Face Recognition in the Presence of Multiple Illumination Sources

Gaurav Aggarwal and Rama Chellappa

Center for Automation Research, University of Maryland, College Park, Maryland, USA.

Abstract

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. 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.

How Important is the Nonlinearity in the Lambert's Law?

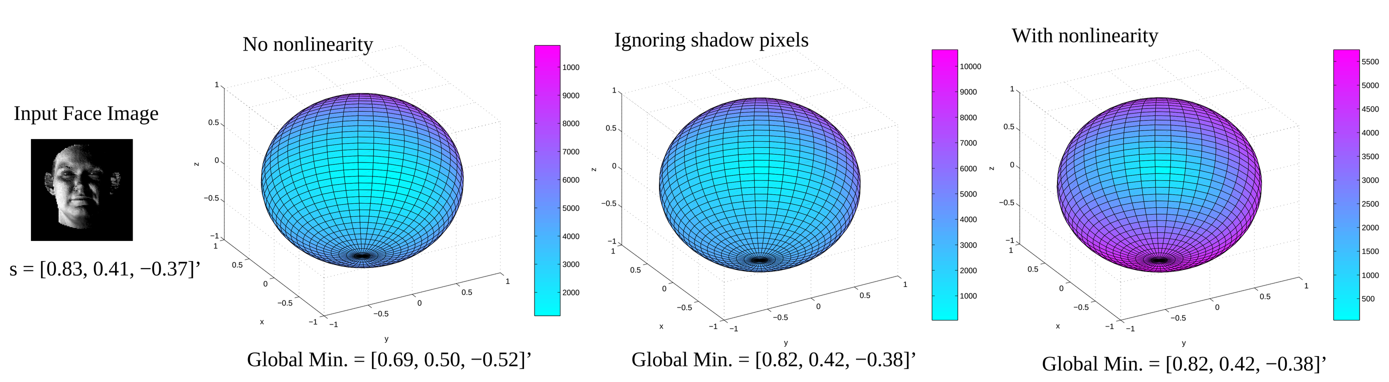
- The nonlinearity accounts for the formation of attached shadows.
- Ignoring attached shadows limits the dimensionality of the subspace of the observed images.
- **Illustration:** Estimate illumination direction from an image given the shape and albedo.
- Three approaches:

Completely linear: $$\varepsilon(s) = \| h - \rho n^T s \|^2 \quad (1)$$

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

Non-linear rule: $$\varepsilon(s) = \| h - \max(\rho n^T s, 0) \|^2 \quad (3)$$

- The accuracy of the global minimum and its ambiguity on the error surface(sphere) is taken as the criterion for the goodness of the method.



Using the *linear* Lambert's law, the image of an object illuminated by k light sources can be written as:

$$h = \sum_{i=1}^{k} \rho n^T s_i = \rho n^T \sum_{i=1}^{k} s_i = \rho n^T s^\star \quad (4)$$

What happens when $k = 2$ and $s1 = -s2$?

Face Recognition: Single Light Source Scenario (1)

Using the Lambert's Law,

$$h_i = (\rho n^T)_i s = t_i^T s \quad (5)$$

Suppose the image has d pixels, then

$$h_{d\times 1} = [h_1, h_2, \ldots, h_d]^T = T_{d\times 3} . s_{3\times 1} \quad (6)$$

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

$$\begin{aligned} T &= f_1 T_1 + f_2 T_2 + \ldots + f_m T_m \\ &= [T_1 T_2 \ldots T_m](f \otimes I_3) \\ &= W (f \otimes I_3) \end{aligned} \quad (7)$$

- W is the class specific (here, face) shape-albedo matrix.
- f contains illumination-free shape and albedo information.

Face Recognition: Single Light Source Scenario (2)

- We generate the shape-albedo matrix W using 3D data for all our experiments.
- Given W, the recovery of the identity vector f and illumination s can be posed as the following optimization problem:

$$\min_{f,s} \ \varepsilon(f, s) \equiv \| h - h_{rec} \|^2 + (1^T f - 1)^2 \quad (8)$$

$$\text{where,} \quad h_{rec} = \sum_{i=1}^{m} f_i \max(T_i s, 0) \quad (9)$$

- Minimization performed using an iterative approach (Linear LS + Nonlinear LS)

Recognition Results on PIE (comparison with the linear approach)

Gallery / Probe	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 Zhou *et al.*
f_{08}	-	100	100	100	96	97	81	72	43/50	100	97	84	90	88
f_{09}	100	-	100	100	100	99	97	96	60/75	100	100	97	97	94
f_{11}	100	100	-	100	100	97	94	78	65/63	100	99	94	94	93
f_{12}	100	100	100	-	100	100	100	99	76/90	100	100	100	99	97
f_{13}	97	100	100	100	-	100	100	100	88/96	100	100	100	99	99
f_{14}	94	100	100	100	100	-	100	100	96/99	100	100	100	99	99
f_{15}	88	97	97	100	100	100	-	100	100/100	97	100	100	98	96
f_{16}	74	90	81	93	100	100	100	-	100/100	76	97	100	93	89
f_{17}	$\frac{44}{59}$	$\frac{60}{74}$	$\frac{51}{63}$	$\frac{71}{87}$	$\frac{84}{99}$	$\frac{91}{99}$	$\frac{99}{100}$	$\frac{100}{100}$	-/-	$\frac{56}{71}$	$\frac{75}{94}$	$\frac{94}{100}$	87	75
f_{20}	99	100	100	100	100	99	96	82	68/71	-	100	97	95	93
f_{21}	97	100	100	100	100	100	100	99	82/96	100	-	100	99	98
f_{22}	93	100	99	100	100	100	100	100	99/99	99	100	-	99	98
Average	92	97	95	98	100	99	97	94	87	95	99	98	96	-
Average from Zhou *et al.*	89	93	92	96	98	99	96	91	80	91	96	98	-	93

Face Recognition: Multiple Light Sources Scenario (1)

- Naive approach:

$$\min_{f,s,k} \varepsilon(f, s) = \| h - h_{rec} \|^2 + (1^T f - 1)^2 \quad (10)$$

$$\text{where,} \quad h_{rec} = \sum_{i=1}^{m} f_i \sum_{j=1}^{k} \max(T_i s_j, 0) \quad (11)$$

- k is the hypothesized number of light sources.
- Disadvantage: Computationally intensive and the optimization problem becomes more and more intractable as k increases.
- Our assumption: 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.*)

$$h = \sum_{i=1}^{9} \alpha_i h_i \quad \text{where,} \quad h_i = \max(\rho n^T s_i, 0) \quad (12)$$

$\{s_1, s_2, \ldots, s_9\}$ are the pre-specified illumination directions. Therefore, we can write

$$h_{rec} = \sum_{i=1}^{m} f_i \sum_{j=1}^{9} \alpha_j \max(T_i s_j, 0) \quad (13)$$

Face Recognition: Multiple Light Sources Scenario (2)

- f and α are estimated by iterative minimization (Linear LS + Linear LS) as follows:

$$f = \begin{bmatrix} G_{d\times m} \\ 1_{1\times m} \end{bmatrix}^{\dagger}_{(d+1)\times m} \begin{bmatrix} h \\ 1 \end{bmatrix}; \qquad \alpha = B^{\dagger} h \quad \text{where,} \quad (14)$$

$$G_{d\times m} = \left[\sum_{j=1}^{9} \alpha_j \max(T_1 s_j, 0) \ \sum_{j=1}^{9} \alpha_j \max(T_2 s_j, 0) \cdots \sum_{j=1}^{9} \alpha_j \max(T_m s_j, 0) \right]_{d\times m} \quad (15)$$

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

Recognition Results and Comparison

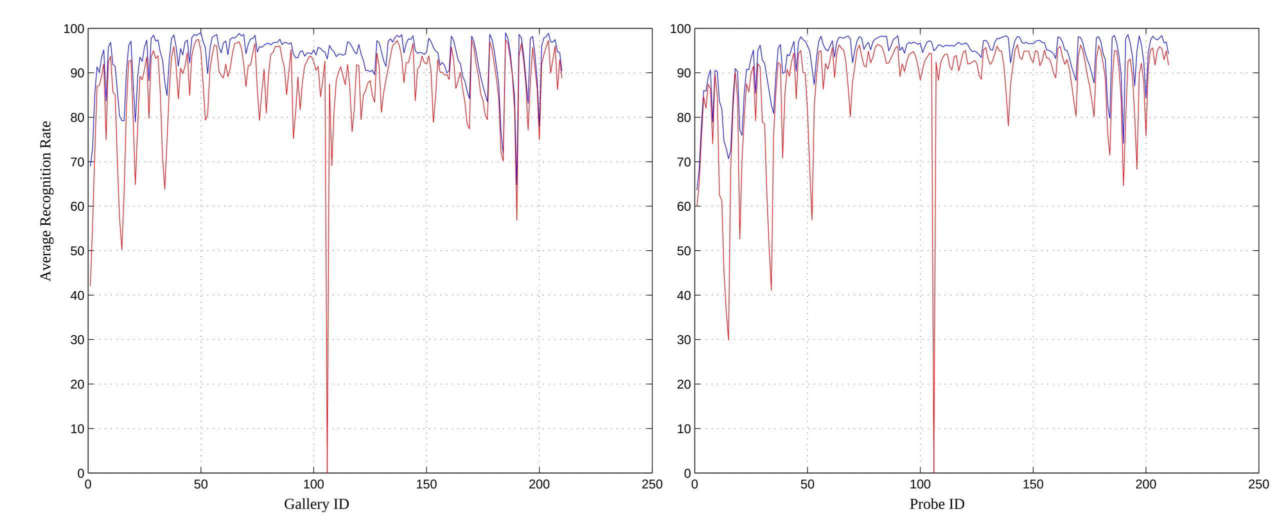


Figure 1: The per-gallery and per-probe set average recognition rates on the 210 doubly-illuminated scenarios generated from the PIE dataset.

Figure 2: The doubly-illuminated yale face data generated using 6 bad illumination conditions. The proposed algorithm achieves 77% recognition accuracy which is up by more than 25% compared to that of Zhou *et al.*'s single light source approach.

Table 1: Recognition results on the multiply-illuminated data generated from the PIE dataset-the various scenarios differ in the number of light sources.

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

Figure 3: Comparison of the reconstructed shapes: Each column displays the 3 components of the reconstructed surface normals. The number of light sources vary from 1 to 5 across the columns.

- In a similar experiment, with the number of randomly selected light sources varying from 1-10, the proposed approach achieves 99.7% recognition rate (Zhou *et al.*'s single light source approach achieves 54% accuracy in this experiment).