On Illumination Effects in PCA-Based Vision Systems

Li Zhao Yee-Hong Yang
Computer Vision & Graphics Lab
Department of Computer Science
University of Saskatchewan
Saskatoon
CANADA, S7N 5A9

Abstract

In this paper, a theoretical method to analyse the illumination effects in principal component analysis (PCA) based vision systems is presented.

In the first part, an analytical closed-formed formula is given to compute the covariance matrix for PCA-based vision systems. This method is fundamentally different from previous work, in a sense, in that an infinite number of images (include every possible image) are used to computed the covariance matrix.

In the second part, a framework is developed to account for the changes of light source type, number, intensity, and other illumination effects such as interreflection, shadow and specular highlight. Finally an equation called the illumination equation is derived.

Two thorough experiments are conducted to test the proposed method. Since the major concern in this paper is representation instead of recognition, the quality of reconstructed images and the examination of residual images are used as the main measurement for the experiments. Experimental results confirm that the new method is able to capture all these illumination effects satisfyingly.

This is research is fundamental different from all the former research in many places. The comparison between the proposed method and former methods is given at the beginning of the paper.

It is believed that this is the best method to account for different illumination effects using only three sample images.

Keywords: illumination effects, interreflection, light intensity change, light sources, Mosaic Image Method, principal component analysis, radiosity method.

1 Introduction

Illumination changes can drastically alter the appearances of objects. Topics related to illumination such as shape from shading, and photometric stereo are well studied in computer vision [7]. However, research on illumination effects in PCA-based vision systems is mainly conducted in an empirical manner. Some theoretical research is mainly for recognition (which is a high level vision problem) [2]. For methods to compute representations which account for lighting changes, the current practice is simply to take sample images in different lighting conditions.

In this paper, a theoretical analysis on illumination effects in the context of PCA-based vision systems is given. This analysis takes into account every illumination effect such as changes of light source type, number, and intensity and other illumination effects such as interreflection, shadow and specular highlight. It gives a new formulation to account for one global illumination effect, interreflection. All these analyses are based on first principles. Therefore they all have clear meanings. Furthermore, since the method to account for lighting intensity change gets its inspiration from observation of human vision systems, therefore it may also have implications for human vision systems.

The organization of this paper is as follows. In the next section, a review of the related research is given. The differences between former methods and the proposed method are pointed out and discussed. In the following section, a closed-form formula to compute covariance matrix in PCA-based vision systems is established. This gives a new way to compute the representation. Then a method to accommodate changes of light source type, number, intensity and interreflection is presented. This method gives an equation to account for these illumination effects based on the new representation. Experimental results are given to show the advantages of the proposed method. The paper concludes with a summary and a discussion on further research.

2 Comparison with Former Research

In [12], Shashua gives a method to deal with illumination changes and viewing position changes. He starts with a generalized K Linear Reflection Model. This generalization is not needed since it adds no further meaning to his research. The research presented in this paper also uses a similar starting theorem, which is derived from Horn's [7] classical treatment of Lambertian surface and simple linear algebra theories. After the K Linear Reflection Model, Shashua does nothing more than directly gives a linear alignment method for recognition. In contrast, the approach presented in this paper is mainly concerned with representation rather than recognition. Besides, the research presented here proceeds in the context of principal component analysis based vision systems. The specific contribution of this paper is the closed-form formula for computing the representation and a framework to account for every aspect of illumination effects. Certainly the starting theorem is not new and it is not claimed as a contribution. In summary, Shashua's research is has little to do with the research presented in this paper except for the starting theorem.

In [2], Belhumeur and Kriegman try to answer what is the set of images of an object under different lighting conditions for the purpose of recognition. They prove that, for a convex Lambertian object, the set of images of this convex object is a cone in $\mathbb{R}^n$ (where n is the number of pixels in an image). This is an interesting theoretical analysis. However, we present different methods to account for illumination effects.

In [1], Belhumeur, Hespanha, and Kriegman propose a post-processing method to accommodate lighting changes. Their method works in the eigenspace to process eigenimages to get better new eigenimages to account for lighting changes. The method to generate eigenimages is the traditional method in which: (1) sample images in different lighting conditions are taken; (2) the probability for every sample image is assumed to be the same. Then the covariance matrix is computed. As demonstrated later, the method proposed here gives a totally different way to compute the covariance matrix. It uses only three images as input yet accounts for all (infinite number of) images in a light pyramid. Furthermore, it establishes a framework to accommodate arbitrary lighting changes. This together with the new method to compute the covariance matrix is fundamentally different from their post-processing (of eigenimages) method.

In [5], Hallinan gives a method to take sample images in different light conditions. He claims that this new method of taking sample images generates good eigenimages which can account for lighting changes. The difference between his method and the proposed method is obvious. The proposed method generates good eigenimages by a formula instead of a better way to sample the image space. The formula is developed by theoretical analysis and by using all images in a light pyramid in its derivation.

In [10], Nayar and Murase give a result on the dimensionality of their parametric manifold in the eigenspace. The method to compute eigenimages is the traditional method, i.e., by sampling the image space and computing the covariance matrix using these sample images. In this aspect, the difference between their research and the research presented in this paper is the same as the difference between Hallinan's work mentioned above and this research. Furthermore the major concern in their research is how to construct the manifold in different lighting conditions in eigenspace. Their method is a new recognition algorithm. The research presented here does not deal with the eigenspace. Instead, the representation is the major concern. Hence there is no intersection here.

In [2, 10], Belhumeur, Kriegman, Nayar and Murase address a similar problem, interreflection. The research presented in this paper also addresses interreflection problem. However, the treatment here is fundamentally different from theirs in that (1) they only consider the interreflection of a *concave object*. In contrast, the research presented in this paper considers all sources of interreflection, not just the interreflection which occurs in a concave object due to its concavity; (2) they form *pseudo surfaces* to account for interreflection. On the contrary, here *pseudo light sources* is used as the modelling technique to account for interreflection. Since the major concern is the appearances of the objects instead of the shapes of the objects, the proposed formulation brings more insight into the illumination effects and establishes a more consistent way to treat illumination effects.

In [9, 10], Murase and Nayar give a method for the purpose of removing light intensity changes. The method is a unit vector normalization. This method only has clear meaning for the same lighting sources with different light intensities. In the general case where light source type, number, and intensity can change at the same time, this method has no clear meaning. Although this method makes all images on a unit sphere in a high dimensional Hilbert-space, the unit sphere or the Hilbert-space brings no more insight to computer vision analysis. On the contrary, a different method to remove the lighting intensity changes is presented in this paper. The method is inspired by an observation of human vision systems. More importantly, this method is proved to be the optimum method to remove the lighting intensity changes. This method works in the general case where light source type, number, and intensity can change at the same time in any way and interreflection can occur at the same time. This new method has been verified by synthetic images and real images.

From the comparison given above, it is clear that nearly all research done so far to account for lighting changes in PCA-based vision systems is in the recognition stage. The only work in the representation stage is an empirical study which gives a better method to sample the image space.

In the following section, a new theoretical method to compute eigenimages given only three sample images, and a new method to account for arbitrary illumination effects are presented.

3 A Theoretical Method to Compute Covariance Matrix of All Images in a Light Pyramid Under Single Directional Light Source

3.1 Preparations

In this section a theoretical method to compute the mean image and the covariance matrix of all images in a light pyramid is given. It is assumed that the object is Lambertian (Definition 2). Furthermore, in this section it is assumed that the light source is a single directional light source. Later, this restriction is removed. Before proceeding, some definitions are given first.

Definition 1 *Directional light source: a light source which transmits light in one direction in the space* [3].

Directional light source is a term used in computer graphics. It is also termed as parallel light source. Sunlight is generally regarded as a directional light source.

Definition 2 *Lambertian surface: surface which scatters light equally in all directions and appears equally bright from all view directions* [7].

Definition 3 *Albedo (denoted as ρ): the fraction of incident light re-emitted by the surface* [7].

According to the definition of Lambertian surface, the following theorem results [7].

Theorem 1 *For any point in a Lambertian surface (Fig. 1), denote its normal as* $\mathbf{N}$ *(*$|\mathbf{N}| = 1$*), the direction of the single directional light source is* $\mathbf{L}$*. Then the radiance of the this point in any viewing direction is:*

$$p = \rho(\mathbf{N} \cdot \mathbf{L}). \tag{1}$$

Theorem 2 *The image of a Lambertian surface illuminated by any possible single directional light source can be generated by any three images of this surface illuminated by different single directional light sources so long as the three directions of these directional light sources are independent of each other.*

Proof: Assume that: (1) the direction of the single directional light source is $\mathbf{L}$; (2) The image of the surface under this light source is $\mathbf{I}$ (the image is raster-scanned to this vector); (3) $\mathbf{I}_i$ is a pixel of the image. According to Eq. (1), the radiance of the corresponding point in the surface is

$$\mathbf{I}_i = \rho_i(\mathbf{N}_i \cdot \mathbf{L}). \tag{2}$$

For this special point, the first two terms are constants under any lighting conditions. The only possible change comes from the light direction $\mathbf{L}$. However, for a single directional light source, $\mathbf{L}$ is

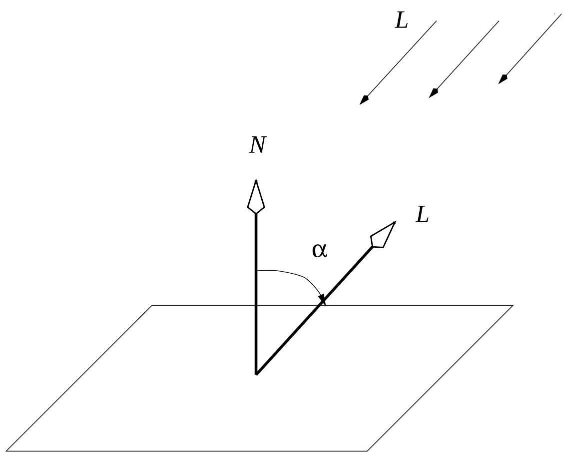


Figure 1: *A Lambertian surface.*

the same for every point on the surface. In other words, $\mathbf{L}$ is constant for any point on the surface when a single directional light source is used. (Note: this is a very important point.)

Since the dimension of $\mathbf{L}$ is 3, according to the theory of linear algebra, $\mathbf{L}$ can be represented as a linear sum of three different vectors $\mathbf{L}_i$ $(i = 1, 2, 3)$ (denote corresponding images as Υ_i),

$$\mathbf{L} = \sum_{i=1}^{i=3} (\mathbf{L} \cdot \mathbf{L}_i)\mathbf{L}_i, \tag{3}$$

as long as these three vectors are independent of each other. Since $\mathbf{L}$ is the same for all points in the surface, then Eq. (3) holds for every point on the surface. Hence,

$$\begin{aligned} \forall i,\ \mathbf{I}_i &= \rho_i(\mathbf{N}_i \cdot \mathbf{L}) \\ &= \rho_i(\mathbf{N}_i \cdot \{\sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})\mathbf{L}_l\}) \\ &= \rho_i \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})(\mathbf{L}_l \cdot \mathbf{N}_i) \\ &= \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})(\rho_i(\mathbf{L}_l \cdot \mathbf{N}_i)). \end{aligned} \tag{4}$$

Note that in the above equation, term $\rho_i(\mathbf{L}_l \cdot \mathbf{N}_i), l = 1, 2, 3$ are pixel values Υ_l^i. Hence,

$$\forall i,\ \mathbf{I}_i = \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})\Upsilon_l^i. \tag{5}$$

In general,

$$\mathbf{I} = \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})\Upsilon_l. \tag{6}$$

Therefore any image taken under illumination of a single directional light source can be composed by three images taken under illumination of single directional light sources $\mathbf{L}_i$ $(i = 1, 2, 3)$.□

It is important to note that the lighting direction should be the same for every point on the surface. This is only true for a directional light source. (Only this research and [2] make it clear that Eq. (3) is true only for a single directional light source.) For a single ideal point light source, Eq. (6) is an approximation (by ideal it means that the point occupies no space).

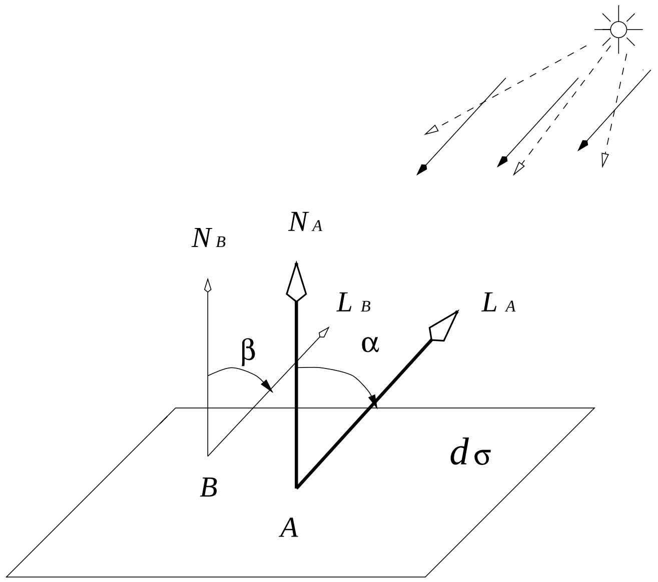


Figure 2: *Approximation.* Eq. (6) is only true when the light source is a single directional light source. For an ideal point light source, Eq. (6) is an approximation.

The accuracy of this approximation for an ideal point light source depends on:

1. the distance between the ideal point light source and surface point;
2. the range of concerned surface points. One measurement for the range of these concerned surface points is the largest distance between any of two concerned surface points.

When the first distance is much larger than the second distance, the approximation is very accurate. One special case is that only one surface point is concerned. In this situation, Eq. (6) is always true. However, if the range of the surface is in the same order of the first distance, the approximation is not accurate.

According to the analysis above on Eq. (6), a local method in which only local points are treated as the same can give a more accurate approximation than a global method in which all points in the surface are treated as the same. Note that this favours Mosaic Image Method.

The three vectors $\mathbf{L}_i$ $(i = 1, 2, 3)$ can be selected to be orthogonal to each other (hence independent of each other). This kind of three vectors forms a space called light pyramid.

Definition 4 *Light pyramid: the smaller part of a unit sphere defined by three unit vectors* $\mathbf{L}_i$ $(i = 1, 2, 3)$ *which are orthogonal to each other. These three vectors are called primary light directions. Images taken under illumination of a single directional light source* $\mathbf{L}_i$ $(i = 1, 2, 3)$ *are called primary images. (see Fig. 3).*

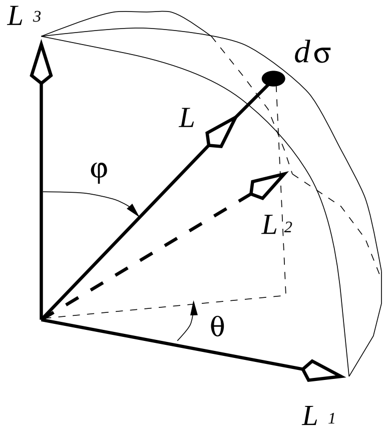


Figure 3: *The light pyramid.*

Every point on this part of the unit sphere defines a direction $\mathbf{L}$ for a single directional light source. In the following, the image corresponding to this direction $\mathbf{L}$ is simply called *an image in the light pyramid*. Denote it as $\mathbf{I}$ (which is raster-scanned to this vector). As discussed above, this image can be represented as a linear sum of the three primary images.

3.2 A New Method to Compute Eigenimages in PCA-Based Vision Systems

In the following a new method to compute eigenimages of a Lambertian object is presented. The fundamental difference between this method and the former methods is:

Every possible image in the light pyramid is used to compute the covariance matrix.

Theorem 3 *Denote the primary images as* Υ_l $(l = 1, 2, 3)$ *and any image in the light pyramid as* $\mathbf{I}$. *The mean image* $\mathbf{m}$ *of all the images in the light pyramid is:*

$$\mathbf{m} = E(\mathbf{I}) = \frac{1}{2}\sum_{l=1}^{l=3} \Upsilon_l. \tag{7}$$

Proof: Assume the probability of the light coming from any direction in the light pyramid is the same. Also assume that the direction corresponding to image $\mathbf{I}$ is $\mathbf{L}$. According to Theorem 2, any pixel in image $\mathbf{I}$ is:

$$\mathbf{I}_i = \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L})\Upsilon_l^i. \tag{8}$$

Since it is assumed that the object is motionless and only the single directional light source in the light pyramid can change, then the mean value of any pixel in the image is

$$\mathbf{m}_i = \sum_{\mathbf{L}} \mathbf{I}_i p_{\mathbf{L}}, \tag{9}$$

where

$$p_{\mathbf{L}} = \frac{d\sigma}{\pi/2}. \tag{10}$$

Here $d\sigma$ is the differential area on the unit sphere. In the special coordinate system shown in Fig. 3,

$$d\sigma = sin\psi \; d\theta d\psi. \tag{11}$$

Hence,

$$\begin{aligned}
\mathbf{m}_i &= \sum_{\mathbf{L}} \mathbf{I}_i p_{\mathbf{L}} = \frac{1}{\pi/2} \sum_{\mathbf{L}} \sum_{l=1}^{l=3} \{(\mathbf{L}_l \cdot \mathbf{L}) sin\psi \; d\theta d\psi\} \Upsilon_l^i \\
&= \frac{1}{\pi/2} \sum_{l=1}^{l=3} \{\sum_{\mathbf{L}} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi \; d\theta d\psi\} \Upsilon_l^i \\
&= \frac{1}{\pi/2} \sum_{l=1}^{l=3} \{\int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi \; d\theta d\psi\} \Upsilon_l^i.
\end{aligned} \tag{12}$$

It is easy to show that,

$$\begin{aligned}
\int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi d\theta d\psi &= \int_0^{\pi/2} \int_0^{\pi/2} cos\psi sin\psi \; d\theta d\psi \\
&= \int_0^{\pi/2} \int_0^{\pi/2} sin^2\psi \begin{pmatrix} sin\theta \\ cos\theta \end{pmatrix} d\theta d\psi \\
&= \frac{\pi}{4}.
\end{aligned} \tag{13}$$

Therefore,

$$\mathbf{m}_i = \frac{1}{2} \sum_{l=1}^{l=3} \Upsilon_l^i. \tag{14}$$

In general,

$$\mathbf{m} = \frac{1}{2} \sum_{l=1}^{l=3} \Upsilon_l. \tag{15}$$

□

Theorem 4 *The element $C_{i,j}$ of the covariance matrix:*

$$C = E((\mathbf{I} - \mathbf{m})(\mathbf{I} - \mathbf{m})^T), \tag{16}$$

of all images in the light pyramid (with the difference of a factor $\pi/2$) is:

$$C_{i,j} = \frac{1}{3} \sum_{l=1}^{l=3} \sum_{l'=1}^{l'=3} (1 + (\frac{\pi}{2} - 1)\delta_{l,l'}) \; \Upsilon_l^i \Upsilon_{l'}^j \; - \frac{\pi}{4}(\mathbf{m}_i \sum_{l=1}^{l=3} \Upsilon_l^j + \mathbf{m}_j \sum_{l=1}^{l=3} \Upsilon_l^i) \; + \frac{\pi}{2} \mathbf{m}_i \mathbf{m}_j, \tag{17}$$

where $\delta_{i,j} = 1$ *if* $i = j$, *otherwise* $\delta_{i,j} = 0$.
Proof: According to the definition of expectation,

$$C = E((\mathbf{I}-\mathbf{m})(\mathbf{I}-\mathbf{m})^T) \ = \sum_{\mathbf{L}} (\mathbf{I_L}-\mathbf{m})(\mathbf{I_L}-\mathbf{m})^T p_{\mathbf{L}}, \tag{18}$$

where $p_{\mathbf{L}}$ is the probability for image $\mathbf{I}$. Since the objective is to compute the eigenimages (eigenvectors), then, the factor $2/\pi$ can be omitted with no effect on the solution of the eigensystem[1], so element $C_{i,j}$ of the covariance matrix is

$$\begin{aligned}
C_{i,j} &= \sum_{\mathbf{L}} (\mathbf{I}^i_{\mathbf{L}} - \mathbf{m}_i)(\mathbf{I}^j_{\mathbf{L}} - \mathbf{m}_j) d\sigma_{\mathbf{L}} \\
&= \sum_{\mathbf{L}} \{ (\sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^i_l - \mathbf{m}_i)(\sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^j_l - \mathbf{m}_j) \} d\sigma_{\mathbf{L}} \\
&= \sum_{\mathbf{L}} \{ \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^i_l \ \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^j_l \ - (\mathbf{m}_i \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^j_l + \mathbf{m}_j \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^i_l) + \mathbf{m}_i \mathbf{m}_j \} d\sigma_{\mathbf{L}} \\
&= \sum_{\mathbf{L}} \sum_{l=1}^{l=3} \sum_{l'=1}^{l'=3} (\mathbf{L}_l \cdot \mathbf{L})(\mathbf{L}_{l'} \cdot \mathbf{L}) \ \Upsilon^i_l \Upsilon^j_{l'} d\sigma_{\mathbf{L}} \\
&\quad -(\mathbf{m}_i \sum_{\mathbf{L}} \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^j_l d\sigma_{\mathbf{L}} \ + \mathbf{m}_j \sum_{\mathbf{L}} \sum_{l=1}^{l=3} (\mathbf{L}_l \cdot \mathbf{L}) \Upsilon^i_l d\sigma_{\mathbf{L}}) \ + \mathbf{m}_i \mathbf{m}_j \sum_{\mathbf{L}} d\sigma_{\mathbf{L}} \\
&= \sum_{l=1}^{l=3} \sum_{l'=1}^{l'=3} \Upsilon^i_l \Upsilon^j_{l'} \{ \sum_{\mathbf{L}} (\mathbf{L}_l \cdot \mathbf{L})(\mathbf{L}_{l'} \cdot \mathbf{L}) d\sigma_{\mathbf{L}} \} \\
&\quad -(\mathbf{m}_i \sum_{l=1}^{l=3} \Upsilon^j_l \{ \sum_{\mathbf{L}} (\mathbf{L}_l \cdot \mathbf{L}) d\sigma_{\mathbf{L}} \} \ + \mathbf{m}_j \sum_{l=1}^{l=3} \Upsilon^i_l \{ \sum_{\mathbf{L}} (\mathbf{L}_l \cdot \mathbf{L}) d\sigma_{\mathbf{L}} \}) \ + \mathbf{m}_i \mathbf{m}_j \{ \sum_{\mathbf{L}} d\sigma_{\mathbf{L}} \} \\
&= \sum_{l=1}^{l=3} \sum_{l'=1}^{l'=3} \{ \int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L})(\mathbf{L}_{l'} \cdot \mathbf{L}) sin\psi \ d\psi d\theta \} \Upsilon^i_l \Upsilon^j_{l'} \\
&\quad -(\mathbf{m}_i \sum_{l=1}^{l=3} \{ \int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi \ d\psi d\theta \} \Upsilon^j_l + \mathbf{m}_j \sum_{l=1}^{l=3} \{ \int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi \ d\psi d\theta \} \Upsilon^i_l) \\
&\quad + \mathbf{m}_i \mathbf{m}_j \{ \int_0^{\pi/2} \int_0^{\pi/2} sin\psi \ d\psi d\theta \}.
\end{aligned} \tag{19}$$

It is trivial to show that the last integral is $\pi/2$. From Theorem 3,

$$\int_0^{\pi/2} \int_0^{\pi/2} (\mathbf{L}_l \cdot \mathbf{L}) sin\psi \ d\psi d\theta = \frac{\pi}{4}.$$

[1] $C\mathbf{x} = \lambda \mathbf{x}$, $(|\mathbf{x}| = 1)$; it is easy to see that a factor to C only affects eigenvalue λ, yet has no effect on *unit* eigenvector $\mathbf{x}$.

For the computation of

$$\int_0^{\pi/2}\int_0^{\pi/2}(\mathbf{L}_l\cdot\mathbf{L})(\mathbf{L}_{l'}\cdot\mathbf{L})sin\psi\ d\psi d\theta, \tag{20}$$

it is easy to prove that when $l = l'$,

$$\int_0^{\pi/2}\int_0^{\pi/2}(\mathbf{L}_l\cdot\mathbf{L})(\mathbf{L}_{l'}\cdot\mathbf{L})sin\psi\ d\psi d\theta\ = \int_0^{\pi/2}\int_0^{\pi/2}cos^2\psi sin\psi\ d\psi d\theta = \frac{\pi}{6}, \tag{21}$$

when $l \neq l'$,

$$\int_0^{\pi/2}\int_0^{\pi/2}(\mathbf{L}_l\cdot\mathbf{L})(\mathbf{L}_{l'}\cdot\mathbf{L})sin\psi\ d\psi d\theta\ = \int_0^{\pi/2}\int_0^{\pi/2}cos\psi sin^2\psi\begin{pmatrix} sin\theta \\ cos\theta \end{pmatrix}d\psi d\theta = \frac{1}{3}. \tag{22}$$

□

The eigenimages Φ_i $(i = 0, 1, 2, ..., n-1)$ (here n is the number of the pixels in a image) generated by solving the eigensystem formed by the covariance by Theorem 4 account for every directional light source in the light pyramid.

For any image $\mathbf{I}$ in the light pyramid,

$$\mathbf{I}\ =\ E(\mathbf{I}) + \sum_{i=0}^{n-1} y_i\Phi_i, \tag{23}$$

where y_i $(i = 0, 1, ..., n-1)$ is the projection of $(\mathbf{I} - E(\mathbf{I}))$ onto Φ_i $(i = 0, 1, ..., n-1)$. $\mathbf{I}$ can be approximated as

$$\mathbf{I} \approx \tilde{\mathbf{I}} = E(\mathbf{I}) + \sum_{i=0}^{m-1} y_i\Phi_i, \tag{24}$$

where $m \ll n$.

This approximation is only true when only one directional light source is involved. In real life, the light source number, type, intensity, and environment where the image is taken can all change. How to account for these illumination effects is a subtle problem which is the topic of the next section.

4 General Illumination Effects

In this section, a generalized definition of light sources is given. This new definition of light source encompasses all kinds of light sources used in computer graphics and computer vision. It serves as the basis for the following analyses. Following this generalized definition, a new formulation of interreflection is presented. Then, a new method to remove light intensity changes is given. Finally by taking account of all these analyses, an equation is established which is called the illumination equation. Since all these analyses are based on first principles, it is expected that the methodology of these analyses is applicable to the analysis of illumination effects in other vision systems or other related areas.

4.1 Light Sources

Light sources are very important to vision systems since without light sources it is impossible to see. It is also a fundamental element in computer graphics. In all light sources, the point light source is the simplest one.

Definition 5 *Point light source: a point in three dimensional space which emits energy in all directions in this space. This point has a position in three dimensional space yet occupies no space. Note, the energy emitted in different directions may be different.*

In a rigorous sense, there is no point light source in the real world. The point light source is only a tool for analysis since a point light source should occupy no space or area yet has a position. Although the point light source is widely used in computer graphics research and practices, it is, in fact, an approximation.

One important observation on light sources is that for visual effects (or "see"), the only important geometry of light sources is the light sources' surfaces. Surfaces in three dimensional space are two dimensional. (Some issues are not important for visual research such as some aspects of the physics of light sources; for example, the nuclear reaction inside the sun and so forth. Of course, the light sources can be modelled in three dimensional space, however at the end it is the light intensities on the light sources' surfaces and the light source surfaces' shapes that matter.)

Given the observation discussed above and the fact that there is no point light source in real world, the following axiom results.

Axiom *All real world light sources are area light sources.*

Therefore all linear light sources and area light sources in computer graphics are, in fact, area light sources. Here a linear light source is actually an approximation in one dimension.

Another commonly light source is the directional light source; this is indeed an approximation in real life - when an area light source is far away enough from a surface (or only a very small area on the surface is concerned), the area light source can be modelled as directional light sources. Of course, when the surface decays to a point, any area light source can be regarded as a collection of directional point light sources. This point is theorized in the following.

To analyse area light sources, the natural way is to model the area light source as the sum of many infinitesimal small area (differential area) light sources. These infinitesimal area light sources can be regarded as point light sources. In Fig. 4, denote the light source surface as ζ, denote an infinitesimal area in ζ as $d\sigma$, then the radiance of a point A in a Lambertian object is:

$$p = \iint_{\zeta} \rho(\mathbf{N} \cdot \mathbf{L})\xi \; d\sigma \qquad (25)$$

$$= \sum_{d\sigma \subset \zeta} \rho(\mathbf{N} \cdot \mathbf{L}_{d\sigma}) \; \xi_{d\sigma} \; d\sigma, \qquad (26)$$

where ξ is the light intensity of $d\sigma$ measured from A . Therefore, it can be considered that point A is illuminated by many different directional light sources. This light source can be denoted as:

$$\mathbf{D} = \sum_{d\sigma \subset \zeta} (\xi_{d\sigma} d\sigma) \; \mathbf{L}_{d\sigma}, \qquad (27)$$

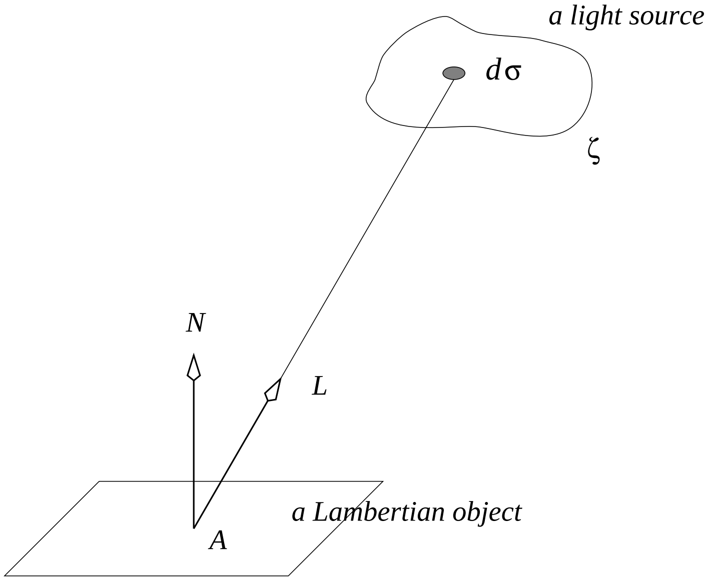


Figure 4: *A light source.*

where $\mathbf{L}_{d\sigma}$ is the direction of a directional light source. $\xi_{d\sigma}d\sigma$ is the intensity of this directional light source.

Therefore, no matter how the light source type, number and intensity change, all these light sources can be modelled by Eq. (27). However, one thing that needs to be emphasized here is that considering $\mathbf{D}$ as a collection of directional light sources is only true for a single point A. When many points on the object surface are taken into account at the same time, light source $\mathbf{D}$ is a cluster of ideal point light sources, where $\mathbf{L}_{d\sigma}$ is the vector from a point of the object's surface to $d\sigma$. $\xi_{d\sigma}d\sigma$ is the intensity of an ideal point light source measured from a point of the object's surface to $d\sigma$. As discussed in preparations section, when Theorem 2 is used under $\mathbf{D}$, there is an approximation.

In summary, all light sources in the real world are area light sources. The geometry which matters for visual research is two dimensional surfaces of light sources. Any area light source can be modelled by Eq. (27).

4.2 Interreflection

Interreflection is a very common lighting phenomenon. In fact, interreflection is widely used in modern lighting engineering to get the so-called soft "natural" lighting. Apart from direct lighting from the light sources, interreflection is an important illumination effect which must be accounted for.

According to Kajiya's rendering equation [3]:

$$I(x,x') = g(x,x')\left[\varepsilon(x,x') + \int_S \rho(x,x',x'')I(x',x'')dx''\right]. \tag{28}$$

The energy reaches x from x' includes: (1) light from light sources at x', (2) light from x'' that is reflected to x'; this latter part of energy transfer is called interreflection. Since the major concern of this research is illumination effects, interreflection is modelled as *pseudo light sources*. This kind of light sources is just like the Moon which itself does not emit light yet reflects light. The light intensity value $I(x, x')$ can be computed by the rendering equation. (However, since this is a computer vision research, this computation is not needed. Indeed it is computed by Nature as the input image.) This formulation is supported by the two-pass rendering method in computer graphics research [11], in which the rendering equation is solved in the view-independent first pass and the image viewed from a special direction is rendered by the second-pass, where the solution of the first pass gives the interreflection which serves as the secondary illumination in the second-pass.

This formulation can be applied to any interreflection sources no matter whether it is due to a concave object or energy reflected by an object. In fact, interreflection of a concave object may be not big enough to affect the function of PCA-based vision systems. For example, one major application area of PCA-based vision systems is human face recognition. Obviously the human face is a concave object, e.g., there exists interreflection between the nose and the cheek. However from the simulation, there are few differences between face images generated when interreflection is accounted for and when it is not. In other words, interreflection caused by a concave object itself may have little effect on its appearances.

Since interreflection is modelled as pseudo light sources, Eq. (27) can be used to model it.

4.3 Light Intensity Changes

Light intensity changes have important effects on images. Therefore how to account for light intensity changes is an important problem. Some researchers [9] try to normalize the image vector into a unit vector to remove the light intensity change (this is the only normalization method in the literature). As pointed out earlier, there is no clear meaning here. The claim that the unified image is in the Hilbert space brings no more insight for vision research. This section provides a new method to account for light intensity changes which is inspired by an observation of human vision systems.

In real life, everyone may have this experience: when one goes into a bright environment from a dim one, one may lose his vision temporarily. The same is true when one goes into a dim place from a bright one. That one can eventually see means that the responses of sensor elements are not all the same. Therefore there exist changes around the mean intensity value. The change around mean intensity gives us the pattern of or the information on the world. When one goes to a bright environment from a dim one, all responses of the sensors overflow. Therefore, one loses the pattern of the world. In other words, there is no change around the mean.

Consider the following example. Assume that the mean responses of the sensors are zero in a dim environment. After entering a bright environment, it takes a little while to adjust to this new environment. In this period of time, the zero mean value is still being used. However, the responses of the sensors in the bright environment are much larger than those in the dim environment. These responses are much larger than the dim mean value, zero. Therefore, the sensors overflow. After a period of time, the sensors are adjusted to the new environment. It is guessed that the adjustment proceeds according to the bright mean response. After the bright mean response is adopted, the responses of the sensors no longer overflow. Instead, the responses of the sensors wave around the new mean response again. Hence it is guessed that the mean value may be a good source to remove the effect of the intensity changes.

In the following, it is proved that the mean of an image is the optimum means to account for light intensity changes.

Theorem 5 *Assume that there are t unit intensity directional light sources $\mathbf{L}_i$ $(i = 0, 1, 2, ..., t-1)$, and that the mean image of all images taken under only one directional light source on is $\mathbf{m}$. Now consider the situation where the directional light sources may turn on at the same time with different intensity values α_i (they may also be 0),*

$$\mathbf{D} = \sum_{i=0}^{i=t-1} \alpha_i \mathbf{L}_i. \tag{29}$$

Assume image $\mathbf{I}$ is taken under this new light source $\mathbf{D}$; the optimum method to estimate the value of $\sum_{i=0}^{i=t-1} \alpha_i$ is by computing the ratio of the mean pixel value of new image $\mathbf{I}$ to the mean pixel value of the mean image $\mathbf{m}$.

Proof: Assume the image taken under $\mathbf{L}_i$ $(i = 0, 1, 2, ..., t-1)$ is $\mathbf{I}^j$ $(j = 0, 1, 2, ..., t-1)$, then the image under light source $\mathbf{D}$ is

$$\mathbf{I} = \sum_{j=0}^{j=t-1} \alpha_i \mathbf{I}^j. \tag{30}$$

Assume that there is n pixel in an image, then the ratio of mean value of new image $\mathbf{I}$ to that of mean-image $\mathbf{m}$ is

$$ratio = \frac{\sum_i \mathbf{I}_i / n}{\sum_i \mathbf{m}_i / n} = \frac{\sum_i \mathbf{I}_i}{\sum_i \mathbf{m}_i}. \tag{31}$$

The expectation of $ratio$ is (note: $\forall i,\ \mathbf{m}_i$ is constant)

$$E(ratio) = E(\frac{\sum_i \mathbf{I}_i}{\sum_i \mathbf{m}_i}) = \frac{E(\sum_i \mathbf{I}_i)}{\sum_i \mathbf{m}_i} = \frac{\sum_i E(\mathbf{I}_i)}{\sum_i \mathbf{m}_i}. \tag{32}$$

Since

$$E(\mathbf{I}_i) = \sum_j \alpha_j E(\mathbf{I}_i^j) = \sum_j \alpha_j \mathbf{m}_i = \{\sum_j \alpha_j\} \mathbf{m}_i, \tag{33}$$

therefore,

$$E(ratio) = \frac{\sum_i E(\mathbf{I}_i)}{\sum_i \mathbf{m}_i} = \frac{\{\sum_j \alpha_j\} \sum_i \mathbf{m}_i}{\sum_i \mathbf{m}_i} = \sum_j \alpha_j. \tag{34}$$

□

4.4 Illumination Equation

In this section, an equation called *illumination equation* is derived. Given eigenimages generated by Theorem 4, this equation accounts for all illumination effects described above.

Theorem 6 *Given eigenimages generated by Theorem 4, an image* $\mathbf{I}$ *taken in any lighting condition (as long as these light sources are all in the light pyramid) can be represented by:*

$$\mathbf{I} \approx \alpha\mathbf{m} + \sum_{i=0}^{i=n-1} p_i \Phi_i, \tag{35}$$

where

$$p_i = (\mathbf{I} - \alpha\mathbf{m}) \cdot \Phi_i, \tag{36}$$

and α *is given by Theorem 5.*

Proof: Since any light source can be represented by Eq. (27), besides, this equation also holds for interreflection, then any light sources $\mathbf{D}$ (including pseudo light sources) can be modelled as:

$$\mathbf{D} = \sum_i \alpha_i \mathbf{L}_i. \tag{37}$$

Hence image $\mathbf{I}$ taken under this lighting condition is:

$$\mathbf{I} = \sum_i \alpha_i \mathbf{I}_i. \tag{38}$$

According to the theory of principal component analysis, $\forall\ \mathbf{I}_i$, the following equation holds:

$$\mathbf{I}_i = \mathbf{m} + \sum_{j=0}^{n-1} \{(\mathbf{I}_i - \mathbf{m}) \cdot \Phi_j\} \Phi_j. \tag{39}$$

Therefore,

$$\begin{aligned} \mathbf{I} = \sum_i \alpha_i \mathbf{I}_i &= \{\sum_i \alpha_i\}\mathbf{m} + \sum_{j=0}^{n-1} \{(\{\sum_i \alpha_i \mathbf{I}_i\} - \{\sum_i \alpha_i\}\mathbf{m}) \cdot \Phi_j\} \Phi_j \\ &= \{\sum_i \alpha_i\}\mathbf{m} + \sum_{j=0}^{n-1} \{(\mathbf{I} - \{\sum_i \alpha_i\}\mathbf{m}) \cdot \Phi_j\} \Phi_j. \end{aligned} \tag{40}$$

In Eq. (40), the only unknown part is $\sum_i \alpha_i$; however, this term can be optimumly estimated by Theorem 5. Therefore, given eigenimages generated by Theorem 4, an image taken in any illumination situation can be represented by:

$$\mathbf{I} \approx \alpha\mathbf{m} + \sum_{i=0}^{i=n-1} \{(\mathbf{I} - \alpha\mathbf{m}) \cdot \Phi_j\} \Phi_i, \tag{41}$$

where α is estimated by Theorem 5. $\square$

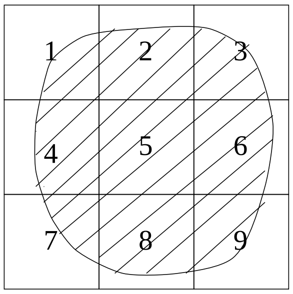
4.5 Implications of Illumination Equation

The rendering equation is the foundation of computer graphics. It describes every light behaviour. The illumination equation accounts for two of three terms in the rendering equation, i.e., $\varepsilon(x, x')$ and $\int\limits_S \rho(x, x', x'')I(x', x'')dx''$. These two terms correspond to light sources and interreflection. The only term which the illumination equation does not account for is the geometry term $g(x, x')$. This term corresponds to occlusion or shadow which is addressed in the next section.

The solution of the rendering equation $I(x, x')$ is very demanding. One simplification of it is the so-called radiosity method. Yet the radiosity method is still the most computationally expensive method in computer graphics. Fortunately for computer vision research, $I(x, x')$ is indeed the input image which is rendered by *nature*. What is needed for computer vision research is to extract information from $I(x, x')$.

5 Combination with Mosaic Image Method

Mosaic Image Method is a method developed mainly for accounting for occlusion [13]. The basic idea behind it is to generate the whole eigenimages by generating them locally. The first step is to slice the image into small mosaic images, then to generate eigenimages for these small mosaic images. These small mosaic-eigenimages then combine to form a representation of the whole image. It can be proved that eigenimages formed by this way are orthogonal to each other. Therefore, they form part of a basis of the image space. Hence, in this way, a global and local representation can be generated (Fig. 5). Since this method is a global and local method, combining it with the framework developed here brings three advantages.



Objects=Features+Topology

Figure 5: *Mosaic Image Method.*

5.1 Better Approximation

As indicated earlier, in a rigorous sense, Theorem 2 (hence Theorem 6) only holds for directional light sources. To apply it to general light source, an approximation is involved (Fig. 2). If the surface involved is too large, the approximation is not accurate. Therefore it is advantageous to have the surface points clustered in a small area. This is exactly what Mosaic Image Method does.

5.2 Tractable Computation

Since the covariance matrix is generated by Eq. (17), the fact that the eigenvectors of AA^T is a linear transformation of eigenvectors of A^TA is not applicable. Therefore, if the image is treated as a whole and used to generate covariance matrix by Eq. (17), the computation of this eigensystem is very demanding. For example, if the training image is in the dimensions of 100×120, then eigenvectors of a matrix of 12000×12000 need to be computed. This is a very demanding undertaking.

On the contrary, if Mosaic Image Method is applied, the covariance matrix is much smaller. For example, if the mosaic image is in dimensions of 20×20, then for the whole image in dimension 100×120, 48 400×400 matrixes' eigenvectors need to be computed. This is a much easier task compared with the original one if the image is treated as a whole. (The computational complexity of solving an eigensystem is in the order of $O(N^3)$ where N is the dimension of the matrix.)

5.3 Accommodating Occlusion, Shadow and Specular Highlight

5.3.1 Occlusion

Since Mosaic Image Method can accommodate a large degree of occlusion, by combining with it, occlusion can be accommodated.

5.3.2 Shadow

Shadow is formed by points in a scenario which cannot see at least some light sources but can see the camera. If the points can see part of the light sources but not all, this kind of point forms a *penumbra* region in an image. If the points cannot see any light sources, this kind of point forms an *umbra* region in an image. The shadow can also be classified as cast-shadow and attached-shadow [12]. Attached-shadow is formed by the points on the object surface whose normals and the directions of the light source form obtuse angles. The cast-shadow, however, is formed by points on object surface which are obstructed from the light sources by other objects or part of the same object.

Attached shadow is an inherent part of the image. This kind of shadow can be reconstructed by the method given by Theorem 2. However, this reconstruction introduces some errors since negative values are assumed to be zero. However this error is small. Cast-shadow can be classified as penumbra and umbra. These two kinds of shadows can change the appearances of objects drastically and make resulting images totally different from images without shadow. For the umbra part, since this part of the image is totally different from the original one, this part can be regarded as an occluded part. This can be processed in the recognition stage in Mosaic Image Method. For the penumbra part, if this part can see the same light sources, then according to illumination equation, this part of the image can be accommodated by the proposed method. However, if the points in this part of the image see different light sources, then according to illumination equation, this part of the image may not be able to be accommodated by the proposed method if this part of the image changes too much. If this happens, this kind of shadows can be treated as occlusion by Mosaic Image Method in recognition stage. Of course, if the variation due to different light sources is small enough, it is able to be accommodated by illumination equation.

As discussed above, by combining with Mosaic Image Method, shadows can be treated as occluded parts in Mosaic Image Method.

5.3.3 Specular Highlight

Specular highlight occurs with objects such as metal, glass, marble. However according to the reflection model [6], specular highlight highly depends on the direction of the light source and the viewing direction. Only a small part of objects can present specular highlight when viewed from a special direction; for most parts of objects, specular highlight can be ignored. Therefore, specular highlight is a local lighting phenomenon. Furthermore the parts of images where specular highlight occur are totally different from the "correct" images. Therefore these parts of images can be treated as occluded parts. By combining with Mosaic Image Method, the current method acquires the ability to account for specular highlight in the recognition stage.

6 Experiments

To test the proposed methods, two major experiments are conducted. In the first experiment, test images are generated by computer graphics simulation in which all ideal conditions such as ideal Lambertian surface and ideal directional light sources are met. Images in different lighting conditions are rendered by a powerful rendering tool [4]. In the second experiment, a real Lambertian object is used to test the proposed methods. Images in different lighting conditions of the object are taken by a CCD camera. Eigenimages are computed by Theorem 4. The reconstructed and residual images are generated by Theorem 6. In the following sections, detailed description of these two experiments and their results are given.

Since in this research, the major objective is to test a new *representation* and a new method to capture the different lighting effects based on this new representation. To evaluate a representation, the best way to do is to apply this new representation to represent the subject which this new representation is for. In this case, this means to reconstruct the original image (use the representation) and estimate the difference between the original data and the reconstructed data by the representation (residual). For the difference (residual), the error per pixel can be computed, however, it is found that this measurement bring little insight. Hence, residual images are used for this purpose. The larger the residual at a pixel, the larger the error.

6.1 Experiment (I): Computer Graphics Simulations

In this experiment, all test images are generated by computer graphics simulation. A polygonal mesh human face model is used as the subject (Fig. 6). The face model is assumed to be ideally Lambertian in the rendering.

The three primary images are generated under ideal directional lighting sources. The primary lighting directions are from the above of the head, the lower left, and the lower right. These three primary images are shown in Fig. 7.

The three primary images are used as the input to compute eigenimages by Theorem 4. The eigenimages generated by the proposed method and the traditional method are shown in Fig. 8.

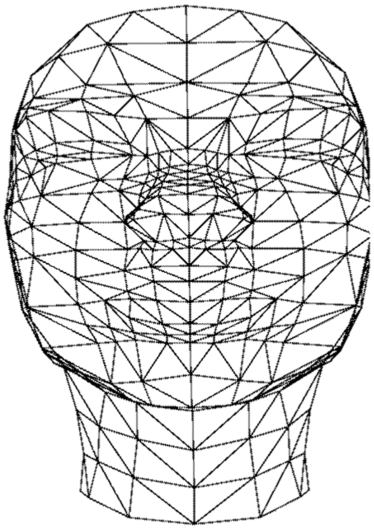


Figure 6: *A polygonal mesh human face.*

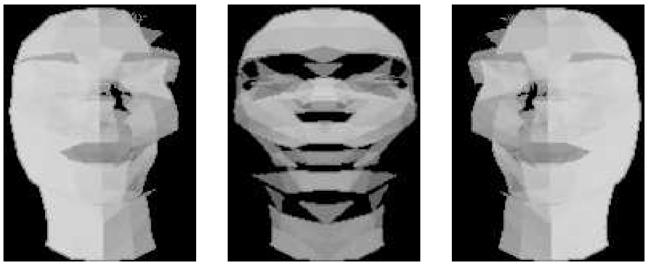


Figure 7: *The three synthesized primary images.* $\gamma = 2$.

To test the new representation and the illumination equation, many test images are rendered in different lighting conditions. In the following sections, these images and the experimental results are presented according to different lighting conditions.

6.1.1 Same Light Sources with Different Light Intensities

In this test, three groups of images are rendered. In every group of images, the light sources are kept the same while the intensities of the light sources are varied in five different levels. As an example, a group of images is shown in Fig. 9. The lighting for this group of images is two directional light sources. The first image is rendered when the the intensity is 0.27 and 0.4 respectively. The intensities are increased by 1.5 times for the followed images (Table 1).

The illumination equation is used to reconstruct these test images. As a comparison, the traditional method is also used to reconstruct these images. The residual images are also compared. As expected, the new method give smaller residuals. Besides, the reconstructed images are better than those by the traditional method. The results for one group are given in Fig. 9. The first row is the rendered images. The second and third rows are images reconstructed by the new method and the traditional method.

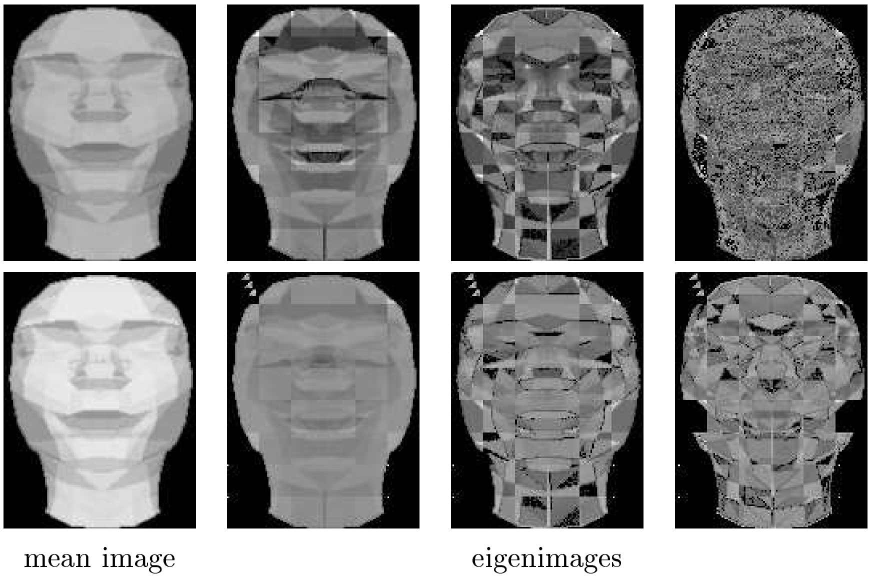


Figure 8: *The mean image and the first three eigenimages.* The top row is eigenimages generated by the traditional method. The second row is eigenimages generated by Theorem 4. The input images are the three primary images. Note the fundamental difference in the third eigenimage. $\gamma = 2$.

Table 1: *Same light sources with different light intensities.*

	intensities of light sources
first image	(0.27, 0.4)
second image	(0.4, 0.6)
third image	(0.6, 0.9)
fourth image	(0.9, 1.35)
fifth image	(1.35, 2.0)

In every test image, ratio α is estimated for every mosaic image. The results are shown in Fig. 10. As expected, the higher the intensity, the larger the estimated ratio α.

6.1.2 Multiple Directional Light Sources

This group of test images is rendered to test the performance of the proposed methods in the illumination of multiple directional light sources (the counterpart of this test for the real images involves an approximation in which several bulb light sources are used.) Fifteen test images are rendered with $2, 3, 4$ directional light sources. The directions of these light sources are in the light pyramid (specification for the setup can be found in Table 2). Note that the interreflection effect is not rendered for these test images.

Five of the 15 test images are shown in Fig. 11. The reconstructed images and the corresponding residual images by the new methods are the second and the fourth rows in the figure. The recon-

Table 2: *Specification for the test images rendered under illumination of multiple directional light sources.*

light source number	intensities of light sources
2	(0.5, 0.5), (0.5, 0.8), (0.5, 1.5), (1.0, 1.0)
3	(0.2, 0.3, 0.5), (0.3, 0.3, 0.3), (0.4, 0.6, 0.8), (1.0, 1.0, 1.0)
4	(0.2, 0.2, 0.2, 0.2), (0.2, 0.3, 0.4, 0.5), (0.3, 0.4, 0.5, 0.6), (0.5, 0.6, 0.7, 0.8), (0.6, 0.7, 0.8, 0.9), (0.5, 0.7, 0.9, 1.1), (1.0, 1.0, 1.0, 1.0)

structed images and the corresponding residual images by the traditional method are the third and the sixth rows in the figure. From this figure, it is clear that the new methods give smaller residual and better reconstructed images.

6.1.3 Area Light Sources

Two groups of images are rendered under area light sources, one group with interreflection and one group without. The general setup is shown in Fig. 12. The polygonal mesh human face is located inside a large cube, parts of the cube surfaces are in the light pyramid and they therefore serve as the media of interreflection. To enforce the interreflection effects, two elliptical spheres and two cubes are positioned near the polygonal mesh human face. Hence these four objects also serve as the media of interreflection. 15 different area light sources located inside the light pyramid are used. The shapes of these area light sources are triangles and rectangles.

In Fig. 13, some images under area light sources which do not account for interreflection are shown (the first row). The reconstructed images by the new methods and the corresponding residual images are also shown (the second and the fourth rows). The residual images by the proposed methods are smaller than those by the traditional method. The quality of the reconstructed images by the proposed methods is also better than those by the traditional method.

To test the proposed method in the presence of interreflection, test images are also rendered by the radiosity method which accounts for interreflection. The interreflection comes from the different parts of the face (e.g., between nose and cheek), from parts of the large cube surfaces (which contain the human face model), from the four small objects near the face model. The rendered test images and the reconstructed and residual images are shown in Fig. 14 (the first, the second, and the third rows). The reconstructed images by the proposed methods are better than those by the traditional method. Furthermore the residual images are also smaller than those by the traditional method.

6.2 Experiment (II): Real Objects

In this experiment, a celebrity model is used. The material painted on the model's face is Lambertian.

The three primary images are taken with three $60W$ bulb light sources. The light directions are from the high left, high right and lower middle. Since the model's face is 1/6 as large as the light source surface and the distance between the light source and the model is much larger than the range of the model face, the approximation of the directional light sources is fairly good. The

Table 3: *Specification of the lighting conditions for the real image experiment.*

first	2 $60W$ bulbs
second	1 medium-size fluorescent table light
third	1 $60W$ bulb, 1 medium-size fluorescent table light
fourth	1 $100W$ bulb
fifth	2 $100W$ bulbs
sixth	1 $100W$ bulb, 1 medium-size fluorescent table light
seventh	all light sources in the office turned on

environment for the image taking is a usual office environment. When taking the three primary images, all light sources are turned down except the primary light sources, the three bulb light sources.

Seven test images are taken in different light conditions. The lighting conditions are shown in Table 3.

All these light sources are in the light pyramid except the last one. However the result is good which means the proposed method is very robust. Furthermore, to add interreflection to the test images, three thick books are put near the model and located in the light pyramid.

The three primary images are shown in Fig. 15.

The eigenimages generated by the proposed method are shown in Fig. 16.

In Fig. 17, the reconstructed images by the traditional method and the proposed methods are shown. The first row is the original images, the second and the third rows are the reconstructed images by the proposed method and the traditional method. It can be seen that the proposed methods give better reconstruction. The comparison of the reconstructed images by the proposed methods with different sizes of mosaic images are shown in Fig. 18.

From Fig. 17, it is clear that the reconstruction is quite good even when shadows appear. Therefore it may be able to compress images in any lighting condition by only three eigenimages generated with only three images as input. Fig. 19 gives the original test images and the reconstructed images with small post processing in which a simple smooth algorithm is used to remove the lines appearing at the edges of the mosaic images. The results are quite good.

6.3 Using More than Three Eigenimages in the Proposed Method

In the proposed method, if the mosaic image is 20×20, there are 400 eigenimages generated. Therefore, it is possible to use more than three eigenimages in the proposed methods. (In contrast, in the traditional method, since there are only three sample images, then only three eigenimages can be generated.) However, it is found that usually three eigenimages are enough for the reconstruction; using more eigenimages only improves the quality of reconstructed images slightly (Fig. 20).

7 Summary

In this paper a new method to compute the covariance matrix using only three input images yet accounting for all images in the light pyramid is proposed. A framework to account for arbitrary illumination effects is given. An equation called illumination equation is derived from the framework. The proposed methods are thoroughly tested by two experiments. The experimental results are very good. In the following, the main points of this paper are summarized.

1. If the object is Lambertian, the image of the object can be represented by just three images taken in three different lighting conditions. The result is actually reached by two approximations: (1) in a rigorous sense, the light source should be a directional light source. However, usually an ideal point light source gives a quite accurate approximation so long as the range of the object surface is far smaller than the distance between the point light source and the surface. By combining with Mosaic Image Method, this approximation can be made more accurate; (2) when the attached-shadow occurs, an approximation is introduced. Again, the error caused by this approximation is small.

2. Since Theorem 2 holds for a Lambertian object, the mean image of all images taken in the light pyramid can be computed by Theorem 3. The covariance matrix of these images can be computed by Theorem 4. In a sense, the fundamental difference between the proposed method and the former PCA-based vision systems is that in the proposed methods, all images are taken into account when computing the mean image and the covariance matrix. The probabilty is the differential area corresponding to the special lighting direction. Furthermore, in a sense the images used by the proposed method are countless while the image used by the former PCA-based vision system is a finite number.

3. The treatment in this research is theoretical. This is a main difference from earlier empirical research.

4. In this paper, it is observed all light sources are, in fact, area light sources in real life. The point light source is used as an analytical means in this paper. Light sources are modelled as clustered point light sources. Therefore a linear relation can be taken advantage here. By using this linear relationship, any light source can be written as a linear sum of a collection of point light sources; the coefficients of these point light sources is the relative intensity of these point light sources.

5. Interreflection can be modelled as pseudo light sources. This modelling method is fundamentally different from the current modelling method for interreflection in the computer vision community in which interreflection is modelled as pseudo surface. It is the author's view that modelling interreflection by pseudo surfaces is not suitable for the research of illumination effects. The modelling technique presented in this paper is derived from the rendering equation and can form a highly consistent analysis for illumination effects.

6. α can be optimumly estimated by computing the ratio of the mean pixel value of the input image to the mean pixel value of the mean image. This method is inspired by an observation of human vision systems and is expected to bear a meaning to human vision systems.

7. All lighting effects are summed up and form an equation called *illumination equation.* This equation can account for any lighting effects analysed above. The relationship between this equation and the rendering equation is emphasized.

8. By combining with Mosaic Image Method, local light phenomena such as occlusion, shadow, and specular highlight can be accounted for. This is mainly for the recognition stage. It cannot help the reconstruction. The reconstruction usually is not able to recover correctly these illumination effects since there is no information available on them in the training stage.

9. By combining with Mosaic Image Method, the computation of the eigenimages is easier.

10. The proposed methods are thoroughly tested by rendered and real images. The results confirm the advantages of the proposed methods.

8 Direction for Further Research

A direct research is how to apply this represenation and illumination equation in the recognition stage. Because light source number, type, intensity, and other illumination effects (e.g., interreflection caused by movement of nearby objects) may change at the same time and change in complex ways, it is impossible to sample all these lighting conditions. Therefore the popular method in which some sample images in special lighting conditions are taken, and the projections are computed and stored for recognition may not be applicable in general lighting conditions. Hence, to find a non-brute-force recognition method to account for every possible lighting change is an interesting and subtle problem which has great application potential.

References

[1] Belhumeur, P., Hespanha, J., and Kriegman., D., Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *European Conference on Computer Vision*, pp. 45-48, April 1996.

[2] Belhumeur, P., and Kriegman., D., What Is the Set of Images of an Object Under All Possible Lighting Conditions? *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 270-277, San Francisco, CA, July 1996.

[3] Foley, J.D., van Dam, A., Feiner, S.K., and Hughes, J.F., *Computer Graphics: Principles and Practice*, Addison-Wesley Press, 1992.

[4] Gritz, L., Blue Moon Rendering Tools, *Http://www.seas.gwu.edu/student/gritz/bmrt.html.*

[5] Hallinan, P.W., A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions, *CVPR '94*, Seattle, Washington, June, 1994.

[6] He, X.D., Torrance, K.E., Sillion, F.X., and Greenberg, D.P., A Comprehensive Physical Model for Light Reflection, *SIGGRAPH 1991*, pp. 175-186, Las Vegas, Nevada, July 1991.

[7] Horn, B.K., *Robot Vision*, MIT Press, 1986.

[8] Kukunaga, K., *Introduction to Statistical Pattern Recognition*, 2nd Ed., Academic Press, 1990.

[9] Murase, H., and Nayar, S., Learning and Recognition of 3D Object from Appearance, *Proceedings of IEEE Workshop on Qualitative Vision*, pp. 39-50, June 1993.

[10] Nayar, S.K., and Murase, H., Dimensionality of Illumination Manifolds in Eigenspace, *Technical Report*, CUCS-021-94, Computer Science Department, Columbia University, Revised in September, 1995.

[11] Reichert, M.C., A Two-Pass Radiosity Method Driven by Lights and Viewer Position, *Master's Thesis*, Program of Computer Graphics, Cornell University, January, 1992.

[12] Shashua, A., On Photometric Issues in 3D Visual Recognition From a Single 2D Image, *International Journal of Computer Vision*, to appear, 1997.

[13] Zhao, L. and Yang, Y.H., Mosaic Image Method: a Local and Global Method, *manuscript*, 1997

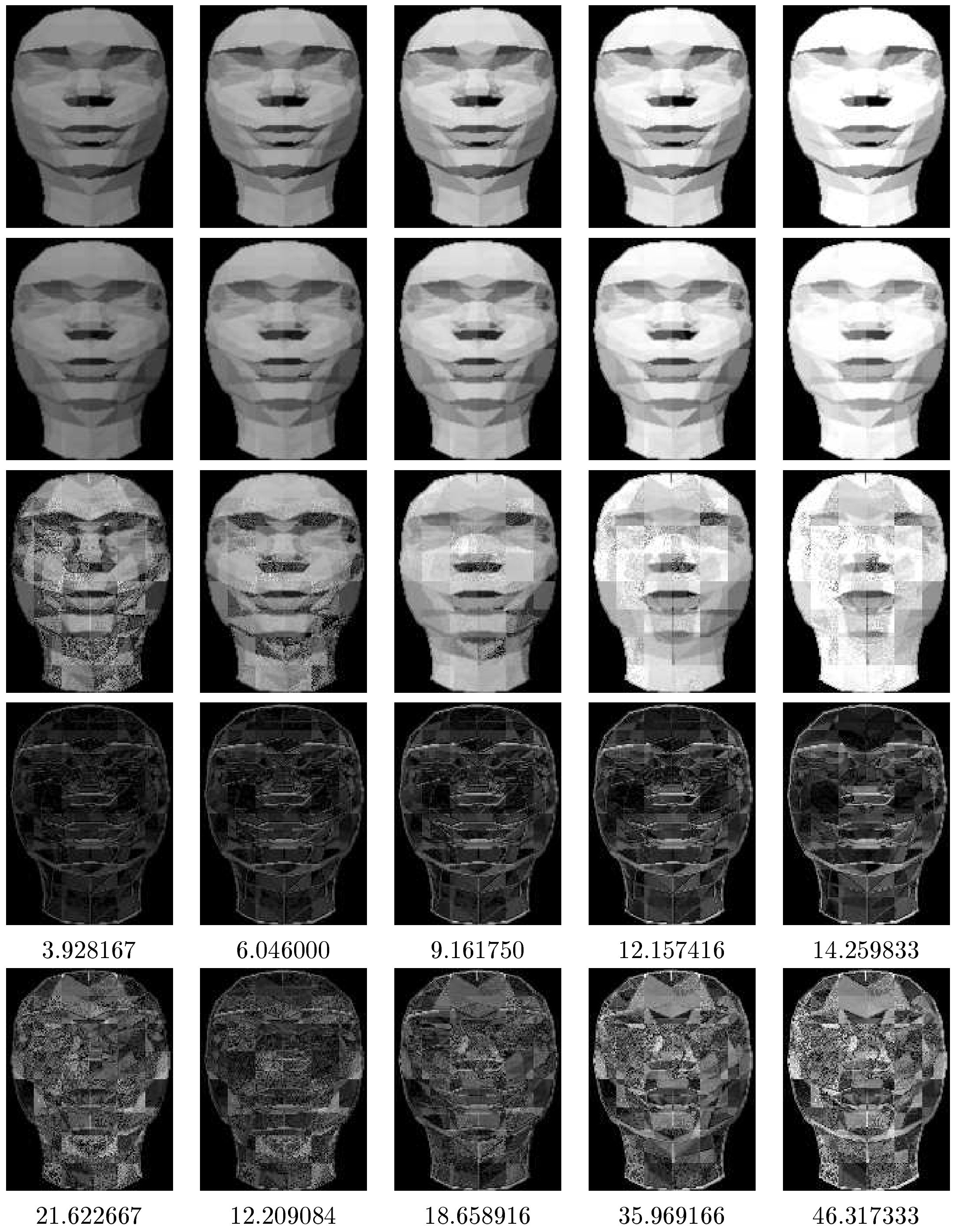


Figure 9: *Same light sources with different intensities.* The first row is the rendered test images. The second row is the image reconstructed by the proposed methods. The third row is the images reconstructed by the traditional method. The fourth and sixth rows are the corresponding residual images. The fifth and seventh rows are errors per pixel of the reconstructed images. $\gamma = 2$ in all these images.

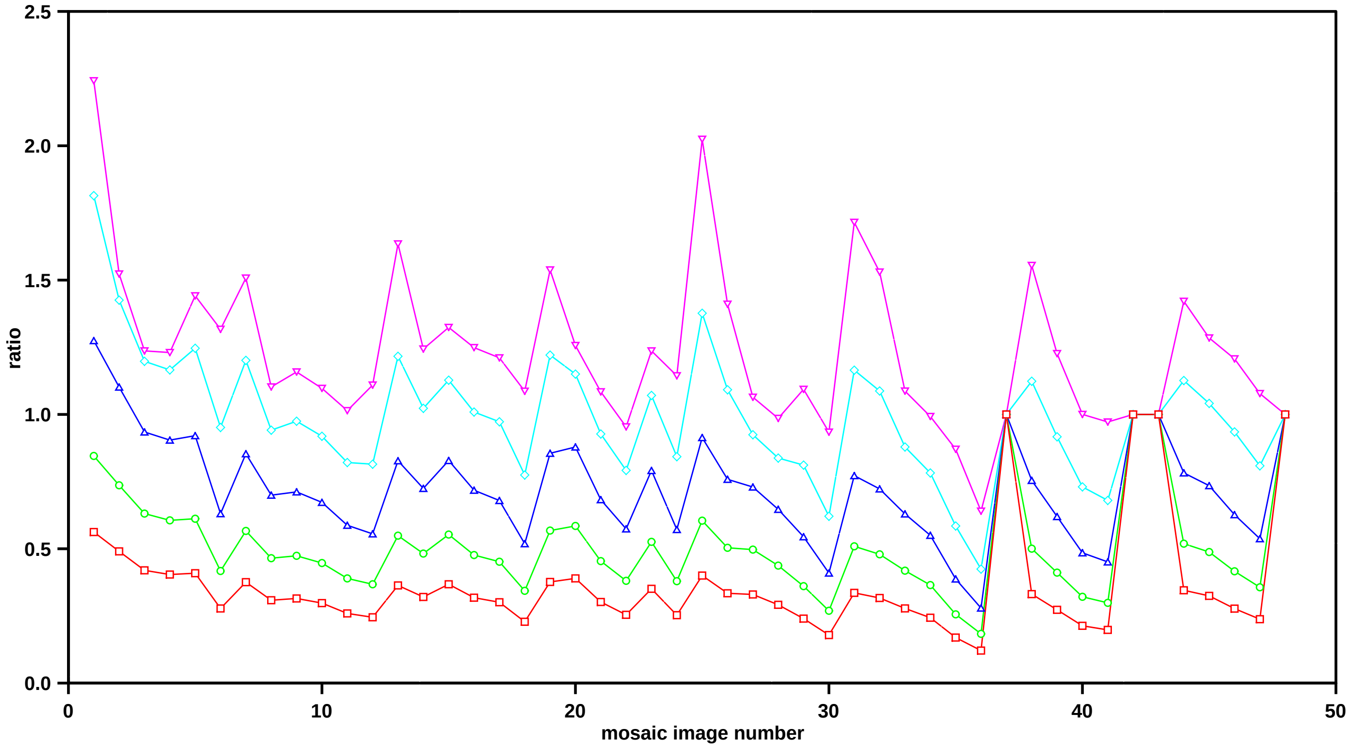


Figure 10: *The ratio α estimated for every mosaic image in every image.*

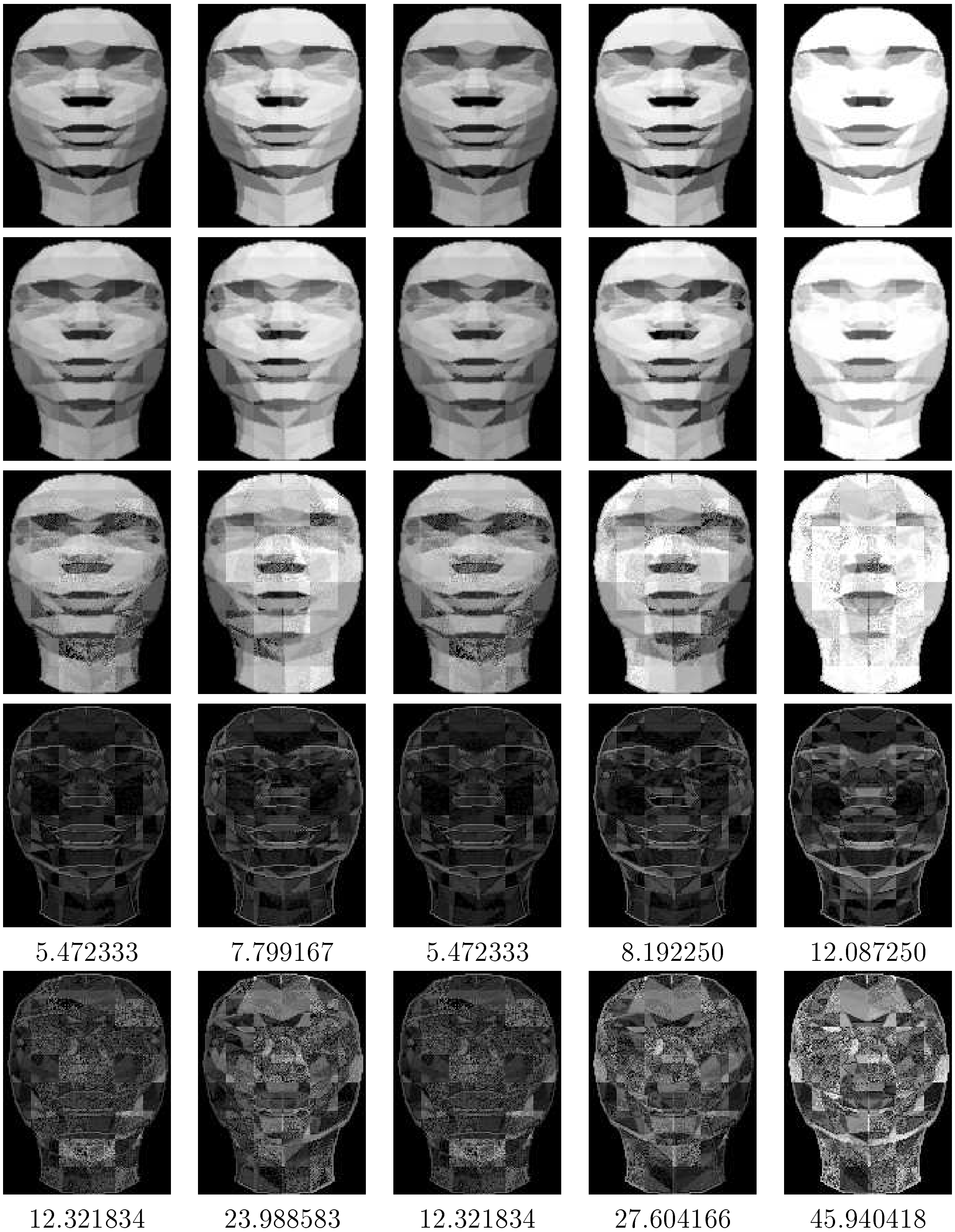


Figure 11: *Multiple light sources test.* The first row is the rendered test image in different light conditions. The second and third row are reconstructed images by the proposed methods and the traditional method. The fourth and sixth rows are corresponding residual images. The fifth and seventh rows are errors per pixel for the reconstructed images. $\gamma = 2$ in all these images.

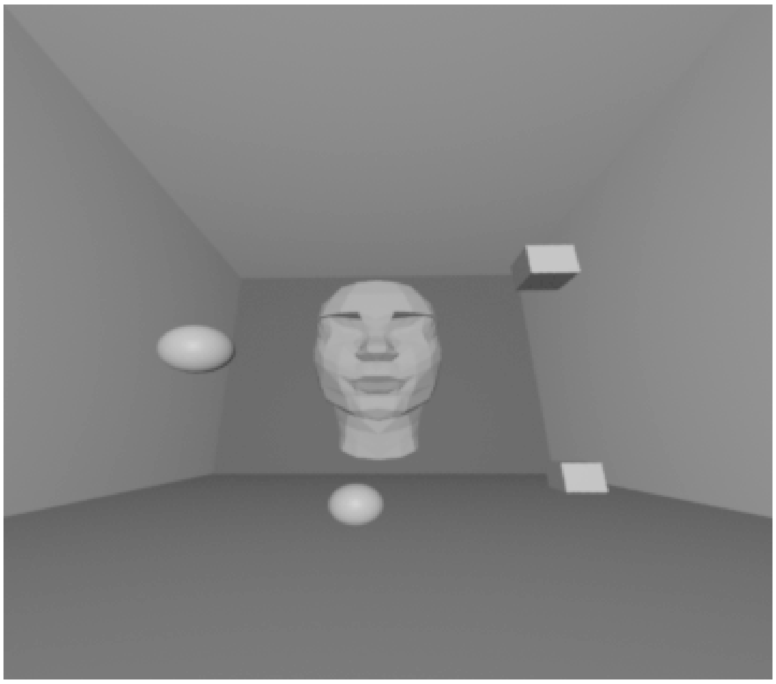


Figure 12: *An image of the experiment setup.*

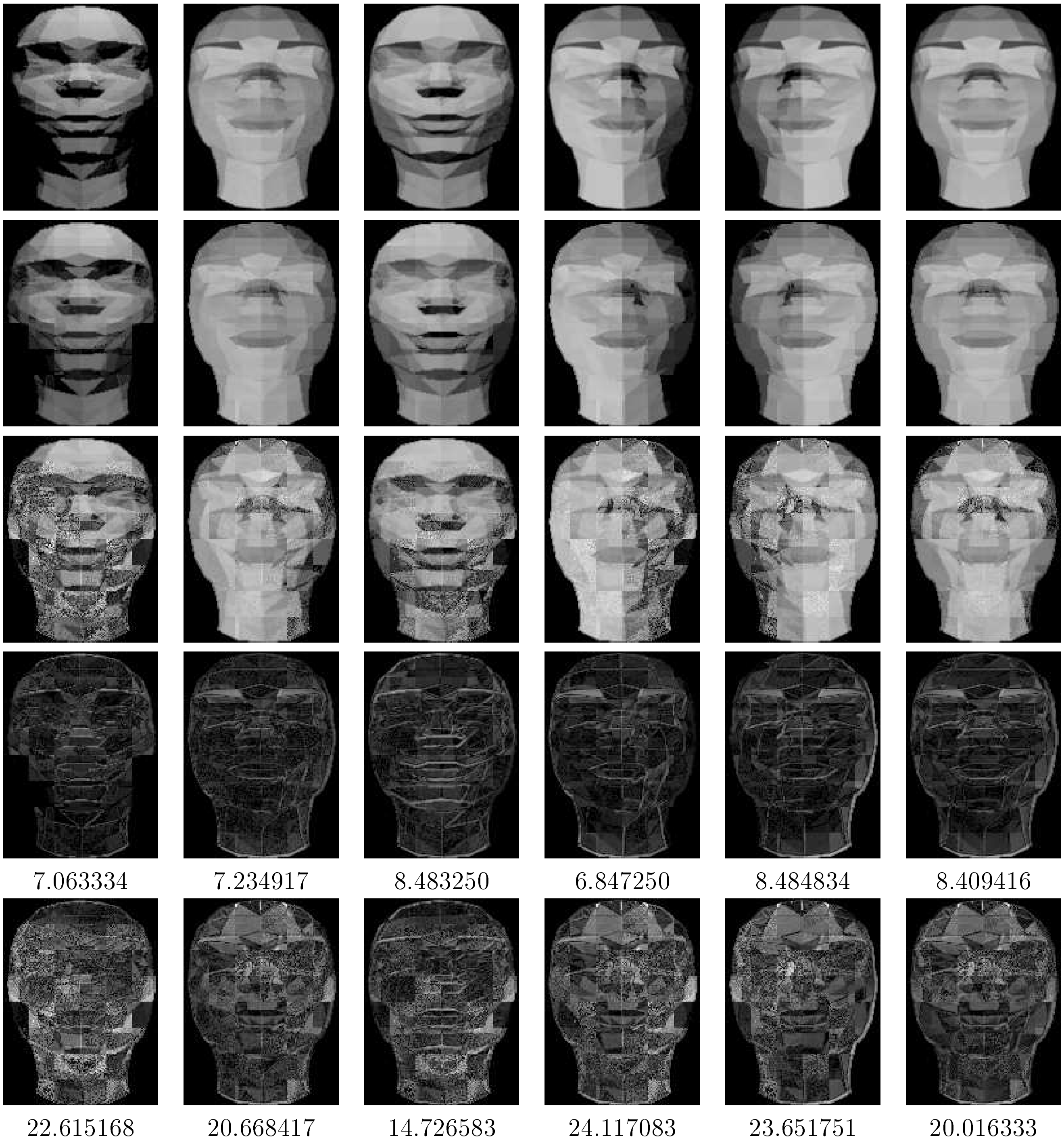


Figure 13: *Test results for area light sources without accounting for interreflection.* The first row is the rendered test images without accounting for interreflection. The second and third rows are the reconstructed images by the proposed methods and the traditional method. The fourth and sixth rows are corresponding residual images. The fifth and seventh rows are errors per pixel of these reconstructed images. $\gamma = 2$ in all these images.

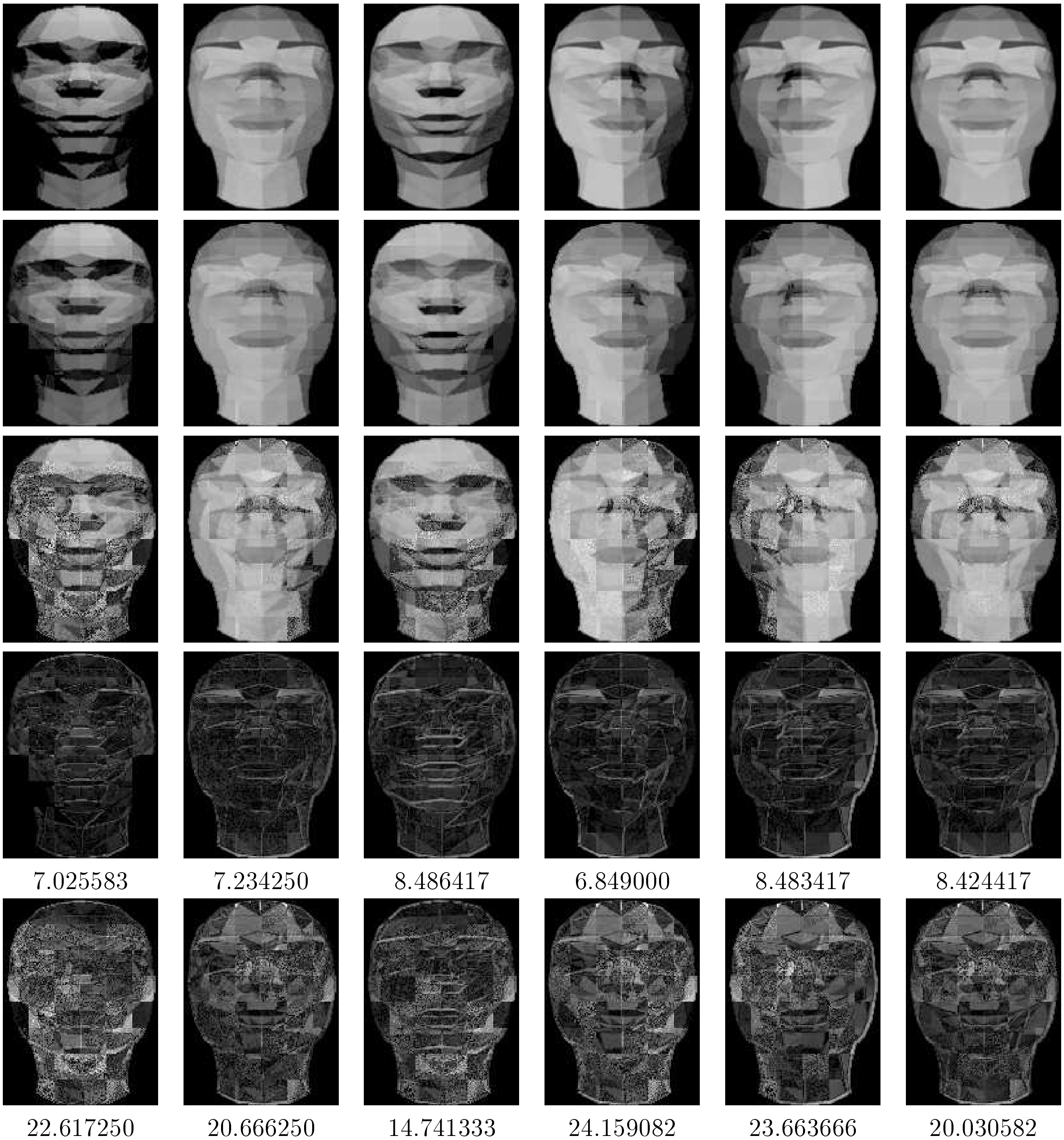
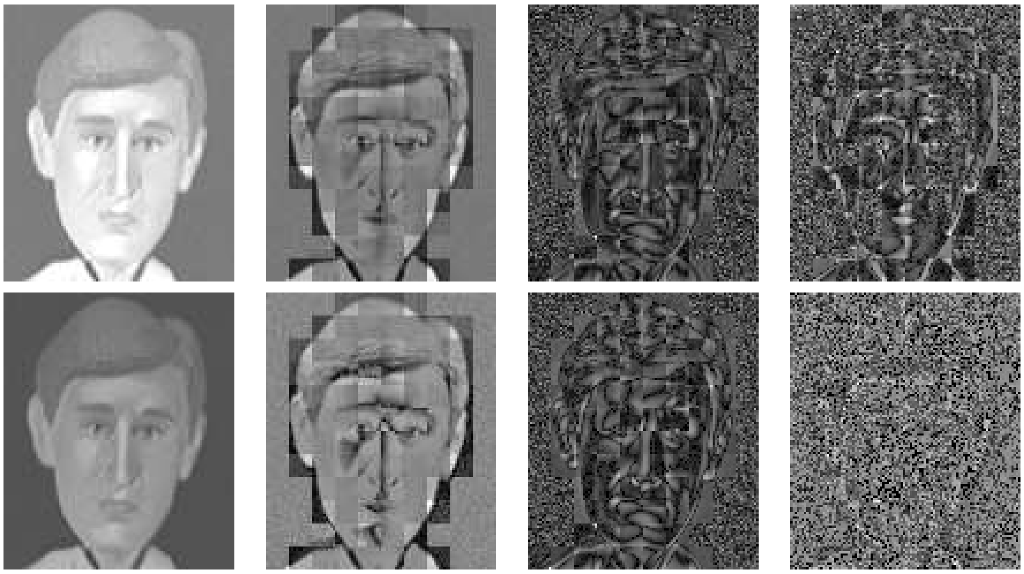


Figure 14: *Test results for area light sources accounting for interreflection.* The first row is the rendered images accounting for interreflection. The second and third rows are reconstructed images by the proposed methods and traditional method. The fourth and sixth rows are corresponding residual images. The fifth and seventh rows are errors per pixel for the reconstructed images. $\gamma = 2$ in all these images.

Figure 15: *The three primary images of a celebrity's face.* $\gamma = 1$.



mean image eigenimages

Figure 16: *The eigenimages and mean images for the real image test.* The images in the first row are generated by the proposed method. The images in the second row are generated by the traditional method. $\gamma = 1$ for the mean images. $\gamma = 2$ for all eigenimages.

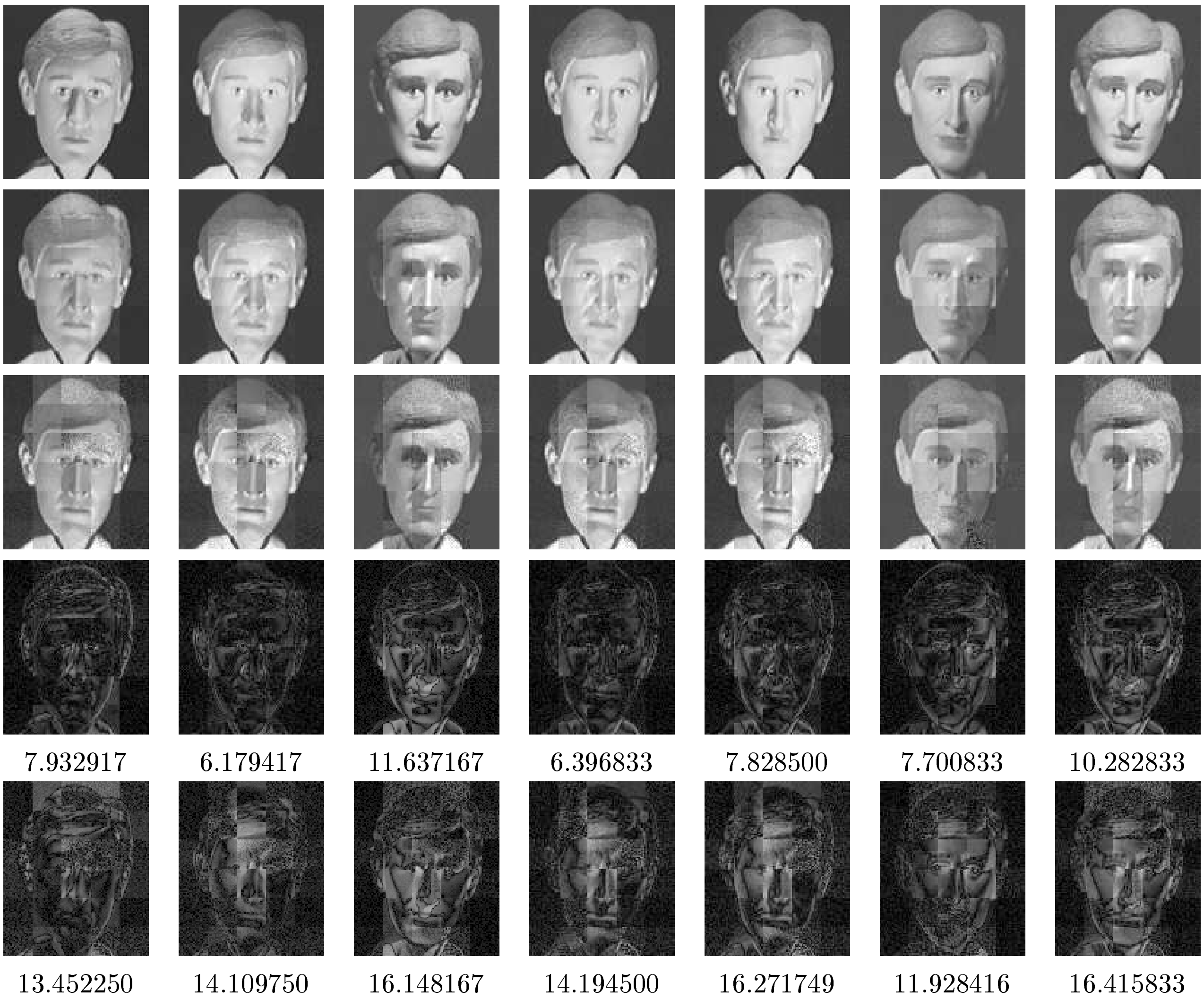


Figure 17: *Test for real images.* The first row is the test images. The second and third rows are reconstructed images by the proposed and traditional methods. The fourth and sixth rows are residual images by the proposed method and the traditional method. The fifth and seventh rows are errors per pixel of the reconstructed images. Three eigenimages are used in the reconstruction since for the traditional method only three eigenimages are available. $\gamma = 2$ for all residual images, $\gamma = 1$ for all other images.

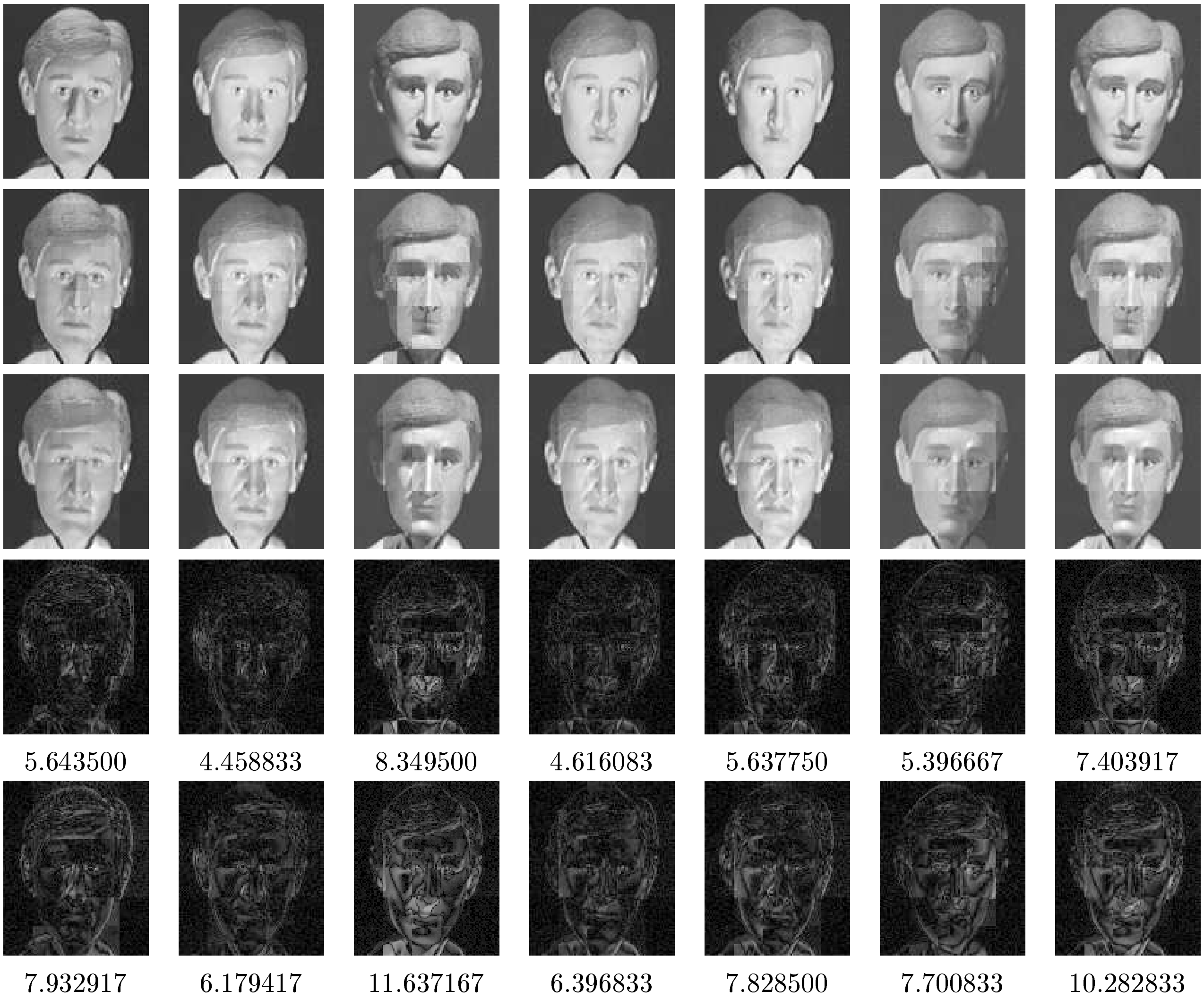


Figure 18: *Comparison of reconstructed images with different sizes of mosaic images.* The first row is the test images. The second and third rows are reconstructed images by the proposed methods with mosaic image sizes of 10×10, and 20×20 respectively. The fourth and sixth rows are corresponding residual images. The fifth and seventh rows are errors per pixel for reconstructed images. $\gamma = 2$ for all residual images, $\gamma = 1$ for all other images.

Figure 19: *The original images and the smoothed reconstructed images by three eigenimages.*

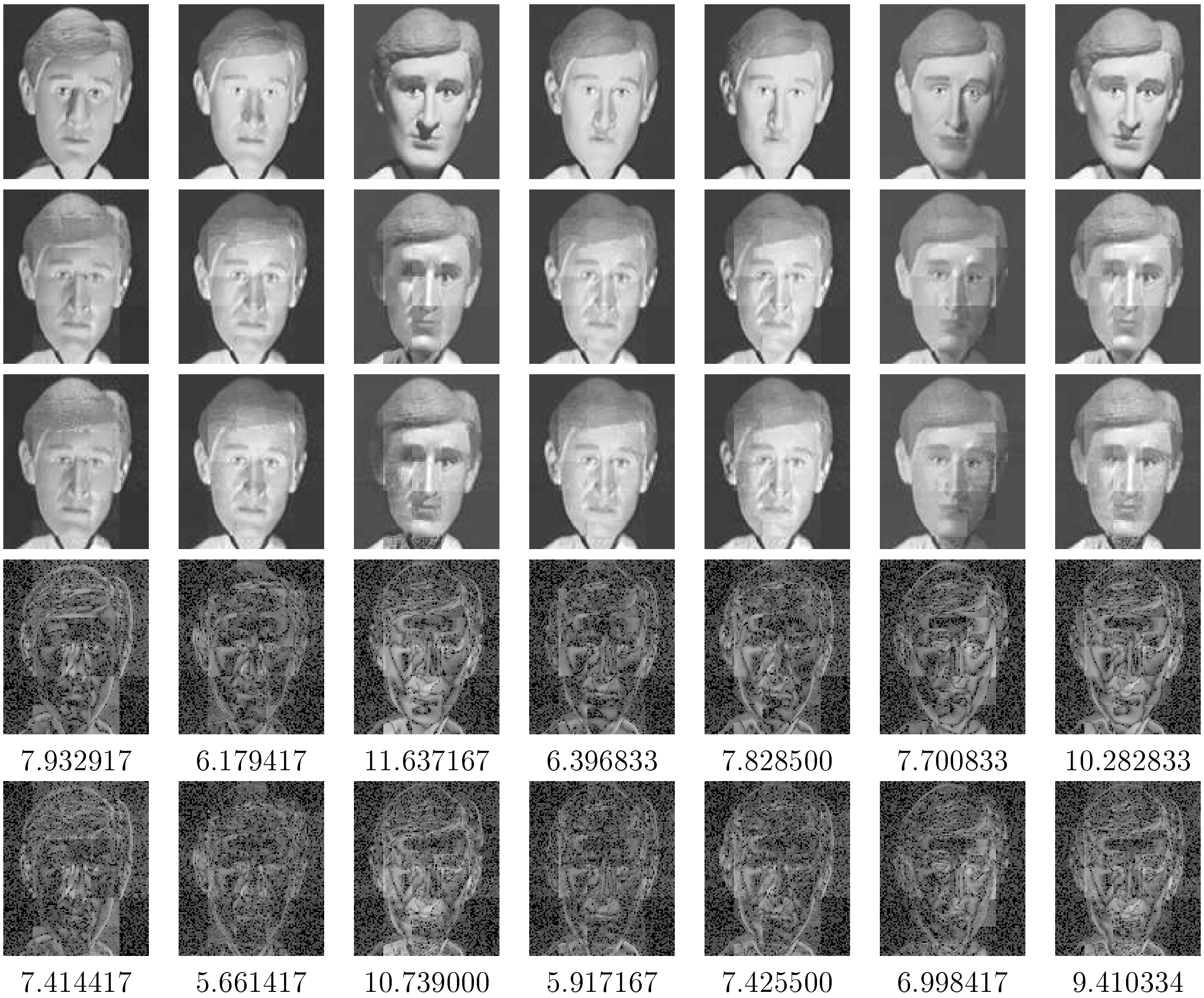


Figure 20: *Using different numbers of eigenimages in the proposed methods.* (The mosaic image is 20×20.) The first row is the original images. The second and third rows are reconstructed images using three and ten eigenimages repectively. The fourth and sixth rows are corresponding residual images. The fifth and the seventh rows are errors per pixel for the reconstructed images. $\gamma = 4$ for all residual images. $\gamma = 1$ for all other images.