Hallucinating the Full Face from the Periocular Region via Dimensionally Weighted K-SVD

Felix Juefei-Xu, Dipan K. Pal, and Marios Savvides
CyLab Biometrics Center, Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213, USA
felixu@cmu.edu, dipanp@cmu.edu, msavvid@ri.cmu.edu

Abstract

Identifying a suspect wearing a mask (where only the suspect's periocular region is visible) is one of the toughest real-world challenges in biometrics that exist. In this paper, we present a practical method to hallucinate the full frontal face given only the periocular region of a face. This is an important problem faced in many law-enforcement applications on almost a daily basis. In such real-world situations, we ***only*** *have access to the periocular region of a person's face. Unfortunately commercial matchers are unable to process these images successfully. We propose in this paper, an approach that will reconstruct the entire frontal face using* ***just*** *the periocular region. We empirically show that our reconstruction technique, based on a modified sparsifying dictionary learning algorithm, can effectively reconstruct faces which we show are actually very similar to the original ground-truth faces. Further, our method is open set, thus can reconstruct any face* ***not seen*** *in training. We show the real-world applicability of method by benchmarking face verification results using the reconstructed faces to show that they still match competitively compared to the original faces when evaluated under a large-scale face verification protocol such as NIST's FRGC protocol where over 256 million face matches are made.*

1. Introduction

Over the past decades, biometric identification and verification using facial features has gained a lot of prominence both in traditional video surveillance/access control systems and in hand-held devices for daily use. Most of these approaches work under the implicit assumption that we are able to capture the entire face of the subject with decent quality. However, there are many real-world scenarios where only a partial face is captured or instances when only the eye region of a face is visible, especially for the cases of uncooperative and non-cooperative subjects. Therefore, the problem of looking into the capabilities of matching subjects using only the periocular region has developed considerable interest. Specifically, we consider the periocular region of the face, which is rich in textural information - eyebrows, eye folds, eyelid contours, *etc*., which could all vary in shape, size and color. Biologically and genetically speaking, more complex structure means more "coding processing" going on during fetal development, and therefore more proteins and genes involved in the determination of appearance. One can speculate that this is why the periocular region should be the most important facial area for distinguishing people. Robust periocular based biometric recognition can lead to very useful applications, for example, identifying criminals wearing masks, where only the eye region is exposed, or in videos containing many occluded faces with the eye region un-occluded, or in other cases as shown in Figure 1.

In addition, commercial matchers, and law enforcement agencies who rely on commercial matchers to perform face matching for identification will run into problems in the case where only the periocular region is available. This is due to the fact that commercial matching algorithms are developed using the entire human face and typically simply cannot deal with partial faces.

In this work, we develop a novel approach that hallucinates the full face from just the periocular region of a subject with high fidelity devoted to the known periocular region. The approach is based on the modification of the problem formulated for sparsely coded dictionary learning. It explicitly focuses on reconstructing the periocular region faithfully while providing a good visual approximation of facial features that can be used for further processing. In hallucinating the rest of the face, our method capitalizes on weak correlations between periocular features and other facial features. These correlations might exist due to specific gender, ethnicity or age, which are soft-biometric in nature. Our approach finds these relations in an unsupervised manner from a large corpus of frontal training images.

Rest of this paper is organized as follows: Section 2 lists

Figure 1. **Examples of scenarios where only periocular region is accessible.** (a1) A suspect in an ATM robbery wearing a mask, (a2) A masked Taliban militant, and (a3) a masked bank robber. (b1)-(b3) Many people in crowds only have their eyes visible from the camera's perspective. (c1) Veiled women, (c2) masked doctors and nurses, and (c3-c4) masked fire fighter and policeman.

several prior work on periocular region recognition. Section 3 details related algorithms and the proposed method. Section 4 discusses in detail the experimental setup and results. Finally we conclude our work in Section 5. Throughout the paper, the term "hallucination" and "reconstruction" are used interchangeably.

2. Related Work

In 2009, Park *et al.* [17] has carried out one of the earliest studies on periocular biometrics identification where the feasibility of using periocular region of an individual as a biometric trait were claimed. They showed 77% rank-1 identification rate on a rather small database (958 images from 30 subjects). In the following year, more studies were carried out on periocular region. Juefei-Xu *et al.* [5] evaluated the performance of periocular biometrics on a large scale FRGC ver2.0 database. They proposed various local feature sets and showed that even without any subspace training, periocular region with their proposed feature sets can still outperform NIST's baseline using PCA on full face on the FRGC Experiment 4 protocol. More detailed analysis were shown in [6, 11]. Also the discriminability of eyebrow was also looked into [8]. Lyle *et al.* [14] studied gender and ethnicity classification based on the periocular features. They used a subset of FRGC and obtained a high classification accuracy for gender and ethnicity. The effect of the quality of the periocular images on recognition performance was studies in [15] where they analyzed the uniqueness of texture between different color channels, and texture information present in different color channels. Woodard *et al.* [19] utilized periocular region appearance cues for biometric identification both on images captured in visible and NIR spectrum while Park *et al.* [16] studied periocular biometrics in the visible spectrum. Hollingsworth *et al.* [3] used NIR periocular images to identify useful features for recognition, while others fused periocular with iris images for recognition [18]. Some more recent work using periocular region include age invariant face recognition [7] and expression tolerance [9] based on periocular region and twin identification using periocular region [10]. To the best of our knowledge, the problem of hallucinating the entire face using only the periocular region has not been approached yet.

3. Algorithmic Approach

In this section we describe our proposed approach to the problem of hallucinating a complete face based on purely the periocular region. However, before we introduce our method, it would helpful to briefly look at the PCA based reconstruction in the context of this problem. Throughout the paper, the data matrix $\mathbf{Y} \in \mathbb{R}^{d \times n}$ is assumed with dimension d. All matrices have their elements arranged columnwise.

3.1. PCA Based Hallucination

PCA has proved to be a hugely popular subspace learning method over the years. It has also found many applications in denoising. In our application, to hallucinate the entire face based on the periocular region, PCA can be applied in a straight forward way. Assume that $\mathbf{D}$ is the global PCA basis of the full face data. We now assume that the periocular region can be obtained by using a mask Λ which is the set of particular dimensions from an image belonging to that region. Given an unseen periocular image of a subject $\mathbf{y}_\Lambda$, our goal is to obtain $\mathbf{y}$. We obtain the PCA projection coefficients $\mathbf{x} = (\mathbf{D}_\Lambda^T \mathbf{D}_\Lambda)^{-1} \mathbf{D}_\Lambda \mathbf{y}_\Lambda$. Here, $\mathbf{D}_\Lambda$ is the dictionary restricted to dimensions or rows of the matrix in the set Λ. Finally we obtain the reconstruction $\mathbf{y}$ using $\mathbf{y} = \mathbf{D}\mathbf{x}$. Note that during reconstruction, we use all dimensions of $\mathbf{D}$. Even though PCA provides a simple approach to this problem, since it learns a single global subspace, each testing sample would tend to have a very similar reconstruction. Thus, very little biometric information is preserved in the reconstruction rendering the problem unsolved.

3.2. K-SVD Based Hallucination

Dictionary learning methods have gained much popularity in the recent decade. One relatively recent such algorithm is the K-SVD [1]. K-SVD aims to be a natural extension of K-means with the analogy that the cluster centers are the elements of the learned dictionary and the memberships are defined by the sparse approximations of the signals in that dictionary. Formally, it provides a solution to the problem $\text{minimize}_{\mathbf{D},\mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2$ such that $\forall i, \|\mathbf{x}_i\|_0 < K$, where $\mathbf{Y}$, $\mathbf{D}$ and $\mathbf{X}$ are the data, the learned dictionary and the sparse approximation matrix respectively. Here $\|.\|_0$ is the pseudo-norm measuring sparsity. The sparse approximations of the data elements are allowed to have some maximum sparsity $\|\mathbf{x}\|_0 \leq K$.

K-SVD has found multiple practical applications such as image denoising [2] and inpainting [12]. However, these applications are ones in which the number of missing or corrupted pixels is not significant and does not contain any explicit spatial structure. In such a scenario, patch based reconstructions using the learned dictionary would be useful. In our application however, where majority of the pixels are missing in a structured manner, patch-based reconstruction does not make sense. Rather, we would have to apply an approach similar in spirit to the PCA based reconstruction.

One method for hallucination using K-SVD is to train a generative dictionary $\mathbf{D}$ using a large number of full faces. Each of the dictionary elements is now the first eigenface of its member training samples. Unlike PCA, K-SVD avoids learning a global subspace and in turn approximates using multiple local subspaces. Given a novel periocular image of an unseen subject $\mathbf{y}_\Lambda$, we would obtain the sparse coding $\mathbf{x}$ in the dimensionally restricted $\mathbf{D}_\Lambda$ using any sparse coding algorithm such as OMP (Orthogonal Matching Pursuit). For reconstruction, we simply use $\mathbf{x}$ with the original dictionary $\mathbf{D}$ to obtain $\mathbf{y} = \mathbf{D}\mathbf{x}$.

However, one assumes a critical fact in this method that the sparse representation of the periocular region alone in $\mathbf{D}_\Lambda$ is very similar to the representation of the entire face in $\mathbf{D}$. This is a questionable assumption since the dictionary learning procedure K-SVD does not explicit optimize in that regard. Indeed, faces can be considered as an ensemble of features, with multiple individuals sharing a particular feature. The assumption of common approximation coefficients between the periocular region and the entire face in a dictionary learning full faces would imply that there exists a one-to-one relationship between the eye features and the other features of a face. Even though the algorithm for our hallucination problem would have to find weak correlations between periocular and facial features, a one-to-one correspondence is too strong an assumption and is unreasonable.

Recall that our problem is to hallucinate, for the sake of visual and practical purposes, the entire face from only the periocular region. Thus, the only true biometric that we observe is the given cropped periocular image. In this light, it is vital that our reconstructed face be faithful in the periocular region. However, standard methods of generating dictionaries, such as the typical use of K-SVD, do not focus on representing particular dimensions or parts of the signal better. A method which weighs errors due to the periocular region more than the rest of the face would tend to generate a reconstruction more faithful (higher PSNR) to that region.

Here one might argue that a simple get-around to this problem would be to train two dictionaries separately, one representing the full face $\mathbf{D}_f$ and the other trained specifically for the periocular region $\mathbf{D}_p$. Thereby, we can specifically optimize for a low reconstruction error in the periocular region. We could then follow a similar protocol for reconstruction by obtaining the sparse representation $\mathbf{x}$ of $\mathbf{y}_\Lambda$ in $\mathbf{D}_p$ and then reconstruct using $\mathbf{y} = \mathbf{D}_f\mathbf{x}$. However, since the training of the two dictionaries is independent, there is no reason to hope that the K-sparse representation $\mathbf{x}$ of $\mathbf{y}_\Lambda$ in $\mathbf{D}_p$ is close to the that of $\mathbf{y}$ in $\mathbf{D}_f$. This is the same problem that we highlighted previously. Reconstruction using such a procedure is not expected to give good and visually appealing results.

This problem can be addressed by designing a dictionary learning procedure which tries to have a consistent sparse representation across the two dictionaries $\mathbf{D}_p$ and $\mathbf{D}_f$ while weighting errors in $\mathbf{D}_p$ more. Such a method would also address the problems that have been brought to light in previous paragraphs. In the next section, we present a simple reformulation of the objective function to arrive at one such procedure.

3.3. Dimensionally Weighted K-SVD Based Hallucination

Our goal is to reconstruct or hallucinate the rest of the face given the periocular region. Keeping in mind the issues related to the dictionary learning, we arrive at the problem of jointly optimizing the learning procedure for the two goals. The first is to learn a dictionary of whole faces so as to include prior knowledge about the spatial relationships between the facial features and the periocular features. The second is to obtain a dictionary in which the reconstruction error for the periocular region is penalized more than the entire face and both are jointly minimized for the same sparse coefficients.

We propose a simple approach which promotes the approximation coefficients to be jointly shared for the periocular region and the entire face. Our first objective is to learn a dictionary by solving

$$\underset{\mathbf{D},\mathbf{X}}{\text{minimize}} \ \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 \ \text{ such that } \ \forall i, \|\mathbf{x}_i\|_0 < K \tag{1}$$

However, we would also like to have a low reconstruction error using the same sparse coefficients restricted to the periocular region set Λ. Thus we also desire to solve

$$\underset{\mathbf{D}_\Lambda,\mathbf{X}}{\text{minimize}} \ \|\mathbf{Y}_\mathbf{\Lambda} - \mathbf{D}_\mathbf{\Lambda}\mathbf{X}\|_F^2 \ \text{ such that } \ \forall i, \|\mathbf{x}_i\|_0 < K \tag{2}$$

Combining the two objectives to solve them jointly allows us to force a common K-sparse representation and also provides a trade-off between errors with an efficient algorithmic solution. Our primary problem is therefore

$$\underset{\mathbf{D},\mathbf{X},\mathbf{D}_\Lambda}{\arg\min} \ \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 + \beta\|\mathbf{Y}_\mathbf{\Lambda} - \mathbf{D}_\mathbf{\Lambda}\mathbf{X}\|_F^2 \tag{3}$$
$$\text{such that } \ \forall i, \|\mathbf{x}_i\|_0 < K$$

Here β provides a trade-off between the reconstruction error of the periocular dimensions versus the entire face.

Obtaining a consistent sparse encoding between the two sets of dimensions allows for a more meaningful reconstruction. This is apparent if one considers the reconstruction procedure. Given a novel periocular image, we would first obtain the sparse representation $\mathbf{x}$ in $\mathbf{D}_\Lambda$. We then obtain the reconstruction using $\mathbf{Dx}$. Using the original K-SVD training method, there was no reason to expect a low reconstruction error in obtaining the entire face. Thus, relationships between periocular and other facial features are not explicitly learned. However, by forcing consistent sparse representations $\mathbf{x}$ during training, we optimize for a low reconstruction error for both regions jointly and simultaneously.

Solving the formulation is achieved by a simple rearrangement before using the standard K-SVD as previously observed [4]:

$$\underset{\mathbf{D},\mathbf{D}_\Lambda,\mathbf{X}}{\arg\min} \left\| \begin{pmatrix} \mathbf{Y} \\ \sqrt{\beta}\mathbf{Y}_\Lambda \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \sqrt{\beta}\mathbf{D}_\Lambda \end{pmatrix} \mathbf{X} \right\|_F^2 \quad (4)$$
$$\text{such that} \quad \forall i, \|\mathbf{x}_i\|_0 \leq K$$

This translates to the standard K-SVD problem where we $\text{minimize}_{\mathbf{D}',\mathbf{X}'} \|\mathbf{Y}' - \mathbf{D}'\mathbf{X}\|_2$ under $\|\mathbf{x}_i\|_0 \leq K$. with $\mathbf{Y}' = (\mathbf{Y}^T, \mathbf{Y}_\Lambda^T)^T$ and $\mathbf{D}' = (\mathbf{D}^T, \mathbf{D}_\Lambda^T)^T$. In effect the formulation is equivalent to re-weighting dimensions belonging to Λ by $(1 + \sqrt{\beta})$. Note that one can easily generalize this framework to include multiple subsets of other dimensions with different weights. Further, this method along with PCA based and K-SVD based methods, is open set thereby enabling reconstruction of any face that is not present in the training set. For convenience, we call this method Dimensionally Weighted K-SVD or DW-KSVD.

4. Experimental Results

4.1. Database

All test experiments were performed on the NIST's Face Recognition Grand Challenge (FRGC) ver2.0 database. It has three components, the first is the generic training set which contains both controlled and uncontrolled images of 222 subjects, and a total of $12,776$ images. Second, the target set containing 466 different subjects with a total of $16,028$ images. Lastly, the probe set containing the same 466 subjects as in target set, with half as many images for each person as in the target set, bringing the total number of probe images to $8,014$. Image examples from the FRGC database are shown in Figure 2.

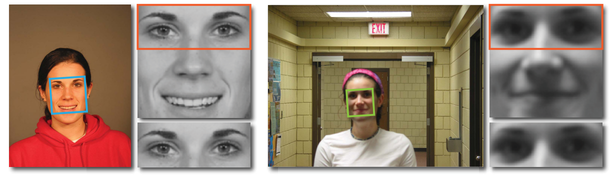

Figure 2. **Example image from the FRGC database:** (a1,a2) controlled and uncontrolled still of the same subject, (b1,b2) cropped full face and periocular region. [11]

4.2. Dictionary Learning and Reconstruction

To learn the dictionary used for reconstruction, we trained using Dimensionally Weighted K-SVD (DW-KSVD) on $500,000$ frontal mugshot images resized to 32 by 32 pixels. A large number of images are necessary to obtain a comprehensive dictionary of weak periocular-to-facial feature relationships. For all experiments, we set the number of dictionary elements to $5,000$ and the maximum allowed sparsity K as 10 to force the dictionary elements to span a smaller local subspace to account for high variation between subjects. We set $\beta = 100$ to strongly emphasize periocular reconstruction. The dictionary was initialized using randomly chosen data elements and K-SVD was run for 20 iterations for learning all dictionaries. We define the periocular region in 32 by 32 images as the top 13 by 32 part of the image. In order to focus our efforts on the reconstruction performance itself, we restrict ourselves from exploring other templates. Our method can handle in a straight-forward way, cases in which the periocular region varies from our defined template.

For reconstruction using DW-KSVD, we first obtain the sparse representation of the periocular image using OMP in the periocular component of the DW-KSVD dictionary. We then reconstruct using the face component of the DW-KSVD dictionary and the same sparse coefficients. Note that we would have a trade-off in choosing sparsity K while using OMP for sparse representation during reconstruction. The reason is that as we increase K, we would keep achieving a lower periocular reconstruction error, however, the full face reconstruction error might increase after a point. This is because OMP is only optimizing for the periocular representation error and not the full face reconstruction error. To learn the optimal reconstruction sparsity for the task, we conduct a pilot experiment in which we measure the PSNR between the unseen original face and the reconstructed face while increasing sparsity. We adopt the peak signal-to-noise ratio (PSNR) as the measurement of reconstruction fidelity between images I and I' as follows: $\text{PSNR} = 10\log_{10}\left(\frac{255^2}{\text{MSE}}\right) = 10\log_{10}\left(\frac{255^2}{\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - I'(i,j)]^2}\right)$.

For the experiment we use 1000 randomly chosen faces from FRGC and compute the PSNR of the reconstruction error for each using DW-KSVD. Figure 3 shows the mean PSNR varying with sparsity. We find that the best full face reconstruction occurs at using $K = 40$ which is what we use for all further experiments. It is also worth noted that the more training samples presented to the dictionary learning algorithm, the higher PSNR it can achieve in hallucinating

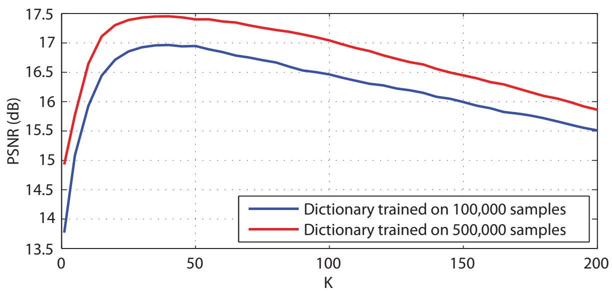

Figure 3. Mean PSNR versus reconstruction sparsity K using DW-KSVD trained dictionaries.

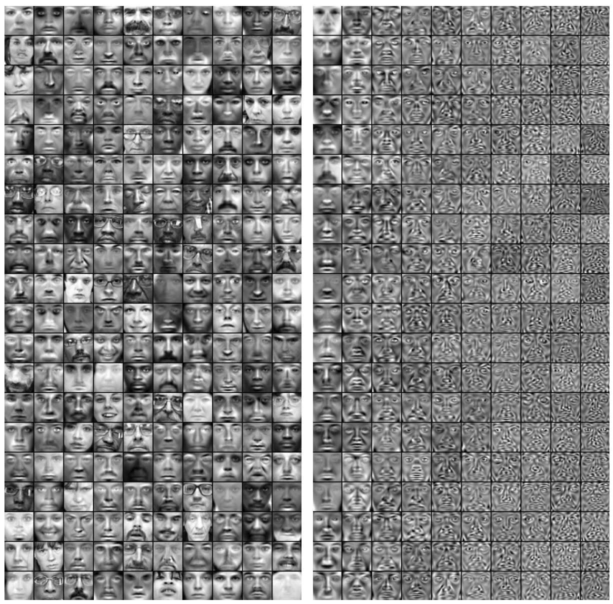

Figure 4. Left: K-SVD dictionary. Right: PCA dictionary. Only the first 200 dictionary elements are shown.

the full face from the periocular region. Figure 4 showcases the first 200 dictionary elements of the K-SVD and PCA dictionaries. In the PCA dictionary, after the first 40 eigenfaces, the dictionary elements start to lose facial structures and shift to capture higher frequency components. While in the K-SVD dictionary, visual appearances of the elements are close to the top eigenfaces where facial structures are well-preserved. This might explain why K-SVD and DW-KSVD leads to hallucinations with higher fidelity as we would see in the next section. Figure 5 shows a part of the full face component as well as the periocular component of a dictionary trained using the proposed DW-KSVD algorithm. For display, the intensities for the two components are both normalized.

4.3. Reconstruction Fidelity

Our primary goal is to provide a practical method for hallucinating a full face from the periocular region to aid further processing such as commercial face matching. However, a natural metric to evaluate methods for reconstruction would be to compare the reconstructed images to the original images using the PSNR metric. In this experiment, we reconstruct the entire target set in the FRGC ver2.0 database ($16,028$ images from 466 subjects) using the three methods and compute the corresponding PSNR for each pair. For all experiments using PCA, we restrict ourselves to the first 40 eigenvectors (same number of dictionary elements that K-SVD and DW-KSVD would use) which can represent over 93.9% of the total energy.

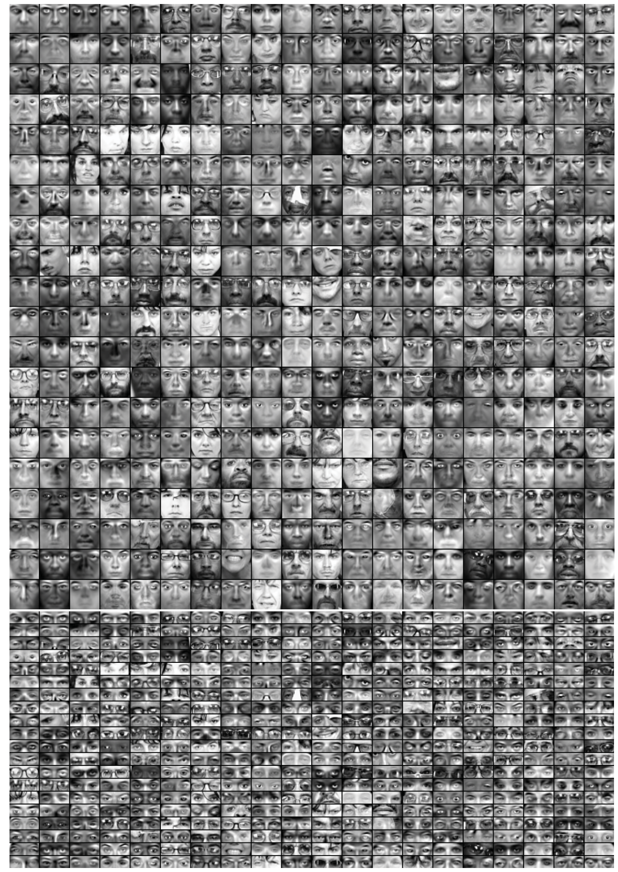

Figure 5. Top: Full face component of the DW-KSVD dictionary trained using 500,000 training samples. Bottom: Periocular component of the DW-KSVD dictionary. Only the first 400 dictionary elements are shown.

Figure 6 shows the overall mean PSNR computed for each subject (multiple images per subject) using DW-KSVD, K-SVD and PCA reconstruction (bold line) along with the mean PSNR for each individual subject for the three methods (markers). In FRGC ver2.0 target set, each individual has on an average 34 images. Figure 7 shows the corresponding histograms. We find that DW-KSVD on average clearly outperforms both K-SVD and PCA by a large margin in PSNR. Table 1 shows the mean and the standard deviation of the distribution of the PSNR values. A few randomly chosen samples and their reconstructions are shown

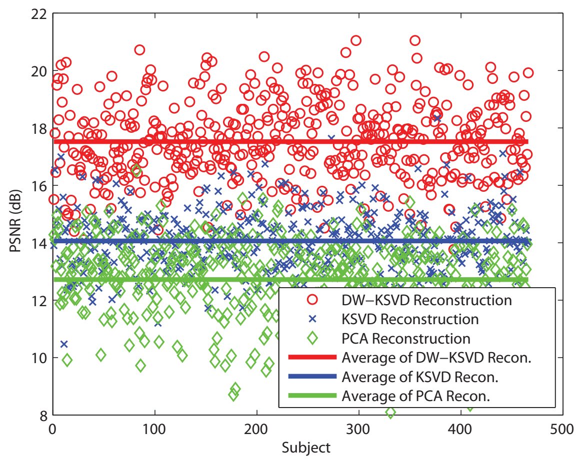

Figure 6. Mean PSNR values of reconstruction errors of individual subjects (multiple images per subject) along with the overall mean for the three methods.

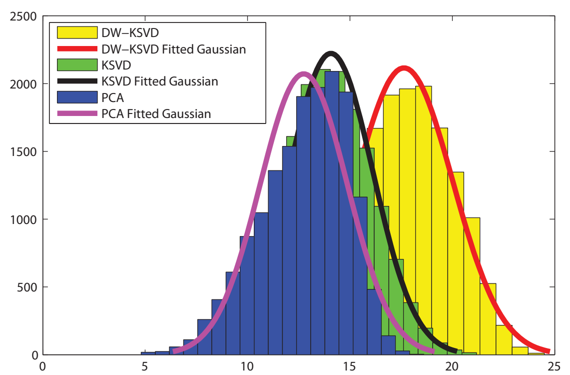

Figure 7. Overall distribution of the PSNR values for the three methods along with the corresponding fitted Gaussian curve.

in Figure 8. It is worth noting that most of the reconstructed faces are neutral in expression. This is because our dictionaries are trained on mugshot images, which typically have neutral expression. This, however, works in our favor because commercial matchers perform better under neutral expressions. Our proposed method actually eliminates expression variations and will be an asset for real-world matching.

We find that DW-KSVD not only provides reconstructions with higher PSNR values on average but the reconstructions are in fact much more visually appealing and similar to the original images than either PCA or K-SVD based reconstructions. This suggests that DW-KSVD and the combined formulation is able to extract the weak correlations and dependencies between the periocular and other facial features. Hence, explicitly penalising reconstruction error in the periocular region more seems to be favorable.

4.4. Face Verification

We now provide a few results which explore face matching using the reconstructed faces. We carry out a large-scale face verification experiment to evaluate whether the hallucinated faces can practically replace the ground-truth full face in face verification.

Table 1. Mean and standard deviations for the distributions of the PSNR values for reconstruction.

Methods	***Mean***	***Standard Deviation***
PCA Recon.	12.7439	2.1288
KSVD Recon.	14.0720	2.0532
DW-KSVD Recon.	17.6402	2.3757

4.4.1 Reconstructed Face vs. Reconstructed Face

In our first verification experiment, we strictly follow NIST's FRGC Experiment 1 protocol which involves 1-to-1 matching of the $16,028$ controlled target images to themselves ($\sim$ 256 million pair-wise face match comparisons). For this experiment, we adopt the normalized cosine distance (NCD) to compute the similarities between images: $d(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$.

The result of each algorithm is a similarity matrix with the size of $16,028 \times 16,028$ whose entry SimM_{ij} is the NCD between the feature vector of query image i and gallery image j. In the case of FRGC Experiment 1, the query set and gallery set are the same. The performance is analyzed using verification rate (VR) at 1% (0.01) false accept rate (FAR), equal error rate (EER) and the receiver operating characteristic (ROC) curves. Table 2 shows the VR at 1% FAR and EER for the FRGC Experiment 1 evaluation. Figure 9 shows the corresponding ROC curves. It can be noted from the table as well as the plot that DW-KSVD can achieve comparable results as the full face evaluation, which, from another angle, shows the fidelity of the hallucination. Further, it clearly outperforms both PCA and K-SVD based reconstructions. We also observe that the periocular region crop performs slightly better than the full face. This is because in the FRGC target set, two facial expressions are presented by each subject, neural and smiling. The periocular region, however, is less affected by such expression variations, thus gives rise to slightly better performance than the full face. This observation brings a valid point that one may focus on periocular region which has higher tolerance for expression variations when matching faces with unconstrained expressions.

Since this protocol matches the reconstructions to themselves, it only goes to show that the reconstructed images preserve biometric and identity information amongst themselves almost as well as the original images. The performance is not expected to drop too much if the entire system is trained on reconstructed images. This however is impractical. It is very hard to retrain commercial matchers, and our original problem was to evaluate the reconstructions using recognition systems trained on original images. To evaluate this, we run a second verification experiment in which we

Figure 8. Original full faces and periocular region crops along with the corresponding reconstructed or hallucinated images using exclusively the periocular crops for various samples from FRGC.

match the original images with the reconstructed images for all methods *i.e*. the targets are the original images and the probes are the reconstructed images.

4.4.2 Original Face vs. Reconstructed Face

For this experiment, we use a face verification algorithm that had good performance in the NIST's FRGC evaluation: the kernel class-dependence feature analysis (KCFA) [13]. In our experiment, KCFA was trained on the original images of the 222 subjects belonging to FRGC ver2.0 training set. We match the original face images of FRGC ver2.0 target set to the corresponding reconstructed images using the KCFA feature vectors extracted. Thus, we simulate a real-world situation, *i.e*. matching the reconstructed images to the original ones with a verification algorithm that has been trained on unseen original images. Our hope is to find that reconstructed images using DW-KSVD perform competitively as compared to matching the original images themselves. Indeed, this is what we observe. Figure 10 shows the ROC curves corresponding to this experiment. We find that among the three methods, DW-KSVD clearly outperforms both PCA and K-SVD reconstructions and in fact the ROC curve shows that the evaluation is indeed competitive to the one using the *original* full faces. Thus, we see that the periocular based full face reconstruction using DW-KSVD seems to be a practical solution in cases where the face verification system cannot adapt to partial faces. Moreover, it clearly outperforms both PCA and K-SVD based reconstructions in all evaluations. One reason that PCA's VR is lower than KSVD in this protocol but higher in the previous one might be that the truncated PCA reconstruction produces faces that lack details and tend towards the mean face. Thus, for matching reconstruction to reconstruction images, PCA actually gains by looking alike each other. However, such a lack of detail hurts PCA in the second protocol when matching reconstructed faces to the original ones. Thus for face hallucination, this provides another reason to favor the sparse representation in an overcomplete basis framework. Note that for this experiment, we essentially arrived at the same problem that motivated this study. We were unable to match only the periocular region to the original face given the trained KCFA based matcher, which is why no periocular ROC curve exists in Figure 10.

5. Conclusion

We present a practical and effective method to hallucinate a full face image using only the periocular region. Such a method would have applications in areas such as commercial face matching and law enforcement where currently algorithms are not adaptive to having only the periocular region. Our algorithm DW-KSVD is a modification of the K-SVD dictionary learning paradigm tailored so as to emphasize on more accurate reconstruction of a subset of dimensions, in this case the periocular region. Our experiments demonstrate that reconstruction using DW-KSVD can be practically used to hallucinate faces from the periocular re-

Table 2. VR at 1% FAR and EER for the FRGC Experiment 1 evaluation. The last three rows are matching reconstructed faces to the reconstructed faces.

Methods	*VR at 1% FAR*	*EER*
Original Full Face	0.524	0.170
Periocular Region	0.561	0.161
DW-KSVD Recon.	0.475	0.188
KSVD Recon.	0.285	0.248
PCA Recon.	0.329	0.236

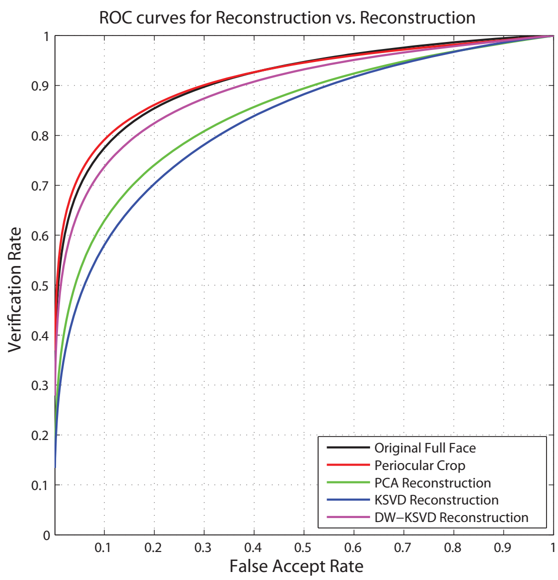

Figure 9. ROC curves obtained by following FRGC experiment 1 protocol (matching all sets to themselves) using raw pixels matched under the NCD metric.

Table 3. VR at 1% FAR and EER for the FRGC Experiment 1 evaluation using KCFA. The last three rows are matching reconstructed faces to the original full faces.

Methods	*VR at 1% FAR*	*EER*
Original Full Face	0.982	0.014
DW-KSVD Recon.	0.826	0.056
KSVD Recon.	0.438	0.165
PCA Recon.	0.046	0.452

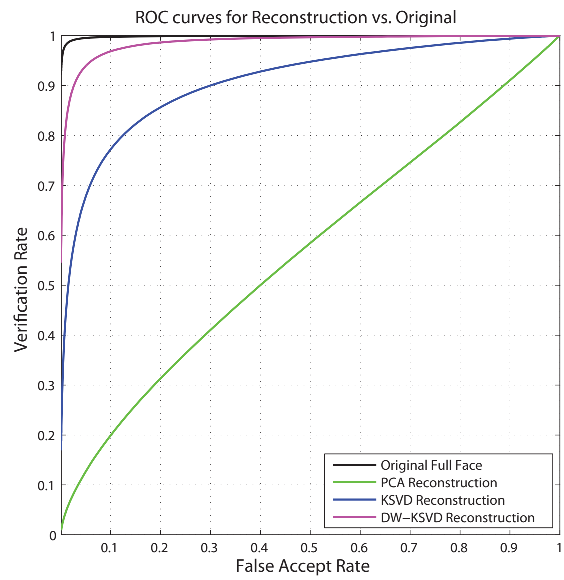

Figure 10. ROC curves obtained by matching all sets to the original full faces using KCFA features matched under the NCD metric.

gion without sacrificing face verification performance too much. Further, our method is open set and can hallucinate faces not present in training. It also outperforms standard K-SVD and PCA based reconstruction schemes in the same tasks. Our method is general in that one can try to reconstruct an entire signal given a part of it, given that weak correlations exist between that part and the rest of the signal. In future work, it would be interesting to explore the method based reconstructions being evaluated using other dictionary and subspace learning techniques and various commercial matchers.

6. Acknowledgement

This work was funded in part by the National Institute of Justice Grant 2013-IJ-CX-K005 and CMU CyLab.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11):4311–4322, Nov 2006. 2

[2] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on*, 15(12):3736–3745, Dec 2006. 3

[3] K. Hollingsworth, K. W. Bowyer, and P. J. Flynn. Identifying useful features for recognition in near-infrared periocular images. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–8, sep 2010. 2

[4] Z. Jiang, Z. Lin, and L. Davis. Label consistent K-SVD: Learning a discriminative dictionary for recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(11):2651–2664, Nov 2013. 4

[5] F. Juefei-Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides. Robust local binary pattern feature sets for periocular biometric identification. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–8, sep 2010. 2

[6] F. Juefei-Xu, M. Cha, M. Savvides, S. Bedros, and J. Trojanova. 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. 2

[7] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen. Investigating age invariant face recognition based on periocular biometrics. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–7, Oct 2011. 2

[8] F. Juefei-Xu and M. Savvides. Can your eyebrows tell me who you are? In *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*, pages 1–8, Dec 2011. 2

[9] F. Juefei-Xu and M. Savvides. Unconstrained periocular biometric acquisition and recognition using cots ptz camera for uncooperative and non-cooperative subjects. In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 201–208, Jan 2012. 2

[10] F. Juefei-Xu and M. Savvides. An augmented linear discriminant analysis approach for identifying identical twins with the aid of facial asymmetry features. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, pages 56–63, June 2013. 2

[11] F. Juefei-Xu and M. Savvides. Subspace based discrete transform encoded local binary patterns representations for robust periocular matching on nist's face recognition grand challenge. *IEEE Trans. on Image Processing*, 2014. 2, 4

[12] M.-S. Koh and E. Rodriguez-Marek. Turbo inpainting: Iterative K-SVD with a new dictionary. In *Multimedia Signal Processing, 2009. MMSP '09. IEEE International Workshop on*, pages 1–6, Oct 2009. 3

[13] B. Kumar, M. Savvides, and C. Xie. Correlation pattern recognition for face recognition. *Proceedings of the IEEE*, 94(11):1963–1976, Nov 2006. 7

[14] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard. Soft biometric classification using periocular region features. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–7, sep 2010. 2

[15] P. E. Miller, J. R. Lyle, S. J. Pundlik, and D. L. Woodard. Performance evaluation of local appearance based periocular recognition. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–6, sep 2010. 2

[16] U. Park, R. R. Jillela, A. Ross, and A. K. Jain. Periocular biometrics in the visible spectrum. *Information Forensics and Security, IEEE Trans. on*, 6(1):96–106, mar 2011. 2

[17] U. Park, A. Ross, and A. K. Jain. Periocular biometrics in the visible spectrum: A feasibility study. In *Biometrics: Theory, Applications, and Systems. BTAS. IEEE 3rd Int'l Conf. on*, pages 1–6, sep 2009. 2

[18] D. L. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross. On the fusion of periocular and iris biometrics in non-ideal imagery. In *Pattern Recognition (ICPR), 20th Int'l Conf. on*, pages 201–204, aug 2010. 2

[19] D. L. Woodard, S. J. Pundlik, J. R. Lyle, and P. E. Miller. Periocular region appearance cues for biometric identification. In *Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society Conf. on*, pages 162–169, jun 2010. 2