Face Recognition in the Presence of Multiple Illumination Sources

Gaurav Aggarwal*
Dept. of Computer Science
University of Maryland, College Park, MD
gaurav@cs.umd.edu

Rama Chellappa
Dept. of Electrical and Computer Engineering
University of Maryland, College Park, MD
rama@cfar.umd.edu

Abstract

Most existing face recognition algorithms work well for controlled images but are quite susceptible to changes in illumination and pose. This has led to the rise of analysis-by-synthesis approaches due to their inherent potential to handle these external factors. Though these approaches work quite well, most of them assume that the face is illuminated by a single light source which is usually not true in realistic conditions. In this paper, we propose an algorithm to recognize faces illuminated by arbitrarily placed, multiple light sources. The algorithm does not need to know the number of light sources and works extremely well even while recognizing faces illuminated by different number of light sources. Results using this algorithm are reported on multiple-illumination datasets generated from PIE [10] and Yale Face Database B [5]. We also highlight the importance of the hard non-linearity in the Lambert's law which is often ignored, probably to linearize the estimation process.

1. Introduction

Recent improvements in the accuracy of the face recognition algorithms for images taken under controlled conditions has shifted the focus to more challenging tasks of achieving the same performance for uncontrolled scenarios. A detailed survey of various face recognition algorithms is presented in [15]. Several researchers attempt to achieve invariance to illumination by using image processing techniques like histogram equalization [11]. Some subspace based methods try to counter illumination variations by discarding the first few principal components [3]. These techniques do improve the accuracies of the respective algorithms but are usually ineffective in the case of a non-trivial change in illumination conditions.

The inability of such heuristics to handle illumination variation has led to the rise of generative (or analysis-by-synthesis) approaches for face recognition [4][16][12][5][9]. Broadly speaking, these techniques try to model the physical process of image formation by taking into consideration quantities like surface albedo, surface normals and illumination source direction. Though the recovery of shape and surface properties (reflectivity or albedo) from image(s) has been studied for a long time, its application to the problem of face recognition is fairly recent. An example in this category is the application of shape from shading (SFS) algorithms. SFS research typically assumes a constant albedo across an object which is usually not true and thus limits the use of the approach. Since then, there have been several advances which have led to the application of SFS for face recognition and rendering. Zhao *et al.* [14] present an SFS approach to recover both shape and albedo for a symmetric object from a single image. [12] uses singular value decomposition (SVD) to learn generative models of objects from a set of images taken under different, and unknown illuminations. Shashua *et al.* [9] perform recognition across varying illumination under an ideal-class assumption. All objects belonging to the ideal class are assumed to have the same shape. [5] uses illumination cone models for illumination-invariant face recognition. They require a small number of training images of each face under different illuminations to recover the shape and albedo of the face. Basri *et al.* [1] propose methods for recovering surface normals in a scene using images taken under general illumination conditions. Their work is based on [2], [7] which prove that the set of all Lambertian reflectance maps obtained with arbitrary distant illumination sources approximately lie in a 9D linear subspace. In [4], Blanz *et al.* perform face recognition across pose and illumination by fitting a 3D morphable model to images. They use a set of textured 3D scans of heads for learning the model. [13] uses harmonic image exemplars to perform face recognition under varying lighting. Zhou *et al.* [16] generalize the traditional photometric approach to handle all the appearances of all the objects in a class. They impose a rank constraint on shapes and albedos in a class to separate the two from illumination using the factorization approach.

*Partially supported by an NSF-ITR Grant 03-25119

Despite the advances made, most of the cited approaches have not been applied for the face recognition problem using a large database. This might be because many techniques require multiple, differently illuminated images of each face which are usually not present in most face datasets. Moreover, most of the approaches make single light source assumption which does not hold in most real conditions. The assumption might be driven by the unavailability of suitable datasets to test approaches which can potentially handle images illuminated by multiple number of light sources.

In this paper, we propose an algorithm to perform face recognition across varying illumination (using a single face image) for images illuminated by multiple number of light sources. The algorithm does not need any prior information about the number or placements of the light sources. We are not aware of any standard controlled dataset containing faces illuminated by two or more number of light sources, which can be used to study the effect of the single light source assumption on face recognition algorithms. We generate such data using faces from PIE [10] and Yale Face Database B [5]. Experimental results are presented to confirm the efficacy of the approach. The proposed algorithm performs much better than its counterpart that makes the single light source assumption.

The algorithm models a face as a Lambertian surface. Therefore, it is worthwhile to address the issue of the often ignored nonlinearity in the Lambert's law before explaining the algorithm. In fact the performance of the algorithm can get adversely affected if the nonlinearity is ignored (thereby allowing pixel intensities to take negative values). Section 2 illustrates the importance of the nonlinearity. The importance is further highlighted in Section 3 by showing an improvement in the recognition accuracy when the nonlinearity is taken into account. The face recognition algorithm to handle faces illuminated by multiple light sources is presented in Section 4. Section 5 describes several challenging experiments performed to rigorously test the approach.

2. How Important is the Non-linearity in the Lambert's Law?

The diffuse component of the reflection of a surface is often modeled using the popular Lambert's Law. For example, Blanz and Vetter[4] use the following equation to model the diffuse component

$$L_{r,k} = R_k . L_{r,dir} . < n_k, l >, \quad (1)$$

where R_k is the red component of the diffuse reflection coefficient, $L_{r,dir}$ is the red channel of the directed light, n_k is the surface normal and l is the light source direction. Similarly, the generalized photometric stereo method [16] uses

$$h = \rho n^T s, \quad (2)$$

where ρ is the surface albedo, n is the surface normal, and s is the light source direction (multiplied by the intensity), as the rule for image formation. A close look at these equations reveals that the Lambert's law is assumed to be linear in both these models. If used in its pure form, the nonlinearity in the Lambert's law would have made (2) to be

$$h = \rho \max(n^T s, 0) \quad (3)$$

Quite clearly, the linearity assumption is perfectly valid as long as the directed light source is in front of the surface for all its points. In general, objects like faces do not have all the surface points facing the illumination source which leads to the formation of shadows (commonly known as form/attached shadows). The cast and attached shadows are often ignored from the analysis to keep the subspace of the observed images in a three [8] or with the addition of an ambient component [12], four dimensional linear subspace. Therefore, several generative approaches either ignore this non-linearity completely or try to somehow ignore the shadow pixels. Here we present a simple illustration to highlight the role attached shadows can play which can be modeled by incorporating the nonlinearity.

2.1. Illustration 1

Suppose the goal is to estimate the illumination source from a single face image given the shape and albedo of the face. We explore three approaches for this task: the first approach ignores the non-linearity completely, the second one uses the linear rule but ignores the shadow pixels and the last one uses the Lambert's Law in its pure form. The accuracy of the global minimum and its ambiguity on the error surface is taken as the criterion for the goodness of the method. The analytical expressions for the error function using the three options can be written as :

Completely linear: $$\varepsilon(s) = \| h - \rho n^T s \|^2 \quad (4)$$

Shadow pixels ignored: $$\varepsilon(s) = \| \tau \circ (h - \rho n^T s) \|^2 \quad (5)$$

Non-linear rule: $$\varepsilon(s) = \| h - \max(\rho n^T s, 0) \|^2 \quad (6)$$

where, $\varepsilon(s)$ is the error with s as the illumination source direction, $h_{d\times 1}$ is the vectorized input image, ρ is the albedo vector, $n_{3\times d}$ contains the surface normals, and $\tau_{d\times 1}$ is the shadow indicator vector which is 0 for the shadow pixels and 1 for the rest. Clearly, the linear method penalizes the correct illumination at the shadow pixels by having non-zero error values for those pixels. On the other hand, when shadows are ignored, the illuminations which produce wrong values for the shadow pixels do not get penalized there. As the set of all possible normals lies on the surface of a unit sphere, we use a sphere to display the computed error functions. Figure 1 shows the error surfaces for

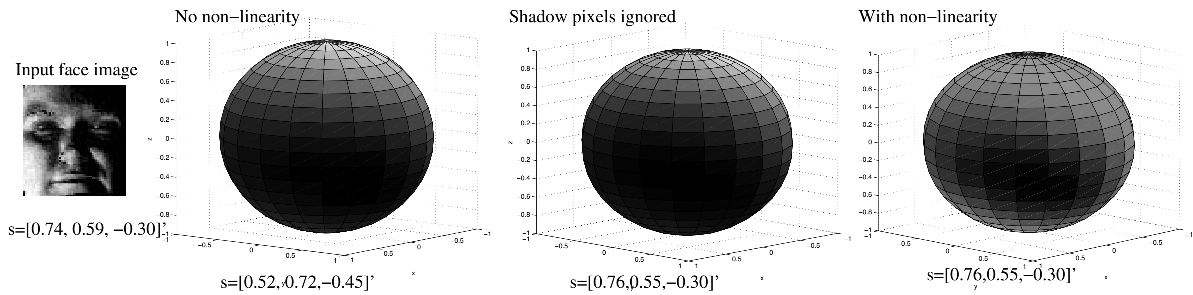

Figure 1. The error surfaces for the estimation of the light source direction for a given face image. The plots correspond to the three approaches described in Illustration 1. The lower the error is for a particular illumination direction, the darker the error sphere looks at the point corresponding to that direction. The true and estimated values of the illumination direction are listed along with the plots.

the three methods for a given face image. The lower the error is for a hypothesized illumination direction s, the darker the surface looks at the corresponding point on the sphere. The global minimum is far from the true value using the first approach but is correct up to a discretization error for the second and third approaches. In fact, the second and third methods will always produce the same global minimum (assuming τ is correct), but the global minimum will always be less ambiguous in the third case because several wrong hypothesized illumination directions do not get penalized enough in the second approach due to the exclusion of the shadow pixels (Figure 1) .

2.2. The Case of Multiple Light Sources

The above analysis implicitly assumes that there is only one distant light source illuminating the face. Though the assumption is valid for datasets like PIE, it does not hold for most realistic scenarios. We now explore the impact of using the *linear* Lambert's law for images illuminated by multiple light sources. Using the *linear* Lambert's law, an image illuminated by k light sources can be represented as:

$$h = \sum_{i=1}^{k} \rho n^T s_i = \rho n^T \sum_{i=1}^{k} s_i = \rho n^T s^\star \tag{7}$$

where, $s^\star = \sum_{i=1}^{k} s_i$. This shows that under the linear assumption, multiple light sources can be replaced by a suitably placed single light source without having any effect on the image. This is a bit counter-intuitive as can be seen in a simple two source scenario where $s_1 = -s_2$

$$h = \rho n^T (s_1 - s_2) = 0 \tag{8}$$

Thus the linear assumption can make the effect of light sources interfere in a destructive manner and give bizarre outcomes. Quite clearly, the harm done by the linearity assumption is proportional to the angle subtended by the light sources at the surface.

Though the above discussion concludes that the Lambert's law in its pure form is better suited for illumination estimation than the other variants, it is only of academic interest if inclusion of the non-linearity does not improve the recognition results. The following section proposes a variant of the generalized photometric approach taking the non-linearity into account. The improvement in the recognition accuracy highlights the importance of including the attached shadows in the analysis.

3. Face Recognition across Varying Illumination (single light source)

We extend the approach proposed in [16]. We first present a quick overview of the method and then highlight the impact of the nonlinearity on the approach. Using Lambert's law (the linear version), the intensity of a pixel can be written in terms of its albedo, shape and an illumination source as

$$h_i = (\rho n^T)_i s = t_i^T s \tag{9}$$

Suppose the image has d pixels, then

$$h_{d\times 1} = [h_1, h_2, \ldots, h_d]^T = T_{d\times 3}.s_{3\times 1} \tag{10}$$

where, $T = [t_1 t_2 \ldots t_d]^T$ is the object specific shape-albedo matrix. If T can be represented as a linear combination of m basis $T_i's$, we get

$$\begin{aligned} T &= f_1 T_1 + f_2 T_2 + ... + f_m T_m \\ &= [T_1 T_2 ... T_m](f \otimes I_3) \\ &= W(f \otimes I_3) \end{aligned} \tag{11}$$

where $W = [T_1 T_2 \ldots T_m]$ is the class specific shape-albedo matrix and I_3 is the 3×3 identity matrix. In this formulation, vector $f = [f_1 f_2 \ldots f_m]^T$ is treated as the illumination-free identity vector. Given n different objects under different (and unknown) illumination conditions, Zhou *et al.* [16] estimate W (up to an invertible matrix) by solving a rank $3m$ problem using the factorization

Table 1. Recognition results on the PIE dataset. The averages from [16] are included for comparison. f_i denotes images taken with a particular flash ON as labeled in PIE. Each $(i,j)^{th}$ entry in the table shows the recognition rate obtained with the images from f_j as gallery while from f_i as probes.

Gallery	f_{08}	f_{09}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{17}	f_{20}	f_{21}	f_{22}	Average	Average from [16]
Probe														
f_{08}	-	100	100	100	96	97	81	72	50	100	97	84	90	88
f_{09}	100	-	100	100	100	99	97	96	75	100	100	97	97	94
f_{11}	100	100	-	100	100	97	94	78	63	100	99	94	94	93
f_{12}	100	100	100	-	100	100	100	99	90	100	100	100	99	97
f_{13}	97	100	100	100	-	100	100	100	96	100	100	100	99	99
f_{14}	94	100	100	100	100	-	100	100	99	100	100	100	99	99
f_{15}	88	97	97	100	100	100	-	100	100	97	100	100	98	96
f_{16}	74	90	81	93	100	100	100	-	100	76	97	100	93	89
f_{17}	59	74	63	87	99	99	100	100	-	71	94	100	87	75
f_{20}	99	100	100	100	100	99	96	82	71	-	100	97	95	93
f_{21}	97	100	100	100	100	100	100	99	96	100	-	100	99	98
f_{22}	93	100	99	100	100	100	100	100	99	99	100	-	99	98
Average	92	97	95	98	100	99	97	94	87	95	99	98	96	-
Average from [16]	89	93	92	96	98	99	96	91	80	91	96	98	-	93

approach. The ambiguity is resolved using symmetry and integrability constraints. The interested reader is referred to [16] for the complete derivation. The average recognition results reported in [16] improve from 67% to 93%, when the W matrix is estimated from Vetter's 3D data [4] instead of the approach mentioned above.

Our main focus here is to highlight the importance of the non-linearity in the Lambert's law and not generalized photometric stereo. Therefore, we generate the shape-albedo matrix W using Vetter's 3D data for all our experiments. As opposed to [16], we take into account the inherent hard non-linearity present in the Lambert's law. Given shape-albedo matrix W, the recovery of the identity vector f and illumination s can be posed as an optimization problem as follows:

$$\min_{f,s} \quad \varepsilon(f,s) \equiv \| h - h_{rec} \|^2 + (1^T f - 1)^2 \qquad (12)$$

$$\text{where,} \qquad h_{rec} = \sum_{i=1}^{m} f_i \max(T_i s, 0) \qquad (13)$$

The second term is included in the error function to take care of scale ambiguity between f and s. Please note that s is not a unit vector as it contains the intensity of the illumination source also. The minimization is performed using an iterative approach, fixing f for optimizing ε w.r.t. s and fixing s for optimization w.r.t. f. In each iteration, f can be estimated by solving a linear least-squares (LS) problem but a non-linear LS solution is required to estimate s. The non-linear optimization is performed using the *lsqnonlin* function in MATLAB which is based on the interior-reflective Newton method. For most faces, the function value did not change much after 4-5 iterations. Therefore, the iterative optimization was always stopped after 5 iterations. The whole process took about 5-7 seconds per image on a normal desktop.

We perform recognition experiments across illumination using the frontal faces from the PIE dataset. The correlation coefficient of the identity vectors is taken as the measure of the similarity between face images. Table 1 shows the recognition results obtained using this approach. Recognition is performed across illumination with images from one illumination condition from the PIE dataset forming the gallery while images from another illumination condition forming the probe set. Each gallery/probe set contains one frontal image per subject taken in the presence of a particular light source (there are 68 subjects in each gallery/probe set). Each entry in the table shows the recognition rate achieved for one such choice of gallery and probe set. The averages from [16] are shown for comparison. For fair comparison, we show results only across the illumination scenarios displayed in [16]. The recognition performance with the inclusion of the non-linearity in the Lambert's law is almost always better or same. The overall average performance is up from 93% to 96%. The improvement is significant in cases involving difficult illumination conditions (with lots of shadows) like the flash f_{17} in the PIE dataset. This shows that though the estimation becomes slightly more difficult, the recognition rate improves with the inclusion of the non-linearity.

4. Illumination-Invariant Face Recognition in the Presence of Multiple Light Sources

One of the issues in handling multiple illumination case is the prior knowledge of the number of light sources. In the absence of this knowledge, one can hypothesize several different cases and choose the one with minimum residual error. This can be done in a manner very similar to the approach described for single illumination case with the following change in the objective function:

$$\varepsilon(f, s) = \| h - h_{rec} \|^2 + (1^T f - 1)^2 \tag{14}$$

$$\text{where,} \qquad h_{rec} = \sum_{i=1}^{m} f_i \sum_{j=1}^{k} \max(T_i s_j, 0) \tag{15}$$

where k is the hypothesized number of light sources. The objective function can be minimized repeatedly for different values of k and the one with minimum error can be taken as the correct hypothesis. Figure 2 shows the variation of the error with k, for an image illuminated by three different light sources. As can be seen, the error more or less stabilizes for $k \geq 3$. As the parameter spaces are nested, ideally the error plot should be a non-increasing function of k, but the increase in complexity of the non-linear optimization can make the plot behave otherwise. Please note that for the *linear* Lambert's law, such a curve will look more or less horizontal due to the equivalence of the single and multi-light source scenarios (Equation 7) under the linear assumption.

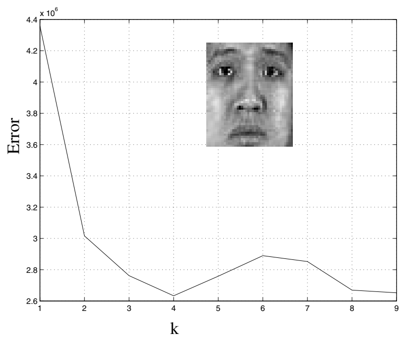

Figure 2. The error obtained for different hypothesized number of light sources. The face was illuminated using 3 light sources.

Though one can use this approach by varying k, it is both inelegant and computationally intensive. In our approach, we avoid the extra computations by making the following assumption. We assume that an image of an arbitrarily illuminated face can be approximated by a linear combination of the images of the same face in the same pose, illuminated by nine different light sources placed at pre-selected positions. Lee *et al.* [6] show that this approximation is quite good for a wide range of illumination conditions. Hence, a face image can be written as

$$h = \sum_{i=1}^{9} \alpha_i h_i \tag{16}$$

$$\text{where,} \qquad h_i = \max(\rho n^T s_i, 0) \tag{17}$$

$\{s_1, s_2, \ldots, s_9\}$ are the pre-specified illumination directions. As proposed in [6], we use the following directions for $\{s_1, s_2, \ldots, s_9\}$:

$$\begin{aligned} \phi &= \{0, 49, -68, 73, 77, -84, -84, 82, -50\}^\circ \\ \theta &= \{0, 17, 0, -18, 37, 47, -47, -56, -84\}^\circ \end{aligned} \tag{18}$$

Under this formulation, h_{rec} in (15) changes to

$$h_{rec} = \sum_{i=1}^{m} f_i \sum_{j=1}^{9} \alpha_j \max(T_i s_j, 0) \tag{19}$$

This way one can potentially recover the illumination-free identity vector f without any prior knowledge of the number of light sources or any need to check different hypotheses for the same.

Now the objective function is minimized with respect to $f = [f_1 f_2 \ldots f_m]$ and $\alpha = [\alpha_1 \alpha_2 \ldots \alpha_9]$. This gives us the illumination-free identity vector f which is used for recognition. The optimization is done in an iterative fashion by fixing one parameter and estimating the other and vice-versa as shown below:

$$h_{rec} = \left[\sum_{j=1}^{9} \alpha_j \max(T_1 s_j, 0) \quad \sum_{j=1}^{9} \alpha_j \max(T_2 s_j, 0) \ldots \right.$$

$$\left. \cdots \sum_{j=1}^{9} \alpha_j \max(T_m s_j, 0) \right]_{d \times m} f_{m \times 1} = G_{d \times m} f_{m \times 1} \tag{20}$$

Therefore,

$$f = A^\dagger \begin{bmatrix} h \\ 1 \end{bmatrix} \tag{21}$$

where $h_{d \times 1}$ is the vectorized input face image and $A^\dagger$ is the Moore-Penrose inverse of $A = \begin{bmatrix} G_{d \times m} \\ 1_{1 \times m} \end{bmatrix}_{(d+1) \times m}$. $1_{1 \times m}$ is the m-dimensional vector of ones, included to handle scale ambiguity between f and α.

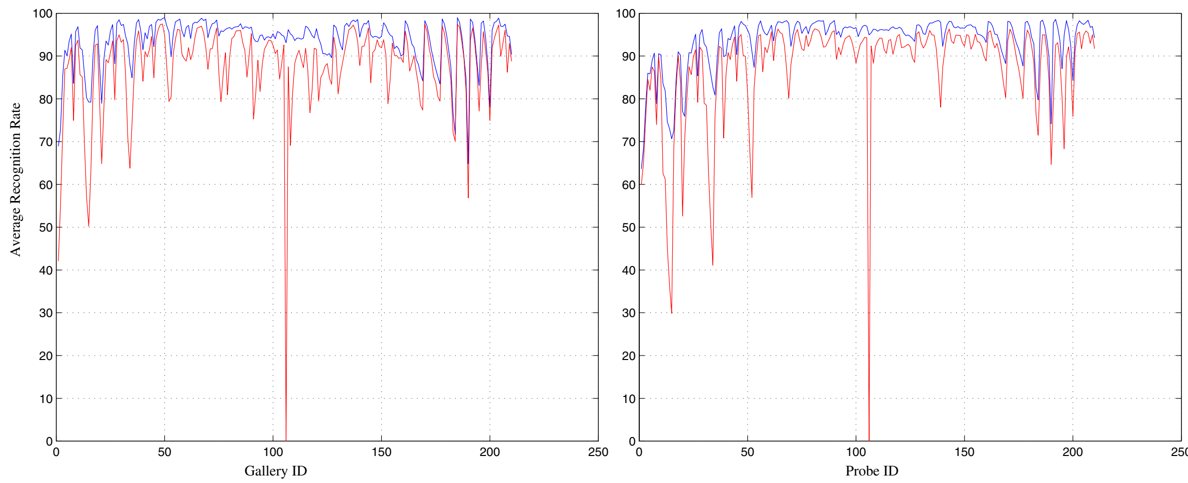

Figure 3. The per-gallery and per-probe set average recognition rates on the 210 doubly-illuminated scenarios generated from the PIE dataset. The blue curve shows the performance of the proposed approach while the red curve corresponds to the linear single light source approach [16].

Looking carefully at the objective function ((14) and (19)), one can easily observe that α too can be estimated by solving a linear LS problem (as $\{s_1 s_2 \dots s_9\}$ is known). This avoids the need for any nonlinear optimization here. Please recall that nonlinear LS was required to estimate s in the approach proposed for the single light source case. The expression for α can be written as:

$$\alpha = B^{\dagger} h \tag{22}$$

where,

$$B = \left[\sum_{i=1}^{m} f_i \max(T_i s_1, 0) \quad \sum_{i=1}^{m} f_i \max(T_i s_2, 0) \dots \right.$$
$$\left. \cdots \sum_{i=1}^{m} f_i \max(T_i s_9, 0) \right]_{d \times 9} \tag{23}$$

For most of the face images, the iterative optimization converged within 5-6 iterations. As there is no non-linear optimization involved, it took just 2-3 seconds to recover f and α from a given face image on a normal desktop. As the identity variable is estimated from an image by separating the effect of all the light sources in the form of α, it is used as the illumination-invariant representation for recognition across varying illumination. The correlation coefficient of the identity vectors is used as the similarity measure for recognition experiments.

5. Experiments and Results

To begin with, we test this algorithm by running the same experiment as we do for the single light source approach. Though the PIE dataset is not suited to test the ability of this algorithm to handle arbitrarily illuminated images, a good performance here can be considered as a proof of concept. The overall average recognition rate for the experiment obtained using this algorithm is 95% which is higher than [16].

Due to the unavailability of a standard dataset containing face images with multiple light sources ON at a time, we generate such data using the PIE and Yale datasets. Due to the controlled nature of the datasets, multiple images of a subject under different illuminations but same pose, are more or less aligned. If we ignore any camera gain, this allows us to add multiple images of a person taken under different illuminations to get one with the effect of an image captured with multiple lights ON. The images generated this way look pretty realistic (see Figure 4).

We perform experiments on the dataset created by adding images from two illumination conditions from PIE at a time. As PIE has 21 different illumination scenarios, we get a total of ${}^{21}C_2 = 210$ different *doubly-illuminated* scenarios. Recognition was done across all 210 scenarios by taking one as the gallery and another one as the probe set at a time to get 210×209 recognition scores. As it is difficult to show the recognition scores by drawing a 210×210 table, we show only the aggregated per-gallery and per-probe set recognition rates (similar to the averages in Table 1) in Figure 3. The blue curve on the top shows the averages obtained by the proposed approach. For comparison, we show the recognition rates obtained on this dataset using Zhou *et al.*'s linear method [16] that ignores shadow pixels under the single light source assumption (red curve). For ease of use, we will call this method as ISP-SLS (Ignores Shadow Pixels under Single Light Source Assumption). There exist a zero in the red curve because for one gallery/probe set, the method ended up ignoring most of the pixels as shadows and thus was unable to recover the identity variable. The recognition rates obtained using the proposed approach

Figure 4. The doubly-illuminated images of a subject from the Yale database. Each image is generated by adding 2 images of the same subject illuminated by different light sources.

Figure 5. The 6 illumination conditions from the Yale face Database B used to generate the doubly-illuminated data.

are always better or same as compared to ISP-SLS. The increase in the recognition accuracy is more prominent for the cases where the two illumination sources combined to generate the doubly-illuminated scenario were far apart. This happens because the destructive interference of two light sources (due to the linearity assumption in ISP-SLS as described in Section 2) increases with an increase in the angle between the two.

We further test the algorithm by generating a similar *doubly-illuminated* data using Yale Face Database B [5]. Figure 5 shows the six challenging illumination conditions used to generate fifteen different scenarios (shown in Figure 4) by pairing two at a time. The average recognition rate achieved on this difficult data (Figure 4 shows images of one subject under the 15 illumination conditions) using our algorithm is 77%. This is up by more that 25% compared to the accuracy achieved both by ISP-SLS method and the method which takes the non-linearity into account under the single light source assumption.

Figure 6. The reconstructed shapes of a face using the ISP-SLS approach. Each column displays the 3 components of the reconstructed surface normals. Columns 1-5 correspond to the five illumination scenarios with the number of light sources varying from 1-5, respectively. The quality of the reconstructed surface degrades as the number of light sources increase.

Figure 7. The reconstructed shapes of a face using our approach. As in Figure 6, each column shows the 3 components of the reconstructed surface normals. There is hardly any difference in the reconstructed surfaces across different illuminations scenarios.

All the above experiments implicitly assume that the faces in the gallery and probe set are illuminated by the same number of light sources. Clearly, the proposed algorithm does not impose any such restriction. Therefore, we perform another experiment to test the ability of the proposed approach to generalize across varying number of light sources. We generate five illumination scenarios using the PIE dataset with the number of light sources (added to create each scenario) ranging from 1-5. To avoid any bias, the combinations of the light sources are selected randomly from the 21 illumination sets in the PIE dataset. Recognition is performed across the five scenarios by considering one among them as the gallery and another one as the probe set at a time. As before, each gallery/probe set contains one image for each of the 68 subjects present in the PIE dataset. While the ISP-SLS approach performs badly in this experiment, the proposed approach does a perfect job as shown in Table 2. Figures 6 and 7 show the reconstructed surfaces for a face illuminated in the presence of the five illumination scenarios using the two approaches. The quality of the reconstructions explains the difference in the recognition accuracy obtained using the two methods. To confirm the authenticity of the results, we perform another similar experiment with 10 different scenarios with the number of randomly selected light sources (added to generate the 10 scenarios) ranging from 1-10. Here, the proposed approach achieves average recognition accuracy of 99.7% (The average recognition rate achieved by ISP-SLS here is 54%).

Table 2. Recognition results on the multiply-illuminated data generated from the PIE dataset. The various scenarios differ in the number of light sources. The flash Ids from PIE randomly selected to generate each scenario are shown in curly braces. The 1st number shows the recognition rate obtained using our approach while the 2nd number shows the performance of the ISP-SLS method.

Gallery	$\{f_{20}\}$	$\{f_{05}, f_{22}\}$	$\{f_{20}, f_{06}, f_{18}\}$	$\{f_{21}, f_{06}, f_{07}, f_{03}\}$	$\{f_{03}, f_{15}, f_{06}, f_{19}, f_{05}\}$
Probe					
$\{f_{20}\}$	- / -	100 / 100	100 / 66	100 / 26	100 / 26
$\{f_{05}, f_{22}\}$	100 / 100	- / -	100 / 62	100 / 28	100 / 25
$\{f_{20}, f_{06}, f_{18}\}$	100 / 93	100 / 91	- / -	100 / 72	100 / 74
$\{f_{21}, f_{06}, f_{07}, f_{03}\}$	100 / 62	100 / 66	100 / 90	- / -	100 / 93
$\{f_{03}, f_{15}, f_{06}, f_{19}, f_{05}\}$	100 / 66	100 / 66	100 / 93	100 / 93	- / -

6. Summary and Conclusions

In this paper, we have proposed an algorithm to recognize *multiply*-illuminated faces from just one image. The algorithm has the capability to generalize across face images taken in the presence of varying number of unknown light sources. The approach performs well and outperforms the ISP-SLS approach which uses the linear version of the Lambert's law. Though the comparison with the ISP-SLS approach might not seem fair as it does not try to model multiple light sources, the comparison does reflect the effect, the single light source assumption might have, for recognizing faces in real conditions. Moreover, we illustrated that if the Lambert's law is assumed to be linear, the single and multiple light scenarios are equivalent. Therefore, one can infer that the relaxation of the non-linearity in the Lambert's law is harmful for recognizing faces illuminated by an arbitrary number of light sources. Almost perfect performance of the proposed approach in experiments involving large number of light sources conforms to the belief that face recognition gets easier with an increase in number of light sources.

References

[1] R. Basri and D. Jacobs. Photometric stereo with general, unknown lighting. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 374–381, 2001.

[2] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25:218–233, 2003.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19:711–720, 1997.

[4] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, September 2003.

[5] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):643–660, June 2001.

[6] K. C. Lee, J. Ho, and D. Kriegman. Nine points of light: acquiring subspaces for face recognition under variable lighting. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 519–526, December 2001.

[7] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of convex Lambertian object. *Journal of the Optical Society of America A*, pages 2448–2459, October 2001.

[8] A. Shashua. On photometric issues in 3d visual recognition from a single 2d image. *International Journal of Computer Vision*, 21:99–122, 1997.

[9] A. Shashua and T. R. Raviv. The quotient image: Class-based re-rendering and recognition with varying illuminations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(2):129–139, February 2001.

[10] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression database. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(12):1615–1618, December 2003.

[11] K. Sung and T. Poggio. Example-based learning for view-based human face detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20:39–51, 1997.

[12] A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur. Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability. *International Journal of Computer Vision*, 35(3):203–222, 1999.

[13] L. Zhang and D. Samaras. Face recognition under variable lighting using harmonic image exemplars. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 19–25, 2003.

[14] W. Zhao and R. Chellappa. Symmetric shape from shading using self-ratio image. *International Journal of Computer Vision*, 45(1):55–75, October 2001.

[15] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.

[16] S. Zhou, R. Chellappa, and D. Jacobs. Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints. In *European Conference on Computer Vision*, 2004.