Hide and Seek: Uncovering Facial Occlusion with Variable-Threshold Robust PCA

Wee Kheng Leow, Guodong Li, Jian Lai, Terence Sim, Vaishali Sharma
Dept. of Computer Science, National University of Singapore
{leowwk, laij, tsim, vaishali}@comp.nus.edu.sg, lguodong@u.nus.edu

Abstract

Face images are very important in human social activities, which can be severely hampered when they are corrupted by occluders such as eyeglasses, face marks, and scarfs. Existing methods for removing occlusions in face images can be grouped into three broad categories, namely PCA, robust PCA (RPCA), and sparse coding. The major weaknesses of these methods are inconsistent performance across test conditions and possible corruption of unoccluded part of the recovered target image. This paper presents variable-threshold RPCA (VRPCA) method based on RPCA with variable thresholding. Comprehensive tests show that VRPCA is able to preserve the unoccluded parts of the target image with practically zero error. Compared to existing methods, it is more accurate, reliable, and consistent across various test conditions.

1. Introduction

Face images are very important in human social activities including security, surveillance, biometric identification, criminal and forensic investigation, etc. These activities are severely hampered when face images are corrupted by occluders such as eyeglasses, face masks, and scarfs. Although there are algorithms for recognizing occluded faces, unoccluded face images of victims and suspects are still needed for investigation, monitoring, and communication to the public. In fact, many occluded face recognition methods directly or indirectly recover unoccluded faces during the recognition process.

Existing methods capable of removing occlusions in face images can be grouped into three broad categories: principal component analysis (PCA), robust PCA (RPCA), and sparse coding. PCA methods map the unoccluded parts of an occluded target image into an eigenspace constructed from unoccluded training images, and use the PCA coefficients to generate the unoccluded target image. RPCA methods decompose a data matrix containing unoccluded training images and an occluded target image into a low-rank matrix containing unoccluded training and target images and a sparse error matrix containing noise and occluders. Sparse coding methods model a face image as a weighted sum of unoccluded training images with sparse weights. Given an occluded target image, they compute the sparse weights that classify the target image. The weights can be used to generate an unoccluded image of the target.

There are two major shortcomings in these existing methods: (1) Their performance are not consistent across test conditions, and (2) the unoccluded parts of the target images can be changed by the algorithms, resulting in unintended corruption to the unoccluded parts.

This paper presents variable-threshold RPCA (VRPCA) method for removing occlusions in a target face image. VRPCA differs from RPCA in the use of soft thresholding or shrinkage operator. Whereas RPCA applies the same soft threshold to the entire image, VRPCA applies different soft thresholds to the occluded and unoccluded part. Comprehensive tests show that VRPCA can preserve the unoccluded parts of the target image with practically zero error. Compared to existing methods, it is more accurate, reliable, and consistent across various test conditions.

This paper focuses on the removal of occlusions in face images, and it studies the various occluder characteristics such as intensity, size, shape, and location that can affect the performance of occlusion removal algorithms. The issue of automatic occlusion detection will be addressed in a separate paper.

2. Related Work

Existing methods capable of removing occlusions in face images can be grouped into three broad categories: PCA, robust PCA, and sparse coding. PCA methods first construct a face eigenspace using unoccluded training images. Given an occluded image, they compute the unoccluded image as the sum of the eigenvectors weighted by linear coefficients, which can be computed in a variety of ways. For example, [28] uses the coefficients computed from projecting

the unoccluded part of the target image into the eigenspace. In contrast, [10, 11, 14, 24, 29] apply linear least-squares to compute the coefficients, whereas [25] applies iteratively reweighted least-squares, and [26] iteratively updates the weights that blend a mean image and the generated image. Instead of applying standard PCA, [30] applies incremental PCA whereas [19, 33] apply probabilistic PCA. In addition, [10, 14, 29] apply random sub-sampling of images to improve the robustness of their methods.

PCA methods generate an entire unoccluded image based on computed coefficients. So, they can corrupt the unoccluded part of the target image. To mitigate this problem, [13] and [35] replace only the occluded pixels and image blocks, respectively, of the target image by the most correlated pixels or blocks in the PCA-generated unoccluded image. Correlation between pixels and blocks are computed based on training images. In particular, [35] constructs eigenspaces of image blocks instead of entire face images, which are used to generate unoccluded image blocks. Such local replacement, however, creates apparent discontinuities or seams that need to be removed by blending the replaced pixels and blocks with their surrounding pixels, which leads to corruption of the surrounding pixels. Most of these PCA methods require a single training set of unoccluded face images. In contrast, [25, 26, 33] require training pairs of occluded and unoccluded face images, making these methods restrictive in real applications.

It is well known that PCA methods can be severely affected by large noise amplitude. Robust PCA (RPCA) and sparse coding methods, on the other hand, are more robust towards large noise amplitude. They are developed for machine recognition of occluded faces rather than face deocclusion per se. Nevertheless, many of these methods can recover unoccluded faces from occluded images. [23] applies RPCA method to decompose a data matrix containing unoccluded training images and an occluded target image into a low-rank matrix $\mathbf{A}$ that contains unoccluded faces and a sparse matrix $\mathbf{E}$ that contains noise and occluders. It solves the matrix decomposition problem by augmented Lagrange multiplier (ALM) method. [31] showed that exact solution of $\mathbf{A}$ can be obtained if $\mathbf{A}$ is low-rank and $\mathbf{E}$ is sparse. Empirical study shows that the rank of face images is around 50, which is not low given, say, 100 training images. Arbitrarily lowering the relative rank by increasing the number of training images may help but will significantly increase computation cost.

Sparse coding methods model a face image as a weighted sum of unoccluded training images with sparse weights. Given a target occluded face image, they compute the sparse weights to classify the subject in the image. The sparse weights can be used to generate an unoccluded image of the target through the weighted sum. Various methods have been proposed for computing the sparse weights. For example, [17, 34] use maximum likelihood estimation whereas [32] applies l_1-norm minimization. [6, 12, 16, 18, 22] use ALM to solve combined RPCA and sparse coding problem, whereas [27] uses alternating direction method of multipliers. As these methods are designed for face recognition, they require the presence of the subject in the training images. When the subject in the target image is not present in the training images, these methods produce a weighted sum of faces similar but not identical to the subject.

As for PCA, RPCA and sparse coding methods operate globally on the entire image. So, the unoccluded image that they generate can differ from the unoccluded part of the target image. Applying local replacement of only the occluded part, as for [13, 35], can mitigate this problem. But blending of the replaced pixels with their surrounding pixels in the unoccluded part to remove apparent seams still causes corruption to the surrounding pixels.

Our proposed method may be applied to video background recovery if the foreground area to be removed is known. In contrast, there are existing RPCA methods for video background recovery that do not require prior knowledge of the foreground area. For example, we have previously developed a fixed-rank RPCA method that does not require variable thresholds, and it has been shown to outperform regular low-rank RPCA for video background recovery [15]. Other variations of RPCA, such as RPCA via principal component pursuit, has also been developed for video background recovery [1].

3. Variable-Threshold Robust PCA

Given a data matrix $\mathbf{D}$, robust PCA (RPCA) decomposes $\mathbf{D}$ into a low-rank matrix $\mathbf{A}$ and a sparse error matrix $\mathbf{E}$ by

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1, \text{ subject to } \mathbf{D} = \mathbf{A} + \mathbf{E}, \quad (1)$$

where $\|\cdot\|_*$ denotes the nuclear norm and $\|\cdot\|_1$ denotes the l_1-norm. Wright et al. [31] show that $\mathbf{A}$ can be exactly recovered if $\mathbf{A}$ is sufficiently low-rank and $\mathbf{E}$ is sufficiently sparse. This minimization problem can be solved in several ways. In particular, the augmented Lagrange multiplier (ALM) method, which reformulates Eq. 1 into

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 + \langle \mathbf{Y}, \mathbf{D} - \mathbf{A} - \mathbf{E} \rangle + \frac{\mu}{2}\|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2,$$
$$(2)$$

has been shown to be among the most efficient and accurate methods [20], and is widely used as discussed in Section 2. In Eq. 2, $\mathbf{Y}$ contains the Lagrange multipliers, $\langle \mathbf{U}, \mathbf{V} \rangle$ is the sum of the product of corresponding elements in $\mathbf{U}$ and $\mathbf{V}$, and λ and μ positive parameters.

An important operator used in various implementations of RPCA, such as iterative thresholding, augmented Lagrange multipliers and principal component pursuit, is the

soft thresholding or *shrinkage* operator [4, 20]:

$$T_\varepsilon(x) = \begin{cases} x - \varepsilon, & \text{if } x > \varepsilon, \\ x + \varepsilon, & \text{if } x < -\varepsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

With this shrinkage operator, [3, 9] show that, for matrix $\mathbf{M}$ with SVD $\mathbf{USV}^\top$,

$$\mathbf{U}\,T_\varepsilon(\mathbf{S})\mathbf{V}^\top = \arg\min_{\mathbf{X}} \varepsilon\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{M} - \mathbf{X}\|_F^2, \quad (4)$$

$$T_\varepsilon(\mathbf{M}) = \arg\min_{\mathbf{X}} \varepsilon\|\mathbf{X}\|_1 + \frac{1}{2}\|\mathbf{M} - \mathbf{X}\|_F^2. \quad (5)$$

Then, comparing Eq. 4 and 5 with Eq. 2 yields

$$\mathbf{A} = \mathbf{U}\,T_\varepsilon(\mathbf{S})\mathbf{V}^\top, \text{ with } \mathbf{USV}^\top = \mathbf{D} - \mathbf{E}, \quad (6)$$

$$\mathbf{E} = T_\varepsilon(\mathbf{D} - \mathbf{A}). \quad (7)$$

Therefore, the shrinkage operator provides a means of computing optimal $\mathbf{A}$ and $\mathbf{E}$ according to Eq. 6 and 7.

For deocclusion of face image, $\mathbf{D} = [\mathbf{T}\ \mathbf{x}]$ contains n unoccluded training face images arranged in columns in $\mathbf{T}$ and an occluded target face image $\mathbf{x}$. So, $\mathbf{E}$ can be split into three parts:

$$\mathbf{E} = \begin{cases} T_{\varepsilon_t}(\mathbf{D} - \mathbf{A}), & \text{for training images } \mathbf{T}, \\ T_{\varepsilon_o}(\mathbf{D} - \mathbf{A}), & \text{for occluded part of } \mathbf{x}, \\ T_{\varepsilon_u}(\mathbf{D} - \mathbf{A}), & \text{for unoccluded part of } \mathbf{x}, \end{cases} \quad (8)$$

The error of the occluded part of target $\mathbf{x}$ is expected to be larger than that of the unoccluded part. So, the soft threshold for the occluded part ε_o should be small whereas that for the unoccluded part ε_u should be large. The soft threshold for the training images ε_t can take on the default value.

Define an $m \times (n+1)$ weight matrix $\mathbf{W} = [w_{ij}]$ with

$$w_{ij} = \begin{cases} w_t, & \text{training images,} \\ w_o, & \text{occluded part of target image,} \\ w_u, & \text{unoccluded part of target image.} \end{cases} \quad (9)$$

Then, applying exact ALM, our proposed *variable-threshold RPCA* method can be summarized as follows:

VRPCA

Input: D, W.

1. $\mathbf{A} = \mathbf{0}$, $\mathbf{E} = \mathbf{0}$, $\lambda > 0$, $\mu > 0$, $\rho > 1$.
2. $\mathbf{Y} = \text{sgn}(\mathbf{D})/J(\text{sgn}(\mathbf{D}))$.
3. Repeat until convergence:
4.   Repeat until convergence:
5. $\mathbf{U}, \mathbf{S}, \mathbf{V} = \text{SVD}(\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu)$.
6. $\mathbf{A} = \mathbf{U}\,T_{1/\mu}(\mathbf{S})\mathbf{V}^\top$.
7. $\mathbf{E} = T_{(\lambda/\mu)\mathbf{W}}(\mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu)$.
8. $\mathbf{Y} = \mathbf{Y} + \mu(\mathbf{D} - \mathbf{A} - \mathbf{E})$, $\mu = \rho\mu$.

Output: A, E.

In Line 2, $\text{sgn}(\cdot)$ computes the sign of each matrix element, and $J(\cdot)$ computes a scaling factor as recommended in [20]:

$$J(\mathbf{X}) = \max\left(\|\mathbf{X}\|_2, \lambda^{-1}\|\mathbf{X}\|_\infty\right). \quad (10)$$

Lines 5 to 7 are derived from Eq. 6 and 8. Line 7 applies a different weight to the soft threshold λ/μ for different elements of $\mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu$. For training images, $w_t = 1$. For the target image, empirical tests show that w_o and w_u can be set quite independently (Section 4.2). When w_o is sufficiently small, the mean-squared error of the recovered image with respect to ground truth for the occluded part is minimized. When w_u is sufficiently large, the corresponding elements in $\mathbf{E}$ of the unoccluded part are close to zero due to the shrinkage operator, and thus the corresponding elements in $\mathbf{A}$ are practically unchanged. When $w_u = w_o = 1$, VRPCA reverts to RPCA for *matrix recovery* via ALM.

In the extreme case with $w_o = 0$ and $w_t = w_u \to \infty$, Line 7 becomes

$$\mathbf{E} = \begin{cases} \mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu, & \text{occluded target,} \\ 0, & \text{unoccluded target,} \\ 0, & \text{training.} \end{cases} \quad (11)$$

Since the elements in $\mathbf{E}$ of the unoccluded part is 0, the corresponding elements in $\mathbf{A}$ remain unchanged. This variation of VRPCA is equivalent to RPCA for *matrix completion* via ALM [20]. Matrix completion is the problem of filling in the missing elements of a matrix given the available elements [5, 20]. By regarding the occluded parts of a face image as missing elements, face occlusion removal can be naturally framed as a matrix completion problem.

Various kinds of matrix completion algorithm have been proposed for solving computer vision problems. For example, [21] applies matrix factorization to depth map enhancement, [37] applies fixed point iteration to illumination compensation, and [2] applies fixed point continuation to image classification. [36] is related to our method, but it applies ALM to decompose images into three parts: common part, low-rank part, and sparse error part. It is actually more similar to matrix recovery than matrix completion as defined in [5, 20]. To our best knowledge, matrix completion has not been applied to face image deocclusion and face recognition. Thus, it is good that VRPCA includes RPCA for matrix completion as a special case.

An alternative to VRPCA is to apply w_t, w_u and w_o as weights to the elements of $\mathbf{E}$ instead of applying different soft thresholds. This method is less desirable because error scaling does not necessarily minimize Eq. 5. For error scaling to work, the elements of $\mathbf{E}$ corresponding to the occluded part has to be scaled to 0 while those of the unoccluded part and training images are scaled by 1 (i.e., unchanged), which is the same as VRPCA2 (Eq. 11).

VRPCA, as for other RPCA methods, work on the whole data matrix, in this case, the whole image including both the occluded and unoccluded parts. A straightforward alternative is to work on only the occluded part, which would be much more efficient. This alternative, however, has several shortcomings. First, it cannot make use of possible correlations among the occluded and unoccluded parts. The methods of [13] and [35], which generate results only for the occluded pixels and image blocks, also need to look for correlated pixels and blocks in the unoccluded parts of the training images. In VRPCA, consideration of correlated pixels is achieved through the SVD of $\mathbf{D} - \mathbf{E}$ in Line 5.

Second, local replacement of pixels and blocks, as performed in [13, 35], does not consider global consistency over the entire image. It results in apparent distortions, discontinuities and seams that need to be removed by smoothing or blending with surrounding unoccluded pixels (Section 4). Smoothing and blending cause unintended corruption of unoccluded pixels (Section 4). In VRPCA, the SVD and alternative updating of $\mathbf{A}$ and $\mathbf{E}$ retains global consistency and remove the need for smoothing and blending.

4. Experiments

4.1. Data Preparation

Face images were prepared for the tests. 120 images, one for a different individual, were randomly selected from CMU Multi-PIE database [8]. 100 of them formed the *training set* and the other 20 formed the *unfamiliar testing set*. In addition, 20 individuals in the training set were randomly selected, and for each individual, an image different from the training image was randomly selected from Multi-PIE database to form the *familiar testing set*. In other words, 2 different images were selected for these 20 individuals, one as training image and the other as familiar testing image. This arrangement allowed for comparing test results of familiar and unfamiliar cases, which could occur in real applications. The face images were resized to $120{\times}100$ pixels. Then, they were spatially aligned by thin plate spline method using feature points extracted from the face images by Face++ toolkit [7].

Rectangular occluders were manually placed on each image in the two testing sets to indicate the occluded regions, while the original unoccluded versions served as ground truth. Various occluder characteristics were applied, including 3 shapes (block, vertical bar, and horizontal bar), 5 size ratios, 9 intensity levels (0 to 255 in intervals of 32), and 6 locations (left eye, right eye, nose, left cheek, right cheek, and mouth). The 5 size ratios were measured as the ratio of the occluder's area over image area, and included 0.05, 0.1, 0.2, 0.3, and 0.4 (Fig. 1).

In practice, a simple way for a user to manually mark real occluders, such as eyeglasses, scarfs, beards, and face masks, in a target image is to place regularly-shaped blocks over the occlusions. So, the above test conditions are relevant to the removal of real occluders.

4.2. Determination Soft Threshold Weights

VRPCA has three weight parameters for soft thresholding. The weight w_t for training images is set to 1 because training images have no error. The weights w_o and w_u for the occluded and unoccluded parts, respectively, should be set appropriately. To determine the best weight values, a test was performed on VRPCA with varying weight values under the test condition of medium-sized black occluder in the middle of the image (Fig. 1(1a)).

Table 1 shows that with sufficiently large w_u and sufficiently small w_o, their actual values do not affect VRPCA's error significantly. In particular, when $w_u \geq 10$, the error of the unoccluded parts is practically zero for a wide range of values of w_o. The error of the occluded parts attains the smallest value when w_o is close to but larger than 0. So, we set $w_o = 0.0001$ and $w_u = 10$ for subsequent tests.

4.3. Test Procedure

Six algorithms, covering the main categories of methods for face deocclusion, namely RPCA, PCA, and sparse coding, were tested:

- VRPCA: our proposed variable-threshold RPCA.
- VRPCA2: variation of VRPCA that is equivalent to RPCA for matrix completion of [20].
- RPCA: RPCA for matrix recovery, similar to [23] except [23] applies hard instead of soft thresholding.
- FW-PCA: fast weighted PCA of [10].
- BC: block PCA with block correlation of [35].
- RSC: robust sparse coding of [34].

For VRPCA, VRPCA2 and RPCA, the parameters ρ and initial μ were set to the default values of 6 and $0.5/\sigma_1$, where σ_1 is the largest singular value of the initial $\mathbf{Y}$. The parameter λ was set to the theoretical optimum of $1/\sqrt{\max(m,n)}$ [4]. For the other methods, their parameter values were set according to recommendation.

In each test case, each algorithm was executed with all the unoccluded training images and an occluded testing image to produce a recovered unoccluded image. Mean-squared errors (MSE) between the recovered images and their ground truth were computed and averaged over all testing images. Mean-squared errors for the occluded parts and unoccluded parts were computed separately.

4.4. Results and Discussions

First, let us examine the effect of varying occluder intensity on the occluded parts of the test images (Fig. 2(a), top row). The errors of VRPCA, VRPCA2, FW-PCA, and BC
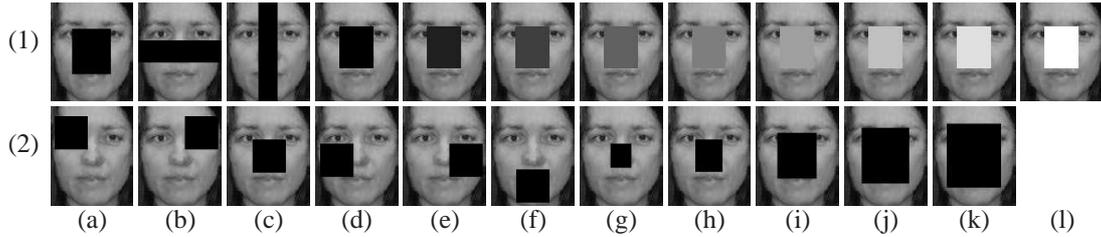
Figure 1. Occluders vary in (1a–c) shape, (1d–l) intensity, (2a–f) location, and (2g–k) size ratio.

Table 1. Effect of weights of soft thresholds on VRPCA's error (MSE).

(a) Occluded parts

w_u	w_o					
	0	0.0001	0.001	0.01	0.1	1
1	74.36	11.16	11.16	11.17	12.28	74.36
5	131.38	11.38	11.39	11.44	13.06	75.22
10	131.38	11.39	11.39	11.44	13.06	75.22
100	131.38	11.39	11.39	11.44	13.06	75.22
1000	131.38	11.39	11.39	11.44	13.06	75.22
10000	131.38	11.39	11.39	11.44	13.06	75.22

(b) Unoccluded parts

w_u	w_o					
	0	0.0001	0.001	0.01	0.1	1
1	10.91	9.98	9.98	9.97	9.93	10.91
5	0	0.2	0.2	0.2	0.2	0
10	0	0	0	0	0	0
100	0	0	0	0	0	0
1000	0	0	0	0	0	0
10000	0	0	0	0	0	0

are independent of occluder intensity for both familiar and unfamiliar images. Among them, VRPCA and VRPCA2 attain the smallest error whereas BC has the largest error. The errors of RSC and RPCA are affected by occluder intensity. For very high or very low occluder intensity, which are sufficiently different from the intensity of the facial features in the image, RSC can attain small error comparable to those of VRPCA and VRPCA2. On the other hand, for mid-range occluder intensity which is similar to facial intensity, the error of RSC can be very large even for familiar faces. RPCA's error is among the largest for small occluder intensity and approaches the error rate of FW-PCA for large occluder intensity.

For the unoccluded parts, VRPCA and VRPCA2 have practically zero error due to the high threshold of unoccluded parts. BC has the second lowest error because it does not change the unoccluded part, except for the pixels that are blended with the replaced blocks. On the other hand, RSC, FW-PCA, and RPCA have large errors for the unoccluded parts because they replace the pixels in the entire images. The intensity effect on RSC for the unoccluded part is similar to that for the occluded part.

Occluder size has varying effects on the tested algorithms. For the occluded parts, VRPCA, VRPCA2, and RSC are least affected and they attain the lowest errors. Surprisingly, FW-PCA has large error on the familiar images that unfamiliar images, and RPCA is severely affected by occluder size. For the unoccluded parts, VRPCA and VRPCA attain practically zero error while BC has very small error. On the other hand, RSC, FW-PCA, and RPCA have large errors for the unoccluded parts. The algorithms' errors on the unoccluded parts under varying occluder size

are consistent with those under varying occluder intensity.

With regard to the effect of occluder location, for the occluded parts, all methods have larger errors at the left and right eyes than at the other areas (Fig. 3(a)). VRPCA, VRPCA2, and RSC attain the smallest errors, whereas BC and RPCA have the largest errors. For the unoccluded parts, VRPCA and VRPCA2 attain practically zero error, while BC attains the second smallest error, and the other methods have large errors. These results are consistent for both familiar and unfamiliar images. Test results on the effect of occluder shape are similar to those of occluder location,

Figure 4 shows sample test results for the unfamiliar test set. The images recovered by VRPCA and VRPCA2 are the most similar to the ground truth. whereas those recovered by RSC and FW-PCA look more like smoothed versions of the ground truth. Moreover, when occluder intensity similar to facial intensity, RSC's result has large error (Fig. 4(row 2, e2). The results of BC and RPCA have clearly visible distortions. In particular, the distortions of BC's results are due to local replacement of occluded blocks.

In summary, VRPCA and VRPCA2 have the most accurate and consistent performance across various test conditions because the different soft thresholds for the occluded and unoccluded parts allow the methods to keep the unoccluded part practically unchanged while recovery accurate pixel values for the occluded part. Moreover, they optimize over the whole target image, and thus produce an unoccluded image that is globally consistent over the whole image. They consistently attain the smallest errors compared to other methods, and have practically no error for the unoccluded parts because they can preserve the unoccluded parts. RSC is strongly affected by occluder intensity
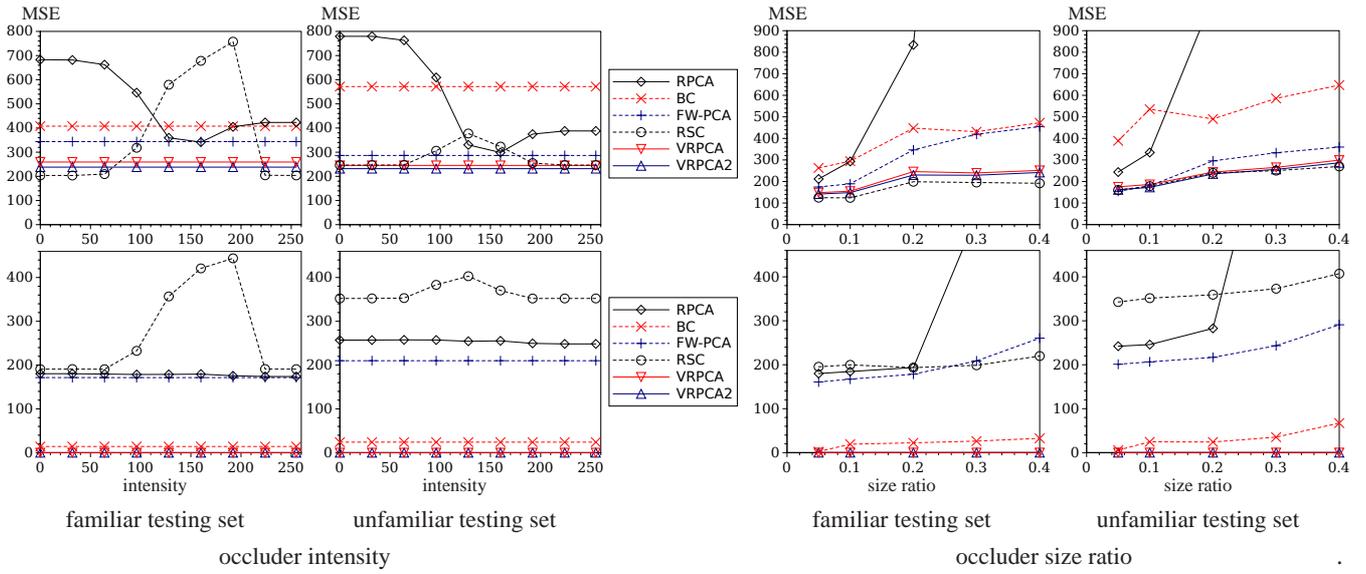
Figure 2. Test results with respect to occluder intensity and occluder size ratio for (top) occluded parts and (bottom) unoccluded parts. The graphs are cropped at large MSE to show the low-error lines more clearly.
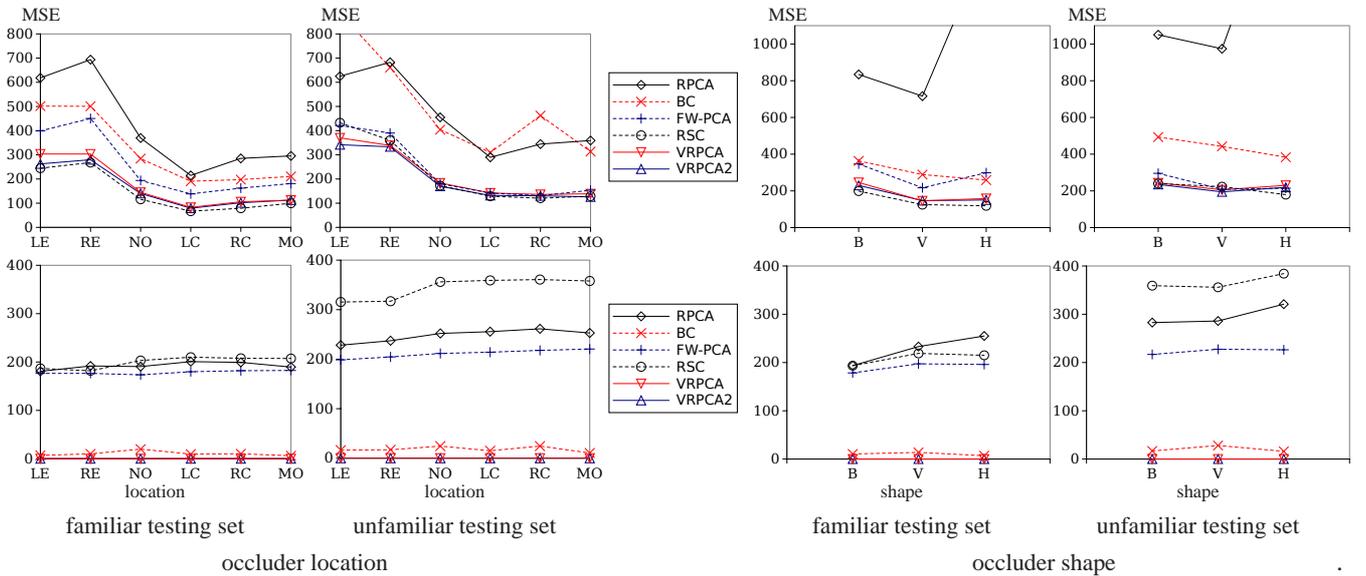

Figure 3. Test results with respect to occluder location and occluder shape for (top) occluded parts and (bottom) unoccluded parts. Locations include (LE) left eye, (RE) right eye, (NO) nose, (LC) left cheek, (RC) right cheek, and (MO) mouth. Shapes include (B) block, (V) vertical bar, and (H) horizontal bar. The graphs are cropped at large MSE to show the low-error lines more clearly. Lines joining data points are meant as visual aid only.

because it has difficulty distinguishing mid-tone occluders and facial features. When the occluder intensity is very different from the facial intensity, RSC can attain good results. However, its error of the unoccluded part is large because it cannot preserve the unoccluded parts. FW-PCA has moderately large error because PCA is not robust to large amplitude noise. BC can also preserve the unoccluded parts but its error is large for the occluded parts because the blocks

are replaced independently without regard to consistency within the entire face image. In other words, the selected blocks may be locally optimal within each block but not globally optimal in the entire image. RPCA has among the largest errors because it corresponds to a version of VR-PCA whose soft thresholds are 1 for both the occluded and unoccluded parts, which are inappropriate soft thresholds.

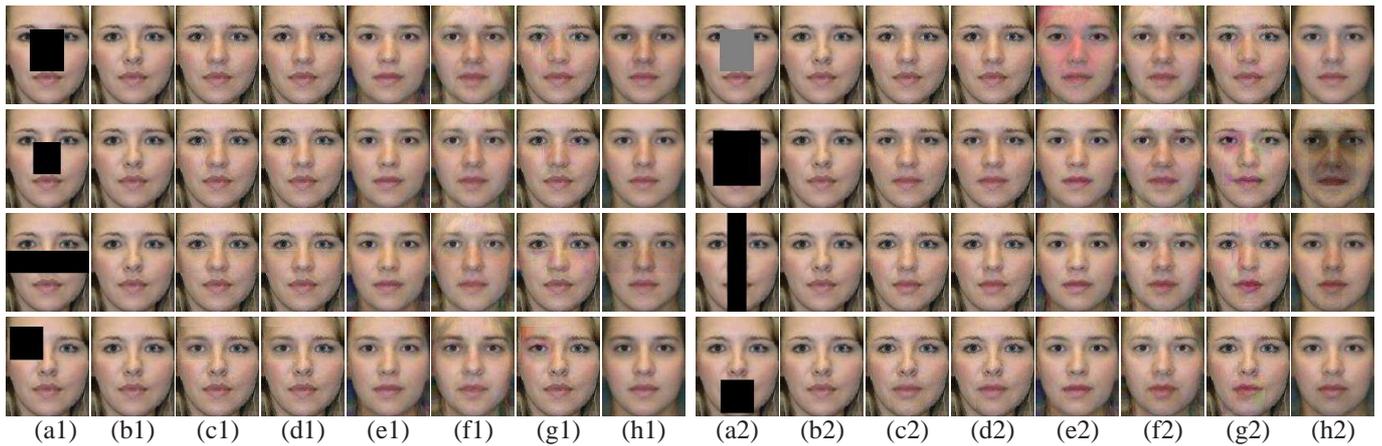(a1) (b1) (c1) (d1) (e1) (f1) (g1) (h1) (a2) (b2) (c2) (d2) (e2) (f2) (g2) (h2)

Figure 4. Sample deocclusion results of color images. Occluders vary in (row 1) intensity, (row 2) size ratio, (row 3) shape, and (row 4) location. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC (f) FW-PCA, (g) BC, and (h) RPCA.

5. Conclusion

This paper presented a variable-threshold RPCA (VR-PCA) method for removing occlusions in face images. VR-PCA is based on RPCA via ALM, which decomposes a data matrix containing unoccluded training images and an occluded target image into a low-rank matrix that contains only unoccluded images and a sparse error matrix that contains noise and occluders. It offers two variations with different soft thresholds for the training images and the occluded and unoccluded parts of the target image. Comprehensive tests show that both variations are consistently more accurate than existing methods across various test conditions. Their accuracies are unaffected by occluder intensity and minimally affected by occluder size and shape. Moreover, unlike existing methods, they can preserve the unoccluded part of the target image with practically zero error. The variable soft thresholds also provide additional constraints that allow VRPCA to perform well even when the data matrix is not exactly low-rank.

Acknowledgment

References

[1] T. Bouwmans and E. Zahzah. Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance. *Computer Vision and Image Understanding*, 122:22–34, 2014.

[2] R. Cabral, F. D. la Torre, J. P. Costeira, and A. Bernardino. Matrix completion for weakly-supervised multi-label image classification. *IEEE Trans. PAMI*, 37(1):121–135, 2015.

[3] J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optimization*, 20(4):1956–1982, 2010.

[4] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of ACM*, 58(3):11, 2011.

[5] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009.

[6] C. F. Chen, C. P. Wei, and Y. C. F. Wang. Low-rank matrix recovery with structural incoherence for robust face recognition. In *Proc. CVPR*, pages 2618–2625, 2012.

[7] Face++ toolkit. www.faceplusplus.com.

[8] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-PIE. *Image & Vision Computing*, 28(5):807–813, 2010.

[9] E. T. Hale, W. Yin, and Y. Zhang. A fixed-point continuation method for l_1-regularized minimization with applications to compressed sensing. Technical Report TR07-07, Dept. of Computational and Applied Math., Rice University, 2007.

[10] T. Hosoi, S. Nagashima, K. Kobayashi, K. Ito, and T. Aoki. Restoring occluded regions using fw-pca for face recognition. In *Proc. CVPR Workshop*, pages 23–30, 2012.

[11] B. W. Hwang and S. W. Lee. Reconstruction of partially damaged face images based on a morphable face model. *IEEE Transactions on PAMI*, 25(3):365–372, 2003.

[12] X. Jiang and J. Lai. Sparse and dense hybrid representation via dictionary decomposition for face recognition. *IEEE Trans. PAMI*, 37(5):1067–1079, 2015.

[13] J.-E. Lee and N. Kwak. Detection and recovery of occluded face images based on correlation between pixels. In *Proc. Int. Conf. PRAM*, pages 568–571, 2012.

[14] A. Leonardis and H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78:99–118, 2000.

[15] W. K. Leow, Y. Cheng, L. Zhang, T. Sim, and L. Foo. Background recovery by fixed-rank robust principal component analysis. In *Proc. CAIP*, 2013.

[16] L. Li, S. Li, and Y. Fu. Discriminative dictionary learning with low-rank regularization for face recognition. In *Proc. CVPR*, 2013.

[17] X.-X. Li, D.-Q. Dai, X.-F. Zhang, and C.-X. Ren. Structured sparse error coding for facing recognition with occlusion. *IEEE. Trans. Image Processing*, 22(5):1889–1900, 2013.

[18] Y. Li, J. Liu, H. Lu, and S. Ma. Learning robust face representation with classwise block-diagonal structure. *IEEE Trans. Info Forensics and Security*, 9(12):2051–2062, 2014.

[19] D. Lin and X. Tang. Quality-driven face occlusion detection and recovery. In *Proc. CVPR*, 2007.

[20] Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. Technical Report UILU-ENG-09-2215, Univ. Illinois at Urbana-Champaign, 2009.

[21] S. Lu, X. Ren, and F. Liu. Depth enhancement via low-rank matrix completion. In *Proc. CVPR*, 2014.

[22] L. Ma, C. Wang, B. Xiao, and W. Zhou. Sparse representation for face recognition based on discriminative low-rank dictionary learning. In *Proc. CVPR*, pages 2586–2593, 2012.

[23] R. Min and J. Dugelay. Inpainting of sparse occlusion in face recognition. In *Proc. ICIP*, pages 1425–1428, 2012.

[24] Z. Mo, J. P. Lewis, and U. Neumann. Face inpainting with local linear representations. In *Proc. BMVC*, 2004.

[25] M. H. Nguyen, J.-F. Lalonde, A. A. Efros, and F. de la Torre. Image-based shaving. *Computer Graphics Forum*, 27(2):627–635, 2008.

[26] J.-S. Park, Y. H. Oh, S. C. Ahn, and S.-W. Lee. Glasses removal from facial image using recursive error compensation. *IEEE Trans. PAMI*, 27(5):805–811, 2005.

[27] J. Qian, J. Yang, F. Zhang, and Z. Lin. Robust low-rank regularized regression for face recognition with occlusion. In *Proc. CVPR Workshop*, 2014.

[28] Y. Saito, Y. Kenmochi, and K. Kotani. Estimation of eyeglassless facial images using principal component analysis. In *Proc. ICIP*, pages 197–202, 1999.

[29] M. Storer, P. Roth, M. Urshiler, and H. Bischof. Fast-robust pca. In *Proc. European Conf. Opt. Commun. (LNCS 5575)*, pages 130–139, 2009.

[30] Z.-M. Wang and J.-H. Tao. Reconstruction of partially occluded face by fast recursive pca. In *Proc. Int. Conf. on Computational Intelligence and Security Workshops*, pages 304–307, 2007.

[31] J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao. Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization. In *Proc. NIPS*, pages 2080–2088, 2009.

[32] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. PAMI*, 31(2):210–227, 2009.

[33] C. Wu, C. Liu, H.-Y. Shum, Y.-Q. Xu, and Z. Zhang. Automatic eyeglasses removal from face images. *IEEE Trans. PAMI*, 26(3):322–336, 2004.

[34] M. Yang, L. Zhang, J. Yang, and D. Zhang. Robust sparse coding for face recognition. In *Proc. CVPR*, pages 625–632, 2011.

[35] D. Yu and T. Sim. Using targeted statistics for face regeneration. In *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2008.

[36] Q. Zhang and B. Li. Joint sparsity model with matrix completion for an ensemble of face images. In *Proc. ICIP*, 2010.

[37] X. Zhang and H. Xiong. Illumination compensation via low rank matrix completion for multiview video coding. In *Proc. ICIP*, 2013.