To the Graduate Council:

I am submitting herewith a thesis written by Andrew R. Wilson entitled "Class-Based Color Image Synthesis and Object Recognition Under Spatially and Spectrally Varying Illuminants." I have examined the final paper copy of this thesis for form and content and ~~recommend~~ that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Electrical Engineering.

Philip W. Smith

Philip W. Smith, Major Professor

We have read this thesis
and recommend its acceptance:

Daniel B. Koch

Hairong Qi

Accepted for the Council:

Vice Provost and Dean of Graduate Studies

CLASS-BASED COLOR IMAGE SYNTHESIS AND OBJECT RECOGNITION UNDER SPATIALLY AND SPECTRALLY VARYING ILLUMINANTS

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Andrew R. Wilson

December 2002

DEDICATION

This thesis is dedicated to my parents, Bob and Cathy, who have taught me to give 100% in everything that I ever do and to my loving wife, Wendy, who is my source of strength and inspiration. I owe you all more than you could ever know. I love you.

ACKNOWLEDGEMENTS

I would like to thank my instructors and advisors, Phil Smith, Dan Koch, Hairong Qi, and Mike Roberts of the University of Tennessee, Knoxville. Their instruction and encouragement had a beneficial effect on my education. I will carry on what I have learned from them for the rest of my career as an electrical engineer.

I would also like to thank all those who allowed me to take their pictures for testing images. Without the cooperation of these people, my work would have been more difficult, and I owe them a great deal of appreciation.

My parents and my wife have provided the encouragement and support necessary to bring me to this point. I will always appreciate their love and faith.

Most importantly, I would like to thank God. I am grateful for His blessings and guidance in everything that I do.

ABSTRACT

The process of creating a synthetic image of a given example image under arbitrary pose and illumination is known as image synthesis or re-rendering. Existing class-based re-rendering methods require large basis image sets and accurate pixel alignment for accurate image synthesis. In addition to this, these synthesis methods have focused primarily on gray-scale images. Based on the novel concept of surface reflectance quotients, two new image synthesis techniques have been developed that produce high quality images given only a small image set constructed using a single basis object. The image synthesis ability of the two new techniques is demonstrated using both synthetic and real image sets.

The concept of surface reflectance quotients has also been applied to object recognition in which a test object is compared and classified against a set of known objects. In such tasks, variations in illumination conditions may cause difficulties in producing a correct match. The strength of the object recognition technique based on the calculation of the surface reflectance quotient will be shown to be the ability to eliminate the illumination component from the image comparison process. This allows the object recognition task to be carried out without the errors associated with variations in lighting. The operation of this object recognition technique is demonstrated through the use of a image set of human faces.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

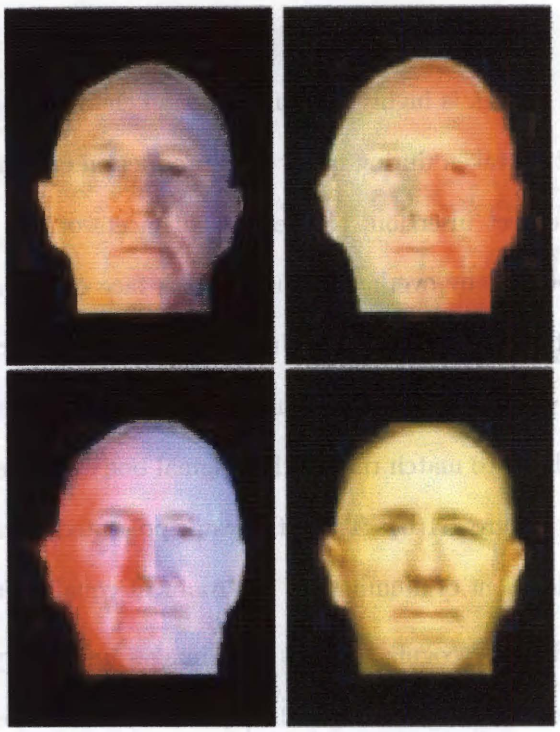
INTRODUCTION

The human visual system is highly sensitive to small changes in photographic lighting. Thus, great care must be taken during many post-production editing tasks, such as image splicing or synthetic object insertion, to preserve all facets of the lighting environment. Figure 1.1(a) was produced by overlaying one person's face onto a second person's body. Because the illumination environments of the new face and the original body are quite different, the average viewer will easily recognize the edit. When the lighting environment of the new face is modified to match that of the original body's, the edit is made much less noticable as shown in Figure 1.1(b). Thus, the work presented in this thesis has focused primarily on the development of illumination environment equalization algorithms.

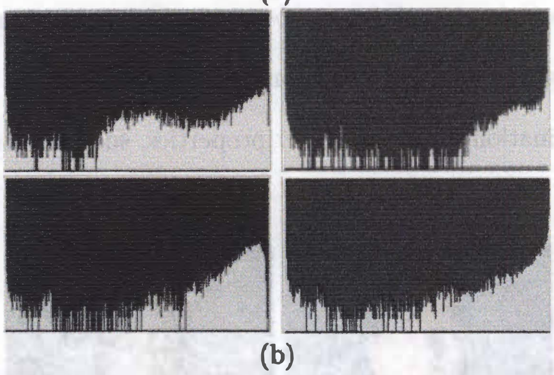
Humans are also able to recognize objects from images under varying illumination conditions. To an observer, the images of Figure1.2(a) are of the same person. The histograms of Figure1.2(b) demonstrate that these images are quite different in a statistical sense. For the human visual recognition system to demonstrate such robustness to changes in lighting, it must be using illumination-invariant object properties, such as relative distance between

(a) (b)

Figure 1.1: A person's face inserted into a scene (a) without illumination matching, and (b) with illumination matching.



(a)



(b)

Figure 1.2: Human recognition examples: (a) Four images of a single face under varying illumination conditions and (b) corresponding image intensity histograms

distinctive features, and/or a mechanism for illumination equalization. Thus, the secondary goal of the work presented was to develop an object recognition algorithm based on illuminant equalization.

The process of generating two-dimensional images of three-dimensional objects under arbitrary pose and illumination conditions is known as image-based re-rendering or image synthesis. There are two methods for performing this task. The first of these methods is to explicitly recover the shape, reflectance, and illumination conditions of the imaged object in order to re-render analytically. The second method takes advantage of the linear nature of the lighting condition by creating arbitraty illumination conditions based on weighted linear combinations of a set of basis illuminants. The second approach is known as the generative approach to image synthesis.

Most existing methods for image synthesis employ a generative approach. A gray scale image may be represented mathematically as

$$I(u,v) = \alpha(u,v)\vec{n}(u,v) \bullet \vec{n}_s(u,v) \tag{1.1}$$

where $\alpha(u,v)$ is the albedo of the imaged object, $\vec{n}(u,v)$ is a 3×1 vector representing the illumination direction, and $\vec{n}_s(u,v)$ is a 3×1 vector representing the surface normal map of the imaged object. An arbitrarily illuminated image of a three-dimensional Lambertian object may be generated via a weighted sum of images of the given object. This new image may be written as

$$\begin{aligned} I_r(u,v) &= \sum_{i=1}^{m} w_i I_i(u,v) \\ &= \sum_{i=1}^{m} w_i \alpha(u,v)\left[\vec{n}_i(u,v) \bullet \vec{n}_s(u,v)\right] \\ &= \alpha(u,v)\left(\sum_{i=1}^{m} w_i \vec{n}_i(u,v)\right) \bullet \vec{n}_s(u,v) \end{aligned} \tag{1.2}$$

where $w_1, w_2, \ldots, w_m$ are the scalar weighting coefficients of images $I_1(u,v) \ldots I_m(u,v)$. Hallinan [1] and others [2, 3] have shown that this image may be generated via a weighted sum of at least three basis images of the object. In each of these cases, it is necessary to create a basis set of images with linearly independent lighting conditions for each object to be re-rendered. Although this process may lead to highly accurate results, all solutions are unique to a single object. The custom illuminated objects in Figure 1.3(b) were created as weighted sums of the images present in the basis set shown in Figure 1.3(a). However, when attempting to reconstruct a new object, Figure 1.3(c), from the same basis set, the resulting image is simply a new custom illuminated image of the original basis object, Figure 1.3(c).

The requirement of creating a basis set of images for each unique imaged object places a limitation on the practical use of Hallinan's re-rendering method. When applying this reillumination technique to human faces, this limitation may be seen. If it is desired to have the ability to reilluminate any human face, it would be necessary to create an image library consisting of three images of each person in the world. Therefore, with a population of 6 billion people, it would require an image library of 18 billion images in order to properly re-render each unique human face. This poses quite a problem due to both limited storage and the inability to obtain all necessary images.

In order to make the generative method of image synthesis more tractable, the basis image concept was further generalized by Vetter and others[4, 5, 6, 7] to allow for the construction of linear object classes. In class-based approaches, objects are assigned to groups based on similar properties, such as shape or texture, and a single basis set is constructed for each group. For example, if all human faces are assumed to have approximately the same shape, one basis set could be constructed for the generation of all faces assuming that the set spans the entire space of possible textures. Unfortunately, if the size of the basis set is too small to span the entire space of facial textures, inaccurate reconstructions result. As an unfortunate result, the performance of class-based re-rendering approaches degrades quickly if basis sets are created that are too small to span the space of similar object properties (i.e. skin texture variations on human faces), and the pixel-wise alignment

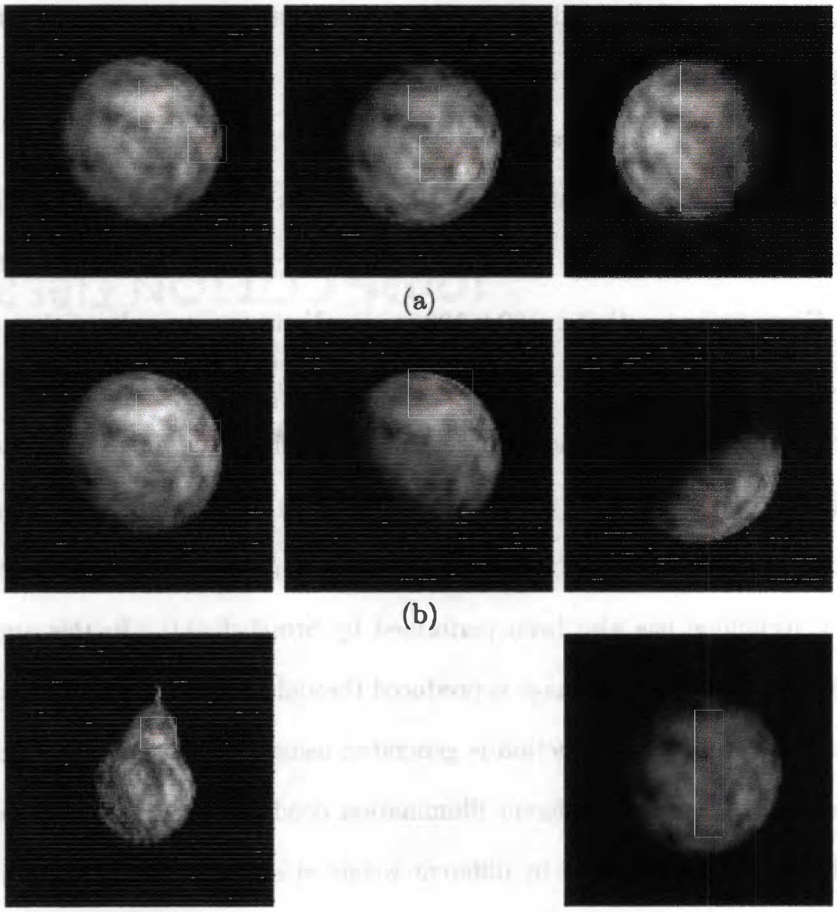


Figure 1.3: Hallinan re-rendering technique: (a) basis set, (b) arbitrarily illuminated images, (c) image of a new object, and (d) reconstruction of new object using the original basis set.

of imaged objects is inadequate. Figure 1.4 illustrates the limitations of this image synthesis technique. The basis set in Figure 1.4(a) has been used to reconstruct the example image in Figure 1.4(b). However, due to the fact that this small basis set does not span the space of similar object properties, this technique fails to accurately reconstruct the example image, as seen in Figure 1.4(c). The need for a large basis set makes this algorithm impractical for many situations in which image re-rendering is required.

The introduction of image classes may dramatically reduce the size of the image set needed from that of Hallinan. Rather than the need for three images of each object in this class, the requirement placed on the size of the basis set is set by the size of the images in the set. Given an image that is 100×100 in size, Vetter's approach requires a basis set of only 10,000 images. Although this basis set is still quite difficult to construct, this is a great improvement over the previous approach that required a set of 18 billion images.

Shashua and Riklin-Raviv [8, 9, 10] recently proposed the use of an illumination invariant image as an alternative to traditional class-based re-rendering. Further research on this re-rendering technique has also been performed by Stoschek [11]. In this approach, an illumination invariant quotient image is produced through the pixel-wise division of a given example image, and its reconstruction is generated using a class-based approach. Images of this example object under different illumination conditions can then be reconstructed through pixel-wise multiplication by different weighted sums of the basis image set. The use of this illumination invariant image helped to dramatically reduce the size of the basis set. In contrast to the large basis sets needed to perform other image synthesis techniques, this method requires a minimum of only three basis images. The perceived quality of images re-rendered via the quotient image technique remains quite high given both small basis sets and poor image alignment. However, the reported quotient image technique is only valid for spectrally flat, or white, illuminants. Given the basis and example images in Figure 1.5(a) and (b), respectively, the reilluminated images in Figure 1.5(c) were created using the same

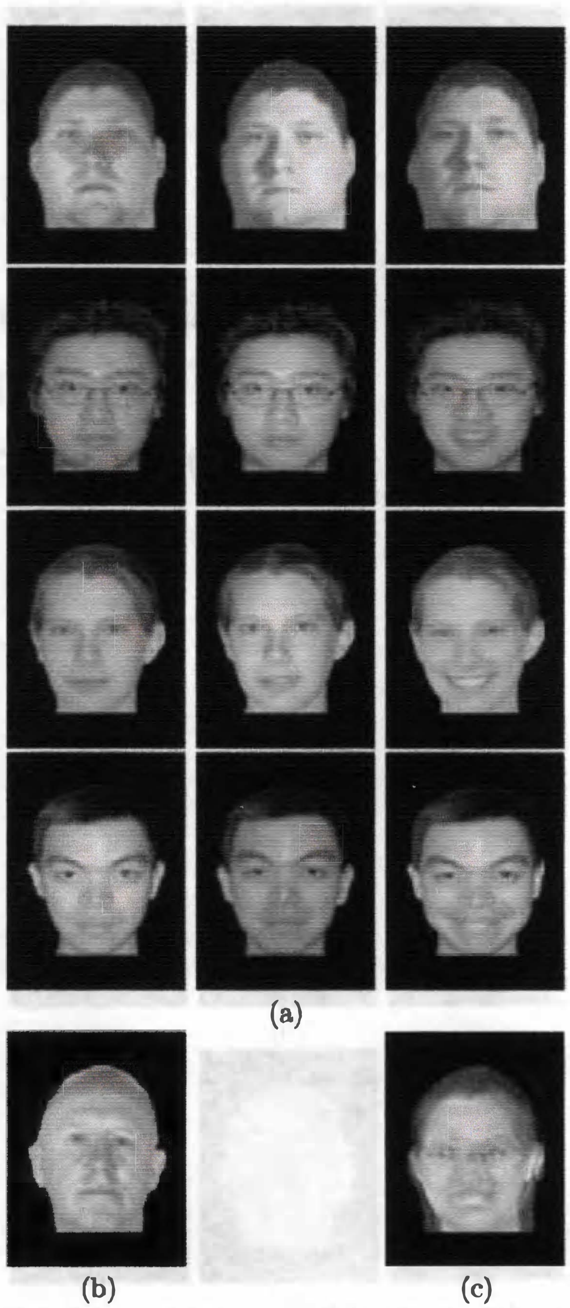


Figure 1.4: Vetter re-rendering technique: (a) basis set, (b) original image, and (c) reconstructed image

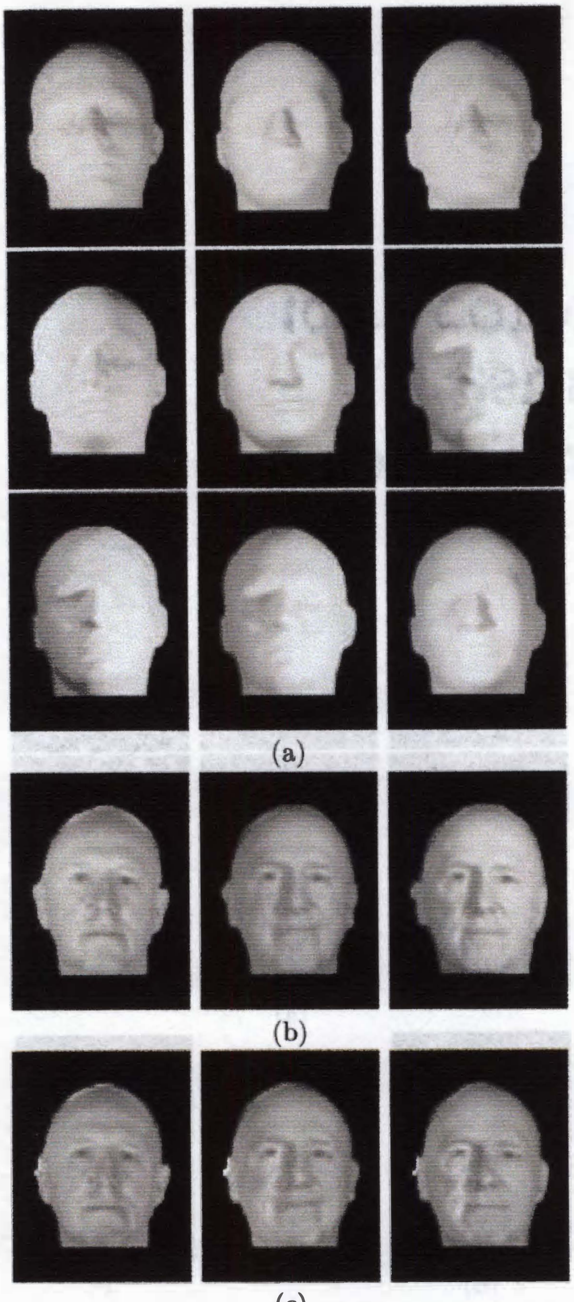


Figure 1.5: Shashua re-rendering technique: (a) basis set, (b) example images, and (c) re-rendered images

weighting coefficients. Although the images in Figure 1.5(b) are of the same person, the resulting reilluminated images show differences in their appearance due to the fact that the illuminants in each of the basis images as well as those of the example images are not flat. Due to the variations in color of the illuminants, each of the original images appears to have a slight variation in surface texture that leads to these reconstruction errors.

In each of the generative algorithms discussed above, one or more of the following assumptions must be made in order for the algorithms to operate:

1) Exact alignment between images
2) Large basis set of images
3) Flat white illuminants

These conditions limit the utility of these algorithms in uncontrolled application environments. In such environments, construction of large basis sets is costly, time consuming, and still does not guarantee success. Also, exact alignment is difficult to enforce in such environments, and it is very rare for flat illuminants to be present. For these reasons, tools must be created in an attempt to overcome the limitations of previous image reillumination techniques.

1.1 RESEARCH GOAL

The objective of this research is to develop generative image re-rendering algorithms based on the calculation of an illumination invariant color image defined as the surface reflectance quotient. These algorithms are to be developed with the purpose of overcoming the need for a large basis set and flat illuminants in order to increase their utility in uncontrolled environments. Two new algorithms for class-based color image synthesis under both spatially and spectrally varying illumination conditions are presented in this thesis. Given a color example image of an object in the class, the first re-rendering technique allows the user to specify a new custom illumination condition for the reconstructed image. The second re-rendering technique matches the illumination of the example image to the lighting environment of a second image of an object in the class.

The utility of this approach to image re-rendering is illustrated through the development of a new color image based facial recognition algorithm. In this algorithm, the surface reflectance quotient is calculated for the given example image. Using it as a signature image, the surface reflectance quotient is then compared to those of the image database to determine the closest match. Due to the illumination invariant nature of the surface reflectance quotient, this comparison is performed without the influence of variations in illumination. Therefore, the comparison is based on the difference in surface textures of the imaged objects.

1.2 UNIQUE CONTRIBUTIONS

The unique contributions of this thesis are as follows:

1) Two new algorithms for re-rendering which employ
   - small basis sets, and
   - colored illuminants.
2) Facial recognition algorithm based on the calculation of the surface reflectance quotient

1.3 ORGANIZATIONAL OVERVIEW

In Chapter 2, I will discuss the mathematical framework for the development of the surface reflectance quotient. This will be followed in Chapter 3 by an algorithm listing for both the custom re-rendering and the re-rendering by example processes and experimental results for each of these algorithms. Chapter 4 will be concerned with an explanation of the object recognition process as well as an algorithm listing and experimental results for the proposed technique. Chapter 5 will contain a summary of the work performed as well as possibilities for future expansion of this work.

CHAPTER 2

THE SURFACE REFLECTANCE QUOTIENT

For color images, the apparent image intensity vector of a point on a three-dimensional Lambertian surface, $\vec{I}(u,v) \in \mathcal{R}^3$, can be modeled using the expression

$$\vec{I}(u,v) = \mathbf{S}(u,v)\mathbf{L}(u,v)\vec{n}(u,v) \tag{2.1}$$

where $\mathbf{S}(u,v) \in \mathcal{R}^{3\times3}$ is a diagonal matrix which represents the surface reflectance function, $\mathbf{L}(u,v) \in \mathcal{R}^{3\times3}$ is the illumination function, and $\vec{n}(u,v) \in \mathcal{R}^3$ is the surface normal map of the imaged object [12]. If the illumination function is held constant over all (u,v), this 3×3 matrix may be written as a weighted sum of $N \geq 9$ known illuminant matrices. In this case, equation 2.1 may be rewritten as

$$\vec{I}(u,v) = \mathbf{S}(u,v)\left(\sum_{i=1}^{N}\alpha_i\mathbf{L}_i\right)\vec{n}(u,v) \tag{2.2}$$

where $\alpha_i \ldots \alpha_N$ are scalar weighting coefficients.

If a class of objects is defined such that the surface reflectance and normal maps must be equivalent across all members, equation 2.2 can be written as the weighted sum of $N \geq dim(\mathbf{L}) = 3 \times 3 = 9$ images

$$\begin{aligned} \vec{I}(u,v) &= \sum_{i=1}^{N}(\alpha_i\mathbf{S}(u,v)\mathbf{L}_i\vec{n}(u,v)) \\ &= \sum_{i=1}^{N}\alpha_i\vec{I}_i(u,v) \end{aligned} \tag{2.3}$$

which is the weighted sum of known images, $\vec{I}_i(u,v)$, of objects in the class. Thus, any image of a class member can be synthesized by determining the proper coefficients, α_i, such that

$$\min_{\alpha_i}\left\|\vec{I}_s - \sum_{i=1}^{N}\alpha_i\vec{I}_i(u,v)\right\|^2 \tag{2.4}$$

is satisfied. To calculate the proper values of α_i that satisfy equation 2.4, the basis images, $\vec{I_i}(u,v)$, are placed into the columns of the $mn \times N$ matrix $\mathbf{A}$ where m and n are the rows and columns of the images, respectively. It is then possible to calculate α_i as the least squares estimate

$$\vec{\alpha} = \mathbf{A}^{\#}\vec{y_s} \tag{2.5}$$

where

$$\mathbf{A}^{\#} = \left(\mathbf{A}^{\top}\mathbf{A}\right)^{-1} \tag{2.6}$$

is the pseudoinverse of $\mathbf{A}$, and $\vec{y_s}$ is the column vector containing the elements of $\vec{I_s}(u,v)$. Approaches to image re-rendering which employ expressions such as equation 2.3 are termed generative because new images are literally 'generated' by weighted combinations of other images.

The generative approach described above is made more complex when the class membership criteria is relaxed to allow for variance in surface texture due to the bi-linear nature of the relation between the surface reflectance and illumination functions. If the surface reflectance is allowed to vary across the objects of the class, an image of an object under a desired illumination condition can be constructed using the expression

$$\vec{I_g}(u,v) = \left(\sum_{j=1}^{M} \gamma_j \mathbf{S}_j(u,v)\right)\left(\sum_{i=1}^{N} \alpha_i \mathbf{L}_i\right)\vec{n}(u,v) \tag{2.7}$$

While a linear solution exists for solving equation 2.7 for the appropriate coefficients γ_j and α_i such that the constraint

$$\min_{\gamma,\alpha} ||\vec{I}(u,v) - \vec{I_g}(u,v)||^2$$

is satisfied [13, 14], numerical instability and over-fitting due to bias result from the need to span the linear solution space of dimension NM. In addition, the basis set must contain N different images of M class objects to produce accurate reconstructions. Thus, the number of basis set images must increase in proportion to the number of objects in the class. This proportional growth can be a prohibitive constraint in applications that require large object classes, such as re-rendering of the human face.

It is possible, however, to construct an illumination invariant representation for each class member in the presence of surface reflectance variations using only a small basis set constructed from a single class object. Defined as the surface reflectance quotient, this quantity can be used to perform high quality color image synthesis.

Let $\vec{I}_s(u,v)$ and $\vec{I}_r(u,v)$ be images of two member objects taken under the same lighting condition, $\bar{\mathbf{L}}$. Also, let $\tilde{I}_r(u,v)$ be an image of the second object taken under a second illumination field, $\tilde{\mathbf{L}}$. Because the two objects belong to the same class, the surface normal maps are assumed equivalent. Noting that equation 2.1 can be rewritten as

$$\mathbf{S}^{-1}(u,v)\vec{I}(u,v) = \mathbf{L}\vec{n}(u,v) \tag{2.8}$$

the relations for $\vec{I}_s(u,v)$ and $\vec{I}_r(u,v)$ may be written as

$$\begin{aligned} \mathbf{S}_s^{-1}(u,v)\vec{I}_s(u,v) &= \bar{\mathbf{L}}\vec{n}_s(u,v) \\ \mathbf{S}_r^{-1}(u,v)\vec{I}_r(u,v) &= \bar{\mathbf{L}}\vec{n}_r(u,v), \end{aligned} \tag{2.9}$$

From this result, the following illumination invariant relation between the two images may be derived,

$$\begin{aligned} \mathbf{S}_s^{-1}(u,v)\vec{I}_s(u,v) &= \mathbf{S}_r^{-1}(u,v)\vec{I}_r(u,v) \\ \vec{I}_s(u,v) &= \mathbf{S}_s(u,v)\mathbf{S}_r^{-1}(u,v)\vec{I}_r(u,v) \\ &= \mathbf{S}_{s,r}^q(u,v)\vec{I}_r(u,v) \end{aligned} \tag{2.10}$$

where $\mathbf{S}_{s,r}^q(u,v) \in \mathcal{R}^{3\times 3}$ is a diagonal matrix defined as the surface reflectance quotient of $\mathbf{S}_s(u,v)$ and $\mathbf{S}_r(u,v)$. Because $\mathbf{S}_{s,r}^q(u,v)$ is diagonal, its entries may be calculated using the expression

$$\mathrm{diag}(\mathbf{S}_{s,r}^q(u,v)) = \vec{I}_s(u,v)/\vec{I}_r(u,v) \tag{2.11}$$

which is the element-by-element division of $\vec{I}_s(u,v)$ and $\vec{I}_r(u,v)$. The images in Figure 2.1 are surface reflectance quotients of various face images constructed using the expression in Equation 2.11. Due to the lack of the illumination component of the original images, these surface reflectance quotient images appear very flat. By introducing a new illumination component to these images, it is now possible to create a synthetic image of the given object under this synthetic lighting condition.

Given the surface reflectance quotient, a new image of the object in $\vec{I}_s(u,v)$ can then be synthesized to appear as it would in the lighting environment of $\tilde{I}_r(u,v)$ using the relation

$$\begin{aligned}
\tilde{I}_s(u,v) &= \mathbf{S}^q_{s,r}(u,v)\tilde{I}_r(u,v) \\
&= \mathbf{S}_s(u,v)\mathbf{S}_r^{-1}(u,v)\mathbf{S}_r(u,v)\tilde{\mathbf{L}}\vec{n}(u,v) \\
&= \mathbf{S}_s(u,v)\tilde{\mathbf{L}}\vec{n}(u,v).
\end{aligned} \tag{2.12}$$

Thus, images of class members can be synthesized under any chosen illumination condition using the surface reflectance quotient, given that images of one class member under all possible lighting conditions are available.

While it may be possible to construct an image set for one object in the class that spans all desired illumination conditions, the generative approach to image synthesis provides a more practical alternative for construction of images that belong to this illumination spanning set. Consider the problem of trying to reconstruct the image, $\vec{I}_s(u,v)$, using the generative approach given a basis set consisting only of images of the second object with surface reflectance $S_r(u,v)$. The weighted-sum solution, $\bar{I}_s(u,v)$, would be such that

$$\begin{aligned}
\bar{I}(u,v) &= \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v) \\
&= \mathbf{S}_r(u,v)\left(\sum_{i=1}^{N} \alpha_i \mathbf{L}_i\right)\vec{n}(u,v) \\
&= \mathbf{S}_r(u,v)\bar{\mathbf{L}}(u,v)\vec{n}(u,v) \\
&= \vec{I}_r(u,v).
\end{aligned} \tag{2.13}$$

Figure 2.1: Surface reflectance quotients for various human faces.

Hence, the basis object image, $\vec{I}_r(u, v)$, required to construct a surface reflectance quotient for a given image $\vec{I}_s(u, v)$ can be computed using a basis image set of at least nine images of a single object under various illumination fields.

The second image of the basis object, $\tilde{I}_r(u, v)$, required to synthesize new images under varying illuminants using the surface reflectance quotient can also be constructed from the basis set. Two algorithms will be presented that employ two alternative methods for the construction of $\tilde{I}_r(u, v)$ using either a user-specified lighting environment or a second, illumination example image of a member object.

CHAPTER 3

RE-RENDERING ALGORITHM AND EXPERIMENTAL RESULTS

The mathematical framework discussed in the previous chapter provides the basis for creating algorithms for the process of image synthesis. Re-rendering using the surface reflectance quotient can be accomplished either through user-specification of a custom illumination field or through illumination matching.

3.1 CUSTOM RE-RENDERING

To perform image synthesis under a custom light-field, the user defines a set of re-rendering coefficients, β_i, from which the required basis object image, $\vec{I}_r(u,v)$, is constructed as

$$\vec{I}_r(u,v) = \sum_{i=1}^{N} \beta_i \vec{I}_i(u,v). \tag{3.1}$$

The synthesized image can then be computed using equation 2.12, given the surface reflectance quotient. This process is known as re-rendering by custom specification.

As an alternative, this algorithm can compute the necessary coefficients, β_i, given a specified illumination function, $\tilde{\mathbf{L}}$, if the illuminant functions, $\mathbf{L}_i$, are known for the basis set. This is done by determining the re-rendering coefficients that satisfy

$$\tilde{\mathbf{L}} = \sum_{i=1}^{N} \beta_i \mathbf{L}_i. \tag{3.2}$$

In order to calculate these coefficients, the basis set illumination functions, $\mathbf{L}_i$, and the desired synthetic illumination function, $\tilde{\mathbf{L}}$, must be vectorized. Given an illumination function of the form

$$\mathbf{L} = \begin{bmatrix} l_{1,1} & l_{1,2} & l_{1,3} \\ l_{2,1} & l_{2,2} & l_{2,3} \\ l_{3,1} & l_{3,2} & l_{3,3} \end{bmatrix}, \tag{3.3}$$

its vectorized form may be calculated as

$$\vec{\mathcal{L}} = \begin{bmatrix} l_{1,1} \\ l_{2,1} \\ l_{3,1} \\ l_{1,2} \\ l_{2,2} \\ l_{3,2} \\ l_{1,3} \\ l_{2,3} \\ l_{3,3} \end{bmatrix}. \tag{3.4}$$

Once the vectorized forms of the illumination functions have been determined, the matrix $\mathcal{I} \in \mathcal{R}^{9 \times N}$ may be formed as

$$\mathcal{I} = \left[\ \vec{\mathcal{L}}_1 \vec{\mathcal{L}}_2 \ldots \vec{\mathcal{L}}_N\ \right] \tag{3.5}$$

with columns containing each of the vectorized basis set illuminants, $\vec{\mathcal{L}}_i$. It may now be shown that the desired illumination condition, $\tilde{\mathcal{L}}$, may be obtained from the matrix multiplication

$$\tilde{\mathcal{L}} = \mathcal{I}\vec{\beta} \tag{3.6}$$

where $\vec{\beta} \in \mathcal{R}^{N \times 1}$ contains the re-rendering coefficients $\beta_1, \beta_2, \ldots, \beta_N$. From this equation, it may be seen that the re-rendering coefficient vector, $\vec{\beta}$, may be determined as

$$\vec{\beta} = \mathcal{I}^{\#}\tilde{\mathcal{L}} \tag{3.7}$$

where $\mathcal{I}^{\#} \in \mathcal{R}^{N \times N}$ is the pseudoinverse of $\mathcal{I}$ calculated as in equation 2.6. The re-rendering coefficients contained in $\vec{\beta}$ may then be used to synthesize a new image with the desired illumination function $\tilde{\mathbf{L}}$.

The custom re-rendering algorithm is outlined below.

Custom Re-rendering Algorithm

1) Calculate the first basis object image

$$\vec{I}_r(u,v) = \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v)$$

such that the constraint

$$\min_{\alpha_i} \left\| \vec{I}_s(u,v) - \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v) \right\|^2$$

is satisfied.

2) Compute the surface reflectance quotient

$$\text{diag}(\mathbf{S}^q_{s,r}(u,v)) = \vec{I}_s(u,v)/\vec{I}_r(u,v).$$

3) If the re-rendering coefficients, β_i, are not specified, compute β_i such that

$$\min_{\beta_i} \left\| \tilde{\mathbf{L}} - \sum_{i=1}^{N} \beta_i \mathbf{L}_i \right\|^2 .$$

4) Construct the second basis object image, $\tilde{I}_r(u,v)$, under the desired illumination condition

$$\tilde{I}_r(u,v) = \sum_{i=1}^{N} \beta_i \vec{I}_i(u,v).$$

5) Construct the re-rendered image, $\tilde{I}_s(u,v)$, under the desired illumination condition using the surface reflectance quotient

$$\tilde{I}_s(u,v) = \mathbf{S}^q_{s,r}(u,v) \tilde{I}_r(u,v).$$

3.2 RE-RENDERING BY EXAMPLE

The surface reflectance quotient can also be used to synthesize a new image of a class object, $\tilde{I}_s(u,v)$, under the same illumination conditions found in a second, or illumination example, image, $\vec{I}_e(u,v)$. In order to recreate the lighting conditions in $\vec{I}_e(u,v)$, the re-rendering coefficients, β_i, for the generation of the second basis object image, $\tilde{I}_r(u,v)$, must be computed such that the constraint

$$\min_{\beta_i} \left\| \vec{I}_e(u,v) - \sum_{i=0}^{N} \beta_i \vec{I}_i(u,v) \right\|$$

is satisfied. Once these coefficients have been determined, the second basis object image, $\vec{I}_r(u,v)$, can be generated using equation 3.1, and the synthesized image, $\tilde{I}_s(u,v)$, can be constructed using equation 2.12. This process is defined as re-rendering by example, or illumination matching. An outline of the re-rendering by example algorithm is provided below.

Re-rendering by Example Algorithm

1) Calculate the first basis object image

$$\vec{I}_r(u,v) = \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v)$$

such that the constraint

$$\min_{\alpha_i} \left\| \vec{I}_s(u,v) - \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v) \right\|^2,$$

is satisfied.

2) Compute the surface reflectance quotient

$$\mathrm{diag}(\mathbf{S}^q_{s,r}(u,v)) = \vec{I}_s(u,v)/\vec{I}_r(u,v).$$

3) Calculate the re-rendering coefficients, β_i, such that

$$\min_{\beta_i} \left\| \vec{I}_e(u,v) - \sum_{i=0}^{N} \beta_i \vec{I}_i(u,v) \right\|.$$

4) Construct the second basis object image, $\tilde{I}_r(u,v)$, under the new illumination condition

$$\tilde{I}_r(u,v) = \sum_{i=1}^{N} \beta_i \vec{I}_i(u,v).$$

5) Construct the re-rendered image, $\tilde{I}_s(u,v)$, under the new illumination condition using the surface reflectance quotient

$$\tilde{I}_s(u,v) = \mathbf{S}^q_{s,r}(u,v)\tilde{I}_r(u,v).$$

3.3 EXPERIMENTAL RESULTS

The two image synthesis algorithms that employ the surface reflectance quotient have been evaluated using two image sets. The first experimental group contains images of synthetic spheres rendered under different known light fields. The basis set for this object was created using a series of untextured spheres as shown in figure 3.1. The second experimental set consists of various real color images of human faces. For this set, two different basis sets were used. The basis set shown in figure 3.2(a) was created using an actual human face under varying illumination conditions. To further generalize the class, the basis set shown in figure 3.2(b) was created using a manequin head under varying illumination conditions. The faces in these images were manually aligned using the tip of the nose as a guide.

Figure 3.3 shows various images synthesized using the custom re-rendering algorithm described in section 3. Note the objects' surface textures are maintained even though the basis sets were constructed using only one object.

Image synthesis results obtained using the re-rendering by example process are shown in figures 3.4 and 3.5. The top rows in each of these figures show the original example images that were re-rendered, while the second rows present the illumination examples employed during the synthesis process. The images produced by the example-driven re-rendering process are displayed in the final row. The illumination fields for the corresponding illumination example and re-rendered images are qualitatively equivalent.

To quantify the synthesis accuracy of the re-rendering by example technique, the algorithm was applied to two images of the same object taken under different illumination conditions. Under these conditions, the re-rendered image should be equivalent to the one used to reconstruct the desired light field. Randomly selected results from these trials are shown in figure 3.6. For each pair of illumination example and synthesized images, both the mean color distance between pixels and the mean pixel error relative to the illumination ex-

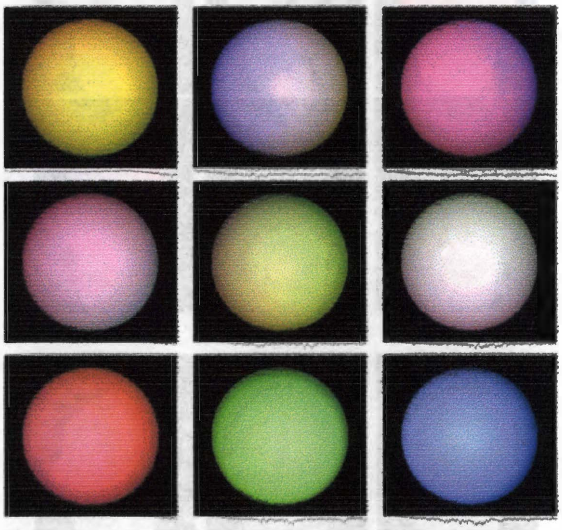
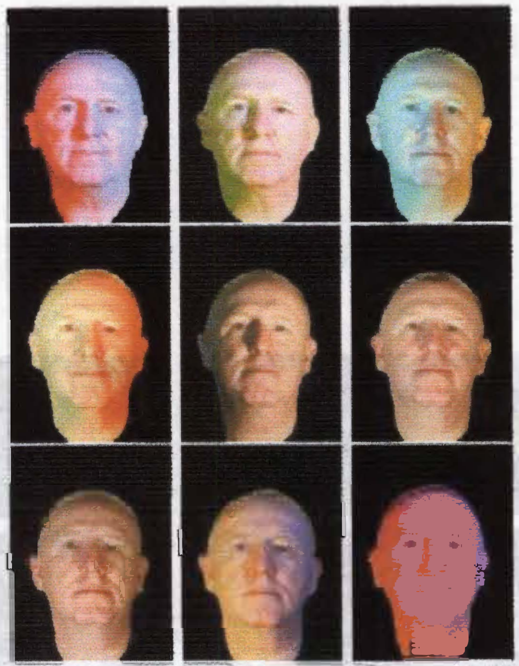


Figure 3.1: Basis image set for sphere class.


(a)


(b)

Figure 3.2: Basis image sets for human face class: (a) real faces and (b) manequin head.

Figure 3.3: Images synthesized using the custom re-rendering algorithm.

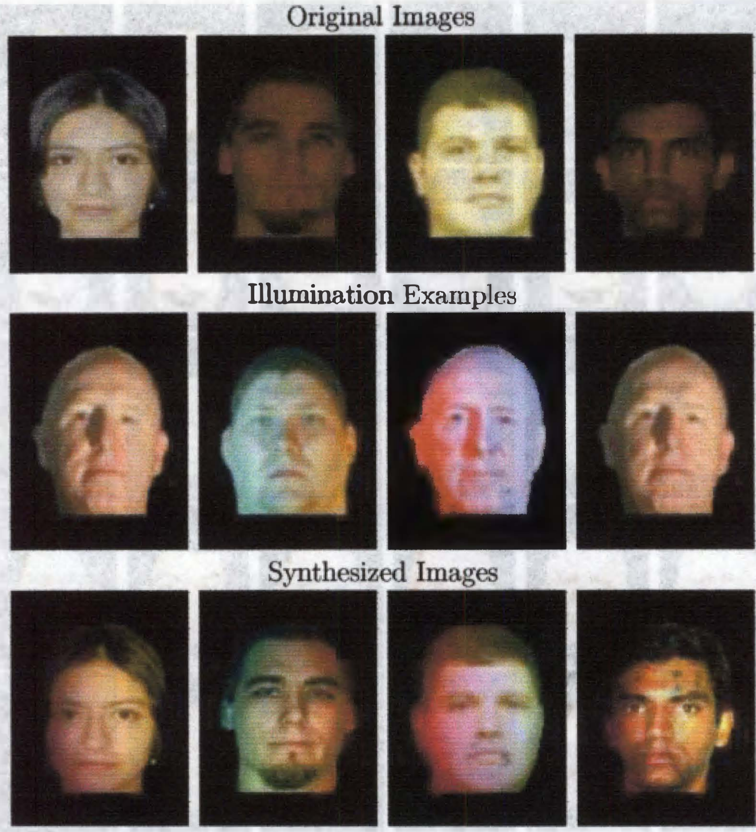


Figure 3.4: Images synthesized using the re-rendering by example algorithm.

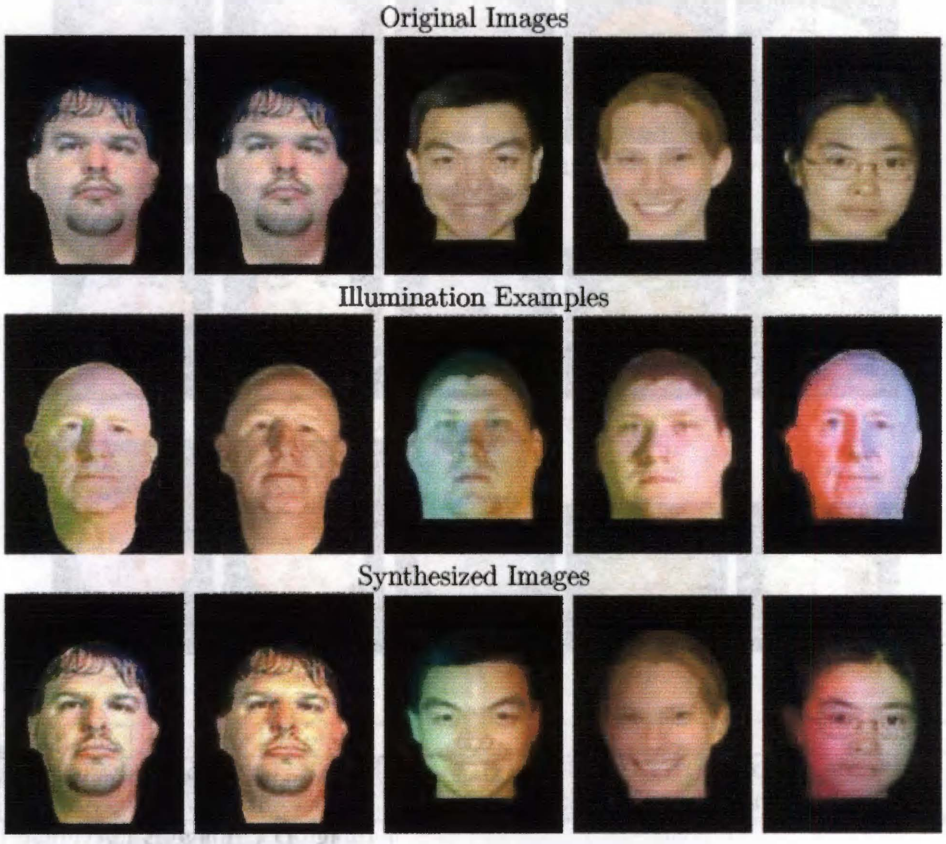


Figure 3.5: Additional images synthesized using the re-rendering by example algorithm.

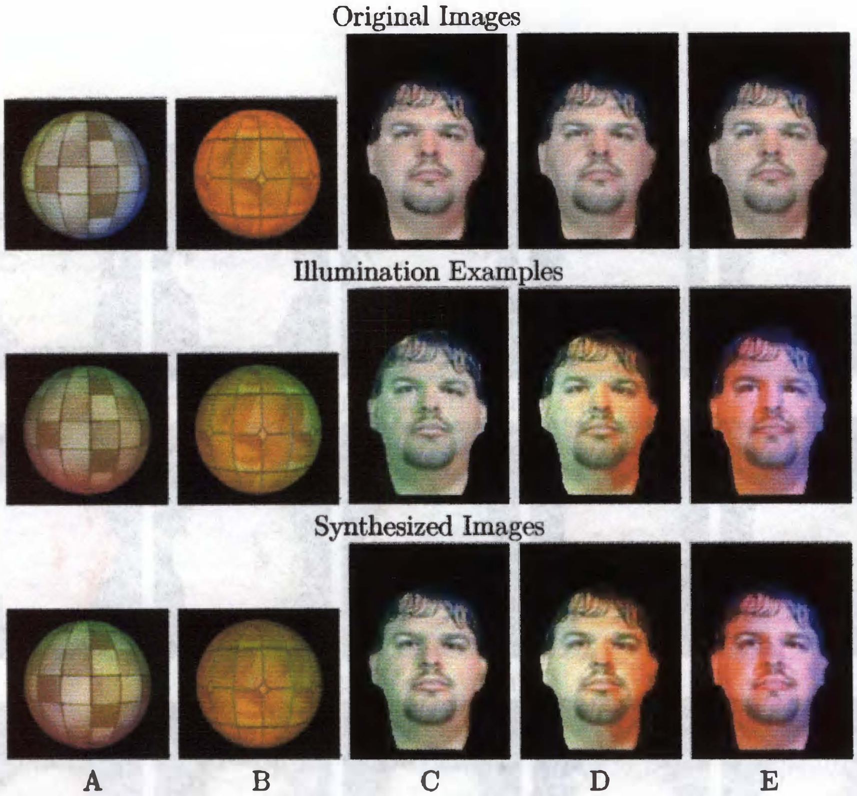


	Image A	Image B	Image C
Mean Color Error (RGB Units)	6.24	2.58	5.54
Difference from Illumination Example Mean	5.83%	2.47%	3.96%

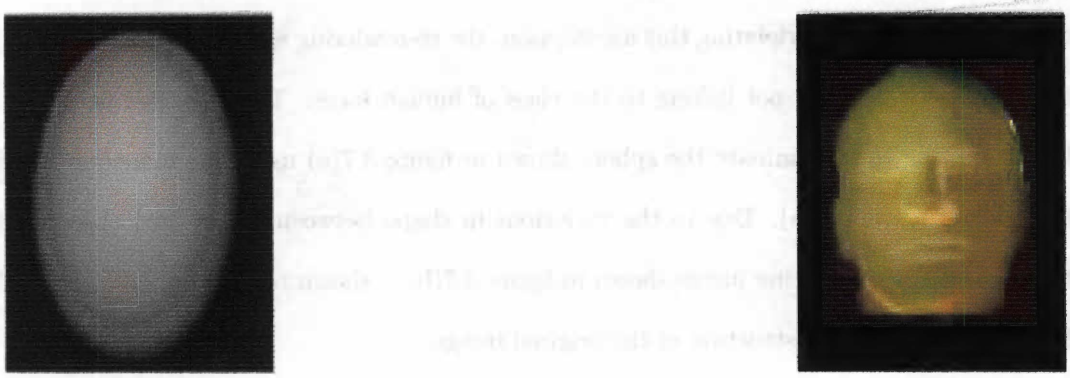
	Image D	Image E	Average
Mean Color Error (RGB Units)	4.59	4.60	4.71
Difference from Illumination Example Mean	3.29%	3.40%	3.79%

Figure 3.6: Images synthesized using the re-rendering by example algorithm with the illumination example and example images being of the same object for mean color error calculation

ample mean value were computed and shown in the table of figure 3.6. The mean color error of 4.71 units in the standard RGB color space and difference from the illumination example image mean of 3.79% across the five tests demonstrate the ability of the new algorithm to accurately synthesize new images.

The development of the surface reflectance quotient, as well as the re-rendering algorithms, was performed with the assumption that the shape of the class object is constant. To determine the effect of violating this assumption, the re-rendering algorithm was performed on an object that does not belong to the class of human faces. This test was performed by attempting to reilluminate the sphere shown in figure 3.7(a) using the manequin head basis set of figure 3.2(b). Due to the variations in shape between the imaged object and this basis set, the resulting image shown in figure 3.7(b) is shown to be an undesirable and highly inaccurate reconstruction of the original image.

The requirement of image alignment still holds for this re-rendering technique as with previous methods. To determine the degree of error caused by poor image alignment, the re-rendering algorithm was applied to a series of images with increasing degrees of misalignment in both the horizontal and vertical directions. These images may be observed in figure 3.8(a). The resulting images, shown in figure 3.8(b), demonstrate that proper image alignment is necessary to obtain proper results. With smaller misalignments, the errors are evident as parts of the image being "erased" due to the fact that these parts of the images are no longer aligned with the basis object. However, as larger degrees of misalignment are encountered, more errors begin to appear as features on the basis object begin creating artifacts in the reilluminated images. The mean color error between the properly aligned synthesized image and the misaligned synthesized images was calculated in RGB units, and the results may be observed in figure 3.9(a). These values were also plotted against the degree of pixelwise misalignment as seen in figure 3.9(b). As noted above, it may be observed on this graph that the error between the properly aligned image



(a) (b)

Figure 3.7: Sphere rerendering test: (a) original image and (b) re-rendered image.

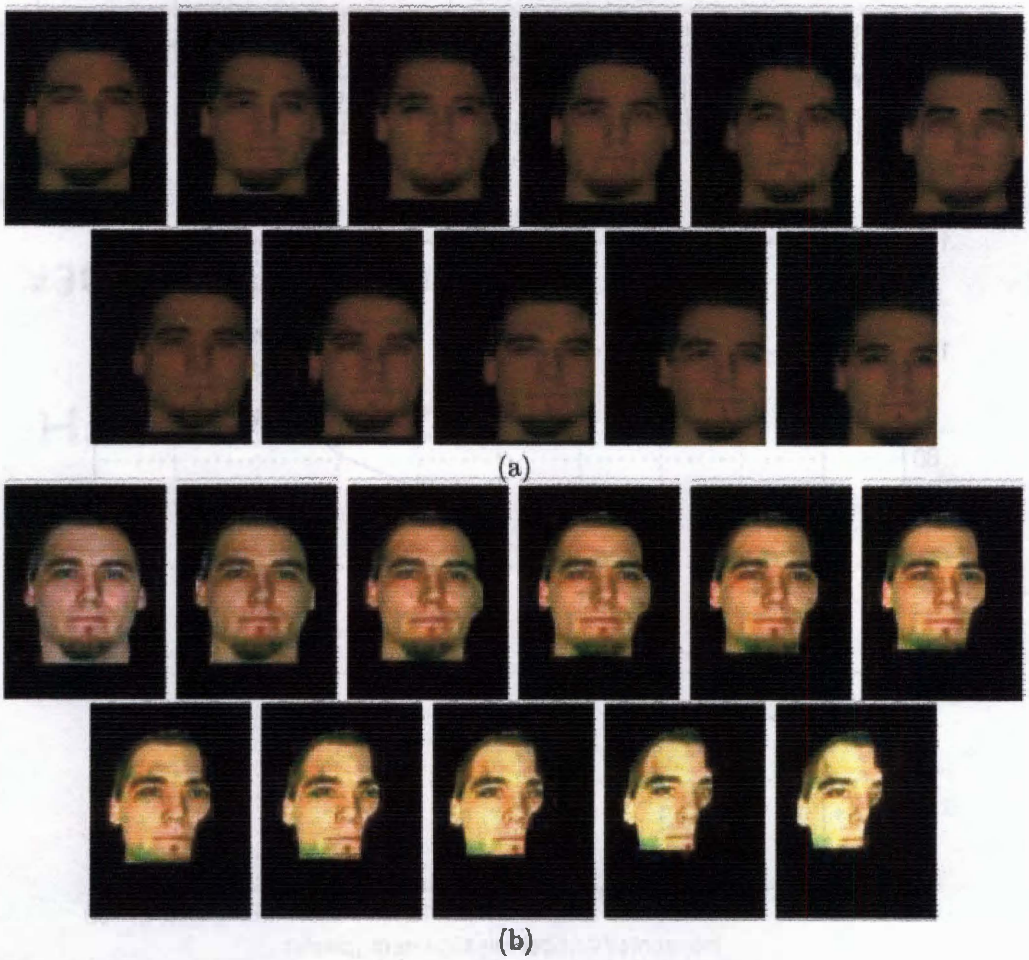
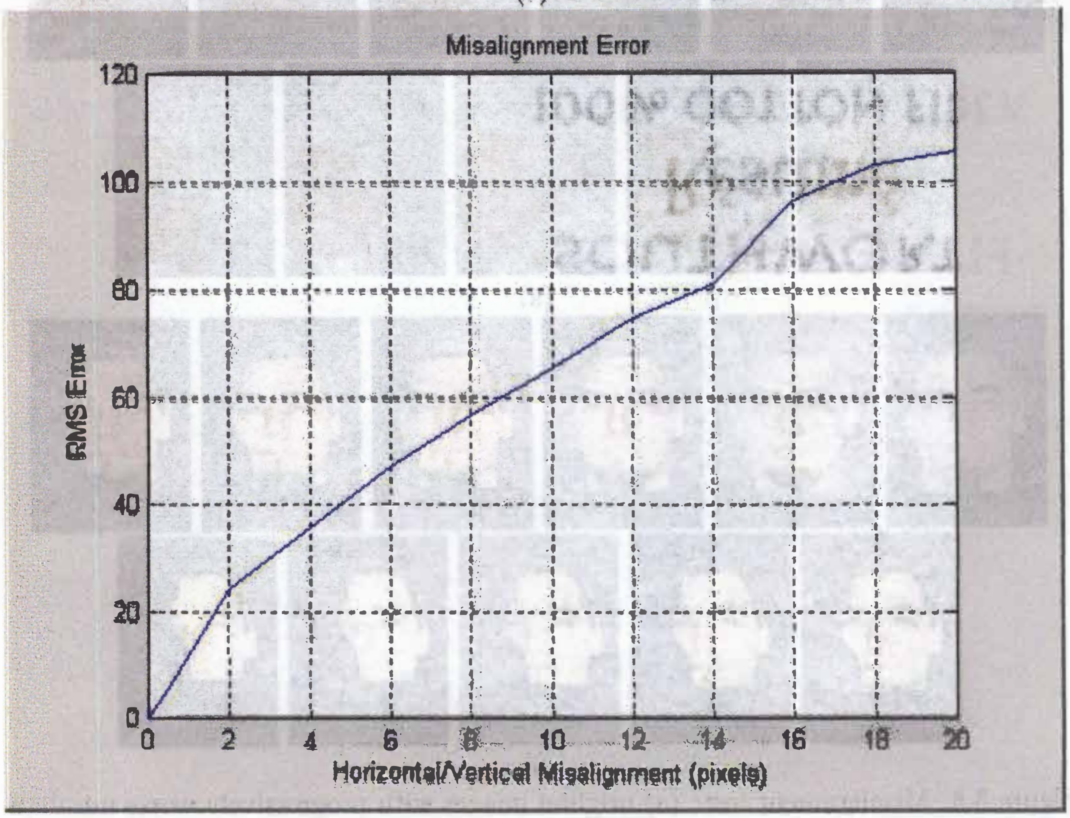


Figure 3.8: Misalignment test: (a) original images with progressively worse misalignment and (b) re-rendered images.

Degree Of Misalignment	2 Pixels	4 Pixels	6 Pixels	8 Pixels
Mean Color Error (RGB Units)	24.3668	35.8509	47.3217	56.5384
Degree Of Misalignment	10 Pixels	12 Pixels	14 Pixels	16 Pixels
Mean Color Error (RGB Units)	65.6092	74.4906	81.0399	96.2496
Degree Of Misalignment	18 Pixels	20 Pixels		
Mean Color Error (RGB Units)	103.031	105.68		

(a)



(b)

Figure 3.9: Misalignment test results: (a) mean color error and (b) plot of error versus pixelwise misalignment.

and the misaligned image increases with the degree of pixelwise misalignment.

In addition to the effect of image misalignment, it is also desirable to examine the way in which a synthesized image may be affected by reducing the size of the basis image set below the minimum number of images required to span the space of all possible illuminations. To show this effect, an image was re-rendered with progressively smaller image databases from 9 images to only a single image. The results of this test may be seen in figure 3.10. The image in figure 3.10(a) is the original image used for this test, and the image shown in figure 3.10(b) is the re-rendered image obtained using a basis set of 9 images. The results of reducing the size of the basis image set may be seen in figure 3.10(c). The number of basis images used to synthesize each image in this set has been reduced from 8 images in the case of the first image to only 1 basis image in the case of the final image in this set. As expected, the synthesized image began deviating farther from the desired result as the number of basis images was reduced. In the case of only 1 basis image, it may also be noted that the resulting image is simply a scaled version of the original image shown in figure 3.10(a). This is shown in figure 3.11. The histogram shown in figure 3.11(a) is that of the original image shown in figure 3.11(a), and the histogram shown in figure 3.11(b) is that of the synthesized image obtained when only a single basis set image is used. The histogram shown in figure 3.11(b) may be seen to be a stretched version of the original histogram shown in figure 3.11(a) showing that the result of the image synthesis process is a contrast stretching operation performed on the original image.

To quantify the error produced by a basis set of insufficient size, the mean color error was calculated between the image shown in figure 3.10(b) and each image shown in figure 3.10(c) in RGB units. These results may be seen in the table in figure 3.12(a). Additionally, these error values have been plotted versus the number of basis images used to synthesize the image. This plot may be seen in figure 3.12(b).

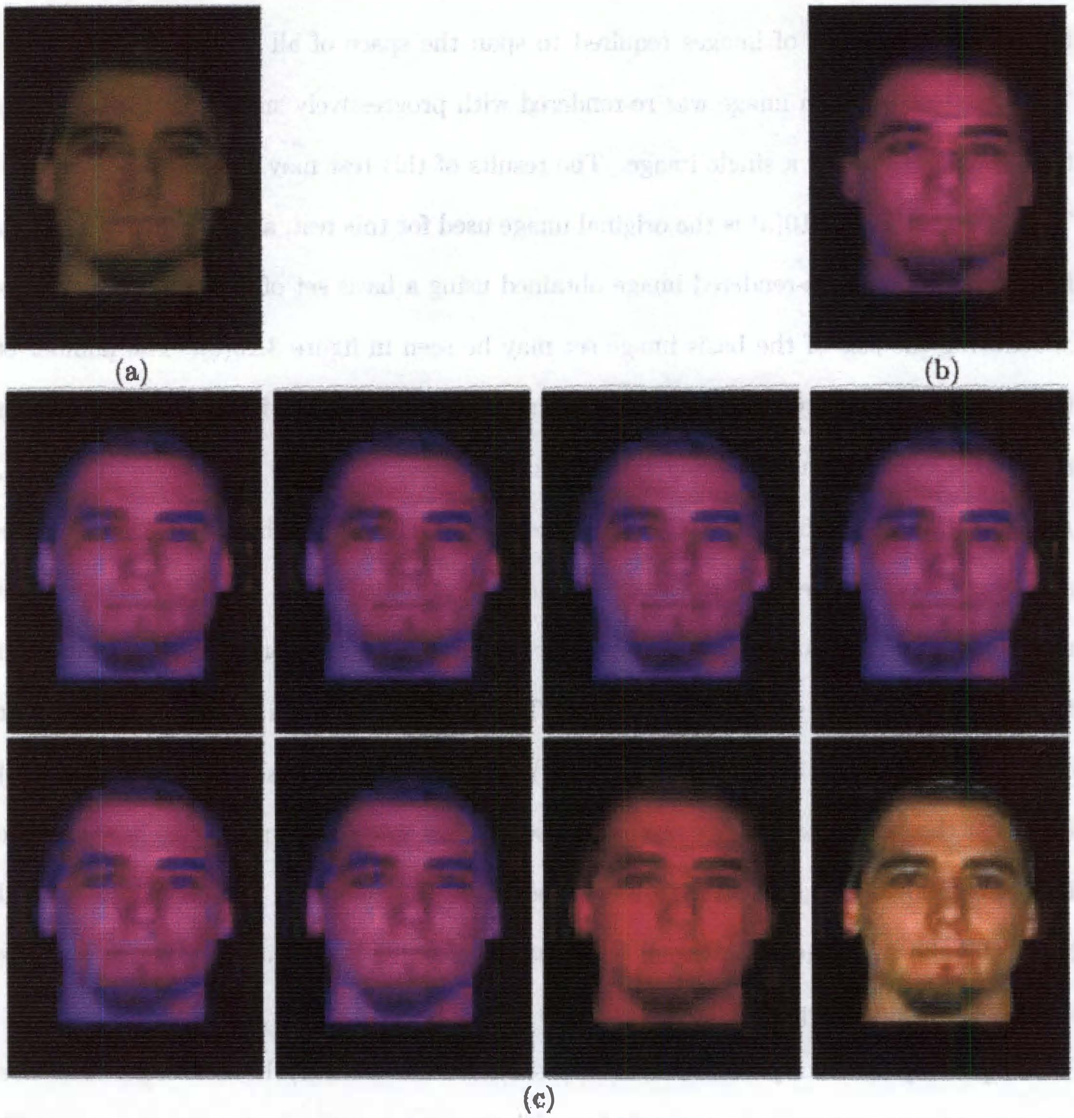


Figure 3.10: Basis set reduction test: (a) Original image from basis set reduction tests, (b) synthesized image obtained when using 9 basis set images, (c) synthesized images obtained by using progressively smaller basis sets.

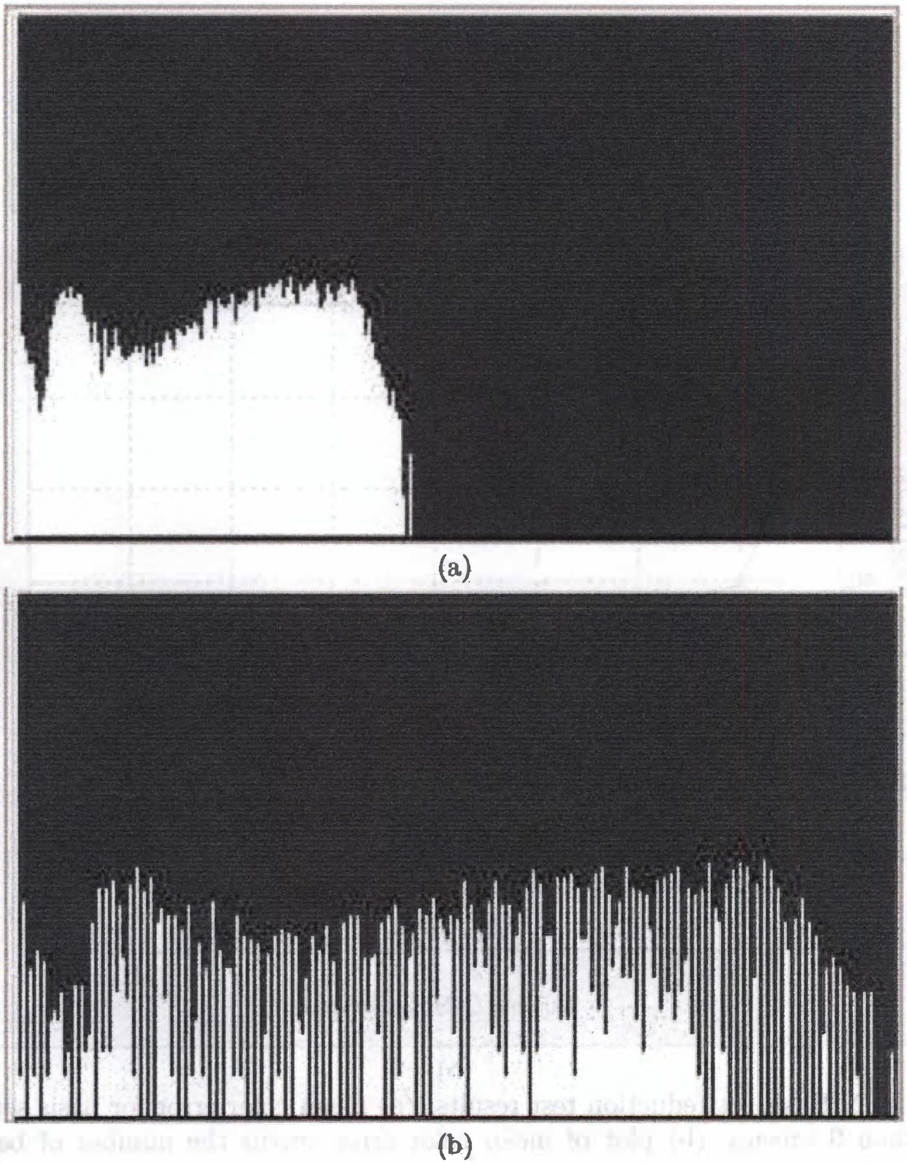
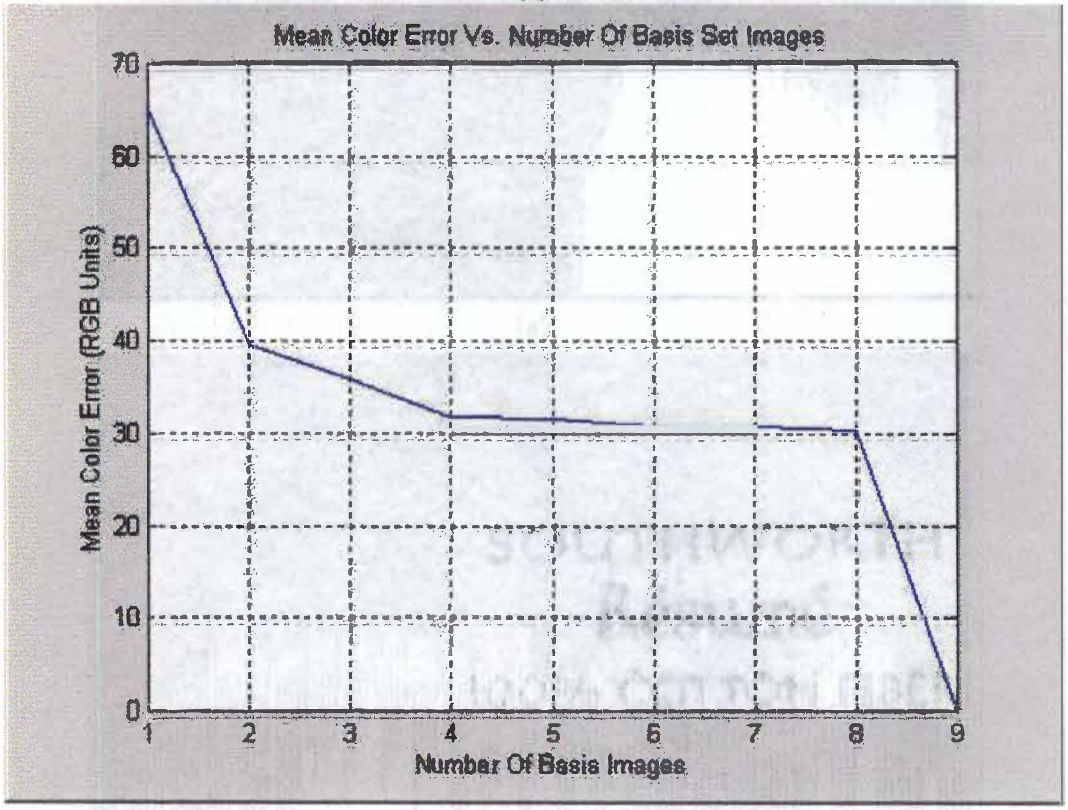


Figure 3.11: Basis set reduction histogram comparison: (a) Histogram of the original image used in the basis set reduction test, (b) histogram of the synthesized image obtained with a basis set of a single object.

Number Of Basis Images	8	7	6	5
Mean Color Error (RGB Units)	30.2908	30.7184	30.8712	31.6063
Number Of Basis Images	4	3	2	1
Mean Color Error (RGB Units)	31.7832	35.8957	39.5609	64.9861

(a)



(b)

Figure 3.12: Basis set reduction test results: (a) mean color error for basis sets with fewer than 9 images, (b) plot of mean color error versus the number of basis set images.

CHAPTER 4

OBJECT RECOGNITION ALGORITHM AND EXPERIMENTAL RESULTS

The mathematical framework discussed in Chapter 2 for the development of the surface reflectance quotient provides the basis for a color object recognition algorithm. The benefit of this recognition algorithm is due to the fact that the surface reflectance quotient allows the task to be performed without the errors associated with variations in illumination.

The process of object recognition is performed by comparing testing object, $\vec{I}_t(u, v)$, with a database of known objects, $\vec{I}_j(u, v)$ for $j = 1 \dots M$ where M is the number of objects in the database. The illumination invariant property of the surface reflectance quotient provides a means of performing this task while avoiding the errors associated with changing illumination conditions. Used as a signature image, the surface reflectance quotient may be used to test the example image may be against this database of known objects for classification.

The surface reflectance quotient of the image of the testing object, $\mathbf{S}^q_{t,r}(u, v)$, as well as those for the image database, $\mathbf{S}^q_{j,r}(u, v)$, may be calculated using equation 2.11. In ideal circumstances, the testing object would be recognized as the object imaged in the database with an identical surface reflectance quotient to that of the testing object. Therefore, the testing image may be taken to be of the same class as the database image that satisfies

$$\mathbf{S}^q_{t,r}(u, v) - \mathbf{S}^q_{j,r}(u, v) = \mathbf{0} \tag{4.1}$$

where $\mathbf{0}$ is the zero matrix. However, due to slight misalignments, the requirement of equation 4.1 can not be met. To accomodate for alignment issues, the root mean squared, or RMS, error between the intensity of the testing image and that of each database image is calculated. The testing image is then taken to be of the same class as the database image providing the minimum RMS error.

The RMS errors between the image of the testing object and those in the image database are calculated using the following technique. The mean squared error between the surface reflectance quotients may be calculated as

$$\mathbf{E}_j = \sum_{u=1,v=1}^{m,n} \frac{\left(\mathbf{S}^q_{t,r}(u,v) - \mathbf{S}^q_{j,r}(u,v)\right)^2}{mn} \tag{4.2}$$

where m is the number of rows in the images and n is the number of columns in the images. The diagonals of $\mathbf{E}_j$ may be vectorized to form

$$\vec{E}^{mse}_j = \mathrm{diag}(\mathbf{E}_j) \tag{4.3}$$

where $\vec{E}^{mse}_j = (r^{mse}_j, g^{mse}_j, b^{mse}_j)^\top$ is the column vector representing the mean squared error in the red, green, and blue channels. The RMS intensity error may then be calculated as

$$e^{rms}_j = \sqrt{\frac{r^{mse}_j + g^{mse}_j + b^{mse}_j}{3}}. \tag{4.4}$$

The following object recognition algorithm employs this method to determine the minimum RMS error. The database object corresponding to the minimum RMS error is then taken to be a representative of the testing object's correct classification.

4.1 OBJECT RECOGNITION ALGORITHM

To perform the task of object recognition, the user must provide an input image, $\vec{I}_t(u,v)$, of the given class object under test. This test image is to be compared to a database of images of known class objects, $\vec{I}_j$ for $j = 1 \ldots M$ and M is the number of images in the database. This is done by first determining the surface reflectance quotients of each of these images. The surface texture of the test object may then be compared to the textures of each database object by calculating the root mean squared, or RMS, error between the surface reflectance quotients of the test image and each database image. The test object is then taken to be the object represented in the database image corresponding to the minimum RMS error.

The object recognition algorithm is outlined below.

Object Recognition Algorithm

1) Calculate the first basis object image based on the input image

$$\vec{I}_r(u,v) = \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v)$$

such that the constraint

$$\min_{\alpha_i} \left\| \vec{I}_t(u,v) - \sum_{i=1}^{N} \alpha_i \vec{I}_i(u,v) \right\|^2$$

is satisfied.

2) Calculate the basis object images based on the database images

$$\vec{I}_r^j(u,v) = \sum_{i=1}^{N} \beta_i^j \vec{I}_i(u,v)$$

such that the constraint

$$\min_{\beta_i^j} \left\| \vec{I}_j(u,v) - \sum_{i=1}^{N} \beta_i^j \vec{I}_i(u,v) \right\|^2$$

is satisfied for j = 1 . . . M, where M is the number of objects in the database.

3) Compute the surface reflectance quotient of the input image

$$\mathrm{diag}(\mathbf{S}_{i,r}^q(u,v)) = \vec{I}_t(u,v)/\vec{I}_r(u,v).$$

4) Compute the surface reflectance quotients of the database images

$$\mathrm{diag}(\mathbf{S}_{j,r}^q(u,v)) = \vec{I}_j(u,v)/\vec{I}_{r,j}(u,v)$$

for j = 1 . . . M.

5) Calculate the mean squared error vectors

$$\vec{E}_j^{mse} = \text{diag}(\mathbf{E}_j)$$

for j = 1 ...M, where $\mathbf{E}_j$ is the mean squared error matrix defined in equation 4.2.

6) Calculate the root mean squared error

$$e_j^{rms} = \sqrt{\frac{(r_j^{mse} + g_j^{mse} + b_j^{mse})}{3}}$$

for j = 1 ...M, where r_j^{mse}, g_j^{mse}, b_j^{mse} are the red, green, and blue channel components of $\vec{E}_j^{mse}$, respectively.

7) Find j_{min} corresponding to $\min(e_{rms,j})$ for j = 1 ...M.

8) Assign image $\vec{I}_s(u,v)$ to class j_{min}.

4.2 EXPERIMENTAL RESULTS

The object recognition algorithm was evaluated using the image sets shown in figures 4.1 and 4.2. Each of these image sets consists of human faces of varied race and sex. The image set shown in figure 4.1 contains a database of images of known classification. The testing set shown in figure 4.2 is made up of unclassified images of faces under varying illumination conditions. Each of the test images corresponds to a face in the image database. Ideally, the error between the input image and the matching database image should be 0. However, due to variations in facial expression and slight misalignments, this is not the case.

This algorithm was first tested by using the database example images as test images. In each of these cases, the error between the input image and only a single database image was calculated to be 0. This test verified the correct operation of the object recognition process.

Once proper operation was determined, the object recognition algorithm was tested using the testing image set. This set contains 66 images of human faces of varying age,

Figure 4.1: Image database for object recognition.

Figure 4.2: Testing image set for object recognition.

race, and sex under arbitrary illumination conditions. The object recognition algorithm was applied to each of the testing images in order to compare these images to those of the database image set for classification. The results of this test showed 60 correct classifications with only 6 errors. Using these results, the accuracy rate of the object recognition algorithm was found to be 90.9%.

Although this object recognition algorithm has been shown to have a hight rate of accuracy, this does not show that this technique has provided any degree of improvement over direct RMS comparisons of the images without illumination compensation. In order to demonstrate this, the outlined object recognition algorithm was performed without the calculation of the surface reflectance quotients. In this test, the RMS error was determined between the input image and each image of the database. This algorithm was tested using the same testing image and database image sets. This resulted in only 44 correct classifications with 22 misclassified images. Using these restults, the accuracy rate of this recognition algorithm is calculated to be only 66.7%. This shows that the object recognition algorithm has provided a 24.2% increase in accuracy over the same algorithm performed without the advantage of illumination invariance.

Errors in this object recognition technique may occur due to misalignments, varying facial expressions, or close similarities between facial features. The images shown in figure 4.3 were misclassified in the test discussed above. The misclassification of image (a) is most likely due to the very close similarity between the test image and the database image for the erroneous classification. Although this image may be recognized as belonging to the class shown in the proper classification, it can also be noted that this image looks quite similar to the erroneous class. The misclassifications in images (b), (e), and (f) appear to be caused by large misalignment between the testing images and the image database. These misalignments cause large offsets in the calculated RMS error values. Due to these offsets, the error between the testing images and the corresponding database images is no

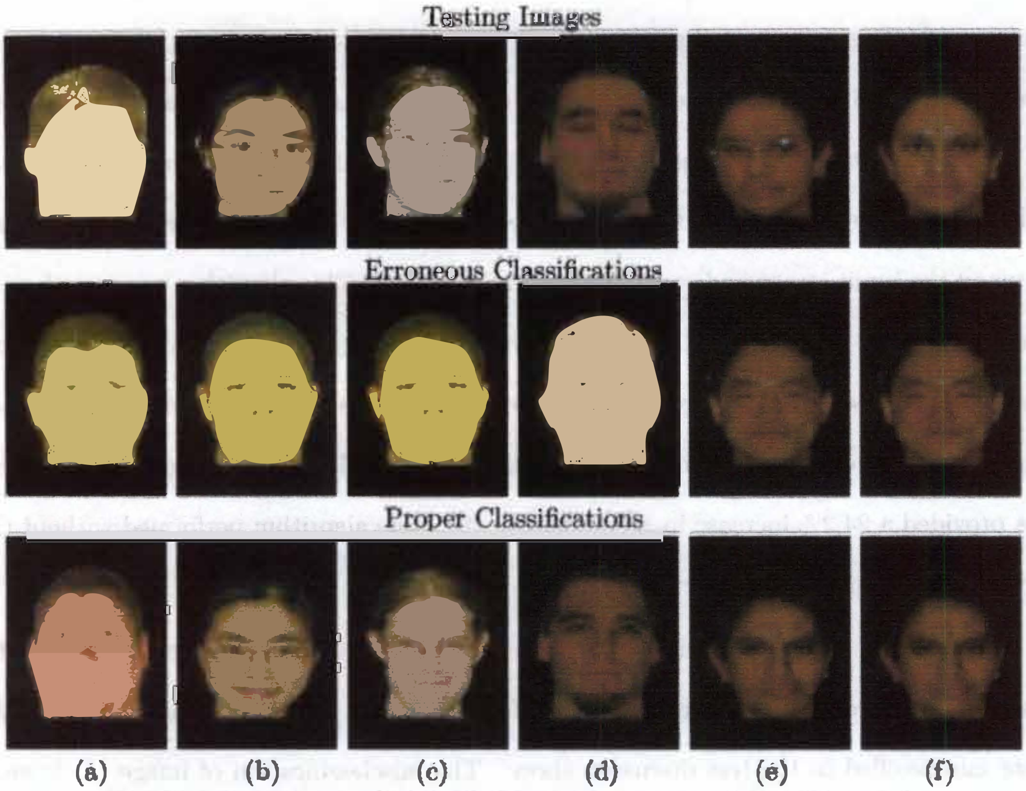


Figure 4.3: Images misclassified by object recognition algorithm.

longer guaranteed to be the minimum calculated error. Image (c) also displays a degree of misalignment due to a turned head, but it also shares the feature of a variation in facial expression with image (d). These variations in facial expression cause misclassifications in the same manner as misalignment.

CHAPTER 5

SUMMARY

Through the construction and use of the surface reflectance quotient, the class-based color image synthesis procedures described in this thesis provide the ability to create new color images of class member objects as they would appear under spatially and spectrally varying illumination conditions. Unlike the common generative approaches to class-based image re-rendering, high quality images can be created using these new techniques given a simple basis set of at least nine images constructed from a single class object, and minimal inter-image object alignment.

Two algorithms for image re-rendering using the surface reflectance quotient have been developed and demonstrated. In the first, or custom re-rendering, algorithm, the desired lighting condition is specified by the user in order to create a synthetic image under a new, custom illumination condition. In the second, or re-rendering by example, process, the original image is re-rendered such that the new image's light field matches that of an illumination example image.

The construction of the surface reflectance quotient hass also provided the basis for developing an illumination invariant technique for color object recognition. In performing this task, the surface reflectance quotient is used as an illumination invariant signature image for each object under test. By eliminating the effects of variations in lighting, the comparison process may be performed based on the surface texture of the test object and those of the database objects. An algorithm for this recognition technique has been developed and demonstrated with highly accurate results.

There are a number of areas in which the work discussed in this thesis may be extended in the future. The proposed re-rendering technique does not allow for proper rendering of specular reflections in synthesized images. Therefore, a linear model for specular reflection would lead to the ability to create more accurate re-rendered images. Also, the ability to add cast shadows into reilluminated images would increase the realism of this re-rendering technique. This could be accomplished through the use of shape recovery methods or possibly through a linear combination of images. Shape recovery techniques would also lead to a method of automating the choice of the proper basis set to be used for re-rendering rather than leaving this choice to the user. The object recognition algorithm may also be compared to other reported techniques.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] P. Hallinan,
"A low-dimensional representation of human faces for arbitrary lighting conditions," in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 995–999, June 1994.

[2] J. Nimeroff, E. Simoncelli, and J. Dorsey,
"Efficient re-rendering of naturally illuminated environments," in *Proceedings of the Fifth Eurographics Symposium on Rerendering*, June 1994.

[3] D. Rowland and D. Perrett,
"Manipulating facial appearance through shape and color," *IEEE Computer Graphics and Applications*, pp. 70 – 76, September 1995.

[4] T. Vetter and T. Poggio,
"Linear object classes and image synthesis from a single example image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 733–742, 1997.

[5] T. Vetter and T. Poggio,
"Image synthesis from a single example image," in *Proceedings of the European Conference on Computer Vision*, 1996.

[6] T. Vetter, M.J. Jones, and T. Poggio,
"A bootstrapping algorithm for learning linear models of object classes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 40–46, 1997.

[7] C. Schoeneman, J. Dorsey, B. Smits, J. Arvo, and D. Greenberg,
"Painting with light," in *Computer Graphics Proceedings, Annual Conference Series*, pp. 143–146, 1993.

[8] A. Shashua and T. Riklin-Raviv,
"The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 129–139, 2001.

[9] T. Riklin-Raviv and A. Shashua,
"The quotient image: Class-based re-rendering and recognition with varying illuminations," in *Computer Science TR99-1*. Hebrew University, 1999.

[10] T. Riklin-Raviv and A. Shashua,
"The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 1999.

[11] A. Stoschek,
"Image-based re-rendering of faces for continuous pose and illumination directions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 582–587, June 2000.

[12] K. Torrance and E. Sparrow,
"Theory for off-specular reflection from roughed surfaces for ray reflection,"
Journal of the Optical Society of America (A), vol. 65, no. 1, pp. 531–536, 1967.

[13] P. Meer and Y. Leedan,
"Estimation with bilinear constraints in computer vision,"
in *Proceedings of the International Conference on Computer Vision*, pp. 733–738, 1998.

[14] W. Freeman and J. Tenenbaum,
"Learning bilinear models for two-factor problems in vision,"
in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 554–560, 1997.

VITA

Andrew Wilson was born on April 20, 1977 and was raised in Elizabethton, TN. He attended grade school at Hunter Elementay School and middle school at Hampton Elementray School in Carter County, TN. He then attended Elizabethton High School in Elizabethton, TN where he graduated in 1995. Upon graduating high school, he attended the University of Tennessee, Knoxville where he received a B.S. in electrical engineering in 2000.

Andrew has had relevant work experience in the field of electrical engineering both as a cooperative engineering student and a graduate research associate. During his time in the engineering co-op program, he spent two semesters working for Philips Consumer Electronics in Knoxville, TN as well as one semester working for Concorde Microsystems also located in Knoxville, TN. Since receiving his B.S. degree, he has worked with the Applied Visualization Research Center at Oak Ridge National Laboratory and, most recently, with the Applied Visualization Center at the University of Tennessee, Knoxville where he is still currently employed.

Andrew is currently pursuing the M.S. degree in electrical engineering degree with an emphasis in Digital Image Processing at the University of Tennessee, Knoxville.