International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Image Recognition of Sparsity, Non-Uniform Motion Blur, Illumination, and Pose

Darja Usharani, Bullarao.D
Department of Electronics and Communication Engineering,Swetha Institute of Technology & Sciecne, JNTUA, Tirupati , India.

***Abstract*—Existing methods for performing face recognition in the presence of blur are based on the convolution model and cannot handle non-uniform blurring situations that frequently arise from tilts and rotations in hand-held cameras. In this paper, we propose a methodology for face recognition in the presence of space-varying motion blur comprising of arbitrarily-shaped kernels. We model the blurred face as a convex combination of geometrically transformed instances of the focused gallery face, and show that the set of all images obtained by non-uniformly blurring a given image forms a convex set. We first propose a nonuniform blur-robust algorithm by making use of the assumption of a sparse camera trajectory in the camera motion space to build an energy function with l_1-norm constraint on the camera motion. The framework is then extended to handle illumination variations by exploiting the fact that the set of all images obtained from a face image by non-uniform blurring and changing the illumination forms a bi-convex set. Finally, we propose an elegant extension to also account for variations in pose.**

***Index Terms*—Face recognition, non-uniform blur, sparsity, illumination, pose.**

I. INTRODUCTION

IT IS well-known that the accuracy of face recognition systems deteriorates quite rapidly in unconstrained settings [1]. This can be attributed to degradations arising from blur, changes in illumination, pose, and expression, partial occlusions etc. Motion blur, in particular, deserves special attention owing to the ubiquity of mobile phones and hand-held imaging devices. Dealing with camera shake is a very relevant problem because, while tripods hinder mobility, reducing the exposure time affects image quality. Moreover, in-built sensors such as gyros and accelerometers have their own limitations in sensing the camera motion. In an uncontrolled environment, illumination and pose could also vary, further compounding the problem. The focus of this paper is on developing a system that can recognize faces across non-uniform (i.e., space-variant) blur, and varying illumination and pose.

Traditionally, blurring due to camera shake has been modeled as a convolution with a single blur kernel, and the blur is assumed to be uniform across the image [2], [3]. However, it is space-variant blur that is encountered frequently in hand-held cameras [4]. While techniques have been proposed that address the restoration of non-uniform blur by local space-invariance approximation [5]–[7], recent methods for image restoration have modeled the motion-blurred image as an average of projectively transformed images [8]–[12].

Face recognition systems that work with focused images have difficulty when presented with blurred data. Approaches to face recognition from blurred images can be broadly classified into four categories. (i) Deblurring-based [13], [14] in which the probe image is first deblurred and then used for recognition. However, deblurring artifacts are a major source of error especially for moderate to heavy blurs. (ii) Joint deblurring and recognition [15], the flip-side of which is computational complexity. (iii) Deriving blur-invariant features for recognition [16], [17]. But these are effective only for mild blurs. (iv) The *direct* recognition approach of [18] and [19] in which reblurred versions from the gallery are compared with the blurred probe image. It is important to note that all of the above approaches assume a simplistic space-invariant blur model. For handling illumination, there have mainly been two directions of pursuit based on (i) the 9D subspace model for face [20] and (ii) extracting and matching illumination insensitive facial features [21], [22]. Tan et al. [23] combine the strengths of the above two methods and propose an integrated framework that includes an initial illumination normalization step for face recognition under difficult lighting conditions. A subspace learning approach using image gradient orientations for illumination and occlusion-robust face recognition has been proposed in [24]. Practical face recognition algorithms must also possess the ability to recognize faces across reasonable variations in pose. Methods for face recognition across pose can broadly be classified into 2D and 3D techniques. A good survey article on this issue can be found in [25].

Although the problem of blur, illumination and pose are individually quite challenging and merit research in their own right, a few attempts have been made in the literature to

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

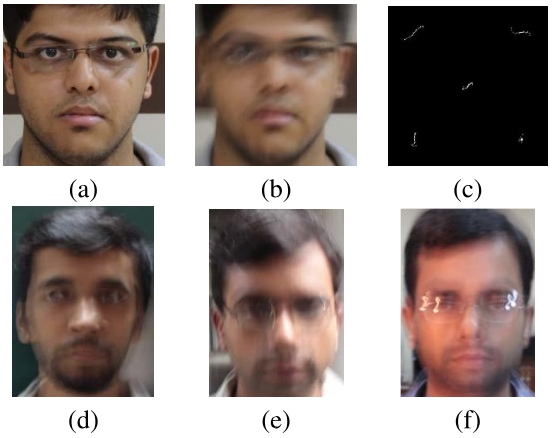


Fig. 1. (a) Focused image, (b) synthetically blurred image obtained by applying random in-plane translations and rotations on the focused image, (c) point spread functions (PSF) at various locations in the image showing the presence of non-uniform blur which cannot be explained by the convolution model (best viewed as PDF), and (d, e, f) real blurred images from the dataset we ourselves captured using a hand-held camera.

jointly tackle some of these issues under one framework. Patel et al. [26] have proposed a dictionary-based approach to recognizing faces across illumination and pose. A sparse minimization technique for recognizing faces across illumination and occlusion has been proposed in [27], while [28], which is based on similar principles, additionally offers robustness to alignment and pose. But these works do not deal with blurred images. A very recent work [19] formally addresses the problem of recognizing faces from distant cameras across both blur and illumination wherein the observed blur can be well-approximated by the convolution model. To the best of our knowledge, the only attempt in the literature at recognizing faces across non-uniform blur has been made in [17] in which the uniform blur model is applied on overlapping patches to perform recognition on the basis of a majority vote. However, they do not explicitly model illumination changes going from gallery to probe. We would like to point out that both [17] and [19] limit their discussion to frontal faces.

In this paper, we propose a face recognition algorithm that is robust to non-uniform (i.e., space-varying) motion blur arising from relative motion between the camera and the subject. Following [19], we assume that only a *single* gallery image is available. The camera transformations can range from in-plane translations and rotations to out-of-plane translations, out-ofplane rotations, and even general 6D motion. An example is shown in Fig. 1. Observe that the blur on the faces can be significantly non-uniform. The simple yet restrictive convolution model fails to explain this blur and a spacevarying formulation becomes necessary. Subsequently, we also show how the proposed method can be elegantly modified to account for variations in illumination and pose.

We assume a planar structure for the face [14], [17], [19] and use the geometric framework proposed in [8]–[10], [29] to model the blurred face as the weighted average of geometrically warped instances (homographies) of the focused gallery image. The warped instances can be viewed as the intermediate images observed during the exposure time. Each warp is assigned a weight that denotes the fraction of the exposure duration for that transformation. The weights corresponding to the warps are referred to as the transformation spread function (TSF) [29] in the literature.

We develop our basic non-uniform motion blur (NU-MOB)-robust face recognition algorithm based on the TSF model. On each focused gallery image, we apply all the possible transformations that exist in the 6D space (3 dimensions for translations and 3 for rotations) and stack the resulting transformed images as columns of a matrix. We extend the convexity result proved for the simple convolution model in [19] to the TSF model and show that the set of all images obtained by blurring a particular gallery image is a convex set given by the convex hull of the columns of the corresponding matrix. To recognize a blurred probe image, we minimize the distance between the probe and the convex combination of the columns of the transformation matrix corresponding to each gallery image. The gallery image whose distance to the probe is minimum is identified as a match. We do not impose any constraints on the nature of the blur. Following [9], [11], we assume that the camera motion trajectory is sparse in the camera motion space. This allows us to construct an optimization function with l_1-norm constraint on the TSF weights. Minimizing this cost function gives us an estimate of the transformations that when applied on the gallery image results in the blurred probe image. Each gallery image, blurred using the corresponding optimal TSF, is compared with the probe in the LBP (local binary pattern) [30] space. This direct method of recognition allows us to circumvent the challenging and ill-posed problem of single image blind-deblurring. The idea of reblurring followed by LBP-based recognition has been suggested in [19], and LBP histograms have been shown to work well on blurred faces too. We have extended the formulation in [19] to space-varying situations.

Furthermore, we propose extensions to the basic framework to handle variations in illumination as well as pose. We approximate the face to a convex Lambertian surface, and use the 9D subspace model in [20] and the bi-convexity property of a face under blur and illumination variations in the context of the TSF model. Our motion blur and illumination (MOBIL)-robust face recognition algorithm uses an alternating minimization (AM) scheme wherein we solve for the TSF weights in the first step and use the estimated TSF to solve for the nine illumination coefficients in the second, and go on iterating till convergence. We finally transform (reblur

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

and relight) each gallery image and compare it with the probe in the LBP space. Using a rough initial estimate of the pose to synthesize gallery images in the new pose, we extend this formulation and propose an algorithm to handle motion blur, illumination and pose (MOBILAP) for non-frontal faces. The new synthesized gallery image is reblurred and relit as before, and compared with the probe using LBP.

Differences With [19]: The DRBF and IRBF methods proposed in [19] are restricted to the simplistic convolution blur model which is valid only when the motion of the camera is limited to in-plane translations. This assumption of uniform blur does not hold true in real settings because camera tilts and rotations occur frequently in the case of hand-held cameras [11]. The algorithms proposed in this paper, in contrast, are capable of handling any general motion of the camera which sets our work distinctly apart from [19]. In addition, we handle pose variations while the discussion in [19] is restricted to frontal faces. Our method allows for arbitrarily-shaped space-varying kernels across the image unlike [19] which seeks to explain the blur using a *single* PSF for the entire image. In fact, our scheme based on the TSF model subsumes the work in [19] - for the special case of only in-plane translational motion, the TSF reduces to a PSF.

The work proposed in this paper advances the state-of-theart in many ways as discussed next.

- This is the first attempt to *systematically* address face recognition under (i) non-uniform motion blur and (ii) the combined effects of blur, illumination and pose.
- We prove that the set of all images obtained by non-uniformly blurring a given image forms a convex set. We also show that the set of all images obtained from a face image by non-uniform blurring and change of illumination forms a bi-convex set.
- We extend our method to non-frontal situations by transforming the gallery to a new pose.
- We propose a multi-scale implementation that is efficient both in terms of computation as well as memory usage.
- We demonstrate superior performance over contemporary methods on standard face databases (FERET, PIE, Extended Yale B) in the presence of blur, illumination and pose variations, as well as a real dataset which contains, in addition to these degradations, small occlusions and expression changes.

The organization of the rest of the paper is as follows: We review the convolution model for uniform blur in Section II and discuss its shortcomings. In Section III, we develop the non-uniform motion blur model for faces and propose an elegant and efficient scheme for face recognition under spacevarying motion blur. Experimental results are given for purpose of validation. In Section IV, we incorporate illumination and pose into the basic formulation discussed in Section III. Section V contains results and comparisons on synthetic and real examples involving blur, illumination and pose as well. Section VI concludes the paper.

II. Convolution Model for Space-Invariant Blur

As discussed in the introduction, while the convolution model is sufficient for describing blur due to in-plane camera translations, a major limitation is that it cannot describe several other blurring effects (including out-of-plane motion and in-plane rotation) arising from general camera motion. In order to demonstrate the weakness of the convolution model in handling images blurred due to camera shake, we synthetically blur the focused gallery image to generate a probe, and provide both the gallery image and the blurred probe image as input to two algorithms- the convolution model which assumes spaceinvariant blur, and the non-uniform motion blur model (to be discussed in Section III) which represents the space-variant blurred image as a weighted average of *geometrically warped* instances of the gallery. Next, we compare the reconstruction

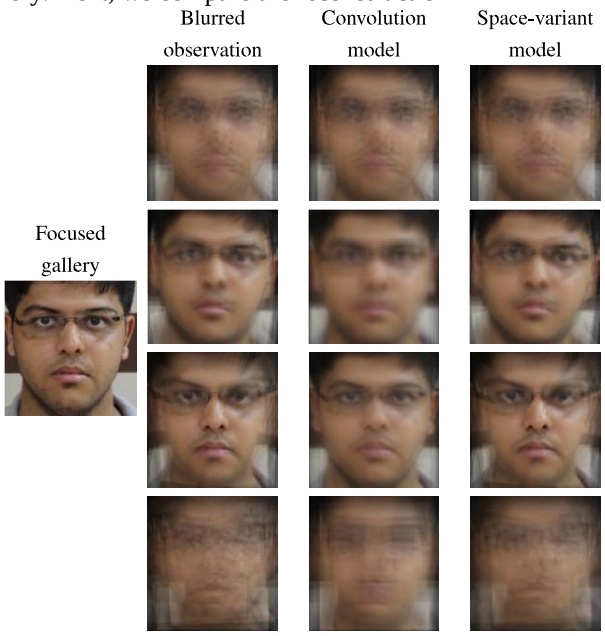


Fig. 2. Comparing the reconstructed faces and the reconstruction errors (RMS) using the convolution model and the space-variant model. The gallery image is shown in column 1. The results are shown in row 1 for space-invariant blur and rows 2 through 4 for space-variant blur. The RMS errors for the convolution model in column 3 and the space-variant model in column 4, respectively, are- row 1: 0.04, 0.04, row 2: 12.48, 0.05, row 3: 14.88, 0.15, and row 4: 15.31, 5.15. (The blurred image in column 2 was used as the reference for RMS computation.)

errors between the probe and the gallery reblurred using the camera motion estimated by both the methods. This experiment is performed for different camera motions as shown in Fig. 2 - row 1: in-plane translations, row 2: in-plane

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

translations and rotations, row 3: out-of-plane rotations, and row 4: full 6D blur. The reconstructed faces as well as the RMS errors are also shown in Fig. 2. Note that there is no change in illumination or pose between the gallery and the probe, and only the blur has to be accounted for. Observe that except for in-plane translations (row 1), where, as expected, the RMS is the same for both the models, in all the other cases, the space-variant motion blur model gives significantly smaller RMS error than the convolution model. Note that the RMS value is smaller than one, except for 6D motion (row 4) for which it is marginally higher as our algorithm needs to search through a very large set of transformations.

III. Motion Blur Model for Faces

The apparent motion of scene points in the image will vary at different locations when the camera motion is not restricted to in-plane translations. In such a scenario, the space-varying blur across the image cannot be explained using the convolution model and with a single blur kernel. In this section, we present the space-variant motion blur model [8]–[10], [29] and illustrate how this model can explain geometric degradations of faces resulting from general camera motion. Later, we propose an optimization algorithm to recover the camera motion.

Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ denote the focused gallery face captured by a still camera. Assume the origin to be the camera center and let $\mathbf{X} = [X\ Y\ Z]^T$ denote the spatial coordinates of a point on the face. Let the corresponding image coordinates be $x = \frac{vX}{Z}$ and $y = \frac{vY}{Z}$, where v denotes the focal length of the camera. The projection of $\mathbf{X}$ on the image plane can be represented as $\mathbf{x} = \mathbf{K_v X}$, where $\mathbf{K_v} = \text{diag}(v,v,1)$. To get the image coordinates (x, y), the standard practice is to express $\mathbf{x}$ in homogeneous form i.e., scale $\mathbf{x}$ by its third element. At each instant of time τ during exposure, the coordinates of the 3D point $\mathbf{X}$ changes to $\mathbf{X}_\tau = \mathbf{R}_\tau \mathbf{X} + \mathbf{T}_\tau$ due to relative motion between the camera and the subject. Here, $\mathbf{T}_\tau = [T_{X\tau}\ T_{Y\tau}\ T_{Z\tau}]^T$ is the translation vector, and $\mathbf{R}_\tau$ represents the rotation matrix parameterized in terms of the angles of rotation ϑ_X, ϑ_Y and ϑ_Z about the three axes using the matrix exponential

$$\mathbf{R}_\tau = e^{\mathbf{A}_\tau} \text{ where } \mathbf{A}_\tau = \begin{bmatrix} 0 & -\vartheta_{Z_\tau} & \vartheta_{Y_\tau} \\ \vartheta_{Z_\tau} & 0 & -\vartheta_{X_\tau} \\ -\vartheta_{Y_\tau} & \vartheta_{X_\tau} & 0 \end{bmatrix} \quad (1)$$

Following prior works in face recognition, [14], [17], [19], we too model the face by a flat surface i.e., all the points are at a distance d_o from the camera. Therefore, the depth is constant, and the point $\mathbf{x}_\tau$, at which $\mathbf{X}_\tau$ gets projected in the camera, can be obtained through a homography $\mathbf{H}_\tau$ as $\mathbf{x}_\tau = \mathbf{H}_\tau \mathbf{x}$ where

$$\mathbf{H}_\tau = \mathbf{K_v} \left(\mathbf{R}_\tau + \frac{1}{d_o} \begin{bmatrix} & & \\ & \mathbf{T}_\tau & \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \right) \mathbf{K_v}^{-1} \quad (2)$$

If g_τ denotes the transformed image (due to relative motion) captured at time instant τ, then we can write $g_\tau(\mathbf{H}_\tau \mathbf{x}) = f(\mathbf{x})$, or alternately, $g_\tau(\mathbf{x}) = f(\mathbf{H}_\tau^{-1}\mathbf{x})$ where $\mathbf{H}_\tau^{-1}$ denotes the inverse of $\mathbf{H}_\tau$. The arguments of f in $f(\mathbf{H}_\tau^{-1}\mathbf{x})$, which are the image coordinates, correspond to the first two elements of $\mathbf{H}_\tau^{-1}\mathbf{x}$ (a 3×1 vector) expressed in homogeneous form. We follow this convention throughout the paper. Now the blurred face g can be interpreted as the average of transformed versions of f during exposure. Therefore, the intensity at an image point $\mathbf{x}$ on the blurred face is given by

$$g(\mathbf{x}) = \frac{1}{T_e} \int_0^{T_e} f(\mathbf{H}_\tau^{-1}\mathbf{x}) d\tau \quad (3)$$

where T_e is the total exposure duration.

The blurred face can be more appropriately modeled in terms of the gallery face by averaging it over the set of possible transformations resulting from the relative motion between the camera and the subject. Let $\mathbf{T}$ denote this set of all possible transformations. Let $h_T : \mathbf{T} \rightarrow \mathbb{R}^+$, called the *transformation spread function* (TSF), denote a mapping from $\mathbf{T}$ to non-negative real numbers. The value of the TSF, $h_T(\lambda)$, for each transformation $\lambda \in \mathbf{T}$, denotes the fraction of the total exposure duration for which the camera stayed in the position that caused the transformation $\mathbf{H}_\lambda^{-1}$ on the image coordinates. Hence, $\sum_{\lambda \in \mathbf{T}} h_T(\lambda) = 1$. Note that the term λ denotes the transformation parameters corresponding to the homography matrix $\mathbf{H}_\lambda^{-1}$. The blurred image can be equivalently written as

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

an average of warped versions of f weighted by the TSF, h_T, i.e.,

$$g(\mathbf{x}) = \int_{\lambda \in \mathbf{T}} h_T(\lambda)\, f(\mathbf{H}_\lambda^{-1}\mathbf{x})\, d\lambda \qquad (4)$$

Observe that a single TSF using (4) is sufficient to describe the observed space-varying blur. When the motion is confined to 2D translations, the TSF will have non-zero weights only for the in-plane translational components and will be identical to the PSF i.e., the convolution model for the blur is a special case of the space-varying motion blur model.

In practice, the TSF is defined on the discrete transformation space **T** and can be considered as a vector in $\mathbb{R}^{N_T}$ where N_T is the total number of transformations present in **T**. N_T is controlled by the number of translation steps along each axis as well as the number of rotation steps about each axis. Hence,

$hN_T(\lambda=\mathbf{k})N_T \geq_x \times 0$ $N$for$T_y \times k$ $N=T_z$ $1\times,2N,\ldots,\vartheta_x \times$ $NN\vartheta_{Ty}\times$. In the discrete domain,$N\vartheta_z$, $k^{N=T}{}_1\, h_T(\lambda\mathbf{k}) = 1$,

(4) can be written as

$$g(r,c) = \sum_{k=1}^{N_T} h_T(\lambda_\mathbf{k})\, f(\mathbf{H}_\lambda^{-1}{}_\mathbf{k}[r\ c\ 1]^T)) \qquad (5)$$

where $g(r,c)$ and $f(r,c)$ represent the pixel intensity at (r,c) for the discretized blurred image and latent image, respectively. If **g**, **f** represent the blurred image and the latent image, respectively, lexicographically ordered as vectors, then (5) can be expressed in matrix-vector notation as

$$\mathbf{g} = \mathbf{A}\mathbf{h}_\mathbf{T} \text{ such that } \mathbf{h}_\mathbf{T} \geq 0, ||\mathbf{h}_\mathbf{T}||_1 = 1. \qquad (6)$$

where $\mathbf{A} \in \mathbb{R}^{N \times N_T}$ is the matrix whose N_T columns contain warped copies of **f**, $\mathbf{h}_\mathbf{T}$ denotes the vector of weights $h_T(\lambda)$, and N is the total number of pixels in the image. The warped versions of **f** are obtained by applying the homography matrix $\mathbf{H}_\lambda^{-1}$ corresponding to each of the N_T transformations on the gallery image. From (6), the set of all blurred images obtained from **f** can be written as

$$\mathtt{B} \{\mathbf{A}\mathbf{h}_\mathbf{T} | \mathbf{h}_\mathbf{T} \geq 0, ||\mathbf{h}_\mathbf{T}||_1 = 1\} \qquad (7)$$

Proposition 1: The set of all images B *obtained by blurring an image* **f** *using the TSF model is a convex set. Moreover, this convex set is given by the convex hull of the columns of the matrix* **A**, *where the columns of* **A** *are warped versions of* **f** *as determined by the TSF.*

Proof: Let $\mathbf{g}_1$ and $\mathbf{g}_2$ be elements from the set B. Then there exist TSFs $\mathbf{h}_{T1}$ and $\mathbf{h}_{T2}$, both satisfying the conditions $\mathbf{h}_{Ti} \geq 0$ and $||\mathbf{h}_{Ti}||_1 = 1$, $i = 1,2$ such that $\mathbf{g}_1 = \mathbf{A}\mathbf{h}_{T1}$ and $\mathbf{g}_2 = \mathbf{A}\mathbf{h}_{T2}$. To show that the set B is convex, we need to show that for any γ satisfying $0 \leq \gamma \leq 1$, $\mathbf{g}_3 = \gamma\mathbf{g}_1 + (1-\gamma)\mathbf{g}_2$ is an element of B. Now

$$\begin{aligned}\mathbf{g}_3 &= \gamma\mathbf{g}_1 + (1-\gamma)\mathbf{g}_2 \\ &= \gamma\mathbf{A}\mathbf{h}_{T1} + (1-\gamma)\mathbf{A}\mathbf{h}_{T2} \\ &= \mathbf{A}(\gamma\mathbf{h}_{T1} + (1-\gamma)\mathbf{h}_{T2}) \\ &= \mathbf{A}\mathbf{h}_{T3} \end{aligned} \qquad (8)$$

Here $\mathbf{h}_{T3}$ is non-negative and sums to one, and hence $\mathbf{g}_3$ is an element of B. Thus, B is a convex set defined as

$$\{\mathbf{A}\mathbf{h}_\mathbf{T} | \mathbf{h}_\mathbf{T} \geq 0, ||\mathbf{h}_\mathbf{T}||_1 = 1\}, \qquad (9)$$

which, by definition, is the convex hull of the columns of **A**.

Note that $\mathbf{h}_\mathbf{T}$ is sparse since for motion blur only a fraction of the total poses N_T will have non-zero weights [29]. We make use of this fact to build the following energy function

$$E(\mathbf{h}_\mathbf{T}) = ||\mathbf{g} - \mathbf{A}\mathbf{h}_\mathbf{T}||_2^2 + \beta||\mathbf{h}_\mathbf{T}||_1 \quad \text{subject to } \mathbf{h}_\mathbf{T} \geq 0. \qquad (10)$$

The optimization problem in (10) can be solved using the *nnLeastR* function of the Lasso algorithm [31] which considers the additional l_1-norm and non-negativity constraints. This energy function when minimized provides an estimate of the transformations that must be applied to the gallery image to produce the blurred image.

A. Multiscale Implementation

Since we are fundamentally limited by the resolution of the images, having a very fine discretization of the transformation space **T** leads to redundant computations. Hence, in practice, the discretization is performed in a manner that the difference in the displacements of a point light source due to two different transformations from the discrete set **T** is at least one pixel. It should be noted that since the TSF is defined over 6 dimensions, doubling their sampling resolution increases the total number of poses, N_T, by a factor of 2^6. As the number of transformations in the space **T** increases, the optimization

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

process becomes inefficient and time consuming, especially since only a few of these elements have non-zero values. Moreover, the resulting matrix **A** will have too many columns to handle. Following [9], we resort to a multiscale framework to solve this problem. We perform multiscaling in 6D (instead of 3D as in [9]). We select the search intervals along each dimension according to the extent of the blur we need to model, which is typically a few pixels for translation and a few degrees for rotation.

The idea is to start from a coarse representation of the image and the TSF, and repeatedly refine the estimated TSF at higher resolutions. Downsampling a blurred image by a certain factor reduces the amount of pixel displacements due to camera translation along *X* and *Y* axes by the same factor, and if the focal length of the camera is large enough, it has the same effect on the pixel displacements due to camera rotation about*X* and *Y* axes. Hence, downsampling the images also reduces the space of allowed transformations.[1]

We first build Gaussian pyramids for both the focused and blurred images. At the coarsest scale, the matrix **A** is built for the whole transformation space **T**. But it is to be noted that the search intervals for the TSF are reduced depending on the downsampling factor. The TSF $\mathbf{h_T}$ is estimated by minimizing equation (10). We then upsample $\mathbf{h_T}$ to the next scale using bilinear interpolation, and find the non-zero elements of this upsampled and interpolated TSF. Also, using a suitably chosen threshold, we remove insignificant values resulting from the interpolation process. This gives us several 6D nonzero regions inside the transformation space. When finding the optimal $\mathbf{h_T}$ at the next scale, we only search for valid homographies which lie within these non-zero regions. This corresponds to discarding many columns of **A**, reducing both the computation and memory demands of the search process. We repeat this procedure at each scale, until the optimal TSF at the finest resolution is found. The improvement in speed that accrues is discussed in Section III-C.

B. Face Recognition Across Blur

Suppose we have *M* face classes with one focused gallery face $\mathbf{f_m}$ for each class *m*, where $m = 1,2,\ldots, M$. Let us denote the blurred probe image which belongs to one of the *M* classes by **g**. Given $\mathbf{f_m}$s and **g**, the task is to find the identity $m^* \in \{1,2,\ldots, M\}$ of **g**. Based on the discussions in

Section III, the first step is to generate the matrix $\mathbf{A_m}$ for each gallery face. Then, since **g** belongs to one of the *M* classes, it can be expressed as the convex combination of the columns of one of these matrices. Therefore, the identity of the probe image can be found by minimizing the projection error of **g** onto $\{\mathbf{A_m}\}$s. The reconstruction error d_m can be obtained by solving

$$d_m = \min_{\mathbf{h_T}} ||\mathbf{g} - \mathbf{A_m h_T}||_2^2 + \beta ||\mathbf{h_T}||_1 \quad \text{subject to } \mathbf{h_T} \geq 0. \tag{11}$$

One could compute d_m for each $m = 1,2,\ldots, M$ and assign **g** the identity of the gallery image with the minimum d_m. Note that in (11), all the pixels receive equal weight and influence the TSF estimation step equally. But not all regions in the face convey the same amount of information. Following [19], we modify the above equation by introducing a weighting matrix **W** (which weighs different regions in the face differently) when computing the reconstruction error between the probe image and the gallery images. Equation (11) then becomes

$$d_m = \min_{\mathbf{h_T}} ||\mathbf{W}(\mathbf{g} - \mathbf{A_m h_T})||_2^2 + \beta ||\mathbf{h_T}||_1 \quad \text{subject to } \mathbf{h_T} \geq 0 \tag{12}$$

where **W** (a diagonal matrix) is learned following the procedure outlined in the appendix of [19]. This matrix has the highest weights for regions around the eyes and de-emphazises the mouth and cheeks.

It must be mentioned that the quantity d_m is not preferable as a metric for face recognition because of its sensitivity to even small pixel misalignments. Instead, we use Local Binary Patterns (LBP) [30], which are reasonably robust to alignment errors, for the recognition task. For this purpose, we first compute the optimal TSF $\mathbf{h_{Tm}}$ for each gallery image by solving (12), i.e.,

$$\mathbf{h_{Tm}} = \operatorname*{argmin}_{\mathbf{h_T}} ||\mathbf{W}(\mathbf{g} - \mathbf{A_m h_T})||_2^2 + \beta ||\mathbf{h_T}||_1 \quad \text{subject to } \mathbf{h_T} \geq 0. \tag{13}$$

Next, we blur each of the gallery images with the corresponding optimal TSFs $\mathbf{h_{Tm}}$. For each blurred gallery image and probe, we divide the face into non-overlapping rectangular patches (details of the patch sizes can be found in [19]), extract LBP histograms independently from each patch and

[1] Translation along and rotation about the *Z*-axis remain unchanged after downsampling the image.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

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

Fig. 3. Sample images from *ba* and *bj* folders in the FERET database. (a) Gallery, (b) probe, (c)-(g) probe blurred synthetically using random transformations from the TSF intervals listed in Setting 1 - Setting 5 of Section III-C.

Algorithm 1 NU-MOB: Non-Uniform Motion Blur-Robust Face Recognition

Input: Blurred probe image $\mathbf{g}$ and a set of gallery images $\mathbf{f_m}, m = 1, 2, ..., M$.

Output: Identity of the probe image.

1: For each gallery image $\mathbf{f_m}$, find the optimal TSF $\mathbf{h_{T_m}}$ by solving equation (13).
2: Blur each gallery image $\mathbf{f_m}$ with its corresponding $\mathbf{h_{T_m}}$ and extract LBP features.
3: Compare the LBP features of the probe image $\mathbf{g}$ with those of the transformed gallery images and find the closest match.

concatenate the histograms to build a global descriptor. The intuition behind dividing the image into blocks is that the face can be seen as a composition of micro-patterns, and the textures of the facial regions are locally encoded by the LBP patterns while the whole shape of the face is recovered by the construction of the global histogram i.e., the spatially enhanced global histogram encodes both the appearance and the spatial relations of facial regions. We then perform recognition with a nearest neighbour classifier using Chi square distance [16] with the obtained histograms as feature vectors. The steps are outlined in Algorithm 1. An alternative approach would be to use the optimal TSFs $\mathbf{h_{Tm}}$ to perform a non-blind deblurring of the probe. However, we found that, deblurring artifacts introduced in this process tend to significantly reduce the recognition accuracy (by almost 15% to 20%) which suggests that reblurring the gallery is preferable to deblurring the probe.

C. Experiments

We evaluate the proposed algorithm NU-MOB on the standard and publicly available FERET database [32]. Since this database contains only focused images, we blur the images synthetically to generate the probes. The camera motion itself is synthesized so as to yield a connected path in the motion space. The resulting blur induced mimics the real blur encountered in practical situations. In all the experiments presented in this paper, we use grayscale images resized to 64 × 64 pixels and we assume only one image per subject in the gallery.

To evaluate our algorithm on different types and amounts of blur, we synthetically blur the face images with the TSF model using the following five blur settings:

Setting 1 (S1): in-plane translations, Setting 2 (S2): in-plane translations and rotations, Setting 3 (S3): out-of-plane rotations, Setting 4 (S4): out-of-plane translations, and Setting 5 (S5): full 6D blur. We select the transformation intervals on the image plane, both for generating synthetically blurred images and recognizing them, as follows: in-plane translations range = $-4 : 1 : 4$ pixels, out-of-plane translations range = $0.8 : 0.01 : 1.2$, in-plane rotations range = $-2° : 1° : 2°$. The focal length is set to 200 pixels which is in line with the width of the image (64 pixels), and out-of-plane rotations range is selected as $= -\frac{4}{3}° : \frac{1}{3}° : \frac{4}{3}°$.

The transformation intervals are chosen such that synthetically blurring a 64 × 64 pixel image using transformations lying in these intervals results in moderate to heavy blur which renders it a challenging problem from a face recognition perspective. Note that the matrix $\mathbf{A_m} \in \mathrm{R}^{N \times N_T}$ in (13) has 4096 rows equal to the number of pixels in the image, while the number of columns N_T is determined by the blur setting. For example, in the case of Setting 2 which has in-plane translations and rotations, N_T = (Number of translation steps along X-axis) × (Number of translation steps along Y-axis) × (Number of rotation steps about Z-axis) =($-4 : 1 : 4$ pixels along X-axis) × ($-4 : 1 : 4$ pixels along Y-axis) × ($-2° : 1° : 2°$ about Z-axis) = 9 × 9 × 5 = 405. The calculation of N_T proceeds along similar lines for the remaining four settings, and the value of N_T is 81 for S1 and S3, 41 for S4, and 1345005 for S5. Sample synthetically blurred probe images for the five different blur settings are shown in Fig. 3. In the *randomly* generated TSFs, the number of non-zero homographies is chosen to be a small fraction of the total number of homographies in the TSF space. This number ranges from 10 for Settings 3 and 4, to 15 for Setting 1, and 25 for Settings 2 and 5.

To evaluate our NU-MOB algorithm, we use the *ba* and *bj* folders in FERET, both of which contain 200 images with one image per subject. We use the *ba* folder as the gallery. Five

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

different probe sets, each containing 200 images, are obtained by blurring the *bj* folder using the settings mentioned above. (See Fig. 3.) The lighting and the pose are the same for both gallery and probe since the objective here is to study our algorithm's capability to model blur. Notice, however, that *small* facial expression changes exist between the gallery and the probe, but the weighting matrix **W** in (13) makes our algorithm reasonably robust to these variations. We set the number of scales in the multiscale implementation to 3 as it offered the best compromise between running time and accuracy.

We also compare our results with several recently proposed face recognition techniques - S.Nos. 1(a), 2 through 9 in Table I. While we select methods 1(a), 2 through 4 for their ability to handle blurred faces, S.Nos. 5 and 6 were chosen because comparisons with existing methods in [26] and [27] suggest that the SRC and the DFR algorithms are among the best for classical face recognition applications. We also compare using a two-step non-uniform deblur [9] + recognize approach in S.Nos. 7 and 8 in Table I since neither SRC nor DFR can cope with blur. Yet another two-step baseline comparison that uses LBP features extracted from the deblurred probe for recognition is provided in S.No. 9. Recognition scores were computed for various blur kernel sizes ranging from 3 to 13 pixels for the DRBF [19] algorithm. We report the best recognition rates in Table II. However, we would like to add that the authors in [19] have, on their data, reported recognition rates that are, on an average, 3 to 4 percent greater using their rDRBF algorithm. For comparison with [17] for the space-varying cases in Settings 2 to 5, following the discussion in Section 4.1.2 of their paper, we divided the image into overlapping patches with sizes 75, 50 and 40 percent of the original image, performed recognition separately on each patch and used a majority vote to calculate the final recognition score. (For Setting 1, the algorithm in 4.1.1 of their paper was used.) This was repeated for various blur kernel sizes ranging from 3 to 13 pixels, and the best recognition rates have been reported in Table II. In our implementation of the FADEIN algorithm, the statistical models for PSF inference were learned from 25 PSFs which included

TABLE I

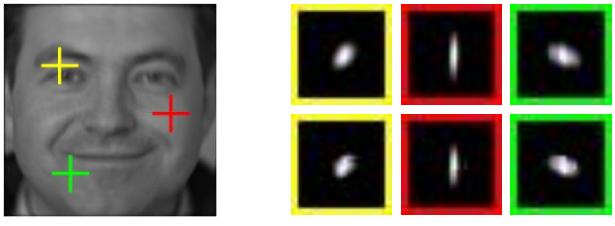
SUMMARY OF COMPARISON TECHNIQUES. UB: UNIFORM BLUR, NUB: NON-UNIFORM BLUR, I: ILLUMINATION, P: POSE

S. No.	Comparison technique	Approach	Methods compared with	Code	Degradations modeled UB	NUB	I	P	Remarks
1(a)	DRBF [19]	Direct recognition using LBP	FADEIN, LPQ, FADEIN+LPQ	Shared by authors	✓	×	×	×	Targeted at recognizing faces acquired from distant cameras where the blur is well-approximated by convolution.
1(b)	IRBF [19]				✓	×	✓	×	
2	[17]	Blur invariants on a manifold for recognition	FADEIN, LPQ, [13]	Shared by authors	✓	✓	×	×	Space-varying blur is handled using overlapping patches, where the blur in each patch is assumed to be uniform.
3	FADEIN [14]	Deblurring using inferred PSF followed by recognition	Eigen faces, Laplacian faces LBP, LPQ, [13]	Our implementation	✓	×	×	×	Limited to learned blur kernels only. Cannot capture the entire space of PSFs.
4	FADEIN + LPQ [14]	Recognition using LPQ features extracted from probe deblurred using FADEIN		LPQ code downloaded from author's webpage	✓	×	✓	×	LPQ's ability to handle illumination is governed by FADEIN correctly inferring the PSF when there is a change in lighting.
5	SRC [27]	l_1-minimization based on sparse representation	Nearest Neighbour, Nearest Subspace, Linear SVM	Our implementation	×	×	✓	×	Dictionary is built using basis images of all subjects. Cannot cope with blur in the images.
6	DFR [26]	Dictionary-based approach	SRC, CDPCA	Shared by authors	×	×	✓	✓	
7	[9] + SRC [27]	Probe deblurred using space-varying blind deconv. code in [9] passed to [27], [26], [30] for recognition	Not applicable	Deblurring code downloaded from the webpage of the first author in [9]	✓	✓	✓	×	Deblurring artifacts are a major source of error.
8	[9] + DFR [26]				✓	✓	✓	✓	
9	[9] + LBP [30]				✓	✓	✓	×	

TABLE II

RECOGNITION RESULTS (%) ON THE FERET DATASET USING NU-MOB ALONG WITH COMPARISONS

Method	S1	S2	S3	S4	S5
NU-MOB	94.5	90.5	94	97	79
DRBF [19]	81	77	78	71	51
[17]	64.5	59.5	59.5	64.5	37
FADEIN [14]	31.5	23	13.5	25.5	12.5
FADEIN + LPQ [14]	39	17	23	50	6
SRC [27]	31.5	33.5	14.5	47.5	13
[9] + SRC [27]	25.5	25	11.5	40.5	13
DFR [26]	34	31	18.5	34.5	16
[9] + DFR [26]	30	26	16.5	27	13
[9] LBP [30]	39.5	30	30	44	9.5



(a) (b)

Fig. 4. (a) Blurred probe, (b) row 1 - true PSFs at three locations (marked by crosshairs in (a)), row 2 - PSFs estimated by our NU-MOB algorithm.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

24 motion blur kernels (length = 3,5,7,9,11,13 pixels, and angle = $0,0.25\pi,0.5\pi,0.75\pi$) and one 'no blur' delta function. Since there is only one image per subject in the current scenario, and SRC and DFR work best in the presence of multiple images for each subject, to be fair, we provide as input to the algorithms in [26] and [27] the nine basis images of each subject (obtained using the relighting technique in [26]) in the database. Table II shows that our method consistently performs better than contemporary techniques, and generalizes satisfactorily to all types of camera motion. A probe image blurred using Setting 2 is shown in Fig. 4(a). The true kernels at three locations marked by crosshairs on Fig. 4(a) are shown in the first row of Fig. 4(b), while the kernels obtained from the TSF estimated by NU-MOB are shown in the second. Observe that the estimated PSFs closely resemble the true PSFs which indicates that the motion has been computed correctly.

Using the multiscale implementation, we could obtain a considerable speed-up of the algorithm. For example, in Setting 2, with 200 images in the gallery, while each query image without the multiscale approach took an average of 230.80 seconds (on an 8Gb linux machine with 8 cores running Matlab), the multiscale implementation (with number of scales set to 3) performed the task in 18.06 seconds. This is an order of speed-up. For Setting 5, where the multiscale implementation is indispensable, the speed-up is around 25 times!

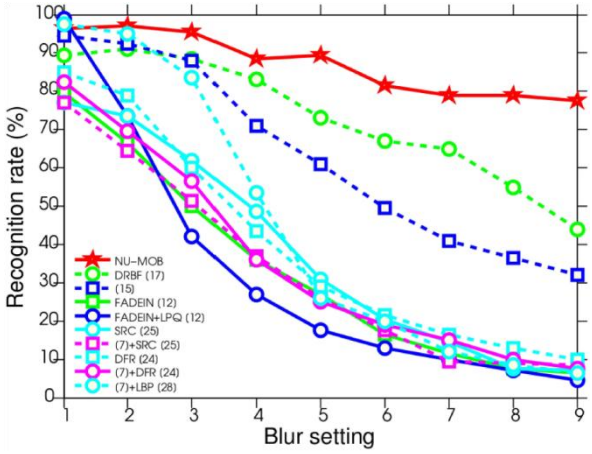


Fig. 5. Effect of increasing the blur. (Refer to the text for blur settings along the *X*-axis.)

1) Effect of Increasing the Blur: We now examine our algorithm's performance as the extent of the blur is increased. The gallery, as before, is the *ba* folder. We select random transformations from the following nine sets of intervals to blur the images in the *bj* folder and generate the probes. The ranges for in-plane translations (in pixels) and in-plane rotations, respectively, are- (1) 0, 0, (2) [−1 : 1 : 1], 0,

(3) [−2 : 1 : 2], [−1° : 1° : 1°], (4) [−3 : 1 : 3],

[−(6)1°[−:5 1:° 1: :1°5], (5)], [−2[−◦ 4: 1:◦1: :2◦4], (7)], [−2[−° 6: :1°1:: 26°]],,

[−2° : 1° : 2°], (8) [−7 : 1 : 7], [−3° : 1° : 3°], and (9) [−8 : 1 : 8],

[−3◦ : 1◦ : 3◦]. It is clear from the plot given in Fig. 5 that our algorithm greatly outperforms all other comparison techniques as the extent of the blur is increased.

2) Effect of Underestimating or Overestimating the TSF Search Intervals: In all the above experiments, we have assumed that the TSF limits are known, and we used the same transformation intervals as the ones used for synthesizing the blur, while attempting recognition. Although in some applications we may know the extent of the blur, in many practical settings, we may not. Hence, we perform the following experiments to test the sensitivity of our algorithm to the TSF search intervals.

As before, the *ba* folder of FERET is chosen as the gallery and the probe images are generated by blurring the *bj* folder using random transformations lying in the intervals- in-plane translations [−2 : 1 : 2] pixels, and in-plane rotations = [−1° : 1° : 1°]. We then perform recognition with the following seven choices of TSF search intervals for in-plane translations and in-plane rotations, respectively- (1) 0, 0,
(2) [−1 : 1 : 1], 0, (3) [−2 : 1 : 2], [−1° : 1° : 1°],

[−(4) [−41◦: :13◦:], (6)4], [−[−2212° :: 11° ::1222]°,], (5)[−4◦ [−: 18◦ :: 14◦]: and8],

3° :
(7) [−16 : 1 : 16], [−5° : 1° : 5°]. The experimental results are shown in the plot of Fig. 6. Observe that, in case (1), where the TSF intervals are set to zero, our method reduces to LBP, and the poor recognition accuracy is further testimony to the fact that the blur, unless accounted for, will cause classical FR algorithms to fail. We note that the recognition rates are

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

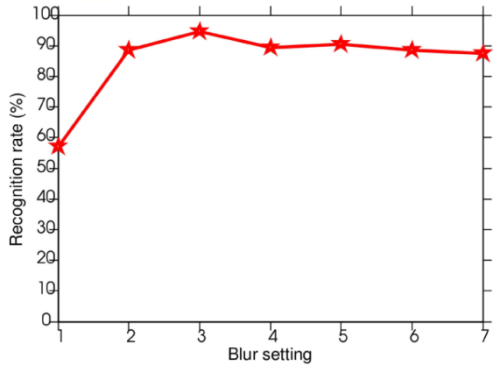


Fig. 6. Effect of underestimating or overestimating the TSF search intervals. (Refer to the text for blur settings along the *X*-axis.)

fairly stable even when the TSF search intervals are chosen to be much larger than the true TSF intervals (cases (4)-(7)) i.e., our algorithm picks the correct TSF even when the search space is large. It can thus be concluded that it is not advisable to underestimate the TSF search intervals.

IV. Face Recognition Across Blur, Illumination, and Pose

Poor illumination is often an accompanying feature in blurred images because larger exposure times are needed to compensate for the lack of light which increases the chances of camera shake. Pose variation is another challenge for realizing the true potential of face recognition systems in practice. This section is devoted to handling the combined effects of blur, illumination and pose.

A. Handling Illumination Variations

To handle illumination variations, we modify our basic blur-robust algorithm (NU-MOB) by judiciously utilizing the following two results:

- In the seminal work of [20], it has been shown that if the human face is modeled as a convex Lambertian surface, then there exists a configuration of nine light source directions such that the subspace formed by the images taken under these nine sources is effective for recognizing faces under a wide range of lighting conditions. Using this "universal configuration" of lighting positions, an image f of a person under any illumination condition can be written as

$$f = \sum_{i=1}^{9} \alpha_i f_i \qquad (14)$$

where $\alpha_i, i = 1,2,...,9$ are the corresponding linear coefficients. The f_is, which form a basis for this 9D subspace, can be generated using the Lambertian reflectance model as

$$f_i(r,c) = \rho(r,c) \max(\mathbf{n}(r,c)^T \mathbf{s}_i, 0) \quad (15)$$

where ρ and $\mathbf{n}$ are the albedo and the surface normal, respectively, at the pixel location (r,c), and $\mathbf{s}$ is the illumination direction. Following [19], we approximate

Algorithm 2 MOBIL: Motion Blur and Illumination-Robust Face Recognition

Input: Blurred and differently illuminated probe image $\mathbf{g}$, and a set of gallery images $\mathbf{f_m}, m = 1, 2, ..., M$.

Output: Identity of the probe image.

1: For each gallery image $\mathbf{f_m}$, obtain the nine basis images $\mathbf{f_{m,i}}, i = 1, 2, ..., 9$.
2: For each gallery image $\mathbf{f_m}$, find the optimal TSF $\mathbf{h_{T_m}}$ and illumination coefficients $\alpha_{m,i}$ by solving equation (16).
3: Transform (blur and re-illuminate) the gallery images $\mathbf{f_m}$ using the computed $\mathbf{h_{T_m}}$ and $\alpha_{m,i}$ and extract LBP features.
4: Compare the LBP features of the probe image $\mathbf{g}$ with those of the transformed gallery images and find the closest match.

the albedo ρ with a frontal, sharp, and well-illuminated gallery image captured under diffuse lighting, and use the average (generic) 3D face normals from [33] for $\mathbf{n}$.

- In [19], it has been shown that for the case of space-invariant blur, the set of all images under varying illumination and blur forms a bi-convex set, i.e., if we fix either the blur or the illumination, the resulting subset is convex. As discussed in Section III, according to the motion blur model for faces, the set of all *motion-blurred* images obtained by blurring a focused gallery image using the TSF model also forms a convex set. Therefore, the result in [19] extends equally well to our situation i.e., the set of all images under varying illumination and non-uniform motion blur also forms a bi-convex set.

Based on these two results, we develop our non-uniform motion blur and illumination (MOBIL)-robust face recognition algorithm. The solution that we seek can be posed as the minimization of the following cost function given by

$$[\mathbf{h_{Tm}}, \alpha_{m,i}] = \underset{\mathbf{h_T}, \alpha_i}{\operatorname{argmin}} \left|\left| \mathbf{W}\left(\mathbf{g} - \sum_{i=1}^{9} \alpha_i \mathbf{A_{m,i}} \mathbf{h_T}\right) \right|\right|_2^2 + \beta \left|\left| \mathbf{h_T} \right|\right|_1$$

$$\text{subject to } \mathbf{h_T} \geq 0 \qquad (16)$$

We adopt the alternating minimization strategy outlined in [19] to solve the above equation. But note that, unlike in [19],

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

we need to solve for the TSF in the motion estimation step. We first obtain the nine basis images $\mathbf{f}_{m,i}$ i = 1,2,...,9 for each gallery image $\mathbf{f}_m$, m = 1,2,..., M. Next, for each gallery image $\mathbf{f}_m$, we estimate the optimal TSF $\mathbf{h}_{Tm}$ and illumination coefficients $\alpha_{m,i}$ by solving equation (16). To determine the identity of the probe, we transform (reblur and re-illuminate) each of the gallery images $\mathbf{f}_m$ using the estimated TSF $\mathbf{h}_{Tm}$ and the illumination coefficients $\alpha_{m,i}$, compute the LBP features from these transformed gallery images and compare them with those from the probe $\mathbf{g}$ to find the closest match. See Algorithm 2.

We now elaborate on the two steps involved in our AM algorithm. For any gallery image $\mathbf{f}_m$, in the first iteration, we assume the blur to be an impulse (i.e., no blur) and estimate the nine illumination coefficients $\alpha_{m,i}$ by solving the linear least squares problem $\mathbf{g} = \mathbf{L}_m\alpha_m$, where $\mathbf{L}_m$ is a matrix whose

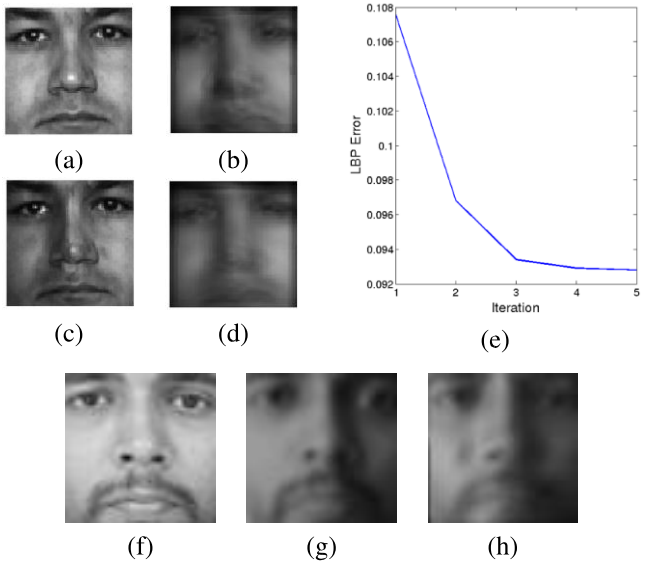


Fig. 7. An example image from the PIE dataset illustrating the alternating minimization algorithm is shown in the first two rows. (a) Gallery, (b) probe, (c) relit gallery image, (d) reblurred and relit gallery image, and (e) a plot of the LBP cost versus iterations. Another example from our real dataset is shown in row 3. (f) Gallery, (g) probe, (h) reblurred and relit gallery image.

nine columns contain the basis images $\mathbf{f}_{m,i}$ corresponding to the subject m lexicographically ordered as vectors, and $\alpha_m = [\alpha_{m,1},\alpha_{m,2},...,\alpha_{m,9}]$ are its corresponding illumination coefficients. Now, we create a new relit gallery image from the basis images using the estimated illumination coefficients $\alpha_{m,i}$. This completes the first step of the alternation wherein we fixed the blur and estimated the illumination. Next, we build $\mathbf{A}_m$ using warped versions of this relit gallery image and estimate the TSF $\mathbf{h}_{Tm}$ by solving equation (13). Observe that the illumination is fixed and we are solving for the blur in this step. Following this, we blur each of the basis images using the estimated TSF $\mathbf{h}_{Tm}$ before proceeding to the second iteration. Note that the matrix $\mathbf{L}_m$ in the second and subsequent iterations is built from the blurred basis images. To summarize, in one step, we minimize over $\alpha_{m,i}$ keeping $\mathbf{h}_{Tm}$ fixed and in the other, we minimize over $\mathbf{h}_{Tm}$ keeping $\alpha_{m,i}$ fixed, and iterate till convergence. An example from the PIE dataset illustrating this sequence of operations is shown in Fig. 7. The optimization problem is convex with respect to blur and illumination considered individually. Therefore, we are guaranteed to converge to at least a local minima by alternately optimizing over $\mathbf{h}_{Tm}$ and $\alpha_{m,i}$. Another example from our real dataset depicting the gallery, the probe, and the reblurred and relit gallery is also shown in Fig. 7.

B. Handling Pose Variations

Most face recognition algorithms are robust to small variations in pose (~ 15°) [25], but the drop in performance is severe for greater yaw and pitch angles. In our experiments, we found this to be true of our MOBIL algorithm also. The reason behind this drop in accuracy is that intra-subject variations caused by rotations are often larger than inter-subject differences. Clearly, there is no overstating the

Fig. 8. Example images of a subject from the PIE database under new poses. The images in (a) and (b) are synthesized from the frontal gallery image using the average face depthmap shown in (c).

formidable nature of the problem at hand - recognizing faces across blur, illumination and pose. To this end, we next propose our MOBILAP algorithm which, using an estimate of the pose, matches the incoming probe with a synthesized non-frontal gallery image. To the best of the authors' knowledge, this is the first ever effort to even attempt this compounded scenario.

Owing to the robustness of face recognition algorithms to small pose variations of upto ~15°, a certain level of quantization of the in-depth rotations is possible [25]. A recent work [34] that unifies face detection, pose estimation, and landmark localization has also adopted this 15° discretization. This method, suited for focused, cluttered images, detects the face(s) and returns a quantized estimate (between −90° to 90° every 15°) of the pose(s). We use this technique to obtain an estimate of the pose of the blurred probe image. We note from our experiments that there are errors in landmark localization

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

due to blur, and the method in [34] can then yield inaccurate pose estimates with the true pose being returned only about 45–55% of the time. However, it almost always returns an estimate which is within ±15° of the true pose. Using this estimate, we synthesize, from each frontal gallery, the image of the subject under the new pose with the help of the average depthmap used in Section IV-A. (See Fig. 8.) These synthesized poses now form the new gallery set. Although the shape of each subject's face may vary from the generic depthmap, the algorithm retains its simplicity and the increase in computational time due to this step is only minimal. The nine illumination basis images are estimated as before using (15) but with ρ now being the new synthesized pose and $\mathbf{n}$ being the surface normals recomputed from the rotated depthmap. Note that the motion blur model for faces discussed in Section III applies even in the event of a pose change. An overview of the algorithm is presented in Algorithm 3.

V. Experiments

In Section V-A, we first demonstrate the effectiveness of our MOBIL algorithm (of Section IV) in recognizing faces across blur and illumination using two publicly available databases - PIE [35] and Extended Yale B [36]. Using the PIE dataset, we further go on to show, in Section V-B, how our MOBILAP algorithm can handle even pose variations. Note that, as before, we blur the images synthetically to generate the probes as these two databases do not contain motion blur. Therefore, these experiments (and the ones already discussed in Section III-C) are synthetic or quasi-real because the blur is synthetically introduced. In Section V-C, we report

Algorithm 3 MOBILAP: Motion Blur, Illumination and Pose-Robust Face Recognition

Input: Blurred and differently illuminated probe image $\mathbf{g}$ under a different pose, and a set of gallery images $\mathbf{f_m}, m = 1, 2, ..., M$.

Output: Identity of the probe image.

1: Obtain an estimate of the pose of the blurred probe image using the method in [34].
2: For each gallery image $\mathbf{f_m}$, synthesize the new pose $\mathbf{f_{syn_m}}$ based on the above estimate.
3: For each synthesized gallery image $\mathbf{f_{syn_m}}$, obtain the nine basis images $\mathbf{f_{syn_{m,i}}}, i = 1, 2, ..., 9$ using normals recomputed from the rotated depthmap.
4: For each synthesized gallery image $\mathbf{f_{syn_m}}$, find the optimal TSF $\mathbf{h_{T_m}}$ and illumination coefficients $\alpha_{m,i}$ by solving equation (16).
5: Transform (blur and re-illuminate) the synthesized gallery images $\mathbf{f_{syn_m}}$ using the computed $\mathbf{h_{T_m}}$ and $\alpha_{m,i}$ and extract LBP features.
6: Compare the LBP features of the probe image $\mathbf{g}$ with those of the transformed gallery images and find the closest match.

MOBILAP's results on the Labeled Faces in the Wild [37] dataset (which is a publicly available real dataset) using the 'Unsupervised' protocol. We also evaluate the performance of MOBILAP on our own real dataset captured using a handheld camera that contains significant blur, illumination and pose variations, in addition to small occlusions and changes in facial expressions.

A. Recognition Across Blur and Illumination

We first run our MOBIL algorithm on the *illum* subset of the PIE database which consists of images of 68 individuals under different illumination conditions. We use faces with a frontal pose (c_{27}) and frontal illumination (f_{11}) as our gallery. The probe dataset, which is also in the frontal pose (c_{27}), is divided into two categories- 1) Good Illumination (GI) consisting of subsets f_{06}, f_{07}, f_{08}, f_{09}, f_{12} and f_{20} (6 different illumination conditions) and 2) Bad Illumination (BI) consisting of subsets f_{05}, f_{10}, f_{13}, f_{14}, f_{19} and f_{21} (6 different illumination conditions). Next, we blur all the probe images using the five different blur settings, and the transformation intervals discussed in Section III-C.

To perform recognition using MOBIL, we first compute the nine illumination basis images for each gallery image as described in Section IV. For comparison, we used the methods S.Nos. 1(b), 2, through 9 mentioned in Table I. Since [17] does not have an explicit formulation to handle variations in lighting, we followed the approach taken in their paper and histogram equalized both the gallery and the probe images before executing their algorithm. The recognition results are provided in Table III. It is clearly evident that MOBIL outperforms the comparison techniques (including NU-MOB which does not explicitly handle illumination) for all blur settings.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

TABLE III

RECOGNITION RESULTS (%) ON THE PIE DATASET USING MOBIL ALONG WITH COMPARISONS

Blur Setting	S1		S2		S3		S4		S5	
Illumination	GI	BI	GI	BI	GI	BI	GI	BI	GI	BI
MOBIL	99.75	99.51	99.26	98.04	99.75	99.51	99.75	99.75	93.38	74.51
NU-MOB	93.87	57.60	85.29	47.30	88.24	50.00	97.79	78.19	45.10	12.25
IRBF [19]	93.87	82.11	93.38	79.17	77.45	61.27	97.30	91.67	52.70	40.69
[17]	56.86	24.51	58.58	24.75	50.98	18.87	75.25	37.75	28.92	12.99
FADEIN [14]	18.87	1.96	17.89	1.72	12.01	1.72	17.16	1.72	9.31	1.72
FADEIN + LPQ [14]	71.32	45.59	37.75	19.12	48.04	26.47	67.16	48.04	23.28	12.50
SRC [27]	41.67	35.05	38.97	32.11	25.49	23.28	82.60	67.40	21.32	17.40
[9] + SRC [27]	38.73	31.86	33.33	32.60	21.32	22.79	77.45	58.58	17.89	16.18
DFR [26]	54.66	46.32	50.25	39.71	32.60	30.88	75.25	64.95	22.30	19.85
[9] + DFR [26]	51.23	43.14	46.81	39.46	32.11	27.70	67.40	55.64	21.81	17.65
[9] LBP [30]	57.60	32.84	43.38	26.72	49.75	28.68	69.36	38.24	24.75	14.71

TABLE IV

RECOGNITION RESULTS (%) ON THE EXTENDED YALE B DATASET USING MOBIL ALONG WITH COMPARISONS

Blur Setting	S1		S2		S3		S4		S5	
Illumination	GI	BI	GI	BI	GI	BI	GI	BI	GI	BI
MOBIL	99.60	91.07	99.60	87.14	99.60	88.57	99.40	87.14	95.27	62.14
NU-MOB	88.29	26.79	82.94	28.21	85.52	26.79	92.66	33.57	51.03	12.98
IRBF [19]	89.29	52.50	89.48	48.57	73.41	40.71	93.45	58.57	51.64	26.67
[17]	62.90	14.29	55.56	13.57	56.15	14.29	75.20	23.21	36.83	13.33
FADEIN [14]	37.70	4.64	35.52	5.00	27.18	4.64	34.13	5.71	22.43	4.81
FADEIN + LPQ [14]	69.44	37.86	49.60	28.93	56.94	33.93	76.59	43.57	32.51	24.07
SRC [27]	59.52	33.21	60.91	34.29	42.06	26.43	81.94	43.21	38.68	17.78
[9] + SRC [27]	56.94	30.36	57.14	32.50	39.09	27.86	78.37	41.43	36.42	16.30
DFR [26]	74.80	44.29	72.82	42.14	55.36	33.21	79.76	46.43	46.30	28.89
[9] + DFR [26]	69.44	39.64	69.05	40.71	49.60	30.36	76.79	41.43	44.65	25.56
[9] LBP [30]	60.91	34.29	50.60	27.86	50.79	31.43	64.68	28.93	26.95	17.78

We also tested our Algorithm 2 on the Extended Yale B dataset. The dataset consists of images of 28 individuals under different illumination conditions. The frontal pose $P00$ and light source position $A{+}000E{+}00$ constitutes the gallery. The probe set is once again divided into two categories - 1) Good Illumination (GI) consisting of subsets $A\,{-}005E\,{-}10$, $A\,{-}005E\,{+}\,10$, $A\,{+}\,005E\,{-}10$, $A\,{+}\,005E\,{+}\,10$, $A\,{+}\,000E\,{+}\,20$, $A\,{+}\,000E\,{-}20$, $A\,{-}010E\,{+}\,00$, $A\,{-}010E\,{-}20$, $A\,{+}\,010E\,{+}\,00$, $A\,{+}\,010E\,{-}20$, $A\,{-}\,015E\,{+}\,20$, $A\,{+}\,015E\,{+}\,20$, $A\,{-}\,020E\,{+}\,10$, $A\,{-}020E\,{-}10$, $A\,{+}\,020E\,{-}10$, $A\,{+}\,020E\,{+}\,10$, $A\,{-}025E\,{+}\,00$, $A\,{+}\,025E\,{+}\,00$ (18 different illumination conditions), and 2) Bad Illumination (BI) consisting of subsets $A\,{-}020E\,{-}40$, $A\,{-}035E\,{-}20$, $A\,{-}035E\,{+}\,15$, $A\,{-}035E\,{+}\,40$, $A\,{+}\,000E\,{+}\,45$, $A\,{+}\,035E\,{+}\,40$, $A\,{+}\,035E\,{+}\,15$, $A\,{+}\,035E\,{-}\,20$, $A\,{+}\,020E\,{-}\,40$, $A\,{+}\,000E\,{-}35$ (10 different illumination conditions.) Here, A denotes the azimuth angle and E denotes the elevation angle of the light source. The camera is fixed at the frontal pose ($P00$) and only the light source positions are varied. The recognition results are presented in Table IV and MOBIL yet again scores over others. The AM algorithm converges in a few iterations, and in our experiments, we normally terminate it after five iterations. Due to the severity of blur, our results for full 6D motion are marginally lower than our scores for the other four blur settings.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

In the experiments conducted so far, the blur was synthesized on pre-aligned images i.e., the gallery and probe images were aligned prior to applying the blur on the probe. To demonstrate our method's robustness to alignment errors, we also performed the following experiment. For the GI subsets of both PIE and Extended Yale B datasets, we used the code in [34] to detect the eye centers in the blurred probes. The probes were then registered with the gallery using these detected eye centers. We note that the method in [34], designed for focused images, does not return very accurate eye centers when a blurred image is presented. However, even a coarse registration between the focused and the blurred faces suffices for our technique because minor alignment errors are implicitly accounted for in the TSF estimation procedure i.e., small pixel misalignments are compensated for by the in-plane translational search intervals in the motion estimation step. We noticed that the drop in recognition accuracy (over pre-aligned images) is quite marginal - only 1 to 2% on an average, whereas the scores of our closest competitors (despite being provided aligned images directly) are much lower as can be seen from Tables III and IV.

B. Recognition Across Blur, Illumination, and Pose

Finally, we take up the very challenging case of allowing for pose variations in addition to blur and illumination. We once again use the PIE dataset. We begin by selecting four *near-frontal* poses (pitch and yaw angles within ~15°) and explore the robustness of MOBIL itself to small variations in pose.

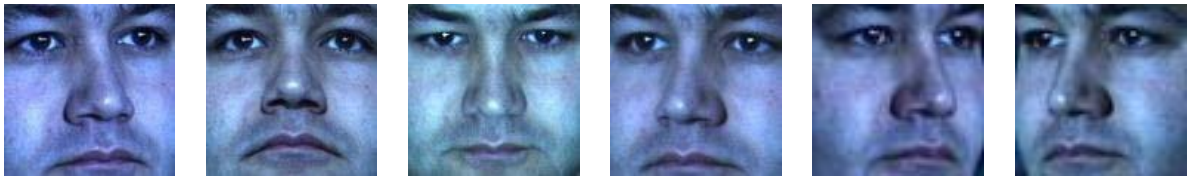


Fig. 9. PIE database. Different camera poses under frontal illumination (f11).

TABLE V

RECOGNITION RESULTS (%) FOR MOBIL ON THE GOOD ILLUMINATION SUBSET OF THE PIE DATABASE ALONG WITH COMPARISONS

Blur Setting	S1				S2				S3				S4			
Pose	C05	C07	C09	C29	C05	C07	C09	C29	C05	C07	C09	C29	C05	C07	C09	C29
MOBIL	98.53	98.52	99.02	97.06	97.06	96.79	99.02	95.34	98.53	97.04	99.51	96.32	96.81	96.79	99.26	89.71
IRBF [19]	69.12	68.63	78.19	40.93	61.52	67.40	75.98	39.71	55.39	51.47	59.31	34.56	65.93	74.02	81.86	42.89
[17]	38.48	39.22	33.33	36.27	34.31	34.80	31.13	36.27	29.41	30.88	28.92	26.47	34.56	38.97	43.63	33.09
FADEIN [14]	14.71	12.99	20.59	7.84	11.76	11.03	18.63	7.35	8.58	8.33	16.42	4.66	11.52	10.05	19.36	6.62
FADEIN LPQ [14]	46.81	56.13	41.18	34.56	22.05	31.86	20.59	19.36	24.75	42.65	28.19	24.75	37.25	55.15	39.22	30.15
SRC [27]	21.08	23.77	28.19	20.34	18.87	22.55	26.47	17.89	12.99	15.69	17.65	13.48	20.83	34.80	50.49	22.06
[9] SRC [27]	14.95	20.10	25.49	17.89	14.46	19.36	23.28	17.65	11.52	14.46	16.18	12.01	14.95	32.84	43.14	19.36
DFR [26]	27.45	30.64	37.01	18.14	25.98	29.66	31.86	15.69	19.61	16.67	23.77	12.50	26.23	32.84	45.59	15.93
[9] DFR [26]	17.65	24.57	27.21	16.18	15.69	18.14	26.96	14.71	12.25	15.44	18.14	11.52	13.97	23.77	28.68	16.42
[9] LBP [30]	46.57	41.91	47.55	40.93	35.29	34.31	39.22	30.88	38.48	35.05	41.18	34.80	43.38	46.08	59.31	40.20

TABLE VI

RECOGNITION RESULTS (%) FOR MOBIL ON THE BAD ILLUMINATION SUBSET OF THE PIE DATABASE ALONG WITH COMPARISONS

Blur Setting	S1				S2				S3				S4			
Pose	C05	C07	C09	C29	C05	C07	C09	C29	C05	C07	C09	C29	C05	C07	C09	C29
MOBIL	77.70	81.13	91.42	79.17	76.47	76.23	91.18	74.02	73.28	77.45	87.99	76.47	77.45	78.19	91.67	70.83
[19] IRBF	50.25	54.17	60.78	39.71	39.71	49.02	52.70	35.05	35.29	37.99	45.59	27.21	43.87	59.80	64.22	32.35
[17]	13.48	14.22	14.46	12.25	11.52	14.95	15.20	11.52	10.78	12.50	10.54	11.03	14.46	16.91	17.16	15.44
FADEIN [14]	1.72	1.72	2.21	1.96	1.72	1.72	2.21	1.96	1.72	1.72	2.21	1.96	1.96	1.96	2.21	1.96
FADEIN LPQ [14]	25.49	37.01	24.26	18.38	10.54	18.63	11.03	6.62	14.95	25.00	15.44	12.75	23.77	37.25	25.74	17.40
SRC [27]	17.89	17.89	23.77	19.36	16.18	18.14	23.53	17.89	12.25	14.71	16.42	12.01	15.69	22.79	33.58	21.81
[9] SRC [27]	13.97	17.16	20.10	17.16	15.69	16.91	20.83	16.18	9.56	11.27	13.73	9.80	11.52	20.10	28.92	18.14
DFR [26]	22.79	27.94	31.62	20.10	17.89	20.59	28.92	16.42	14.22	17.65	22.79	13.24	18.14	24.75	36.76	18.38
[9] DFR [26]	22.55	29.90	30.88	16.42	22.30	28.19	29.41	12.75	18.14	15.44	21.32	11.27	21.32	31.13	38.48	13.73
[9] LBP [30]	24.02	25.00	28.68	23.04	19.36	20.10	20.83	14.22	19.36	20.10	23.28	20.83	25.25	28.92	31.86	25.49

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

As before, the camera position c_{27} (frontal pose) and flash position f_{11} (frontal illumination) constitute the gallery. In this experiment, however, the probe set, divided into good and bad illumination subsets, contains the four nearfrontal poses c_{05} (−16° yaw), c_{07} (0° yaw and −13° tilt), c_{09} (0° yaw and 13° tilt) and c_{29} (17° yaw). Note that both camera poses and flash positions are varied. Example images are shown in Fig. 9 columns 1 to 4. The recognition results for GI and BI, presented in Tables V and VI, respectively, are clear indicators of MOBIL's robustness to small variations in pose. However, we note that there is some drop in our algorithm's recognition accuracy in going from GI to BI due to the combined effects of blur, illumination and pose.

Next, we select differently illuminated probes in two non-frontal poses c_{37} (−31° yaw) and c_{11} (32° yaw). See Fig. 9 columns 5 and 6. Once again, the frontal camera position c_{27} and flash position f_{11} constitute the gallery. For such large changes in pose, we found that MOBIL returned recognition rates less than 15%. Hence, as discussed in Section IV-B, we first obtain an estimate of the pose of the blurred probe image using the code of [34]. Recognition is performed using the steps outlined in Algorithm 3 MOBILAP. The average success rate of the algorithm in [34] in determining the exact true pose (−30° for c_{37} and +30° for c_{11}) was only 52.70%. However, the estimated pose was almost always within ±15° of the true pose. We have seen from our previous experiment that our method is robust to pose variations of upto ±15°. Therefore, our recognition rates remain largely unaffected despite some errors in pose estimation, and our MOBILAP algorithm exhibits stable performance even under considerable pose variations and illumination changes. This can be seen from our results in Table VII. As this is the first attempt of its kind at recognizing faces across blur, illumination and pose, there are no comparison results to report.

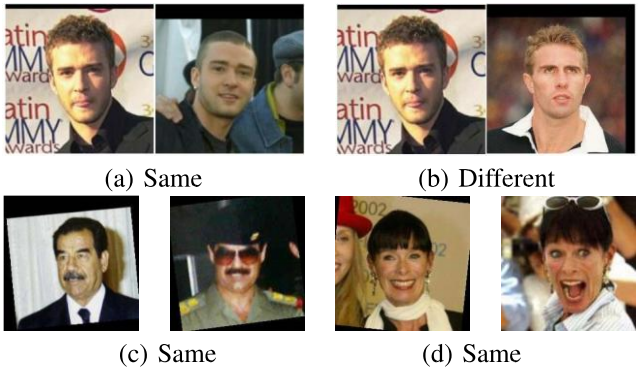
C. Real Datasets

We also report MOBILAP's performance on two real datasets - the publicly available Labeled Faces in the Wild (LFW) [37] dataset and our own real dataset captured under unconstrained settings. First, we evaluate our algorithm's performance on LFW which is a very challenging dataset containing 13,233 images of 5,749 subjects in which the face images have large variations in illumination, pose, expression and age, in addition to other degradations such as occlusion, low resolution etc. However, as pointed out in [38] and [39] the images in LFW are typically posed and framed by professional photographers, and are known to contain very little or no blur. Even so, an evaluation on this dataset is quite useful because in real applications, the extent of blur in the probe images is not known a priori. The database was designed to study the *face verification* problem in which a pair of two face images are presented, and it is required to classify the pair as either 'same' or 'different' depending upon whether the images are of the same person or not. Fig. 10

shows examples of 'same' and 'different' pairs of face images from this dataset. We evaluate the proposed approach on 'View 2' of the dataset (as per LFW protocol) consisting of 3,000 matched and 3,000 mismatched pairs divided into 10 sets. Since our method does not involve any training, we report results under the 'Unsupervised' protocol. Note that this protocol is considered the most difficult [40] of all since no training data is available.

TABLE VII
RECOGNITION RESULTS (%) OF MOBILAP ACROSS BLUR, ILLUMINATION, AND POSE ON THE PIE DATABASE

Pose	c37					c11				
Illumination	f6	f7	f8	f11	f20	f6	f8	f11	f12	f20
S1	85.29	80.88	75.00	92.65	83.82	75.00	88.24	91.18	82.35	86.76
S2	79.41	63.24	73.53	86.76	72.06	64.71	73.53	91.18	60.29	79.41
S3	80.88	75.00	76.47	91.18	76.47	67.65	82.35	97.06	72.06	86.76
Average	81.86	73.04	75.00	90.20	77.45	69.12	81.37	93.14	71.57	84.31



(a) Same (b) Different

(c) Same (d) Same

Fig. 10. Examples of same and different pairs of face images from the LFW dataset.

In the Unsupervised paradigm, the area under the ROC curve (AUC) has been selected as a suitable scalar-valued measure of accuracy according to the latest score reporting procedures for LFW. We use the LFW-aligned version [41] of the database to report our scores with Chi square distance as our similarity measure. Given the challenging nature of this dataset, we use multi-resolution LBP histograms [40] (uniform LBP histograms are extracted in 10 different resolutions instead of just one single resolution) for this particular experiment only. In order to minimize the effects of the background, the faces are first closely cropped. The search intervals for the TSF were kept small (in-plane translations of [−2 : 1 : 2] pixels, and in-plane rotations of [−1° : 1° : 1°]) because the images in this dataset contain very little or no

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

blur. Since frontal well-illuminated gallery images may not be available in all cases, we first reblur, relight and change the pose of the first image with respect to the second, and match them. The pose is compensated for using the estimate of the pose returned by [34]. Following [40], we then exchange the roles of the first and second image and match again. This procedure is then repeated for horizontally mirrored versions of both images. The final similarity measure for a given image pair is considered as the minimum of the distances thus obtained from these four combinations. Table VIII reports the AUC values obtained by MOBILAP along with other approaches that follow the Unsupervised protocol. The scores have been reproduced from the LFW results page http://viswww.cs.umass.edu/lfw/results.html#Unsupervised.
The ROC curves are also shown in the plot of Fig 11 in order to better evaluate the performance. Note that MOBILAP ranks close to MRF-MLBP [40] which is next only to PAF [42]. LFW contains significant amounts of occlusion, expression and age variations (see Fig. 10(c)-(d)) which we do not model in our approach, whereas competing methods handle one or more of these effects explicitly. While our framework can also handle blur when introduced in LFW (see supplementary material) the competing methods are not tailored to deal with blur.

TABLE VIII

PERFORMANCE COMPARISON FOR DIFFERENT METHODS ON LFW DATABASE UNDER THE 'UNSUPERVISED' PROTOCOL

Method	AUC
SD-MATCHES, 125x125, funneled	0.5407
H-XS-40, 81x150, funneled	0.7547
GJD-BC-100, 122x225, funneled	0.7392
LARK unsupervised, aligned	0.7830
LHS, aligned	0.8107
Pose Adaptive Filter [42]	0.9405
MRF-MLBP, aligned [40]	0.8994
MOBILAP, aligned	0.8410

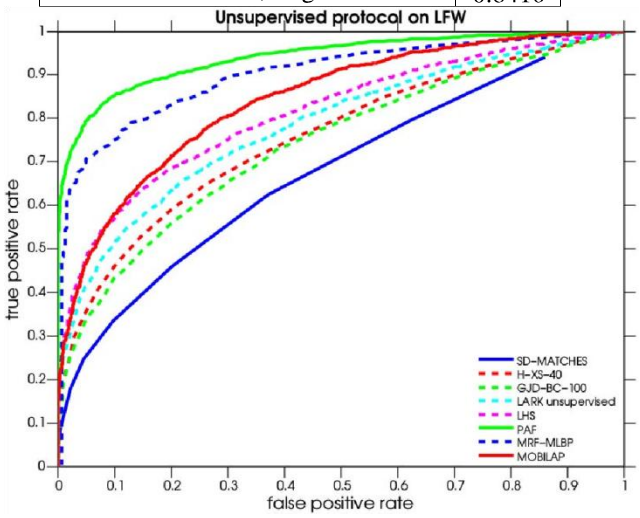


Fig. 11. ROC curves of different approaches on the LFW dataset for the Unsupervised protocol.

Finally, we report recognition results on a real dataset that contains face images that we ourselves captured in unconstrained settings. There are 50 subjects in the dataset. The gallery contains one frontal, sharp and well-illuminated image taken outdoor under diffuse lighting. The probe images, 2,200 in number, were captured using a hand-held camera under indoor and outdoor lighting conditions. The probe images suffer from varying types and amounts of blur, variations in illumination and pose, and even some occlusion and facial expression changes. Although the blur was predominantly due to camera shake, no restriction was imposed on the

Fig. 12. Cropped faces of four subjects from the real dataset. The gallery images are shown in column one. The probe images have variations in illumination, facial expressions changes (column four in row one), small occlusions (missing spectacles - column four, row two) and differences in pose (columns two through four in row three).

TABLE IX

RECOGNITION RESULTS FOR MOBILAP ON OUR REAL DATASET ALONG WITH COMPARISONS

Method	Recognition rate (%)
MOBILAP	76.27
IRBF [19]	58.32
[17]	42.50
FADEIN [14]	20.45
FADEIN + LPQ [14]	36.36
SRC [27]	54.09
[9] + SRC [27]	43.59
DFR [26]	43.00

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

[9] + DFR [26]	33.82
[9] + LBP [30]	43.14

movement of the subjects during image capture, and, therefore, a subset of these images could possibly have both camera and object motion. Following [19], we manually cropped the faces and resized them to 64×64 pixels. Some representative images from the gallery and probe are given in Fig. 12. Observe that, as compared to the gallery, the probes can be either overlit or underlit depending on the setting under which they were captured. We generate the nine illumination basis images for each image in the gallery and then run MOBILAP. It has been pointed out in [9]–[11] that in most practical scenarios, a 3D TSF is sufficient to explain the general motion of the camera. In view of this observation and in consideration of computation time, we select the search intervals for the TSF as [−4 : 1 : 4] pixels for in-plane translations, and [−2° : 1° : 2°] for in-plane rotations. The recognition results are presented in Table IX. Although the accuracy of all the methods drop due to the unconstrained and challenging nature of this dataset, the effectiveness of the proposed technique in advancing the state-of-the-art in handling non-uniform blur, illumination, and pose in practical scenarios is reaffirmed yet again.

VI. Conclusions

We proposed a methodology to perform face recognition under the combined effects of non-uniform blur, illumination, and pose. We showed that the set of all images obtained by non-uniformly blurring a given image using the TSF model is a convex set given by the convex hull of warped versions of the image. Capitalizing on this result, we initially proposed a non-uniform motion blur-robust face recognition algorithm NU-MOB. We then showed that the set of all images obtained from a given image by non-uniform blurring and changes in illumination forms a bi-convex set, and used this result to develop our non-uniform motion blur and illumination-robust algorithm MOBIL. We then extended the capability of MOBIL to handle even non-frontal faces by transforming the gallery to a new pose. We established the superiority of this method called MOBILAP over contemporary techniques. Extensive experiments were given on synthetic as well as real face data. The limitation of our approach is that significant occlusions and large changes in facial expressions cannot be handled.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, Dec. 2003.

[2] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, Jul. 2006.

[3] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 73:1–73:10, Aug. 2008.

[4] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2354–2367, Dec. 2011.

[5] M. Šorel and F. Šroubek, "Space-variant deblurring using one blurred and one underexposed image," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 157–160.

[6] H. Ji and K. Wang, "A two-stage approach to blind spatially-varying motion deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 73–80.

[7] S. Cho, Y. Matsushita, and S. Lee, "Removing non-uniform motion blur from images," in *Proc. Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[8] Y.-W. Tai, P. Tan, and M. S. Brown, "Richardson-Lucy deblurring for scenes under a projective motion path," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1603–1618, Aug. 2011.

[9] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186, 2012.

[10] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 171–184.

[11] Z. Hu and M.-H. Yang, "Fast non-uniform deblurring using constrained camera pose subspace," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 1–11.

[12] C. Paramanand and A. N. Rajagopalan, "Non-uniform motion deblurring for bilayer scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1115–1122.

[13] H. Hu and G. de Haan, "Adaptive image restoration based on local robust blur estimation," in *Proc. 9th Int. Conf. Adv. Concepts Intell. Vis. Syst.*, 2007, pp. 461–472.

[14] M. Nishiyama, A. Hadid, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi, "Facial deblur inference using subspace analysis for recognition of blurred faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 838–845, Apr. 2011.

[15] H. Zhang, J. Yang, Y. Zhang, N. M. Nasrabadi, and T. S. Huang, "Close the loop: Joint blind image restoration and recognition with sparse representation prior," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 770–777.

[16] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkila, "Recognition of blurred faces using local phase quantization," in *Proc. 19th Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.

[17] R. Gopalan, S. Taheri, P. Turaga, and R. Chellappa, "A blur-robust descriptor with applications to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1220–1226, Jun. 2012.

[18] I. Stainvas and N. Intrator, "Blurred face recognition via a hybrid network architecture," in *Proc. 15th Int. Conf. Pattern Recognit.*, vol. 2. Sep. 2000, pp. 805–808.

[19] P. Vageeswaran, K. Mitra, and R. Chellappa, "Blur and illumination robust face recognition via set-theoretic characterization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1362–1372, Apr. 2013.

[20] K.-C. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.

[21] S. Biswas, G. Aggarwal, and R. Chellappa, "Robust estimation of albedo for illumination-invariant matching and shape recovery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 5, pp. 884–899, May 2009.

[22] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2599–2606, Nov. 2009.

[23] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Analysis and Modeling of Faces and Gestures* (Lecture Notes in Computer Science), vol. 4778. Berlin, Germany: Springer-Verlag, 2007, pp. 168–182.

[24] G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "Subspace learning from image gradient orientations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 12, pp. 2454–2466, Dec. 2012.

[25] X. Zhang and Y. Gao, "Face recognition across pose: A review," *Pattern Recognit.*, vol. 42, no. 11, pp. 2876–2896, Nov. 2009.

[26] V. M. Patel, T. Wu, S. Biswas, P. J. Phillips, and R. Chellappa, "Dictionary-based face recognition under variable lighting and pose," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 954–965, Jun. 2012.

[27] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[28] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, and Y. Ma, "Towards a practical face recognition system: Robust registration and illumination by sparse representation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 597–604.

[29] C. Paramanand and A. N. Rajagopalan, "Inferring image transformation and structure from motion-blurred images," in *Proc. Brit. Mach. Vis. Conf.*, 2010, pp. 1–12.

[30] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[31] J. Liu, S. Ji, and J. Ye, *SLEP: Sparse Learning With Efficient Projections*. Tempe, AZ, USA: Arizona State Univ., 2009.

[32] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[33] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1063–1074, Sep. 2003.

[34] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2879–2886.

[35] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.

[36] A. S. Georghiades, P. N. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[37] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Dept. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 07-49, 2007.

[38] R. Chellappa, J. Ni, and V. M. Patel, "Remote identification of faces: Problems, prospects, and progress," *Pattern Recognit. Lett.*, vol. 33, no. 14, pp. 1849–1859, Oct. 2012.

[39] N. Pinto, J. J. DiCarlo, and D. D. Cox, "How far can you get with a modern face recognition test set using only simple features?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2591–2598.

[40] S. R. Arashloo and J. Kittler, "Efficient processing of MRFs for unconstrained-pose face recognition," in *Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Sep./Oct. 2013, pp. 1–8.

[41] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 1978–1990, Oct. 2011.

[42] D. Yi, Z. Lei, and S. Z. Li, "Towards pose robust face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3539–3545.