FACE IMAGE DE-OCCLUSION WITH VARIABLE-THRESHOLD ROBUST PCA

LI GUODONG

(B.Sc., University of Science and Technology of China, China, 2014)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

NATIONAL UNIVERSITY OF SINGAPORE

2016

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

Signature: Li Guodong **Date:** 2016/9/15

Li Guodong

2016

Acknowledgements

First and foremost, it is hard to overstate my heartfelt and sincere gratitude towards my supervisor, Professor Leow Wee Kheng for the guidance, encouragement and advice he has provided throughout my time as his student. I am extremely gratified and consider myself very fortunate for being able to work with an extremely conscientious and responsible supervisor who cared deeply about my work, and responded to my questions and queries so promptly. I have learnt a lot from him not only scientifically but in every aspect of life. Prof. Leow helped me choose an exciting research topic and trained me to strengthen the thinking and innovation abilities which will benefit throughout my life. More than just being my teacher, he has been approachable like a friend and extending a helping hand whenever I had problems. His positive outlook, his inspirational guidance, his enthusiasm in research, his pearls of wisdom and his incredibly unselfish support has tremendously helped me in obtaining the master's degree.

Thanks to Prof. Terence Sim and Prof. Michael Brown for their guidance and advice.

I would also like to extend many thanks to Dr. Li Jian for his kind and thoughtful help.

Huge appreciation to all my brilliant lab mates and friends, their company really makes my day here.

Finally, I want to thank my family for their endless love and support.

Contents

Summary

The human face is one of the most commonly seen and used biometric traits in our daily life. With the development of cameras and digital electronics, face images are more easily generated and collected. As a result, face images are becoming more and more important in human social activities such as security, surveillance, identification, criminal and forensic investigation. These activities can be severely hampered when face images are corrupted by occluders such as hairs, eyeglasses, face masks, and scarves. Although some algorithms can handle face identification or recognition with occlusion, they still suffer from performance degradation due to face occlusion. Moreover, unoccluded face images of victims and suspects are still needed for investigation, monitoring, and communication to the public. Therefore, removal of occlusions in face images is a very important task.

Unfortunately, removing facial occlusions is extremely difficult and challenging. The reconstruction method has to predict the unoccluded parts. Although human faces generally have similar shapes and appearances, the feature details may differ greatly among people from different races, genders and ages, which makes recovering unoccluded face images a very difficult task.

Existing methods for removing occlusions in face images can be grouped into three broad categories, namely principal component analysis (PCA), robust PCA (RPCA), and sparse coding. There are two major shortcomings in these existing methods: (1) their performances are not consistent across varying test conditions, and (2) the unoccluded parts of target images can be changed by the algorithms, resulting in unintended corruption to the unoccluded parts.

This thesis aims to overcome the difficulties of existing methods for face occlusion removal by the proposed variable-threshold RPCA (VRPCA) method. Given prior knowledge of the occluded area, VRPCA separates the error matrix into three different parts: the training part, the testing unoccluded part and the testing occluded part. By applying different thresholds to these three parts, VRPCA ensures that the errors are mostly from the occluded part and that unintended corruptions is minimized in the reconstruction of the unoccluded parts.

To evaluate the proposed method, this thesis studies various occluder characteristics such as intensity, size, shape, and location that can affect the performance of occlusion removal algorithms. Test results show that VRPCA is consistently more accurate than existing methods across various test conditions for recovering occluded parts. Moreover, VRPCA is able to preserve the unoccluded parts of the target image with practically zero error.

List of Publications

W. K. Leow, **G. Li**, J. Lai, T. Sim and V. Sharma. Hide and Seek: Uncovering Facial Occlusion with Variable-Threshold Robust PCA. In *Proceedings of IEEE Winter Conference on Application of Computer Vision*, pages 1–8, 2016.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The human face is one of the most commonly seen and used biometric traits in our daily life. As an important means of identity, friends and acquaintances can recognize each other in a distance. The police posts wanted posters to find criminals. Nowadays, with the development of cameras and digital electronics, face images can be acquired easily. In fact, the government and private organizations keep face photo collection of people in the form of ID cards, passports and student cards. Face images are becoming more and more important in human social activities including security, surveillance, identification, criminal and forensic investigation.

These activities, however, are severely hampered when face images are occluded. In real-world environments, faces are easily occluded either actively or passively. For example, the face can be covered by other people in the crowd or human parts such as hairs and hands. It is also very common for people to wear facial accessories such as sunglasses, scarves, hats, masks and veils for personal or cultural purposes (Figure 1.1). One can require subjects to remove these accessories for face identification. But this will give rise to inconvenience. Besides, in security related scenarios, criminals tend to use disguise to hide their identities and subject cooperation is not applicable at all. Existing face recognition systems can recognize some occluded

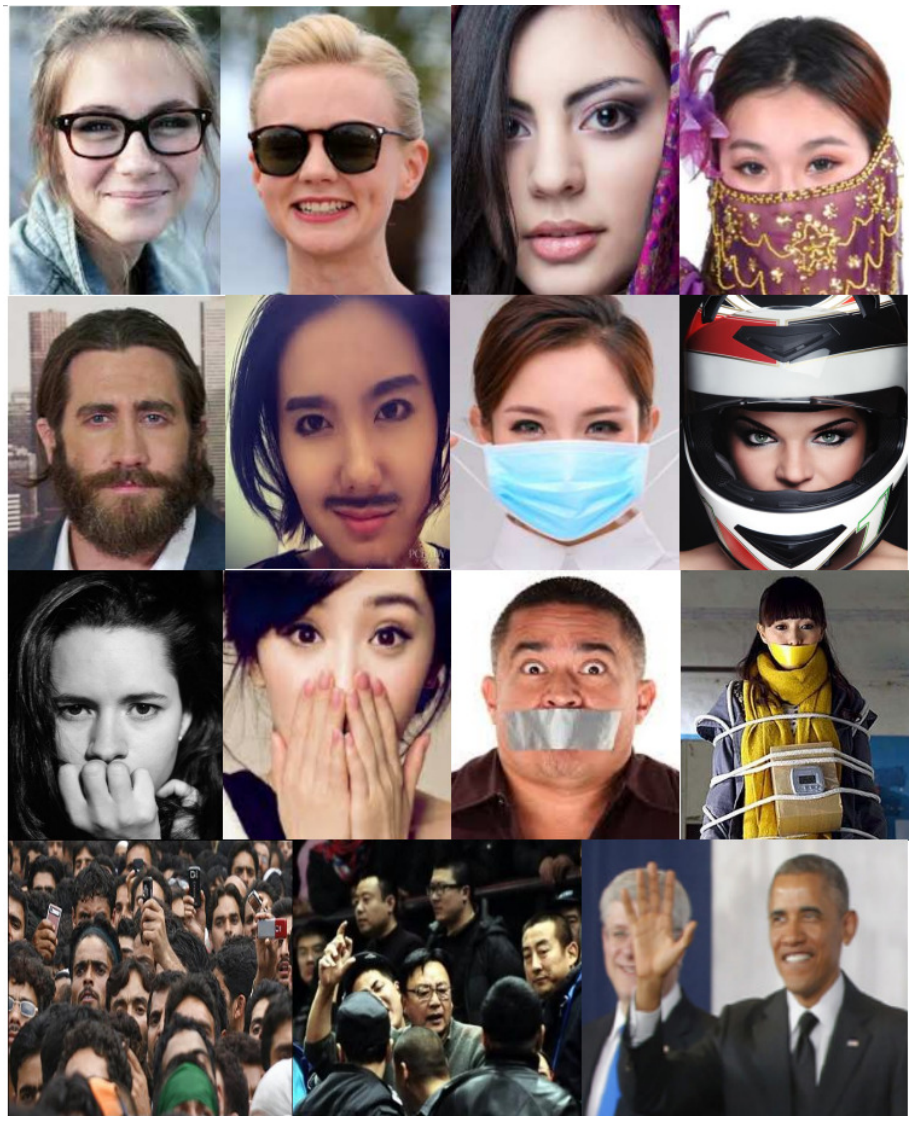


Figure 1.1: Examples of face occlusion in real life.

face images. However, they suffer significant performance degradation with unconstrained occlusions. Moreover, unoccluded face images of victims and suspects are still needed for investigation, monitoring, and communication to the public. Therefore, uncovering facial occlusions is an important problem to solve. In this thesis, we call this problem face image de-occlusion.

Face image de-occlusion is a challenging problem. The original shape and texture of occluded parts are unknown. When the unknown parts have complex textures, it is difficult and complex to predict and generate. For a face image with large occlusions, it is also very difficult to recover the original shape and texture. The shape and texture of human face varies greatly across different age, gender and racial groups. Moreover, the pose, lighting and illumination of faces will also influence the computation process. During computation, the reconstruction algorithm should make the reconstruction as close as possible to the original unoccluded images and robust under various occlusion conditions. Furthermore, the reconstruction algorithm should ensure that the original unoccluded parts are unchanged and that there are no clear seams between the recovered parts and the original unoccluded parts. All the above issues make face image de-occlusion a challenging computational problem.

There are some different problems that are related to face image de-occlusion, for example image denoising and inpainting [DFKE07, LCN14, BSCB00, XS10]. They are meant for recovering face images with noise or scratches. In contrast, this thesis focuses on recovering face images with large and continual occlusions.

Existing face de-occlusion methods can be grouped into three broad categories: principal component analysis (PCA), robust PCA (RPCA), and sparse coding. PCA methods [HNK+12, HL03, LB00, MLN04, SRUB09, SKK99, NLET08, POAL05, WLS+04] 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 [MD12] 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. Sparse coding methods [LDZR13, YZYZ11] 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 performances are not consistent across varying test conditions, and (2) the occluded parts of target images can be changed by the algorithms, resulting in unintended corruption of the unoccluded parts.

1.2 Thesis Objective

The objective of this research is to develop an accurate and robust method for recovering unoccluded face images. To achieve this objective and overcome the difficulties of existing methods, this thesis develops an algorithm based on RPCA. It uses aligned training face images to predict the occluded part of the testing face image. With the prior knowledge of the occluded area, different soft thresholds can be applied to the occluded and unoccluded part of the RPCA error matrix. As a result, the errors can be constrained to the occluded part. To evaluate the proposed method, related state-of-the-art methods are compared under different occlusion conditions.

1.3 Outline of Thesis

The rest of this thesis is structured as follows: Chapter 2 discusses existing methods for face image de-occlusion, including PCA (Section 2.1), RPCA (Section 2.2) and sparse coding (Section 2.3). A summary and a comparison of all these methods are presented (Section 2.4). Chapter 3 presents the proposed solution to the face image de-occlusion problem. RPCA technique is reviewed first (Section 3.1), followed by the proposed VRPCA and its advantages (Section 3.2). Chapter 4 discusses experimental studies of VRPCA. Chapter 5 briefly describes some potential future work and Chapter 6 concludes the thesis.

Chapter 2

Related Work

Existing methods capable of removing occlusions in face images can be grouped into three broad categories: principal component analysis (PCA), robust PCA, and sparse coding. These methods will be discussed in Section 2.1, 2.2 and 2.3, respectively.

2.1 PCA

PCA finds a set of orthonormal vectors that maximize the variances of the data points that they represent. For face image de-occlusion problem, PCA methods first construct a face eigenspace using unoccluded training images. Given an occluded target image, they map the unoccluded part of the target image into the eigenspace to recover linear coefficients, and use the coefficients to generate unoccluded target image (e.g., [SKK99]).

Instead of applying standard PCA, [WT07] applies incremental PCA whereas [LT07, WLS$^+$04] apply probabilistic PCA. [ANS12] reconstructs unoccluded face images with asymmetric PCA. [LB00, SRUB09] apply random sub-sampling of images and [HNK$^+$12] applies weighted PCA. [NLET08] applies iteratively reweighted least-squares to compute the coefficients. In addition, some models based on PCA have been developed for removing occlusion. [HL03] proposes a morphable face

model estimating the optimal coefficients for linear combinations of prototypes of shape and appearance from the unoccluded region. The same linear combination is then applied to recover the occluded region. Most of these PCA methods required a single or multiple training set of unoccluded face images. In contrast, [NLET08, POAL05, WLS+04] require training pairs of occluded and unoccluded face images of some people, making these methods restrictive in real applications. Moreover, these methods recover a linear representation of the training face images, which is a smooth unoccluded face image.

PCA methods usually generate an entire unoccluded image based on computed coefficients. So, they can corrupt the unoccluded part of the target image. To mitigate this problem, [YS08] applies block PCA, and the coefficients of the occluded blocks are generated from those of the most correlated unoccluded block of the same face. Correlation between blocks is computed based on training images. Such local replacement, however, create apparent discontinuities or seams that need to be removed by blending the replace blocks with their surrounding pixels, which lead to corruption of the surrounding pixels.

2.2 Robust PCA

It is well known that PCA methods can be severely affected by large amplitude noise. Many methods have been proposed to make PCA more robust [XY95, DlTB01, SBL02, KK05, DZHZ06, Kwa08, WGR+09]. Among them, Robust PCA (RPCA) [WGR+09] is one of the most well-known robust methods which separates a data matrix into a low-rank matrix and sparse error matrix. [WGR+09] shows that the low-rank matrix can be exactly recovered by solving a convex optimization problem. The detailed discussion will be carried out in Chapter 3. RPCA has been successfully applied to many computer vision problems, such as video surveillance [LCZ+13], background recovering [BZ14], robust face alignment [PGW+12], subspace clustering [LLY10], textures transform invariant representation [ZLGM10] and 3D man-made objects alignment [JWL12].

RPCA is developed for machine recognition of occluded faces [LL14, WX10] rather than face de-occlusion per se. Nevertheless, some of these methods can recover unoccluded faces from occluded images, especially when the occluded part is sparse. [MD12] gives a definition of dense or sparse occlusion and solves the face occlusion removal when the occlusions are not dense. Examples of sparse occlusions include face dirt, and the face behind the fence and so on. Then RPCA method is applied to decompose a data matrix containing unoccluded training images and an occluded target image into a low-rank matrix containing unoccluded faces and a sparse matrix containing noise and occluders. However, the general RPCA method cannot solve the problem of dense occlusion removal. Moreover, the original unoccluded part can also be altered which causes unintended corruption.

2.3 Sparse Coding

Sparse coding is another robust method for processing multi-dimensional data. In contrast to PCA, sparse coding aims at representing the input image as a linear combination of training images whose coefficients are sparse. In the ideal case, only the coefficients that correspond to the training images from the same class as the input image are non-zero; the others are zero. Sparse coding methods can be applied for face recognition, and some of them can also handle the occluded face images hence there is an occlusion removal process. Given a target occluded face image, they compute the sparse coefficients to classify the subject in the image. If the individuals of the training classes look different and the training samples for each class are sufficiently large, most non-zero values for the coefficient vector fall in one class and the vector is naturally sparse. The sparse coefficients vector can be used to generate an unoccluded image of the target using the weighted sum of training images.

In practice, to find the sparsest solution, various methods have been proposed for computing the sparse coefficients. [WYG+09] applies l_1-norm minimization technique to find the sparsest representation. [YZYZ11] solves the face recognition

problem by a variation of the l_2-norm minimization technique, which adds a weight matrix to give the outliers low weight values. Based on the learnt class-specific and non-specific dictionary, [JL15] proposes to reconstruct the identity and variation component by l_1-norm and l_2-norm minimization respectively. In all these methods, a large number of unoccluded training images are required in the training sets. However, in many practical face recognition scenarios, the training images of each subject are often insufficient, especially when the face identity is not inside training datasets. Moreover, the unoccluded part can also be distorted during the separation due to global processing of the image.

2.4 Summary

Most of PCA, RPCA and sparse coding methods operate on the whole image. Therefore, the unoccluded part of the reconstructed image can differ from the unoccluded part of the target image. Applying local replacement of only for the occluded part, as for [YS08], can mitigate this problem. However, blending of the replaced pixels with their surrounding pixels in the unoccluded part to remove apparent seams still causes corruption to the surrounding pixels. Table 2.1 compares of the reviewed face image de-occlusion methods. Although the proposed method works on the whole image, it uses variable soft thresholds to constrain the errors of the unoccluded part of the target image to very small values. Therefore, the unoccluded part is practically unchanged in the recovered unoccluded image.

Table 2.1: Comparison of face image de-occlusion methods.

Methods	Citation	Approach	No corruption of unoccluded parts	Consistent to various cases
PCA	[LT07] [HNK+12]	Global Processing	No	No
	[YS08]	Local Replacement	Yes (except the surrounding)	Yes
RPCA	[MD12] [WGR+09]	Global Processing	No	No
SC	[WYG+09] [YZYZ11]	Global Processing	No	No
Proposed Method	NA	Global Processing	Yes	Yes

Chapter 3

Variable Threshold Robust PCA

3.1 Robust PCA

Given an observed data matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$, PCA seeks to find a low-rank matrix $\mathbf{A}$ such that the error matrix $\mathbf{E} = \mathbf{D} - \mathbf{A}$ is minimized:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{E}\|_F \text{ subject to } \operatorname{rank}(\mathbf{A}) \leq r, \mathbf{D} = \mathbf{A} + \mathbf{E}, \tag{3.1}$$

where $\|\cdot\|_F$ is the Frobenius norm and $r \leq \min(m, n)$ indicates the target rank of $\mathbf{A}$. This problem can be easily solved using Singular Value Decomposition (SVD). However, the solution will be vastly inaccurate when the error entries in $\mathbf{E}$ are arbitrarily large.

To tackle the problem, Robust PCA (RPCA) [WGR+09] is proposed whose main idea is a matrix decomposition problem:

$$\min_{\mathbf{A},\mathbf{E}} \operatorname{rank}(\mathbf{A}) + \lambda \|\mathbf{E}\|_0 \text{ subject to } \mathbf{D} = \mathbf{A} + \mathbf{E}, \tag{3.2}$$

where $\mathbf{A}$ is the recovered low-rank matrix, while $\mathbf{E}$ is a sparse error matrix capturing the outliers. λ is a positive balancing parameter. $\|\cdot\|_0$ is l_0-norm, which is defined as the number of non-zero entries. This rank operation and l_0-norm minimization problem is NP-hard. Under some mild conditions, however, the rank operation and

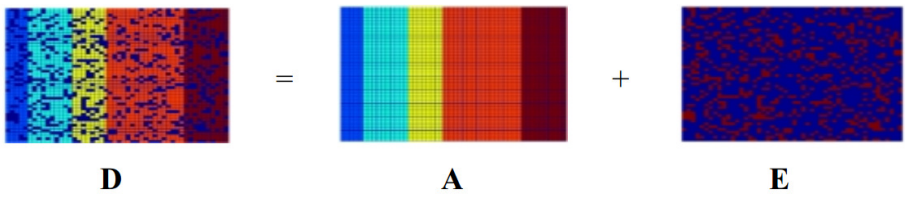


Figure 3.1: RPCA schematic diagram. RPCA separates the data matrix $\mathbf{D}$ to low-rank matrix $\mathbf{A}$ and sparse error matrix $\mathbf{E}$.

l_0-norm can be replace by nuclear norm and l_1-norm, respectively. As a result, the revised problem becomes:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 \text{ subject to } \mathbf{D} = \mathbf{A} + \mathbf{E}, \tag{3.3}$$

where l_1-norm of $\|\mathbf{E}\|_1$ is defined as the sum of absolute values of all entries in $\mathbf{E}$, while the nuclear norm $\|\mathbf{A}\|_*$ is the sum of singular values of $\mathbf{A}$.

This minimization problem can be solved in several ways. In particular, the augmented Lagrange multiplier (ALM) method has been shown to be one of the most efficient and accurate methods [LCM09], which reformulates the Formula 3.3 into:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 + \langle \mathbf{D} - \mathbf{A} - \mathbf{E}, \mathbf{Y} \rangle + \frac{\mu}{2} \|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2. \tag{3.4}$$

In Formula 3.4, $\mathbf{Y}$ contains the Lagrange multipliers, λ and μ are positive scalar which are parameters to be specified, and $\langle \cdot, \cdot \rangle$ is the element-wise product, with $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{tr}(\mathbf{X}^\top \mathbf{Y}) = \text{tr}(\mathbf{Y}^\top \mathbf{X})$. So Formula 3.4 is the same as:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 + \text{tr}(\mathbf{Y}^\top (\mathbf{D} - \mathbf{A} - \mathbf{E})) + \frac{\mu}{2} \|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2 . \tag{3.5}$$

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

is *soft thresholding* or *shrinkage* operator [CLMW11, LCM09]:

$$T_\varepsilon(x) = \begin{cases} x - \varepsilon, & \text{if } x > \varepsilon, \\ x + \varepsilon, & \text{if } x < -\varepsilon, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

With this shrinkage operator, [CCS10, HYZ07] show that, for matrix $\mathbf{M}$ with SVD $\mathbf{USV}^\top$,

$$\mathbf{U}T_\varepsilon(\mathbf{S})\mathbf{V}^\top = \arg\min_{\mathbf{X}} \varepsilon\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{M} - \mathbf{X}\|_F^2, \tag{3.7}$$

$$T_\varepsilon(\mathbf{M}) = \arg\min_{\mathbf{X}} \varepsilon\|\mathbf{X}\|_1 + \frac{1}{2}\|\mathbf{M} - \mathbf{X}\|_F^2. \tag{3.8}$$

The matrices $\mathbf{A}$ and $\mathbf{E}$ are optimized using alternating optimization method. Optimizing $\mathbf{A}$ with other variables fixed reduces Formula 3.5 to:

$$\begin{aligned} &\min_{\mathbf{A}} \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 + \mathrm{tr}(\mathbf{Y}^\top(\mathbf{D} - \mathbf{A} - \mathbf{E})) + \frac{\mu}{2}\|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2 \\ \Rightarrow &\min_{\mathbf{A}} \|\mathbf{A}\|_* + \mathrm{tr}(\mathbf{Y}^\top(\mathbf{D} - \mathbf{A} - \mathbf{E})) + \frac{\mu}{2}\|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2 \\ \Rightarrow &\min_{\mathbf{A}} \|\mathbf{A}\|_* + \mathrm{tr}(-\mathbf{Y}^\top\mathbf{E}) + \frac{\mu}{2}\mathrm{tr}((\mathbf{D} - \mathbf{A} - \mathbf{E})^\top(\mathbf{D} - \mathbf{A} - \mathbf{E})) \\ \Rightarrow &\min_{\mathbf{A}} \|\mathbf{A}\|_* + \frac{\mu}{2}\mathrm{tr}((\mathbf{E} - (\mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu))^\top(\mathbf{A} - (\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu))) \\ \Rightarrow &\min_{\mathbf{A}} \|\mathbf{A}\|_* + \frac{\mu}{2}\|\mathbf{A} - (\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu)\|_F^2 \\ \Rightarrow &\min_{\mathbf{A}} \frac{1}{\mu}\|\mathbf{A}\|_* + \frac{1}{2}\|(\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu) - \mathbf{A}\|_F^2. \end{aligned} \tag{3.9}$$

Comparing with Equation 3.7, by replacing ε with $\frac{1}{\mu}$, $\mathbf{X}$ with $\mathbf{A}$ and $\mathbf{M}$ with $(\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu)$, the optimized solution for Formula 3.9 is:

$$\mathbf{A} = \mathbf{U}T_{\frac{1}{\mu}}(\mathbf{S})\mathbf{V}^\top, \text{with } \mathbf{USV}^\top = \mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu. \tag{3.10}$$

Similarly, optimizing $\mathbf{E}$ with other variables fixed implies:

$$\begin{aligned}
&\min_{\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 + \mathrm{tr}(\mathbf{Y}^\top(\mathbf{D}-\mathbf{A}-\mathbf{E})) + \frac{\mu}{2}\|\mathbf{D}-\mathbf{A}-\mathbf{E}\|_F^2 \\
\Rightarrow &\min_{\mathbf{E}} \lambda \|\mathbf{E}\|_1 + \mathrm{tr}(\mathbf{Y}^\top(\mathbf{D}-\mathbf{A}-\mathbf{E})) + \frac{\mu}{2}\|\mathbf{D}-\mathbf{A}-\mathbf{E}\|_F^2 \\
\Rightarrow &\min_{\mathbf{E}} \lambda \|\mathbf{E}\|_1 + \mathrm{tr}(-\mathbf{Y}^\top\mathbf{E}) + \frac{\mu}{2}\mathrm{tr}((\mathbf{D}-\mathbf{A}-\mathbf{E})^\top(\mathbf{D}-\mathbf{A}-\mathbf{E})) \\
\Rightarrow &\min_{\mathbf{E}} \lambda \|\mathbf{E}\|_1 + \frac{\mu}{2}\mathrm{tr}((\mathbf{E}-(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu))^\top(\mathbf{E}-(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu))) \\
\Rightarrow &\min_{\mathbf{E}} \lambda \|\mathbf{E}\|_1 + \frac{\mu}{2}\|\mathbf{E}-(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu)\|_F^2 \\
\Rightarrow &\min_{\mathbf{E}} \frac{\lambda}{\mu}\|\mathbf{E}\|_1 + \frac{1}{2}\|(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu)-\mathbf{E}\|_F^2.
\end{aligned} \tag{3.11}$$

Comparing with Equation 3.8, by replacing ε with $\frac{\lambda}{\mu}$, $\mathbf{X}$ with $\mathbf{E}$ and $\mathbf{M}$ with $(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu)$, the optimized solution for Formula 3.11 is:

$$\mathbf{E} = T_{\frac{\lambda}{\mu}}(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu). \tag{3.12}$$

3.2 Variable-Threshold RPCA

For de-occlusion 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 $\mathbf{x}$. Given the prior knowledge of the occluded part, VRPCA separates the error matrix $\mathbf{E}$ into three parts for training images, occluded part of target image and unoccluded part of target image. By separating $\mathbf{E}$ it into three different parts, variable-threshold is introduced to calculate the error matrix:

$$\mathbf{E} = \begin{cases} T_{w_t\lambda/\mu}(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu), & \text{for training images } \mathbf{T}, \\ T_{w_o\lambda/\mu}(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu), & \text{for occluded part of } \mathbf{x}, \\ T_{w_u\lambda/\mu}(\mathbf{D}-\mathbf{A}+\mathbf{Y}/\mu), & \text{for unoccluded part of } \mathbf{x}. \end{cases} \tag{3.13}$$

Define the matrix $\mathbf{W} = [w_{ij}] \in \mathbb{R}^{m\times(n+1)}$ such that

$$w_{ij} = \begin{cases} w_t, & \text{for training images } \mathbf{T}, \\ w_o, & \text{for occluded part of } \mathbf{x}, \\ w_u, & \text{for unoccluded part of } \mathbf{x}. \end{cases} \tag{3.14}$$

Then Equation 3.13 can be written as $\mathbf{E} = T_{(\lambda/\mu)\mathbf{W}}(\mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu)$.

The error of the occluded part of target $\mathbf{x}$ is expected to be larger than that of the unoccluded part. So the weight of soft-threshold for the occluded part w_o should be small whereas that for the unoccluded part w_u should be large. The weight of soft-threshold for the training images w_t takes the default value of 1. Then, applying exact ALM, the proposed variable-threshold RPCA method can be summarized as follows:

VRPCA

Input: $\mathbf{D}$, $\mathbf{W}$.

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: $\mathbf{A}$, $\mathbf{E}$.

In Line 2, $\text{sgn}(\cdot)$ computes the sign of each matrix element, and $J(\cdot)$ computes a scaling factor as recommended in [LCM09]:

$$J(\mathbf{X}) = \max(\|\mathbf{X}\|_2, \lambda^{-1}\|\mathbf{X}\|_\infty). \tag{3.15}$$

Lines 5 to 7 are derived from Formula 3.12, Formula 3.10. Line 7 applies different weights 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

$$\mathbf{E} = \begin{cases} \mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu, & \text{occluded target,} \\ 0, & \text{unoccluded target,} \\ 0, & \text{training.} \end{cases} \tag{3.16}$$

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 [LCM09]. Matrix completion is the problem of filling in the missing elements of a matrix given the available elements [CR09, LCM09]. 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, [LRL14] applies matrix factorization to depth map enhancement, [ZL10] applies fixed point iteration to illumination compensation, and [CTCB15] applies fixed point continuation to image classification. Similar to the proposed method, [ZL10] applies ALM but decomposes 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 [CR09, LCM09]. To our best knowledge, matrix completion has not been applied to face image de-occlusion and face recognition. Thus, it is good that VRPCA includes RPCA for matrix completion as a special case.

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 method of [YS08], which generates results only for the occluded pixels and image blocks, also need to look for correlated pixels and blocks in unoccluded parts of the training images. In VRPCA, consideration of correlated pixels is achieved through SVD of $\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu$ in Line 5. Second, local replacement of pixels and blocks, as performed in [YS08], does not consider global consistency over the whole image. It results in apparent discontinuities and seams that need to be removed by smoothing or blending with surrounding unoccluded pixels. Smoothing and blending cause unintended corruption of unoccluded pixels. In VRPCA, the SVD and alternative updating of $\mathbf{A}$ and $\mathbf{E}$ retains global consistency and remove the need for smoothing and blending.

Chapter 4

Experiments and Discussions

4.1 Data Preparation

Two face databases were employed to test the algorithm, namely FERET [PMRR00] and CMU Multi-PIE [GMC+10]. The FERET database contained 1564 sets with 14,126 gray scale images in total, included 1199 individuals and 365 duplicate sets. A duplicate set contained images of a person taken at a different day. The CMU Multi-PIE face database contained 337 individuals, more than 750,000 RGB images in total. Each subject had 15 view points and 19 illumination conditions. In addition, high resolution frontal images were acquired as well. The database was collected in four sessions over a month. Each session included about 300 different individuals, which means duplicate individuals appeared among these sessions.

In the tests, two types of testing images, namely familiar images and unfamiliar images, were collected from each database for validation. The individuals of familiar images also appear in the training images, while the individuals of unfamiliar images are different from those of the training images. For each database, 100 face images formed the training set and the other 20 face images 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 to form the familiar testing set. In other words, 2 different images were selected for

(a) (b) (c)

Figure 4.1: Feature points and TPS alignment result. (a) Reference image with 83 feature points, (b) face image with 83 feature points and (c) aligned face by TPS.

Table 4.1: Various occluder characteristics for generating test images.

Characteristics	Number	Range
shape	3	block, vertical/horizontal bar
size	5	0.05, 0.1, 0.2, 0.3, 0.4
intensity	9	0 to 255 in intervals of 32
location	6	left/right eye, nose, left/right cheek, mouth

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.

All face images were resized to 120 $\times$ 100 pixels. Then, 83 feature points were extracted for every face image by employing the Face++ toolkit [Inc13]. Some examples are shown in the Figure 4.1. All face images were aligned to the mean face by thin plate spline (TPS) [Boo89] method with the extracted feature points.

Rectangular occluders were placed on each testing image to synthesize the occluded images, while the original unoccluded versions served as ground truth. Various occluder characteristics were applied, which are shown in Table 4.2. Note that the 5 size ratios were measured as the ratio of occluder's area over the image area. Examples of synthesized testing images are shown in Figure 4.2. In practice, a simple way for a user to manually mark real occluders, such as eyeglasses, scarves, beards, and face masks, in a target image is to place regularly-shaped blocks over the occlusions. So, the above test scenarios are relevant to the removal of real occluders.

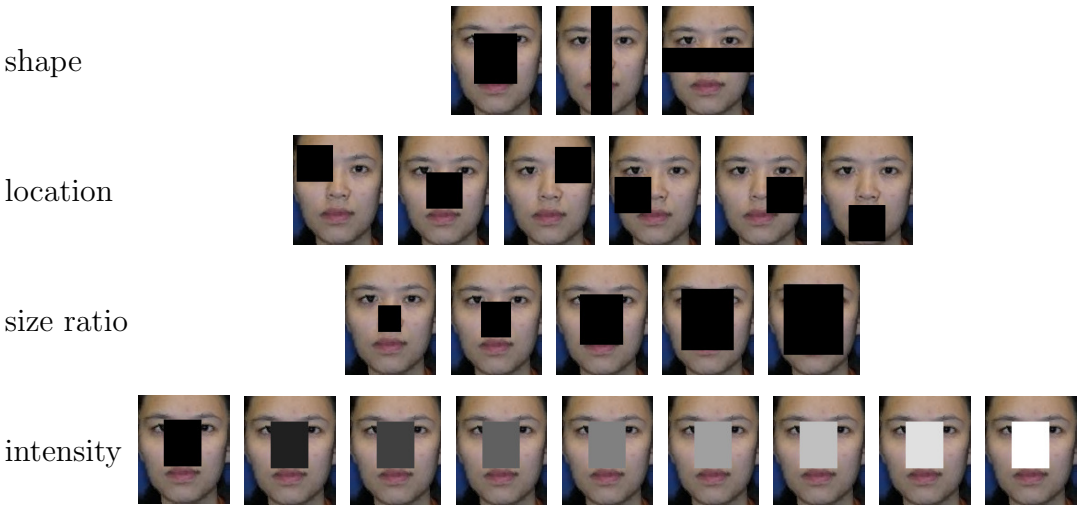


Figure 4.2: Characteristics of synthetic occluders.

4.2 Parameter Setting

VRPCA has three weight parameters for soft thresholds. The weight w_t of training images was set to 1 because training images are unoccluded. The weights w_o and w_u for the occluded and unoccluded parts of a target image, 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. Table 4.2 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. Therefore, for subsequent tests, w_o and w_u were set to 0.0001 and 10 respectively.

4.3 Test Procedure

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

Table 4.2: Effect of weights of soft thresholds on VRPCA's performance (RMSE).

(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

(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**

- VRPCA: the proposed variable-threshold RPCA.
- VRPCA2: variation of VRPCA that is equivalent to RPCA for matrix completion of [LCM09].
- RPCA: RPCA for matrix recovery, similar to [MD12] except [MD12] applies hard instead of soft thresholding.
- FW-PCA: fast weighted PCA of [HNK+12].
- BC: block PCA with block correlation of [YS08].
- RSC: robust sparse coding of [YZYZ11].

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 Y. The parameter λ was set to the theoretical optimum of $1/\sqrt{\max(m,n)}$ [ZLW+10]. For the other methods, their parameter values were set according to recommendation. [LCM09]

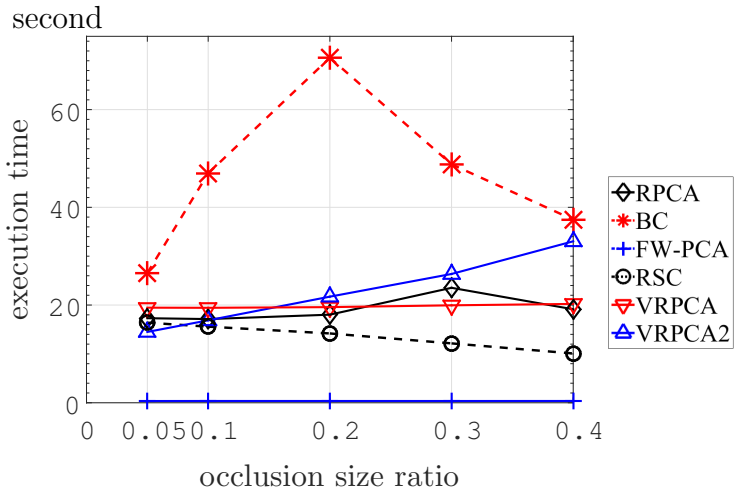


Figure 4.3: Comparison of execution time of various algorithms across different occlusion size ratios.

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. Root mean-squared errors (RMSE) between the recovered images and their ground truth images were computed and averaged over all testing images. RMSE for the occluded parts and RMSE unoccluded parts were computed separately. Intuitively, RMSE can be regarded as the average error of every pixel. Furthermore, for color images, the proposed method was performed on each of the RGB channels and the RMSE results were averaging over the three channels' results.

4.4 Execution Time

In this experiment, all test programs were implemented in Matlab and ran in an Ubuntu 14.04 x64 PC with Intel Core i7-4770 CPU and 7.7GB RAM. The methods were tested on the test images with varying size ratios as shown in Figure 4.2.

Figure 4.3 shows that fast-weighted PCA is the fastest because PCA requires only one SVD step. Moreover, its running time is independent of the occlusion size ratio. RPCA, VRPCA and VRPCA2 have similar longer running time, because they run multiple iterations, each requiring a SVD step. Among them, VRPCA2's

execution time increases with occlusion size ratio because it needs to fill in more matrix elements. In contrast, the running time of VRPCA and RPCA are independent of occlusion size because they work on the entire matrix. Surprisingly, the running time of robust sparse coding decreases with increasing occlusion size. Block PCA has the longest running time because every occluded block should be matched with every unoccluded block to determine the most correlated block for filling the occluded block. At occlusion size ratio of 0.2, it preforms the most amount of block comparisons, which results in the longest running time.

4.5 Effect of Occluder Intensity

This test evaluates the effect of occluder intensity on the algorithm's performance. As shown in Figure 4.4 and Figure 4.5, familiar test sets have less error compared with unfamiliar test sets, and this result is consistent with different datasets. For the occluded part of different algorithms, the errors of VRPCA, VRPCA2, FW-PCA, and BC 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 in 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. Sample images of test results are shown in Figure 4.6 to Figure 4.9.

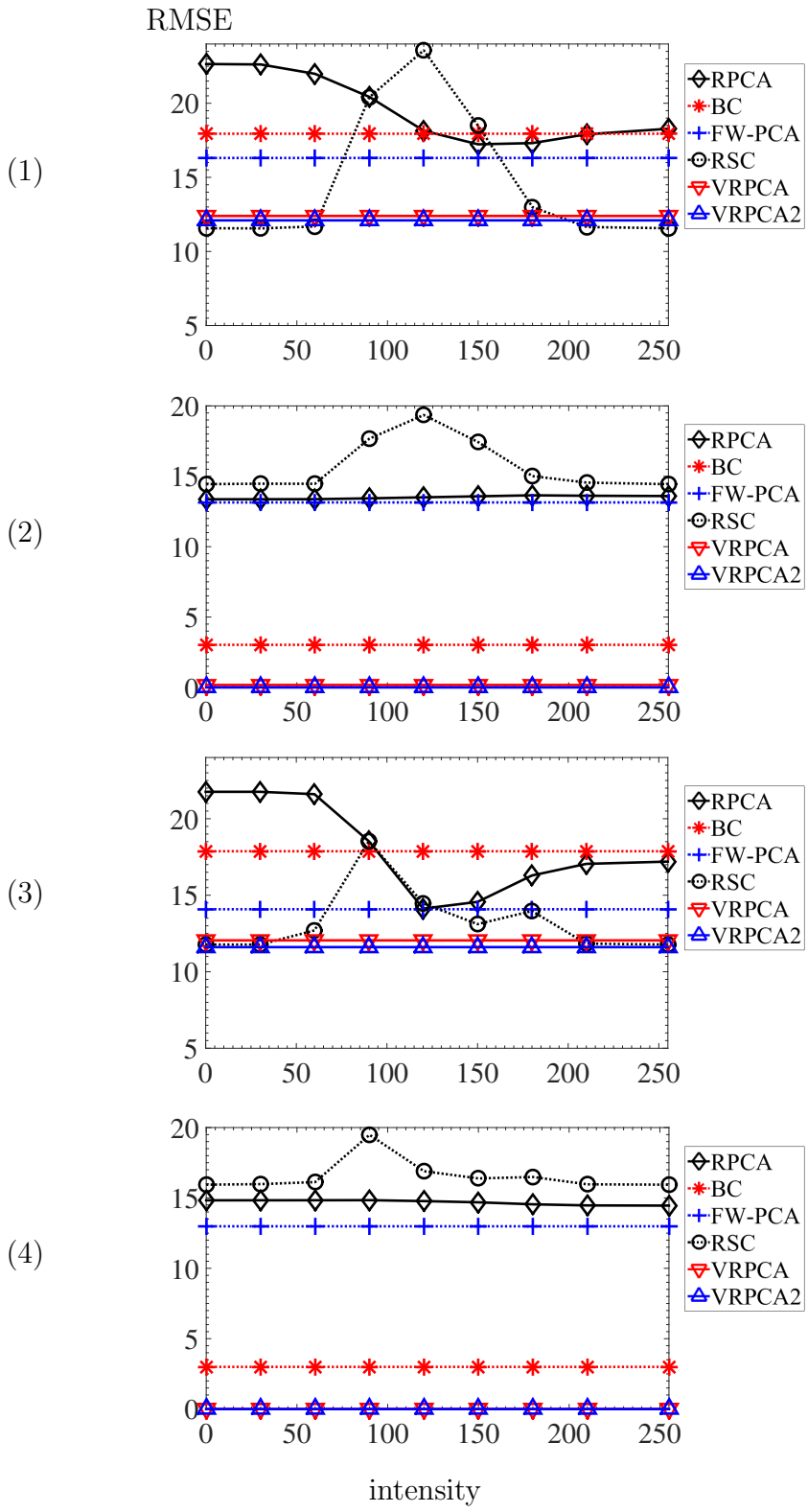


Figure 4.4: Effect of occluder intensity on Multi-PIE test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts.

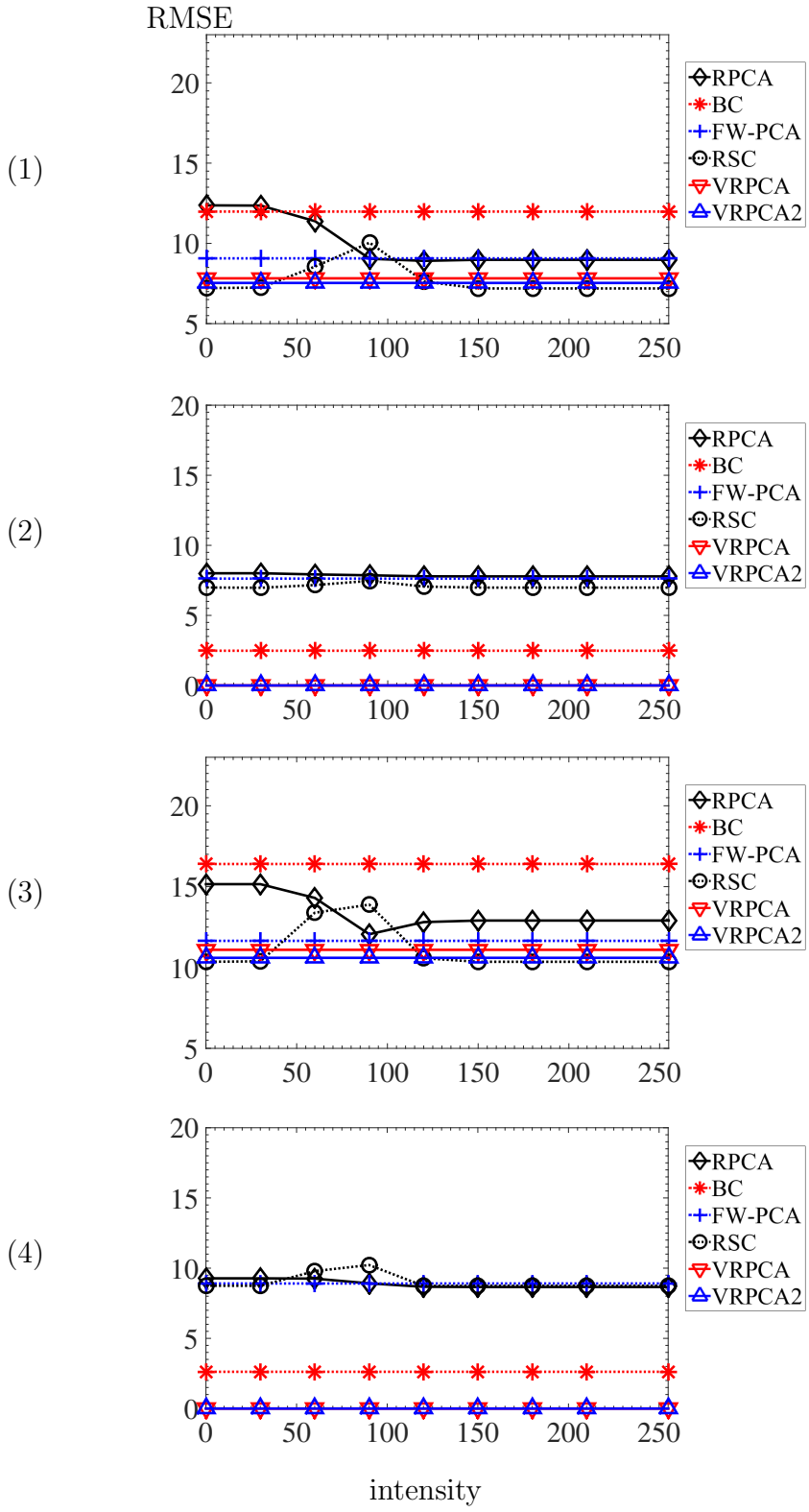


Figure 4.5: Effect of occluder intensity on FERET test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts.

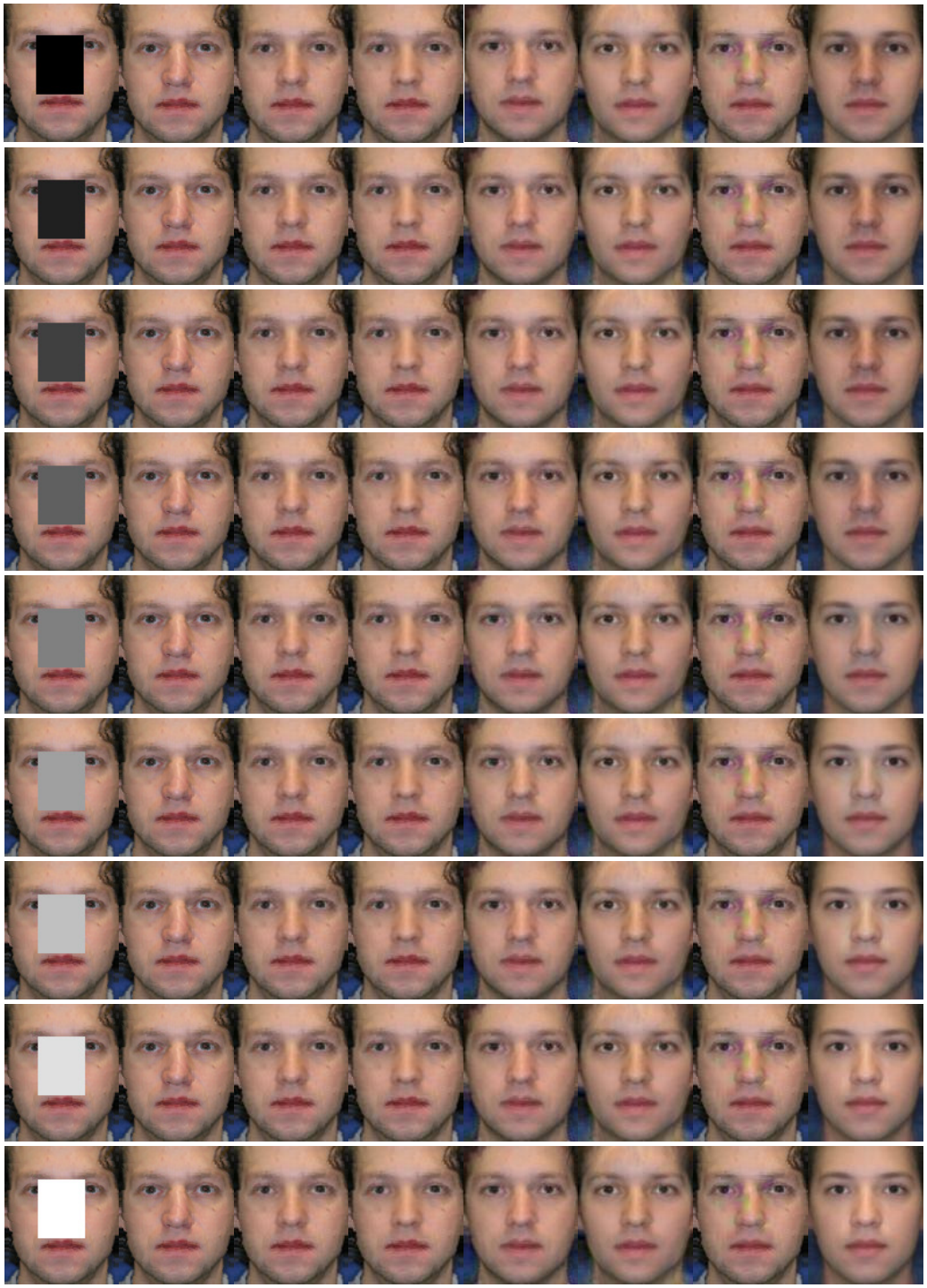


Figure 4.6: Sample results of occluder intensity test for Multi-PIE familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

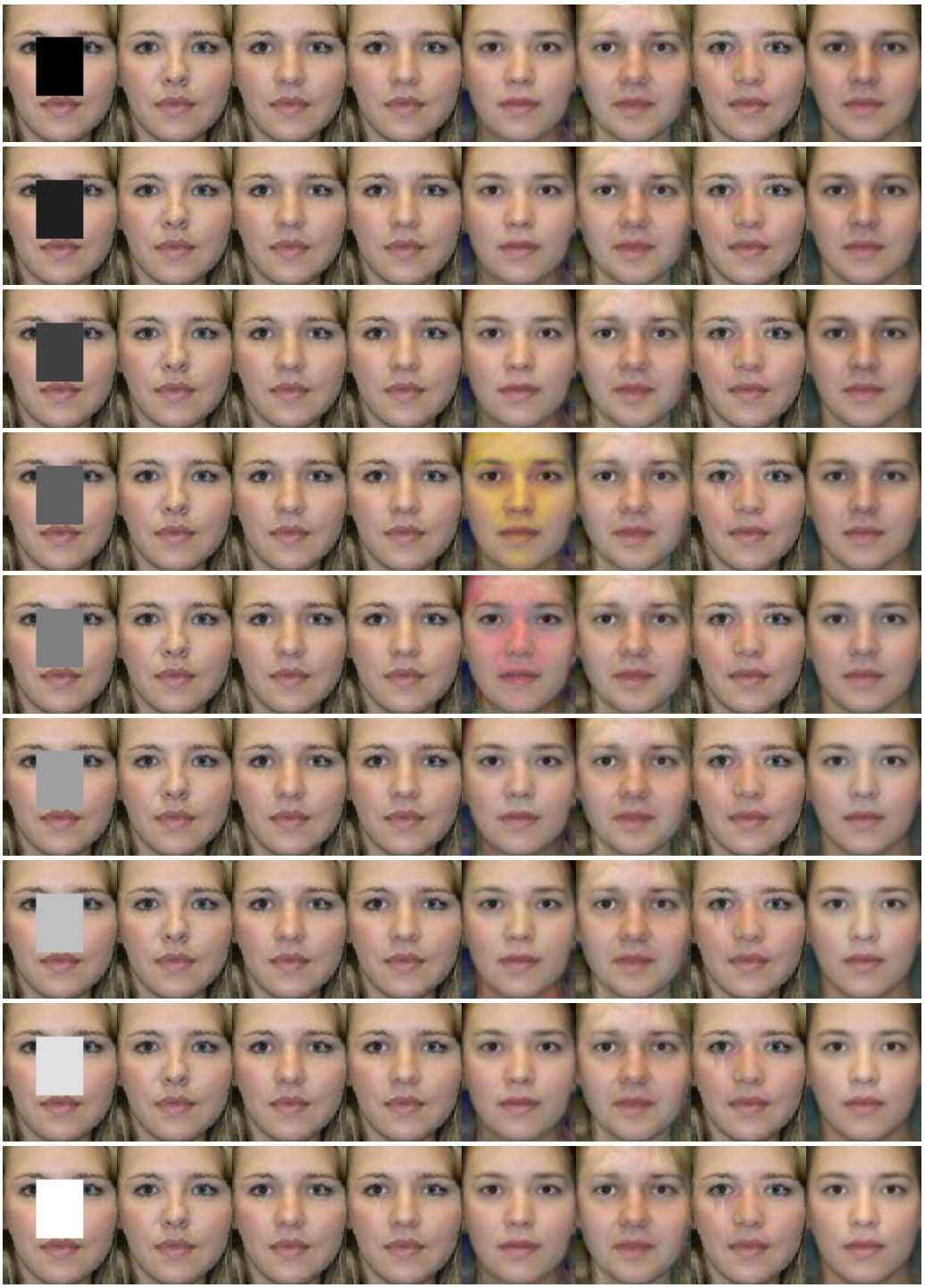


Figure 4.7: Sample results of occluder intensity test for Multi-PIE unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

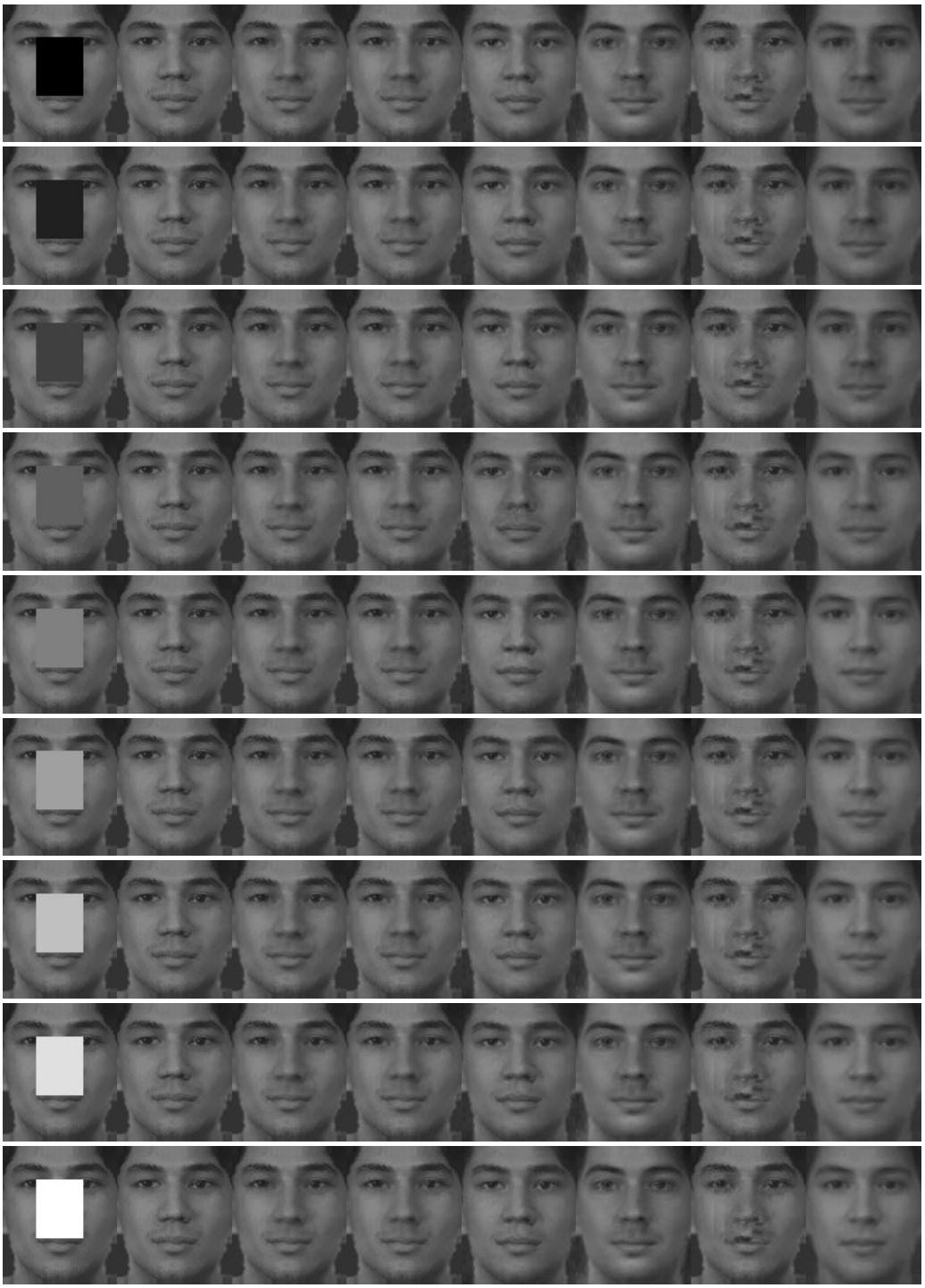


Figure 4.8: Sample results of occluder intensity test for FERET familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

Figure 4.9: Sample results of occluder intensity test for FERET unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

4.6 Effect of Occluder Size

This test evaluates the effect of occluder size ratio on the algorithm's performance. As shown in Figure 4.10 and Figure 4.11, familiar test sets have less error compared with unfamiliar test sets, and this result is consistent with different datasets. 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. For the unoccluded parts, VRPCA and VRPCA2 attain practically zero error while BC has very small error. On the other hand, RSC, FW-PCA, and RPCA have large errors in the unoccluded parts. The algorithms' errors on the unoccluded parts under varying occluder size are consistent with those under the varying occluder intensity. The images recovered by VRPCA and VRPCA2 are the most similar to the ground truth. Whereas those recovered by RSC and FW-PCA looks more like smoothed versions of the ground truth. Moreover, when occluder intensity similar to facial intensity, RSC's result has large error. 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. Sample images of test results are shown in Figure 4.12 to Figure 4.15.

4.7 Effect of Occluder Location

This test evaluates the effect of occluder location on the algorithm's performance. As shown in Figure 4.16 and Figure 4.17, familiar test sets have less error compared with unfamiliar test sets. With regard to the effect of occluder location, for the occluded parts, all methods have larger errors on the left and right eyes than in the other areas. 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 familiar or unfamiliar images and different databases images. Sample images of testing results are shown in Figure 4.18 to Figure 4.21.

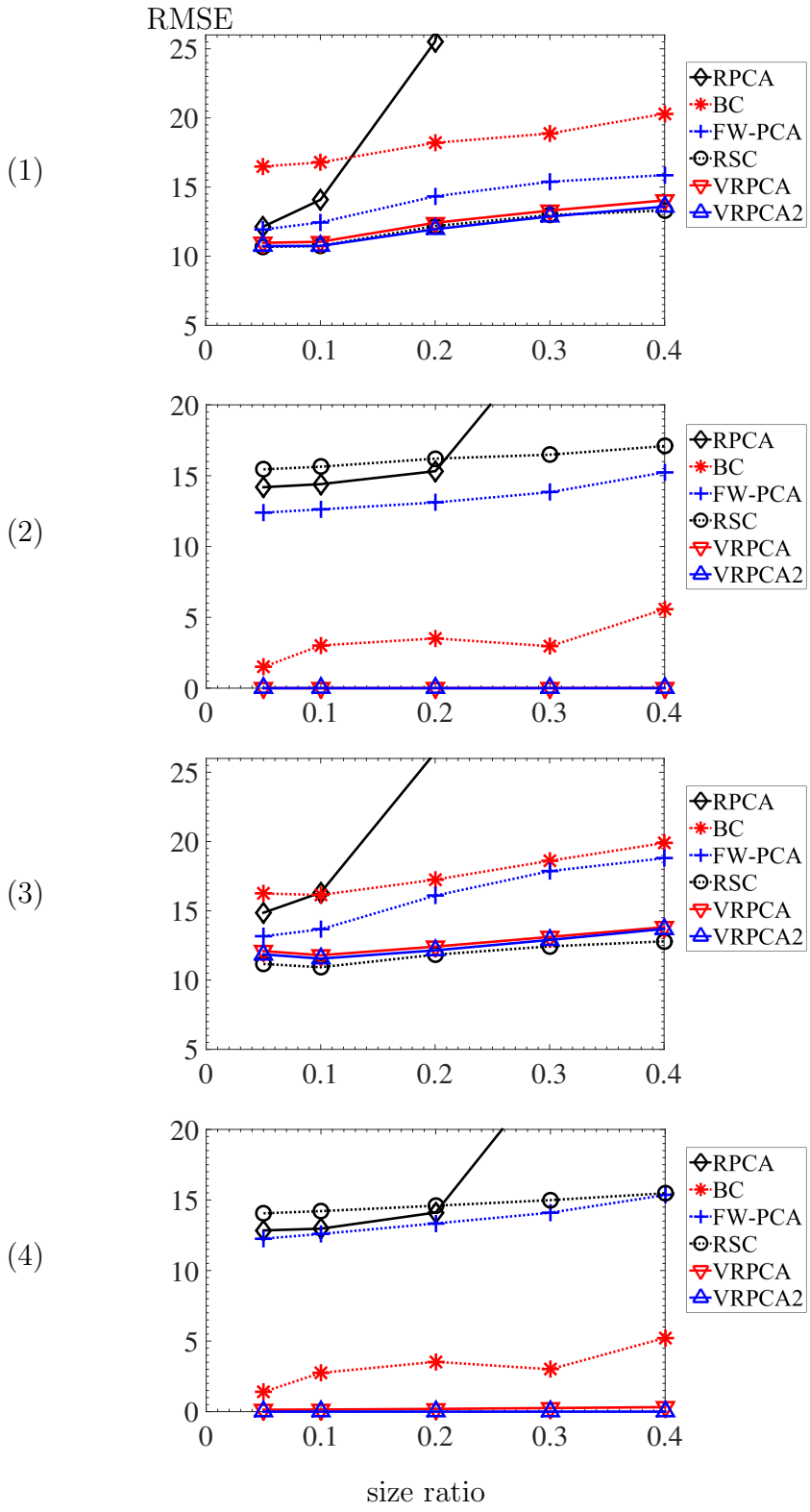


Figure 4.10: Effect of occluder size ratio on Multi-PIE test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts.

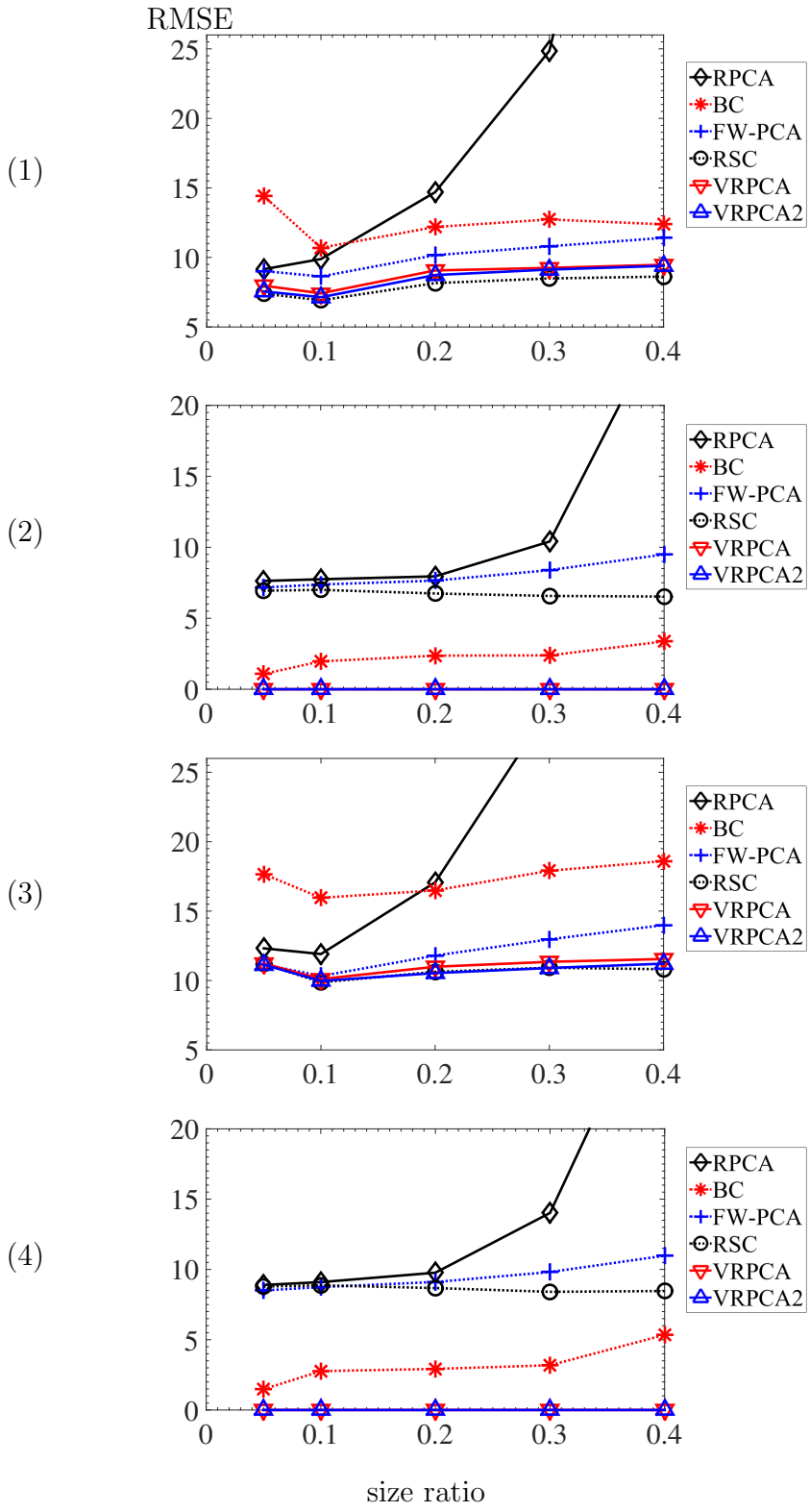


Figure 4.11: Effect of occluder size ratio on FERET test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts.

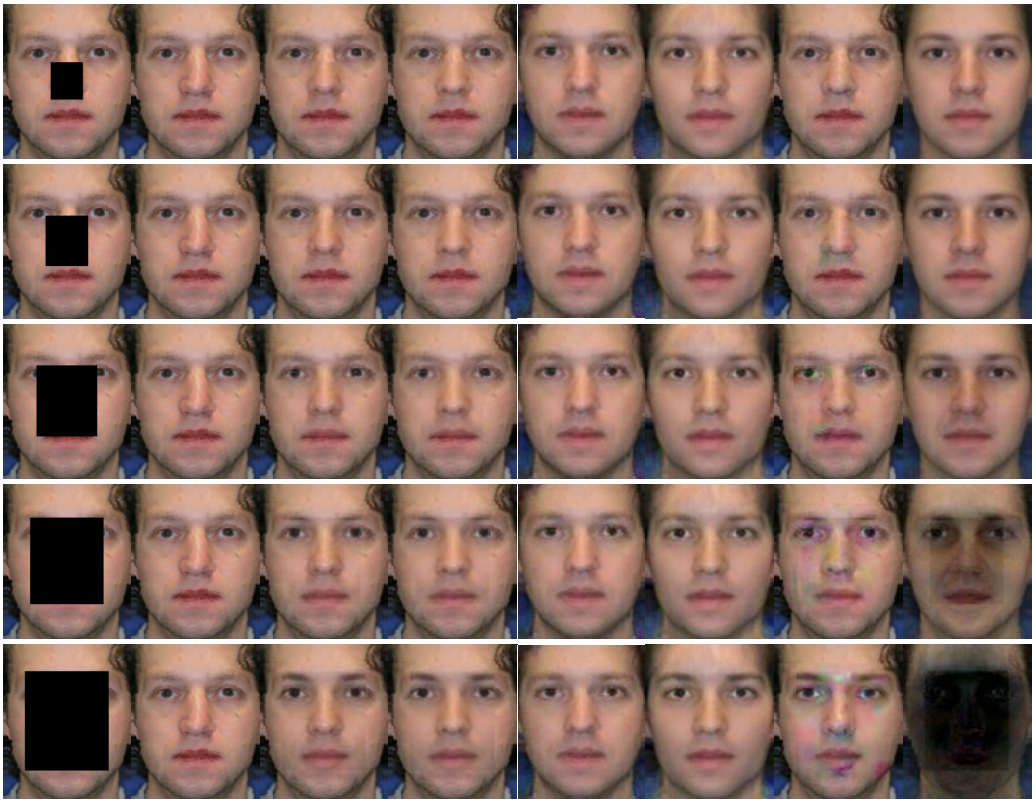


Figure 4.12: Sample results of occluder size ratio for Multi-PIE familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

4.8 Effect of Occluder Shape

This test evaluates the effect of occluder shape on the algorithm's performance. As shown in Figure 4.22 and Figure 4.23, familiar test sets have less error compared with unfamiliar test sets. Among the three occlusion shapes, the vertical occlusion has the least RMSE for almost all methods, while the horizontal occlusion has the largest RMSE. For the unoccluded part, VRPCA still preserves the original information. This result is consistent with all the familiar or unfamiliar and different databases test images. Sample images of testing results are shown in Figure 4.24, to Figure 4.27.

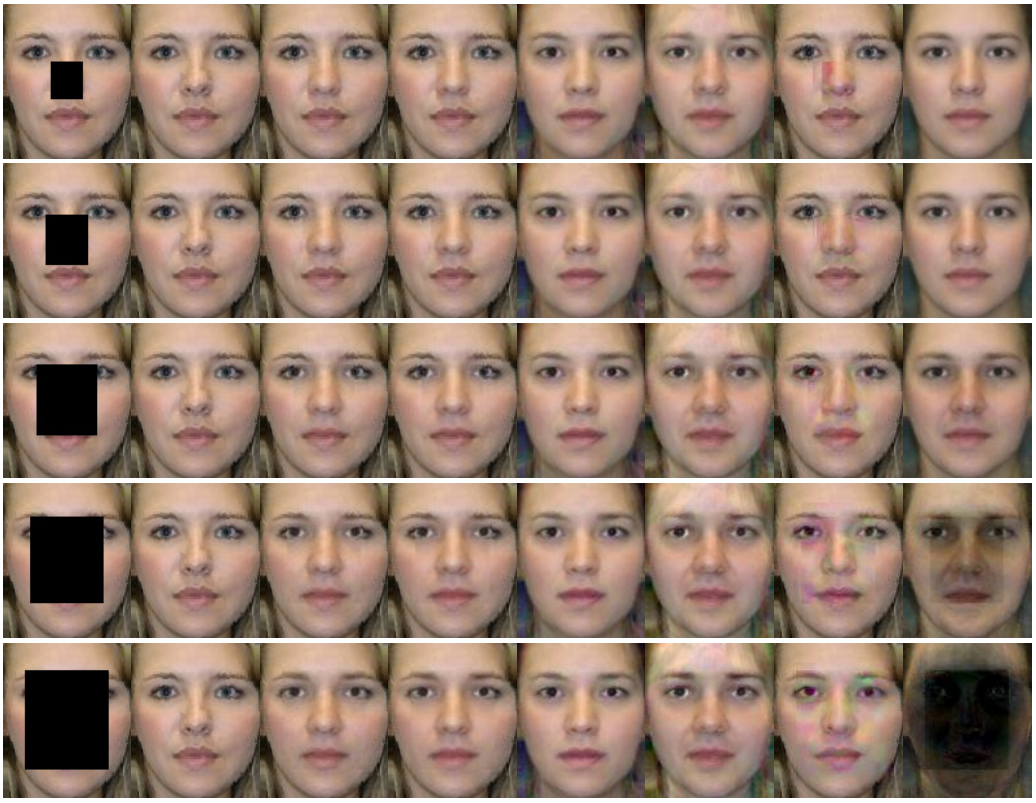


Figure 4.13: Sample results of occluder size ratio for Multi-PIE unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

4.9 Summary

In summary, VRPCA and VRPCA2 have the most accurate and consistent performance across various test scenarios 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 the occluding intensity because it has difficulty distinguishing mid-tone occluders and facial features. When the occluder intensity is very different

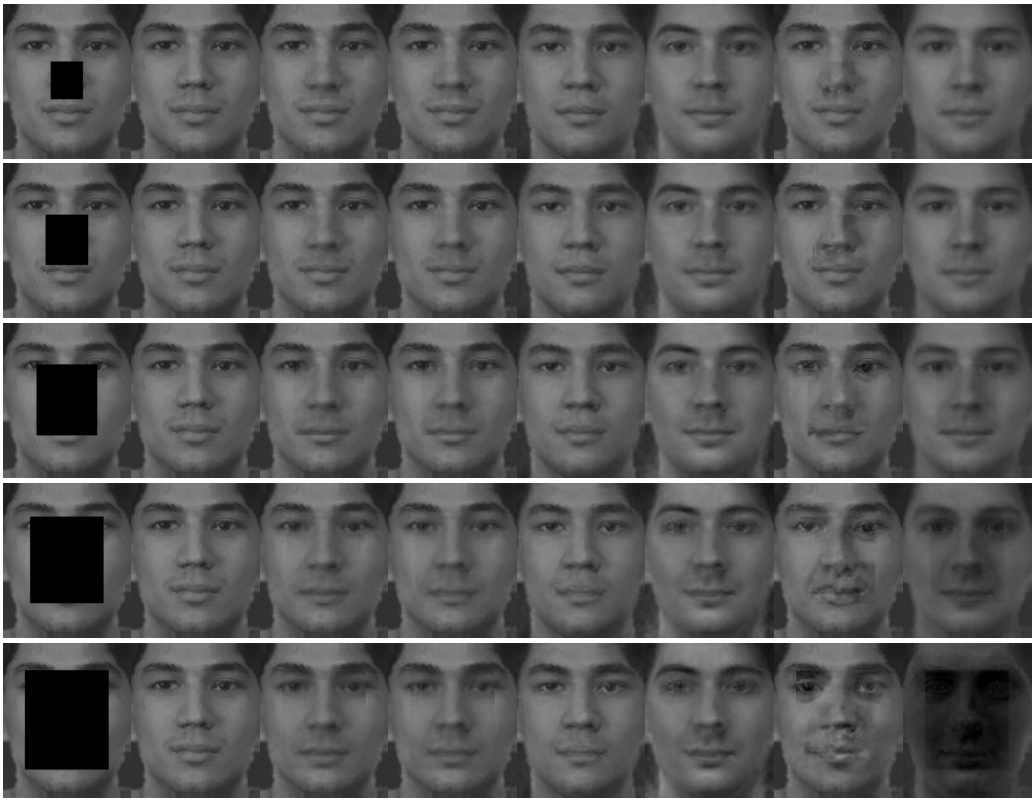


Figure 4.14: Sample results of occluder size ratio for FERET familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

from the facial intensity, RSC can attain good results. However, its error on 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 VRPCA whose soft thresholds are 1 for both the occluded and unoccluded parts, which are inappropriate soft thresholds.

Figure 4.15: Sample results of occluder size ratio for FERET unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

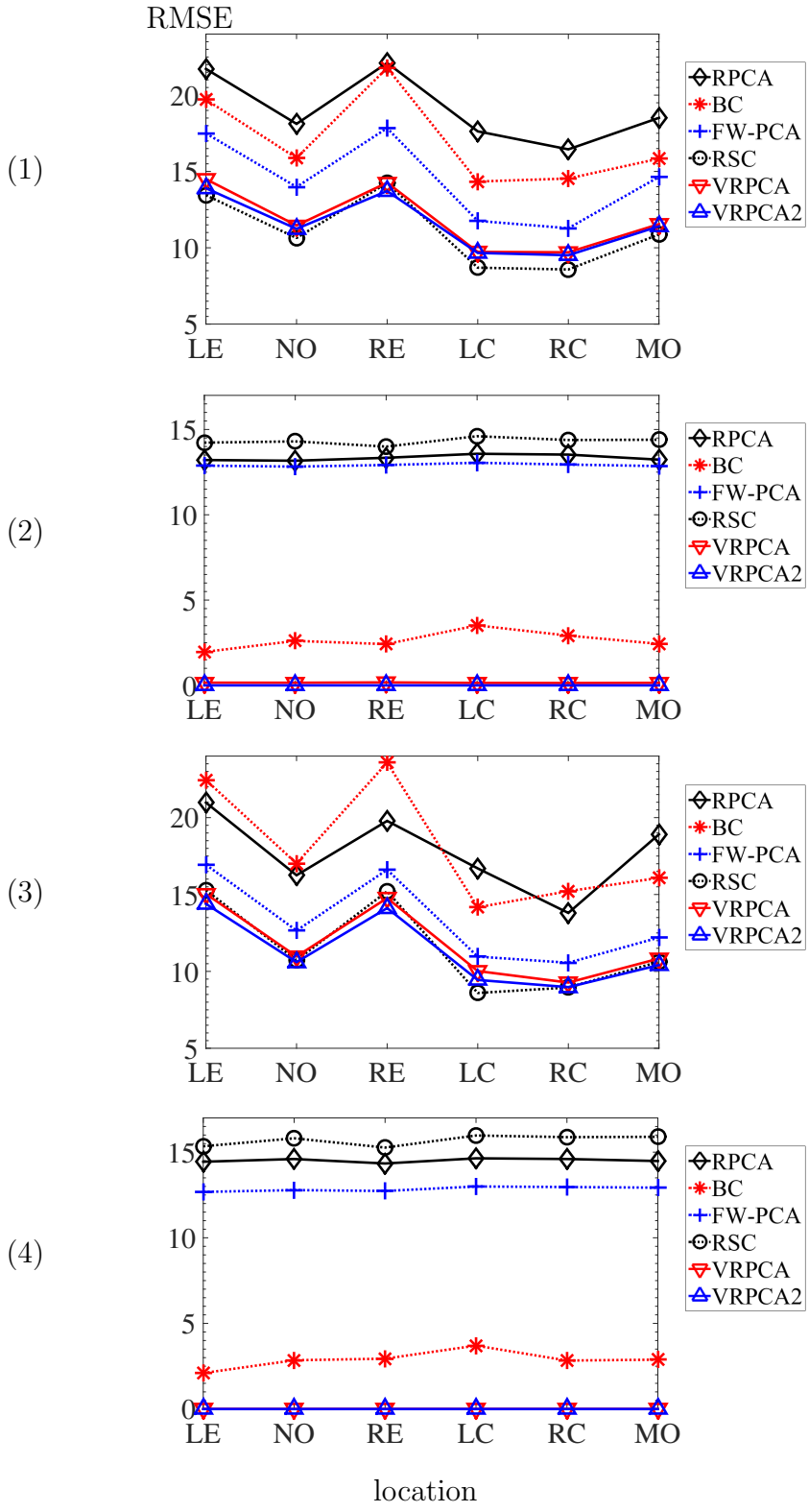


Figure 4.16: Effect of occluder location on Multi-PIE test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts. Locations include (LE) left eye, (RE) right eye, (NO) nose, (LC) left cheek, (RC) right cheek, and (MO) mouth. Lines joining data points are meant as visual aid only.

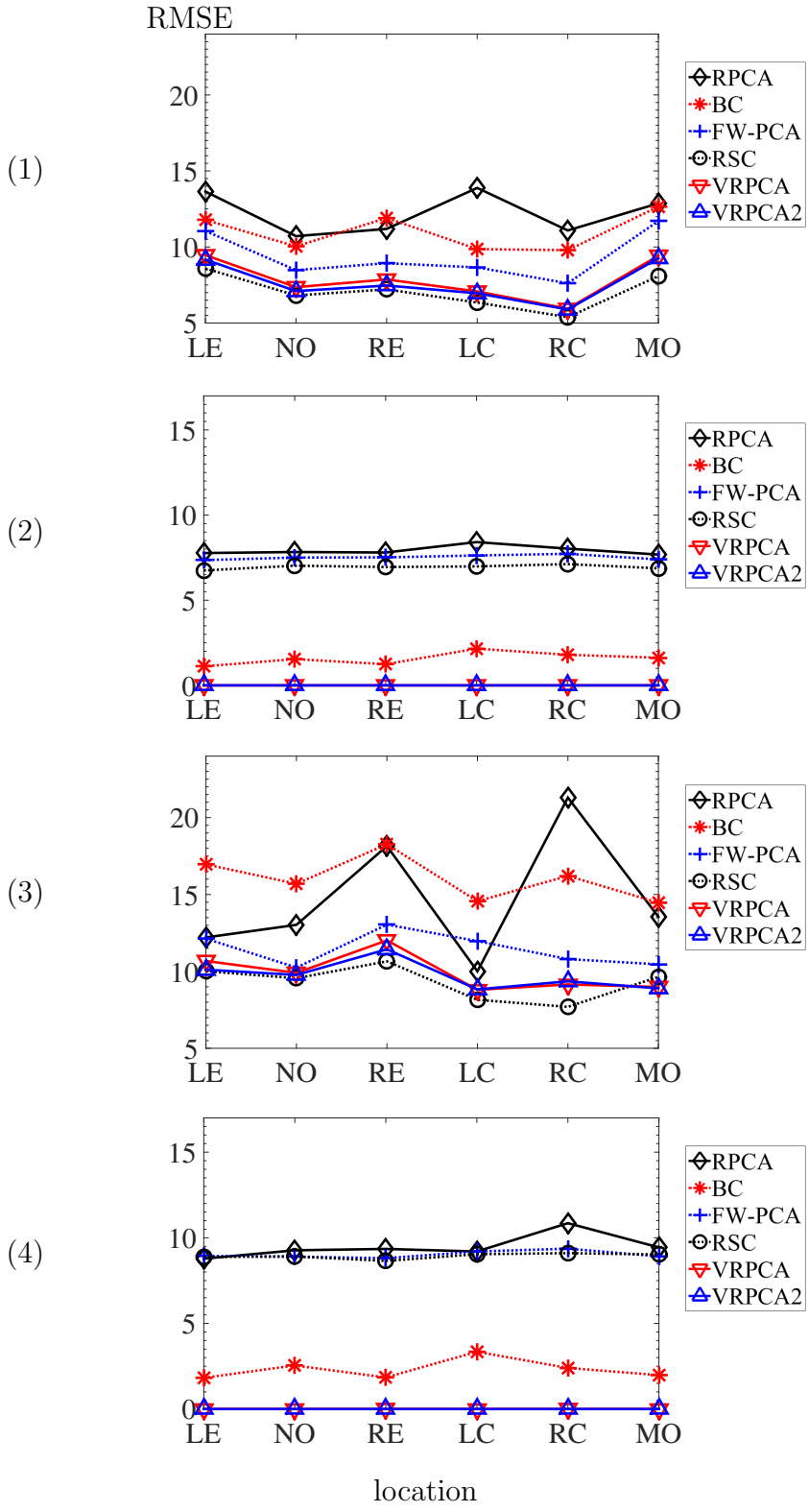


Figure 4.17: Effect of occluder location on FERET test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts. Locations include (LE) left eye, (RE) right eye, (NO) nose, (LC) left cheek, (RC) right cheek, and (MO) mouth. Lines joining data points are meant as visual aid only.

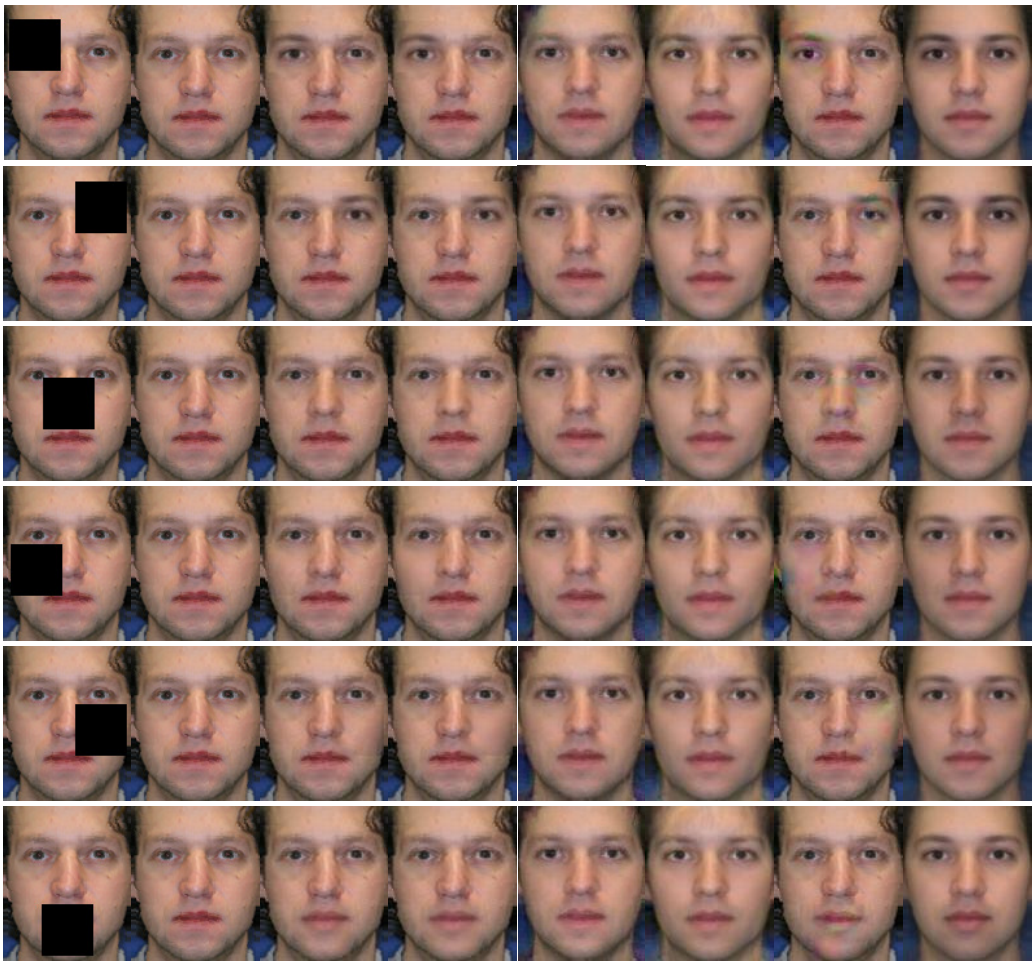


Figure 4.18: Sample results of occluder location test for Multi-PIE familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

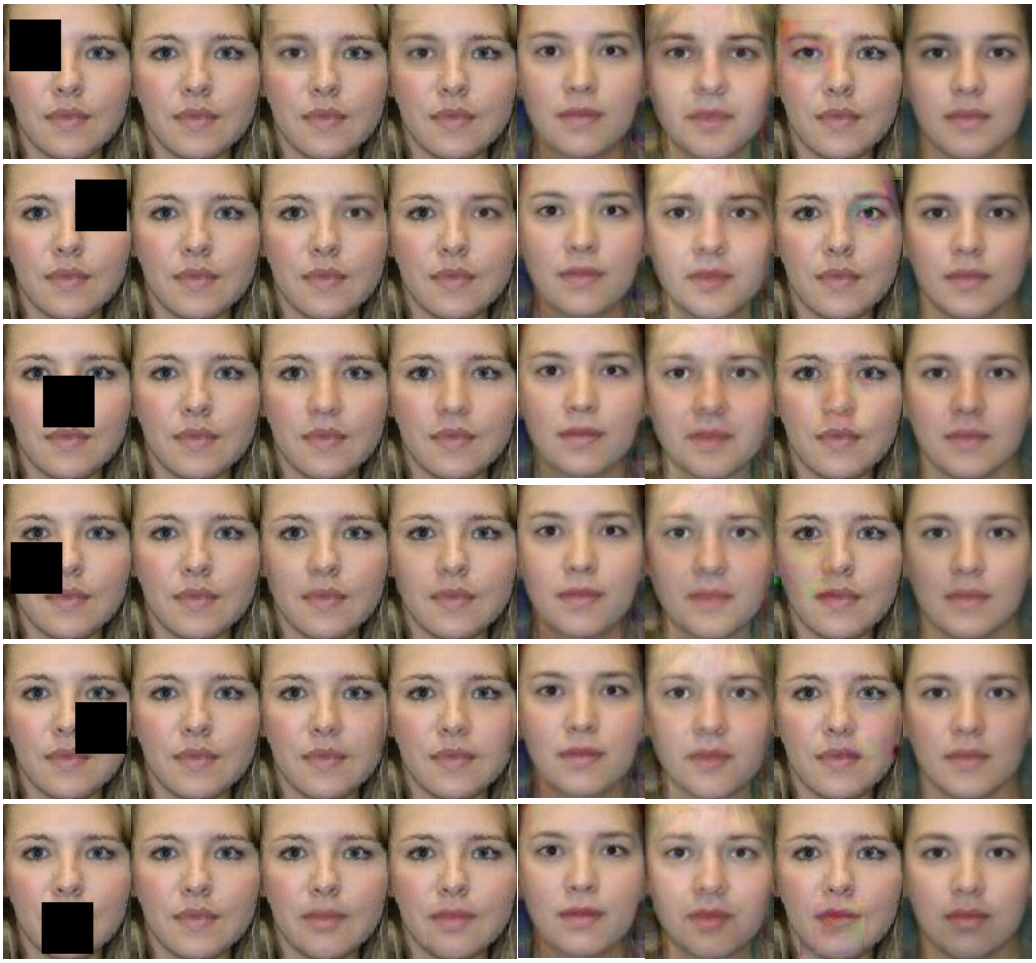


Figure 4.19: Sample results of occluder location test for Multi-PIE unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

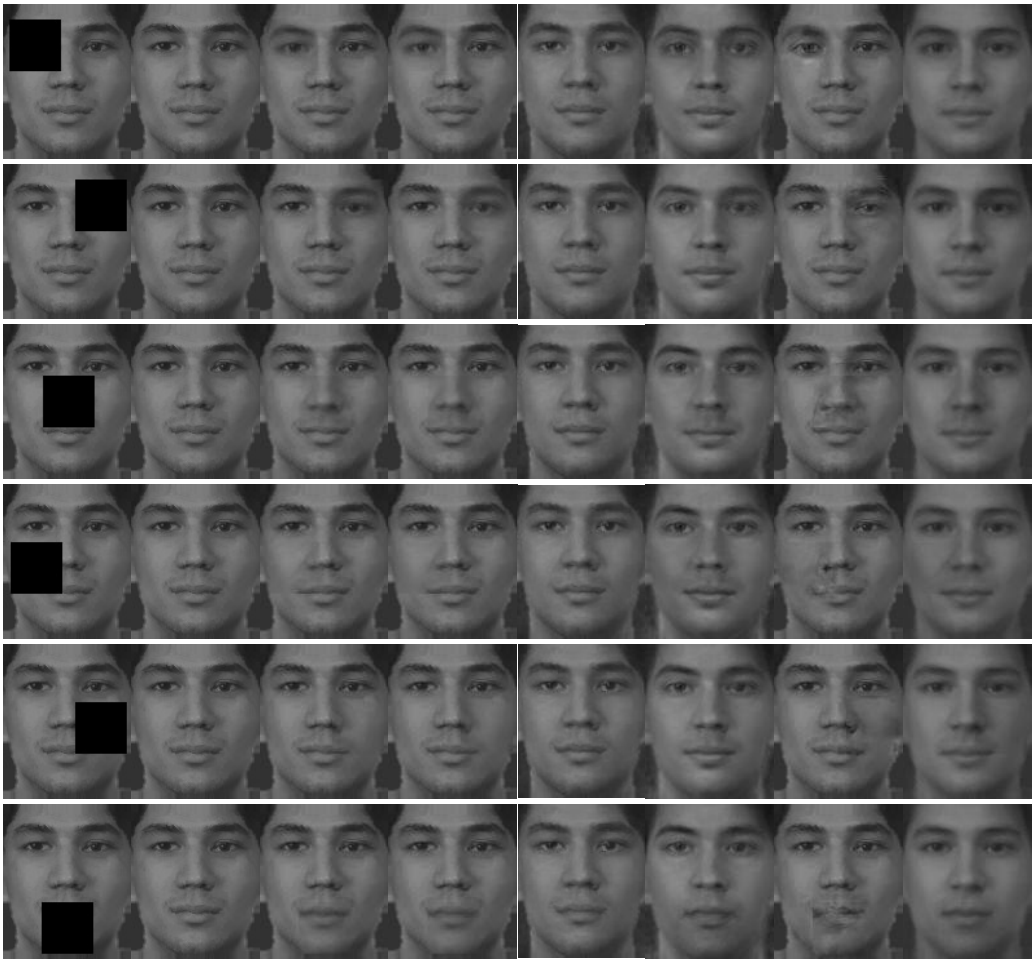


Figure 4.20: Sample results of occluder location test for FERET familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

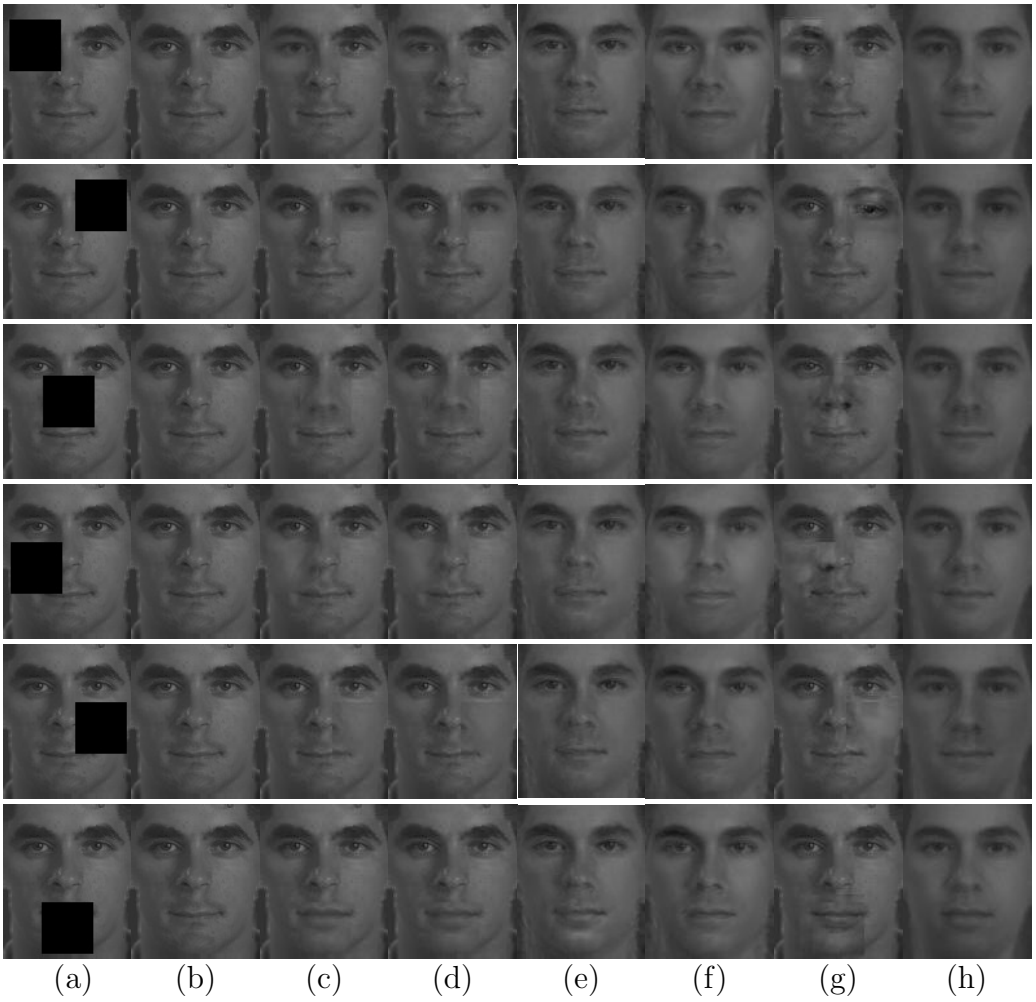


Figure 4.21: Sample results of occluder location test for FERET unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

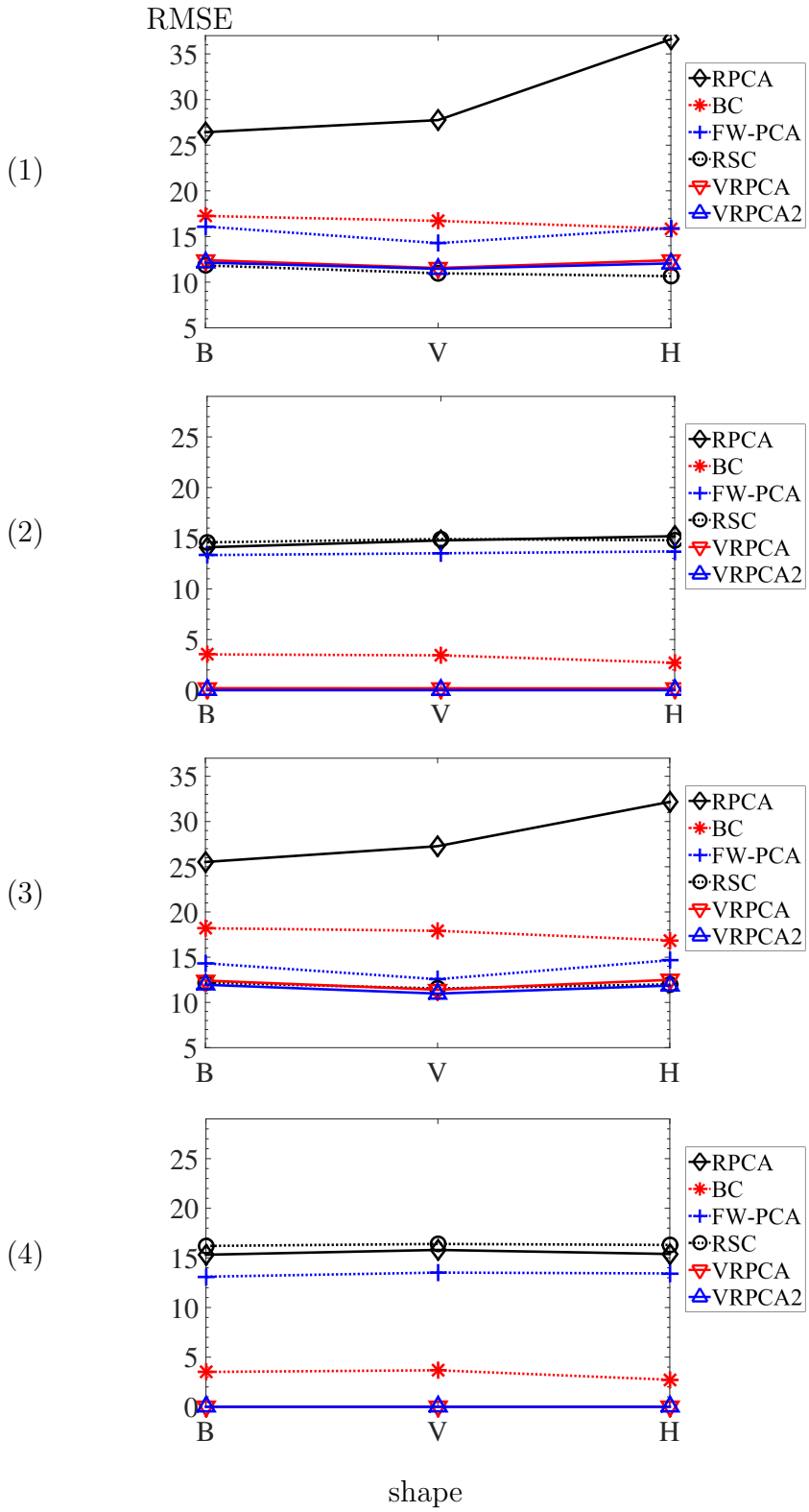


Figure 4.22: Effect of occluder shape on Multi-PIE test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts. Shapes include (B) block, (V) vertical bar, and (H) horizontal bar. Lines joining data points are meant as visual aid only.

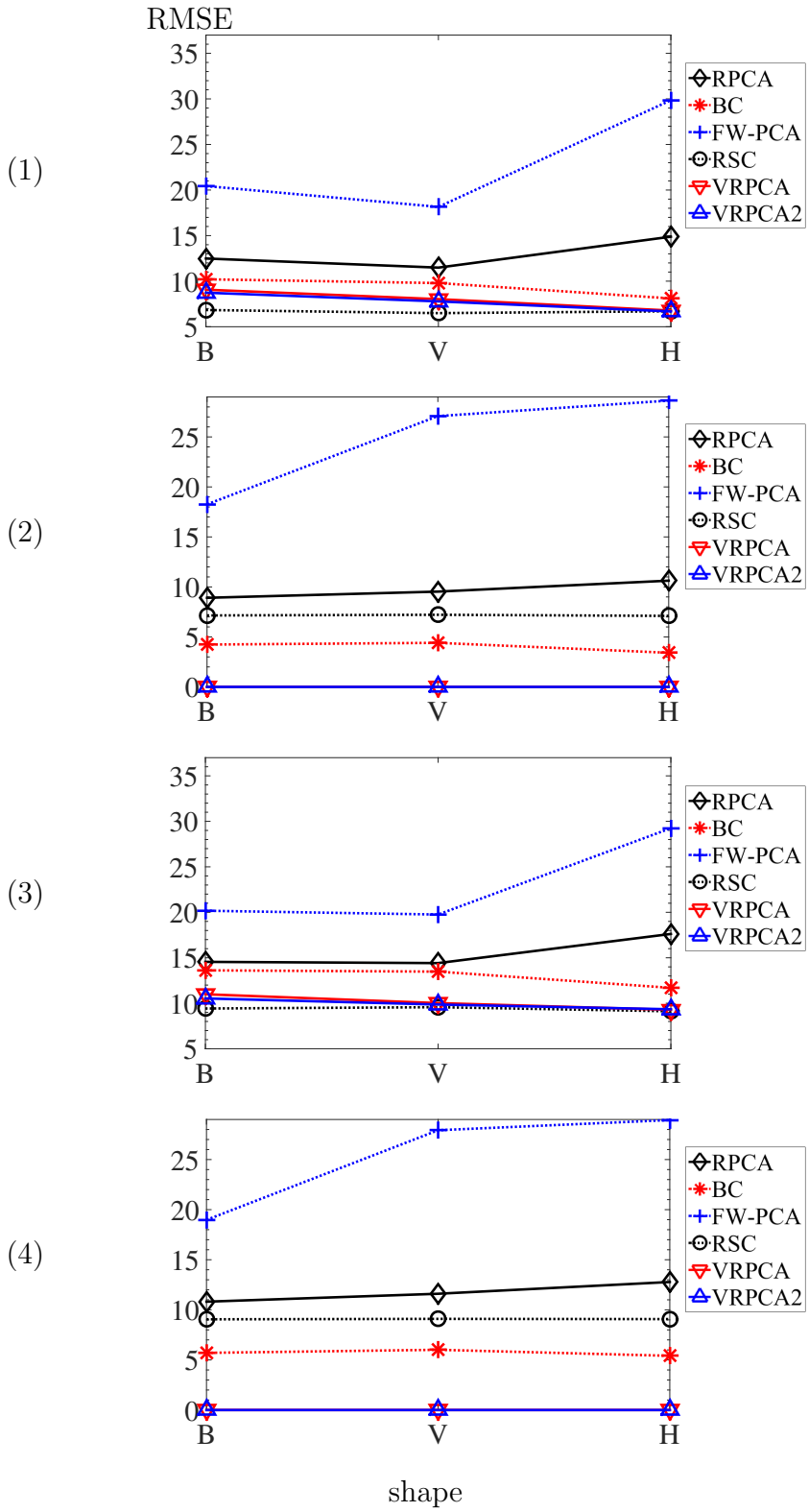
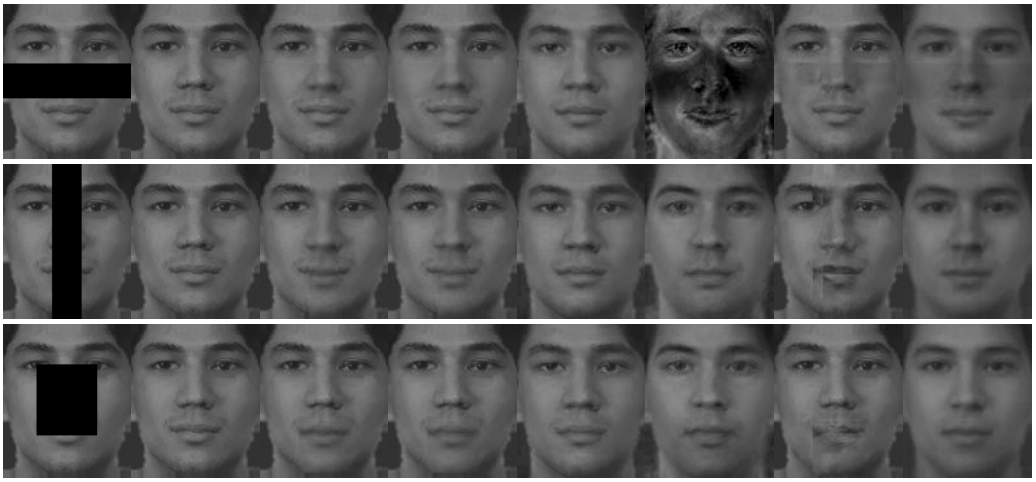


Figure 4.23: Effect of occluder shape on FERET test images. (1, 2) Familiar test sets, (3, 4) unfamiliar test sets. (1, 3) Occluded parts, (2, 4) unoccluded parts. Shapes include (B) block, (V) vertical bar, and (H) horizontal bar. Lines joining data points are meant as visual aid only.

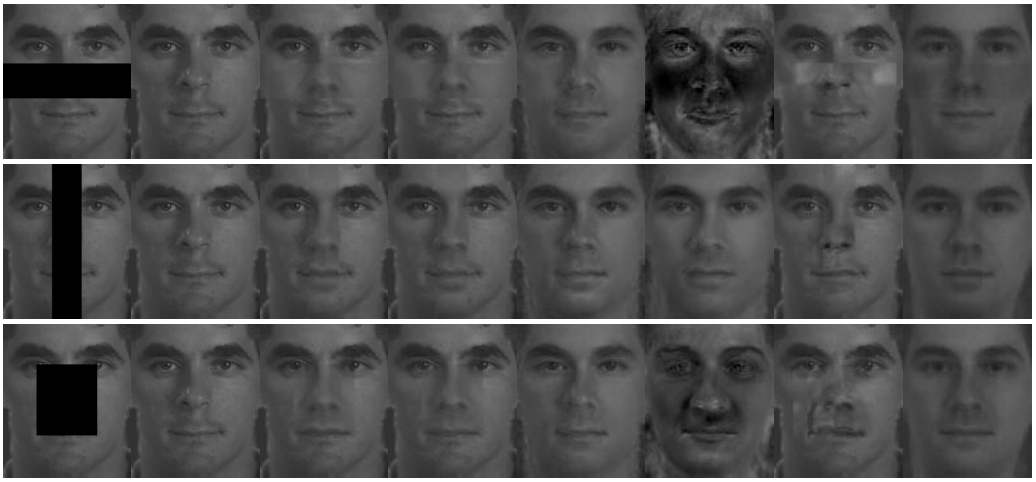
Figure 4.24: Sample results of occluder shape test for Multi-PIE familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

Figure 4.25: Sample results of occluder shape test for Multi-PIE unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.



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

Figure 4.26: Sample results of occluder shape test for FERET familiar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.



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

Figure 4.27: Sample results of occluder shape test for FERET unfamiliar set. (a) Occluded image, (b) ground truth, (c) VRPCA, (d) VRPCA2, (e) RSC, (f) FW-PCA, (g) BC, and (h) RPCA.

Chapter 5

Future Work

There are some problems to be solved in the continuing research related to face image de-occlusion. The following Section 5.1 and Section 5.2 highlight possible future work about robust face alignment and automatic occlusion detection respectively.

5.1 Robust Face Alignment

Face alignment aims at locating facial feature points automatically. In this work, the facial feature points were extracted by Face++ toolkit and the images were aligned by TPS. The testing face images were also aligned and occluders were placed on them to generate synthetic occluded images. However, in real applications, detection of face features in images with large occlusion may fail, resulting in misalignment of face images.

One possible way to solve the misalignment problem is to manually mark the feature points. However, this approach is tedious and error-prone. With the development of robust face alignment, many recent studies can handling the detection and alignment of facial images with various occlusions, expressions, poses and lightings.

Recent face image landmark detection studies employ regression-based technique. [WJ15] proposes a cascade framework which aims at handle landmark detec-

tion under face occlusions or head poses conditions. Two regressors are employed for estimating the probability of landmark visibility and the location of landmarks respectively. [YLBM14] develops a consensus of regression method by constrained local model (CLM) [CC08, SLC11]. Assuming that the local regions are occluded, a set of regressors are learnt to estimate the location of face landmarks, while the consensus is achieved by combining the set of regressors and maximizing the overall alignment likelihood.

5.2 Automatic Occlusion Detection

The proposed VRPCA requires the user to mark the occluded parts of a black rectangular block. Many occluded face recognition methods include the automatic occlusion detection. Some methods detect the occlusion part and then recover the face images by iterative updating. Nevertheless, there are some works that focus on occlusion detection [ZSCG07, OLL08, SUB10, MHD14]. [ZSCG07] employs local Gabor binary patterns (LGBP). [OLL08] proposes to conduct binary occlusion detection in local patches based on PCA face subspace. [SUB10] presents an approach based on color techniques using the H-channel of the hue, saturation and value (HSV) color space, then a combination of active shape model (ASM) and a component-based PCA subspace reconstruction is used to improve the method. [MHD14] approaches the occlusion detection problem using Gabor wavelets, PCA and support vector machines (SVM).

Chapter 6

Conclusion

This thesis has presented the proposed method variable-threshold robust PCA (VRPCA) for face image occlusion removal. This method 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. Given the knowledge of occlusion area, VRPCA separates the error part to three different parts, the training part, the testing occluded part and the testing unoccluded part, then variable-threshold is applied to the error parts instead of the same soft-threshold. As a result, the variable-threshold method preserves the original unoccluded part and constrains that the separation errors are mostly from the testing occluded part. Comprehensive tests show that both variations are consistently more accurate than existing methods across various test conditions for two face image databases. 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. In contrast to original robust PCA (RPCA), the variable soft thresholds also provide additional constraints that allow VRPCA to perform well even when the data matrix is not exactly low-rank.

References

[ANS12] Mohammad Al-Naser and Ulrik Söderström. Reconstruction of occluded facial images using asymmetrical principal component analysis. *Integrated Computer-Aided Engineering*, 19(3):273–283, 2012.

[Boo89] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 11(6):567–585, 1989.

[BSCB00] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *Proceedings of Annual Conference on Computer Graphics and Interactive Techniques*, pages 417–424, 2000.

[BZ14] Thierry Bouwmans and El Hadi 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.

[CC08] David Cristinacce and Tim Cootes. Automatic feature localisation with constrained local models. *Pattern Recognition*, 41(10):3054–3067, 2008.

[CCS10] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

[CLMW11] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM*, 58(3):11:1–11:37, 2011.

[CR09] Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.

[CTCB15] Ricardo Cabral, F. D. la Torre, João P Costeira, and Alexandre Bernardino. Matrix completion for weakly-supervised multi-label image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(1):121–135, 2015.

[DFKE07] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collabora-

tive filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.

[DlTB01] Fernando De la Torre and Michael J Black. Robust principal component analysis for computer vision. In *Proceedings of International Conference on Computer Vision*, pages 362–369, 2001.

[DZHZ06] Chris Ding, Ding Zhou, Xiaofeng He, and Hongyuan Zha. R1-PCA: rotational invariant l_1-norm principal component analysis for robust subspace factorization. In *Proceedings of International Conference on Machine Learning*, pages 281–288, 2006.

[GMC+10] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, and Simon Baker. Multi-pie. *Image and Vision Computing*, 28(5):807–813, 2010.

[HL03] Bon-Woo Hwang and Seong-Whan Lee. Reconstruction of partially damaged face images based on a morphable face model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):365–372, 2003.

[HNK+12] Tomoki Hosoi, Sei Nagashima, Koji Kobayashi, Koichi Ito, and Takafumi Aoki. Restoring occluded regions using FW-PCA for face recognition. In *Proceedings of Workshops at IEEE Conference on Computer Vision and Pattern Recognition*, pages 23–30, 2012.

[HYZ07] Elaine T Hale, Wotao Yin, and Yin Zhang. A fixed-point continuation method for l_1-regularized minimization with applications to compressed sensing. *CAAM Technical Report TR07-07, Department of Computational and Applied Mathematics, Rice University*, 43:44, 2007.

[Inc13] Megvii Inc. Face++ research toolkit. www.faceplusplus.com, December 2013.

[JL15] Xudong Jiang and Jian Lai. Sparse and dense hybrid representation via dictionary decomposition for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(5):1067–1079, 2015.

[JWL12] Yong Jin, Qingbiao Wu, and Ligang Liu. Unsupervised upright orientation of man-made models. *Graphical Models*, 74(4):99–108, 2012.

[KK05] Qifa Ke and Takeo Kanade. Robust l_1-norm factorization in the presence of outliers and missing data by alternative convex programming. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 739–746, 2005.

[Kwa08] Nojun Kwak. Principal component analysis based on l_1-norm maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9):1672–1680, 2008.

[LB00] Aleš Leonardis and Horst Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1):99–118, 2000.

[LCM09] Zhouchen Lin, Minming Chen, and Yi Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *Technical Report UILU-ENG-09-2215, University Illinois at Urbana-Champaign*, 2009.

[LCN14] Enming Luo, Stanley H Chan, and Truong Q Nguyen. Image denoising by targeted external databases. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pages 2450–2454, 2014.

[LCZ+13] Wee Kheng Leow, Yuan Cheng, Li Zhang, Terence Sim, and Lewis Foo. Background recovery by fixed-rank robust principal component analysis. In *Proceedings of International Conference on Computer Analysis of Images and Patterns*, pages 54–61, 2013.

[LDZR13] Xiao-Xin Li, Dao-Qing Dai, Xiao-Fei Zhang, and Chuan-Xian Ren. Structured sparse error coding for face recognition with occlusion. *IEEE Transactions on Image Processing*, 22(5):1889–1900, 2013.

[LL14] Xiao Luan and Weisheng Li. Facial feature extraction based on robust PCA and histogram. *Pattern Recognition*, 484:296–302, 2014.

[LLY10] Guangcan Liu, Zhouchen Lin, and Yong Yu. Robust subspace segmentation by low-rank representation. In *Proceedings of International Conference on Machine Learning*, pages 663–670, 2010.

[LRL14] Si Lu, Xiaofeng Ren, and Feng Liu. Depth enhancement via low-rank matrix completion. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 3390–3397, 2014.

[LT07] Dahua Lin and Xiaoou Tang. Quality-driven face occlusion detection and recovery. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7, 2007.

[MD12] Rui Min and Jean-Luc Dugelay. Inpainting of sparse occlusion in face recognition. In *Proceedings of International Conference on Image Processing*, pages 1425–1428, 2012.

[MHD14] Rui Min, Abdenour Hadid, and Jean-Luc Dugelay. Efficient detection of occlusion prior to robust face recognition. *The Scientific World Journal*, 2014, 2014.

[MLN04] Zhenyao Mo, John P Lewis, and Ulrich Neumann. Face inpainting with local linear representations. In *Proceedings of British Machine Vision Conference*, pages 347–356, 2004.

[NLET08] Minh Hoai Nguyen, Jean-Francois Lalonde, Alexei A Efros, and F. D. la Torre. Image-based shaving. In *Proceedings of Annual Conference of the European Association for Computer Graphics*, pages 627–635, 2008.

[OLL08] Hyun Jun Oh, Kyoung Mu Lee, and Sang Uk Lee. Occlusion invariant face recognition using selective local non-negative matrix factorization basis images. *Image and Vision Computing*, 26(11):1515–1523, 2008.

[PGW+12] Yigang Peng, Arvind Ganesh, John Wright, Wenli Xu, and Yi Ma. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2233–2246, 2012.

[PMRR00] P Jonathon Phillips, Hyeonjoon Moon, Syed A Rizvi, and Patrick J Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.

[POAL05] Jeong-Seon Park, You Hwa Oh, Sang Chul Ahn, and Seong-Whan Lee. Glasses removal from facial image using recursive error compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5):805–811, 2005.

[SBL02] Danijel Skočaj, Horst Bischof, and Aleš Leonardis. A robust PCA algorithm for building representations from panoramic images. In *Proceedings of European Conference on Computer Vision*, pages 761–775, 2002.

[SKK99] Yasuyuki Saito, Yukiko Kenmochi, and Kazunori Kotani. Estimation of eyeglassless facial images using principal component analysis. In *Proceedings of International Conference on Image Processing*, pages 197–201, 1999.

[SLC11] Jason M Saragih, Simon Lucey, and Jeffrey F Cohn. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision*, 91(2):200–215, 2011.

[SRUB09] Markus Storer, Peter M Roth, Martin Urschler, and Horst Bischof. Fast-robust PCA. In *Proceedings of Scandinavian Conference on Image Analysis*, pages 430–439, 2009.

[SUB10] Markus Storer, Martin Urschler, and Horst Bischof. Occlusion detection for icao compliant facial photographs. In *Proceedings of Workshops at IEEE Conference on Computer Vision and Pattern Recognition*, pages 122–129, 2010.

[WGR+09] John Wright, Arvind Ganesh, Shankar Rao, Yigang Peng, and Yi Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *Proceedings of Advances in Neural Information Processing Systems*, pages 2080–2088, 2009.

[WJ15] Yue Wu and Qiang Ji. Robust facial landmark detection under significant head poses and occlusion. In *Proceedings of International Conference on Computer Vision*, pages 3658–3666, 2015.

[WLS+04] Chenyu Wu, Ce Liu, Heung-Yueng Shum, Ying-Qing Xy, and Zhengyou Zhang. Automatic eyeglasses removal from face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):322–336, 2004.

[WT07] Zhi-Ming Wang and Jian-Hua Tao. Reconstruction of partially occluded face by fast recursive PCA. In *Proceedings of International Conference on Computational Intelligence and Security Workshops*, pages 304–307, 2007.

[WX10] Ziheng Wang and Xudong Xie. An efficient face recognition algorithm based on robust principal component analysis. In *Proceedings of International Conference on Internet Multimedia Computing and Service*, pages 99–102, 2010.

[WYG+09] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

[XS10] Zongben Xu and Jian Sun. Image inpainting by patch propagation using patch sparsity. *IEEE Transactions on Image Processing*, 19(5):1153–1165, 2010.

[XY95] Lei Xu and Alan L Yuille. Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE Transactions on Neural Networks*, 6(1):131–143, 1995.

[YLBM14] Xiang Yu, Zhe Lin, Jonathan Brandt, and Dimitris N Metaxas. Consensus of regression for occlusion-robust facial feature localization. In *Proceedings of European Conference on Computer Vision*, pages 105–118, 2014.

[YS08] Dan Yu and Terence Sim. Using targeted statistics for face regeneration. In *Proceedings of International Conference on Automatic Face & Gesture Recognition*, pages 1–8, 2008.

[YZYZ11] Meng Yang, Lei Zhang, Jian Yang, and David Zhang. Robust sparse coding for face recognition. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 625–632, 2011.

[ZL10] Qiang Zhang and Baoxin Li. Joint sparsity model with matrix completion for an ensemble of face images. In *Proceedings of International Conference on Image Processing*, pages 1665–1668, 2010.

[ZLGM10] Zhengdong Zhang, Xiao Liang, Arvind Ganesh, and Yi Ma. TILT: transform invariant low-rank textures. In *Proceedings of Asian Conference on Computer Vision*, pages 314–328, 2010.

[ZLW+10] Zihan Zhou, Xiaodong Li, John Wright, Emmanuel Candes, and Yi Ma. Stable principal component pursuit. In *Proceedings of International Symposium on Information Theory Proceedings*, pages 1518–1522, 2010.

[ZSCG07] Wenchao Zhang, Shiguang Shan, Xilin Chen, and Wen Gao. Local Gabor binary patterns based on Kullback-Leibler divergence for partially occluded face recognition. *IEEE Signal Processing Letters*, 14(11):875–878, 2007.