the same eigenvectors $\mathbf{w}$ and the eigenvalue of one class is exactly the complement of the eigenvalue of the other class. Because of the complementary property of the eigenvalues, the eigenvectors that are the most dominant in one class, is the least dominant in the other class. So in the traditional FKT method as applied to any two-class problem, a discriminative subspace is created by selecting a few of the most dominant eigenvectors for one class and the least dominant ones for the other class. By ignoring the eigenvectors in the middle range, the subspace we obtain contains basis that are very discriminative and will yield discriminative feature selection after projection as shown in Figure 2.

4.1 Dual Subspace Modeling

The aforementioned case is only true for the ideal scenario where the covariance matrices $\mathbf{\Sigma}_{\widehat{\mathbf{X}}}$ and $\mathbf{\Sigma}_{\widehat{\mathbf{Y}}}$ are full rank and have non-zero eigenvalues. But in the

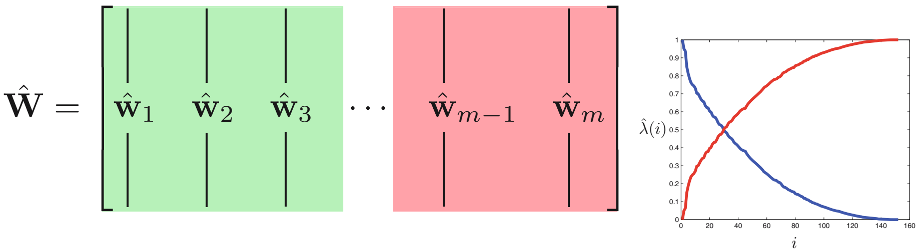


Fig. 2. (left) The most and least dominant vectors for one class are used for better classifying two classes; (right) complementary property of the eigenvalues, where they sum up to 1

real application, this is seldom the case. When the covariance matrix is not full rank, there will be k least dominant eigenvectors in class 1 that all have 0 eigenvalues. According to the FKT, there will be k most dominant eigenvectors with eigenvalues being 1. That means, in the eigenvector $\mathbf{w}$, there are $2k$ eigenvectors that are not properly ranked, and thus the complementary paring of eigenvalues and the sharing of eigenvectors is no longer valid.

In this case, instead of decomposing the covariance of only one class, we propose to decompose both the covariance matrices $\boldsymbol{\Sigma}_{\widehat{\mathbf{X}}}$ and $\boldsymbol{\Sigma}_{\widehat{\mathbf{Y}}}$, and instead of modeling both covariance matrices using the same eigenvector $\mathbf{w}$, we propose a dual subspace model using class-specific eigenvectors $\mathbf{w}_x$ and $\mathbf{w}_y$ for decomposition:

$$\boldsymbol{\Sigma}_{\widehat{\mathbf{X}}}\mathbf{w}_x = \lambda_x \mathbf{w}_x \tag{16}$$

$$\boldsymbol{\Sigma}_{\widehat{\mathbf{Y}}}\mathbf{w}_y = \lambda_y \mathbf{w}_y \tag{17}$$

The complementary relationship now becomes:

$$\boldsymbol{\Sigma}_{\widehat{\mathbf{Y}}} = \mathbf{w}_x - \boldsymbol{\Sigma}_{\widehat{\mathbf{X}}} = \mathbf{w}_x - \lambda_x \mathbf{w}_x = (1-\lambda_x)\mathbf{w}_x \tag{18}$$

$$\boldsymbol{\Sigma}_{\widehat{\mathbf{X}}} = \mathbf{w}_y - \boldsymbol{\Sigma}_{\widehat{\mathbf{Y}}} = \mathbf{w}_y - \lambda_y \mathbf{w}_y = (1-\lambda_y)\mathbf{w}_y \tag{19}$$

Here, the most dominant eigenvector in $\mathbf{w}_x$ and the least dominant eigenvectors in $\mathbf{w}_y$ are not necessarily the same. Instead of keeping the first and the last tier of eigenvectors from $\mathbf{w}$ for subspace modeling, we now take the most dominant eigenvectors from $\mathbf{w}_x$ as well as from $\mathbf{w}_y$ to create the subspace. In this way, we essentially remove the eigenvectors corresponding to 0 eigenvalues, while still keeping high discriminative power.

5 Experiments

In this section, we first describe the experimental setup for the aforementioned facial ethnic appearance synthesis techniques: (1) the ℓ_2 minimization, (2) PCA reconstruction, (3) the ℓ_1 minimization, and (4) the dual subspace modeling using FKT. Second, we show and analyze the experimental results using all four synthesis techniques.

5.1 Experimental Setup

For the single subspace reconstruction methods (the ℓ_2 minimization, PCA reconstruction, and the ℓ_1 minimization), only the subspace obtained from the target ethnic group is needed in the reconstruction process. The query image from the source ethnic group is reconstructed and synthesized to best match the target ethnic group.

In the dual subspace reconstruction method using FKT, two subspaces from both the source and target ethnic group are acquired. By linearly combining the two subspaces using α blending, the query image from the source ethnic group is gradually transformed to the target ethnic group.

All the color images are in the **RGB** format, so the facial ethnic appearance synthesis is done by reconstructing each color channel individually and finally combined together to display the color synthesis images.

One important characteristic of our proposed facial ethnic appearance synthesis is that the subject identity is very well preserved during the synthesis, and at the same time, the subject's facial ethnic features such as skin tone and eye contours are altered to best match the target ethnic group. In this way, we can transform people from one ethnic group to another, by keeping their own uniqueness. This is very good in the application of synthesizing new subjects from particular ethnic groups that are unique.

5.2 Experimental Results

We have trained our optimal projection coefficient $\mathbf{w}^*$ for each of the single subspace methods using the images only from the target ethnic group, and synthesize the facial appearance of the query image that best matches the target group.

Figure 3 and 4 show the synthesis using the ℓ_2 minimization. In this experiment, we pick query images of celebrities from the Internet and synthesize them to another ethnic group in terms of: (a) black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian, for both genders.

As can be seen, the reconstruction is of bad quality with many artifacts. This is partially due to the fact that the number of images in the target set is limited and when an unseen query image looks quite different from the images in the target set, the reconstruction would be jeopardized.

Figure 5 and 6 show the synthesis using PCA reconstruction. The same query images are selected as in the ℓ_2 minimization case to show the comparisons. The reconstruction, still not satisfactory. The results using PCA look similar to the ones using the ℓ_2 minimization because the way PCA finds the optimal projecting directions (principal components) is actually minimizing the variance: $\mathrm{Var}(\mathbf{w}^\top \mathbf{x})$, and both PCA and ℓ_2 minimization is dealing with second-order statistics of the data. Thus their optimal results are similar.

Figure 7 and 8 show the synthesis using the ℓ_1 minimization. The same query images and target ethnic groups are used. We can see that the reconstruction still is not as good as we have expected. Many of the synthesized faces are bluish.

(a) (b) (c) (d)

Fig. 3. Facial ethnic appearance synthesis using ℓ_2 minimization on female subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.

(a) (b) (c) (d)

Fig. 4. Facial ethnic appearance synthesis using ℓ_2 minimization on male subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.

This is because the ℓ_1 minimization gives a sparse solution, and since we are dealing each color channel individually, the **R** and **G** channel are overwhelmed by the **B** channel. Moreover, some target ethnic set does not have enough images to create an over-complete dictionary so the reconstruction quality using the ℓ_1 minimization is still questionable.

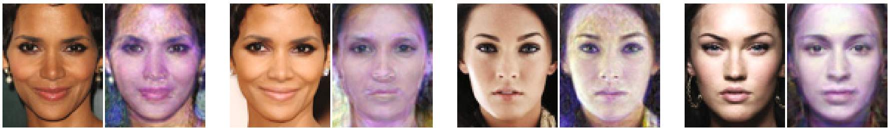
By applying our proposed dual subspace modeling using FKT, the facial ethnic appearance synthesis results are much better than the previously discussed single subspace methods. Figure 9 and 10 show the synthesis. The query images are on the far left and the following five images are reconstructed using the dual subspaces (one subspace is built using target ethnic group, and the other sub-

(a) (b) (c) (d)

Fig. 5. Facial ethnic appearance synthesis using PCA reconstruction on female subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.

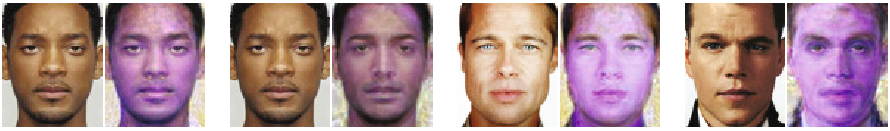
(a) (b) (c) (d)

Fig. 6. Facial ethnic appearance synthesis using PCA reconstruction on male subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.



(a) (b) (c) (d)

Fig. 7. Facial ethnic appearance synthesis using ℓ_1 minimization on female subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.



(a) (b) (c) (d)

Fig. 8. Facial ethnic appearance synthesis using ℓ_1 minimization on male subjects. (a) Black to east asian, (b) black to south asian, (c) white to east asian, and (d) white to south asian. In each subfigure, the input is on the **left** and the reconstructed image is on the **right**.

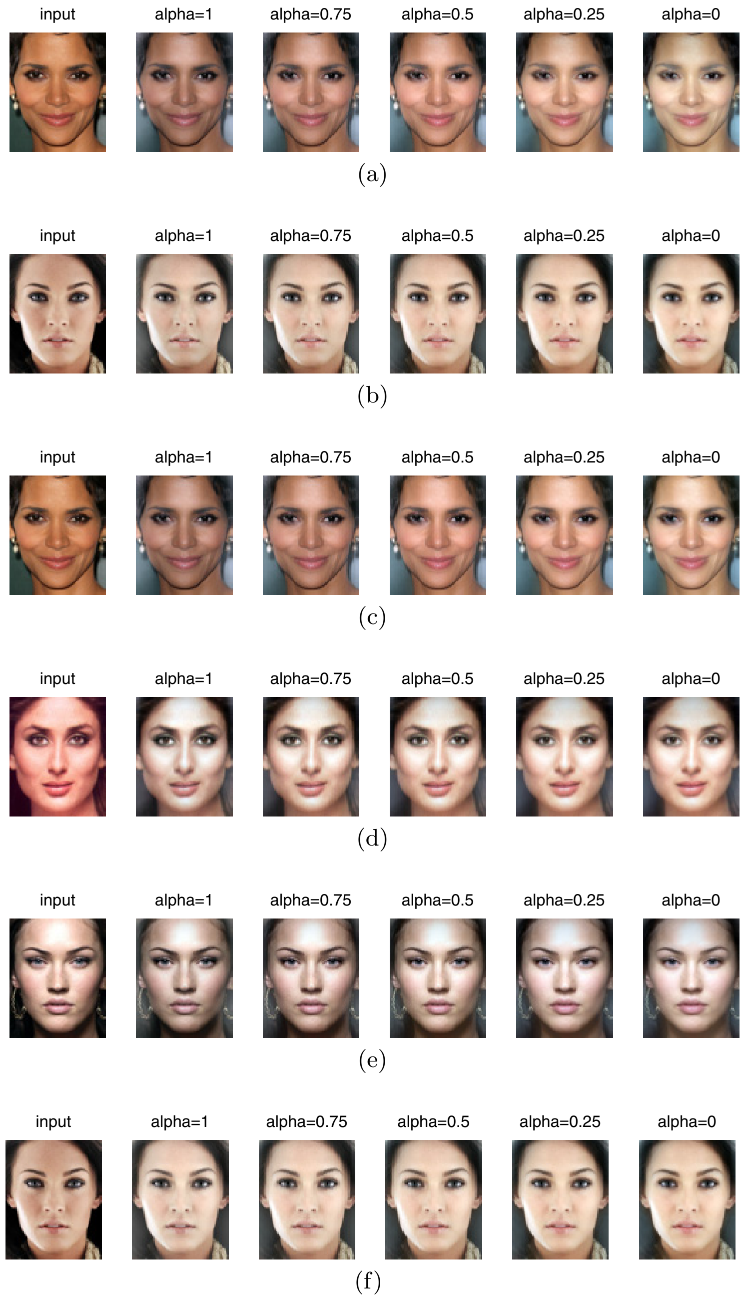


Fig. 9. Facial ethnic appearance synthesis using FKT with dual subspace modeling on female subjects. (a) Black to east asian, (b) black to south asian, (c) black to white, (d) south asian to east asian, (e) white to east asian, and (f) white to south asian. In each subfigure, the input is on the far left and the following five images are the synthesis with α blending of the dual subspaces.

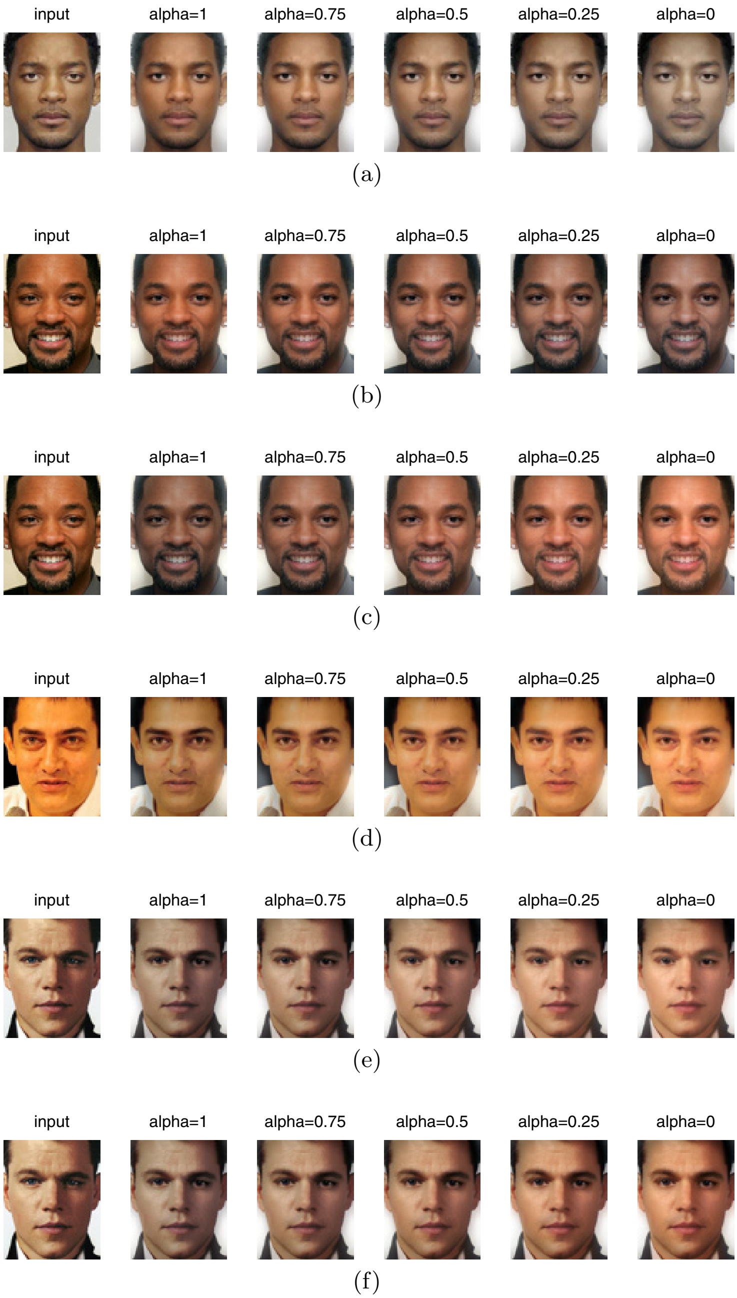


Fig. 10. Facial ethnic appearance synthesis using FKT with dual subspace modeling on male subjects. (a) Black to east asian, (b) black to south asian, (c) black to white, (d) south asian to east asian, (e) white to east asian, and (f) white to south asian. In each subfigure, the input is on the far left and the following five images are the synthesis with α blending of the dual subspaces.

space is built using the source ethnic group, the same as the query image to be ethnically altered). The α blending coefficient is shown in the 2 figures. When $\alpha = 1$, pure source group subspace is utilized and when $\alpha = 0$, pure target group subspace is used. By changing α from 1 to 0, a gradual transformation from the source ethnic group to the target ethnic group can be shown. The synthesis quality of FKT dual subspace modeling is much better than the single subspace method with no identifiable artifacts at all.

The query image is transformed to the target ethnic group while keeping his or her identity to the largest extent, meaning the distinctive identity features are well preserved. As are shown in Figure 9 and 10, the photometric features of the query images are also well kelp. For example, the highlights on the forehead and the shadow on the cheek and so forth are still well preserved in the ethnic appearance synthesized image.

From the mean faces of our database as shown earlier in Figure 1, the eye region is the best registered region on the faces. Compared with eye region, the mouth region is not as well aligned. This is due to the fact that images in our database may have different expression and the mouth region are not perfectly aligned and registered. Therefore, we should expect a better synthesis quality around the eye region and less around the mouth region. Even with that, we are still able to achieve a much better ethnic appearance synthesis than the single subspace techniques such as the ℓ_2 minimization, PCA and the ℓ_1 minimization.

In our experiments with the FKT dual subspace modeling, we apply our facial ethnic appearance synthesis to unseen facial images from the web. These celebrity images are actually very different from the images in our source and target database. The skin is usually highly polished due to makeups and photo re-touching. So, the synthesis results is not as genuine as the query images that are actually from the source data set. We cannot disclose the query images from our database in this paper, but the supplementary material available to the reviewers actually show more genuine results.

6 Conclusions

In this work, we have explored several subspace reconstruction methods for facial ethnic appearance synthesis (FEAS). In our experiments, our proposed dual subspace modeling using the Fukunaga Koontz transform (FKT) yields much better facial ethnic synthesis results than the ℓ_1 minimization, the ℓ_2 minimization and the principal component analysis (PCA) reconstruction method. With that, we are able to automatically and efficiently synthesize different facial ethnic appearance and alter the facial ethnic appearance of the query image to any other ethnic appearance as desired. Our technique well preserves the facial structure of the query image and simultaneously synthesize the skin tone and ethnic features that best matches target ethnicity group. Facial ethnic appearance synthesis can be applied to synthesizing facial images of a particular ethnicity group for unbalanced database, and can be used to train ethnicity invariant classifiers by generating multiple ethnic appearances of the same subject in the training stage.

References

1. Dhamecha, T.I., Sankaran, A., Singh, R., Vatsa, M.: Is gender classification across ethnicity feasible using discriminant functions? In: 2011 International Joint Conference on Biometrics (IJCB), pp. 1–7 (September 2011)
2. Ekanadham, C., Tranchina, D., Simoncelli, E.P.: Recovery of sparse translation-invariant signals with continuous basis pursuit. IEEE Transactions on Signal Processing **59**(10), 4735–4744 (2011)
3. Fukunaga, K., Koontz, W.L.G.: Application of the karhunen-loève expansion to feature selection and ordering. IEEE Transactions on Computers **C-19**(4), 311–318 (1970)
4. Gill, P.R., Wang, A., Molnar, A.: The in-crowd algorithm for fast basis pursuit denoising. IEEE Transactions on Signal Processing **59**(10), 4595–4605 (2011)
5. Guo, G., Mu, G.: A study of large-scale ethnicity estimation with gender and age variations. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), pp. 79–86 (June 2010)
6. Hoang, T.V., Smith, E.H.B., Tabbone, S.: Edge noise removal in bilevel graphical document images using sparse representation. In: 2011 18th IEEE International Conference on Image Processing (ICIP), pp. 3549–3552 (September 2011)
7. Hosoi, S., Takikawa, E., Kawade, M.: Ethnicity estimation with facial images. In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG), pp. 195–200 (May 2004)
8. Jain, A.K., Park, U.: Facial marks: Soft biometric for face recognition. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 37–40 (November 2009)
9. Juefei-Xu, F., Bhagavatula, C., Jaech, A., Prasad, U., Savvides, M.: Gait-id on the move: pace independent human identification using cell phone accelerometer dynamics. In: 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 8–15. IEEE (2012)
10. Juefei-Xu, F., Cha, M., Heyman, J.L., Venugopalan, S., Abiantun, R., Savvides, M.: Robust local binary pattern feature sets for periocular biometric identification. In: 2010 Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS), pp. 1–8. IEEE (2010)
11. Juefei-Xu, F., Cha, M., Savvides, M., Bedros, S., Trojanova, J.: Robust periocular biometric recognition using multi-level fusion of various local feature extraction techniques. In: IEEE 17th International Conference on Digital Signal Processing (DSP) (2011)
12. Juefei-Xu, F., Luu, K., Savvides, M., Bui, T.D., Suen, C.Y.: Investigating age invariant face recognition based on periocular biometrics. In: 2011 International Joint Conference on Biometrics (IJCB), pp. 1–7. IEEE (2011)
13. Juefei-Xu, F., Pal, D.K., Savvides, M.: Hallucinating the full face from the periocular region via dimensionally weighted K-SVD. In: IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW). IEEE (2014)
14. Juefei-Xu, F., Savvides, M.: Can your eyebrows tell me who you are? In: 2011 5th International Conference on Signal Processing and Communication Systems (ICSPCS), pp. 1–8. IEEE (2011)
15. Juefei-Xu, F., Savvides, M.: Unconstrained periocular biometric acquisition and recognition using COTS PTZ camera for uncooperative and non-cooperative subjects. In: 2012 IEEE Workshop on Applications of Computer Vision (WACV), pp. 201–208. IEEE (2012)

16. Juefei-Xu, F., Savvides, M.: An augmented linear discriminant analysis approach for identifying identical twins with the aid of facial asymmetry features. In: IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW) (2013)
17. Juefei-Xu, F., Savvides, M.: An image statistics approach towards efficient and robust refinement for landmarks on facial boundary. In: 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–8. IEEE (2013)
18. Juefei-Xu, F., Savvides, M.: Subspace based discrete transform encoded local binary patterns representations for robust periocular matching on NIST's face recognition grand challenge. IEEE Transactions on Image Processing **23**(8), 3490–3505 (2014)
19. Juefei-Xu, F., Savvides, M.: Weight-optimal local binary patterns. In: European Conference on Computer Vision (ECCV) Workshops. Springer (2014)
20. Kashyap, A.L., Tulyakov, S., Govindaraju, V.: Facial behavior as a soft biometric. In: 2012 5th IAPR International Conference on Biometrics (ICB), pp. 147–151 (April 2012)
21. Li, Y., Savvides, M.: Kernel fukunaga-koontz transform subspaces for enhanced face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007, pp. 1–8 (June 2007)
22. Liu, C., Zakharov, Y.V., Chen, T.: Broadband underwater localization of multiple sources using basis pursuit de-noising. IEEE Transactions on Signal Processing **60**(4), 1708–1717 (2012)
23. Lu, W., Vaswani, N.: Modified basis pursuit denoising (modified-BPDN) for noisy compressive sensing with partially known support. In: 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 3926–3929 (March 2010)
24. Lu, W., Vaswani, N.: Exact reconstruction conditions for regularized modified basis pursuit. IEEE Transactions on Signal Processing **60**(5), 2634–2640 (2012)
25. Lyle, J.R., Miller, P.E., Pundlik, S.J., Woodard, D.L.: Soft biometric classification using periocular region features. In: 2010 Fourth IEEE International Conference on Biometrics: Theory Application and Systems (BTAS), pp. 1–7 (September 2010)
26. Manesh, F.S., Ghahramani, M., Tan, Y.P.: Facial part displacement effect on template-based gender and ethnicity classification. In: 2010 11th International Conference on Control Automation Robotics and Vision (ICARCV), pp. 1644–1649 (December 2010)
27. Mota, J.F.C., Xavier, J.M.F., Aguiar, P.M.Q., Puschel, M.: Distributed basis pursuit. IEEE Transactions on Signal Processing **60**(4), 1942–1956 (2012)
28. Niinuma, K., Park, U., Jain, A.K.: Soft biometric traits for continuous user authentication. IEEE Transactions on Information Forensics and Security **5**(4), 771–780 (2010)
29. Park, U., Jain, A.K.: Face matching and retrieval using soft biometrics. IEEE Transactions on Information Forensics and Security **5**(3), 406–415 (2010)
30. Reid, D.A., Nixon, M.S.: Using comparative human descriptions for soft biometrics. In: 2011 International Joint Conference on Biometrics (IJCB), pp. 1–6 (October 2011)
31. Savvides, M., Juefei-Xu, F.: Image matching using subspace-based discrete transform encoded local binary patterns (September 2013). http://www.google.com/patents/US20140212044

32. Seshadri, K., Savvides, M.: An analysis of the sensitivity of active shape models to initialization when applied to automatic facial landmarking. IEEE Transactions on Information Forensics and Security **7**(4), 1255–1269 (2012)
33. Wu, B., Ai, H., Huang, C.: Facial image retrieval based on demographic classification. In: 17th International Conference on Pattern Recognition (ICPR), vol. 3, pp. 914–917 (August 2004)
34. Yan, S., Wang, H., Tang, X., Liu, J., Huang, T.S.: Regression from uncertain labels and its applications to soft biometrics. IEEE Trans. on Information Forensics and Security **3**(4), 698–708 (2008)
35. Zhang, G., Wang, Y.: Multimodal 2D and 3D facial ethnicity classification. In: 2009 Fifth International Conference on Image and Graphics (ICIG), pp. 928–932 (September 2009)