ABSTRACT

Title of thesis: BLUR AND ILLUMINATION-INVARIANT FACE RECOGNITION VIA SET-THEORETIC CHARACTERIZATION

Priyanka Vageeswaran, Master of Science, 2013

Thesis directed by: Professor Rama Chellappa
Department of Electrical and
Computer Engineering

In this thesis we address the problem of unconstrained face recognition from remotely acquired images. The main factors that make this problem challenging are image degradation due to blur, and appearance variations due to illumination and pose. In this thesis we address the problems of blur and illumination. We show that the set of all images obtained by blurring a given image forms a convex set. Based on this set-theoretic characterization, we propose a blur-robust algorithm whose main step involves solving simple convex optimization problems. We do not assume any parametric form for the blur kernels, however, if this information is available it can be easily incorporated into our algorithm. Further, using the low-dimensional model for illumination variations, we show that the set of all images obtained from a face image by blurring it and by changing the illumination conditions forms a bi-convex set. Based on this characterization we propose a blur and illumination-robust algorithm. Our experiments on a challenging real dataset obtained in uncontrolled settings illustrate the importance of jointly modeling blur

and illumination.

Blur and Illumination Robust Face Recognition via Set Theoretic
Characterization

by

Priyanka Vageeswaran

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2013

Advisory Committee:
Professor Rama Chellappa, Chair/Advisor
Professor Larry S. Davis
Professor Min Wu

Acknowledgments

I would like to thank my family, friends, lab-mates and my advisor Dr. Rama Chellappa for all the support and guidance I have recieved through these years. I would also like to thank my committee members for agreeing to be part of my committee despite their very busy schedules. Lastly I would like to thank my collaborator Dr. Kaushik Mitra for all his technical help.

Table of Contents

List of Figures

Chapter 1:   Introduction

Face recognition has been an intensely researched field of computer vision for the past couple of decades [1]. Though significant strides have been made in tackling the problem in controlled domains (as in recognition of passport photographs) [1], significant challenges remain in solving it in the unconstrained domain. One such scenario occurs while recognizing faces acquired from distant cameras. The main factors that make this a challenging problem are image degradations due to blur and noise, and variations in appearance due to illumination and pose [2] (see Figure 1.1). In this thesis, we specifically address the problem of recognizing faces across blur and illumination.

An obvious approach to recognizing blurred faces would be to deblur the image first and then recognize it using traditional face recognition techniques [3]. However, this approach involves solving the challenging problem of blind image deconvolution [4, 5]. We avoid this unnecessary step and propose a direct approach for face recognition. We show that the set of all images obtained by blurring a given image forms a convex set, and more specifically, we show that this set is the convex hull of shifted versions of the original image. Thus with each gallery image we can associate a corresponding convex set. Based on this set-theoretic characterization, we propose a blur-robust face recognition algorithm. In the basic version of our algorithm, we compute the distance of a given probe image (which we want to recognize) from each of the convex sets, and assign it the identity of the closest gallery image. The distance-computation steps are formulated as convex optimization problems over the space of blur kernels. We do not assume any parametric or symmetric form for the blur kernels; however, if this information is available, it can be easily incorporated into our algorithm, resulting in improved recognition performance. Further, we make our algorithm robust to outliers and small pixel mis-alignments

Figure 1.1: *Face images captured by a distant camera in unconstrained settings.* The main challenges in recognizing such faces are variations due to blur, pose and illumination. In this thesis we specifically address the problems of blur and illumination.

by replacing the Euclidean distance by weighted L_1-norm distance and comparing the images in the LBP (local binary pattern) [6] space.

It has been shown in [7] and [8] that all the images of a Lambertian convex object, under all possible illumination conditions, lie on a low-dimensional (approximately nine-dimensional) linear subspace. Though faces are not exactly convex or Lambertian, they can be closely approximated by one. Thus each face can be characterized by a low-dimensional subspace, and this characterization has been used for designing illumination robust face recognition algorithms [7,9]. Based on this illumination model, we show that the set of all images of a face under all blur and illumination variations is a bi-convex one. That is- if we fix the blur kernel then the set of images obtained by varying the illumination conditions forms a convex set; and if we fix the illumination condition then the set of all blurred images is also convex. Based on this set-theoretic characterization, we pro-

pose a blur and illumination robust face recognition algorithm. The basic version of our algorithm computes the distance of a given probe image from each of the bi-convex sets, and assigns it the identity of the closest gallery image. The distance computations steps can be formulated as 'quadratically constrained quadratic programs' (QCQPs), which we solve by alternately optimizing over the blur kernels and the illumination coefficients. Similar to the blur-only case, we make our algorithm robust to outliers and small pixel mis-alignments by replacing the Euclidean norm by the weighted L_1-norm distance and comparing the images in the LBP space.

To summarize, the main technical contributions of this thesis are:

- We show that the set of all images obtained by blurring a given image forms a convex set. More specifically, we show that this set is the convex hull of shifted versions of the original image.

- Based on this set-theoretic characterization, we propose a blur-robust face recognition algorithm, which avoids solving the challenging and unnecessary problem of blind image deconvolution.

- If we have additional information on the type of blur affecting the probe image, we can easily incorporate this knowledge into our algorithm, resulting in improved recognition performance and speed.

- We show that the set of all images of a face under all blur and illumination variations forms a bi-convex set. Based on this characterization, we propose a blur and illumination robust face recognition algorithm.

We have published a portion of this work in [10].

1.0.1 Related Work

Face recognition from blurred images can be classified into four major approaches. In the first approach, the blurred image is first deblurred and then used for recognition. This is the approach taken in [11] and [3]. The drawback of this approach is that we first need to solve the challenging problem of blind image deconvolution. Though there have been many attempts at solving the blind deconvolution problem [4, 5, 12–14], it is an avoidable step for the face recognition problem. Also, in [3] statistical models are learned for each blur kernel type and amount; this step might become infeasible when we try to capture the complete space of blur kernels.

In the second approach, blur invariant features are extracted from the blurred image and then used for recognition; [15] and [16] follow this approach. In [15], the local phase quantization (LPQ) [17] method is used to extract blur invariant features. Though this approach works very well for small blurs, it is not very effective for large blurs [3]. In [16], a (blur) subspace is associated with each image and face recognition is performed in this feature space. It has been shown that the (blur) subspace of an image contains all the blurred version of the image. However, this analysis does not take into account the convexity constraint that the blur kernels satisfy, and hence the (blur) subspace will include many other images apart from the blurred images. The third approach is the direct recognition approach. This is the approach taken in [18] and by us. In [18], artificially blurred versions of the gallery images are created and the blurred probe image is matched

to them. Again, it is not possible to capture the whole space of blur kernels using this method. We avoid this problem by optimizing over the space of blur kernels. Finally, the fourth approach is to jointly deblur and recognition the face image [19]. However, this involves solving for the original sharp image, blur kernel and identity of the face image, and hence it is a computationally intensive approach.

Set theoretic approaches for signal and image restoration have been considered in [20–22]. In these approaches the desired signal space is defined as an intersection of closed convex sets in a Hilbert space, with each set representing a signal constraint. Image de-blurring has also been considered in this context [21], where the non-negativity constraint of the images has been used to restrict the solution space. We differ from these approaches as our primary interest lies in recognizing blurred and poorly illuminated faces rather than restoring them.

There are mainly two approaches for recognizing faces across illumination variation. One approach is based on the low-dimensional linear subspace model [7, 8]. In this approach, each face is characterized by its corresponding low-dimensional subspace. Given a probe image, its distance is computed from each of the subspaces, and it is then assigned to the face image with the smallest distance [7,9]. The other approach is based on extracting illumination insensitive features from the face image and using them for matching. Many features have been proposed for this purpose such as self-quotient images [23], correleration filters [24], Eigenphases method [25], image preprocessing algorithms [26], gradient direction [27, 28] and albedo estimates [29].

The organization of the rest of the thesis is as follows: In chapter 2 we provide a set-theoretic characterization of the space of blurred images and subsequently propose

our approach for recognizing blurred faces, in section 2.1 we incorporate the illumination model in our approach and in chapter 3 we perform experiments to evaluate the efficacy of our approach on many synthetic and real datasets.

Chapter 2:   Direct Recognition of Blurred Faces

We first review the convolution model for blur. Next, we show that the set of all images obtained by blurring a given image is convex and leveraging on this set-theoretic characterization, we present our algorithm for recognizing blurred faces. Since, this model is inadequate to handle variations in illumination, we propose to model illumination using a low-dimensional linear subspace model. Finally we present our algorithm for jointly handling blur and illumination.

2.0.2 Convolution Model for Blur

A pixel in a blurred image is a weighted average of the pixel's neighborhood in the original sharp image. Thus, blur is modeled as a convolution operation between the original image and a blur filter kernel which represents the weights [30]. Let I be the original image and H be the blur kernel of size $(2k + 1) \times (2k + 1)$, then the blurred image I_b is given by

$$I_b(r, c) = I * H(r, c) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} H(i, j) I(r - i, c - j) \tag{2.1}$$

where $*$ represents the convolution operator and r, c are the row and column indices of the image. Blur kernels also satisfy the following properties- their coefficients are non-negative, $H \geq 0$, and sum up to 1 (i.e. $\sum_{i=-k}^{k} \sum_{j=-k}^{k} H(i, j) = 1$). The blur kernel may possess additional structure depending on the type of blur (such as circular-symmetry for out-of-focus blurs), and these structures could be exploited during recognition.

2.0.3 The Set of All Blurred Images

We want to characterize the set of all images obtained by blurring a given image I. To do this we re-write (2.1) in a matrix-vector form. Let $\mathbf{h} \in \mathbb{R}^{(2k+1)^2}$ be the vector obtained by concatenating the columns of H, i.e. $\mathbf{h} = H(:)$ in MATLAB notation, and similarly $\mathbf{i_b} = I_b(:) \in \mathbb{R}^N$ be the representation of I_b in the vector form, where N is the number of pixels in the blurred image . Then we can write (2.1), along with the blur kernel constraints, as

$$\mathbf{i_b} = \mathbf{Ah} \text{ such that } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1 \tag{2.2}$$

where $\mathbf{A}$ is a $N \times (2k+1)^2$ matrix, obtained from I, with each row of $\mathbf{A}$ representing the neighborhood pixel intensities about the pixel indexed by the row. From 2.2, it is clear that the set of all blurred images obtained from I is given by

$$\mathcal{B} \triangleq \{\mathbf{Ah}|\mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1\} \tag{2.3}$$

We have the following result about the set $\mathcal{B}$.

Proposition 2.0.1. *The set of all images $\mathcal{B}$ obtained by blurring an image I is a convex set. Moreover, this convex set is given by the convex hull of the columns of matrix $\mathbf{A}$, where the columns of $\mathbf{A}$ are various shifted versions of I as determined by the blur kernel.*

Proof. Let $\mathbf{i_1}$ and $\mathbf{i_2}$ be elements from the set $\mathcal{B}$. Then there exists $\mathbf{h_1}$ and $\mathbf{h_2}$, with both satisfying the conditions $\mathbf{h} \geq 0$ and $\|\mathbf{h}\|_1 = 1$, such that $\mathbf{i_1} = \mathbf{Ah_1}$ and $\mathbf{i_2} = \mathbf{Ah_2}$. To show that the set $\mathcal{B}$ is convex we need to show that for any λ satisfying $0 \leq \lambda \leq 1$,
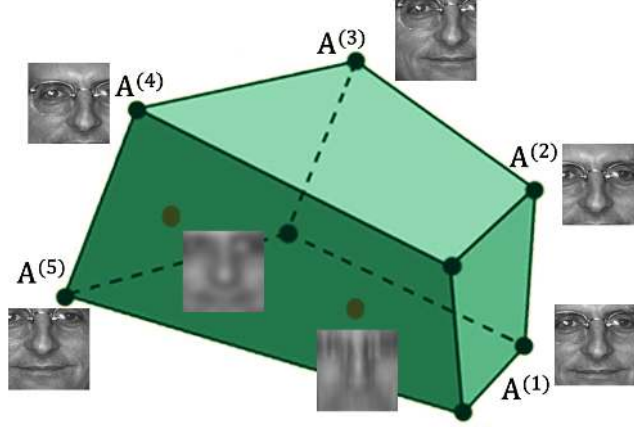
Figure 2.1: *The set of all images obtained by blurring an image I is a convex set.* Moreover, this convex set is given by the convex hull of the columns of the matrix $\mathbf{A}$, which represents the various shifted versions of I as determined by the blur kernel.

$\mathbf{i_3} = \lambda \mathbf{i_1} + (1 - \lambda)\mathbf{i_2}$ is an element of $\mathcal{B}$. Now

$$\begin{aligned}
\mathbf{i_3} &= \lambda \mathbf{i_1} + (1 - \lambda)\mathbf{i_2} \\
&= \mathbf{A}(\lambda \mathbf{h_1} + (1 - \lambda)\mathbf{h_2}) \\
&= \mathbf{A}\mathbf{h_3}.
\end{aligned} \tag{2.4}$$

Note that $\mathbf{h_3}$ satisfies both the non-negativity and sum conditions and hence $\mathbf{i_3}$ is an element of $\mathcal{B}$. Thus, $\mathcal{B}$ is a convex set. $\mathcal{B}$ is defined as

$$\{\mathbf{A}\mathbf{h} | \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1\}, \tag{2.5}$$

which, by definition, is the convex hull of the columns of $\mathbf{A}$. $\qquad\square$

2.0.4 A Geometric Face Recognition algorithm

We first present the basic version of our blur-robust face recognition algorithm. Let $I_j, j = 1, 2, \ldots, M$ be the set of M sharp gallery images. From the analysis above, every

gallery image I_j has an associated convex set of blurred images $\mathcal{B}_j$. Given the probe image I_b, we find its distance from the set $\mathcal{B}_j$, which is the minimum distance between I_b and the points in the set $\mathcal{B}_j$. This distance r_j can be obtained by solving:

$$r_j = \min_{\mathbf{h}} ||\mathbf{i_b} - \mathbf{A_j h}||_2^2 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1 \tag{2.6}$$

This is a convex quadratic program which can be solved efficiently. For $\mathbf{i_b} \in \mathbb{R}^N$ and $\mathbf{h} \in \mathbb{R}^K$, the computational complexity is $\mathcal{O}(NK^2)$. We compute r_j for each $j = 1, 2, \ldots, M$ and assign I_b the identity of the gallery image with the minimum r_j. If there are multiple gallery images per class (person), we can use the *k*-nearest neighbor rule, i.e. we arrange the $r'_j s$ in ascending order and find the class which appears the most in the first *k* instances. In this algorithm we can also incorporate additional information about the type of blur. The most commonly occurring blur types are the out-of-focus, motion and the atmospheric blurs [30]. The out-of-focus and the atmospheric blurs are circularly-symmetric, i.e. the coefficients of H at the same radius are equal; whereas the motion blur is symmetric about the origin, i.e. $H(i, j) = H(-i, -j)$ [30]. Thus, having knowledge of the blur type, we solve (2.6) with an additional constraint on the blur kernel:

$$\begin{aligned} r_j &= \min_{\mathbf{h}} ||\mathbf{i_b} - \mathbf{A_j h}||_2^2 \\ &\text{subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \mathcal{C}(\mathbf{h}) = 0, \end{aligned} \tag{2.7}$$

where $\mathcal{C}(\mathbf{h}) = 0$ represents equality constraints on $\mathbf{h}$. Imposing these constraints reduces the number of parameters in the optimization problem giving better recognition accuracy and faster solutions.

Algorithm *Direct Recognition of Blurred Faces*

Input: (Blurred) probe image I_b and a set of gallery images I_j

Output: Identity of the probe image

1. For each gallery image I_j, find the optimal blur kernel $\mathbf{h_j}$ by solving either (2.9) or its robust version (A.7).

2. Blur each gallery image I_j with its corresponding $\mathbf{h_j}$ and extract LBP features.

3. Compare the LBP features of the probe image I_b with those of the gallery images and find the closest match.

Figure 2.2: *Direct Recognition of Blurred Faces (DRBF/rDRBF) Algorithm:* Our proposed algorithm for recognizing blurred faces.

2.0.5 Making the Algorithm Robust to Outliers and Misalignment

By making some minor modifications to the basic algorithm, we can make it robust to outliers and small pixel misalignments between the gallery and probe images. It is well known in face recognition literature [31] that different regions in the face have different amounts of information. To incorporate this fact we divide the face image into different regions and weigh them differently when computing the distance between the probe image I_b and gallery sets $\mathcal{B}_j$. That is, we modify the distance functions r_j as

$$r_j = \min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{A_j h})||_2^2$$

$$\text{subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \tag{2.8}$$

We learn the weight W, a diagonal matrix, using a training dataset. The training procedure is described in the APPENDIX.

Face recognition is also sensitive to small pixel mis-alignments and, hence, the general consensus in face recognition literature is to extract alignment insensitive features, such as Local Binary Patterns (LBP) [6, 15], and then perform recognition based on these features. Following this convention, instead of doing recognition directly from r_j, we first compute the optimal blur kernel $\mathbf{h_j}$ for each gallery image by solving (A.1), i.e.

$$\mathbf{h_j} = \arg\min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{A_j h})||_2^2$$

$$\text{subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \tag{2.9}$$

We then blur each of the gallery images with the corresponding optimal blur kernels $\mathbf{h_j}$ and extract LBP features from the blurred gallery images. And finally, we compare the

LBP features of the probe image with those of the gallery images to find the closest match.

To make our algorithm robust to outliers, which could arise due to variations in expression, we propose to replace the L_2 norm in (2.9) by the L_1 norm, i.e. we solve the problem:

$$\mathbf{h_j} = \arg\min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{A_j h})||_1$$

$$\text{subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \tag{2.10}$$

Note that the above optimization problem is a convex L_1-norm problem, which we formulate and solve as a Linear Programing (LP) problem. The computational complexity of this problem is $\mathcal{O}((K + N)^3)$. The overall algorithm is summarized in Figure 2.2.

2.1 Incorporating the Illumination Model

The facial images of a person under different illumination conditions can look very different, and hence for any recognition algorithm to work in practice, it must account for these variations. First, we discuss the low-dimensional subspace model for handling appearance variations due to illumination. Next, we use this model along with the convolution model to define the set of images of a face under all possible lighting conditions and blur. We then propose a recognition algorithm based on minimizing the distance of the probe image from such sets.

2.1.1 The Low-Dimensional Linear Model for Illumination Variations

It has been shown in [7, 8] that when an object is convex and Lambertian, the set of all images of the object under different illumination conditions can be approximately

represented using a nine-dimensional subspace. Though the human face is not exactly convex or Lambertian, it is often approximated as one; and hence the nine-dimensional subspace model captures its variations due to illumination quite well [32]. The nine-dimensional linear subspace corresponding to a face image I can be characterized by 9 basis images. In terms of these nine basis images $I_m, m = 1, 2, \ldots, 9$, an image I of a person under any illumination condition can be written as

$$I = \sum_{m=1}^{9} \alpha_m I_m \qquad (2.11)$$

where $\alpha_m, m = 1, 2, \ldots, 9$ are the corresponding linear coefficients. To obtain these basis images, we use the "universal configuration" of lighting positions proposed in [9]. These are a set of 9 lighting positions $s_m, m = 1, 2, \ldots, 9$ such that images taken under these lighting positions can serve as basis images for the subspace. These basis images are generated using the Lambertian reflectance model:

$$I_m(r, c) = \rho(r, c) \max(\langle s_m, n(r, c) \rangle, 0) \qquad (2.12)$$

where $\rho(r, c)$ and $n(r, c)$ are the albedo and surface-normal at pixel location (r, c). We use the average 3-D face normals from [33] for n and we approximate the albedo ρ with a well-illuminated gallery image under diffuse lighting.

If the gallery image is not well-illuminated under diffuse lighting, then the nine basis images cannot be obtained due to the lack of a proper albedo image. However, we believe that in a majority of face recognition/verification protocols such as the FRGC [34], the FERET [35] and and also others like the Yale Face Database B [36] and our own REMOTE dataset [37], the gallery is assumed to be a clear well-lit image. We were working under this setting. However, in the case of unconstrained acquisition settings for

the gallery, we could proceed by recovering the albedo from the poorly lit image using some existing work as in [29], [38] and [39]. These have been shown to work well in recovering the albedo map from poorly lit images by working on average surface normals and rough initial estimates of the albedo and source direction [29, 39]. In particular, [29] recovers the albedo map as the Linear Minimum Mean Square Error (LMMSE) estimate of the true albedo. On the other hand, [38] also accounts for unknown surface normals. In all cases, the recovery is noisy in some (non-Lambertian) regions of the face (like the nose). But these effects could potentially be mitigated by using weights that have been trained as described in the APPENDIX .

2.1.2 The set of all images under varying lighting and blur

For a given face characterized by the nine basis images $I_m, m = 1, 2, \ldots, 9$, the set of images under all possible lighting conditions and blur is given by

$$\mathcal{BI} \triangleq \{\sum_{m=1}^{9} \alpha_m \mathbf{A_m} \mathbf{h} | \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1, \mathcal{C}(\mathbf{h}) = 0\}, \tag{2.13}$$

where the matrix $\mathbf{A_m}$ is constructed from I_m and represents the pixel neighborhood structure. This set is not a convex set though if we fix either the filter kernel $\mathbf{h}$ or the illumination condition α_m the set becomes convex, see figure 2.3.

2.1.3 Illumination-robust Recognition of Blurred Faces (IRBF)

Corresponding to each sharp well-lit gallery image $I_j, j = 1, 2, \ldots, M$, we obtain the nine basis images $I_{j,m}, m = 1, 2, \ldots, 9$. Given the vectorized probe image $\mathbf{i_b}$, for each gallery image I_j we find the optimal blur kernel $\mathbf{h_j}$ and illumination coefficients $\alpha_{j,m}$ by

17
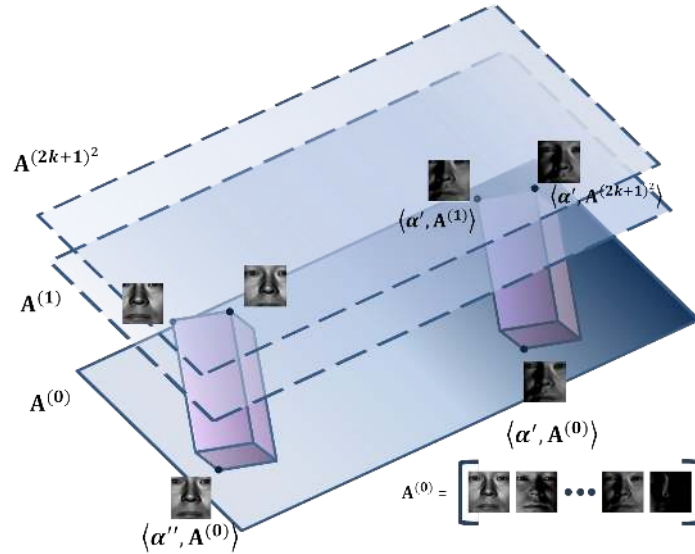
Figure 2.3: *The set of all images under varying lighting and blur for a single face image.* This set is a bi-convex set, i.e. if we fix either the filter kernel **h** or the illumination condition α, the resulting subset is convex. Each hyperplane in the figure represents the illumination subspace at different blur. For example, all points on the bottom-most plane are obtained by fixing the blur kernel at $\mathbf{h}^{(0)}$ (the impulse function centered at 0, i.e. the no-blur case) and varying the illumination conditions α. On this plane two data-points (faces), corresponding to illumination conditions α' and α'', are explicitly marked. Both these data points are associated with their corresponding blur convex hulls, see figure 2.1.

Algorithm *Illumination-robust Recognition of Blurred Faces*

Input: (Blurred and poorly illuminated) probe image I_b and a set of gallery images I_j

Output: Identity of the probe image

1. For each gallery image I_j, obtain the nine basis images $I_{j,m}, m = 1, 2, \ldots, 9$.

2. For each gallery image I_j, find the optimal blur kernel $\mathbf{h_j}$ and illumination coefficients $\alpha_{j,m}$ by solving either (A.5) or its robust version (2.15).

3. Transform (blur and re-illuminate) the gallery images I_j using the computed $\mathbf{h_j}$ and $\alpha_{j,m}$ and extract LBP features.

4. Compare the LBP features of the probe image I_b with those of the transformed gallery images and find the closest match.

Figure 2.4: *Illumination-Robust Recognition of Blurred Faces (IRBF/rIRBF) Algorithm:* Our proposed algorithm for jointly handling variations due to illumination and blur.

solving:

$$[\mathbf{h_j}, \alpha_{j,m}] = \arg\min_{\mathbf{h},\alpha_m} \|W(\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m \mathbf{A_{j,m}h})\|_2^2$$

$$\text{subject to } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \tag{2.14}$$

We then transform (blur and re-illuminate) each of the gallery images I_j using the computed blur kernel $\mathbf{h_j}$ and the illumination coefficients $\alpha_{j,m}$. Next, we compute the LBP features from these transformed gallery images and compare it with those from the probe image I_b to find the closest match, see Figure 2.4. The major computational step of the algorithm is the optimization problem of (A.5), which is a non-convex problem. To solve this problem we use an alternation algorithm in which we alternately minimize over $\mathbf{h}$ and α_m, i.e. in one step we minimize over $\mathbf{h}$ keeping α_m fixed and in the other step we minimize over α_m keeping $\mathbf{h}$ fixed and we iterate till convergence. Each step is now a convex problem: the optimization over $\mathbf{h}$ for fixed α_m reduces to the same problem as (2.9) and the optimization of α given $\mathbf{h}$ is just a linear least squares problem. The complexity of the overall alternation algorithm is $\mathcal{O}(T(N + K^3))$ where T is the number of iterations in the alternation step, and $\mathcal{O}(N)$ is the complexity in the estimation of the illumination coefficients. We also propose a robust version of the algorithm by replacing the L_2-norm in (A.5) with the L_1-norm:

$$[\mathbf{h_j}, \alpha_{j,m}] = \arg\min_{\mathbf{h},\alpha_m} \|W(\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m \mathbf{A_{j,m}h})\|_1$$

$$\text{subject to } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \tag{2.15}$$

Again, this is a non-convex problem and we use the alternation procedure which reduces each step of the algorithm to a convex L_1-norm problem. We formulate these L_1-norm

problems as Linear Programing (LP) problems. The complexity of the overall alternation algorithm is $\mathcal{O}(T(N^3 + (K + N)^3))$. The algorithm is summarized in Figure 2.4.

2.2 Computational Complexity

In this section we analyze the theoretical complexity of the algorithms we have proposed in our thesis. To solve the problem of recognizing blurred images we propose the DRBF algorithm. This involves solving a series of convex programs of the form,

$$r_j = \min_{\mathbf{h}} ||\mathbf{i_b} - \mathbf{A_j}\mathbf{h}||_2^2 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1 \qquad (2.16)$$

where $\mathbf{i_b} \in \mathbb{R}^N$ is the blurred image, $\mathbf{h} \in \mathbb{R}^K$ are the unknown blur kernel coefficients (note K could be less than the size of the blur kernel because of symmetry constraints on the blur kenel). The above problem can be formulated as a quadratic program which can be efficiently solved using interior-point methods or preconditioned conjugate gradients. The *quadprog* function in the Optimization Toolbox of MATLAB uses the later to converge to a global optimum with a worst-case complexity of $\mathcal{O}(\text{NK}^2)$. This is quite fast in real-life scenarios. For example the solver takes $2.3s$ for matching 1 query from the PIE dataset against 68 galleries for a blur of size $\sigma = 2$ on a standard 2.3GHz machine. rDRBF replaces the 2-norm cost with the 1-norm cost;

$$\mathbf{h_j} = \arg\min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{A_j}\mathbf{h})||_1 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1$$

This helps increase robustness of the algorithm to misalignment and variability induced due to expressions. The robust cost can be formulated as a linear programming

problem. We use the MATLAB solver *linprog* to solve this problem, which uses the primal-dual interior-point method with a time complexity of $\mathcal{O}((\text{K+N})^3)$. For 1 query in the PIE dataset (against 68 galleries for $\sigma = 2$) this takes about $3.5s$. the

The IRBF algorithm was formulated to handle illumination variation in the query. It alternates between estimating the blur coefficients using Equation 2.16 and estimating the illumination coefficients using the illumination subspace model. Thus the combined cost function becomes-

$$[\mathbf{h_j}, \alpha_{j,m}] = \arg \min_{\mathbf{h}, \alpha_m} \|W(\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m \mathbf{A_{j,m}h})\|_2^2 \text{ subject to } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1$$

The complexity of the algorithm is then $\mathcal{O}(\text{T}(\text{N} + \text{K}^3))$ where T is the number of iterations in the alternation step, and $\mathcal{O}(\text{N})$ is the complexity in the estimation of the illumination coefficients. For 6 iterations with $\sigma = 2$, we can match a query in PIE against 68 galleries in $8.85s$. Lastly, analogous to the 'blur-only' formulation rIRBF is the robust formulation of IRBF. This is again solved using *linprog* with a complexity of $\mathcal{O}(\text{T}(\text{N}^3 + (\text{K} + \text{N})^3))$ and an implementation time of $17.7s$. All the results on theoretic complexity of the convex programs we have mentioned here are conservative estimates. The actual MATLAB solvers are even more efficient.

Our algorithm is based on generative modeling followed by nearest-neighbor classification between the query image and the gallery space; which makes it difficult to scale it to real-life datasets with millions of images. However, this is a common problem with most algorithms based on generative modeling (for instance SRC claim [40] to take a few seconds for a query in the Extended Yale B database on a reduced feature space). Broadly

speaking, only classifier-based architectures have been shown to scale well to very large datasets since the size of the gallery largely affects the training stage; the testing phase remains relatively fast. Hence we believe that incorporating a discriminative-learning based approach like SVM into this formulation, would be a very promising direction for future work.

Chapter 3:   Experimental Evaluations

We evaluate the proposed algorithms: the 'blur-only' formulation DRBF of chapter 2 and the 'blur and illumination' formulation IRBF of section 2.1 on synthetically blurred datasets- FERET [35] and PIE [41], and a real dataset of remotely acquired faces with significant blur and illumination variations [2]- which we will refer to as the REMOTE dataset, see Figure 1.1. In section 3.0.1, we evaluate the performance of the DRBF algorithm in recognizing faces blurred by different types and amounts of blur. In section 3.0.2, we evaluate the effectiveness of the IRBF algorithm in recognizing blurred and poorly illuminated faces. Finally, in section 3.0.1, we evaluate our algorithms, DRBF and IRBF, on the real and challenging dataset of REMOTE.

3.0.1 Face Recognition across Blur

To evaluate our algorithm DRBF on different types and amounts of blur, we synthetically blur face images from the FERET dataset with four different types of blurs: out-of-focus, atmospheric, motion and general non-parametric blur. We use Gaussian kernels of varying standard deviations to approximate the out-of focus and atmospheric blurs [30], and rectangular kernels with varying lengths and angles for the motion blur. For the general blur we use the blur kernels used in [5]. Figure 3.1 shows some of the blur kernels and the corresponding blurred images. We compare our algorithm with the FADEIN approach [3] and the LPQ approach [17]. As discussed in chapter 1.0.1, the FADEIN approach first infers the deblurred image from the blurred probe image and then uses it for face recognition. On the other hand in the LPQ (local phase quantization) approach a blur insensitive image descriptor is extracted from the blurred image and recognition is

done on this feature space. We also compare our algorithms with 'FADEIN+LPQ' [3], where LPQ features extracted from the deblurred image produced by FADEIN is used for recognition.

(a) No Blur   (b) $\sigma = 8$ (c) M(25,90)   (d) M(21,0)   (e) M(9,45)   (f) General-1   (g) General-2
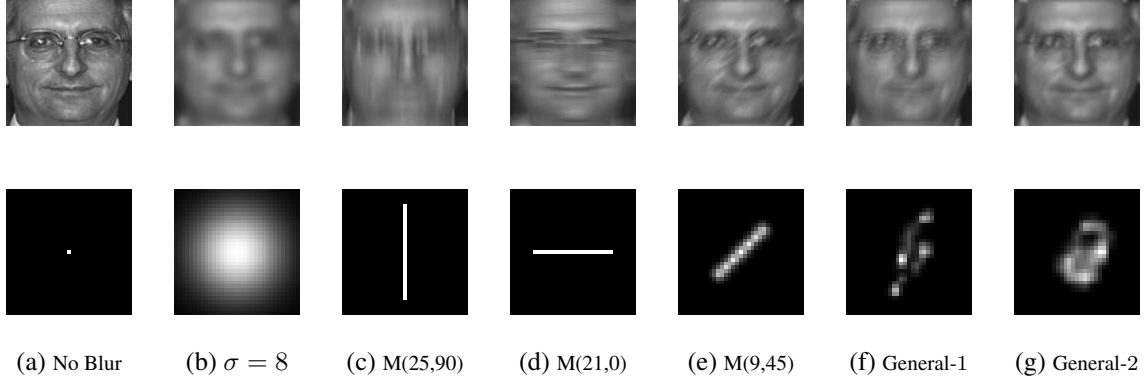
Figure 3.1: Examples of blur kernels and images used to evaluate our algorithms. The *General* blurs shown above have been borrowed from [5]

3.0.1.1 Out-of-Focus and Atmospheric Blurs

We synthetically generate face images from the FERET dataset using Gaussian kernels of various standard deviations for evaluation. We use the same experimental set-up as used in FADEIN [3], i.e. we chose our gallery set as 1001 individuals from the *fa* folder of the FERET dataset. The gallery set so constructed has one face image per person and the images are frontal and well-illuminated. We construct the probe set by blurring images of the same set of 1001 individuals from the *fb* folder of FERET (the images in this folder has slightly different expressions from the fa folder). We blur each individual image by Gaussian kernels of σ values $0, 2, 4,$ and 8 and kernel size $4\sigma + 1$.

To handle small variations in illumination, we histogram-equalize all the images
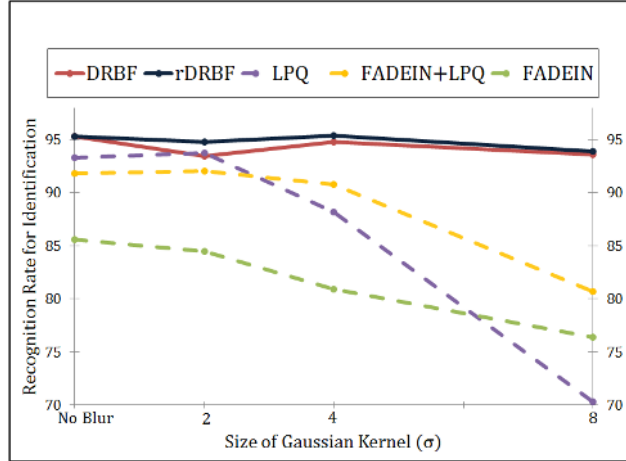
Figure 3.2: *Face recognition across Gaussian blur.* Recognition results by different algorithms as the amount of Gaussian blur is varied. Our algorithms, DRBF and its robust (L_1-norm) version rDRBF, shows significant improvement over the algorithms FADEIN, LPQ and FADEIN+LPQ, especially, for large blurs.

in the gallery and probe datasets. We then perform recognition using the DRBF algorithm and its robust (L_1) version rDRBF, with the additional constraint of circular symmetry imposed on the blur kernel. Figure 3.2 shows the recognition results obtained using the above approach along-side the recognition results from the FADEIN, LPQ and FADEIN+LPQ algorithms. Our algorithms-DRBF and its robust version rDRBF, show significant improvement over the other algorithms, especially for large blurs. rDRBF performs even better than DRBF owing to the more robust modeling of expressions and misalignment, as shown in Figure 3.3
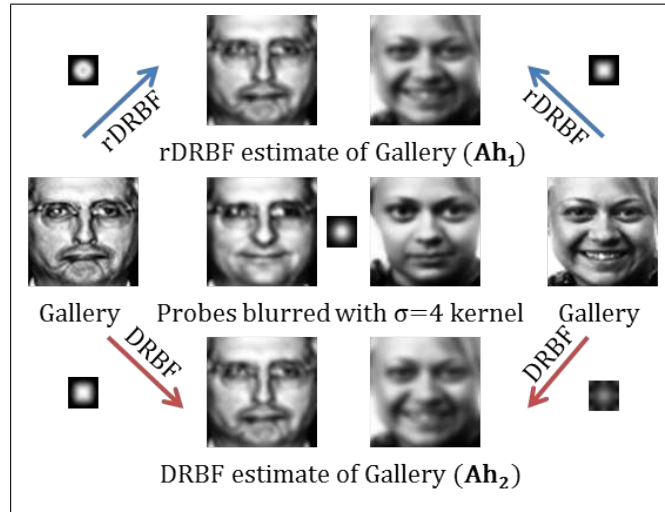
Figure 3.3: *Comparison of DRBF with its robust version rDBRF*: The robust version rDBRF can handle outliers, such as those due to expression variations, more effectively. Two gallery images along with their corresponding probes are shown in the center row. The probes have been blurred by a Gaussian blur of $\sigma = 4$. Note that the probe images have a different expression than the gallery images. The blur kernels estimated by the two algorithms rDRBF and DRBF are shown on the top and bottom rows respectively. As can be seen from the figure, the kernels estimated by rDBRF is closer to the actual kernel (at the center). The gallery images blurred by the estimated kernels further illustrate this fact, as the blurred gallery on the bottom row (corresponding to DRBF) looks significantly more blurred than the blurred gallery images in the top row (corresponding to rDBRF).

3.0.1.2 Motion and General Blurs

We now demonstrate the effectiveness of our algorithms, DRBF and rDRBF, on datasets degraded by motion and general non-parametric blurs. For this experiment we use the *ba* and *bj* folders in FERET, both of which contains 200 subjects with one image per subject. We use the *ba* folder as the gallery set. The probe set is formed by blurring the images in the *bj* folder by different motion and general blur kernels, some of them are shown in Figure 3.1. When we perform recognition using DRBF and rDRBF, we impose appropriate symmetry constraints for the blur types. That is, when we solve for the motion blur case, we impose the 'symmetry about the origin' constraint on the blur kernel, whereas, when we solve for the general or non-parametric blur case we do not impose any constraint. Figure 3.4 shows that DRBF and rDRBF perform consistently better than LPQ and FADEIN+LPQ. Hence, we can say that our method generalizes well to all forms of blur.

3.0.1.3 Effect of Blur Kernel-Size and Symmetry Constraints On DRBF

In all the experiments described above we have assumed that we know the type and size of the blur kernel, and have used this information while estimating the blur kernel in (2.9) or (A.7). For example, for images blurred by a Gaussian blur of standard deviation σ, we impose a kernel-size of $4\sigma + 1$ and circular symmetry. Though in some applications we may know the type of blur or the amount of blur, it may not be known for all applications. Hence, to test the sensitivity of our algorithm to blur kernel-size and blur type (symmetry constraint), we perform a few experiments.
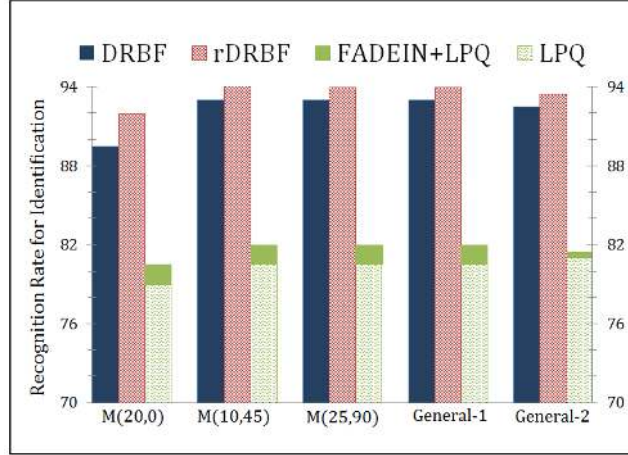
(a)

Figure 3.4: *Recognition result for motion and general blurs:* Performance of different algorithms on some selected motion and non-parametric blurs, see Figure 3.1. Our algorithms, DRBF and rDRBF, perform much better than LPQ and FADEIN+LPQ, with the robust version rDRBF always better than DRBF.

We use the *ba* folder of FERET as the gallery set and we create the probe set by blurring the images in the *bj* folder by a Gaussian kernel of $\sigma = 4$ and size $4\sigma + 1 = 17$. We then perform recognition via DRBF with choices of kernel size ranging from 1 to $32\sigma + 1$. We consider both the cases of imposing the circular symmetry constraint and not imposing any constraint. The experimental results are shown in Figure 3.5. We can see that the recognition rates are fairly stable for the case when we impose appropriate symmetry constraints. This is because imposing symmetry constraints reduces the solution space and makes it a more regularized problem. We also show the mean estimated kernels, from which it is clear that the algorithm works well even for large kernel sizes. On the other hand when no symmetry constraints are imposed, the recognition rate falls
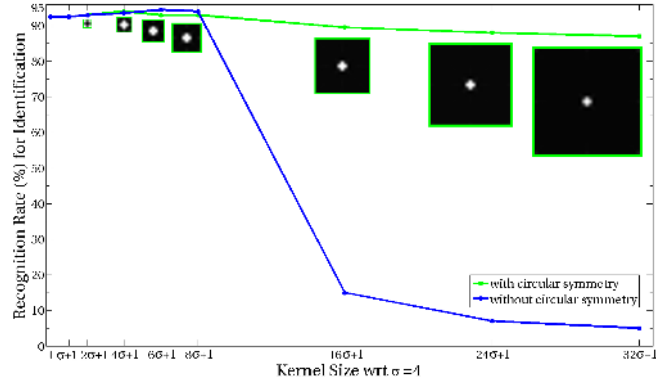
Figure 3.5: *Effect of blur kernel size and symmetry constraints on DRBF*. For this experiment, we use probe images blurred by a Gaussian kernel of $\sigma = 4$ and size $4\sigma + 1 = 17$ and perform recognition using DRBF with choices of kernel size ranging from 1 to $32\sigma + 1$. When we impose appropriate symmetry constraints, the recognition rate remains high even when we over-estimate the kernel size by a large margin (we have also shown the estimated blur kernels). This is because imposing symmetry constraints reduces the solution space, and makes it a more regularized problem. On the other hand when no symmetry constraints are imposed, the recognition rate falls drastically after a certain kernel size. However, as long as we do not over-estimate the kernel size by a large margin, we can expect a good performance from the algorithm.

drastically after a certain kernel size. However, as long as we do not over-estimate the kernel size by a large margin, we can expect good results from the algorithm. We conclude from these experiments that: $1)$ our algorithm exhibits a stable performance for a wide range of kernel-sizes as long as we do not over-estimate them by a large margin, $2)$ it is better to under-estimate the kernel size than over-estimate it and $3)$if we know the blur type then we should impose the corresponding symmetry constraints because imposing them further relaxes the need for an accurate estimate of the kernel-size.

Kernel Size(σ)	0		0.5		1.0		1.5		2.0		2.5		3.0	
Illumination	GI	BI	GI	BI	GI	BI	GI	BI	GI	BI	GI	BI	GI	BI
DRBF	99.63	95.10	99.63	84.55	**99.63**	78.67	**99.63**	77.95	**99.63**	77.45	97.79	58.58	95.58	42.40
IRBF	99.63	93.56	99.63	91.42	**99.63**	90.44	**99.63**	90.68	**99.63**	85.78	98.9	81.13	96.32	77.69
rIRBF	**99.7**	95.1	**99.7**	92.7	**99.63**	92.7	**99.63**	**91.6**	**99.63**	**88.2**	**99.63**	**84.78**	**97.45**	**81.36**
LPQ	99.63	**99.1**	99.63	**97.79**	99.63	**96.08**	99.63	88.97	97.05	73.04	79.42	58.08	46.32	27.7
FADEIN+LPQ	98.53	91.5	95.6	87.7	93.6	81.8	91.2	69.11	89.8	62.74	88.60	56.37	87.13	44.61

Table 3.1: *Recognition across Blur and Illumination on the PIE dataset.* GI and BI represent the 'good illumination' and 'bad illumination' subsets of the probe-set. IRBF and its robust version rIRBF out-perform the other algorithms for blurs of sizes greater than $\sigma = 1$. LPQ performs quite well for small blurs, but for large blurs its performance degrades significantly. This experiment clearly validates the need for modeling illumination and blur in a principled manner.

(a) *Good Illumina-* (b) *Bad Illumination (*BI*)* (c) *Illumination Basis Images*
*tion(*GI*)*

Figure 3.6: To study the effect of blur and illumination we use the PIE dataset which shows significant variation due to illumination. For each of the 68 subjects in PIE, we choose a well illuminated and frontal image of the person as the gallery set. The probe set, which is obtained from all the other frontal images, is divided into two categories: $1)$ Good Illumination (*GI*) set and $2)$ Bad Illumination (*BI*) set. Figures 3.6(a) and (b) shows some images from the *GI* and *BI* sets respectively. Figure 3.6(c) shows the 9 illumination basis images generated from a gallery image.

3.0.2   Recognition across Blur and Illumination

We study the effectiveness of our algorithms in recognizing blurred and poorly illuminated faces. We use the PIE dataset, which consists of images of 68 individuals under different illumination conditions. To study the effect of blur and illumination together, we synthetically blur the images with Gaussian kernels of varying σ's. We use face images with a frontal pose (c_{27}) and good illumination (f_{21}) as our gallery and the rest of the images in c_{27} as probe. We further divide the probe dataset into two categories- $1)$ Good Illumination (*GI*) consisting of subsets f_{09}, f_{11}, f_{12} and f_{20} and $2)$ Bad Illumination (*BI*)

consisting of $f_{13}, f_{14}, f_{15}, f_{16}, f_{17}$ and f_{22}, see Figure 3.6. We then blur all the probe images with Gaussian blurs of σ 0, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.

To perform recognition using our 'blur and illumination' algorithm IRBF, we first obtain the nine illumination basis images for each gallery image as described in section 2.1.1. We impose the circular symmetry constraints while solving the recognition problem using DRBF and IRBF. For comparison, we use LPQ and a modified version of FADEIN+LPQ. Since FADEIN does not model variations due to illumination, we preprocess the intensity images with the self-quotient method [23] and then run the algorithm. Table 3.1 shows the recognition results for the algorithms. We see that our algorithms IRBF and rIRBF out-perform the comparison algorithms for blurs of sizes $\sigma = 1.5$ and greater. Moreover, with a 8-core 2.9GHz procesor running MATLAB, it takes us $2.28s$, $9.53s$ and $17.74s$ per query image with a blur of $\sigma = 2$, for a gallery of 68 images for the DRBF, IRBF and rIRBF algorithms respectively. Thus we can conclude that our algorithms are able to maintain a consistent performance across increasing blur in a reasonable amount of time.

As discussed is section 2.1.3, the main optimization step in the IRBF (A.5) and rIRBF (2.15) is a bi-convex problem, i.e. it is convex w.r.t. to blur and illumination variables individually, but it is not jointly convex. Thus, the global optimality of the solution is not guaranteed. However, since we alternately optimize over the blur kernel and illumination coefficients, we are guaranteed to converge to a local minimum. Figure 3.7 plots the average residual error of the cost function in (A.5) with increasing number of iterations. Note that the algorithm converges in a few iterations. Based on this plot, in our experiments, we terminate the algorithm after six iterations.

34
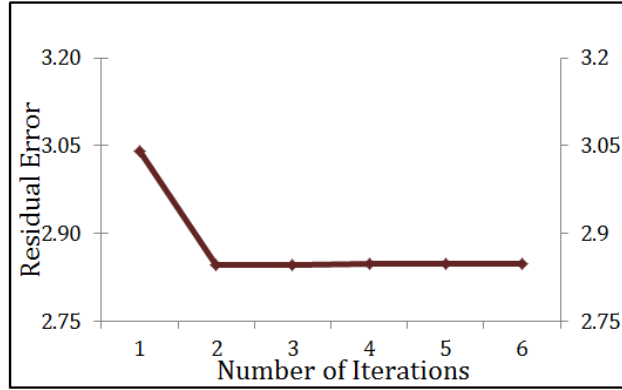
Figure 3.7: *Convergence of the IRBF algorithm*- Note that the IRBF algorithm minimizes a bi-convex function which, in general, is a non-convex problem. However, since we alternately optimize over the blur kernel and illumination coefficients, we are guaranteed to converge to a local minimum. The plot shows the average convergence behavior of the algorithm. Based on this plot, we terminate the algorithm after six iterations.

3.0.3   Recognition in Unconstrained Settings

Finally, we report recognition experiments on the REMOTE dataset where the images have been captured in an unconstrained manner [2,37]. The images were captured in two settings: from ship-to-shore and from shore-to-ship. The distance between the camera and the subjects ranges from 5meters to 250 meters. Hence, the images suffer from varying amounts of blur, variations in illumination and pose, and even some occlusion. This dataset has 17 subjects. We design our gallery set to have one frontal, sharp and well-illuminated image. The probe is manually partitioned into three categories: 1) the *illum* folder containing 561 images with illumination variations as the main problem, 2) the *blur* folder containing 75 images with blur as the main problem and 3) the *illum-blur* folder containing 128 images with both problems, see Figure 3.8(a). All three subsets

(a) *Sample Images*   (b) *Basis Images*
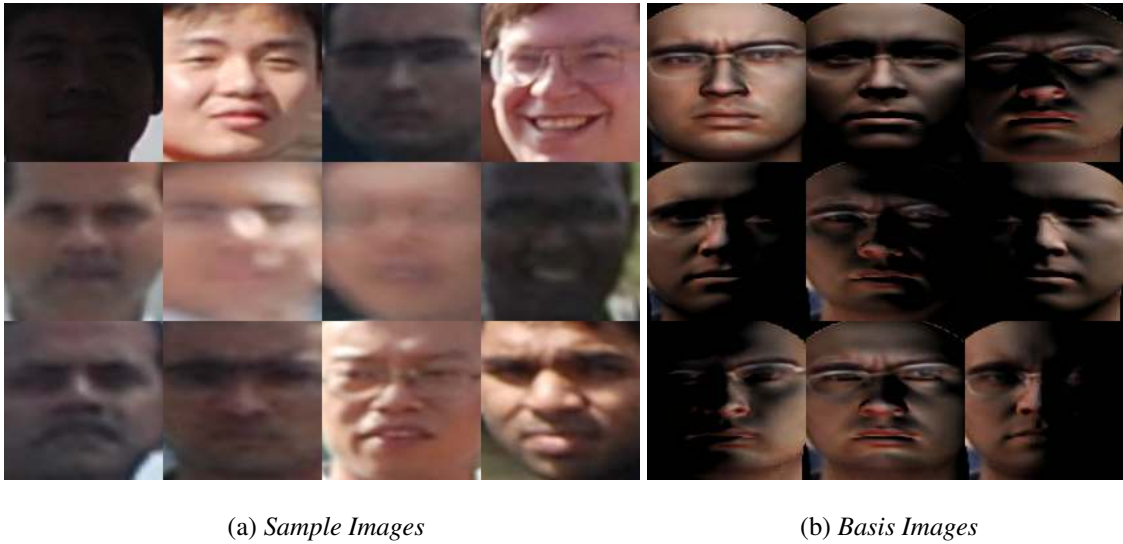
Figure 3.8: Examples from the REMOTE dataset showing three probe partitions in subplot (a). The top row shows images from the *Illum* folder where variations in illumination are the only problem. The middle and bottom rows have images which exhibit variations in blur alone (from the *Blur* folder), and variations in blur and illumination (from the *Illum-Blur* folder) respectively. Subplot (b) shows the basis images generated from a gallery.

contain near-frontal images of the 17 subjects, as set in the protocol in [37]. We register the images as a pre-processing step and normalize the size of the images to 120×120 pixels. We then run our algorithms-DRBF and IRBF, on the dataset. We assume symmetry about the origin as most of the blur arises due to out-of-focus, atmospheric and motion blur, all of which satisfy this symmetry constraint. For the *illum* folder we assume a blur kernel size of 5, and for the other two folders we assume kernel size of 7. For the IRBF and rIRBF algorithms, we generate the 9 illumination basis images for each image in the gallery, see Figure 3.8(b). We compare our algorithm with LPQ, modified FADEIN+LPQ (as described in the previous section 3.0.2). Apart from these algorithms, we also com-

pare our algorithms with the algorithms presented in [37]. These algorithms are- sparse representation based face recognition algorithm [40] (SRC), PCA+LDA+SVM [2] and a PLS-based (Partial least squares) face recognition algorithm [37]. The results are shown in Figure 3.9. The good performances by rIRBF and IRBF further confirms the importance of jointly modeling blur and illumination variations.

Figure 3.9: *Recognition results on the unconstrained dataset REMOTE*. We compare our algorithms, DRBF and IRBF, with LPQ, modified FADEIN+LPQ, a sparse representation-based face recognition algorithm [40] (SRC), PCA+LDA+SVM [2] and a PLS-based (Partial least squares) face recognition algorithm [37]. The good performance by rIRBF and IRBF further confirms the importance of jointly modeling blur and illumination variations.
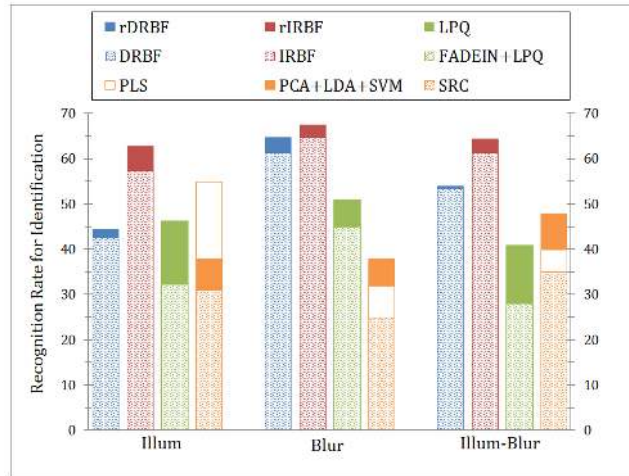
Chapter 4:   Conclusion and Discussion

Motivated by the problem of remote face recognition, we have addressed the problem of recognizing blurred and poorly-illuminated faces. We have shown that the set of all images obtained by blurring a given image is a convex set given by the convex hull of shifted versions of the image. Based on this set-theoretic characterization, we proposed a blur-robust face recognition algorithm DRBF. In this algorithm we can easily incorporate prior knowledge on the type of blur as constraints. Using the low-dimensional linear subspace model for illumination, we then showed that the set of all images obtained from a given image by blurring and changing its illumination conditions is a bi-convex set. Again, based on this set-theoretic characterization, we proposed a blur and illumination robust algorithm IRBF. We also demonstrated the efficacy of our algorithms in tackling the challenging problem of face recognition in uncontrolled settings.

Our algorithm is based on a generative model followed by nearest-neighbor classification between the query image and the gallery space, which makes it difficult to scale it to real-life datasets with millions of images. This is a common issue with most algorithms based on generative models. Broadly speaking, only classifier-based methods have been shown to scale well to very large datasets; this is because the size of the gallery largely affects the training stage, the testing stage remains relatively fast. Hence we believe that incorporating a discriminative-learning based approach like SVM into this formulation would be a very promising direction for future work. We would also like to model pose-variation under the same framework.

Chapter A: The Use of Weights in DRBF and IRBF

In Chapter 2 we address the issue of robustness of the proposed algorithm to mis-alignment and expressions. We propose three operations to that end. The first one is doing the matching in the LBP space. The second was to define the cost of matching using the L1 norm. Lastly, we incorporate weights into the formulation resulting in the following equation.

$$r_j = \min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{A_j h})||_2^2 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \mathcal{C}(\mathbf{h}) = 0.$$

The weights have a two-fold purpose. Firstly, they help reduce the importance of pixels in the low-frequency regions of the face in the kernel-estimation step. This is desirable as the effects of blur are not really perceivable in these regions. Secondly, we can give low weights to regions that are more likely to show non-rigid variability, like the ears and the mouth. With the low weights, the error in estimation due to pixels in these regions contribute less to the overall matching cost, thereby making the algorithm robust to outliers.

To train the weights, we used a method similar to the one used in [42]. We used the 'ba' and 'bj' folders of FERET as gallery and probe, respectively. We blurred the query with a Gaussian kernel of $\sigma = 4$ and partitioned the face images into patches

as shown in A.1(a). We then used DRBF/rDRBF to get the recognition rate for each patch independently. Finally the weights were assigned to each patch proportional to the recognition rate observed. A.1(b) shows the weights obtained by this method for DRBF. These weights were then used for all 3 datasets, namely FERET('fa' and 'fb'), PIE(in the kernel estimation step of IRBF) and REMOTE. The weights obtained using DRBF and the robust cost were experimentally found to be the same to 2 decimal places, and we refer to these using the common notation $\mathbb{W}$.

Hence, the weights are crucial in handling facial expressions. This is especially evident from Figure A.1(b) where we can see that the weights obtained for the outer regions of the face(hair, ears, neck etc) and the mouth are very small; as these regions are more prone to show non-rigid variability. The weights for the regions with the cheeks are also relatively small, validating our hypothesis that less textured (low-frequency) regions

(a)   Image   (b) $\mathbb{W}$

Patches
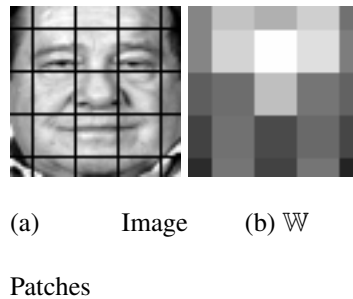
Figure A.1: The weights have been trained on the 'ba' and 'bj' folders of FERET to allow for different regions of the face (shown in A.1(a)) to contribute differently to the overall cost function. This enables us to give low weights to regions of the face that show high variability like the ears and the mouth. The trained weights are shown in Figure A.1(b), with white representing the most weight.

of the face should contribute less towards the estimation problem. Lastly, regions around the eyes are weighed the most which re-affirms the common understanding that they are the more distinguishable features of the human face.

The efficacy of the weights can also be empirically verified with the results observed in Table A.1. We can see from the table that our algorithms show a significant improvement in the performance between the non-weighted and weighted versions for the FERET and REMOTE dataset, The difference for the PIE dataset is much smaller as it has been acquired in a perfectly constrained setting, with almost no misalignment or expression variations.

Dataset	Subset	Algorithm	Experimental Setup	No Weights(%)	Weighted(%)
FERET	fa-fb	DRBF	$\sigma = 4$	86.11	92.51
PIE	GI	IRBF	$\sigma = 2$	97.42	99.26
PIE	BI	IRBF	$\sigma = 2$	85.53	85.8
REMOTE	illum	IRBF	$KS = 7 \times 7$	49.6	57.45
REMOTE	blur	DRBF	$KS = 11 \times 11$	52	61.33
REMOTE	illum-blur	IRBF	$KS = 11 \times 11$	58.7	61.40

Table A.1: *The Effect of Weights on the Recognition Rate.* We see that the weights cause a significant improvement in the performance of our DRBF/IRBF algorithms for the FERET, PIE and REMOTE datasets. For this experiment, we either choose DRBF or IRBF depending on whether illumination is a factor for the corresponding datasets. We see that weights are well-suited to handle expressions and texture-less regions of the face. However, the difference isn't large in the case of PIE, which has been acquired in a very controlled setting.

Chapter A: The Cost Formulation

In this APPENDIX we discuss how the optimzation costs formulated in Chapter 2 can be recast in the standard quadratic program form. The cost function for the blur only case DRBF is written as:

$$r_j = \min_{\mathbf{h}} ||\mathbf{i_b} - \mathbf{A_j}\mathbf{h}||_2^2 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1 \tag{A.1}$$

where $\mathbf{i_b} \in \mathbb{R}^N$ is the blurred image, $\mathbf{h} \in \mathbb{R}^K$ are the unknown blur kernel coefficients (note K could be less than the size of the blur kernel because of symmetry constraints on the blur kenel).

To see that consider the cost function:

$$
\begin{aligned}
J & \triangleq (W(\mathbf{i_b} - \mathbf{Ah}))^T(W(\mathbf{i_b} - \mathbf{Ah})) \\
& \triangleq (\mathbf{i_b} - \mathbf{Ah})^T W^T W (\mathbf{i_b} - \mathbf{Ah}) \\
& \triangleq \mathbf{i_b}^T W^T W \mathbf{i_b} + \mathbf{h}^T \mathbf{A}^T W^T W \mathbf{Ah} - 2\mathbf{i_b}^T W^T W \mathbf{Ah}
\end{aligned}
\tag{A.2}
$$

On the other hand, a quadratic program in its standard form can be written as [43]:

$$QP : \arg\min_{x} \quad \frac{1}{2}x'Hx + f^T x \quad s.t. \quad A_{in}x \leq b_{in}, A_{eq} = b_{eq} \tag{A.3}$$

Comparing equations A.2 and A.3 we get:

44

$$H \triangleq 2(\mathbf{A}^T W^T W \mathbf{A}) \quad f^T \triangleq -2\mathbf{i_b}^T W^T W \mathbf{A}$$

$$A_{in} \triangleq -I_{K \times K} \qquad b_{in} \triangleq \mathbf{0}_{K \times 1} \qquad \text{(A.4)}$$

$$A_{eq} \triangleq \mathbf{1}_{1 \times K} \qquad b_{eq} \triangleq 1$$

where $\mathbf{0}_{K \times 1}$ and $\mathbf{1}_{1 \times K}$ are the *all-zeros*, *all-ones* vectors respectively and $I_{K \times K}$ is the identity matrix. This formulation can now be solved using interior-point methods or preconditioned conjugate gradients. In the case of the *blur-plus-illumination*, the cost function is given by,

$$[\mathbf{h}, \alpha_m] = \arg \min_{\mathbf{h}, \alpha_m} \|W(\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m \mathbf{A_m h})\|_2^2$$

$$\text{subject to } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1, \mathcal{C}(\mathbf{h}) = 0. \qquad \text{(A.5)}$$

This can be solved by the process of alternation. In step I, we fix the blur $\mathbf{h}$ and optimize for α

$$
\begin{aligned}
\hat{\alpha} &= \underset{\alpha_m}{argmin} \|W\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m W \mathbf{A_m h}\|_2^2 \\
&= \underset{\alpha}{argmin} \|W\mathbf{i_b} - W\mathbf{A}^h \alpha\|_2^2 \qquad \text{(A.6)} \\
\therefore \hat{\alpha} &= ((W\mathbf{A}^h)(W\mathbf{A}^h)^T)^{-1}(W\mathbf{A}^h)^T(W\mathbf{i_b})
\end{aligned}
$$

where $\mathbf{A}^h = [\mathbf{A}_1\mathbf{h}, \cdots, \mathbf{A}_9\mathbf{h}]$. Hence in the absence of any linear constraints on α, the step involving its estimation can be analytically solved using the pseudo-inverse.

In the second step, we fix α and try to estimate $\mathbf{h}$. This step is then analogous to the *blur-only* formulation DRBF described above.

A.1 Robust Formulation

The DRBF and IRBF algorithms are sensitive to outliers and misallignment. Hence we propose robust versions of these- i.e. rDRBF and rIRBF. The cost formulation for rDRBF can be written as:

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} ||W(\mathbf{i_b} - \mathbf{Ah})||_1 \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1$$

This is a L1-norm minimization problem. We can recast this as a linear program in the following way:

Let $\lambda_i = |W(i_b - \mathbf{Ah})|_i$ i.e the ith row of the error vector. Thus $IRBF$ can be written as

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h},\lambda_i} \sum_{i=1}^{N} \lambda_i \text{ subject to } \mathbf{h} \geq 0, ||\mathbf{h}||_1 = 1, \lambda_i = |W(i_b - \mathbf{Ah})|_i$$

But $\lambda_i = |W(i_b - \mathbf{Ah})|_i \Rightarrow -\lambda \leq W(i_b - \mathbf{Ah}) \leq \lambda$. Moreover, a linear program in the standard form is written as [43]:

$$\arg\min_{x} \quad c^T x \quad \text{subject to } A_{in}x \leq b_{in}, \qquad A_{eq}x = b_{eq} \tag{A.7}$$

Hence, comparing the two forms A.7 and A.7 we get

$$c^T \triangleq [\mathbf{h}^T, \lambda^T]^T$$

$$A_{in} \triangleq \begin{bmatrix} -W\mathbf{A} & -I_{N,N} \\ W\mathbf{A} & -\mathbf{I}_{N,N} \\ -\mathbf{I}_{K,K} & \mathbf{0}_{K,N} \end{bmatrix} \quad b_{in} \triangleq \begin{bmatrix} -W\mathbf{i}_b \\ W\mathbf{i}_b \\ \mathbf{0}_{K\times 1} \end{bmatrix} \tag{A.8}$$

$$A_{eq} \triangleq [\mathbf{1}_{1,K}, \mathbf{0}_{1,N}] \qquad b_{eq} \triangleq 1$$

The robust formulation for the *blur-and-illumination* model IRBF is given by :

$$[\mathbf{h_j}, \alpha_{j,m}] = \arg \min_{\mathbf{h}, \alpha_m} \|W(\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m \mathbf{A_{j,m}h})\|_2^2 \text{ subject to } \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1$$

This too can be solved using the alternation formulation described above. Hence in STEP I, we fix the blur kernel and optimize over the illumination coeffecients:

$$
\begin{aligned}
\hat{\alpha} &= \arg \min_{\alpha} \|W\mathbf{i_b} - \sum_{m=1}^{9} \alpha_m W \mathbf{A_m h}\|_1 \\
&= \arg \min_{\alpha} \|W\mathbf{i_b} - W\mathbf{A}^h \alpha\|_1 \qquad (A.9) \\
&= \arg \min_{\alpha, \lambda_i} \sum_{i=1}^{N} \lambda_i
\end{aligned}
$$

where $\lambda_i = |W(i_b - \mathbf{Ah})|_i \Rightarrow -\lambda \leq W(i_b - \mathbf{Ah}) \leq \lambda$. Thus comparing with the standard form of the linear programming equation A.7, we get:

$$
\begin{aligned}
c^T &\triangleq [\alpha^T, \lambda^T]^T \\
A_{in} &\triangleq \begin{bmatrix} -W\mathbf{A^h} & -\mathbf{I}_{N,N} \\ W\mathbf{A^h} & -\mathbf{I}_{N,N} \end{bmatrix} \quad b_{in} \triangleq \begin{bmatrix} -W\mathbf{i}_b \\ W\mathbf{i}_b \end{bmatrix} \qquad (A.10)
\end{aligned}
$$

In the second step, we fix the illumination coeffecients α and solve for $\mathbf{h}$.

Bibliography

[1] Anat Levin, Yair Weiss, Frédo Durand, and William T. Freeman. Understanding blind deconvolution algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(12):2354–2367, 2011. iv, 2, 5, 25, 26

[2] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Yi Ma. Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2):210 –227, feb. 2009. vi, 22, 37

[3] Jie Ni and Rama Chellappa. Evaluation of state-of-the-art algorithms for remote face recognition. In *ICIP*, pages 1581–1584, 2010. vi, 2, 25, 35, 37

[4] Rama Chellappa, Jie Ni, and Vishal M. Patel. Remote identification of faces: Problems, prospects, and progress. *Pattern Recognition Letters*, 2011. vi, 16, 35, 36, 37

[5] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, dec 2003. 2

[6] Masashi Nishiyama, Abdenour Hadid, Hidenori Takeshima, Jamie Shotton, Tatsuo Kozakaya, and Osamu Yamaguchi. Facial deblur inference using subspace analysis for recognition of blurred faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):838–845, 2011. 2, 5, 25, 26

[7] D. Kundur and D. Hatzinakos. Blind image deconvolution revisited. 2, 5

[8] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, 2002. 3, 14

[9] Ronen Basri and David W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):218–233, feb 2003. 3, 6, 15

[10] Ravi Ramamoorthi and Pat Hanrahan. A signal-processing framework for reflection. *ACM Trans. Graph.*, 23(4):1004–1042, 2004. 3, 6, 15

[11] Kuang-Chih Lee, Jeffrey Ho, and David J. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:684–698, 2005. 3, 6, 16

[12] P. Vageeswaran, K. Mitra, and R. Chellappa. Blur and illumination robust face recognition via set-theoretic characterization. *IEEE Trans. Image Processing*, 22(4):1362–1372, 2013. 5

[13] Hao Hu and Gerard De Haan. Adaptive image restoration based on local robust blur estimation. In *Proceedings of the 9th international conference on Advanced concepts for intelligent vision systems*, ACIVS'07, pages 461–472, 2007. 5

[14] William Hadley Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.*, 62(1):55–59, Jan 1972. 5

[15] Anat Levin. Blind motion deblurring using image statistics. In *NIPS*, pages 841–848, 2006. 5

[16] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 787–794, 2006. 5

[17] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä. Recognition of blurred faces using local phase quantization. In *Proc. 19th International Conference on Pattern Recognition (ICPR 2008), Tampa, FL, 4 p*, 2008. 5, 14

[18] Raghuraman Gopalan, Sima Taheri, Pavan K. Turaga, and Rama Chellappa. A blur-robust descriptor with applications to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(6):1220–1226, 2012. 5

[19] Ville Ojansivu and Janne Heikkilä. Blur insensitive texture classification using local phase quantization. In *Proceedings of the 3rd international conference on Image and Signal Processing*, ICISP '08, pages 236–243, 2008. 5, 25

[20] Inna Stainvas and Nathan Intrator. Blurred face recognition via a hybrid network architecture. In *In Proc. ICPR*, pages 809–812. 5

[21] Haichao Zhang, Jianchao Yang, Yanning Zhang, Nasser M. Nasrabadi, and Thomas S. Huang. Close the loop: Joint blind image restoration and recognition with sparse representation prior. *Computer Vision, IEEE International Conference on*, 0:770–777, 2011. 6

[22] PL Combettes. The convex feasibility problem in image recovery. *Advances in Imaging and Electron Physics*, 95:155–270, 1996. 6

[23] H. Trussell and M. Civanlar. The feasible solution in signal restoration. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 32(2):201 – 212, apr 1984. 6

[24] P.L. Combettes and J.C. Pesquet. Image restoration subject to a total variation constraint. *IEEE Trans. Image Process.*, 13(9):1213 –1222, sept. 2004. 6

[25] Haitao Wang, S.Z. Li, Yangsheng Wang, and Jianjun Zhang. Self quotient image for face recognition. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 2, pages 1397 – 1400 Vol.2, oct. 2004. 6, 34

[26] B.V.K.V. Kumar, M. Savvides, and Chunyan Xie. Correlation pattern recognition for face recognition. *Proceedings of the IEEE*, 94(11):1963 –1976, nov. 2006. 6

[27] Marios Savvides, B. V. K. Vijaya Kumar, and P. K. Khosla. "eigenphases vs. eigenfaces". *Pattern Recognition, International Conference on*, 3:810–813, 2004. 6

[28] Ralph Gross and Vladimir Brajovic. An image preprocessing algorithm for illumination invariant face recognition. In *Proceedings of the 4th international conference on Audio- and video-based biometric person authentication*, AVBPA'03, pages 10–18, 2003. 6

[29] H.F. Chen, P.N. Belhumeur, and D.W. Jacobs. In search of illumination invariants. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 254–261. IEEE, 2000. 6

[30] Margarita Osadchy, David W. Jacobs, and Michael Lindenbaum. Surface dependent representations for illumination insensitive image comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:98–111, 2007. 6

[31] Soma Biswas, Gaurav Aggarwal, and Rama Chellappa. Robust estimation of albedo for illumination-invariant matching and shape recovery. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(5):884–899, 2009. 6, 17

[32] Reginald L. Lagendijk and Jan Biemond. Basic methods for image restoration and identification. In Alan C. Bovic, editor, *The Essential Guide to Image Processing*, pages 323 – 348. Academic Press, Boston, second edition, 2009. 9, 12, 25

[33] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell. Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, 94(11):1948 –1962, nov. 2006. 14

[34] R. Epstein, P.W. Hallinan, and A.L. Yuille. 5 plus or minus 2 eigenimages suffice: An empirical investigation of low-dimensional lighting models. In *Physics-Based Modeling in Computer Vision, 1995., Proceedings of the Workshop on*, page 108, jun 1995. 16

[35] Volker Blanz and Thomas Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(9):1063–1074, 2003. 16

[36] P. Jonathon Phillips, Patrick J. Flynn, W. Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William J. Worek. Overview of the face recognition grand challenge. In *CVPR (1)*, pages 947–954, 2005. 16

[37] 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*, 22:1090–1104, 2000. 16, 25

[38] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Kriegman. From few to many: Generative models for recognition under variable pose and illumination. In *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*, 2000. 16

[39] W.A.P. Smith and E.R. Hancock. Estimating the albedo map of a face from a single image. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III – 780–3, sept. 2005. 17

[40] Miroslav Hamouz Xuan Zou, Josef Kittler and Jose R. Tena. Robust albedo estimation from face image under unknown illumination. In *SPIE, Proceedings of*, page 69440, 2008. 17

[41] Terence Sim, Simon Baker, and Maan Bsat. The cmu pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1615 – 1618, December 2003. 25

[42] Timo Ahonen, Abdenour Hadid, and Matti Pietik?inen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:2037–2041, 2006. 40

[43] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, March 1996. 44, 46