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: _____________       Date: _____________

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

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>List of Publications</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Thesis Objective</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Outline of Thesis</td>
<td>4</td>
</tr>
<tr>
<td>2 Related Work</td>
<td>5</td>
</tr>
<tr>
<td>2.1 PCA</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Robust PCA</td>
<td>6</td>
</tr>
<tr>
<td>2.3 Sparse Coding</td>
<td>7</td>
</tr>
<tr>
<td>2.4 Summary</td>
<td>8</td>
</tr>
<tr>
<td>3 Variable Threshold Robust PCA</td>
<td>10</td>
</tr>
<tr>
<td>3.1 Robust PCA</td>
<td>10</td>
</tr>
</tbody>
</table>
Abstract

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

List of Figures

1.1 Examples of face occlusion in real life. .............................. 2

3.1 RPCA schematic diagram. .................................................. 11

4.1 Feature points and TPS alignment result. ............................ 18

4.2 Characteristics of synthetic occluders. ............................... 19

4.3 Comparison of execution time of various algorithms across different occlusion size ratios. ............................................ 21

4.4 Effect of occluder intensity on Multi-PIE test images. ............... 23

4.5 Effect of occluder intensity on FERET test images. .................. 24

4.6 Sample results of occluder intensity test for Multi-PIE familiar set. . 25

4.7 Sample results of occluder intensity test for Multi-PIE unfamiliar set. 26

4.8 Sample results of occluder intensity test for FERET familiar set. .... 27

4.9 Sample results of occluder intensity test for FERET unfamiliar set. . 28

4.10 Effect of occluder size ratio on Multi-PIE test images. ............... 30

4.11 Effect of occluder size ratio on FERET test images. ................. 31

4.12 Sample results of occluder size ratio for Multi-PIE familiar set. ...... 32

4.13 Sample results of occluder size ratio for Multi-PIE unfamiliar set. ... 33

4.14 Sample results of occluder size ratio for FERET familiar set. ........ 34
4.15 Sample results of occluder size ratio for FERET unfamiliar set. . . . . 35
4.16 Effect of occluder location on Multi-PIE test images. . . . . . . . . . 36
4.17 Effect of occluder location on FERET test images. . . . . . . . . . . . 37
4.18 Sample results of occluder location test for Multi-PIE familiar set. . 38
4.19 Sample results of occluder location test for Multi-PIE unfamiliar set. 39
4.20 Sample results of occluder location test for FERET familiar set. . . . 40
4.21 Sample results of occluder location test for FERET unfamiliar set. . 41
4.22 Effect of occluder shape on Multi-PIE test images. . . . . . . . . . . . 42
4.23 Effect of occluder shape on FERET test images. . . . . . . . . . . . . 43
4.24 Sample results of occluder shape test for Multi-PIE familiar set. . . . 44
4.25 Sample results of occluder shape test for Multi-PIE unfamiliar set. . 44
4.26 Sample results of occluder shape test for FERET familiar set. . . . . 45
4.27 Sample results of occluder shape test for FERET unfamiliar set. . . . 45
List of Tables

2.1 Comparison of face image de-occlusion methods. . . . . . . . . . . . . . . . . 9

4.1 Various occluder characteristics for generating test images. . . . . . . 18

4.2 Effect of weights of soft thresholds on VRPCA’s performance. . . . . 20
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, YZZY11] 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 im-
age. 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, DITB01, 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.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Citation</th>
<th>Approach</th>
<th>No corruption of unoccluded parts</th>
<th>Consistent to various cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>[LT07] [HNK+12]</td>
<td>Global Processing</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>[YS08]</td>
<td>Local Replacement</td>
<td>Yes (except the surrounding)</td>
<td>Yes</td>
</tr>
<tr>
<td>RPCA</td>
<td>[MD12] [WGR+09]</td>
<td>Global Processing</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SC</td>
<td>[WYG+09] [YZYZ11]</td>
<td>Global Processing</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>NA</td>
<td>Global Processing</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Chapter 3

Variable Threshold Robust PCA

3.1 Robust PCA

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

\[
\min_{A,E} \|E\|_F \quad \text{subject to} \quad \text{rank}(A) \leq r, \quad D = A + E, \tag{3.1}
\]

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

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

\[
\min_{A,E} \text{rank}(A) + \lambda \|E\|_0 \quad \text{subject to} \quad D = A + E, \tag{3.2}
\]

where \( A \) is the recovered low-rank matrix, while \( E \) is a sparse error matrix capturing the outliers. \( \lambda \) 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 $D$ to low-rank matrix $A$ and sparse error matrix $E$.

$L_0$-norm can be replace by nuclear norm and $L_1$-norm, respectively. As a result, the revised problem becomes:

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

(3.3)

where $L_1$-norm of $\|E\|_1$ is defined as the sum of absolute values of all entries in $E$, while the nuclear norm $\|A\|_*$ is the sum of singular values of $A$.

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:

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

(3.4)

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

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

(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 \(M\) with SVD \(USV^\top\),

\[
UT_{\varepsilon}(S)V^\top = \arg \min_X \varepsilon \|X\|_* + \frac{1}{2} \|M - X\|_F^2, \tag{3.7}
\]

\[
T_{\varepsilon}(M) = \arg \min_X \varepsilon \|X\|_1 + \frac{1}{2} \|M - X\|_F^2. \tag{3.8}
\]

The matrices \(A\) and \(E\) are optimized using alternating optimization method. Optimizing \(A\) with other variables fixed reduces Formula 3.5 to:

\[
\begin{align*}
\min_A \|A\|_* + \lambda \|E\|_1 + \text{tr}(Y^\top(D - A - E)) + \frac{\mu}{2} \|D - A - E\|_F^2 & \\
\Rightarrow \min_A \|A\|_* + \text{tr}(Y^\top(D - A - E)) + \frac{\mu}{2} \|D - A - E\|_F^2 & \\
\Rightarrow \min_A \|A\|_* + \frac{\mu}{2} \text{tr}((D - A - E)^\top(D - A - E)) & \\
\Rightarrow \min_A \|A\|_* + \frac{\mu}{2} \text{tr}((E - (D - A + Y/\mu))^\top(A - (D - E + Y/\mu))) & \\
\Rightarrow \min_A \|A\|_* + \frac{\mu}{2} \|A - (D - E + Y/\mu)\|_F^2 & \\
\Rightarrow \min_A \frac{1}{\mu} \|A\|_* + \frac{1}{2} \|D - E + Y/\mu - A\|_F^2.
\end{align*}
\]

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

\[
A = UT_{\frac{1}{\mu}}(S)V^\top, \text{ with } USV^\top = D - E + Y/\mu. \tag{3.10}
\]
Similarly, optimizing $E$ with other variables fixed implies:

$$
\min_E \|A\|_* + \lambda \|E\|_1 + \mu \|D - A - E\|_F^2 \\
\Rightarrow \min_E \lambda \|E\|_1 + \text{tr}(Y^\top(D - A - E)) + \frac{\mu}{2} \|D - A - E\|_F^2 \\
\Rightarrow \min_E \lambda \|E\|_1 + \text{tr}((-Y^\top E) + \frac{\mu}{2} \text{tr}((D - A - E)^\top(D - A - E))) \\
\Rightarrow \min_E \lambda \|E\|_1 + \frac{\mu}{2} \text{tr}(E - (D - A + Y/\mu)) \top(E - (D - A + Y/\mu))) \\
\Rightarrow \min_E \lambda \|E\|_1 + \frac{\mu}{2} \|E - (D - A + Y/\mu)\|_F^2 \\
\Rightarrow \min_E \frac{\lambda}{\mu} \|E\|_1 + \frac{1}{2} \|(D - A + Y/\mu) - E\|_F^2.
$$

(3.11)

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

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

### 3.2 Variable-Threshold RPCA

For de-occlusion of face image, $D = [T \ x]$ contains $n$ unoccluded training face images arranged in columns in $T$ and an occluded target face image $x$. Given the prior knowledge of the occluded part, VRPCA separates the error matrix $E$ into three parts for training images, occluded part of target image and unoccluded part of target image. By separating $E$ it into three different parts, variable-threshold is introduced to calculate the error matrix:

$$
E = \begin{cases} 
T_{w,\lambda/\mu}(D - A + Y/\mu), & \text{for training images } T, \\
T_{w_o,\lambda/\mu}(D - A + Y/\mu), & \text{for occluded part of } x, \\
T_{w_u,\lambda/\mu}(D - A + Y/\mu), & \text{for unoccluded part of } x.
\end{cases} \tag{3.13}
$$
Define the matrix $W = [w_{ij}] \in \mathbb{R}^{m \times (n+1)}$ such that

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

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

The error of the occluded part of target $x$ 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:** $D$, $W$.

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

**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 [LCM09]:

$$J(X) = \max(\|X\|_2, \lambda^{-1} \|X\|_\infty). \quad (3.15)$$
Lines 5 to 7 are derived from Formula 3.12, Formula 3.10. Line 7 applies different weights to the soft-threshold $\lambda/\mu$ 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

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

(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 $D - E + Y/\mu$ 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 $A$ and $E$ 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
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 × 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.

Table 4.1: Various occluder characteristics for generating test images.

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

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.
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

<table>
<thead>
<tr>
<th>$w_u$</th>
<th>$w_o$</th>
<th>0</th>
<th>0.0001</th>
<th>0.001</th>
<th>0.01</th>
<th>0.1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>74.36</td>
<td>11.16</td>
<td>11.16</td>
<td>11.17</td>
<td>12.28</td>
<td>74.36</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>131.38</td>
<td>11.38</td>
<td>11.39</td>
<td>11.44</td>
<td>13.06</td>
<td>75.22</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>131.38</td>
<td>11.39</td>
<td>11.39</td>
<td>11.44</td>
<td>13.06</td>
<td>75.22</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>131.38</td>
<td>11.39</td>
<td>11.39</td>
<td>11.44</td>
<td>13.06</td>
<td>75.22</td>
</tr>
</tbody>
</table>

(b) Unoccluded parts

<table>
<thead>
<tr>
<th>$w_u$</th>
<th>$w_o$</th>
<th>0</th>
<th>0.0001</th>
<th>0.001</th>
<th>0.01</th>
<th>0.1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>10.91</td>
<td>9.98</td>
<td>9.98</td>
<td>9.97</td>
<td>9.93</td>
<td>10.91</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- VRPCA: the proposed variable-threshold RPCA.
- VRPCA2: variation of VRPCA that is equivalent to RPCA for matrix completion of [LCM09].
- 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 $\rho$ and initial $\mu$ were set to the default values of 6 and $0.5/\sigma_1$, where $\sigma_1$ is the largest singular value of the initial $Y$. The parameter $\lambda$ 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]
Chapter 4. Experiments and Discussions

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
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
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.
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.

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


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


