LAW: LOCALITY-AWARE WHITENING

Felix Juefei-Xu and Marios Savvides

Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

ABSTRACT

A recent work on 2D whitening reconstruction has been developed for enhancing the robustness of the PCA [1]. The gist is to perform whitening on a single image, where each row of that image is treated as the data point, as compared to the commonly adopted whitening method that treats each vectorized image as a data point, and performs whitening in a large data matrix. In this work, we have shown the connection between the two and proposed a more generalized whitening process called locality-aware whitening (LAW). The LAW partitions the image data into more meaningful sub-data by keeping the locality of the neighboring pixels, instead of just taking the rows of the image to be the sub-data which results in weak semantic links between them. In addition, the parameterization in LAW allows us to tune the granularity of the whitening algorithm, leading towards a special-purpose pre-processing step for each individual image. Experiments on illumination invariant face recognition on Extended YaleB, CMU PIE, and AR Face datasets have shown the effectiveness of the proposed method. In a larger sense, being able to parameterize the whitening reconstruction brings us one step closer towards an end-to-end trainable system with data pre-processing component included.

Index Terms— PCA, 2D Whitening Reconstruction

1. INTRODUCTION

Principal component analysis (PCA) is perhaps the most widely exercised machine learning algorithm for dealing with high dimensional data. However, PCA does not perform robustly when the matrix is grossly corrupted [2]. In face recognition applications, gross corruption could be due to illumination variations, facial expression variations, or occlusions, and PCA usually exhibits inferior performance when the faces are under these gross degradations. Two-dimensional whitening reconstruction (TWR) [1] was established as a pre-processing step to significantly reduce the influence of gross errors on PCA and its variants via reducing the pixel redundancy of the internal image and maintaining important intrinsic features, and meanwhile making each face image to be close to a Gaussian signal. The gist of TWR is to perform whitening on a single image, where each row of that image is treated as the data point, as compared to the commonly adopted whitening method that treats each vectorized image as a data point, and performs whitening in a large data matrix. In this work, we have shown the connection between the two, and proposed a more generalized whitening process called locality-aware whitening (LAW). The LAW reconstruction method partitions the image data into more meaningful sub-data by keeping the locality of the neighboring pixels, instead of just taking the rows of the image to be the sub-data which results in weak semantic links between them. Being able to look beyond just the horizontal dimension in the image allows the algorithm to capture more intrinsic structural details. In addition, the LAW allows us to tune the granularity of the whitening algorithm, leading towards a special-purpose pre-processing step for each individual image. Experiments on illumination invariant face recognition on both the Extended YaleB, CMU PIE, and AR Face datasets have shown the effectiveness of the proposed methods over its competitors.

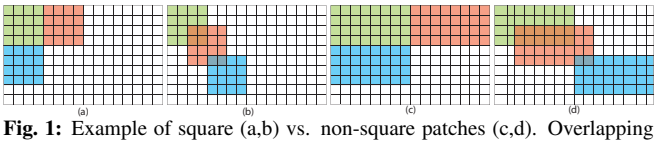

Fig. 1: Example of square (a,b) vs. non-square patches (c,d). Overlapping (b,d) vs. non-overlapping patches (a,c).

Related Work: The earliest application of PCA on human face images can be traced back to [3, 4]. The authors have argued that any face image could be approximately reconstructed by a weighted sum of a small collection of images. Turk and Pentland [5] have presented the well-known Eigenfaces method for face recognition, which marked the beginning of extensive exploration of PCA on all kinds of problems in the next couple decades. Vertically centered PCA process (PCA I) and horizontally centered PCA process (PCA II) have been discussed in [6, 7], as a major performance booster for the superiority of both Architecture I and II in ICA approach [8, 9] over PCA. Both PCA I and PCA II still transform the 2D face matrix into a concatenated 1D vector before feature extraction, which leads to a high-dimensional image vector and a high-dimensional covariance matrix as well. In order to overcome this problem, two-dimensional PCA (2DPCA) [10] has been proposed, allowing the evaluation of the covariance matrix to be both accurate and efficient. PCA and its variants are sensitive to the aforementioned gross degradations. Thus, improving the robustness of PCA and its variants are important tasks.

2. ALGORITHMIC APPROACH

Overview of TWR: Let us begin by briefly reviewing the two-dimensional whitening reconstruction (TWR) algorithm proposed in [1]. The gist of the algorithm is to perform whitening on a single image, where each row of that image is treated as the data point, as compared to the commonly adopted whitening method that treats each vectorized image as a data point, and performs whitening in a large data matrix.

More specifically, given an image $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$, the row-mean vector $\mathbf{u} \in \mathbb{R}^n$ is first calculated, and then the row-mean vector is subtracted from each row in the original image, leading towards X_v. Next, SVD is performed on this image matrix: $X_v = UDV^T$, and finally, the two-dimensional whitening reconstruction is obtain by $\widetilde{X}_v = \sqrt{n}U(:, 1:m)V(:, 1:m)^T$, where m is usually the rank of the row-mean-removed image matrix, or the number of significant eigenvectors kept, as specified by the user.

LAW Reconstruction: The gist of TWR is to whiten each single image by treating rows of the image as the data points, instead of treating an entire vectorized image as a data point and performing whitening transform on the whole data matrix. Although taking rows of the image and treat them as data points for whitening purposes has been inspired by [6], [7], there is no reason not to look beyond just rows of the image. In fact, local facial structure are usually contiguous. For example, eyes, nose, mouth, eyebrows, etc. are all shown as block regions on human faces [11, 12, 13, 14]. In addition, local corruptions are also usually contiguous. This intuition allows us to think beyond just collecting rows of the face images to be the data points. Perhaps a better way of partitioning the face images can lead to better whitening reconstruction than TWR. Another way to think of this is as follows. In traditional image whitening procedure, as is usually performed before PCA is carried out, each training image is vectorized first, and the whitening is performed on the entire data matrix, where the columns consist of the vectorized images. When vectorizing the image, we are actually losing the spatial locality information of the image, because after the vectorization, the algorithm is agnostic about where each pixel comes from. Permuting the pixels in the vectorized image for the entire data matrix will lead to the same whitening results, namely it is a permutation-invariant algorithm. Similarly, in TWR, permuting rows within the image would lead to the same whitening results. This is due to the fact that when collecting local data points within each image, the algorithm does not look beyond horizontal dimension.

In this regard, we propose to generalize the TWR algorithm by taking into account local image information in both horizontal and vertical dimensions, making it locality-aware such that aforementioned permutation will not lead to the same whitening results. This is advantageous over TWR because it allows us to partition the image data into more meaningful sub-data by keeping the locality of the neighboring pixels, instead of just taking the rows of the image to be the sub-data which results in weak semantic links between them. In addition, it allows us to tune the granularity of the whitening algorithm, leading towards a special-purpose preprocessing step for each individual image. For simplicity, we call the proposed method a locality-aware whitening (LAW) reconstruction. Instead of just taking the rows, we now partition the image into various patches of the same size. As shown in Figure 1, we can partition the image using overlapping patches, or non-overlapping patches. We can also choose to select the aspect ratio of the patch. For example, the patches can be square-shaped with various sizes, or it can be rectangular-shaped with specified width and height.

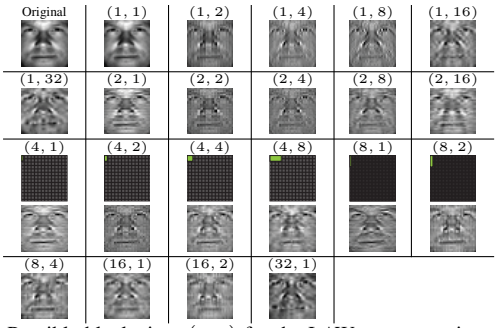

Table 1: Possible block sizes (p, q) for the LAW reconstruction with non-overlapping local patches. Patches are only shown for the 3rd row.

Overlapping vs. Non-Overlapping Patch: Taking overlapping patches obvious leads to a more redundant representation, and sometimes it can lead to a more stable performance, especially when the face image is corrupted. If there is no other constraint, one should always favor overlapping patches over non-overlapping patches. However, there is one slight caveat that some sort of pooling or averaging procedure is required after the whitening and during the reconstruction of the image, because each pixel has been used by multiple patches and should account for contributions from multiple patches when being reconstructed.

The other good thing about taking overlapping patches is that, one can have an extra free parameter, step size, to allow for more flexibility in the algorithm, especially when dealing with face images of different sizes, conditions, etc.

On the other hand, taking non-overlapping patches is quite straight-forward, and the reconstruction will be exact because each pixel is only used by one patch. However, it also comes with one caveat. One needs to carefully select the size of the patch, with respect to the size of the entire image. A less thoughtful selection of the dimensions will lead to the following two cases: either the image is not fully covered by the patches, or some zero-padding is need.

Square vs. Non-Square Patch: The intuition for using square patches is again straight-forward because the face image crops are usually in square shape, and putting up a grid to equally partition the image seems very natural and reasonable. However, it is also interesting to see how different patch shapes could affect the whitening performance. In fact, the

previously discussed TWR method is a special case of LAW, where the patch is taken to be of $1 \times q$, where q coincides with the image width. Perhaps, one should favor one dimension to the other, leading to a rectangular patch.

Patch Size Determination: The selection of patch size is greatly affected by the image size. In this work, we will follow the protocols set forth by the work of TWR [1], and their experiments are heavily based on images of size 32×32. Hence, this is the size of the whole image we start with.

To facilitate the following experiments, we add two constraints: (1) use non-overlapping patches for simplicity, and they have to cover the image entirely without the further need of zero-padding. (2) the obtained per-image data matrix after partitioning the image needs to be a square matrix or a fat matrix, thus avoiding the small-sample-size problem.

To satisfy the first constraint, a patch of size $p \times q$ will have to be selected from the possible factors of 32. In other words, $p, q \in \{1, 2, 4, 8, 16, 32\}$. For the second constraint, the per-image data matrix whose size is $r \times s$ will need to satisfy the following relationship: $r \leq s \text{ and } r \times s = 1024 \Rightarrow r \leq \frac{1024}{r} \Rightarrow r \leq 32$ where $r = p \times q$ is the dimension of each patch, and s is the number of non-overlapping patches. This established relationship $p \times q \leq 32$ has eliminated many patch size possibilities. The remaining valid patch sizes are listed in Table 1 and Table 2. Among these 21 (p, q) choices, we choose the best patch size according to Fisher ratio criterion, as practiced in linear discriminant analysis (LDA) [15] which is a good indicator of class separation. The validation dataset we use is the "target" portion of the FRGC v2 dataset [16] which contains 16,028 frontal images from 466 subjects [17, 18, 19]. Images are aligned by the two eye locations [20, 21, 22] and cropped to be 32×32, in order to be consistent with the following main experiments, to be discussed in the next section.

We will whiten all the images in the validation set using the aforementioned 21 patch sizes and then check the multi-class separability using Fisher ratio. The best performing patch size will be used in the subsequent main experiments. It is worth noted that, our major experiments will be on face classification. Evaluating class separability using Fisher ratio at this point serves a similar purpose while still maintains some differences between the evaluation criterion at hand and the ones to be used in the following experiments, with the hope that algorithmic biases can be mitigated.

The Fisher ratio is defined as follows. We know that LDA aims to find the projection such that the ratio of the between-class scatter $\mathbf{S}_B = \sum_{i=1}^{C} N_i(\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^\top$ and the within-class scatter $\mathbf{S}_W = \sum_{i=1}^{C} \sum_{\mathbf{x}_k \in C_i} (\mathbf{x}_k - \boldsymbol{\mu}_i)(\mathbf{x}_k - \boldsymbol{\mu}_i)^\top$ is maximized, where $\boldsymbol{\mu}_i$ is the mean image of class C_i, $\boldsymbol{\mu}$ is the mean image of all the images, C is the total class number, and N_i is the number of images in class C_i. In light of this, the multi-class Fisher ratio can be calculated through $\text{Fisher Ratio} \triangleq \frac{\text{trace}(\mathbf{S}_B)}{\text{trace}(\mathbf{S}_W)}$. The Fisher ratios are consolidated in Table 2. It can be seen that $(4, 8)$ achieves the highest Fisher ratio and will be used for the following experiments. Table 1 illustrates some visual results for carrying out whitening reconstruction using various patch sizes.

Discussion on the Patch Size Parameters: We think that perhaps the most important contribution of this work is the parameterization of the whitening process. The proposed method looks beyond just the horizontal dimension (as done in TWR), and allows for more meaningful partition of the image, for better whitening performance. We have seen that choosing an optimal patch size can lead to superior performance using LAW approach. In a broader sense, such parameters in the whitening stage themselves can be even updated or determined automatically through data and the given tasks. In a recent work called the spatial transformer networks [23], the affine transformation parameters are updated along with the update of the convolutional filters, making such a CNN capable of dealing with images under such transformations much more robustly because the model has learned invariance to translation, scale, rotation and more generic warping, through such parameterization. Similarly, for our case, being able to parameterize the whitening reconstruction brings us one step closer towards an end-to-end trainable system with data pre-processing component included.

3. EXPERIMENTS

We will carry out two major experiments: face classification and face clustering. We follow the protocols set forth in [1]. The purpose of these experiments is to study whether the proposed locality-aware whitening reconstruction, which is a generalization of TWR [1] can further improve the robustness of PCA and its variants. Three databases are used:

(1) Extended YaleB Face Database [24]: We choose the frontal pose with all 64 illumination conditions for the 38 subjects from the YaleB database, yielding a total of $64 \times 38 = 2432$ face images. All faces are aligned and cropped to be 32×32 in size, with the exception that for Experiment I (2), faces are cropped to be 96×96 in size. *(2) CMU PIE database* [25]: We also choose the frontal pose with all 43 illumination conditions and 4 different expressions, for the 68 subjects from the CMU PIE database, yielding a total of $43 \times 4 \times 68 = 11696$ face images. All faces are aligned and cropped to be 32×32 in size. *(3) AR database* [26]: We choose 100 subjects (50 male and 50 female) out of 126 subjects from the AR database randomly, each under 26 various facial expressions, illumination conditions, and occlusions, yielding a total of $100 \times 26 = 2600$ face images. All faces are aligned and cropped to be 32×32 in size.

Face Classification Experiments: We carry out illumination invariant face classification tasks, and apply the proposed LAW reconstruction approach to PCA, PCA I, PCA II, and 2DPCA. Performance will be compared against TWR applied to PCA and its variants [1].

(p, q)	$(1, 1)$	$(1, 2)$	$(1, 4)$	$(1, 8)$	$(1, 16)$	$(1, 32)$	$(2, 1)$	$(2, 2)$	$(2, 4)$	$(2, 8)$	$(2, 16)$
Fisher Ratio	4.1114	6.5770	7.8623	8.4254	11.4345	10.6504	7.5850	7.7059	9.5171	11.7768	10.9778
(p, q)	$(4, 1)$	$(4, 2)$	$(4, 4)$	$(4, 8)$	$(8, 1)$	$(8, 2)$	$(8, 4)$	$(16, 1)$	$(16, 2)$	$(32, 1)$	
Fisher Ratio	8.3810	9.3931	12.7406	**13.2605**	8.4901	12.0694	12.8768	11.4538	9.7955	10.5877	

Table 2: For 32×32 face images, the following are the possible block sizes (p, q) we choose to carry out the locality-aware whitening reconstruction. Local patches are obtained in a non-overlapping fashion in this work. Corresponding Fisher ratios are also shown.

Method	$q = 2$	$q = 4$	$q = 6$	$q = 8$
Fisherface	38.5 ± 1.3	63.8 ± 2.2	74.7 ± 1.4	80.0 ± 1.5
CRC	52.7 ± 2.6	76.4 ± 2.2	86.1 ± 1.0	90.1 ± 0.8
PCA	11.2 ± 1.0	11.9 ± 1.1	11.6 ± 1.0	11.7 ± 1.1
PCA I	14.9 ± 1.2	17.0 ± 1.3	17.9 ± 1.1	19.2 ± 1.2
PCA II	48.1 ± 2.0	67.3 ± 1.2	75.4 ± 1.6	80.6 ± 1.7
2DPCA	19.4 ± 1.5	24.3 ± 1.9	26.8 ± 1.8	29.3 ± 2.0
ULR+PCA	48.0 ± 2.3	62.9 ± 1.1	70.0 ± 1.7	74.0 ± 1.9
ULR+PCA I	48.0 ± 2.3	68.8 ± 1.1	70.0 ± 1.7	74.0 ± 1.9
ULR+PCA II	52.2 ± 2.1	70.0 ± 1.0	78.3 ± 1.7	84.0 ± 1.4
ULR+2DPCA	49.4 ± 2.2	64.7 ± 1.3	71.8 ± 2.0	75.9 ± 2.2
IN+PCA	46.1 ± 2.5	59.9 ± 2.9	66.2 ± 3.0	71.7 ± 2.7
IN+PCA I	46.1 ± 2.5	60.0 ± 2.9	66.3 ± 3.0	71.8 ± 2.8
IN+PCA II	68.4 ± 2.0	76.3 ± 2.4	90.1 ± 1.1	94.2 ± 1.4
IN+2DPCA	56.6 ± 2.3	71.8 ± 2.7	78.1 ± 2.8	83.3 ± 2.1
DFD+PCA	48.7 ± 2.9	80.4 ± 2.7	85.5 ± 1.7	88.6 ± 1.5
DFD+PCA I	48.6 ± 2.9	80.3 ± 2.7	85.5 ± 1.7	88.6 ± 1.5
DFD+PCA II	46.7 ± 2.2	80.3 ± 2.2	83.5 ± 1.5	85.6 ± 1.8
DFD+2DPCA	46.0 ± 2.8	80.1 ± 2.8	85.5 ± 1.7	88.8 ± 1.4
TWR+PCA	62.6 ± 2.5	77.9 ± 2.0	84.1 ± 1.8	88.6 ± 2.1
TWR+PCA I	62.4 ± 2.5	77.9 ± 2.1	84.0 ± 1.8	88.6 ± 2.1
TWR+PCA II	72.6 ± 2.2	88.8 ± 1.3	93.3 ± 0.7	96.0 ± 0.9
TWR+2DPCA	63.3 ± 2.7	78.5 ± 2.0	85.0 ± 1.8	89.6 ± 2.0
LAW+PCA	69.3 ± 2.1	83.9 ± 2.1	89.3 ± 2.0	92.3 ± 0.9
LAW+PCA I	68.9 ± 2.1	83.6 ± 1.7	89.6 ± 2.1	92.4 ± 1.8
LAW+PCA II	78.8 ± 2.2	94.1 ± 1.3	97.8 ± 1.5	97.9 ± 1.5
LAW+2DPCA	70.1 ± 2.2	86.0 ± 2.2	91.2 ± 1.6	93.8 ± 1.8

Method	$q = 1$	$q = 2$	$q = 3$	$q = 4$
Fisherface	28.4 ± 1.8	87.9 ± 1.4	95.2 ± 1.1	95.2 ± 1.1
CRC	88.1 ± 1.0	99.3 ± 0.3	99.9 ± 0.1	100
PCA	24.1 ± 1.2	28.8 ± 1.4	31.3 ± 1.3	32.7 ± 1.9
PCA I	37.0 ± 1.2	43.9 ± 1.1	46.7 ± 1.5	48.6 ± 1.5
PCA II	78.8 ± 1.3	93.9 ± 1.8	96.9 ± 0.6	97.7 ± 0.6
2DPCA	45.1 ± 1.3	58.3 ± 1.6	63.7 ± 2.0	67.6 ± 2.1
ULR+PCA	88.2 ± 1.8	95.3 ± 0.7	97.3 ± 0.5	98.4 ± 0.4
ULR+PCA I	88.3 ± 1.8	95.3 ± 0.6	97.3 ± 0.5	98.4 ± 0.4
ULR+PCA II	90.6 ± 1.3	99.0 ± 0.3	99.6 ± 0.2	99.7 ± 0.2
ULR+2DPCA	88.7 ± 1.7	95.5 ± 0.7	97.6 ± 0.5	98.5 ± 0.4
IN+PCA	63.6 ± 3.0	72.1 ± 2.2	78.0 ± 2.2	80.4 ± 1.9
IN+PCA I	63.6 ± 3.0	72.1 ± 2.2	78.0 ± 2.1	80.5 ± 1.9
IN+PCA II	76.4 ± 2.4	95.3 ± 0.6	97.0 ± 0.4	97.6 ± 0.5
IN+2DPCA	68.5 ± 2.8	79.2 ± 1.6	84.3 ± 1.7	86.5 ± 1.3
DFD+PCA	45.1 ± 4.0	95.6 ± 0.7	97.0 ± 0.7	97.8 ± 0.6
DFD+PCA I	44.9 ± 4.0	95.6 ± 0.7	97.0 ± 0.7	97.8 ± 0.6
DFD+PCA II	54.7 ± 3.2	97.9 ± 0.4	98.8 ± 0.3	99.2 ± 0.3
DFD+2DPCA	44.5 ± 3.7	96.0 ± 0.6	97.2 ± 0.7	97.8 ± 0.5
TWR+PCA	93.3 ± 1.1	97.0 ± 0.7	98.4 ± 0.4	99.2 ± 0.3
TWR+PCA I	93.1 ± 1.0	97.0 ± 0.7	98.4 ± 0.4	99.2 ± 0.3
TWR+PCA II	96.4 ± 0.5	99.7 ± 0.2	99.9 ± 0.1	100
TWR+2DPCA	93.4 ± 1.0	96.9 ± 0.7	98.4 ± 0.4	99.2 ± 0.4
LAW+PCA	97.5 ± 0.9	98.8 ± 0.6	99.9 ± 0.3	100
LAW+PCA I	98.0 ± 1.1	98.5 ± 0.3	99.7 ± 0.3	100
LAW+PCA II	99.6 ± 0.6	100	100	100
LAW+2DPCA	98.8 ± 0.8	100	100	100

Table 3: Classification results: mean accuracy $\pm$ std. $(\%)$ on Extended YaleB (top) and CMU PIE Database (bottom) under lighting variations.

Training Faces under Illumination Changes: For Extended YaleB database, we randomly select $q = [2, 4, 6, 8]$ images as the training set, following the same protocol as in [1], and the remaining images are used for testing, where q is the number of training images of each subject.

Similarly, for CMU PIE database, we are only using face images with frontal pose. For each subject, there are 43 illumination variations, including one with neutral lighting. In this experiment, we use the 42 images with illumination changes by randomly choosing $q = [1, 2, 3, 4]$ images as the training set and the remaining images for testing. We repeat this process for 20 times and report the mean recognition accuracy as well as its standard deviation.

The baseline algorithms we adopt are the Fisherface [15] and collaborative representation based classification (CRC) [27] approach, which will give us an idea on how well PCA and its variants (PCA I, PCA II, and 2DPCA) perform. Since we are dealing with illumination-invariant face recognition tasks, we also study 3 illumination pre-processing techniques. They are unsupervised low-rank representation (ULR) [28], illumination normalization (IN) [29] , and discriminant face descriptor (DFD) [30], all of which can improve face recognition performance under illumination variations.

Method	S2	S3	S4	S5	Method	S2	S3	S4	S5
LRC	100	100	87.6	42.2					
CRC	100	100	88.0	35.7					
SRC	98.5	93.4	78.4	28.8					
RSRC	100	100	80.3	36.7					
HQ-A	99.8	96.3	67.9	31.3					
HQ-M	99.8	96.3	75.8	36.8					
NMR	100	100	90.2	47.9					
TWR+PCA II	100	100	92.2	76.4	LAW+PCA II	100	100	94.2	81.3
TWR+LRC	100	100	92.2	76.4	LAW+LRC	100	100	94.2	80.8
TWR+CRC	100	100	95.1	76.8	LAW+CRC	100	100	95.0	76.8
TWR+SRC	100	100	97.9	86.8	LAW+SRC	100	100	96.4	88.9
TWR+RSRC	100	100	97.2	85.2	LAW+RSRC	100	100	98.1	87.6
TWR+HQ-A	100	99.6	90.7	60.2	LAW+HQ-A	100	100	93.7	64.4
TWR+HQ-M	100	99.6	90.5	58.0	LAW+HQ-M	100	100	93.4	63.5
TWR+NMR	100	100	98.5	88.2	LAW+NMR	100	100	100	92.5

Table 4: Classification accuracy $(\%)$ on faces under illumination variation changes from slight (Subset 2) to severe (Subset 5) on Extended YaleB.

Classification will be carried out using the nearest centroid classifier [31], except for CRC. Table 3 shows the mean classification accuracies and standard deviation for all the preprocessing methods applied to PCA and its variants on Extended YaleB and CMU PIE databases. We can observe that, all of ULR, IN, DFD and TWR can boost the classification results. In addition, the proposed LAW can further improve upon the TWR results, when applied to PCA and its variants.

Training Faces under Neutral Lighting Condition: In this experiment, we partition the Extended YaleB database into five subsets with increasing level of harshness in illumination changes. We select Subset 1 as the training set (least harsh set), and used Subsets 2-5 for testing. The face image was cropped and resized to 96×96 in size. We pick the following baseline algorithms: LRC [32], SRC [33], CRC [27], RSC [34], NMR [35], HQ A and HQ M [36]. Together with the best performing PCA II (see Table 3), we study how TWR and LAW can further improve upon these aforementioned algorithms. The classification results are shown in Table 4. As can be seen, TWR does largely improve the classification performance on these methods, while our proposed LAW can take it one step further, by showing even higher performance.

Conclusion: In this work, we have proposed a locality-aware whitening (LAW) reconstruction as a generalization of the TWR [1] approach, for improving the robustness of PCA. The LAW reconstruction method partitions the image data into more meaningful sub-data by keeping the locality of the neighboring pixels, instead of just taking the rows of the image to be the sub-data which results in weak semantic links between them [1]. In addition, the parameterization of LAW allows us to tune the granularity of the whitening algorithm, leading towards a special-purpose pre-processing step for each individual image. Experiments on illumination invariant face recognition on both the Extended YaleB, CMU PIE, and AR Face datasets have shown the effectiveness of the proposed methods over its competitors. In a larger sense, being able to parameterize the whitening reconstruction brings us one step closer towards an end-to-end trainable system with data pre-processing component included.

4. REFERENCES

[1] X. Shi, Z. Guo, F. Nie, L. Yang, J. You, and D. Tao, "Two-dimensional whitening reconstruction for enhancing robustness of principal component analysis," *TPAMI*, vol. 38, no. 10, pp. 2130–2136, 2016.

[2] E.J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 11, 2011.

[3] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. of Opti. Soc. of America*, vol. 4, no. 3, pp. 519–524, 1987.

[4] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *TPAMI*, vol. 12, no. 1, pp. 103–108, Jan 1990.

[5] M. Turk and A. Pentland, "Eigenfaces for recognition," *Cognitive Neuroscience, Journal of*, vol. 3, no. 1, pp. 71–86, Jan 1991.

[6] J. Yang, D. Zhang, and J. Yang, "Is ICA significantly better than PCA for face recognition?," in *ICCV*. IEEE, 2005, vol. 1, pp. 198–203.

[7] J. Yang, D. Zhang, and J. Yang, "Constructing PCA baseline algorithms to reevaluate ICA-based face-recognition performance," *TSMC*, vol. 37, no. 4, pp. 1015–1021, 2007.

[8] A.J. Bell and T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[9] A.J. Bell and T.J. Sejnowski, "The "independent components" of natural scenes are edge filters," *Vision research*, vol. 37, no. 23, pp. 3327–3338, 1997.

[10] J. Yang, D. Zhang, A.F. Frangi, and J. Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *TPAMI*, vol. 26, no. 1, pp. 131–137, 2004.

[11] F. Juefei-Xu and M. Savvides, "Fastfood Dictionary Learning for Periocular-Based Full Face Hallucination," in *Biometrics: Theory, Applications and Systems (BTAS), 2016 IEEE Seventh International Conference on*, Sept 2016, pp. 1–8.

[12] F. Juefei-Xu, Dipan K. Pal, and M. Savvides, "Hallucinating the Full Face from the Periocular Region via Dimensionally Weighted K-SVD," in *Computer Vision and Pattern Recognition (CVPR) Workshops, 2014 IEEE Conference on*, June 2014, pp. 1–8.

[13] 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*, vol. 23, no. 8, pp. 3490–3505, aug 2014.

[14] F. Juefei-Xu, K. Luu, and M. Savvides, "Spartans: Single-sample Periocular-based Alignment-robust Recognition Technique Applied to Non-frontal Scenarios," *IEEE Trans. on Image Processing*, vol. 24, no. 12, pp. 4780–4795, Dec 2015.

[15] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *TPAMI*, vol. 19, no. 7, pp. 711–720, 1997.

[16] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *CVPR*. IEEE, 2005, vol. 1, pp. 947–954.

[17] F. Juefei-Xu, V. N. Boddeti, and M. Savvides, "Local Binary Convolutional Neural Networks," in *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, July 2017.

[18] D. K. Pal, F. Juefei-Xu, and M. Savvides, "Discriminative Invariant Kernel Features: A Bells-and-Whistles-Free Approach to Unsupervised Face Recognition and Pose Estimation," in *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, June 2016.

[19] F. Juefei-Xu and M. Savvides, "Multi-class Fukunaga Koontz Discriminant Analysis for Enhanced Face Recognition," *Pattern Recognition*, vol. 52, pp. 186–205, apr 2016.

[20] Felix 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*, sep 2010, pp. 1–8.

[21] F. Juefei-Xu and M. Savvides, "Single Face Image Super-Resolution via Solo Dictionary Learning," in *IEEE International Conference on Image Processing (ICIP)*, Sept 2015, pp. 2239–2243.

[22] F. Juefei-Xu and M. Savvides, "Encoding and Decoding Local Binary Patterns for Harsh Face Illumination Normalization," in *IEEE International Conference on Image Processing (ICIP)*, Sept 2015, pp. 3220–3224.

[23] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," in *NIPS*, 2015, pp. 2017–2025.

[24] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *TPAMI*, vol. 23, no. 6, pp. 643–660, 2001.

[25] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," in *FG*. IEEE, 2002, pp. 46–51.

[26] A.M. Martinez, "The AR face database," *CVC Technical Report*, vol. 24, 1998.

[27] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?," in *ICCV*. IEEE, 2011, pp. 471–478.

[28] Y. Wang, V.I. Morariu, and L.S. Davis, "Unsupervised feature extraction inspired by latent low-rank representation," in *WACV*. IEEE, 2015, pp. 542–549.

[29] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *TIP*, vol. 19, no. 6, pp. 1635–1650, 2010.

[30] Z. Lei, M. Pietikainen, and S.Z. Li, "Learning discriminant face descriptor," *TPAMI*, vol. 36, no. 2, pp. 289–302, 2014.

[31] R. Tibshirani, T. Hastie, B. Narasimhan, and G. Chu, "Diagnosis of multiple cancer types by shrunken centroids of gene expression," *PNAS*, vol. 99, no. 10, pp. 6567–6572, 2002.

[32] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *TPAMI*, vol. 32, no. 11, pp. 2106–2112, 2010.

[33] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *TPAMI*, vol. 31, no. 2, pp. 210–227, 2009.

[34] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in *CVPR*. IEEE, 2011, pp. 625–632.

[35] J. Yang, J. Qian, L. Luo, F. Zhang, and Y. Gao, "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes," *arXiv preprint arXiv:1405.1207*, 2014.

[36] R. He, W.-S. Zheng, T. Tan, and Z. Sun, "Half-quadratic-based iterative minimization for robust sparse representation," *TPAMI*, vol. 36, no. 2, pp. 261–275, 2014.