5-2001

Color image-based shape reconstruction of multi-color objects under general illumination conditions

Ambrose E. Ejiofor Ononye

To the Graduate Council:

I am submitting herewith a dissertation written by Ambrose E. Ejiofor Ononye entitled "Color image-based shape reconstruction of multi-color objects under general illumination conditions." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Philip W. Smith, Major Professor

We have read this dissertation and recommend its acceptance:

Donald W. Bouldin

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Ambrose H. Ejiofor Ononye entitled "**Color Image-Based Shape Reconstruction of Multi-Color Objects Under General Illumination Conditions.**" I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Philip W. Smith, Major Professor

We have read this dissertation
and recommend its acceptance:

Donald W. Bouldin

Hairong Qi

John Chrastil

Jens Gregor

Accepted for the Council:

Vice Provost and Dean of Graduate Studies

COLOR IMAGE-BASED SHAPE RECONSTRUCTION OF MULTI-COLOR OBJECTS UNDER GENERAL ILLUMINATION CONDITIONS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Ambrose H. Ejiofor Ononye

December 2001

DEDICATION

To the Memory of my Beloved Grandmother

Mrs. Elizabeth Obura Odiari

ACKNOWLEDGEMENTS

A complete acknowledgement of my indebtedness for this dissertation is impractical. My acknowledgement first goes to God for all His blessings. To Him be the Glory. I am very much indebted to my advisor, Dr. Philip Smith for his advice, support and encouragement. I would also like to thank the rest of the members of my doctoral committee – Dr. Donald Bouldin, Dr. John Chiasson, Dr. Jens Gregor and Dr. Hairong Qi for their valuable suggestions and insights regarding the ideas presented in the dissertation and above all for accepting the responsibility to be on the committee.

I am particularly grateful to my colleagues Joao Pinto, Jason Rudisill, Stephen Jesse, Mengwei Li, Kaiyn Wang and Zhong Du who allowed me to use the picture of their faces and many others too numerous to be mentioned individually who contributed in one way or another.

I would like to thank my wife, Lawretta for her love, patience and support. I share all my achievements with her. At this juncture, I would like to thank my two sons Ifechukwu and Nnaebuka for their moral support and understanding while I was away from them to the laboratory most of the time.

Finally, I wish to thank my brother, Dr. Ikechukwu Ononye for his support and advice and also Dr. Roberto Benson for his advice.

ABSTRACT

Humans have the ability to infer the surface reflectance properties and three-dimensional shape of objects from two-dimensional photographs under simple and complex illumination fields. Unfortunately, the reported algorithms in the area of shape reconstruction require a number of simplifying assumptions that result in poor performance in uncontrolled imaging environments. Of all these simplifications, the assumptions of non-constant surface reflectance, globally consistent illumination, and multiple surface views are the most likely to be contradicted in typical environments. In this dissertation, three automatic algorithms for the recovery of surface shape given non-constant reflectance using a single-color image acquired are presented. In addition, a novel method for the identification and removal of shadows from simple scenes is discussed.

In existing shape reconstruction algorithms for surfaces of constant reflectance, constraints based on the assumed smoothness of the objects are not explicitly used. Through explicit incorporation of surface smoothness properties, the algorithms presented in this work are able to overcome the limitations of the previously reported algorithms and accurately estimate shape in the presence of varying reflectance. The three techniques developed for recovering the shape of multi-color surfaces differ in the method through which they exploit the surface smoothness property. They are summarized below:

- **Surface Recovery using Pre-Segmentation** - this algorithm pre-segments the image into distinct color regions and employs smoothness constraints at the color-change boundaries to constrain and recover surface shape. This technique is computationally efficient and works well for images with distinct color regions, but does not perform well in the presence of high-frequency color textures that are difficult to segment.

- **Surface Recovery via Normal Propagation** - this approach utilizes local gradient information to propagate a smooth surface solution from points of known orientation. While solution propagation eliminates the need for color-based image segmentation, the quality of the recovered surface can be degraded by high degrees of image noise due to reliance on local information.

- **Surface Recovery by Global Variational Optimization** - this algorithm utilizes a normal gradient smoothness constraint in a non-linear optimization strategy to iteratively solve for the globally optimal object surface. Because of its global nature, this approach is much less sensitive to noise than the normal propagation is, but requires significantly more computational resources.

Results acquired through application of the above algorithms to various synthetic and real image data sets are presented for qualitative evaluation. A quantitative analysis of the algorithms is also discussed for quadratic shapes. The robustness of the three approaches to factors such as segmentation error and random image noise is also explored.

Contents

List of Tables

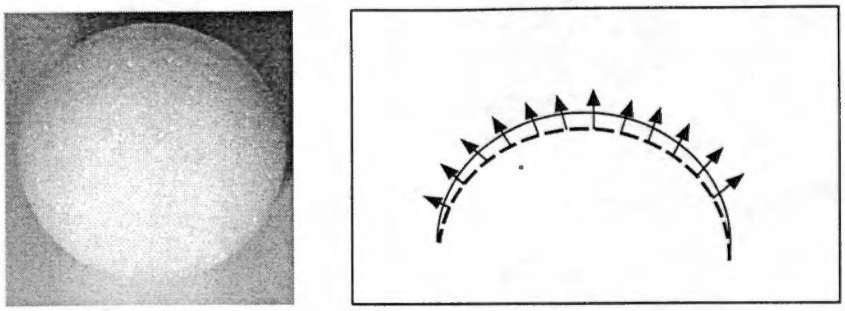
List of Figures

CHAPTER 1

Introduction

Humans have the ability to infer the surface reflectance properties and three-dimensional shape of objects from two-dimensional photographs under simple and complex illumination fields which include visual effects such as specular highlights and shadows. A central theme of research in computer vision has focused on understanding and mimicking this ability of humans to reconstruct object shape from two-dimensional visual cues. The significant interest in this field stems not only from a desire to comprehend the human visual system, but also from the vast number of practical applications in which shape recovery technology can be employed. Reconstruction of object shape can improve the performance of automated visual tasks ranging from military applications, in which object recognition and pose estimation are often critical, to artistic systems, where re-illumination and view re-rendering are often desired. Unfortunately, most of the reported algorithms in the area of shape reconstruction require a number of simplifying assumptions that result in poor performance in uncontrolled imaging environments. Thus, the focus of the research effort discussed in this dissertation is to develop automatic algorithms which are able to recover surface shape under realistic illumination and surface reflectance conditions from single photographs.

Shape estimation is formally defined as the recovery of the surface normals of localized patches along a continuous object surface as shown in figure 1.1. Due to the physics of light reflection at material surfaces, the apparent intensity of a surface patch in an image is a function of the angle between the surface normal and the direction cosines of the incident illuminants. Thus, gray-scale image intensity information can be used to estimate object shape. Techniques which employ intensity data in this manner are known



a b

Figure 1.1: *Surface Normals: (a) is the input image of a spherical styrofoam . (b) illustrates the surface normals extending from various points.*

as shape-from-shading methods. The primary difficulty with using intensity information to estimate the angle between surface normal and illumination direction is that surface orientation has three degrees of freedom while intensity information and normal conditions can only provide two constraints. The under-constrained nature of this problem results in an inherent ambiguity in the recovered surface shape due to the cone of orientations that can map to the same intensity for a given imaged point [1]. Thus, the primary challenge in the development of shape from shading algorithms lies in uncovering other sensory or heuristic data which can be used to further constrain the shape estimation process as was attempted by authors like [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Traditional shape-from-shading algorithms also require objects to have a constant surface albedo, or apparent color, severely restricting their application under realistic imaging conditions.

Horn, in his seminal work [12], employs a smoothness condition to provide the final constraint required to recover a unique shape from gray-scale image data. Surface orientation is recovered by assuming a known orientation for one surface patch and propagating the solution from that point using the relationship between intensity derivatives and surface normal components. Object shapes obtained using this technique are often very irregular because shading gradients are approximated using discrete, noisy image data. Ikeuchi *et. al.* [7] introduced a relaxation procedure for shape recovery to avoid

the problem introduced by point estimation of derivatives. Given known boundary conditions, this technique uses variational methods to iteratively solve equations arising from global smoothness constraints and the surface normal/illuminant direction relation to estimate shape. Global optimization via iteration is computationally expensive, however, and requires knowledge of surface normals at the object boundary to ensure convergence.

Multiple images of the same object can also be used to resolve the ambiguity inherent in surface orientation recovery from intensity variation. Termed photometric stereo [13], at least three registered images of the same scene are obtained from a constant viewpoint but with a light source which is moved between image acquisitions. Using these multiple images results in a system of linear equations in terms of the surface normal elements which has a unique solution if the light source positions are known relative to the viewing angle for each image. Although this technique results in a straightforward linear solution to the shape from shading problem, the use of a controlled mobile light source and multiple images make this technique impractical outside of the laboratory or other highly controlled environments. This approach provides little insight into the human ability to reconstruct such shape information from a *single* image as well.

As reported by Drew [3], the problem of acquiring multiple gray-scale images can be avoided by employing color images because RGB color cameras provide three registered, spectrally distinct images at the same instant. If an object is imaged using a color camera under non-degenerate illumination conditions, the same equations used in traditional photometric stereo algorithms can be used to estimate the surface normal of each observed object point if each channel is treated as a separate input image. Algorithms which use this approach are known as *shape-from-color* techniques. In addition, Drew [3] and others [13],[14] have also shown that it is possible to recover the both positions and color of the illuminants in the scene; and the object surface normals up to an unknown rotation *simultaneously*, eliminating the requirement that illumination conditions be known prior

to shape reconstruction. While allowing for unknown lighting conditions makes such algorithms more adaptable to realistic lighting conditions, these methods still demand the surface reflectance of the object be constant.

All of the above mentioned algorithms utilize the following assumptions to simplify and constrain the shape estimation process:

1. **Minimal Inter-reflection** - the above algorithms assume that the power of the direct light source is significantly stronger than the light reflected off of the object surface onto itself.

2. **Variant Light Source/Viewing Positions** - the above techniques assume the light source and viewing positions are significantly different..

3. **Distant Light Sources** - the methods discussed assume light sources are located far enough from the object surface so that the incident illumination is parallel across the surface.

4. **Continuous Surface Shape** - the above mentioned algorithms assume the object surface can be modelled as a spatially continuous function.

5. **Global Illumination Conditions** - the techniques described above require each surface point to be illuminated by all light sources present.

6. **Constant Surface Reflectance** - the above algorithms assume objects do not vary in color.

The first four assumptions are fairly weak and present little difficulty to the construction of systems for use outside the laboratory. For most non-mirror like surfaces, the power of inter-reflected light is negligible compared to the illuminating source. In most

typical indoor and outdoor scenes, multiple illumination sources are present that are orders of magnitude distant from the object when compared with the object size. A large subset of objects of interest for shape recovery are also smooth and continuous from a given view point. The penultimate and the last assumptions pose the primary difficulties for the construction of shape-from-shading algorithms which can be used in uncontrolled environments. Of primary concern is the assumption of *constant*, non-specular surface reflectance, since most everyday objects exhibit surface color variations. Shadows caused by both self-occlusion and other objects in the illumination field are also common in photographs of everyday scenes. To effectively move shape-from-shading from the realm of the laboratory into practice, tools must be constructed which either do not require or can significantly relax these two requirements.

1.1 The Research Goal

The primary goal of this research effort is to delineate automatic algorithms for estimation of object surface shape with *non-constant surface reflectance properties* (multi-color objects), under general, non-degenerate illumination conditions, from single-color images as shown in figure 1.2. Specifically, three algorithms have been developed for this task which explicitly employ the fundamental equations that relate surface normals to apparent image colors and smoothness conditions to recover the shape of spectrally varying objects. The first, or *pre-segmentation*, approach segments the image into distinct color regions and employs smoothness constraints at the color-change boundaries to constrain and recover surface shape. This technique is fast and works well for images with large color regions, but does not perform well in the presence of high-frequency color textures that are difficult to segment. The second, or *normal propagation*, approach utilizes a smoothness-constrained propagation method in the spirit of the previous work done by

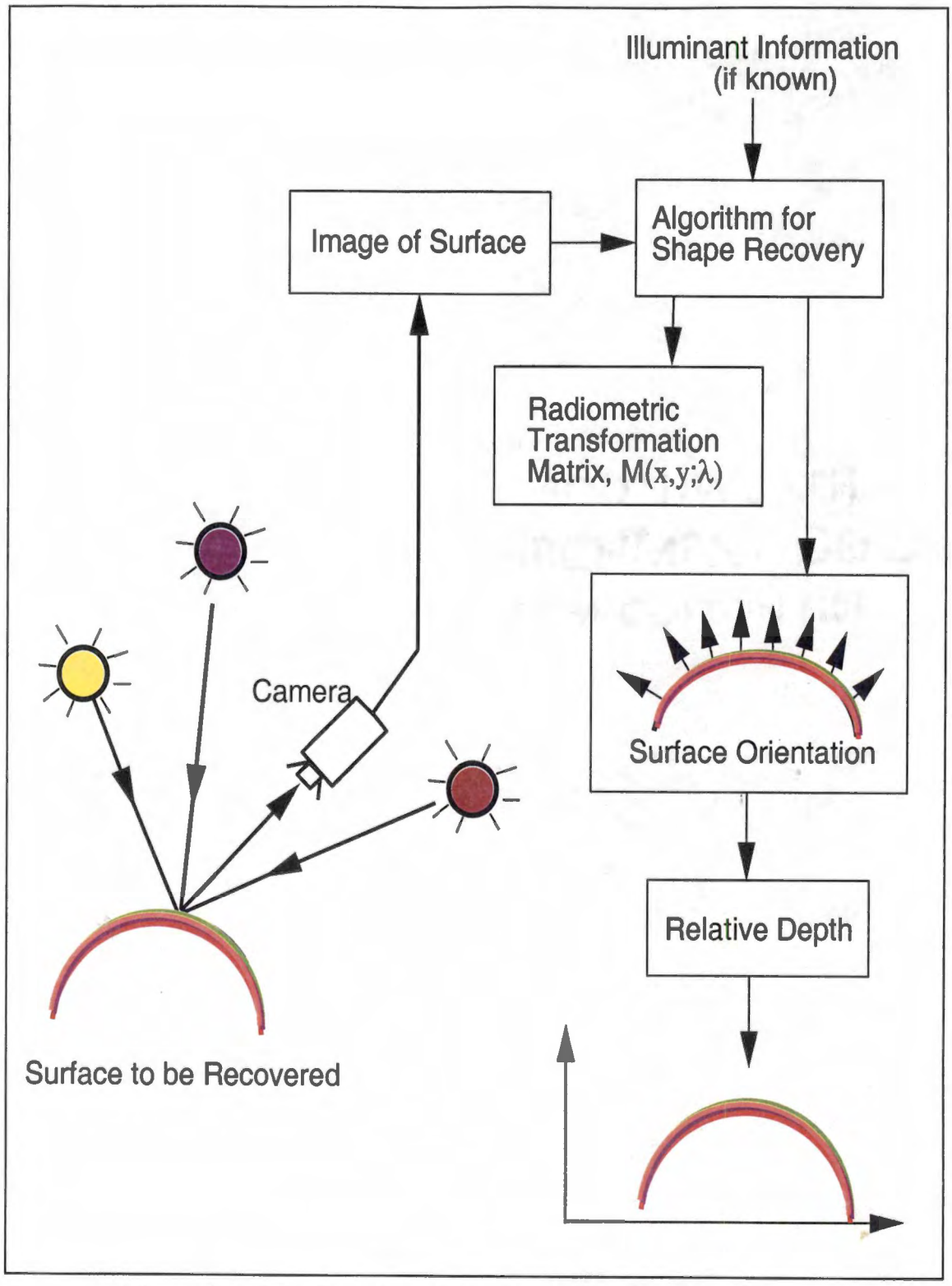


Figure 1.2: *The Overview of the shape reconstruction problem*

Horn [12] to recover surface shape without pre-segmentation. The accuracy of surfaces recovered using the propagation technique is still affected by image noise, however. The third, or *variational*, technique utilizes a global smoothness constraint to iteratively solve for the optimal object surface. While requiring no pre-segmentation and offering resistance to random image noise, this third approach is computationally expensive when compared to the other two algorithms. Thus, a potential user is given the opportunity to choose between the three different algorithms for a given set of application conditions.

Along with the development of the three algorithms for surface shape estimation given multi-color objects, various tools were also constructed. These include software for color-based image segmentation in the presence of significant shading variations, for shadow identification and extraction, and for validating shape estimation results. Critical issues regarding the design and production of these tools will also be discussed in this dissertation as appropriate.

1.2 Summary of Unique Contributions

In the course of reconstructing the shape of multi-colored objects under general illumination conditions, this dissertation makes two unique contributions to the field of automated shape reconstruction. These are:

1. The development of three algorithms for recovering the shape of objects under general illumination conditions using single color images,
    - with constant and non-constant surface reflectance, and
    - with/without cast shadows.
2. The design of a process for identifying and eliminating shadows in color imagery without using the darkest region assumption or requiring a linear camera.

1.3 Organizational Overview

Chapter 2 provides a detailed review of relevant previous research in the area of shape-from-shading. An algorithm for shape reconstruction given single-color objects is discussed in chapter 3. The detrimental effects of non-constant surface reflectance and shadowing on this single-color algorithm are also presented and examined. In chapter 4, three algorithms for estimating the shape of spectrally varying surfaces are developed and empirically examined. In chapter 5, a novel shadow detection and removal technique is discussed. Chapter 6 summarizes the work presented, distinctly lists the unique contributions of this research, and suggests possible areas for future work.

CHAPTER 2

Review of Previous Work

2.1 Introduction

As described in the previous chapter, the primary goal of this dissertation is to construct a computer-based system which estimates object shape from a single-color image under realistic illumination conditions. The construction of such a platform requires specific knowledge from the fields of surface reflectance theory, shape estimation, and illumination artifact detection. In this chapter, a detailed review of previous work in the field of shape from shading will be presented. A discussion of previous work in shadow detection and removal will be presented in chapter 5.

2.2 Surface Reflectance and Radiometric Models

Before discussing shape from shading techniques, a brief review of standard reflection and radiometric measurement models is presented, because such algorithms rely heavily on the underlying analytical structure of such models. When light from a source (or sources) impinges on a material surface, it becomes partially reflected at the interface and partially absorbed and scattered within the material. The spectral energy that is scattered, interacts with the body colorant of the material and eventually emerges from the material surface is called the body or *diffuse* reflectance component (see figure 2.1). The energy that is reflected at the interface is referred to as the interface or *specular* reflectance component. Assuming superposition is valid for light reflectance, the dual nature of the surface reflectance phenomenon can be expressed analytically as

$$I_t(x, y; \lambda) = I_d(x, y; \lambda) + I_s(x, y; \lambda) \tag{2.1}$$

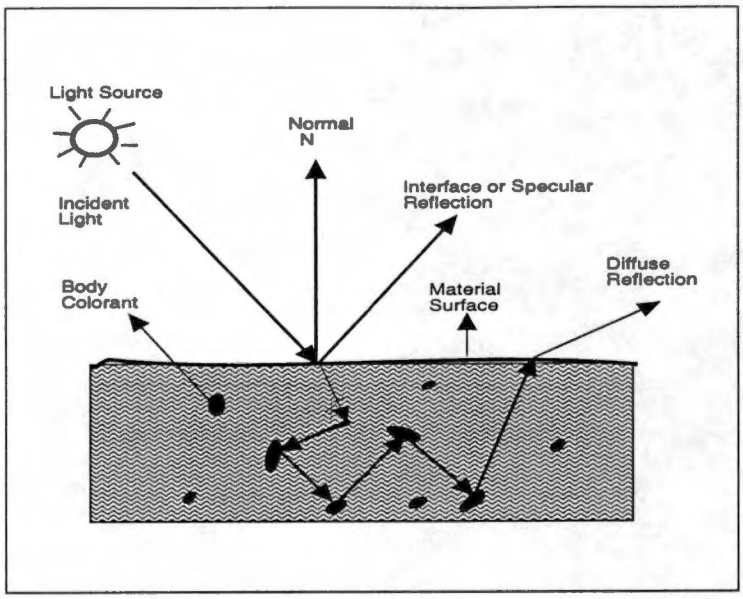


Figure 2.1: *Phenomenon of light reflection, showing specular and body reflections*

where:

I_d is the diffuse component,

I_s is the specular component, and

I_t is the total light reflected from a point on the object surface.

The diffuse component represents the spectral composition of the material and is assumed to radiate in all directions, while the specular component represents the mirror-like properties of the object and tends to reflect the illumination colorant in a distinct direction. While various parameterizations have been developed for modeling both diffuse and specular reflections [6, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27], most algorithms for estimating shape from intensity changes employ the Torrance-Sparrow model with $I_S(x, y; \lambda) = 0$ (see Appendix A). Using this reflectance model and assuming m distinct

illuminants, equation 2.1 can be rewritten as

$$I_t(x,y;\lambda) = \sum_{i=1}^{m} g_i(x,y) L_i(\lambda) S_d(x,y;\lambda) \tag{2.2}$$

where $g_i(x,y) = \vec{n}_s(x,y) \cdot \vec{n}_i = \cos(\theta_i)$ is the cosine of the angle between the surface normal at surface point (x,y) and the direction cosine of the ith illuminant, $L_i(\lambda)$ is the intensity of the ith illuminant at wavelength λ, and $S_d(x,y;\lambda)$ is the diffuse reflectance at surface point (x,y).

While equation 2.2 represents the strength of the light at a specific frequency impinging on a viewer from a specific surface point, the radiometric effects of the camera must also be considered if shape from shading techniques are to be applied to real images. The apparent intensity of a pixel at a given wavelength is often modeled using the expression

$$\begin{aligned} E_t(x,y;\lambda) &= \left(\int_{\lambda_1}^{\lambda_2} C(x,y;\lambda) \left(\sum_{i=1}^{m} \vec{n}_s(x,y) \cdot \vec{n}_i L_i(\lambda) \right) S(x,y;\lambda) d\lambda \right)^{\gamma(\lambda)} \\ &= \left(\int_{\lambda_1}^{\lambda_2} C(x,y;\lambda) \vec{L}^T(\lambda) \vec{N}(x,y) S_b(x,y;\lambda) d\lambda \right)^{\gamma(\lambda)} \end{aligned} \tag{2.3}$$

where $C(x,y;\lambda)$ and $\gamma(\lambda)$ represent the linear and non-linear gain of the detection device, respectively. Assuming a typical color camera with red, green, and blue channels; and narrow bandwidth detectors, equation 2.4 can be written in vector form as

$$\vec{E}_t(x,y;\lambda) = (\mathbf{B}(x,y;\lambda)\mathbf{N}\vec{n}_s(x,y))^{\vec{\gamma}(\lambda)} \tag{2.4}$$

where

$$\begin{aligned} \mathbf{B}(x,y;\lambda) \in \mathcal{R}^{3\times m}, \mathbf{B}(x,y;\lambda) &= \mathbf{C}(x,y;\lambda)\mathbf{S}(x,y;\lambda)\mathbf{L}(\lambda), \\ \mathbf{C}(x,y;\lambda) \in \mathcal{R}^{3\times 3}, \mathbf{C}(x,y;\lambda) &= \mathrm{diag}(C(x,y;\lambda_1), C(x,y;\lambda_2), C(x,y;\lambda_3)) \\ \mathbf{S}(x,y;\lambda) \in \mathcal{R}^{3\times 3}, \mathbf{S}(x,y;\lambda) &= \mathrm{diag}(S(x,y;\lambda_1), S(x,y;\lambda_2), S(x,y;\lambda_3)) \\ \mathbf{L}(\lambda) \in \mathcal{R}^{3\times m}, \mathbf{L}(\lambda) &= \left[\vec{L}(\lambda_1) \;\; \vec{L}(\lambda_2) \;\; \vec{L}(\lambda_3)\right]^T, \\ \mathbf{N} \in \mathcal{R}^{m\times 3}, \mathbf{N} &= [\vec{n}_0 \ldots \vec{n}_{m-1}]^T, \\ \vec{\gamma}(\lambda) \in \mathcal{R}^3, \vec{\gamma}(\lambda) &= [\gamma(\lambda_1), \gamma(\lambda_2), \gamma(\lambda_3)] \end{aligned}$$

and the vector operation, $\vec{v}^{\vec{a}}$ is defined as $\vec{v}^{\vec{a}} = [v_1^{a_1} v_2^{a_2} v_3^{a_3}]$.

As demonstrated by equation 2.4, the apparent color of an imaged pixel can be represented as a linear function of the camera gain, surface reflectance, illuminant colorants, and illuminant directions followed by a non-linear gamma correction. In the algorithms considered in the remainder of this dissertation, it is assumed that the linear camera gain matrix and the non-linear gain vector are known and their effects are removed. In this case of camera gain compensation, equation 2.4 can be rewritten as

$$\vec{E}(x, y; \lambda) = \mathbf{S}(x, y; \lambda)\mathbf{L}(\lambda)\mathbf{N}\vec{n}_s(x, y). \tag{2.5}$$

2.3 Shape Estimation from Shading

Shape from shading techniques employ gray-scale images acquired using a single point light source to recover object shape. The apparent intensity equation thus becomes

$$E(x, y) = \alpha \vec{n}^T \vec{n}_s(x, y)$$

where $\vec{n}_s^T(x, y)\vec{n}_s(x, y) = 1$, and α, known as the surface albedo, is assumed to be constant across the entire surface. Because the above expression and normality condition only provide two constraints on possible solutions to this three dimensional problem, variational constraints based on the assumption of smooth object surfaces are employed to obtain unique shape estimates. In general, there are two approaches to using the smoothness constraint to recover surface shape from intensity images. The first employs incremental propagation from surface points of known height/orientation, while the second uses relaxation techniques to recover optimal surface shape estimates using global information.

2.3.1 Incremental Propagation from Surface Points of Known Height

Originally developed by Horn [12], this method is sometimes referred to as the characteristic strip method because the algorithm essentially develops surface solutions along space

curves. This algorithm assumes that the reflectance function and the lighting parameters be perfectly known [28].

Consider the coordinates of a surface point $(x, y, z)^T$ on a space curve that falls along the object of interest, and let δx, δy be incremental steps along the surface in the x and y directions. If δz is the corresponding change in height, then the surface gradients p and q are given by

$$p = \frac{\delta z}{\delta x}$$
$$q = \frac{\delta z}{\delta y},$$

the change in height is given by

$$\delta z = p\delta x + q\delta y, \tag{2.6}$$

and changes in p and q are calculated using second derivatives of height,

$$\delta p = r\delta x + s\delta y$$
$$\delta q = s\delta x + t\delta y \tag{2.7}$$

where

$$r = \frac{\delta^2 z}{\delta x^2}$$
$$s = \frac{\delta^2 z}{\delta x \delta y}$$
$$t = \frac{\delta^2 z}{\delta y^2}.$$

In the characteristic strip method, the apparent intensity of an image pixel is rewritten in terms of the surface gradients simply as

$$E(x, y) = R(p(x, y), q(x, y)) \tag{2.8}$$

where $R(p, q)$ is called the reflectance map. Differentiating equation 2.8 gives

$$\frac{\delta E(x, y)}{\delta x} = rR_p + sR_q \tag{2.9}$$
$$\frac{\delta E(x, y)}{\delta y} = sR_p + tR_q \tag{2.10}$$

with $f_x = \frac{\delta f}{\delta x}$.

Because the direction of δx and δy may be arbitrarily chosen, δx and δy can be rewritten as

$$\delta x = R_p \xi$$

$$\delta y = R_q \xi$$

The change in the parameter ξ is along the solution curves whose surface orientation is known, and the changes in δp and δq can be expressed as functions of image intensities. Differentiating equations (2.7 and 2.10) w.r.t. ξ yields

$$\dot{x} = R_p$$

$$\dot{y} = R_q$$

$$\dot{z} = pR_p + qR_q$$

$$\dot{p} = E_x$$

$$\dot{q} = E_y$$

where $\dot{f}$ signifies derivative w.r.t. ξ. By expressing the changes of gradients δp and δq as functions of image intensity changes, a scaled depth map can be estimated.

While the above technique works quite well under ideal laboratory conditions, it suffers from several difficulties which makes it inadequate for real imaging scenarios. First and foremost, the algorithm requires that the surface reflectance of the object remain constant or else the reflectance map will vary across the image. Second, this algorithm assumes a single point light source and an image that demonstrates no surface self-occlusion. Third, derivatives estimated using discrete noisy image data are explicitly used to recover incremental changes in surface depth δz. This estimation often results in malformed shape reconstructions due to numerical uncertainties.

2.3.2 Shape Estimation by Relaxation

To overcome the difficulties associated with the explicit use of estimated derivatives during the shape reconstruction process, Ikeuchi [29] developed a variational method for reconstructing shape using gray-scale intensity variations.

Recall that the image irradiance $E(x,y)$ is related to the surface reflectance map, $R(p(x,y),q(x,y))$ by

$$E(x,y) = R(p(x,y),q(x,y))$$

To recover surface shape, the variational algorithm attempts to minimize the expression

$$A(x,y) = [E(x,y) - R(p(x,y),q(x,y))]^2 + \mu\left[(\nabla^2 p)^2 + (\nabla^2 q)^2\right]$$

over $p(x,y)$ and $q(x,y)$ where the smoothness constraint is incorporated by the Lagrange multiplier μ [30]. Minimizing $A(x,y)$ with respect to p and q using numerical approximation yields the following expressions for the surface gradients,

$$p(x,y) = p_{av} + (1/\mu)\left[E(x,y) + R(p,q)\right]\frac{\partial R}{\partial p} \tag{2.11}$$

$$q(x,y) = q_{av} + (1/\mu)\left[E(x,y) - R(p,q)\right]\frac{\partial R}{\partial q}, \tag{2.12}$$

where

$$p_{av}(x,y) = \frac{1}{4}\left[p(x+1,y) + p(x-1,y) + p(x,y+1) + p(x,y-1)\right]$$

$$q_{av}(x,y) = \frac{1}{4}\left[q(x+1,y) + q(x-1,y) + q(x,y+1) + q(x,y-1)\right].$$

Given known initial conditions for the surface orientation at the boundaries of the surface patch of interest, equations 2.11 and 2.12 can be iteratively refined to estimate the best smooth surface for the regions. For further details on the algorithm, the reader is referred to either [1] or [31].

While the global minimization approach employed by this technique significantly reduces the effect of image noise on the resultant shape, this algorithm still assumes that

the object surface is a constant color, is illuminated by a single point light source, and does not cast shadows upon itself. These assumptions severely restrict the variational technique's application outside of laboratory or other highly controlled conditions.

2.3.3 Photometric Stereo

Photometric Stereo was pioneered by Woodham [13, 14]. This technique uniquely recovers surface orientation using at least three gray-scale images with different illumination geometry. A typical photometric stereo experimental setup consists of a camera and $k \geq 3$ point illumination sources with known intensities and directions $n_1, ...n_k$ as shown in figure 2.2. The pixel intensities in each of the images acquired from a fixed direction under

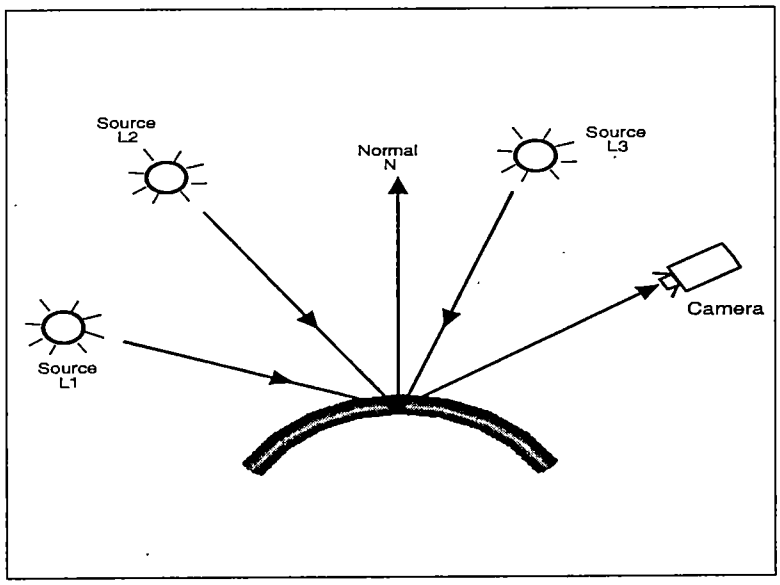


Figure 2.2: *A typical photometric stereo experimental setup with multiple light sources L1, L2 and L3*

varying illumination directions are given by

$$E_k(x,y) = \alpha \vec{n}_k^T(x,y)\vec{n}_s(x,y)$$

where $\vec{n}_k(x,y)$ is the direction cosine of the light source in image k. Combining these equations in matrix form yields

$$\vec{E}(x,y) = \alpha \mathbf{N}(x,y)\vec{n}_s(x,y)$$

where

$$\begin{aligned} \vec{E}(x,y) \in \mathcal{R}^k; \quad \vec{E}(x,y) &= [E_0(x,y) \;\; E_1(x,y) \;\; \ldots \;\; E_{k-1}(x,y)]^T \text{ and} \\ \mathbf{N}(x,y) \in \mathcal{R}^{k\times 3}; \quad \mathbf{N}(x,y) &= [\vec{n}_0(x,y) \;\; \vec{n}_1(x,y) \;\; \ldots \;\; \vec{n}_{k-1}(x,y)]^T. \end{aligned}$$

The surface orientation at each image point can be recovered by inverting $\mathbf{N}$ to obtain

$$\vec{n}_s(x,y) = \frac{(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\vec{E}(x,y)}{\|(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\vec{E}(x,y)\|}. \tag{2.13}$$

The photometric stereo technique can provide a unique computational solution to the surface estimation problem [14]. Unfortunately, this algorithm assumes that the relative position of the camera and the object remain fixed while a point light source is moved to distinct positions between the acquisition of multiple frames. This requirement makes application of this technique problematic outside of the laboratory. In addition, this technique provides little in the way of understanding the human ability to obtain the same information from a single image and still assumes a single-color surface without self-occlusion.

2.3.4 Shape from Color Photometric Stereo

Christensen and Shapiro [2] attempted to overcome some of the shortcomings of traditional photometric stereo by employing two or more color images of a given object acquired using flat white, point illumination. The camera position remains fixed while the illuminant position is varied in the two images. The primary advantage of using two *color* images is that shape can be recovered for regions which exhibit both specular and diffuse reflection.

The color photometric stereo algorithm assumes that a mapping, $\mathcal{M}_i$ exists which maps surface orientations to apparent colors in a given image, $\vec{E}_i(x,y)$, such that

$$\vec{E}_i(x,y) = \mathcal{M}_i(n_s(\vec{x},y)), \quad i = 1 \ldots k \tag{2.14}$$

where k is the number of acquired images. Given this relation, the surface normal for a given imaged point, $\vec{n}_s(x,y)$ that satisfies the apparent colors, $\vec{E}_i(x,y)$ must lie in the intersection of the set of all possible normals for that point which map to the given apparent color. Hence, the surface normal for a given imaged surface point can be recovered using the expression

$$\vec{n}_s(x,y) = \bigcap_i \mathcal{M}_i^{-1}(\vec{E}_i(x,y)), \quad i = 1 \ldots k \tag{2.15}$$

where $\mathcal{M}_i^{-1}$ is a set-valued function that maps a color to the set of normals that correspond to that color such that

$$\mathcal{M}_i^{-1}(\vec{E}_i(x,y)) = \{\vec{n}_s(x,y) \mid \mathcal{M}_i(\vec{n}_s(x,y)) = \vec{E}_i(x,y)\}$$

The color photometric stereo technique has several advantages over the previous techniques which employ gray-scale images. First, any illumination model can be employed due to the general nature of the mapping, $\mathcal{M}$. Second, the technique allows for shape recovery in surface regions which exhibit both specular and diffuse reflection again due to the generic nature of the normal to apparent color mapping. Third, it can be used to recover surface shape given multi-colored objects.

In spite of these significant advances, the color photometric stereo technique suffers from several significant difficulties when applied outside of a controlled environment. The most problematic of these drawbacks is the *a priori* construction of the set of all possible image appearances for surfaces in the image. It is often not possible to construct such sets for typical objects without the use of absolute colorimeters. In addition, the assumption of a flat white light source limits the algorithm's applicability outside of tightly controlled

conditions. It is also possible that under certain circumstances, more than two images are required to recover surface shape, or that the set of all possible normals, $M_{i,m}^{-1}$, could be empty. Finally, the method requires more than a single image to recover object shape.

2.3.5 Shape from Color

Drew *et. al.* [3] and others [32, 33, 34, 35] demonstrate that surface orientation can be recovered directly from a single image by exploiting the linear relationship between color and surface normal given in equation 2.5. If all of the surface reflectance, illuminant color, and illuminant positions are known for a given image and object, then the surface normal of a specific imaged point is given by

$$\vec{n}_s(x,y) = (\mathbf{L}(\lambda)\mathbf{N})^{-1}\mathbf{S}^{-1}(x,y;\lambda)\vec{E}(x,y;\lambda).$$

Drew [3] also demonstrated that it is possible to recover both the surface shape and the linear mapping relating normals to apparent colors up to an unknown rotation from a single image given constant surface reflectance. Drawing from this work, a number of other algorithms for estimating this unknown rotation from a single image have also been reported [33, 34].

Estimating shape using color imagery eliminates the need to acquire multiple images in order to provide a closed form solution for the recovery of surface normals. If a constant surface reflectance is assumed, it also allows for the recovery of both the surface-to-color mapping and the surface shape from a single image up to an unknown rotation. These two properties make this algorithm applicable to a broad range of images which have been acquired under unknown illumination conditions. It still requires constant surface reflection and a lack of shadows to be effective, however.

This dissertation focuses on the development of algorithms for shape estimation given single images of multi-color objects acquired under complex illuminations by building upon and significantly expanding the principles introduced by previous efforts to extract

shape from color. In the next chapter, surface reconstruction via shape-from-color will be discussed in more detail in order to demonstrate and verify the capabilities and illustrate the limitations of existing techniques when applied to surfaces with constant color, varying color and shadows.

CHAPTER 3

Shape Reconstruction for Single-Color Objects

3.1 Introduction

As discussed in chapter 2, when a scene is globally illuminated by multiple sources, the apparent color of the reflected light at an image pixel is given by

$$\vec{E}(x, y; \lambda) = \mathbf{S}(x, y; \lambda)\mathbf{L}(\lambda)\mathbf{N}\vec{n}_s(x, y) \tag{3.1}$$

Letting

$$\mathbf{M}(x, y; \lambda) = \mathbf{S}(x, y; \lambda)\mathbf{L}(\lambda)\mathbf{N}(x, y)$$

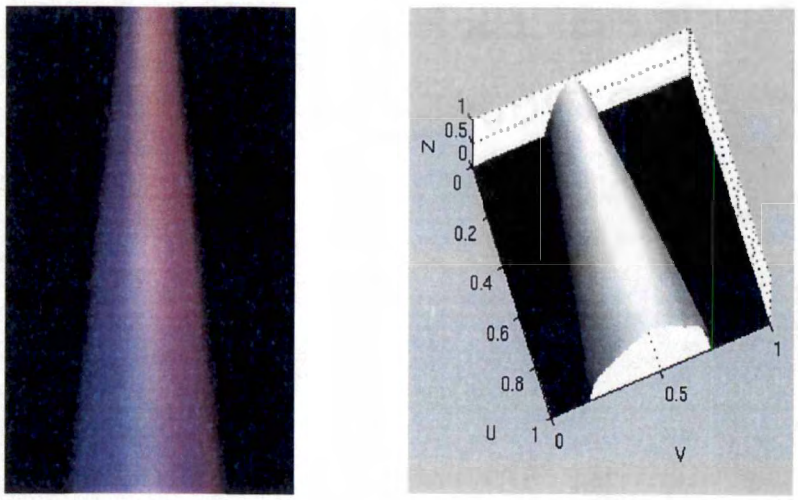
equation 3.1 can be rewritten in terms of apparent color as

$$\vec{n}_s(x, y) = \mathbf{M}(x, y; \lambda)^{-1}\vec{E}(x, y; \lambda) \tag{3.2}$$

This equation represents the fundamental shape from color mapping — a linear transformation which relates the apparent color to surface orientation. Under controlled, laboratory-like conditions when the light source positions and spectra, and the surface reflectance of the object are known, surface orientation can be recovered from apparent color via a simple matrix inversion.

3.2 Estimating the Shape-from-Color Transform

When the surface reflectance and illuminant properties are not known, as is often the case for unknown objects in outdoor scenes or uncontrolled imaging environment, Drew et al [3] demonstrated that surface orientation could be recovered up to a global rotation (as shown in figure 3.1) by exploiting the normality constraint of the surface.



a b

Figure 3.1: *Recovery of Surface Orientation: (a) is real image of a single-color cone illuminated by three lights of different spectral content. (b) is the corresponding 3D shape reconstruction up to a rotation.*

When the shape-from-color mapping, $\mathbf{M}$, is not known, both $\mathbf{M}$ and $\vec{n}_s$ can be recovered up to an orthogonal transformation, $\mathbf{R}$. Let the matrix $\mathbf{G}$ be defined such that

$$\mathbf{G} = \mathbf{M}^{-1} \tag{3.3}$$

$$\vec{n}_s = \mathbf{G}\vec{E} \tag{3.4}$$

$$\Longrightarrow 1 = \vec{n}_s^T \vec{n}_s = \vec{E}^T \mathbf{G}^T \mathbf{G} \vec{E} = \vec{E}^T \mathbf{D} \vec{E} \tag{3.5}$$

where $\mathbf{D} = \mathbf{G}^T\mathbf{G}$. For each pixel,

$$d_{11}E_1^2 + d_{22}E_2^2 + d_{33}E_3^2 + 2d_{12}E_1E_2 + 2d_{13}E_1E_2 + 2d_{23}E_2E_3 = 1 \tag{3.6}$$

where

$$\vec{E} = [E_1, E_2, E_3]^T$$

Equation(3.6) is an equation for an ellipsoid centered on the color-space origin. Given l pixels, there are l such equations which can be used to form the linear system

$$\mathbf{F}\vec{z} = 1 \qquad \text{where} \tag{3.7}$$

$$
\begin{aligned}
\mathbf{F} &= \left(E_1^2, E_2^2, E_3^2, 2E_1E_2, 2E_1E_3, 2E_2E_3\right) \\
\vec{z} &= (d_{11}, d_{22}, d_{33}, d_{12}, d_{13}, d_{23})^T
\end{aligned}
\tag{3.8}
$$

Solving the above expression for $\vec{z}$ results in an estimate for $\mathbf{D}$ that is unique up to an orthogonal transformation.

The entries of the matrix $\mathbf{G}$ can then be computed from $\mathbf{D}$ using the decomposition algorithm [36] described below:

1. Set $g_{11} = \sqrt{d_{11}}$, $g_{12} = g_{13} = 0$
2. Set $g_{21} = d_{21}/g_{11}$, $g_{22} = \pm\sqrt{d_{22} - g_{21}^2}$, $g_{23} = 0$. The $\pm$ sign is chosen so as to make the columns of $\mathbf{G}$, $(\vec{g}_1, \vec{g}_2, \vec{g}_3)$ a right-handed system.
3. Set $g_{31} = d_{31}/g_{11}$, $g_{32} = (d_{32} - g_{21}g_{31})/g_{22}$, $g_{33} = \pm\sqrt{d_{33} - g_{31}^2 - g_{32}^2}$.
4. Compute $\mathbf{M} = \mathbf{G}^{-1}$.

The recovery of $\mathbf{G}$ up to a rotation using the above algorithm implies that only an estimate of $\mathbf{G}$ is known . It also implies that only the rotated version of the surface normals are found. Let $\mathbf{G}_e$ be an estimate for $\mathbf{G}$ and $\vec{\mathcal{N}}_{rot}$ be the rotated surface normal estimates. Then

$$
\vec{\mathcal{N}}_{rot} = \mathbf{G}_e^{-1}E = \mathbf{G}_e^{-1}\mathbf{F}\vec{n}_s \equiv \mathbf{R}\vec{n}_s \tag{3.9}
$$

$$
\text{where } \mathbf{R} = \mathbf{G}_e^{-1}\mathbf{F}
$$

The estimate of the unrotated normal is

$$
\vec{n}_s = \mathbf{R}^{-1}\vec{\mathcal{N}}_{rot} = \mathbf{R}^T\vec{\mathcal{N}}_{rot} \tag{3.10}
$$

This rotation matrix $\mathbf{R}$ can be recovered by integrability condition on the normals. Further information on how this approach works, could be found from [3].

3.3 Reconstructing Relative Depth Maps from Surface Normals

Given the surface normal $\vec{n}_s$ at a point in the image, the relative depth of each point can be determined. Let (u, v) be the image space coordinates of (x, y, z) on the image projection plane. Assuming orthographic projection, the depth of each point is given by

$$z = \vec{n}_s \cdot \begin{bmatrix} u \\ v \\ k \end{bmatrix} = -f_u u - f_v v + f_z k \tag{3.11}$$

up to a scale. The constant k is a scale factor. If $\vec{n}_s = [-f_u \quad -f_v \quad f_z]^T$ then

$$z = -f_u u - f_v v + f_z k$$

Where f_u, f_v and f_z are the partial derivatives of f w.r.t. u,v and z respectively. If the surface is explicitly given by $z = f(x, y)$, then $f_z = 1$. Substituting for $\vec{n}_s$ using equation(3.2) gives:

$$z(x, y) = (\mathrm{M}^{-1}(x, y; \lambda)\vec{E}(x, y; \lambda)) \cdot \begin{bmatrix} u \\ v \\ k \end{bmatrix} \tag{3.12}$$

3.4 Outlier Detection

If a given pixel does not see all the illuminants, it is considered an outlier to the process and is removed from the surface estimation process. Often such pixels have lower average intensity values than other pixels in the image. Hence, *Least Median of Squares* LMS regression outlier detection as proposed by Rousseeuw and Leroy [37] is used for outlier rejection during the surface reconstruction procedure. The LMS algorithm yields a robust dispersion estimate given by

$$s_o = p\left(1 + \frac{5}{N-2} median(|r_i|)\right) \tag{3.13}$$

where r_i is the residue for the ith pixel, $r_i = \vec{E}_i^T \mathbf{D} \vec{E}_i - 1$ (see equation(3.5)) and $p = 1/\Phi^{-1}(0.75) \approx 1.4826$ and Φ being the assumed probability function. The color pixel is

considered valid if it satisfies the constraint

$$\left|\frac{r_i}{s_o}\right| \leq \zeta. \tag{3.14}$$

where ζ has empirically been taken as 2.5.

3.5 Experimental Results

A flow chart describing the entire shape reconstruction for uniformly colored objects process is shown in figure 3.2. Figure 3.3 through 3.7 demonstrate typical shape reconstructions obtained for single-color objects using both synthetic and real images. In all cases, the illuminant mapping, **M**, was unknown and was estimated using the technique described in section 3.2. Note that the algorithm is able to accurately recover the shape of the imaged surface in a qualitative visual sense in all cases.

3.6 Effect of Non-Constant Surface Reflectance on Shape Reconstruction

As previously discussed, one of the prevailing assumptions for shape reconstruction is that of *constant* surface reflectance. This implies that the surface reflectance term, **S**, is assumed to vary only with wavelength, λ, and not with pixel coordinates (x, y). Thus, even if the illuminant mapping, **M**, is known up to a surface reflectance, the proper

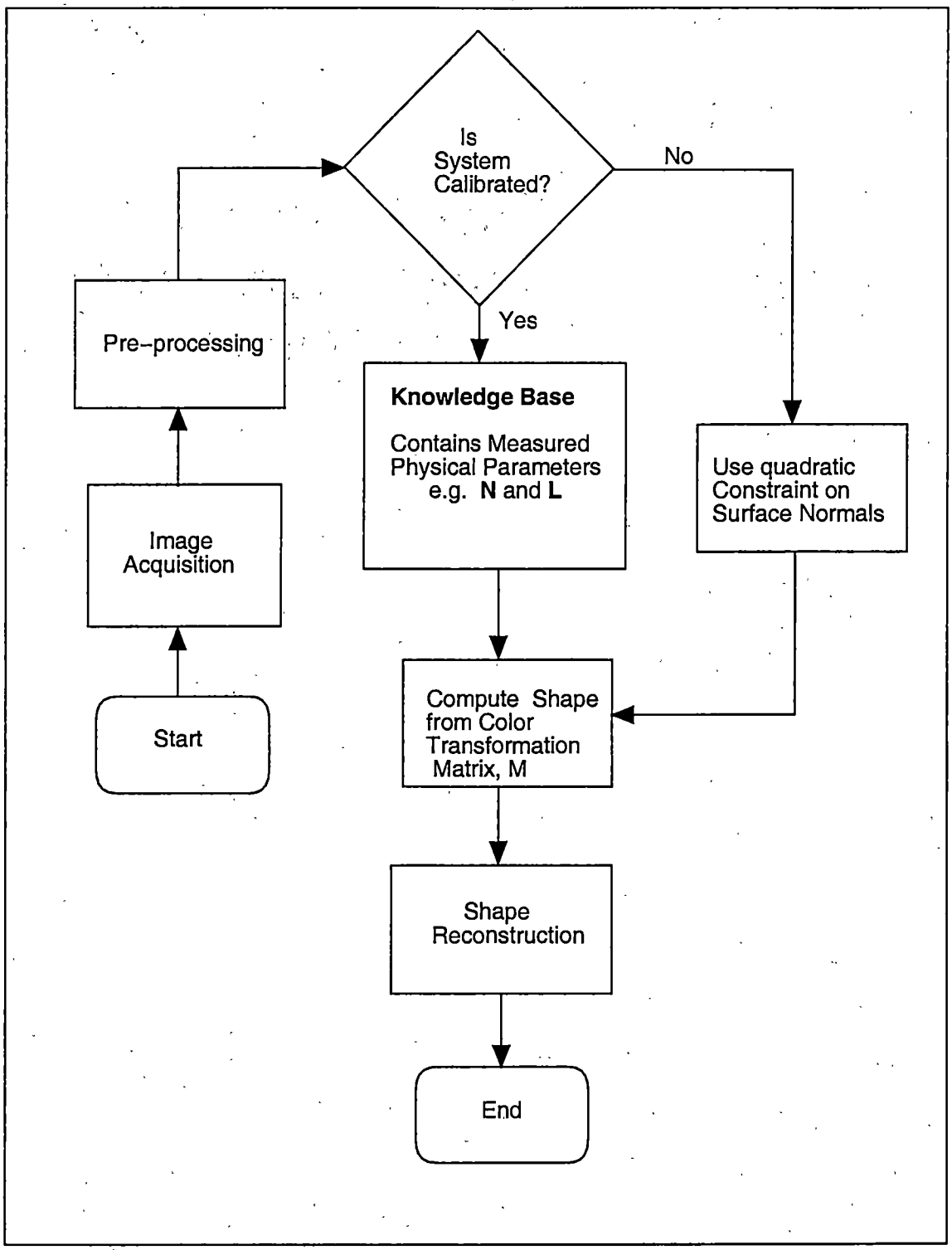
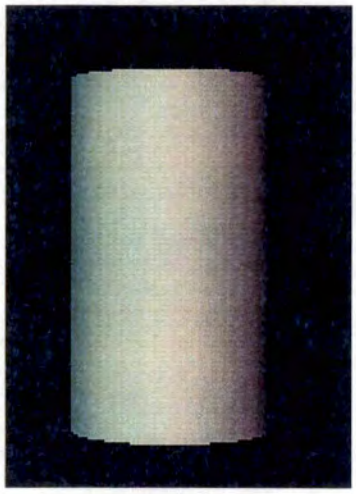
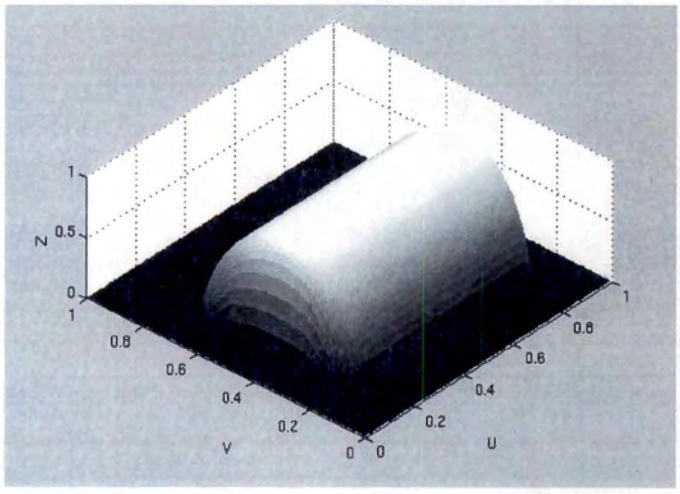


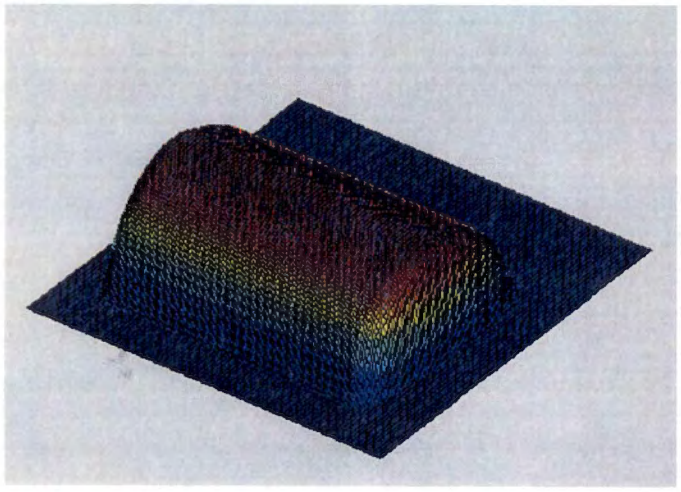
Figure 3.2: *The flow chart for 3D shape reconstruction for single-color objects*



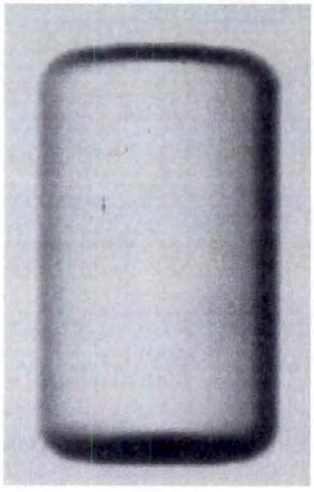
a Input image



b

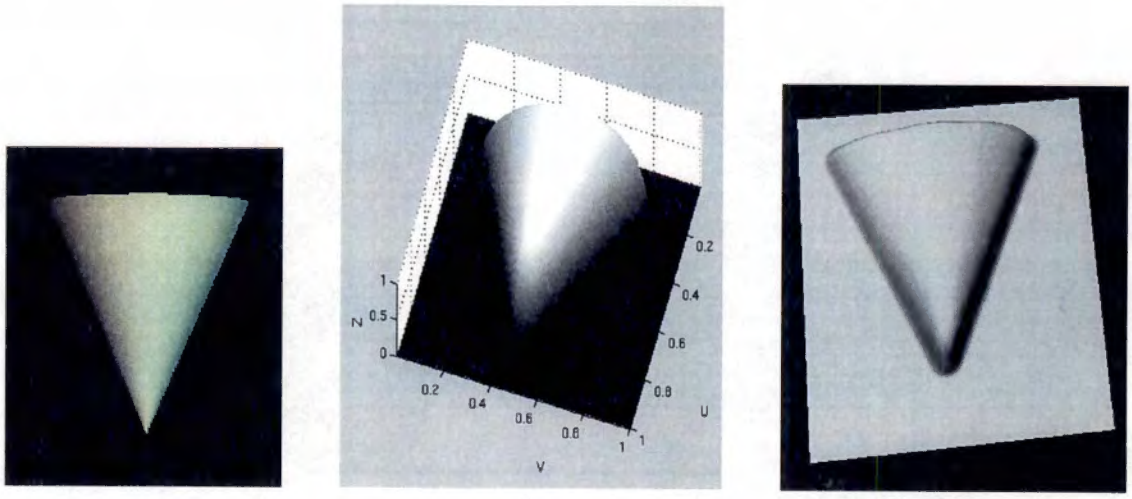


c



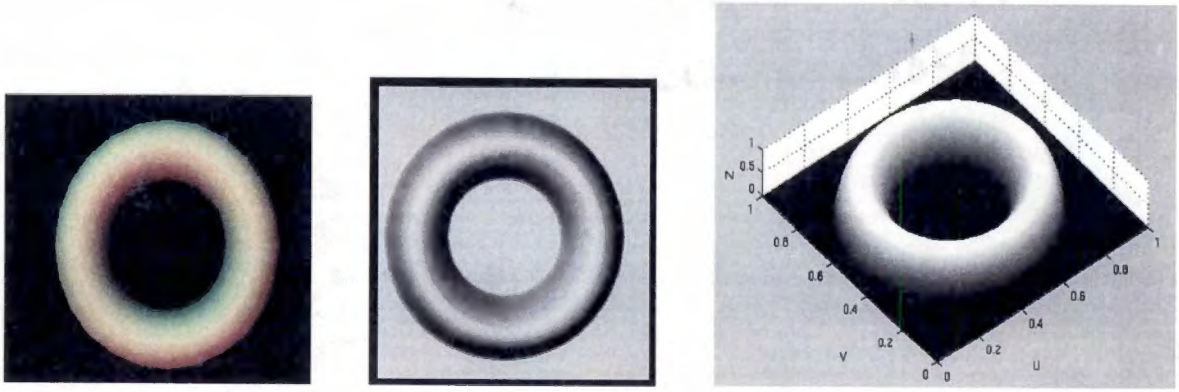
d

Figure 3.3: *Shape reconstruction of a single-color cylinder: (a) is synthetic input image. (b), (c) and (d) are results of different views of the reconstructed shape. (d) is a view displayed using an inventor.*



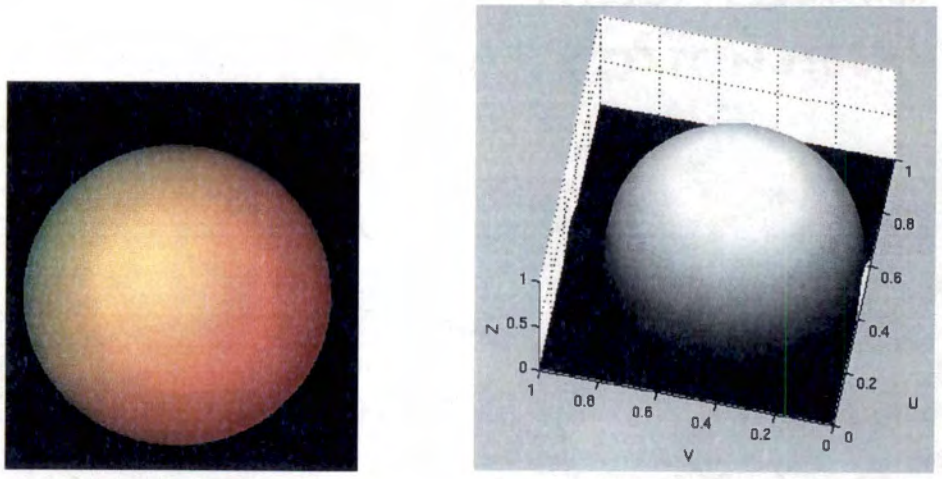
a b c

Figure 3.4: *Shape reconstruction of a single-color cone: (a) is synthetic input image. (b) and (c) are the results of shape reconstruction displayed at different views.*



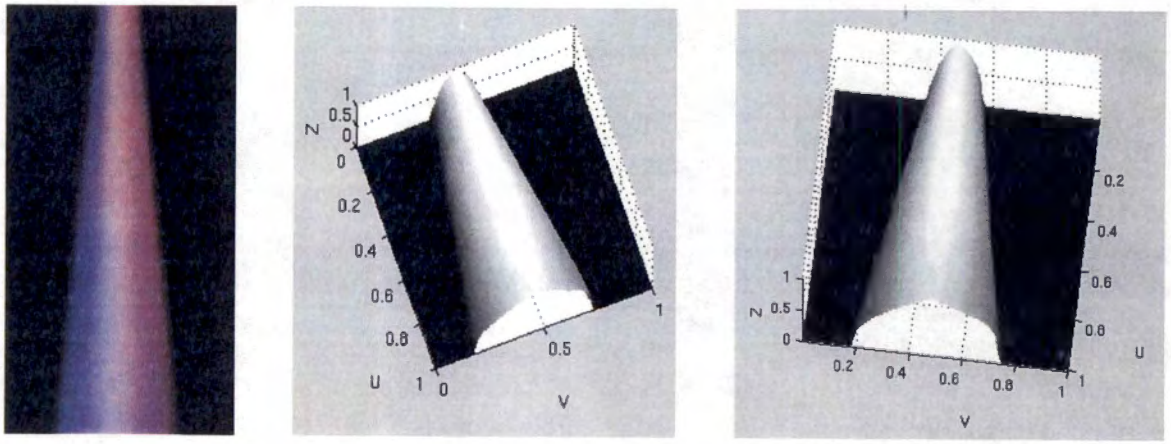
a b c

Figure 3.5: *Shape reconstruction of a single-color torus: (a) is synthetic input image. Results of shape reconstructions for different views are shown in (b) and (c).*



a b

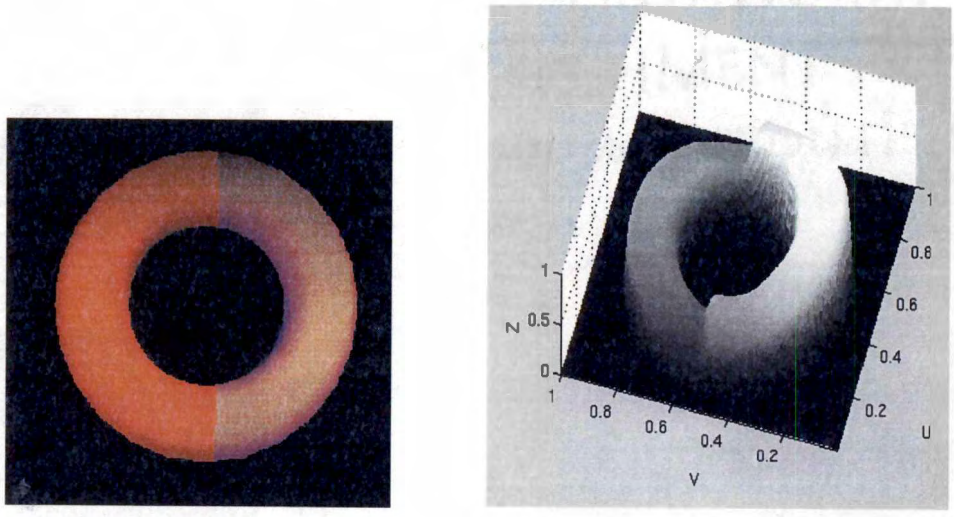
Figure 3.6: *Shape reconstruction of a single-color sphere: (a) is synthetic input image. (b) is the corresponding shape reconstruction.*



a b c

Figure 3.7: *Shape reconstruction of a single-color real cone image: (a) is real input image captured with Kodak DC120 Zoom. Results of different views of the reconstructed shape are shown in (b) and (c).*

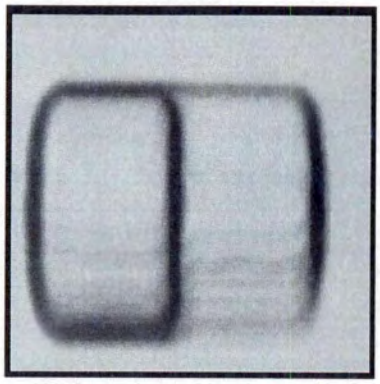
object shape cannot be recovered as is shown in figures 3.8 and 3.9. In figure 3.8, the shape of a multi-color torus was estimated using the existing shape-from-color technique. The reconstructed shape shown in figure 3.8b is a qualitatively poor match for a torus. This in accurate reconstruction results because the surface shapes corresponding to different colors in the image domain have been recovered to varying degrees. The inability to handle varying surface colors is one of the major limitations of existing shape-from-color techniques. In chapter 4, a detailed analysis of the behavior of the existing shape-from-color algorithm will be presented and used to develop three new algorithms for recovering shape-from-color in the presence of varying surface reflectance.



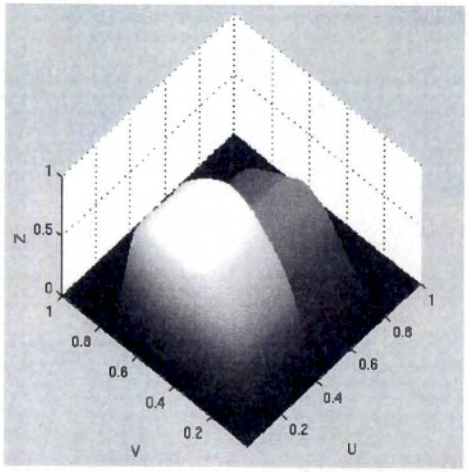
a b

Figure 3.8: *(a)Effect of non-constant surface reflectance on shape reconstruction: (a) is a multi-color torus whose 3D shape is to be reconstructed. (b) is the resulting shape reconstruction using a uniform single surface color approach. The shape falls short of a torus.*

a



b

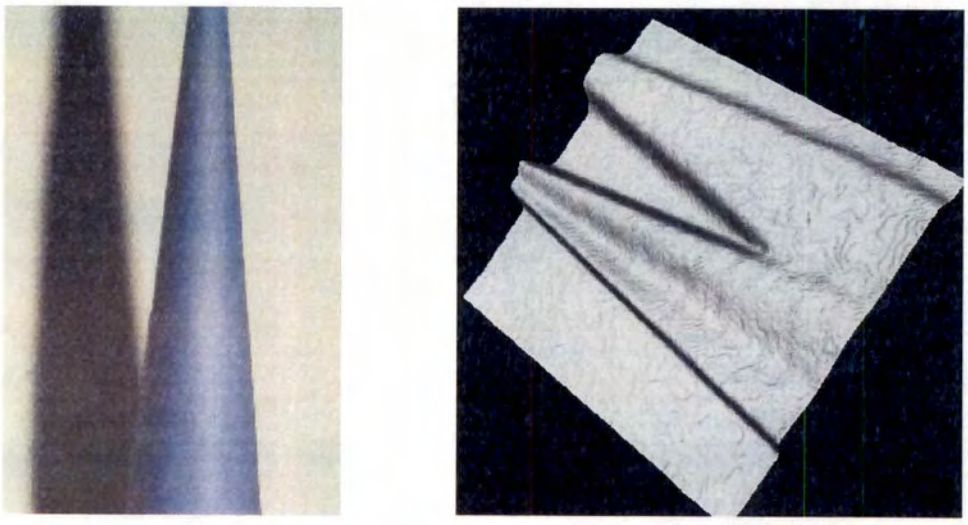


c

Figure 3.9: *Effect of non-constant surface reflectance on shape reconstruction: (a) is synthetic cylindrical image illuminated by three lights of different spectral content. (b) and (c) are different views of the corresponding 3D shape reconstruction using a single uniform surface color approach.*

3.7 Effect of Shadow on Shape Reconstruction

Self-occlusion and shadows pose the same problem as *non-constant* surface reflectance. The major distinction here is that the illuminant parameters in matrix,$\mathbf{A} = \mathbf{LN}$, vary and may be written as $\mathbf{A}(x, y; \lambda) = \mathbf{L}(x, y, \lambda)\mathbf{N}(x, y)$. The existence of the cast shadow in the scene confuses the vision system which consequently leads to a false shape reconstruction as shown in figures 3.10 and 3.11. Hence, a means of identifying and removing shadows in color imagery is desired. In chapter 5, a novel algorithm for identifying and eliminating shadows based on color-space clustering will be presented.



a b

Figure 3.10: *Effect of shadow on shape reconstruction: (a) is a real image of a cone with a shadow, captured with Kodak DC120 Zoom camera. (b) is the corresponding 3D shape reconstruction. Shadow mars shape reconstruction.*

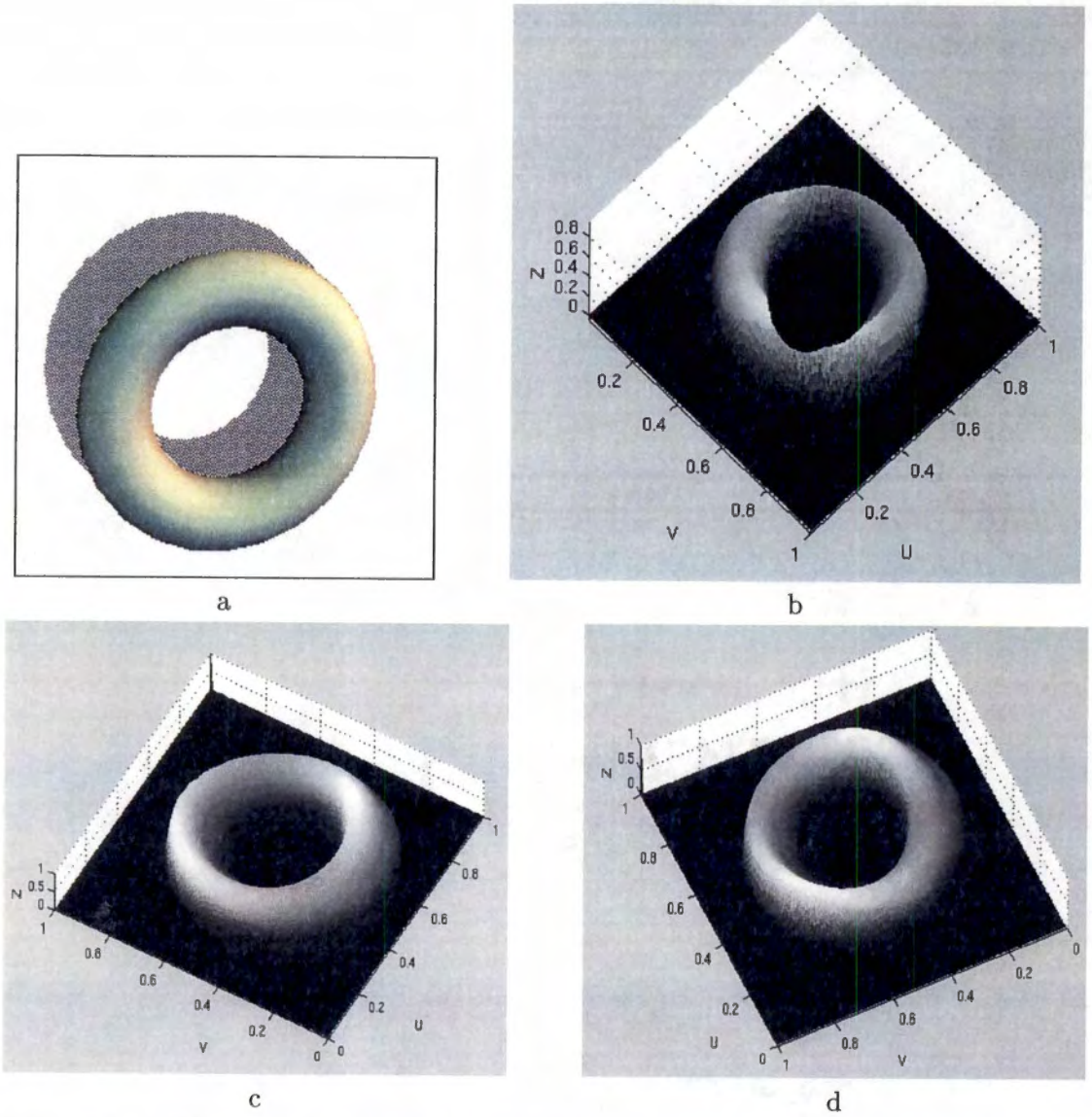


Figure 3.11: *Effect of shadow on shape reconstruction: (a) is synthetic image of a torus with a shadow. Results of the shape reconstruction are shown in (b) through (d). Shadow has a negative effect on shape reconstruction.*

3.8 Conclusion

In this chapter, shape reconstruction for single-color objects was discussed. The limitations of existing techniques in the presence of *non-constant* surface reflectance and shadow were also examined.

This dissertation is built upon Drew's work on single-color objects but extended to multi-color objects by relaxing the major assumptions. In the next chapter, three different techniques for reconstructing shape for multi-color objects are presented.

CHAPTER 4

Shape Reconstruction for Multi-Color Objects

In the previous chapter, the applicability and limitations of the existing shape-from-color techniques were presented. It was shown that the algorithm tends to perform poorly in the presence of non-constant surface reflectance and shadows. In this chapter, three new algorithms for recovering surface shape from color images with surfaces that display non-constant surface reflectance are presented.

4.1 Problem Overview

Reconsider the apparent color to surface normal mapping

$$\vec{n}_s(x,y) = (\mathbf{L}(\lambda)\mathbf{N})^{-1}\mathbf{S}^{-1}(x,y;\lambda)\vec{E}(x,y;\lambda) = \mathbf{A}^{-1}(\lambda)\mathbf{S}^{-1}(x,y;\lambda)\vec{E}(x,y;\lambda). \tag{4.1}$$

Where $\mathbf{A}(\lambda) = \mathbf{L}(\lambda)\mathbf{N}$. If a constant surface reflectance is assumed, then this expression is rewritten as

$$\tilde{n}_s(x,y) = \mathbf{A}^{-1}(\lambda)\mathbf{S}_c(\lambda)^{-1}\vec{E}(x,y;\lambda). \tag{4.2}$$

In this expression, the illumination mapping has no dependence on the pixel location (x,y).

Now, assume equation 4.2 is used to estimate surface shape for a surface with non-constant surface reflectance as given by equation 4.2. The actual surface normal at a given point, $\vec{n}_s$, is related to the recovered surface normal,

$$\vec{n}_s(x,y) = \Gamma(x,y)\tilde{n}_s(x,y) \tag{4.3}$$

where

$$\Gamma(x,y) \in \mathcal{R}^{3\times 3}; \Gamma(x,y) = \mathrm{diag}(\gamma_1, \gamma_2, \gamma_3).$$

Recovering accurate shape estimates from color in the presence of non-constant surface reflectance can, therefore, be viewed as a six dimensional problem in which the three elements of the approximate surface normal, $\tilde{n}_s$, and the three diagonal elements of the correction matrix, $\Gamma(x,y)$, must be calculated. Equation 4.2 can be used to provide three constraints for the solution of this six dimensional problem, while the normality constraint of the surface orientation vectors provides a fourth. To solve the shape-from-color problem for multi-color surfaces, two more constraining equations must be identified.

Note, however, that the traditional shape-from-color approaches have not made explicit use of the surface smoothness they assume. By requiring the object to demonstrate $\mathcal{C}^2$ smoothness, the surface normal gradients, $\frac{\delta \vec{n}_s(x,y)}{\delta x}$ and $\frac{\delta \vec{n}_s(x,y)}{\delta y}$, can be used to provide the two additional constraints required to estimate surface shape for multi-color objects.

In the remainder of this chapter, three different algorithms will be presented which make explicit use of the surface smoothness to estimate shape-from-color for multi-color surfaces. In the first technique, the image is segmented into regions of constant surface reflectance and shape is estimated for each region. The global surface is then constructed by aligning the various region shapes using boundary smoothness constraints. The second algorithm employs a similar method as Horn's original shape from shading technique by propagating solutions from points of known orientation. The third algorithm uses a variational approach to fit an optimal smooth surface to the given color image data given the initial solution assuming a constant surface orientation, $\tilde{n}_s$.

4.2 Shape Reconstruction by Pre-Segmentation

As discussed above, when a surface displays non-constant reflectance, $\mathbf{S}(x, y; \lambda)$ varies both spatially and spectrally. Thus, each distinct color region of the surface displays a different shape to apparent color mapping. If these regions of constant surface reflectance can be identified prior to shape estimation, boundary conditions can be used to recover the shape. Pre-identification of constant surface reflectance regions requires the image to be segmented based on color. In the work presented, both K-Means [38, 39] and a robust algorithm reported by Comaniciu, et. al. [40] are used to identify regions of constant color. Boundary pixels are identified after the color segmentation process using an 8-neighbor connectedness process.

Recall from equation 4.1 that

$$\vec{n}_s(x, y) = \mathbf{A}^{-1}(\lambda)\mathbf{S}^{-1}(x, y; \lambda)\vec{E}(x, y; \lambda)$$

If the surface is assumed to be continuous at the boundary of two color patches, then adjacent surface normals on either side of the boundary between regions i and j are such that $\vec{n}_{s,i}(x, y) \approx \vec{n}_{s,j}(x, y)$. If it is assumed that the entire surface is the same color as region i such that $\mathbf{S}_i = \mathbf{S}_c$, then the correction factor for every normal in region j, $\mathbf{\Gamma}_j(x, y)$ can be found by entry-wise division of $\vec{n}_{s,i}(x, y)$ by $\tilde{n}_{s,j}(x, y)$ such that

$$\mathbf{\Gamma}_j(x, y) = \begin{bmatrix} \frac{n^x_{s,i}}{\tilde{n}^x_{s,j}} & 0 & 0 \\ 0 & \frac{n^y_{s,i}}{\tilde{n}^y_{s,j}} & 0 \\ 0 & 0 & \frac{n^z_{s,i}}{\tilde{n}^z_{s,j}} \end{bmatrix}. \tag{4.4}$$

If this same procedure is employed between all region boundaries, the entire object surface can be recovered up to an unknown affine transformation represented by $\mathbf{S}_c(x, y; \lambda)$. If the surface reflectance of one region is known, then the surface can be recovered completely.

The pre-segmentation algorithm for shape reconstruction for multi-color surfaces is summarized below.

- Identify background and outlier pixels. They are not used in the shape reconstruction.
- Segment the multi-color image into its main representative colors and keep track of pixels in each segment.
- Extract the color boundaries based on the result of the segmentation.
- Compute the preliminary surface orientation map, $\tilde{n}_s(x, y)$ using equation 4.2.
- Determine the correction matrix, $\mathbf{\Gamma}_j$ given in equation 4.4 above for all color regions.
- Compute the corrected surface orientation map, $\vec{n}_s(x, y)$.
- Compute the smooth depth map shape.

4.3 Shape Reconstruction by Normal Propagation

While the pre-segmentation approach works well for surfaces with distinct regions of constant surface reflectance, it is problematic for surfaces with either slowly varying or high frequency color textures, such as human faces or bricks, or any object which is difficult to segment based on color. As observed in figure 4.20(e), the segmentation algorithm had difficulty recognizing only two distinct color regions, and thus the quality of the shape reconstruction is reduced. The second algorithm for recovering shape-from-color eliminates the need for image segmentation by employing a propagation approach based on surface smoothness constraints.

Initially, the surface orientation, $\tilde{n}_s$, is computed for each surface pixel as if it were a single color using the standard shape-from-color technique. Any background and outlier pixels are identified and are not used for the shape reconstruction process. The relative depth map of the surface is then computed using equation 3.12.

In the second phase of the algorithm, the assumed smoothness of the imaged surface is employed to construct two constraints

$$\begin{aligned} \tilde{n}_s(x,y) &\approx \tilde{n}_s(x+i,y+j) \text{ for } i=-1,0,1; j=-1,0,1, \text{and} \\ z(x,y) &\approx z(x+i,y+j) \text{ for } i=-1,0,1; j=-1,0,1. \end{aligned} \tag{4.5}$$

Assuming $\tilde{n}_s(x,y)$ and $z(x,y)$ are correct for the current surface point (x,y), then the neighboring surface orientations must be such that

$$\begin{aligned} \cos^{-1}\left(\tilde{n}_s(x,y)^T\tilde{n}_s(x+i,y+j)\right) &< \epsilon_1 \text{ for } i=-1,0,1; j=-1,0,1 \text{ and} \\ z(x,y)-z(x+i,y+j) &< \epsilon_2 \text{ for } i=-1,0,1; j=-1,0,1 \end{aligned} \tag{4.6}$$

where ϵ_1 and ϵ_2 are user-defined thresholds. In the experimental results shown in this chapter, $\epsilon_1 = 0.003$ radians and $\epsilon_2 = 0.04$.

If both of the above conditions are violated for a given neighbor, then the algorithm assumes that neighbor is a different color than the current pixel and a correction factor, $\mathbf{\Gamma}(x,y)$, is computed. Again, using the same constraint between adjacent surface normals as employed by the pre-segmentation technique, $\vec{n}_s(x,y) \approx \vec{n}_s(x+i,y+j)$ for $i = -1,0,1; j=-1,0,1$, the correction factor is again given as

$$\mathbf{\Gamma}(x+i,y+j) = \begin{bmatrix} \frac{n_s(x,y)^x}{\tilde{n}_s(x+i,y+j)^x} & 0 & 0 \\ 0 & \frac{n_s(x,y)^y}{\tilde{n}_s(x+i,y+j)^y} & 0 \\ 0 & 0 & \frac{n_s(x,y)^z}{\tilde{n}_s(x,i,y+j)^z} \end{bmatrix}. \tag{4.7}$$

Hence, if the correct surface orientation is known for at least one point on the surface, then the entire surface can be reconstructed by propagating the correct surface orientation using the constraints of equation 4.6.

As is the case for shape from shading techniques which employ local gradient information, image noise can reduce the quality of surface shapes recovered using propagation approach. To reduce the effect of image noise on the shape reconstruction process, the input image is smoothed using an edge preserving anisotropic diffusion smoothing algorithm [41, 42, 43]. To further reduce the effects of image noise, a 4-connected averaging filter is applied to the surface orientation map, $\tilde{n}_s(x, y)$, prior to the shape reconstruction process. As a final step, a median filter [44, 45] is applied to the relative depth map computed with the corrected surface normals.

The normal propagation algorithm for shape reconstruction described above is summarized below:

1. Apply anisotropic smoothing filter to image.

2. Identify and remove background and outlier pixels.

3. Compute the surface orientation, $\tilde{n}_s(x, y)$ for each surface point as if the object were a single-color.

4. Apply the 4-connected averaging filter to the surface orientation map, $\tilde{n}_s(x, y)$.

5. Compute the angle between normals for adjacent surface points.

6. Compute the relative depth between adjacent surface points.

7. If the results in *steps 5 and 6* above are less than ϵ_1 and ϵ_2 respectively, *go to step 9.*

8. If the results in *steps 5 and 6* are greater than ϵ_1 and ϵ_2 respectively, determine $\mathbf{\Gamma}(x,y)$ and compute $\vec{n}_s(x,y) = \mathbf{\Gamma}(x,y)\tilde{n}_s(x,y)$.

9. Compute the shape corresponding to the image point using equation(3.12).

10. Post-process the result by median filtering of the mapped depth.

4.4 Shape Reconstruction Using a Constrained Variational Approach

The accuracy of surface shapes recovered via the normal propagation technique is susceptible to high degrees of image noise. As an alternative to solution propagation using local gradient estimates, variational techniques are often developed that employ global smoothness constraints to iteratively solve for optimal solutions given a set of input data. In this section, a variational approach to recovering shape-from-color for non-constant surface reflectance is presented.

From equation(3.2) the color vector associated with a three-channel camera is given by

$$\vec{E}(x,y;\lambda) = \mathbf{M}(x,y;\lambda)\vec{n}_s(x,y) \tag{4.8}$$

Given initial estimates of both the illuminant mapping, $\mathbf{M}(x,y;\lambda)$, and the surface orientation map, $\vec{n}_s(x,y)$, the error associated with those estimates is expressed as

$$\delta(x,y) = \vec{E}(x,y;\lambda) - \mathbf{M}(x,y;\lambda)\vec{n}_s(x,y) \tag{4.9}$$

Thus, a possible solution to the shape-from-color problem is given by the illuminant mapping and surface orientation map which minimizes the functional

$$\delta = \int\int \left\|\vec{E} - \mathbf{M}\vec{n}\right\|^2 dxdy \tag{4.10}$$

where dependence on x, y, and λ have been dropped for simplicity.

Many possible solutions to equation 4.10 could exist, however, that violate (1) the quadratic constraint on the recovered surface normals, $\vec{n}^T\vec{n} = 1$, and (2) the assumed smoothness of the recovered orientation map. To insure that each normal is unity, a new term is incorporated into the functional using a Lagrangian multiplier, μ, yielding

$$\delta = \int\int \left(||\vec{E} - \mathbf{M}\vec{n}||^2 + \mu(||\vec{n}||^2 - 1)^2 \right) dxdy. \tag{4.11}$$

It is of interest to note that Drew's method for recovering shape given unknown lighting conditions is essentially a closed-form solution to the functional of equation 4.11 over both $\mathbf{M}$ and $\vec{n}$. In the presence of non-constant surface reflectance, however, an additional constraint based on surface smoothness must be added to the functional to insure a unique solution (or a solution up to an unknown rotation in the case of an unknown illuminant mapping). Using the smoothness penalty constraint,

$$||\vec{n}_x||^2 + ||\vec{n}_y||^2,$$

the functional is rewritten as

$$\delta = \int\int \left(\left\|\vec{E} - \mathbf{M}\vec{n}\right\|^2 + \beta(||\vec{n}_x||^2 + ||\vec{n}_y||^2) + \mu(||\vec{n}||^2 - 1)^2 \right) dxdy \tag{4.12}$$

where β is another multiplier for the smoothness constraint.

Assuming the illuminant mapping, $\mathbf{M}$, is known, the functional of equation 4.12 must be minimized with respect to the surface orientation map, $\vec{n}$. Using the calculus of variations, the Euler equation corresponding to equation 4.12 is given by

$$\mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta\nabla^2\vec{n} - \mu\vec{n} = 0. \tag{4.13}$$

Solving equation 4.13 for the scaled surface orientation $\mu\vec{n}$ yields

$$\mu\vec{n} = \mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta\nabla^2\vec{n} \tag{4.14}$$

Pre-multiplying equation 4.14 by $\vec{n}^T$ produces the following expression for the Lagrangian multiplier, μ,

$$\begin{aligned}
\mu &= \vec{n}^T \mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta \vec{n}^T \nabla^2 \vec{n} \\
&= \vec{n}^T \mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta(\nabla^2 \vec{n} \cdot \vec{n}) \\
&= (\mathbf{M}\vec{n})^T(\vec{E} - \mathbf{M}\vec{n}) + \beta(\nabla^2 \vec{n} \cdot \vec{n}) \qquad (4.15)
\end{aligned}$$

Substituting this expression for μ into equation 4.13 yields

$$(\mathbf{I} - \vec{n}\vec{n}^T)\mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + (\mathbf{I} - \vec{n}\vec{n}^T)\beta\nabla^2\vec{n} = 0 \qquad (4.16)$$

Letting $\mathbf{K} = (\mathbf{I} - \vec{n}\vec{n}^T) \in \mathcal{R}^{3\times 3}$ reduces the above expression to

$$\mathbf{K}\left(\mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta\nabla^2\vec{n}\right) = 0 \qquad (4.17)$$

To avoid the trivial solution to the above expression where

$$\vec{n} = \mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) + \beta\nabla^2\vec{n}$$

let the updated solution to the above expression be given by $\vec{m}$, such that $\vec{m} = \vec{v} + \kappa\vec{n}$ where $\vec{v}$ is orthogonal to $\vec{n}$. This leads to the expression

$$\mathbf{K}\vec{m} = \mathbf{K}(\vec{v} + \kappa\vec{n}) = \mathbf{K}\vec{v} = \vec{v}$$

where κ is a scalar $= \sqrt{1 - ||\vec{v}||^2}$. Hence, by substitution, equation(4.17) can be written as

$$\mathbf{K}\vec{m} = \vec{v} = \mathbf{K}\beta\nabla^2\vec{n} + \mathbf{K}\mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}) \qquad (4.18)$$

Using the above expressions for $\mathbf{K}, \vec{m}$ and $\vec{v}$, the following sets of equations can be used to iteratively update a globally optimal surface orientation map using color image data in the presence of non-constant surface reflectance,

$$\begin{aligned}
\vec{v}^{(k+1)} &= \mathbf{K}\left(\beta\nabla^2\vec{n}^{(k)} + \mathbf{M}^T(\vec{E} - \mathbf{M}\vec{n}^{(k)})\right) && (4.19) \\
\vec{n}^{(k+1)} &= \vec{v}^{(k+1)} + \vec{n}^{(k)}\sqrt{1 - \left\|\vec{v}^{(k+1)}\right\|^2} && (4.20) \\
\mathbf{K}^{k+1} &= \mathbf{I} - \vec{n}^{(k)}\vec{n}^{(k)^T}. && (4.21)
\end{aligned}$$

The iterative procedure is stopped when the change in the *RMS* error, $\Delta\delta$ is less than some user specified criteria, ϵ_3.

As with any iterative scheme, values for $\vec{n}$ must be supplied to initialize the process. In addition, an estimate for the shape-from-color mapping, $\mathbf{M}$, must be provided. Instead of supplying arbitrary values, Drew's single-color approach is used to initialize $\vec{n}$ assuming a constant surface reflectance and supply an estimate for the matrix, $\mathbf{M}$. By using the single color surface orientation map estimate, $\tilde{n}_s(x,y)$, the iterative scheme is starting from the quadratically constrained solution and proceeding to a solution which is both quadratically and smoothness constrained. It is of interest to note that the final solution for the surface orientation at step k, $\vec{n}^k(x,y)$, should be such that $\vec{n}^k = \mathbf{\Gamma}(x,y)\tilde{n}_s = \mathbf{\Gamma}(x,y)\vec{n}^0$.

4.5 Experimental Results

All three algorithms have been evaluated using numerous images of objects of different shapes and colors acquired under various illumination conditions. The first set of tests were conducted using synthetic images created with *SGI Showcase* using OpenGL. In these cases, the illuminant matrix, $\mathbf{A}$ is known, but the actual surface reflectance of the object is not. The surfaces recovered using all three approaches are shown in figures 4.1 - 4.5 for a subset of these objects. As can be observed from these figures, all three approaches performed quite well given these simple color variations and surface geometries.

The next data set includes real images of simple objects such as cylinders, spheres, and cones, a subset of which are shown in figures 4.6 - 4.8. Again, all three algorithms recovered qualitatively accurate shape reconstructions for the imaged objects.

The next set of images is made up of human faces of various ethnic backgrounds, and thus differing shapes and skin tones. Results from a subset of these images are shown in figures 4.9 through 4.19. Although varying in quality, it is possible to match the

recovered shape of the persons face to the given color image. It is interesting to note that the difference between the three algorithms become more apparent using this data set. In general, human faces are more difficult for segmentation algorithms to group into coherent regions, and thus the quality of the shapes recovered using the pre-segmentation approach is reduced in a few cases (see figure 4.11).

The final three figures 4.20 - 4.22 demonstrate the effects of poor color segmentation on the recovered shape using the pre-segmentation approach. In the first figure, the segmentation algorithm over-segmented the input image thereby resulting to poor shape reconstruction. In the second figure, the lower color region of the sphere is improperly segmented, creating a large 'hole' in the recovered surface. In the case of the third object, the highly textured (marbleized) surface results in many pixels being misclassified by color, leading to a less visually accurate shape reconstruction.

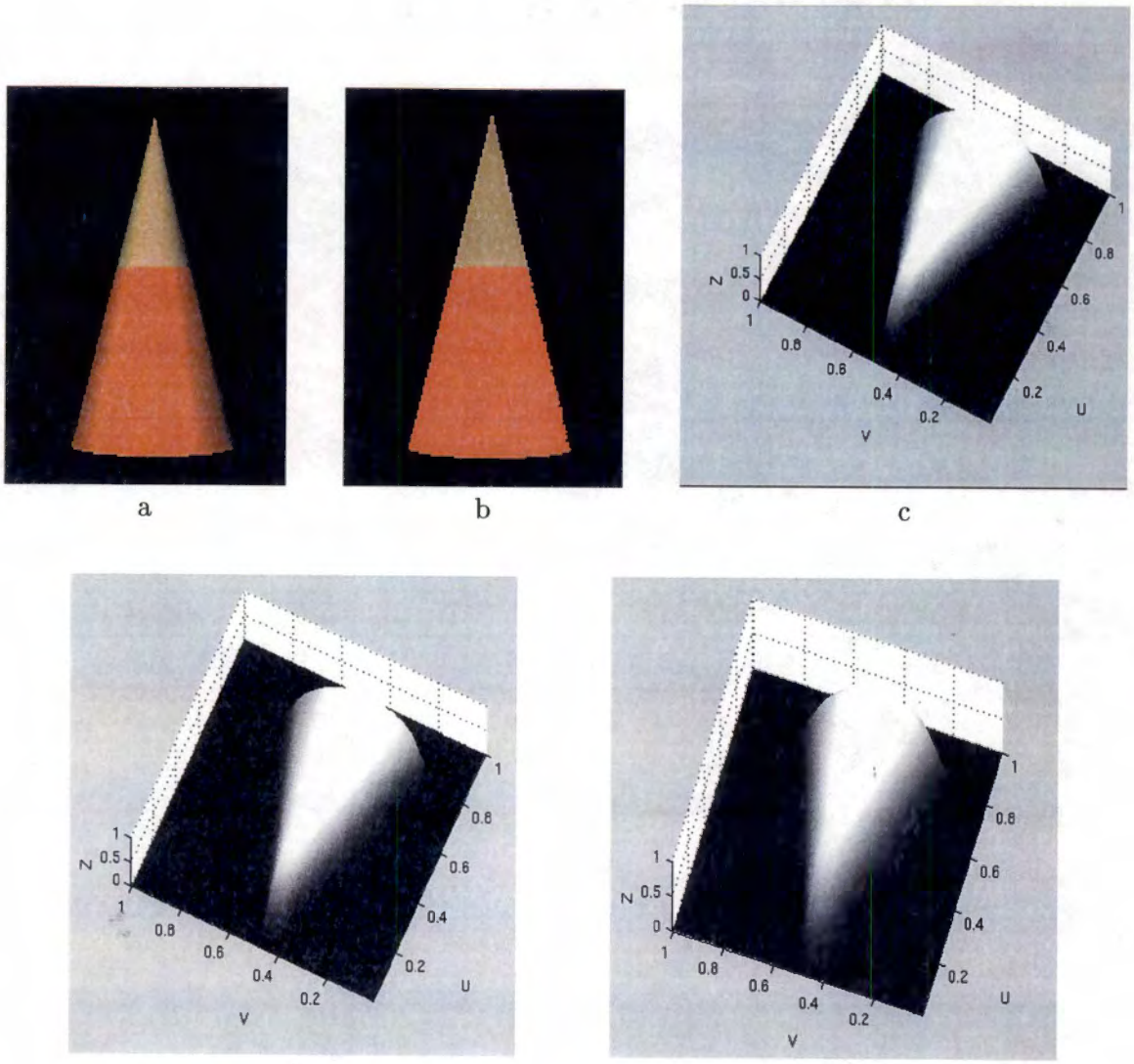


Figure 4.1: *Shape reconstruction of a multi-color cone: (a) is input synthetic image. (b) is the segmentation result. Results of the shape reconstruction by the three techniques are compared in (c) through (e) using (c) Pre-segmentation approach (d) Normal propagation approach and (e) Variational approach.*

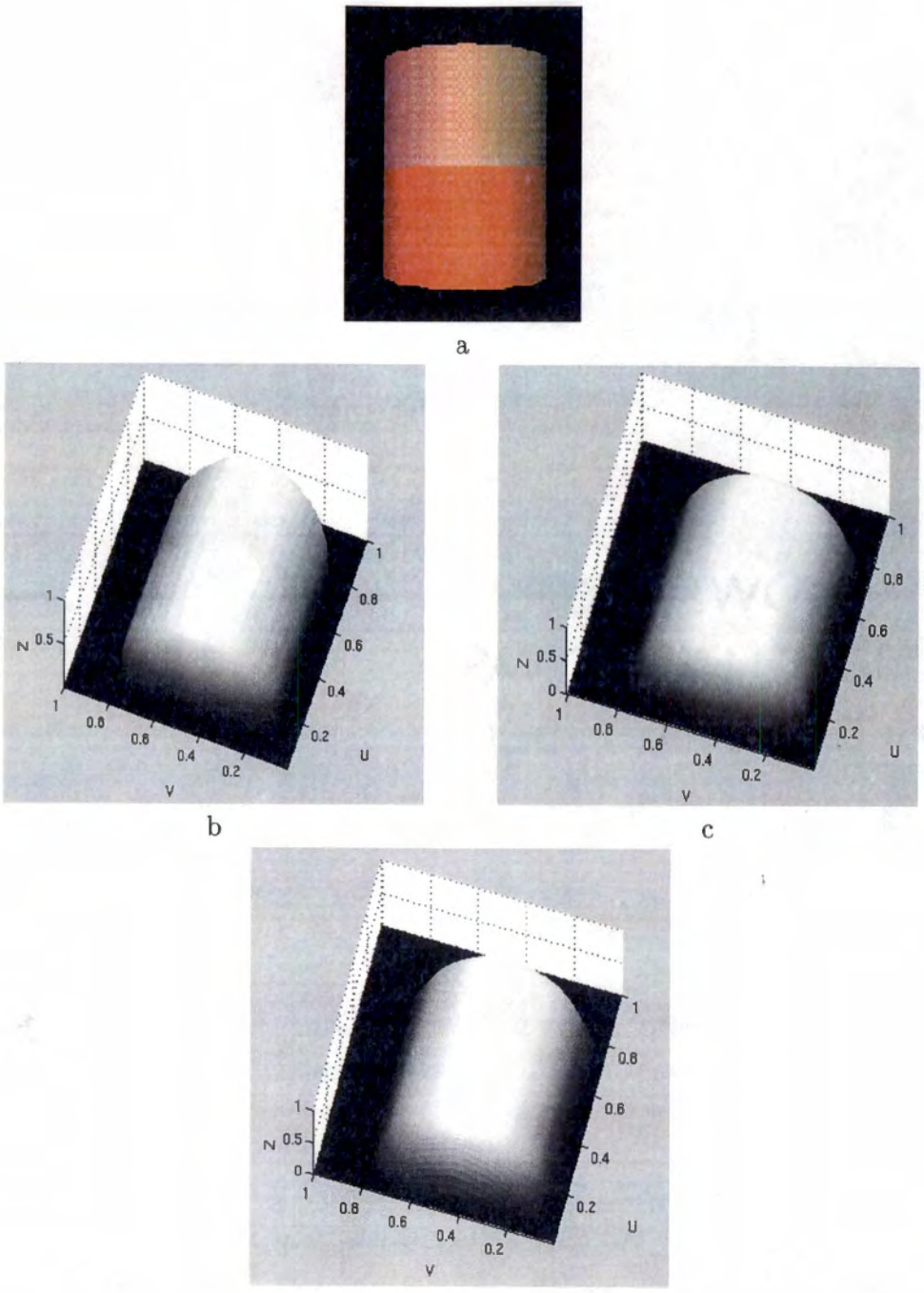


Figure 4.2: *Shape reconstruction of a multi-color cylinder: (a) is synthetic input image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

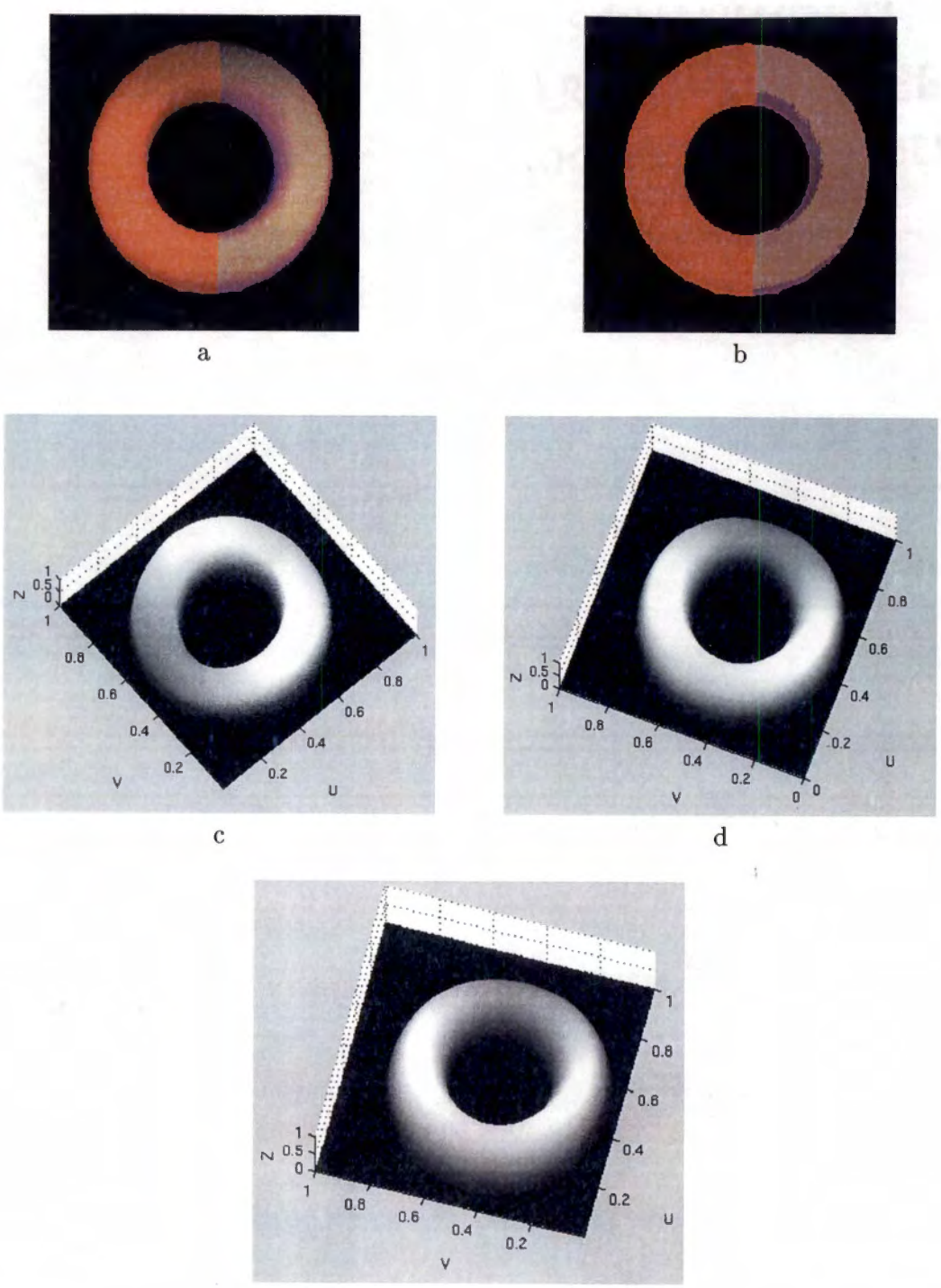
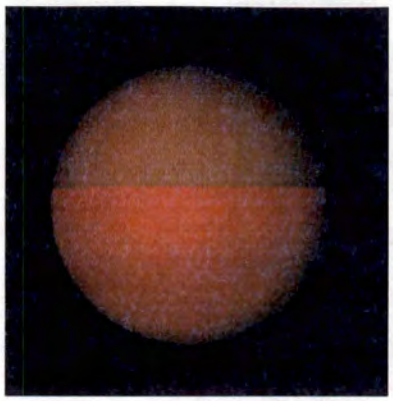
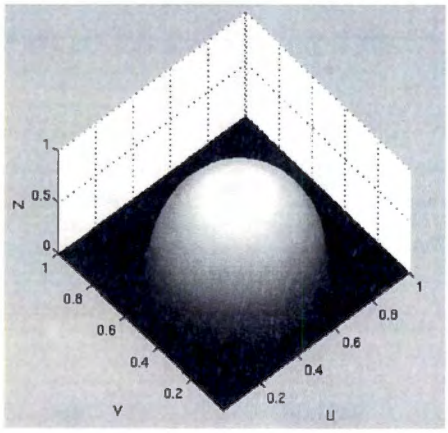


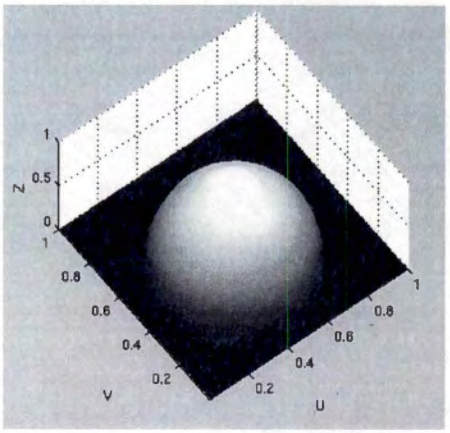
Figure 4.3: *Shape reconstruction of a multi-color torus: (a) is synthetic input image. (b) is the segmented result. Results of the shape reconstruction by the three techniques are compared in (c) through (e) using (c) Pre-segmentation approach (d) Normal propagation approach and (e) Variational approach.*



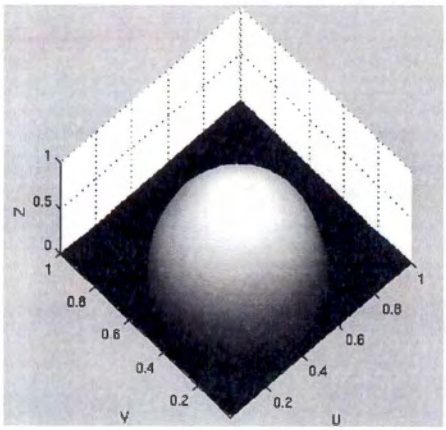
a



b



c



d

Figure 4.4: *Shape reconstruction of a multi-color sphere: (a) is a synthetic input image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

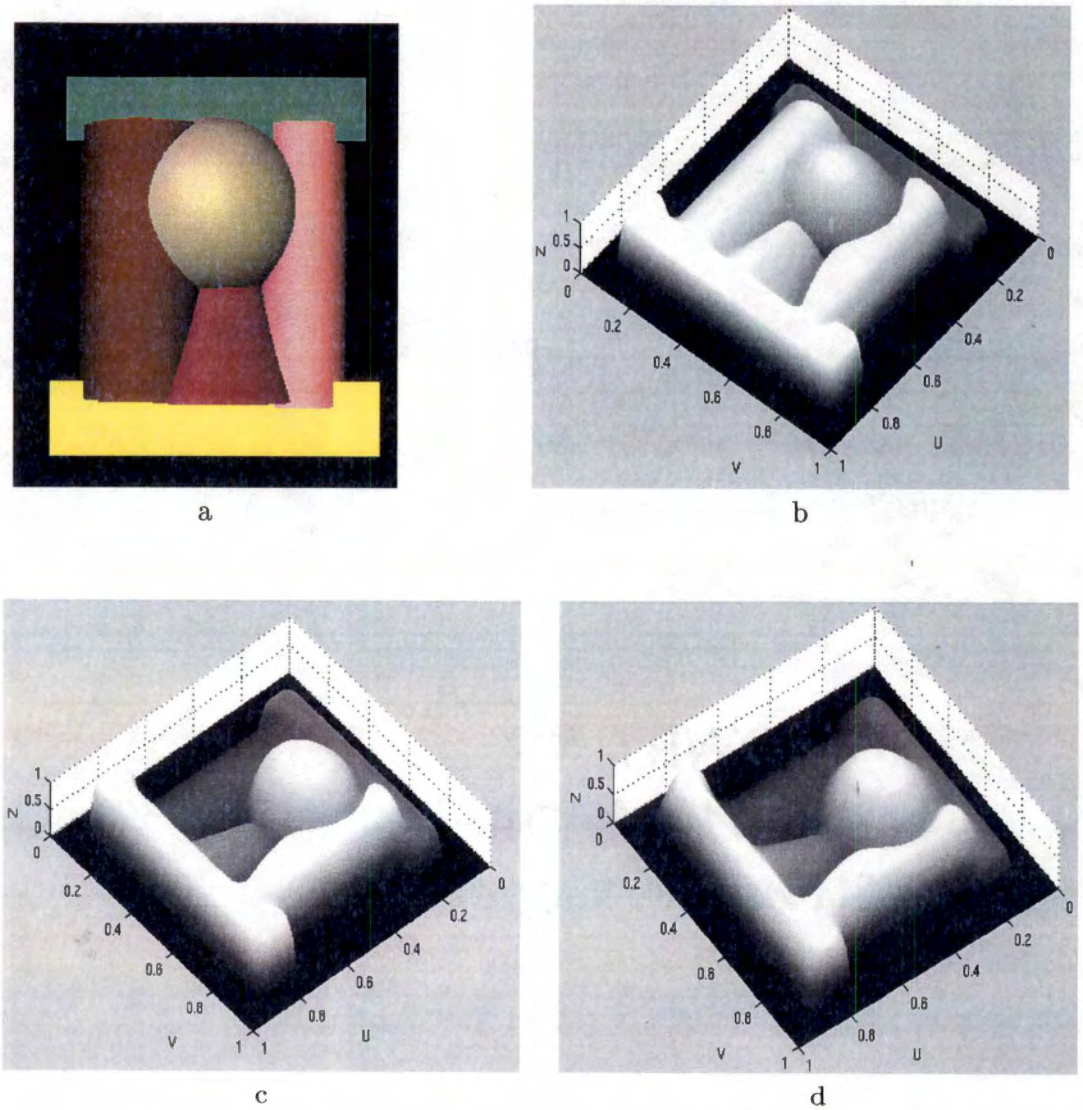


Figure 4.5: *Shape reconstruction of a multi-color complex object: (a) is the input image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

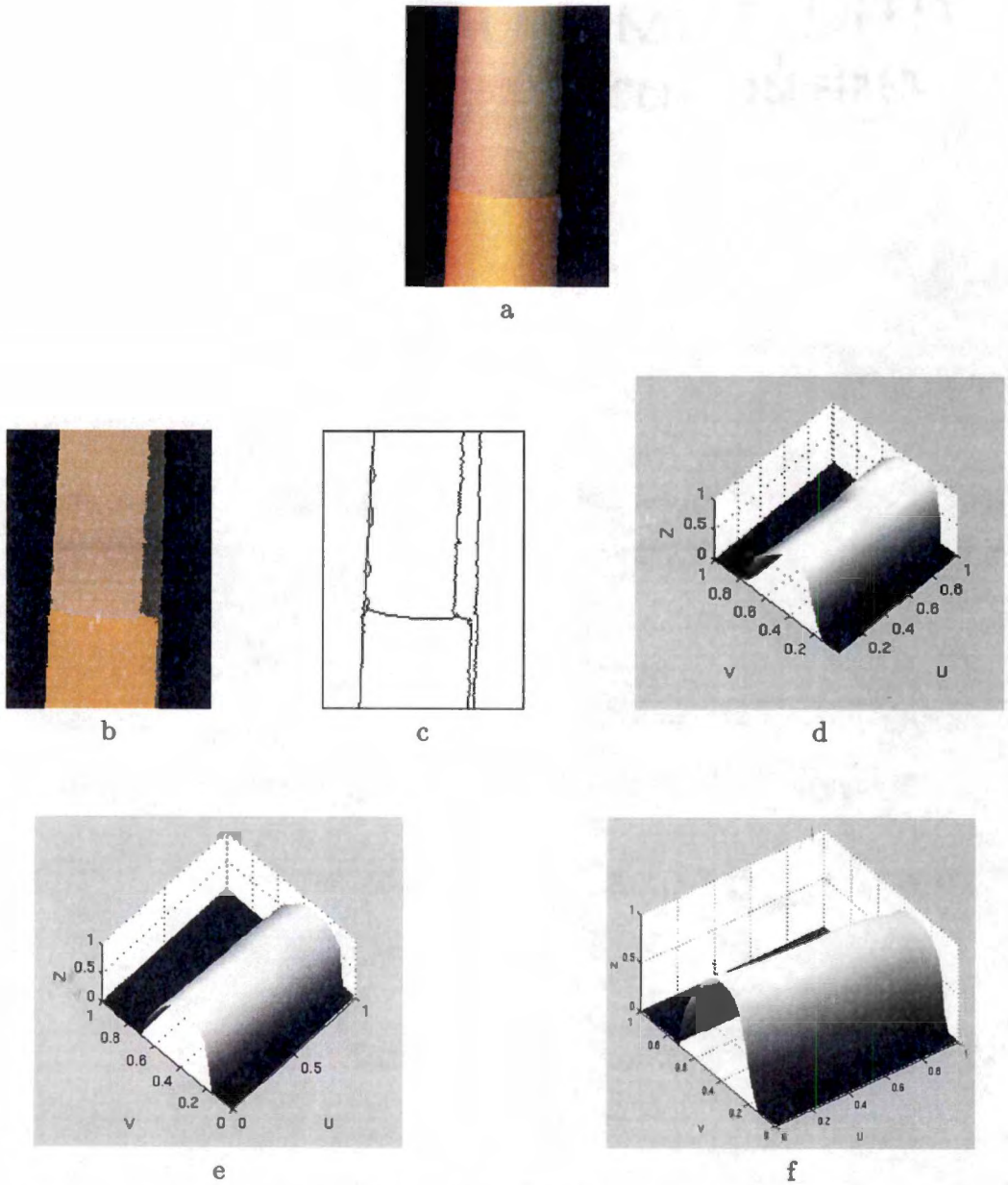
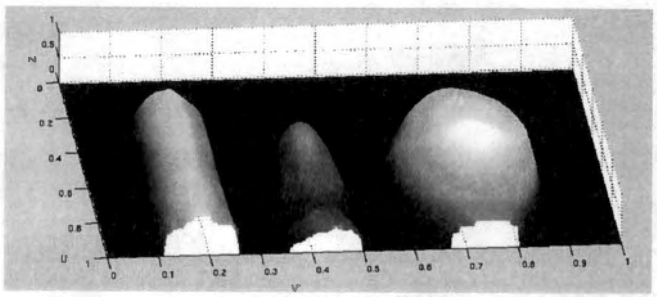
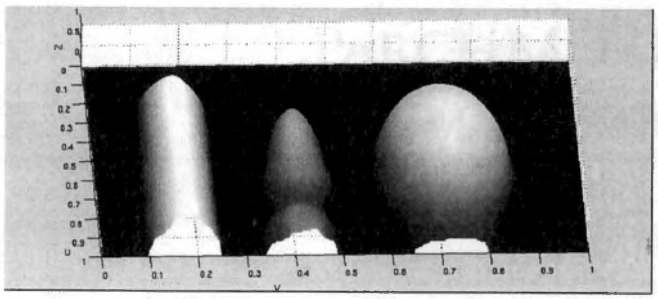


Figure 4.6: *Shape reconstruction of a real image of a cylinder: (a) is the input image. (b) is the segmentation result. (c) shows the extracted edges for color boundary identification. Results of the shape reconstruction by the three techniques are compared in (d) through (f) using (d) Pre-segmentation approach (e) Normal propagation approach and (f) Variational approach.*

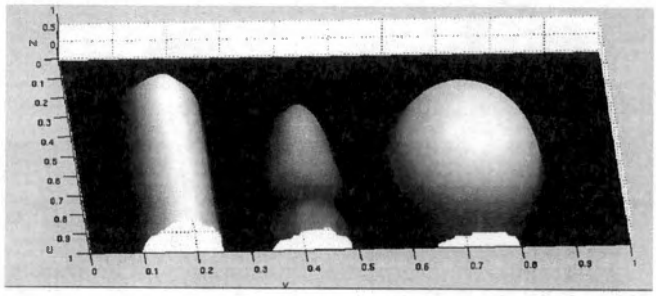
a



b

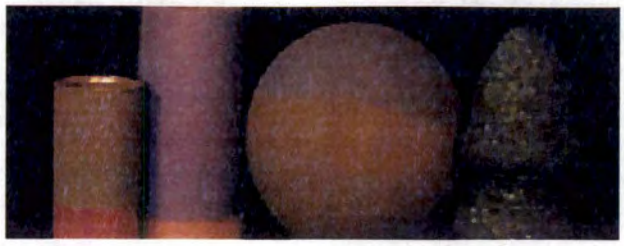


c

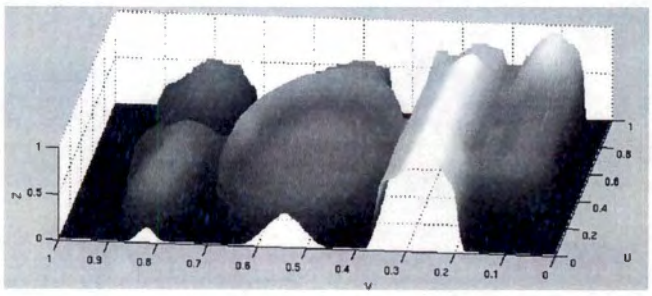


d

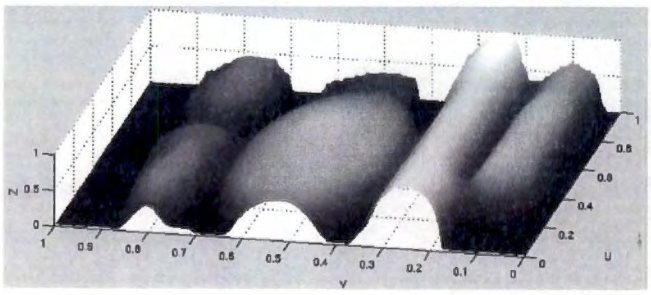
Figure 4.7: *Shape reconstruction of a multi-object (not in contact), multi-color image: (a) is the input image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*



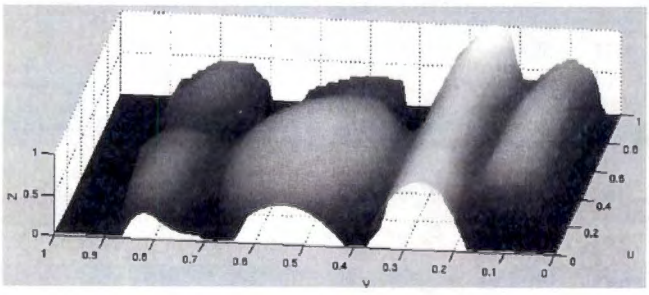
a



b



c



d

Figure 4.8: *Shape reconstruction of a multi-object (in contact), multi-color image: (a) the input image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

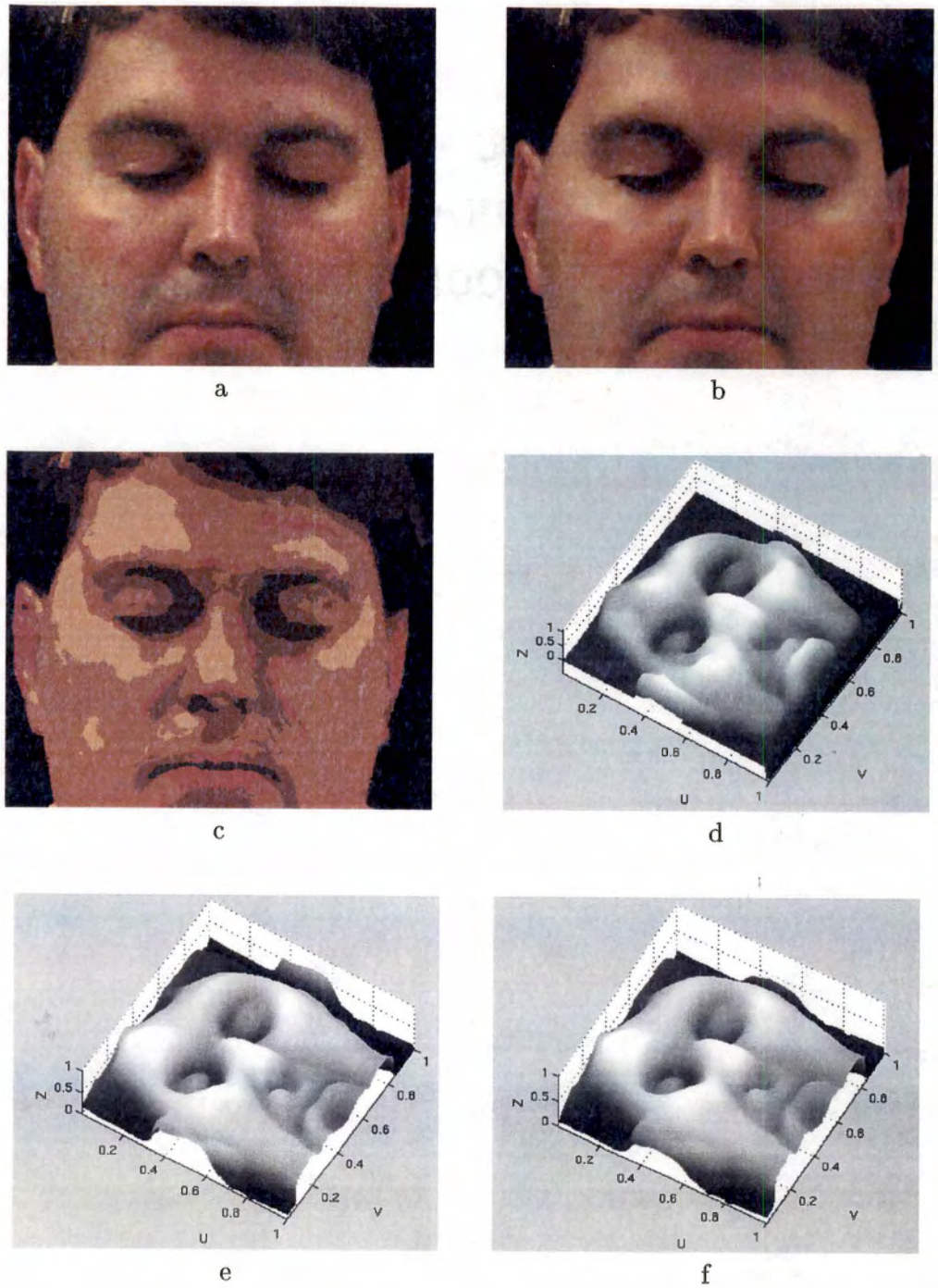


Figure 4.9: *Shape reconstruction of a human face: (a) the input image (Dr. Smith's face). (b) is the pre-processed image. (c) is the segmentation result. Results of the shape reconstruction by the three techniques are compared in (d) through (f) using (d) Pre-segmentation approach (e) Normal propagation approach and (f) Variational approach.*

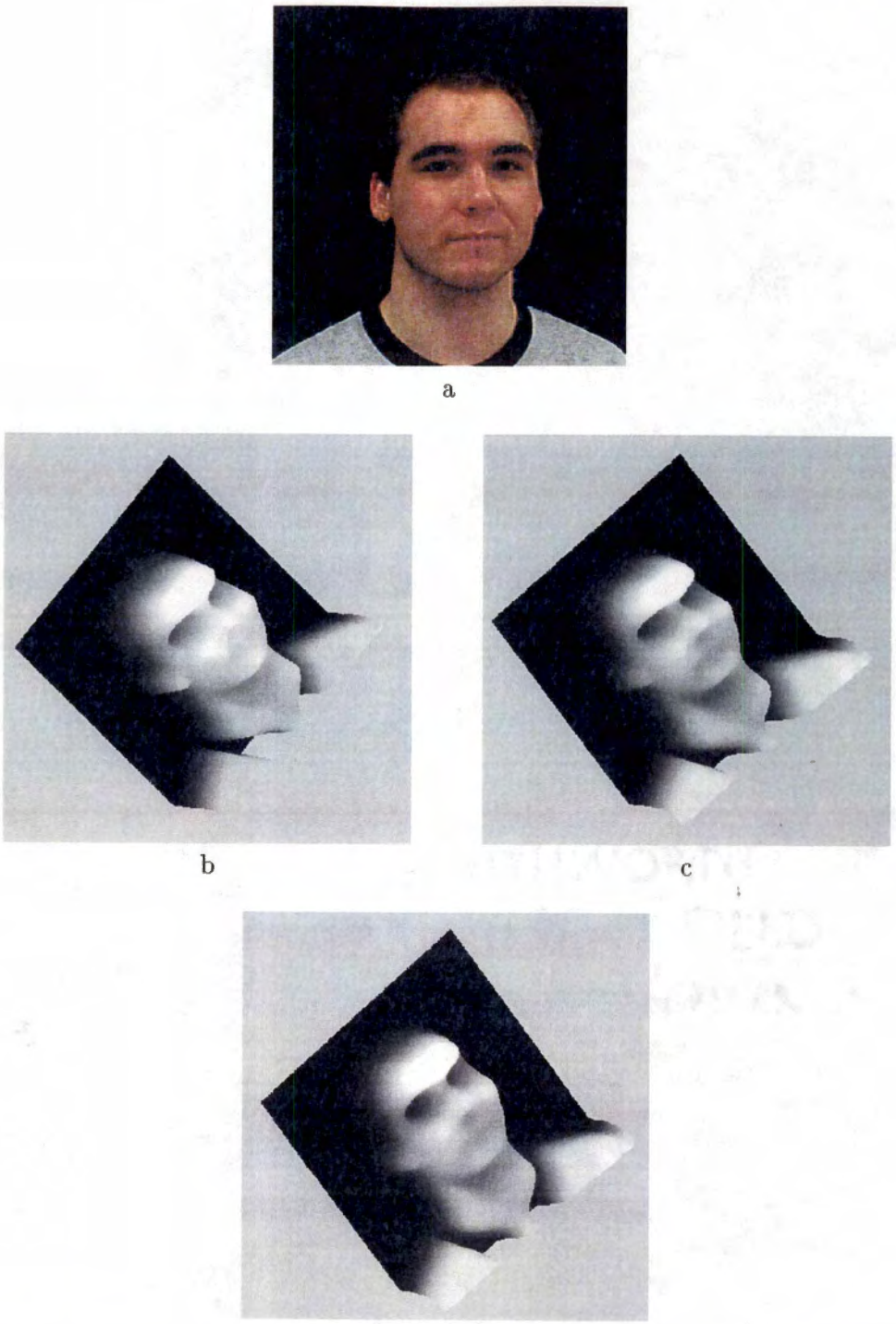
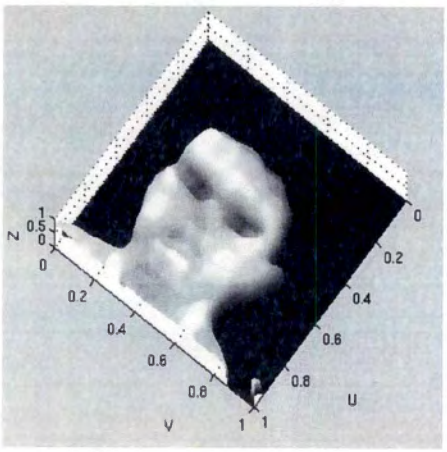
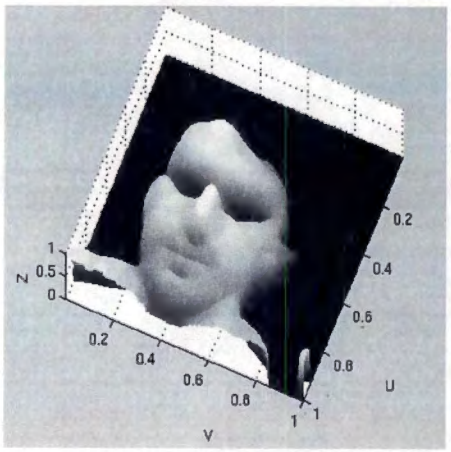


Figure 4.10: *Shape reconstruction of a human face: (a) the input image (Jason Rudisill's face). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

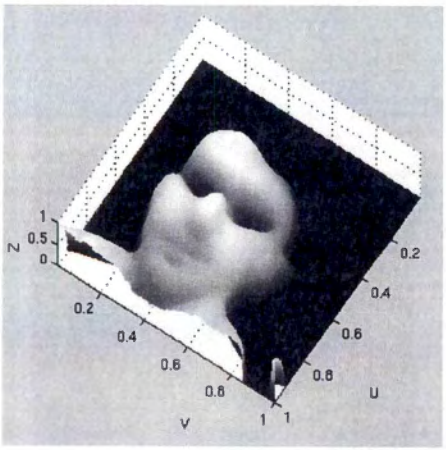
a



b



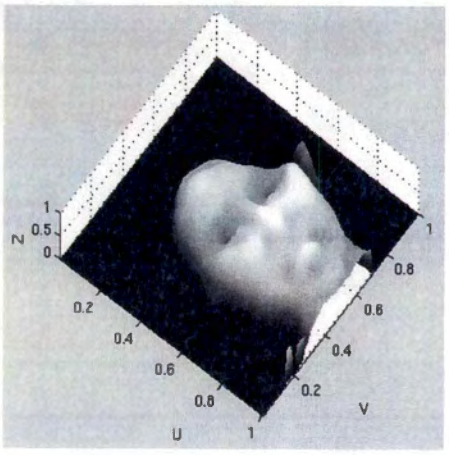
c



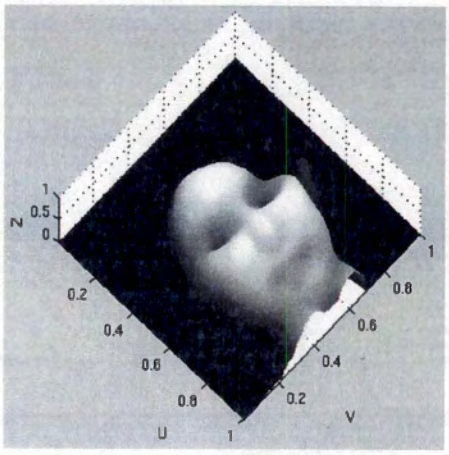
d

Figure 4.11: *Shape reconstruction of a human face: (a) the input image (Stephen Jesse's face). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

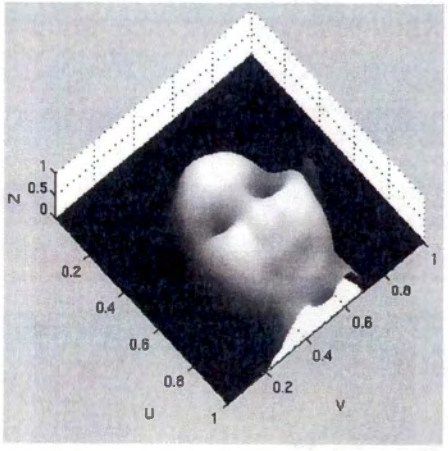
a



b



c



d

Figure 4.12: *Shape reconstruction of a human face: (a) the input image (Mengwei Li's face). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

a

b

c

d

Figure 4.13: *Shape reconstruction of a human face: (a) the input image (Mengwei Li's face). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

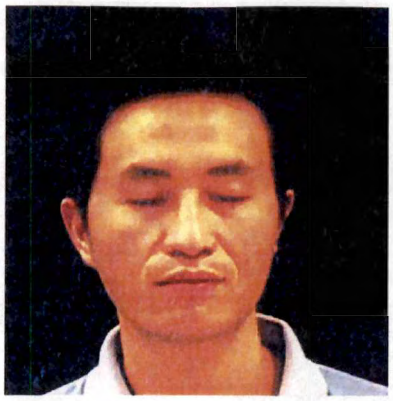
a

b

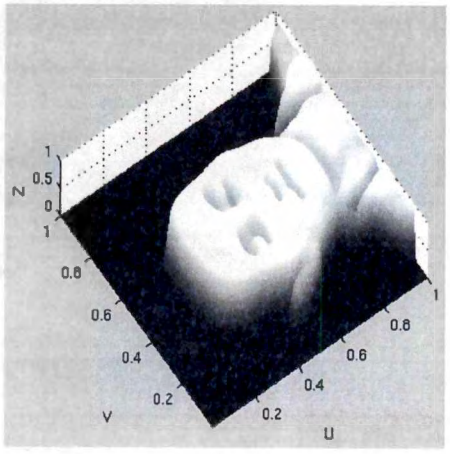
c

d

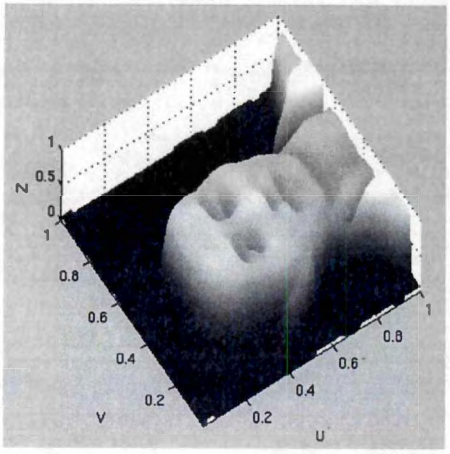
Figure 4.14: *Shape reconstruction of face: (a) is a real image of a human face (Kaiyn Wang). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*



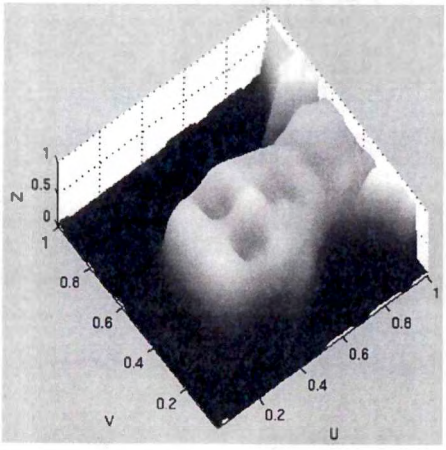
a



b



c



d

Figure 4.15: *Shape reconstruction of face: (a) is a real image of a human face (Zhong Du). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

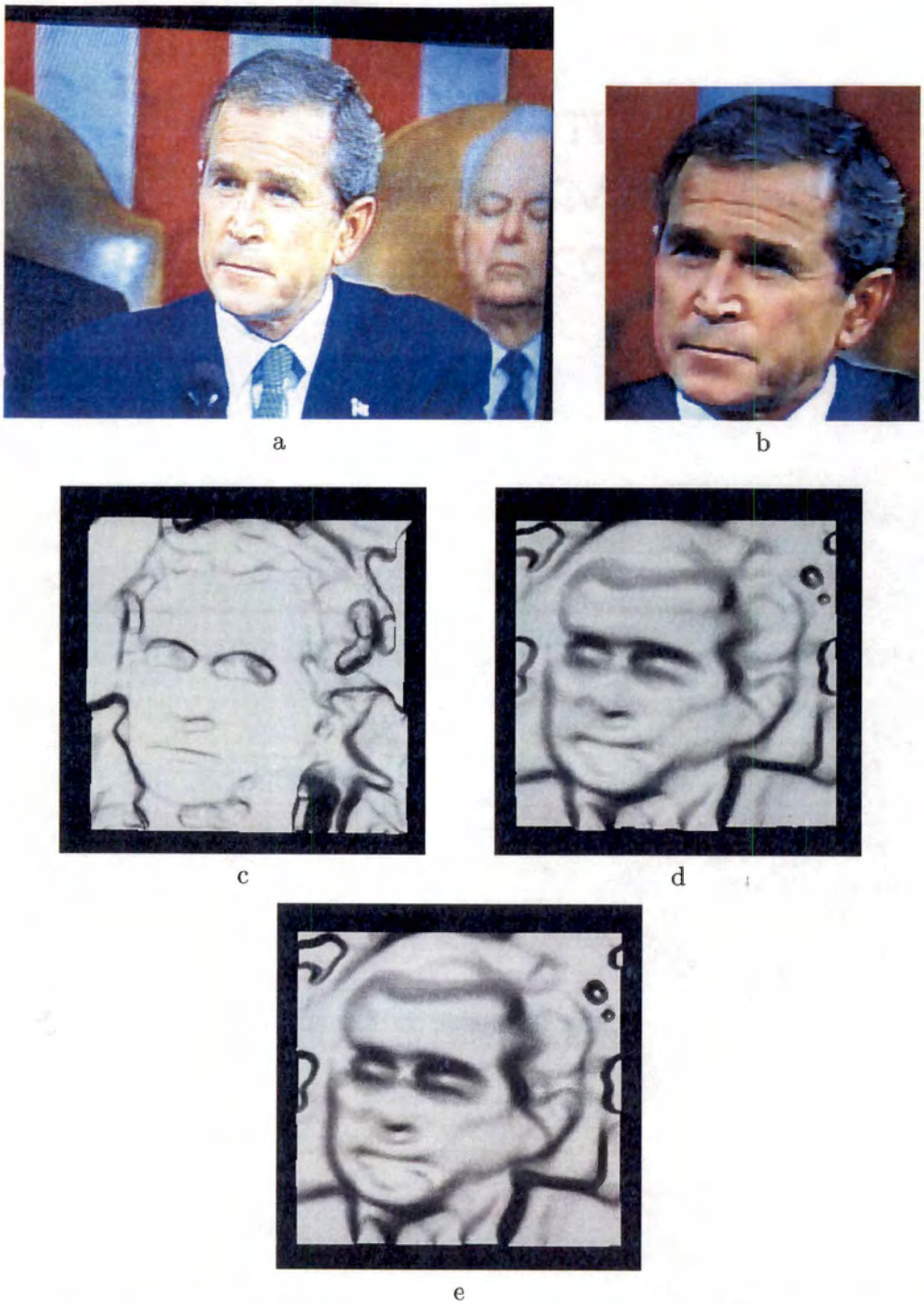
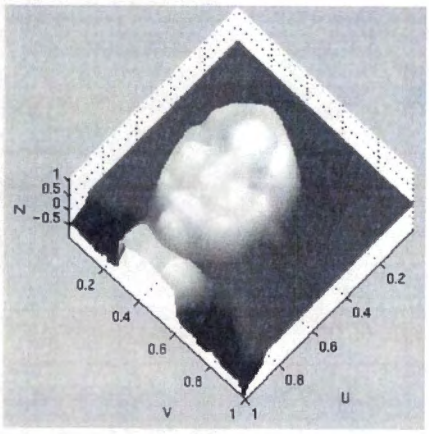
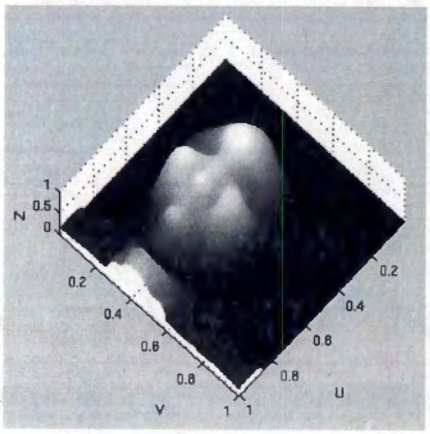


Figure 4.16: *Shape reconstruction of face: (a) is President Bush captured from TV during his address to the nation on Thursday September 20, 2001 after the World Trade Center Bombing. (b) is the cropped and pre-processed image. Results of the shape reconstruction by the three techniques are compared in (c) through (e) using (c) Pre-segmentation approach (d) Normal propagation approach and (e) Variational approach.*

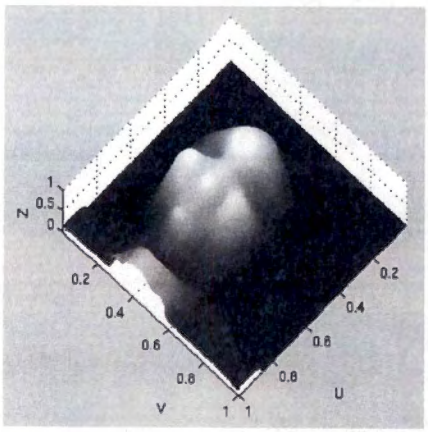
a



b



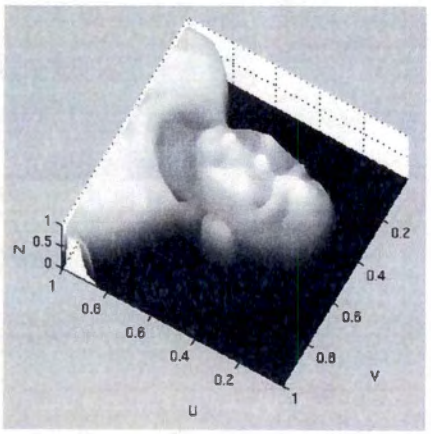
c



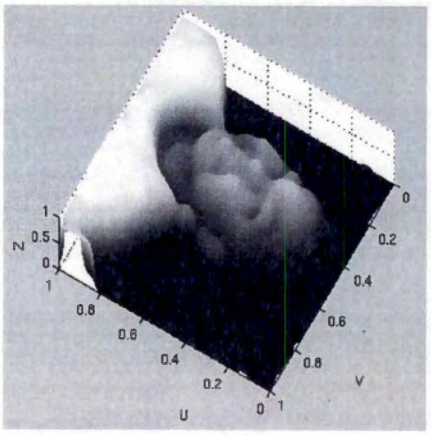
d

Figure 4.17: *Shape reconstruction of face: (a) is a real image of a human face (Lawretta Ononye, author's wife). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

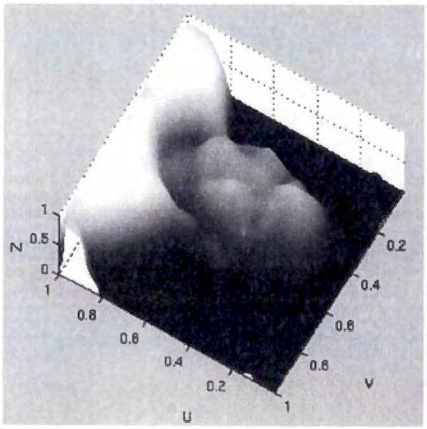
a



b



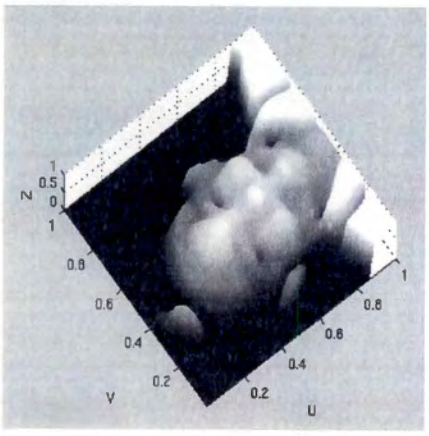
c



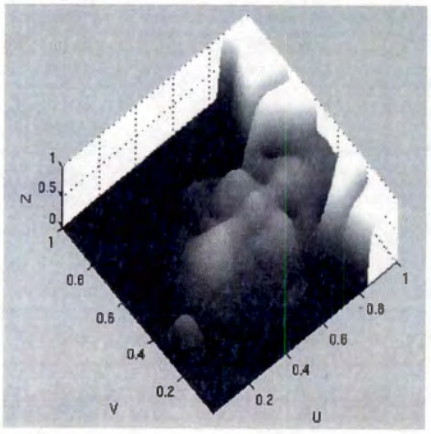
d

Figure 4.18: *Shape reconstruction of face: (a) is a real image of a human face (Ifechukwu Ononye, author's son). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

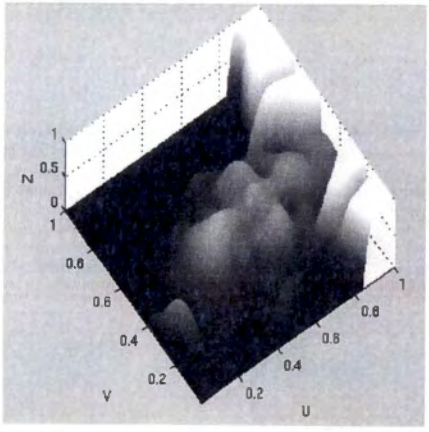
a



b



c



d

Figure 4.19: *Shape reconstruction of face: (a) is a real image of a human face (Nnaebuka Ononye, author's son). Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

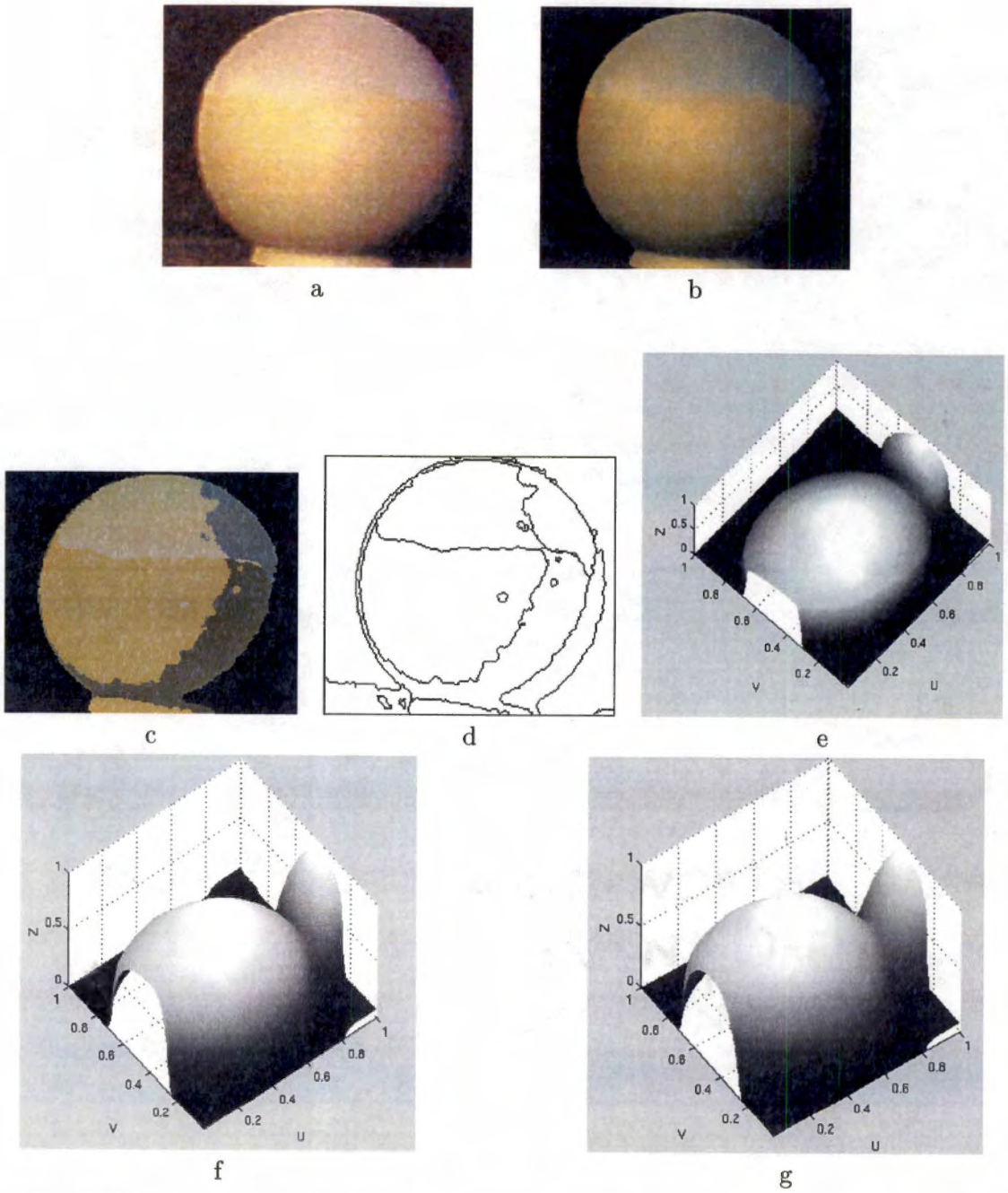


Figure 4.20: *Shape reconstruction of sphere - poor segmentation case: (a) is real image of a multi-color spherical object (b) pre-processed image by smoothing (c) is the segmentation result. (d) shows the extracted edges for color boundary identification. Results of the shape reconstruction by the three techniques are compared in (e) through (g) using (e) Pre-segmentation approach (f) Normal propagation approach and (g) Variational approach.*

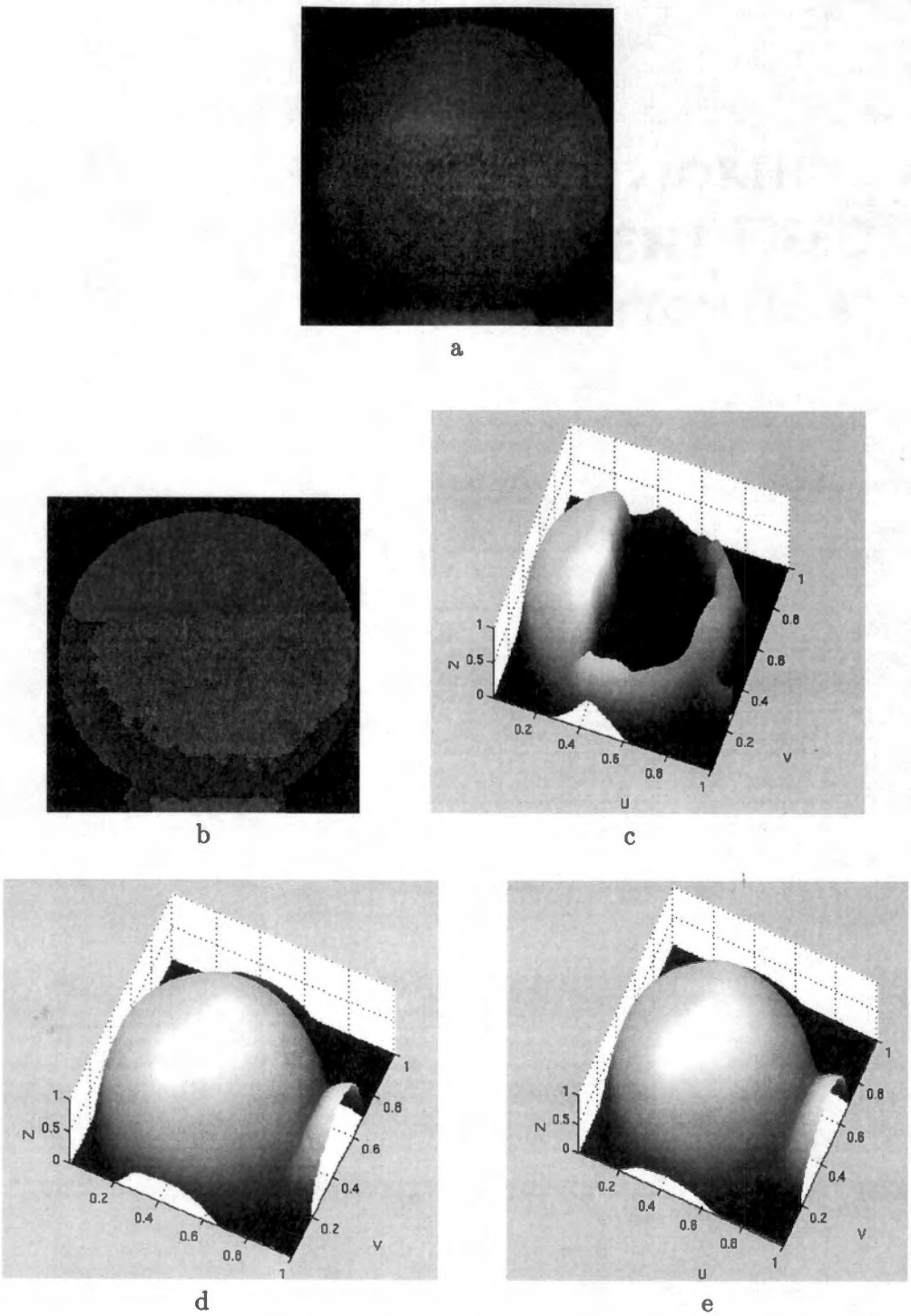
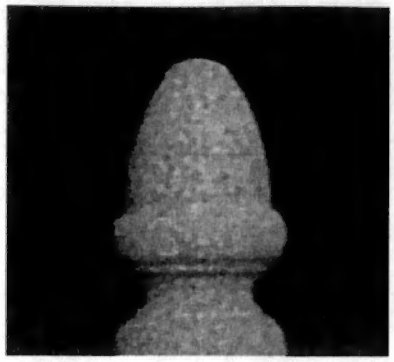
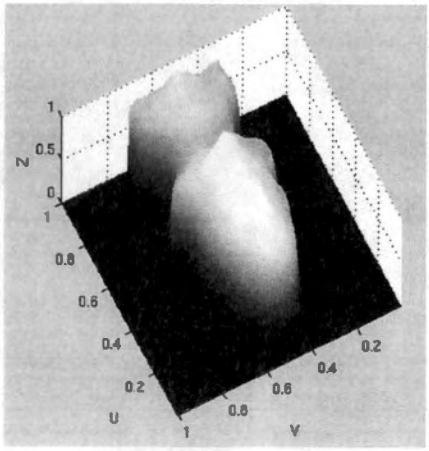


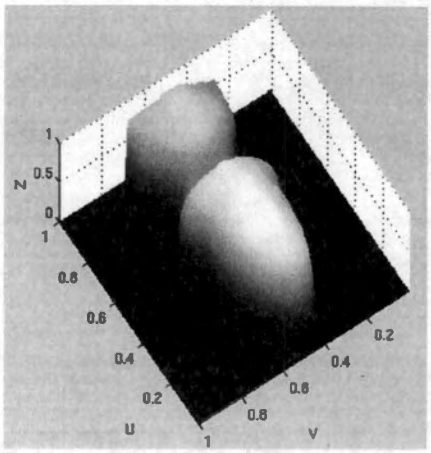
Figure 4.21: *Shape reconstruction of sphere - failure of pre-segmentation approach: (a)is real image of a multi-color(three distinct colors) of a spherical object and (b) is the segmentation result. Results of the shape reconstruction by the three techniques are compared in (c) through (e) using (c) Pre-segmentation approach (d) Normal propagation approach and (e) Variational approach.*



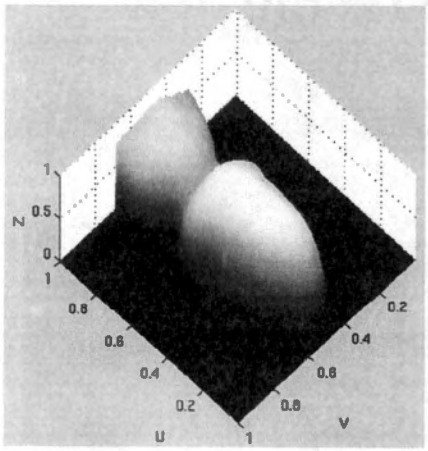
a



b



c



d

Figure 4.22: *Shape reconstruction of a textured real image: (a) is a textured real image. Results of the shape reconstruction by the three techniques are compared in (b) through (d) using (b) Pre-segmentation approach (c) Normal propagation approach and (d) Variational approach.*

4.5.1 Effect of Noise on Shape Reconstruction

The robustness of the normal propagation and variational algorithms to noise was tested by adding various levels of salt and pepper noise to the synthetic image shown in figure 4.23a. The noise densities, $\sigma = 0.1, 0.3$ and 0.5 were applied to the image to construct noisy examples. The time, t taken for each algorithm, as well as the number of iterations, n required (for the case of variational algorithm) to process were determined. The estimated shapes for each of the tests are shown in figures 4.23 and 4.24. It is qualitatively apparent that the variational technique provides better shape reconstruction in the presence of significant noise as would be expected. These processing time and number of iterations for the variational approach also increased with increasing levels of noise.

4.5.2 Quantitative Evaluation of Reconstructed Shapes

Unfortunately, the development of shape metrics that accurately mimic a human's interpretation of similar surfaces is still an open research problem. Often numerical shape validations are misleading and therefore the best validation tool for shape reconstruction might be visual observations. However, if the input data is limited to simple shapes, then numerical measures of shape similarity often correlate well with human observations.

To provide a quantitative measure of the ability of the three algorithms to reconstruct the shape of multi-color objects, two image sets were constructed to which quadratic equations could be fit to the reconstructed surface maps, $z(x, y)$, using least squares techniques[46, 47, 48]. The first set of images contained cylinders of varying textures while the second was constructed using spheres.

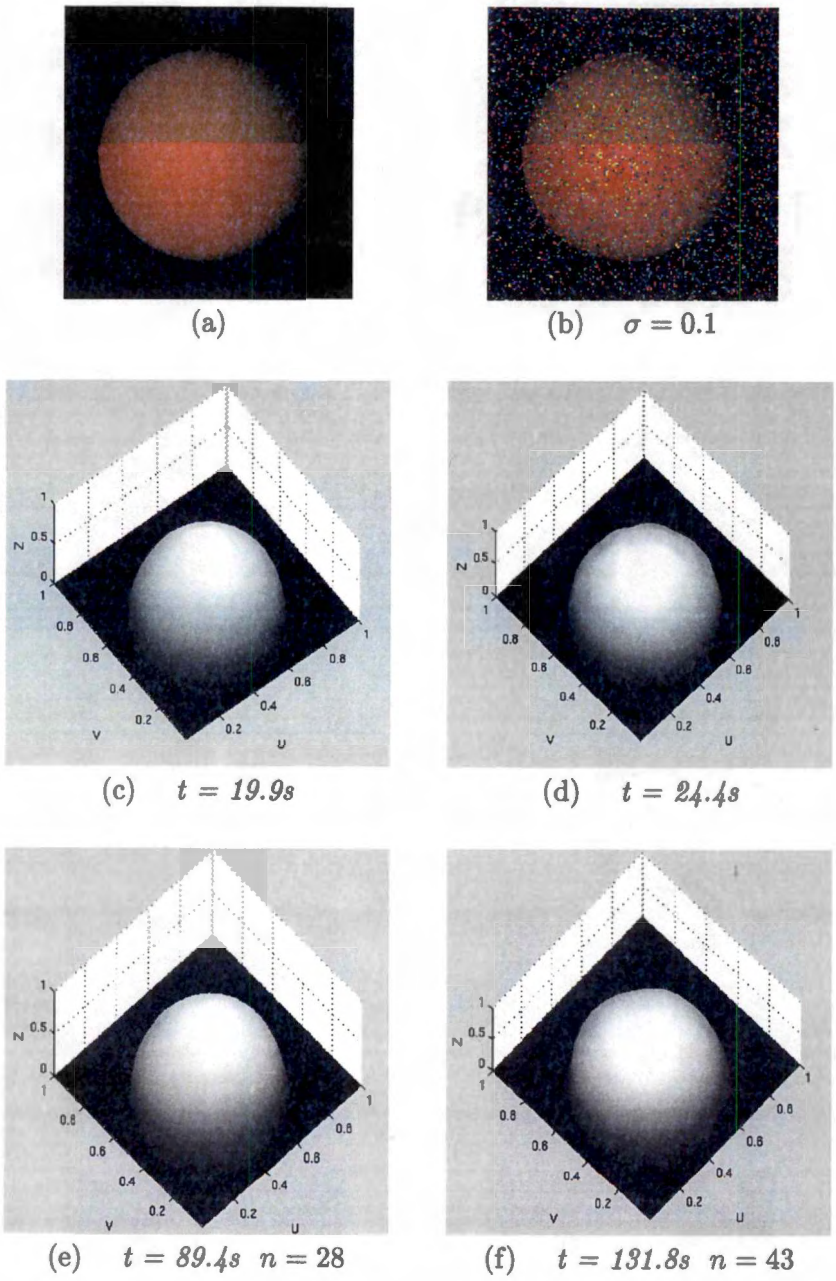


Figure 4.23: *Effect of noise on shape reconstruction: (a) is a noiseless synthetic image (b) generated by adding "salt and pepper" noise density of 0.1. Results of the shape reconstruction by Normal propagation approach are shown in (c and d) and the Variational approach are shown in (e and f).*

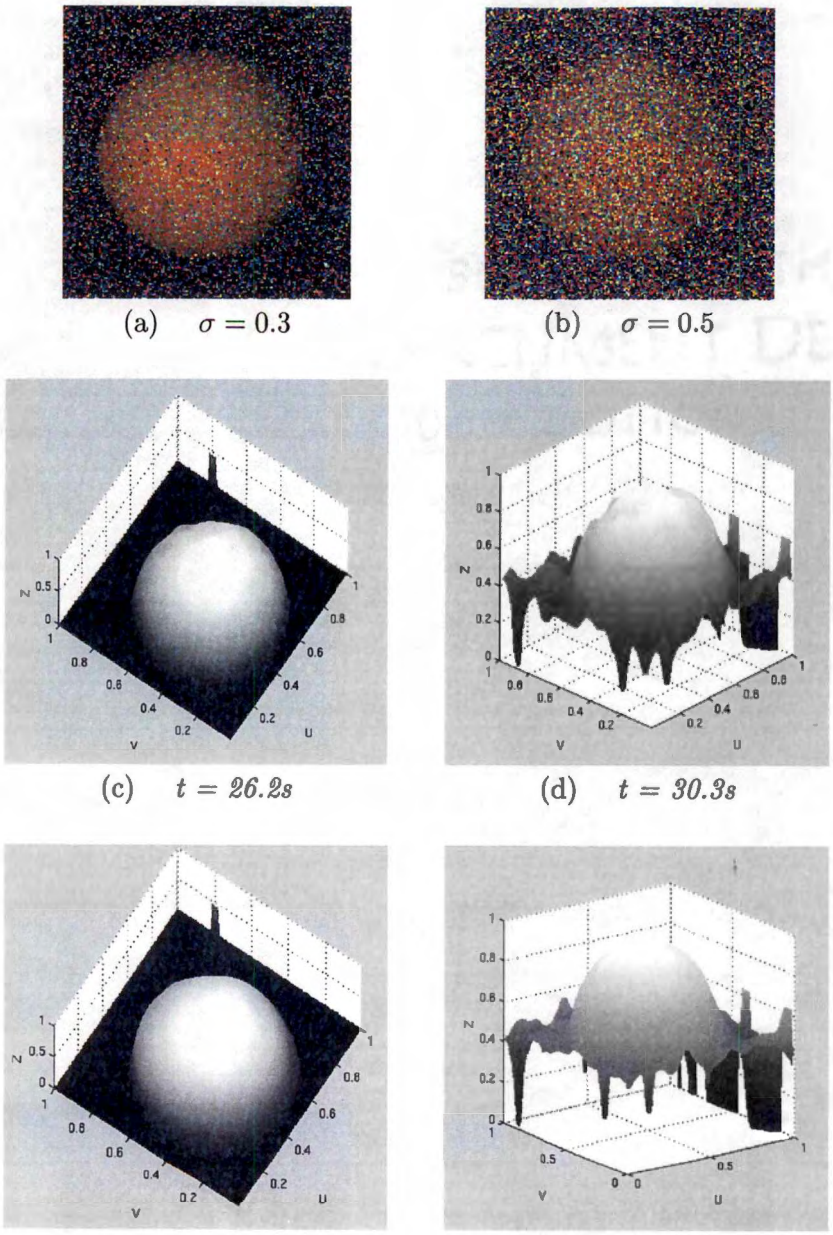


Figure 4.24: *Effect of noise on shape reconstruction: (a) and (b) are noisy images of figure 4.23a. The noise densities are 0.3 and 0.5 respectively. Results of the shape reconstruction by Normal propagation approach are shown in (c and d) and the Variational approach are shown in (e and f).*

The functional representation of the recovered surface is to be determined by the quadratic

$$f_i(x,y) = a_0x_i^2 + a_1y_i^2 + a_2x_iy_i + a_3x_i + a_4y_i + a_5 \tag{4.22}$$

for some choice of a_i, $i = 0, 1, ...5$

At each data point, the difference between the surface elevation, $z(x,y)$ and $f_i(x,y)$ defines the error. Thus, the coefficients a_i must be chosen such that the mean square error δ is minimized, where

$$\delta = \frac{1}{N}\sum_{i=0}^{N}\Big(z_i(x,y) - f_i(x,y)\Big)^2 \tag{4.23}$$

and N is the number of data points. The function δ is a minimum when

$$\frac{\partial\delta}{\partial a_i} = 0 \qquad\qquad i = 0, 1, ...5 \tag{4.24}$$

To solve the above minimization problem, the entities $\mathbf{X} \in \mathcal{R}^{N\times 6}$, $\vec{z}\ in\mathcal{R}^{N\times 1}$, and $\vec{a} \in \mathcal{R}^{6\times 1}$ are constructed, where

$$\mathbf{X} = \begin{bmatrix} x_0^2 & y_0^2 & x_0y_0 & x_0 & y_0 & 1.0 \\ x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 & 1.0 \\ x_2^2 & y_2^2 & x_2y_2 & x_2 & y_2 & 1.0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_N^2 & y_N^2 & x_Ny_N & x_N & y_N & 1.0 \end{bmatrix} \tag{4.25}$$

$$\vec{z} = \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \qquad \vec{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}, \tag{4.26}$$

and the expression

$$\mathbf{X}\vec{a} = \vec{z}$$

is solved for the quadratic coefficients a such that

$$\begin{aligned} \mathbf{X}^T\mathbf{X}\mathbf{a} &= \mathbf{X}^T\mathbf{z} \\ \Longrightarrow \mathbf{a} &= \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{z} \end{aligned} \tag{4.27}$$

Figures 4.25 and 4.26 show quadratic surfaces recovered using the above procedure.

The two test image sets of cylinders and spheres used for the evaluation are shown in tables 4.1 and 4.2 respectively. The LMS error (in pixel units) between the best quadratic fit and the actual recovered shape are also shown in the tables for each of the three algorithms. The techniques were also evaluated when salt and pepper noise of density 0.3 was added to each of the test image sets. The results are equally shown in tables 4.1 and 4.2.

4.6 Conclusion

In this chapter, three approaches to recovering the shape of multi-color objects from color images were developed and experimental evidence of their efficacy was presented.

The first, or pre-segmentation, algorithm works well for surfaces with distinct regions of constant surface reflectance which are by definition easier to segment. Its inadequacies become apparent for surfaces with either slowly varying or high frequency color textures

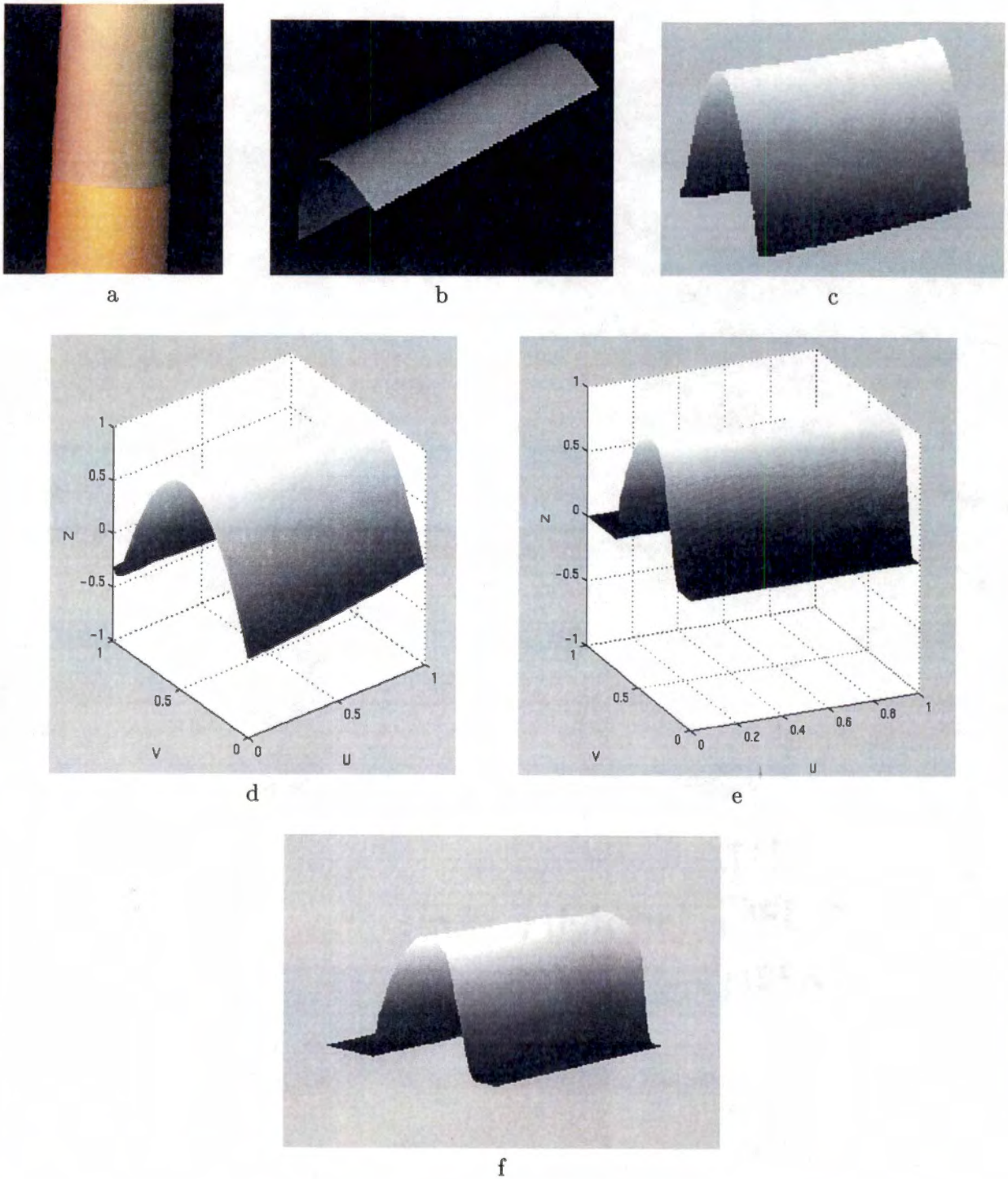
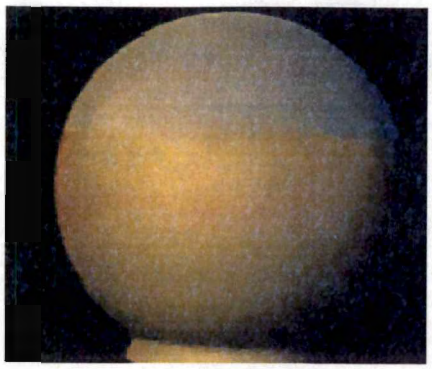
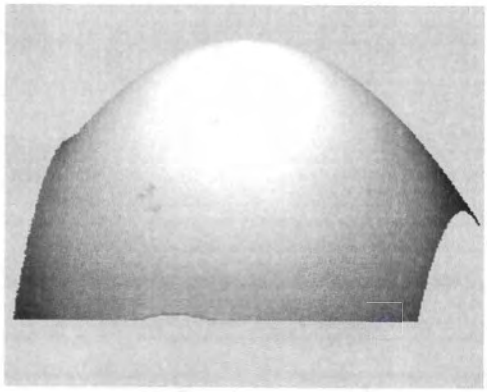


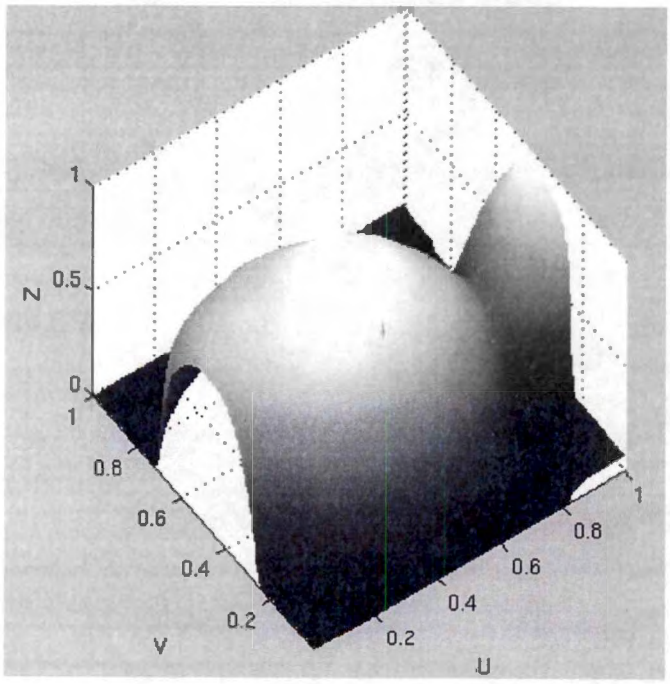
Figure 4.25: *Evaluation of reconstructed shape of a multi-color cylinder: (a) real input image. (b) - (d) are different views of the shape generated by quadratic surface fitting while (e) and (f) are the ones obtained directly from our technique.*



a



b



c

Figure 4.26: *Evaluation of reconstructed shape of a multi-color spherical object: (a) real input image. (b) shape generated using quadratic surface fitting and (c) the reconstructed shape by variational approach.*

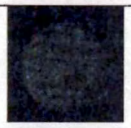
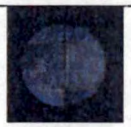
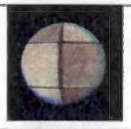
Table 4.1: *Quantitative evaluation of the techniques using textured cylindrical test image sets.*

PS = *Pre-segmentation* **NP** = *Normal Propagation* **V** = *Variational*
WON = *Test image with no noise* **WN** = *Test image with noise*

Test image	**PS Approach**	**NP Approach**		**V Approach**	
	WON	WON	WN $\sigma = 0.3$	WON	WN $\sigma = 0.3$
	0.0594	0.0564	0.0633	0.0644	0.0647
	0.00825	0.1152	0.1257	0.0799	0.0817
	0.1952	0.2141	0.2158	0.01858	0.1904
	0.1952	0.1351	0.1688	0.1327	0.1656
	0.1239	0.1930	0.1983	0.1697	0.1787
	0.0543	0.0681	0.0871	0.0530	0.0651
	0.0737	0.1067	0.0994	0.0732	0.0717
	0.0737	0.0861	0.0968	0.2015	0.2189
	0.0641	0.0821	0.0912	0.0719	0.0754
	0.0450	0.0545	0.0632	0.0553	0.0587
Average Error	0.0967	0.1111	0.1210	0.1087	0.1171

Table 4.2: *Quantitative evaluation of the techniques using textured spherical test image sets.*

PS = *Pre-segmentation* **NP** = *Normal Propagation* **V** = *Variational*
WON = *Test image with no noise* **WN** = *Test image with noise*

Test image	**PS Approach**	**NP Approach**		**V Approach**	
	WON	WON	WN $\sigma = 0.3$	WON	WN $\sigma = 0.3$
	0.0403	0.0637	0.0717	0.0591	0.0562
	0.0572	0.0499	0.0951	0.0436	0.0563
	0.0448	0.0695	0.0844	0.0660	0.0557
	0.0528	0.0412	0.0788	0.0549	0.0749
	0.0763	0.0558	0.0987	0.0858	0.0718
	0.0755	0.0555	0.1079	0.0789	0.0817
	0.1309	0.0451	0.0824	0.0544	0.0636
	0.0676	0.0498	0.0684	0.0671	0.0475
	0.0676	0.0487	0.0863	0.0644	0.0620
	0.0577	0.0454	0.0850	0.0585	0.0623
Average Error	0.0671	0.0525	0.0858	0.0636	0.0632

as could be seen from figures 4.7, 4.8 and 4.20. Hence, the performance of this algorithm is limited to that of the segmentation scheme used. The use of this technique is therefore recommended for surfaces with distinct regions of constant surface reflectance. The pre-segmentation approach failed completely when salt and pepper noise of density 0.3 was added to the image sets. This is because the segmentation could not find a consistent region of constant surface reflectance. Therefore, the performance of this technique could not be quantitatively compared with the other two techniques under this level of noise.

In the second, or normal propagation, technique, the short-coming of pre-segmentation is avoided. It worked quite well in the cases where the pre-segmentation failed. Its major set back is its sensitivity to high degrees of image noise.

The third, or variational, approach uses the global smoothness constraint to iteratively solve for the optimal object surface. While it often provides the best visual result, it is also more robust to noise than the normal propagation as evident from figures 4.23 and 4.24 as well as in tables 4.1 and 4.2, but also requires more computation time.

CHAPTER 5

Shadow Detection and Removal

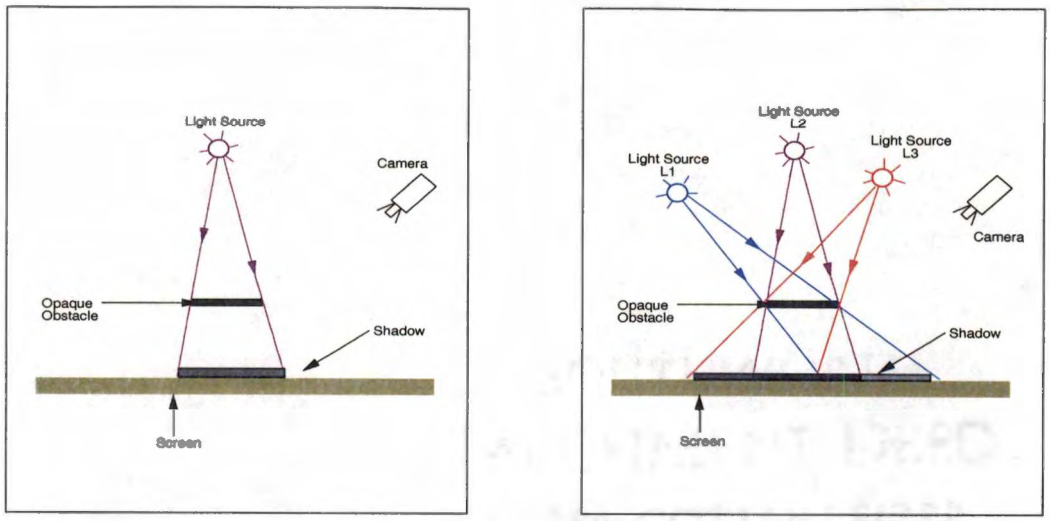
5.1 Introduction

From theory of classical physics, light propagates in straight lines. When the rays of light encounter an obstacle, they graze through the edges of the obstacle and consequently, cast a sharp shadow on the screen as shown in figure 5.1 below. For an extended light source, the shadow consists of two parts, the darker region called the umbra surrounded by a slowly darkening region, the penumbra.

Shadows pose a similar problem to the shape estimation process as non-constant surface reflectance. Reconsider the apparent color to surface normal mapping of equation 4.1

$$\vec{n}_s(x,y) = \mathbf{A}^{-1}(\lambda)\mathbf{S}^{-1}(x,y;\lambda)\vec{E}(x,y;\lambda).$$

In the presence of shadows, the illuminant matrix, $\mathbf{A}$, becomes a function of image coordinates (x,y), violating the assumption of constant illumination across the scene. Unfortunately, apparent color variations caused by shadows mimic those produced by continuous changes in surface shape due to the physics of shadow formation. Therefore, shadows must be identified and eliminated before the surface shape can be reconstructed using the smoothness constrained algorithms of the previous chapter, regardless of whether or not the surface reflectance is constant. In this chapter, a novel technique is presented which employs simple segmentation procedures along with color normalization to both identify and remove shadows without requiring a linear camera or assuming the darkest image regions are shadows.



a b

Figure 5.1: *Schematics of shadow formation: (a) is an illustration of shadow formation by a single illuminant. (b) illustrates physics of shadow formation by multiple illuminants.*

5.2 Previous Work in Shadow Detection

Shadow recognition in a scene by an artificial vision system is a difficult task and therefore a number of cues that suggest its presence have been employed in other efforts. The four most prominent cues are summarized below.

1. **Darkest Region Gambit** - the image region with the lowest average intensity is likely to represent pixels in shadow.

2. **Hue/Saturation Invariance** - the apparent hue and saturation of surfaces in and out of shadows remains constant [49].

3. **Illuminant Direction Dependence** - the shape, size, and position of shadows are directly dependent on scene illuminant directions [50].

4. **Surface Texture Invariance** - cues for recovering texture that do not depend on absolute intensity are preserved across shadow boundaries [51].

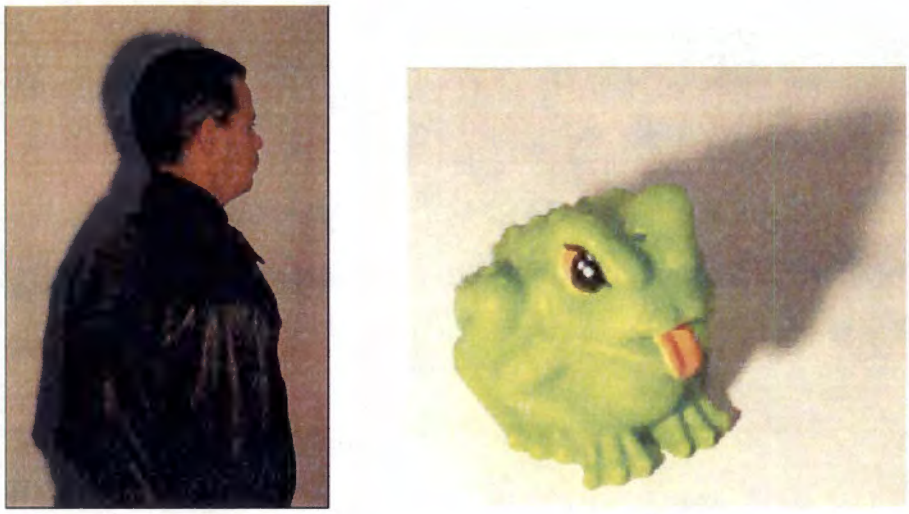
Due to the nature of imaging sensors, it is often safe to assume that the darkest regions of images are cast shadows. Thus, a number of approaches, primarily developed for gray-scale imagery, assume shadows are located in the areas of the image with the lowest intensities [52, 53]. This assumption does not always hold as could be seen from figures 5.2(a and b). In each of these figures, the darkest region does not lie in the shadow. In figure 5.2(a), the darkest region is the hair with an average intensity, $\bar{E}_o = 8$, while the average intensity of a point in the shadow region is $\bar{E}_s = 69$. A similar situation is shown in figure 5.2(b), where the darkest region is the eye of the frog, not the shadow.

The hue/saturation invariance property has been employed in various research efforts [49, 54, 52, 53] to eliminate shadows in scenes. Given an image pixel on a constant color surface outside the shadow, $\vec{E}_o(x, y)$, the invariance property states that any pixel on the same surface inside the shadow, $\vec{E}_s(x, y)$, is such that

$$\vec{E}_s(x, y) = \alpha \vec{E}_o(x, y)$$

where α is a constant such that $\alpha < 1$. Thus, shadows can be eliminated from the image via normalization. Unfortunately, image normalization removes apparent color changes that arise from surface shape variation as well. In addition techniques based on normalization provide only for removal, not identification, of pixels in shadow.

In his work on visual recognition of shadows, Funka-Lee [50] used an active observer equipped width an extendible probe for casting its own shadows on the scene. This allowed the observer to experimentally determine the number, location, and spectral content of the light sources present. The hue/saturation invariance property was then used to identify shadows in a unique way. Constrained by the information obtained regarding the illumination environment and assuming a sensor with linear gain, the color space of the image is searched for radial line features that result from the slowly darkening nature of shadow pixels in the penumbra. Pixels which fall along these lines are assumed to be



a b

Figure 5.2: *Illustration of break-down of darkest region theory: (a), is a man with a cast shadow on the wall. The darkest region corresponds to his hair with RGB distribution of about 8 0 0. The RGB of the shadow region is about 72 66 69. (b) shows a plastic frog with a cast shadow on the background. The darkest region corresponds to the eye and and not the shadow.*

in shadows. While this approach works well in both controlled and uncontrolled environments, many imaging sensors, including most off-the-shelf color cameras, have a non-linear gain component that distorts the expected radial lines, creating radial curves which are often undetected.

In the remainder of this chapter, a novel algorithm for shadow identification and removal is presented. Employing a simple variation on the hue/saturation invariance based method described above, this technique is able to identify and remove shadows from color images acquired using non-linear gain sensors without employing the often unrealistic darkest region assumption.

Before continuing, it should be noted that shadows may be classified as *cast* or *self*. A **cast shadow** arises whenever a free space exists between the shadow and the agent

obstructing the light source while in a **self shadow**, such a free space does not exist. This part of the dissertation focuses on cast shadows on backgrounds. The dissertation does not address self shadows in color imagery.

5.3 A Shadow Detection Algorithm Assuming Linear Camera Gain

The primary disadvantage of previously reported shadow removal techniques that employ the hue/saturation invariance property is that such algorithms only remove shadows from the image. In this section, a new method is presented which uses the hue/saturation properties of images with shadows, along with color-based image segmentation, to not only remove, but identify image pixels in shadowed regions.

The RGB is the feature space used. The input image is transformed from its image domain to the feature space using Mean Shift Algorithm (MSA) [55] described in appendix B.2. Pixels with similar RGB distribution cluster together and high density clusters represent prominent features in the image. The number of such clusters denotes the number of significant distinct color regions including the shadow(s) in the image. Since under typical illumination conditions, the normalized color of two points on any given surface of the same material will be the same even if one of these points is directly illuminated and the other is in a shadow. To further illustrate this, let (r_s, g_s, b_s) be the RGB values of the pixels in the shadow region and (r_i, g_i, b_i) be the ones at a point on the corresponding surface not in a shadow. If normalized color scheme is applied to these two distinct regions, the same normalized color result will be obtained provided the color space is linear. For our case, the input image is normalized and the same clustering algorithm applied to both normalized and unnormalized images. Let m_u and m_n be the number of clusters corresponding to the unnormalized and normalized images respectively. If $m_u = m_n$, it suggests there is no shadow in the scene. If m_n is less than m_u, then the difference m_u

- m_n shadow clusters have vanished as a result of the process. The shadow clusters have consequently merged with other color clusters with identical normalized color. To detect which of the two clusters belongs to the shadow, we examine the two clusters that merged. The one whose cluster center has a lower RGB values belongs to the shadow. Figure 5.3a is a synthetic input image with shadow and 5.3bs are the normalized images. The clusters corresponding to figures 5.3a and 5.3b are respectively shown in figures 5.3d and 5.3e. The cluster labeled 3 which corresponds to the shadow which is completely eliminated after normalization. Figure 5.3b2 is a post-processed image of figure 5.3b1.

The realism of the vanishing clusters after normalization holds if we have linear color space. In other words, the above technique works good if the camera is linear. In the next section, we will discuss the case where the camera is not linear.

5.4 Shadow Detection in a Non-Linear Color Space

If the color space is not linear, the shadow detection technique described above will perform below par. In this case, the color space and of course the camera will have to be linearized by doing some gamma correction (see appendix A.3). Even with gamma correction, normalization may not cause the shadow to vanish completely (see figure 5.8b) as discussed in the previous section. When this is the case, we find the normalized clusters that are closest to each other and form pairs of such clusters, called the **shadow candidate pairs**. The formation of these pairs may result to some invalid ones. The invalid shadow candidate pairs can be eliminated by a user defined Euclidean distance measure constraint, ϵ. Clusters are said to have merged if ϵ is zero or close to being zero. Hence, any shadow candidate pair whose Euclidean distance measure is less than ϵ is regarded as a valid pair. For any valid shadow candidate pair with clusters, $\mathbf{C}_1$ and $\mathbf{C}_2$, the darker of the two is the shadow and by keeping track of the pixels that reside in such cluster, the shadow image

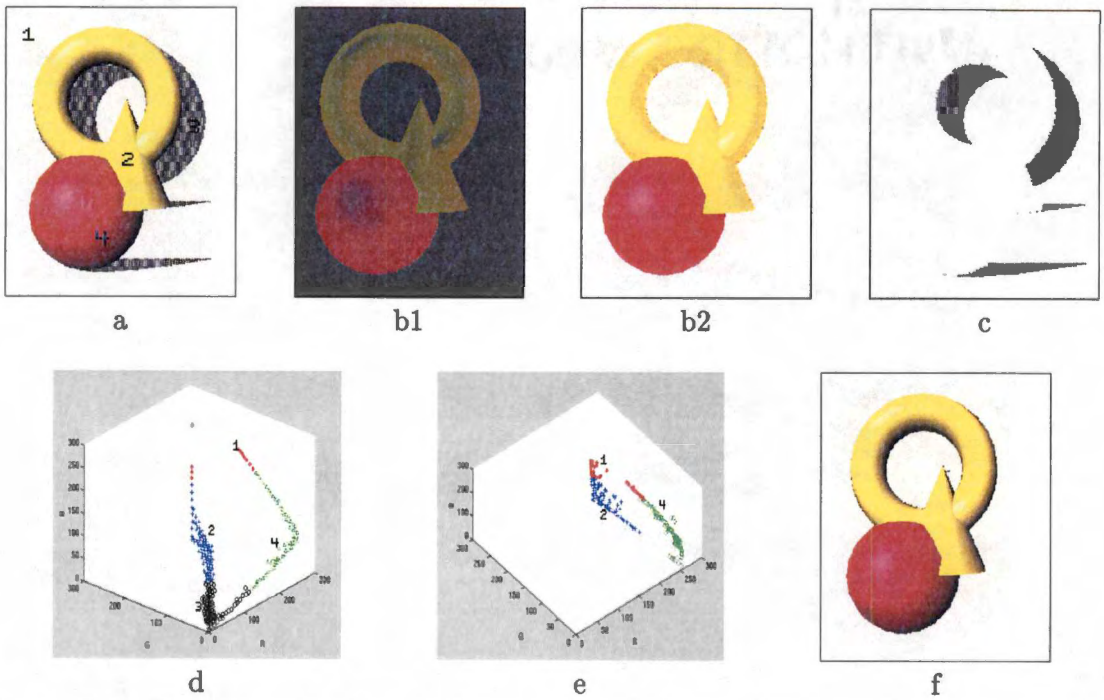


Figure 5.3: *Shadow extraction and removal procedure: (a) is synthetic image with shadows. (b1) is the result of the normalization and (b2) is the post-processed image of (b1). (c) is the extracted shadow and (f) is the restored image. (d and e) show the clusters corresponding to images in (a and b) above. The cluster labeled 3 corresponds to the shadow. (e) shows the clusters corresponding to (b). The shadow is gone after normalization*

could be extracted as shown in figures 5.3c, 5.4c and so on.

5.5 Shadow Removal

We have so far discussed how to recognize and extract shadows. The main issue to be addressed is how does one restore the given image? In other words, how is the shadow removed from the image so that it (image) looks as if it never had one (shadow)? This issue (of shadow removal) has not been addressed at length in literature. To address the above issue, we consider a pair of merging clusters $\mathbf{C}_1$ and $\mathbf{C}_2$ say. If $\mathbf{C}_2$ is the darker of the two clusters, it is therefore regarded as shadow. The corresponding pixels that are in the $\mathbf{C}_2$ are then mapped to the mean of $\mathbf{C}_1$. This type of mapping results to a noticeable edge effect as shown in figures 5.4e and 5.8d. The reason for this phenomenon is that real illuminants do not cast sharp shadows because they are not point sources and may only be partially obstructed. The umbra (which is the darker region of a shadow is due to complete obstruction of the light source) and the penumbra (which results from partial obstruction of illuminant) regions are shown in figure 5.5 as **DE** and **EG** respectively. Consequently, the transition from shadow to non-shadow is not a step function but a slowly varying function as shown in figure 5.5. In this figure, $\mathbf{M}_1$ and $\mathbf{M}_2$ are means of the shadow and non-shadow clusters that merged together. The decision boundary **B** tends to equalize the distance between the two means. After the preliminary mapping of shadow pixels to $\mathbf{M}_2$, the region to the left of the decision boundary, **B** gets mapped to $\mathbf{M}_2$ leaving the partial penumbra region, **FG** unmapped as illustrated in figure 5.6. As a final phase of the shadow removal, the pixels in the partial penumbra region **FG** are then mapped to the mean, $\mathbf{M}_2$. This eliminates the edge effect as shown in figure 5.4f.

5.6 Results of Shadow Detection and Removal

Further results of shadow detection and removal are shown in figures 5.4, 5.8 through 5.10 for single-color background. In figure 5.4, the input image is a real image of a synthetic frog captured with Kodak Zoom 120DC camera with a cast shadow on gray background. Figures 5.4b and 5.4c are the normalized and extracted shadow images respectively. The background is shown in figure 5.4d which is set to white while the main object in the image is set to black. The restored image is shown in figure 5.4e which exhibits some edge effects. This image is further processed to get rid of the edge effect due to the penumbra in the shadow. The result is shown in 5.4f. In figures 5.7 and 5.8, a human head and a man with a cast shadow on a wall are shown. The final restored images are shown in figures 5.7f and 5.8e respectively. Similar results are shown in figures 5.9 and 5.10. Figure 5.10a is the author of this dissertation with a cast shadow on a wall and the restored image is shown in figure 5.10e.

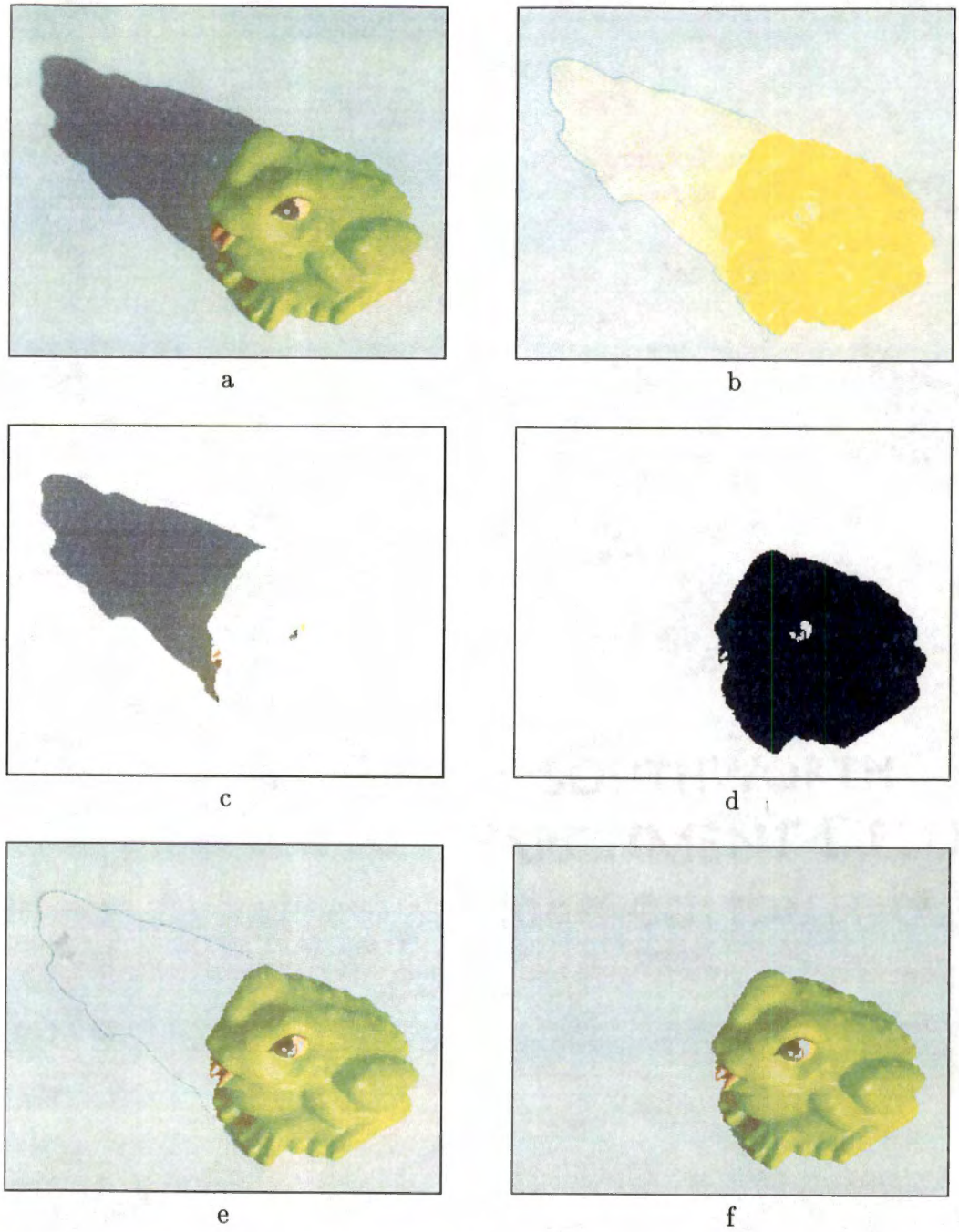


Figure 5.4: *Illustration of shadow extraction: (a) is real image of an artificial frog. (b) is the normalized image (c) shows the extracted shadow and the background is shown in (d). (e) is the restored image with edge effects due to penumbra and (f) is the final image after post-processing.*

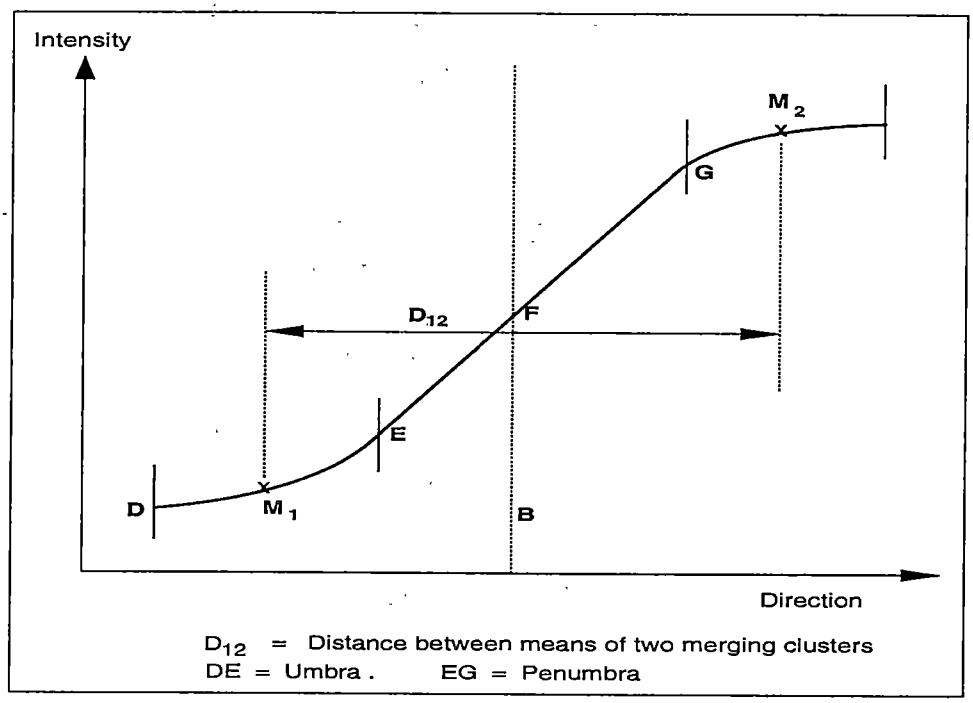


Figure 5.5: *Transition from shadow to non-shadow.* M_1 *is the mean of the shadow cluster*
M_2 *is the mean of the non-shadow cluster which the shadow cluster merged with*
B *is the boundary between* M_1 *and* M_2

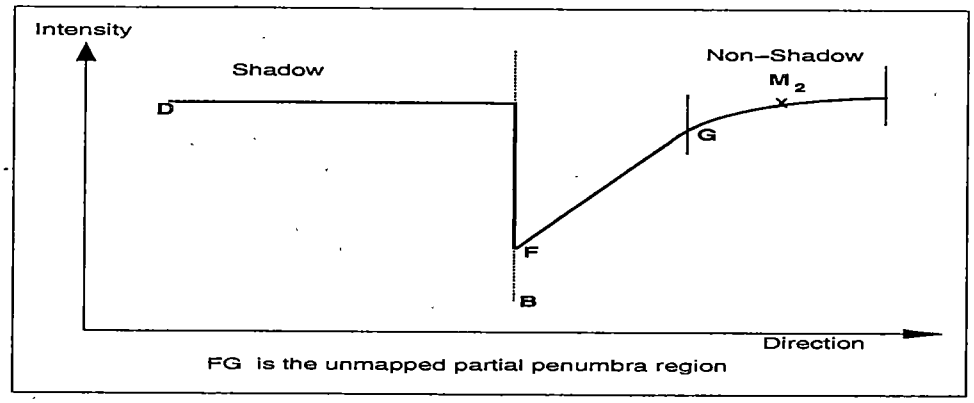


Figure 5.6: *Mapping of shadow to the corresponding non-shadow mean: The result of mapping the shadow pixels in region* DE *and* EF *shown in figure 5.5 to the mean,* M_2. *The edge effect is due the partial penumbra region* FG *not mapped to* M_2.

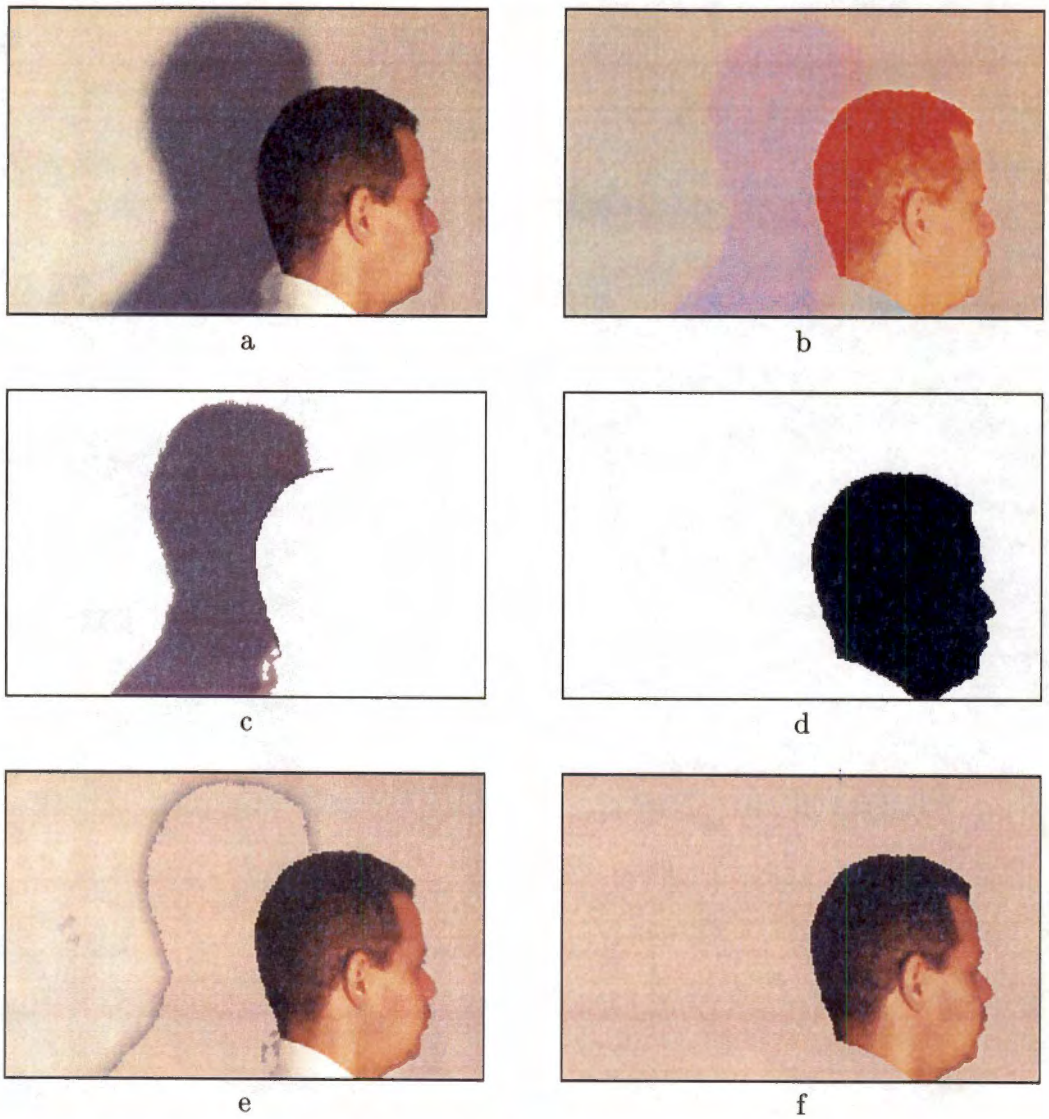


Figure 5.7: *Extraction of shadow cast by a human head: (a) is real image of a human head with a cast shadow on a wall. (b) is the normalized image (c) shows the extracted shadow and the background is shown in (d). (e) is the restored image with edge effects due to penumbra and (f) is the final image after post-processing.*

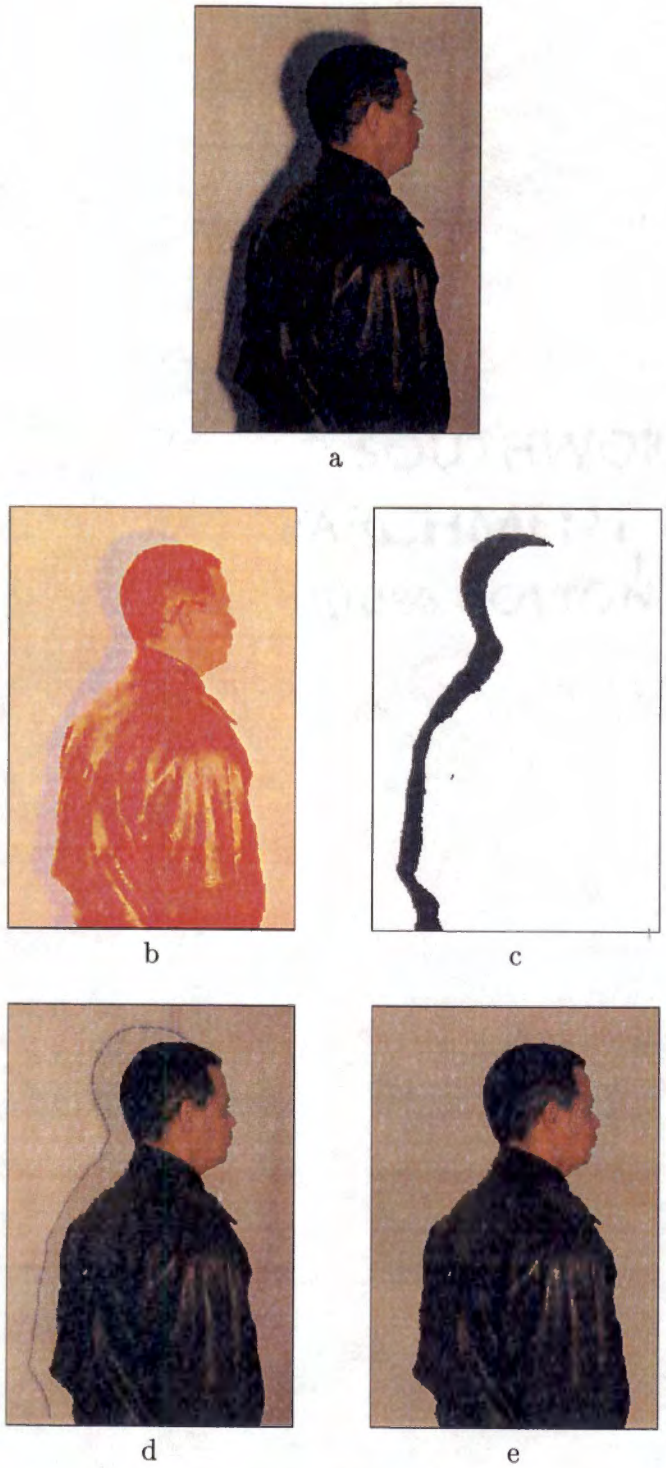


Figure 5.8: *Extraction of cast shadow on wall: (a) is real image of a man with a cast shadow on a wall. (b) is the normalized image (c) shows the extracted shadow). (d) is the restored image with edge effects due to penumbra and (e) is the final image after post-processing.*

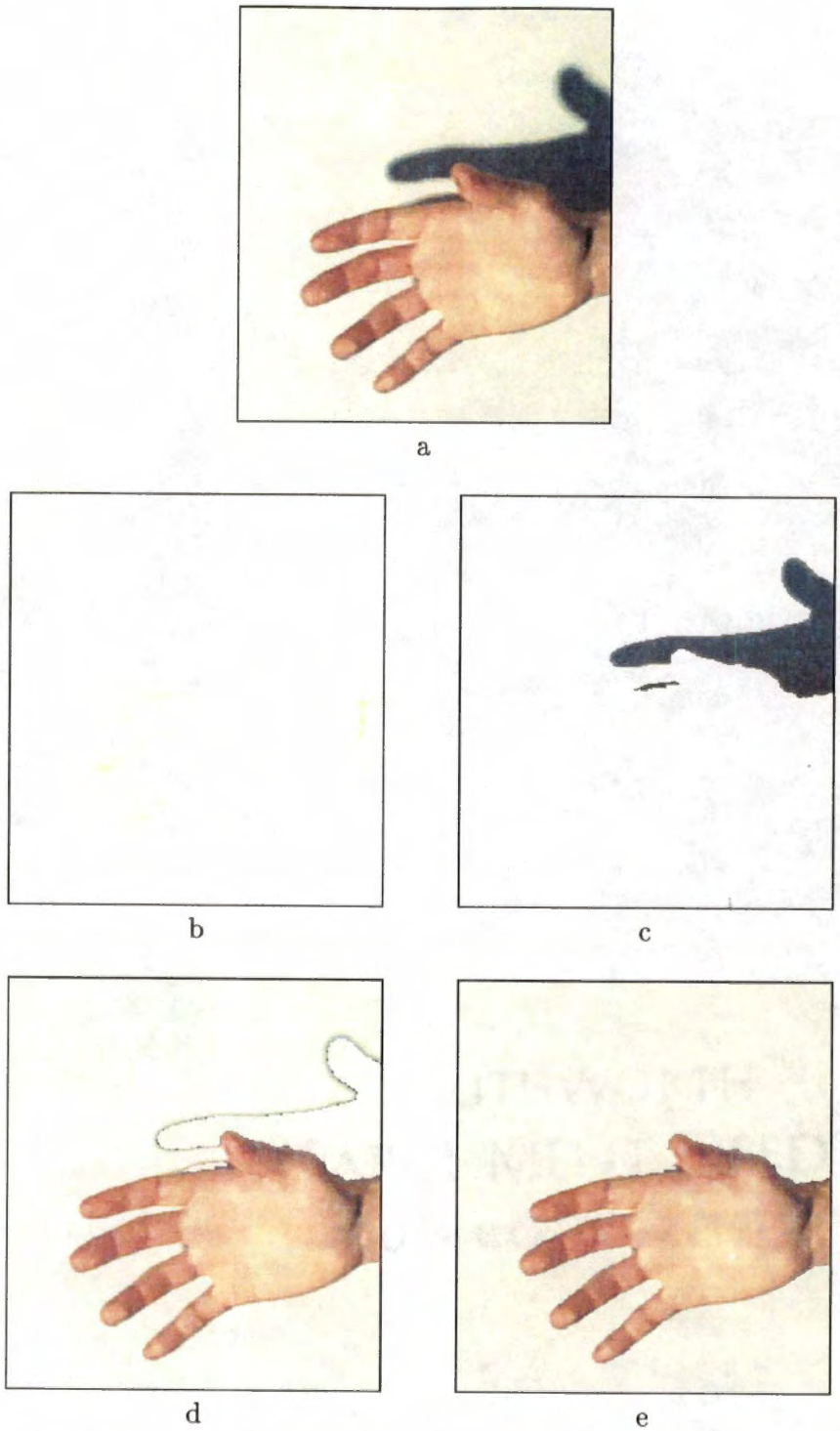


Figure 5.9: *Extraction of shadow cast by human fingers: (a) is the input real image. (b) is the normalized image (c) shows the extracted shadow). (d) is the restored image with edge effects due to penumbra and (e) is the final image after post-processing.*

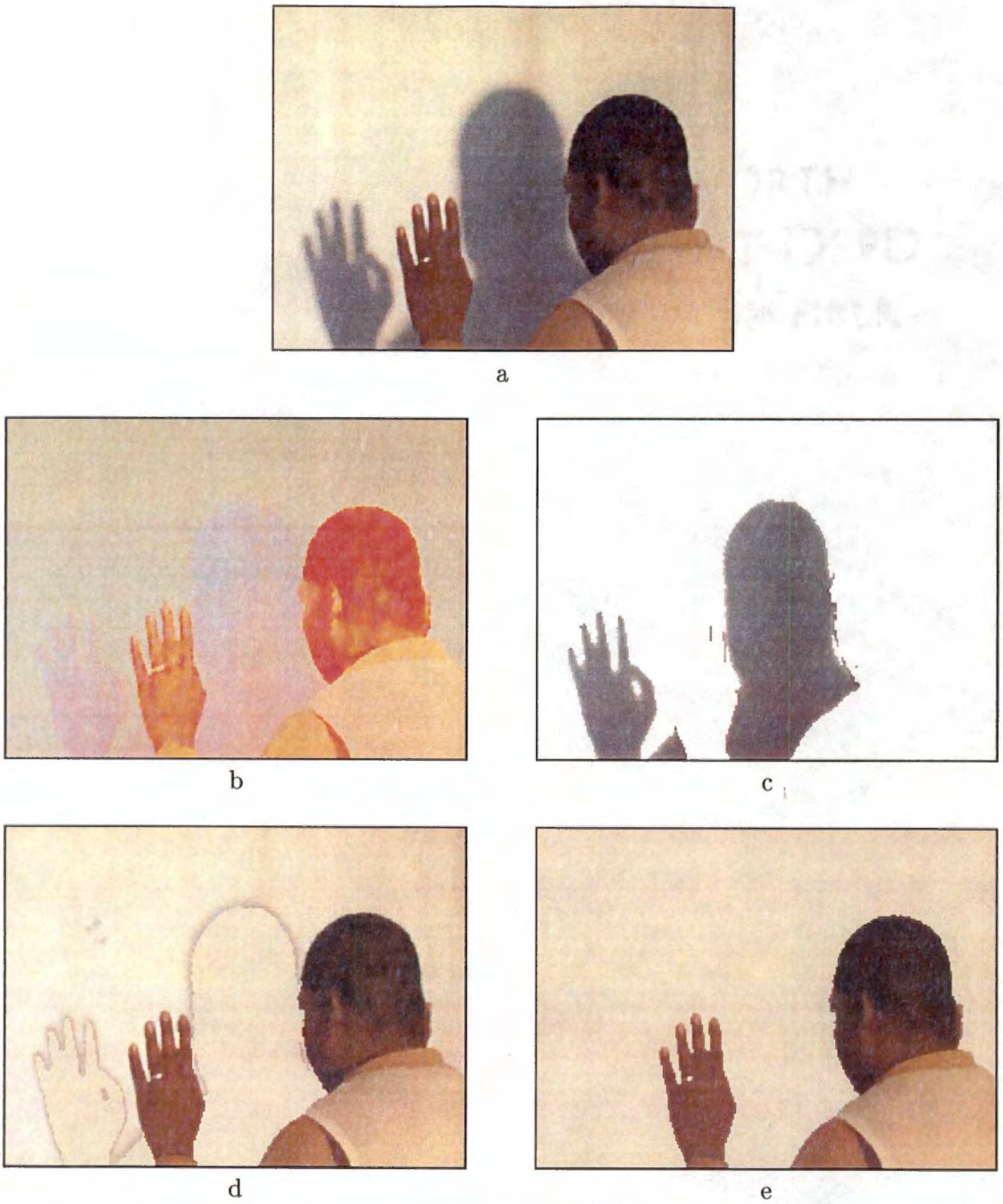


Figure 5.10: *Extraction of shadow cast on a wall: (a) is real image of the author of the dissertation with a cast shadow on a wall. (b) is the normalized image (c) shows the extracted shadow). (d) is the restored image with edge effects due to penumbra and (e) is the final image after post-processing.*

5.7 Shadow on Multi-Color Background

When the cast shadow drops on a multi-color background such as the case illustrated in figures 5.11 below, its detection and removal become more challenging especially if the material properties of any given surface of the background is similar to the spectral properties of any shadow segment. The procedure is similar to the case of single-color background. The main difference is the existence of multiple shadow and background clusters. Once again, the input image is normalized and then mapped to the RGB feature space. The next step is to pair clusters that are closest to each other using the user defined Euclidean distance measure constraint. For example using figure 5.11 below, after the normalization, the shadow cluster corresponding to **S1** will be paired with the background labeled **B1**. Similarily, **S2** will be paired with background **B2**. There is no doubt that false pairing of clusters will exist but the Euclidean distance measure constraint will help to eliminate the most of the false ones.

Results of the shadow detection and removal are shown in figures 5.12 and 5.13 below. Figure 5.12 has a two-color background while figure 5.13 has a three-color background.

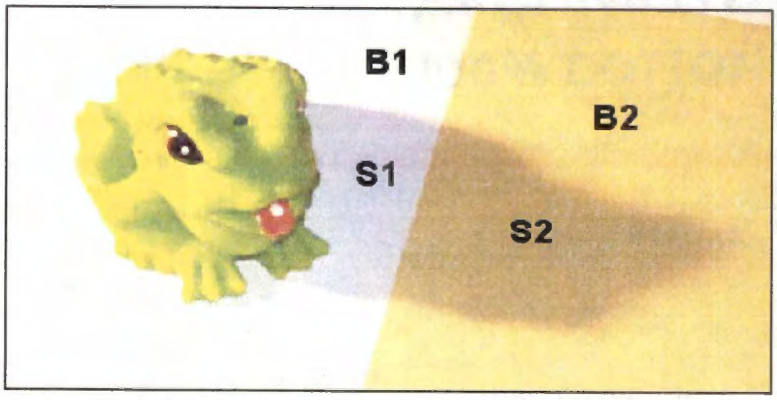


Figure 5.11: *Transition from shadow to non-shadow.*

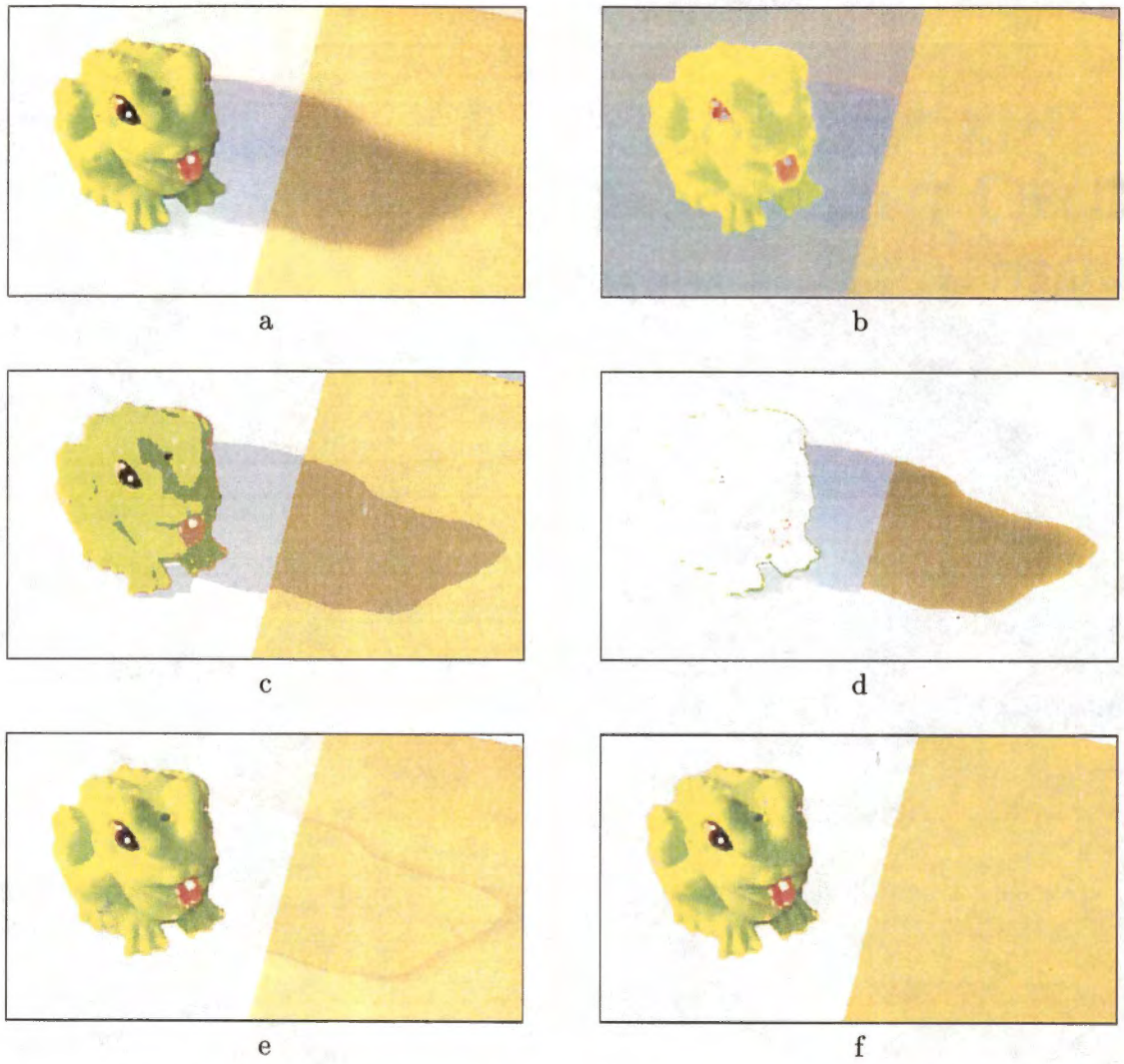


Figure 5.12: *Shadow removal on two-color background: (a) is real image of an artificial frog with a cast shadow on a two-color background. (b) is the normalized image (c) shows the segmented image while the extracted shadow is shown in (d). (e) is the restored image with edge effects due to penumbra and (f) is the final image after post-processing.*

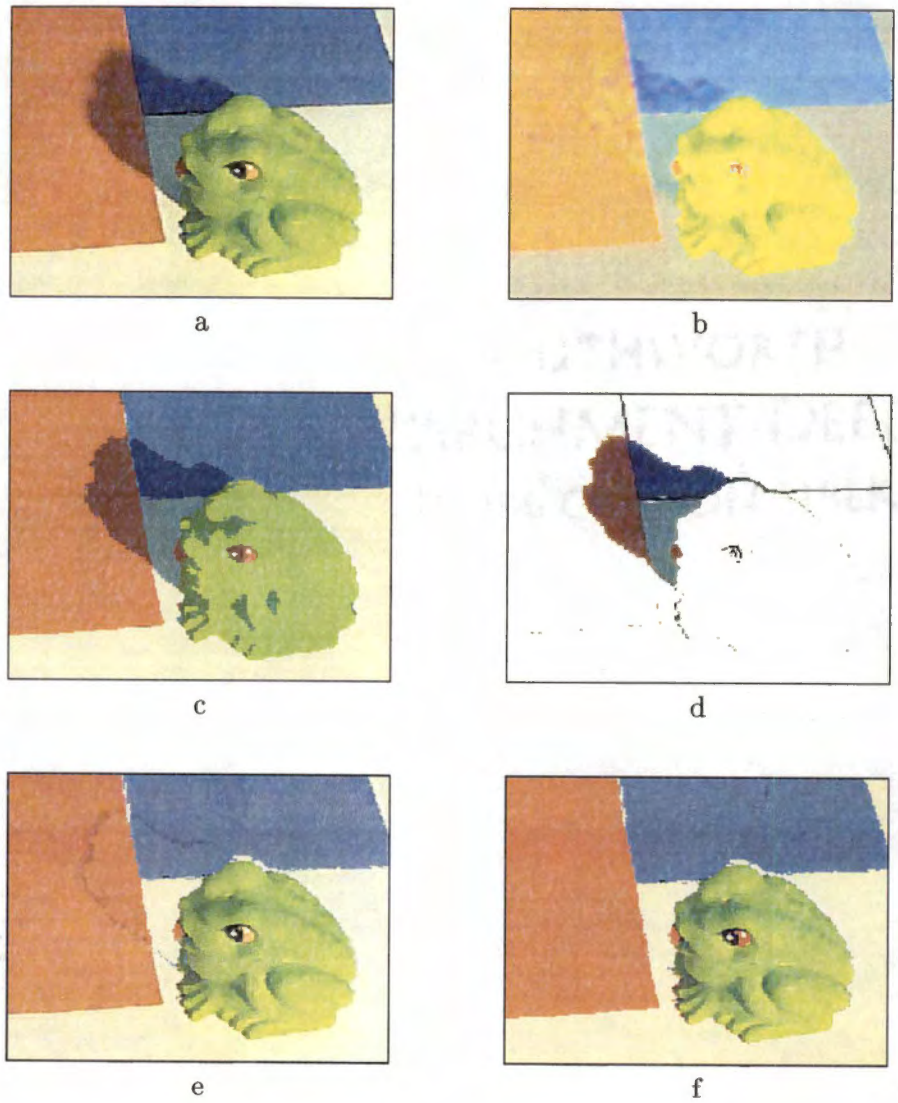


Figure 5.13: *Removal of shadow on multi-color background: (a) is real image of an artificial frog with a cast shadow on a multi-color background. (b) is the normalized image (c) shows the segmented image while the extracted shadow is shown in (d). (e) is the restored image with edge effects due to penumbra and (f) is the final image after post-processing.*

5.8 Algorithm for Shadow Detection and Removal

The following algorithm is used to detect and remove cast shadow(s) in color images.

1. Pre-process the given image by smoothing.
2. Segment/Cluster the image and determine the number of clusters. Call this m_u.
3. Normalize the pre-processed input image and segment/cluster. Let the number of clusters in this case be m_n.
4. If $m_u > m_n$, determine the clusters that have merged and go to step 8.
5. If $m_u = m_n$, form pairs of normalized clusters that are closest to one another. Call them $\mathbf{C}_1$ and $\mathbf{C}_2$ with means $\mathbf{M}_1$ and $\mathbf{M}_2$ respectively.
6. If the distance between $\mathbf{M}_1$ and $\mathbf{M}_2 < \epsilon$, stop - there is no shadow else go to 7.
7. Extract the darker of $\mathbf{C}_1$ and $\mathbf{C}_2$.
8. Locate the shadow pixels in the image domain and map them to the mean of the pixels to which they merged with.
9. Post-process to eliminate the edge effect.
10. End

5.9 Some Results of Shape Reconstruction in the face of Shadow

In chapter 3, we discussed the problem of shadow on shape reconstruction and we have just discussed how to eliminate this problem. By using the combination of our shadow removal technique and the shape reconstruction developed in chapter 4, we obtained some results some of which are shown in the figures 5.14 and 5.15. In figure 5.14, (a) is a

synthetic image with a shadow, (b) and (c) are different views of the reconstructed shape in the presence of the shadow. The corresponding results after shadow has been removed are shown in (d), (e) and (f) respectively. Figure 5.15(a) shows a real image of a human fingers with a cast shadow. Result of shape reconstruction using the unprocessed image is shown at two different views in (b) and (c). Figure 5.15(d) is the image resulting from shadow removal. The Two different views of the reconstructed shape are shown in (e) and (f).

5.10 Conclusion

The shadow detection and removal algorithm detects a cast shadow that drops on the image background. It also recovers the background that is partly in shadow and partly lit and works equally well on a multi-color background. The algorithm is based on normalization and segmentation/clustering technique and it uses user defined Euclidean distance measure constraint tool to validate the shadow candidate pairs.

Since shadow removal is based on mapping the shadow pixels to the mean of the corresponding background the shadow resides, the technique will not work well on a textured and random pixel backgrounds. Another limitation of this approach is the assumption of a linear color space. This is not a serious limitation as the problem can be overcome by linearization of the color space by gamma correction. The algorithm is also dependent on the segmentation/clustering algorithm used. Since no segmentation scheme is perfect, the shadow identification is therefore restricted by the limitation of the segmentation or clustering algorithm used.

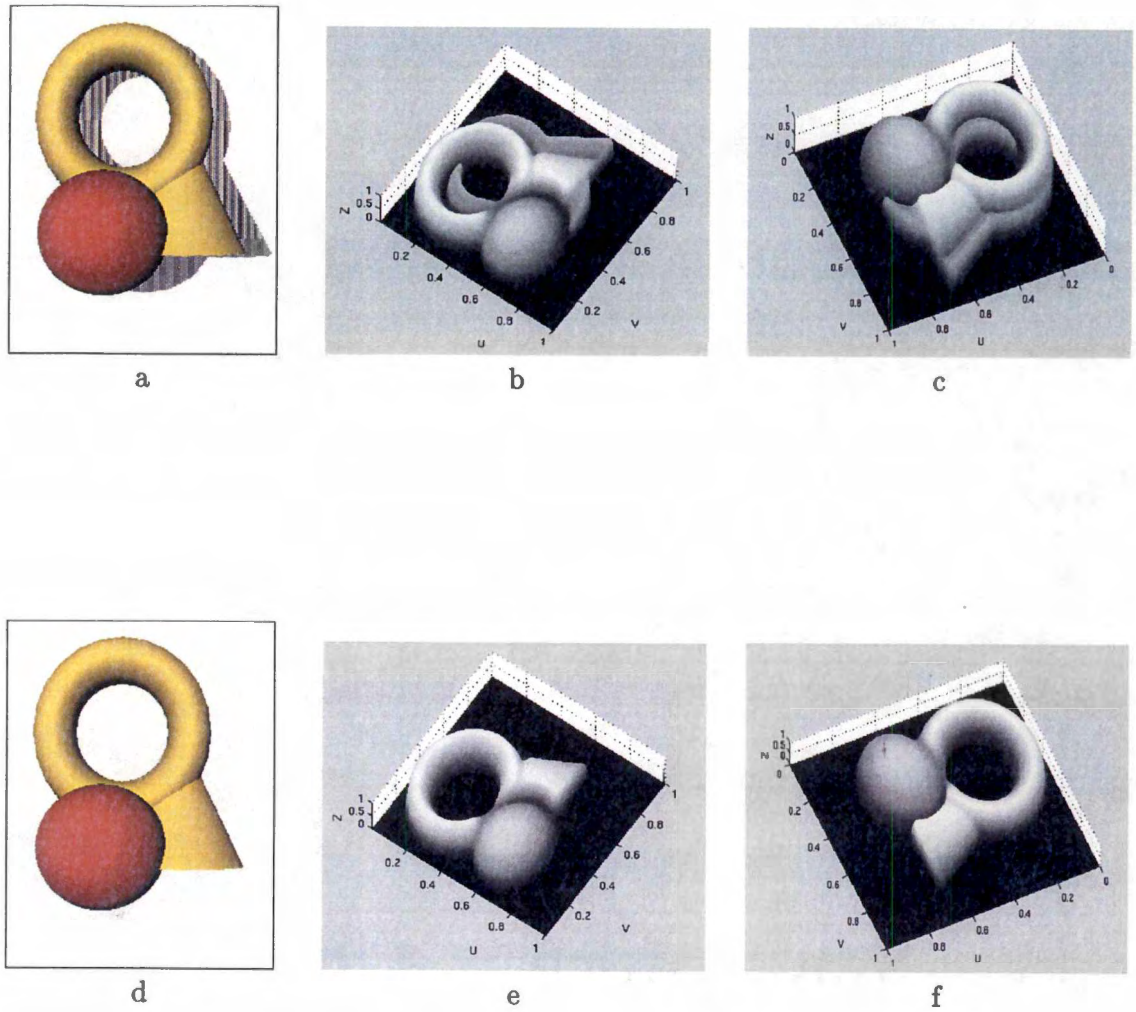


Figure 5.14: *Shape reconstruction of an object in shadow: (a) is synthetic image with a shadow. Results of the shape reconstruction are displayed at different views in (b) and (c). (d) is the restored image and the corresponding results of the shape reconstruction are shown in (e) and (f).*

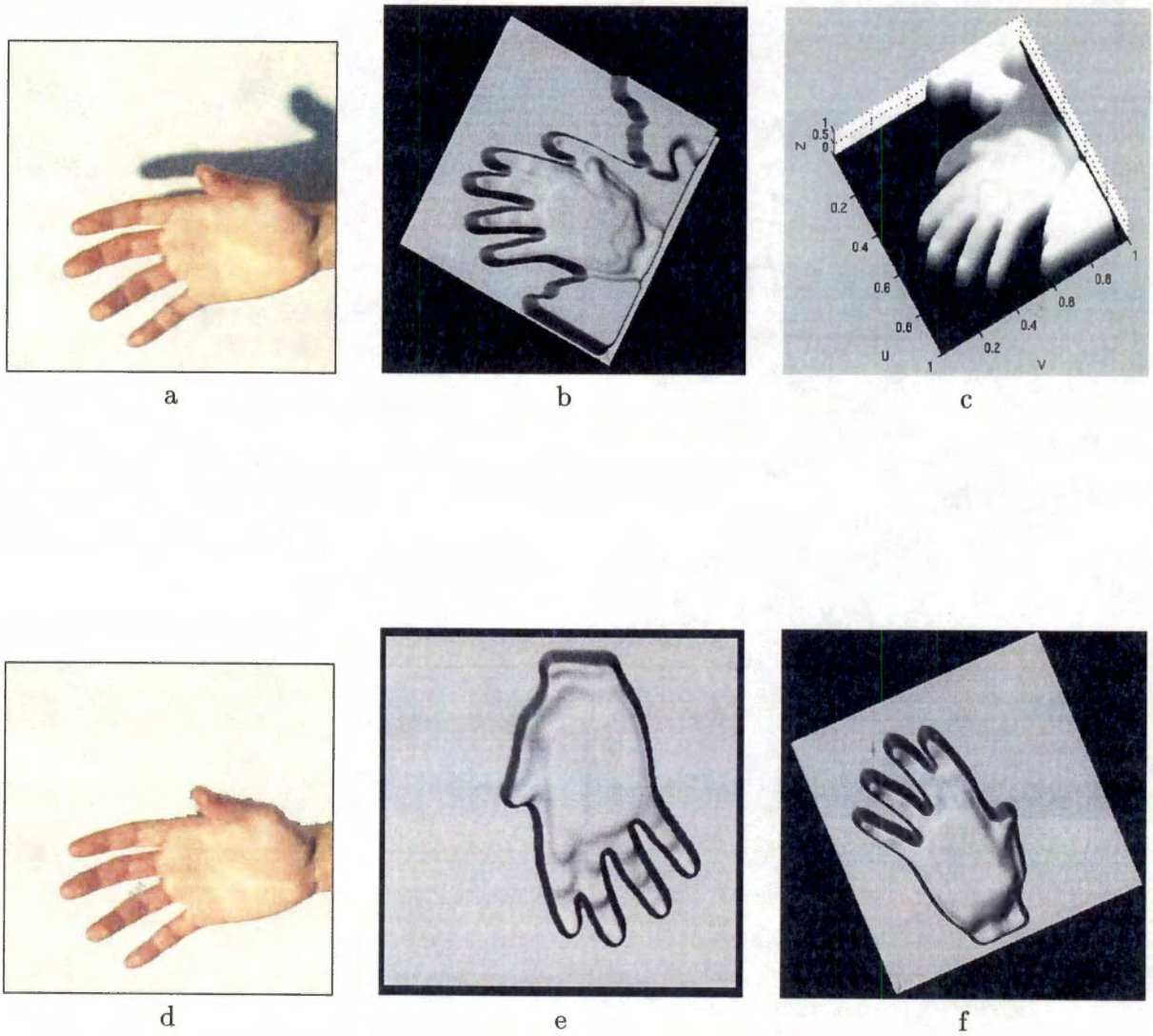


Figure 5.15: *Shadow reconstruction of human fingers with shadow: (a) is a real image of a human fingers. Results of the shape reconstruction are displayed at different views in (b) and (c). (d) is the restored image and the corresponding results of the shape reconstruction are shown in (e) and (f).*

A combination of our shadow and shape reconstruction schemes was employed to reconstruct shape in the presence of shadow. These combined schemes worked well as evident from the results shown in figures 5.14 and 5.15. Although the fusion of these schemes work well enough, it is still limited by the segmentation algorithm which is used for shadow identification and removal.

In the next chapter, we suggest how our results could be improved and other future related work.

CHAPTER 6

Conclusions and Suggestions for Future Work

6.1 Conclusions

The research goal was to:

1. Develop an algorithm for recovering shape of objects using single-color images under general illumination conditions

   - with constant and non-constant surface reflectance,
   - with/without cast shadows

   for calibrated and uncalibrated systems, while relaxing the assumptions given in chapter 1.

2. Design a tool for identifying and eliminating shadows in color imagery without the darkest region assumption for a linear camera.

3. Implement a tool based on least-mean-squared distance shape metric for testing the validity of the shape estimation process.

Each of these tasks was accomplished and the first two are the major contributions of this dissertation. For task one, three different techniques were developed:

- the pre-segmentation approach for identifying regions of constant surface reflectance and the use of boundary conditions,
- normal propagation which utilizes a smoothness constraint and transformation correction matrix and

- the variational approach that utilizes a global smoothness constraint to iteratively solve for the optimal object surface.

The first approach is dependent upon the segmentation algorithm. The rest of the two techniques were designed to overcome dependency on segmentation. The extensive experimental tests show that in general the *variational approach* returned the best results followed by the *normal propagation* and also robust to high degree of image noise.

The main difficulties with these techniques are:

- the pre-segmentation approach is dependent upon the segmentation algorithm and failed when the segmentation algorithm was not able to locate all the true regions of constant surface reflectance,

- the normal propagation is quite sensitive to noise and it requires the input image to be well pre-processed by smoothing,

- the variational approach while offering resistance to noise is relatively slower than the previous ones. It takes between 1 to 10 minutes, depending on the complexity of the image to run.

On the issue of shadow identification and removal, a clustering algorithm was designed without the usual darkest region assumption. The experimental results were nice as seen in the previous chapter. However, the main difficulty is that it is dependent on the segmentation algorithm used. It will fail whenever the segmentation scheme fails.

A validation tool based on least-mean-square distance shape metric was used to validate our results. The results show that our scheme performed reasonably well.

6.2 Suggestions for Future Work

There are several ways this work can be extended to give better results. We observed that segmentation played a key role in this work. It was used in the shape reconstruction (in the pre-segmentation approach) and in shadow detection and removal problem. These schemes failed anytime the segmentation algorithm performed below par. To improve on the performance of these schemes, we suggest two different techniques, segmentation algorithm could be designed.

From the shape reconstruction schemes, a transformation correction matrix $\mathbf{\Gamma}$ was computed whenever a change in the spectral surface-reflectance was detected. By grouping pixels that exhibit same $\mathbf{\Gamma}$, the entire image may be segmented. Another segmentation technique being suggested will use fewer dimensional feature space (azimuth and elevation). A sketch of the algorithm is given below:

1. Normalize the input image and map into the feature space to obtain the parent cluster, which is the union of all the clusters in the feature space.

2. Compute the mean and variance (m_p, σ_p) of the parent cluster.

3. Compute the distance between the mean and the farthest pixel from mean, r_{max}.

4. Reduce the radius to r_1 and compute the new mean and variance (m_1, σ_1) of all the pixels residing inside this circle. If $m_1 \neq m_p$, shrink/inflate the circle by δr. If m_1 is fairly constant and $m_1 \neq m_p$ (call the new radius r_2), break up the cluster into two: child-in and child-out.

5. Using the child-out as the parent cluster go to step 2. If in step 4, $m_1 = m_p$, then a terminal case is reached in which case all the pixels in the cluster will belong to child-in at that level and the iterative process ends.

We called this technique, *Color Image Segmentation by Hierarchical Mean Shift Approach* and the hierarchical structure is obtained by iteratively splitting the child-out cluster until a terminal point is reached. At the end of the iterative clustering procedure, all the pixels in the image must have been classified into groups or classes. The mean of each of the classes (child-in) is computed and all the pixels in each class are assigned to the value of the mean of their class to obtain segmentation result. The algorithm will be slightly modified for shadow removal.

Finally, the dissertation did not address the issue of texture mapping of original color of the input image unto the reconstructed shape. It is therefore being suggested to explore the possibility of doing that as future work.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, An International Thomson Publishing Company, second ed., 1999.

[2] P. H. Christensen and L. G. Shapiro, "Three-dimensional shape from color photometric stereo," *International Journal of Computer Vision* **13**(2), pp. 213–227, 1994.

[3] M. S. Drew and L. L. Kontsevich, "Closed-form attitude determination under spectrally varying illumination," *IEEE International Conference on Computer Vision and Pattern Recognition* , pp. 985–990, 1994.

[4] M. Brooks and B. Horn, "Shape and source from shading," *Artificial Intelligence* **A.I. Memo 820**, pp. 1–11, 1985.

[5] Y. Tian and H. Tsui, "Shape recovery from a color image for non-lambertian surfaces," *Journal of Optical Society of America A* **14**(2), pp. 397–404, February 1997.

[6] A. Yuille, "A method for computing spectral reflectance," *Biol. Cybern.* **56**, pp. 195–201, 1987.

[7] K. Ikeuchi and B. Horn, "Numerical shape from shading and occluding boundaries," *Artificial Intelligence* **17**, pp. 141–184, 1981.

[8] A. P. Pentland, "Local shading analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(2), pp. 170–187, 1984.

[9] A. P. Pentland, "Linear shape from shading," *Int. J. Computer Vision* **4**, pp. 153–162, 1990.

[10] T. Kanade, "Recovery of the three-dimensional shape of an object from a single view," *Artificial Intelligence* **17**, pp. 409–460, 1981.

[11] S. Bakshi and Y. H. Yang, "Shape from shading for non-lambertian surfaces," *IEEE International Conference on Image Processing* **2**, pp. 130–135, Nov 1994.

[12] B. Horn, *Obtaining shape from shading information, in Psychology of Computer Vision, P.H. Winston*, McGraw-Hill, New York.

[13] R. J. Woodham, "Photometric stereo: A reflectance map technique for determining surface orientation from image intensity," *Proc. SPIE* **155**, pp. 136–143, 1978.

[14] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering* **19**, pp. 139–144, 1980.

[15] G. A. Agoston, *Color Theory and its Application in Art and Design*, Spring-Verlag, Berlin, 1979.

[16] M. D'Zmura and P. Lennie, "Mechanisms of color constancy," *Journal of Optical Society of America A* **3**(10), pp. 1662–1672, November 1994.

[17] R. M. Evans, *Eye, Film and Camera in Color Photography*, Wiley, New York, 1959.

[18] G. Healey and T. O. Binford, "Color algorithms for a general vision system," *International Joint Conference on Artificial Intelligence* **2**, pp. 759–762, August 1987.

[19] R. S. Hunter, *Measurement of Appearance*, Wiley, New York, 1975.

[20] H. C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *Journal of Optical Society of America A* **3**(10), pp. 1694–1699, October 1986.

[21] H. C. Lee, *Estimating the illuminant color from the shading of a smooth surface*, no. Al Memo. 1068, MIT, Cambridge, MA, May, 1988.

[22] S. Shafer, "Using color to separate reflection components," *Color Research and Application* **10**(4), pp. 210–218, 1985.

[23] L. Maloney and B. Wandell, "Color constancy: A method for recovering surface spectral reflectance," *Journal of Optical Society of America A* **3**(1), pp. 29–33, January 1986.

[24] B. Wandell, "The synthesis and analysis of color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**(1), pp. 2–13, January 1987.

[25] D. A. Brainard, B. A. Wandell, and W. Cowan, "Black light: How sensors filter spectral variation of the illuminant," *IEEE Trans. Biomed. Eng.* **36**, pp. 140–149, 1989.

[26] G. West and M. H. Brill, "Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy," *J. Math. Biol.* **15**, pp. 249–258, 1982.

[27] J. Ho, B.V.Funt, and M. Drew, "Separating a color signal into illumination and surface reflectance components: Theory and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(10), pp. 966–977, October 1990.

[28] B. K. P. Horn, "Height and gradient from shading," *International Journal of Computer Vision* **5**(1), pp. 37–75, 1990.

[29] K. Ikeuchi, "Determining surface orientations of specular surfaces using the photometric stereo method," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **3**(6), pp. 661–669, 1981.

[30] D. Russell, *Calculus of Variations and Control Theory*, Academic Press, New York, 1976.

[31] D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, 1982.

[32] A. Petrov and L. L. Kontsevich, "Color image field rank and segmentation," 1994.

[33] M. H. Brill, "Photometric models in multispectral machine vision," *SPIE* **1453**, pp. 369–380, 1991.

[34] L. L. Kontsevich, A. Petrov, and I. Vergelskaya, "Reconstuction of shape from shading in color images," 1994.

[35] K. Ikeuchi and B. Horn, "A new approach to surface reconstruction: the coupled depth/slope model," *ICCV* , pp. 227–283, 1987.

[36] Y. I. R. J. Woodham and R. Barman, "Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength," *Technical Report TR 91-18. University of British Columbia Department of Computing Science, 1991* , pp. –10, 1991.

[37] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*, Addison-Wiley, 1987.

[38] V. Kravtchenko, *Pixel Clustering Methods in RGB space*, www.cs.ubc.ca/~ vk/516.html, December 1998.

[39] J. Theiler and G. Gisler, "A contguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation," *SPIE* , pp. 108–118, 1997.

[40] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: Color image segmentation," *IEEE International Conference on Computer Vision and Pattern Recognition* , pp. 750–755, 1997.

[41] P. Perona and J. Malik, "Scale-space and edge detection using anistropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(7), pp. 629–639, 1990.

[42] M. J. Black, D. M. G. Sapiro, and D. Heeger, "Robust anisotropic diffusion," *IEEE Transactions on Image Processing* **7**(3), pp. 421–432, 1998.

[43] M. Black and G. Sapiro, "Edges as outliers: Anisotropic smoothing using local image statistics," *www.cs.brown.edu/people/black/papers.html* .

[44] R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley, third edition ed., 1992.

[45] E. Davies, *Machine Vision: Theory, Algorithms and Practicalities*, Academic Press, 1990.

[46] M. P. Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, 1976.

[47] P. Lancaster and K. Šalkauskas, *Curve and Surface Fitting - An Introduction*, Academic Press, 1986.

[48] P. Dierckx, *Curve and Surface Fitting with Splines*, Oxford, 1993.

[49] E. Hall, *Computer Image Processing and Recognition*, Academic Press, 1979.

[50] G. Funka-Lea, "The visual recognition of shadows by an active observer," *A Dissertation in Computer and Information Science, University of Pennsylvania* .

[51] A. Witkin, "Intensity-based edge classification.," *In Proceedings of the National Conference on AI* , pp. 36–41, August 1982.

[52] M. Ali, W. Martin, and J. Aggarwal, "Color-based computer analysis of aerial photographs," *Computer Graphics and Image Processing* **9**(4), pp. 282–293, 1979.

[53] B. Schachter, L. Davis, and A. Rosenfeld, "Scene segmentation by cluster detection in color space," *University of Maryland Computer Science Department* (424), 1975.

[54] A. Elgammal, D. Harwood, and L. Davis, *Non-parametric Model for Background Subtraction*, www.eecs.lehigh.edu/FRAME/Elgammal/bgmodel.html.

[55] Y. Cheng, "Mean shift, mode seeking and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**, pp. 790–799, 1995.

[56] L. B. Wolff, "Diffuse reflection," *IEEECV* , pp. 472–478, 1992.

[57] K. Torrance and E. Sparrow, "Theory for offspecular reflection from roughned surface for ray reflect ion," *Journal of Optical Society of America A* **65**(1), pp. 531–536, 1967.

[58] S. Tominaga and B. Wandell, "Standard surface-reflection model and illuminant estimation," *Journal of Optical Society of America A* **6**(4), pp. 576–584, April 1989.

[59] V. Shao, *Implementing the Gamma Correction Algorithm Using the TMS320C2xx DSP*, no. SPRA361, Texas Instruments, September 1997.

[60] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, Interscience Publishers, New York, 1953.

[61] A. K. Jain and R. C. Dudes, *Algorithms for Clustering Data*, Prentice-Hall, 1988.

[62] S. Kaski, *Clustering Methods*, www.cis.hut.fi/~ sami/thesis/node9.html.

[63] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, second ed., 1990.

[64] G. Wyszechi and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, Wiley, New York, second ed., 1982.

[65] S. Kaski, *Clustering Methods*, www.cis.hut.fi/~ sami/thesis/node9.html.

APPENDICES

APPENDIX A

Surface Reflectance and Radiometric Camera Models

A.1 Lambertian Model

A Lambertian surface, which is also referred to as a perfectly diffuse reflector provides radiance which is independent of the viewing direction. Hence, the surface appears equally bright from all viewing directions. The amount of light reflected, I$(x, y; \lambda)$ from a small area, dA toward the viewer is directly proportional to the cosine of the angle between the viewing direction and the normal vector. This may be expressed as

$$I(x, y; \lambda) = \alpha I_o(\lambda) \left(\vec{n}(x, y) \cdot \vec{v}(x, y)\right) \tag{A.1}$$

where

I_o is the intensity of the light source, α is the albedo ($0 \leq \alpha \leq 1$)and λ is the wavelength of the illuminant.

The popularity and simplicity of this model has lead to most common assumption over the years in computer vision. Most reflectance dependent algorithms assume the intensity distribution of diffuse reflection from inhomogeneous dielectric is Lambertian [56]. If a surface is Lambertian, it follows that there is absence of specularity. The main difficulty with this model is that most common objects display some specularity at a given viewing/illumination angle combinations. This is the major motivation for other models.

A.2 Torrance and Sparrow Reflection Model

One of the most commonly used models, Torrance and Sparrow Reflection Model [57], [58], assumes that the intensity of the reflected light $I(x, y; \lambda)$ from a given surface consists of two parts. The part which is dependent on the angle of reflection is called specular

reflection, $I_S(x, y; \lambda)$ and the other which is independent of angle of reflection is called the diffuse or body reflection, $I_B(x, y; \lambda)$. Using this model, the radiance of the reflected light which is a function of geometric scale factors and wavelength is given by

$$I(x, y; \lambda) = I_S(x, y; \lambda) + I_B(x, y; \lambda) \tag{A.2}$$

where $I(x, y; \lambda)$ = radiance of the reflected light, Since the spectral composition of reflected light is not dependent on the geometric conditions [58] for single light, eq(A.2) can be rewritten as

$$I(x, y; \lambda) = g_S(x, y)R_S(\lambda) + g_B(x, y)R_B(\lambda) \tag{A.3}$$

where $g_S(x, y)$ and $g_B(x, y)$ are the geometric scale factors for specular and diffuse reflections and

$$I_S(x, y; \lambda) = g_S(x, y)R_S(\lambda)$$

$$I_B(x, y; \lambda) = g_B(x, y)R_B(\lambda)$$

where $R_S(\lambda)$ and $R_B(\lambda)$ may be defined by

$$R_S(\lambda) = L(\lambda)S_S(\lambda)$$

$$R_B(\lambda) = L(\lambda)S_B(\lambda)$$

Where $L(\lambda)$ is defined as the spectral power distribution of the incident light and S the spectral surface reflectance. Hence, eq(A.3) may be re-written as:

$$I(x, y; \lambda) = g_S(x, y)L(\lambda)S_S(\lambda) + g_B(x, y)L(\lambda)S_B(\lambda) \tag{A.4}$$

For multiple light sources of about the same spectral content, eq(A.4) can be rewritten as a superposition of all light sources as;

$$\sum_{i=1}^{n} I_i(x,y;\lambda) = \sum_{i=1}^{n}(g_S(x,y)L(\lambda)S_S(\lambda))_i + \sum_{i=1}^{n}(g_B(x,y)L(\lambda)S_B(\lambda))_i \tag{A.5}$$

where n = the number of light sources.

Figure A.2 below illustrates image formation using multiple light sources. We have chosen to use this model in this work. The motivation for this choice includes the simple and flexible nature of the model. It works well on a wide range of materials and offers easy separability of specular and diffuse components. Although the model has a disadvantage that not all materials can be modeled, the vast majority of materials can be effectively represented.

A.3 Radiometric Camera Models

The camera spectral responsivity, $\mathbf{C}(\lambda)$ has the tendency to modify reflection model. Using the form of Torrance and Sparrow Reflection Model given in eq(A.4), the model modifies to:

$$\vec{I}(x,y;\lambda) = g_S(x,y)\mathbf{C}(\lambda)\mathbf{L}(\lambda)\vec{S}_S(\lambda) + g_B(x,y)\mathbf{C}(\lambda)\mathbf{L}(\lambda)\vec{S}_B(\lambda) \tag{A.6}$$

The diagonal entries of $\mathbf{C}$ are c_r,c_g and c_b. The above equation (eq(A.6)) may be referred to as the quantum catch of the camera. The quantum catch of the camera is affected by the camera non-linearity, which is also known as camera or system gamma. This brings about the issue of gamma correction.

Gamma Correction

There is a non-linear relationship between a pixel value and its displayed intensity. Computer monitors have intensity to electrical response curve which is about 2.2 power function. To illustrate this, consider a pixel of intensity i, $(0 \leq \mathrm{i} \leq 1)$, sent to a computer monitor. The monitor displays this intensity as $\mathrm{i}^{2.2}$. The device (monitor) is therefore

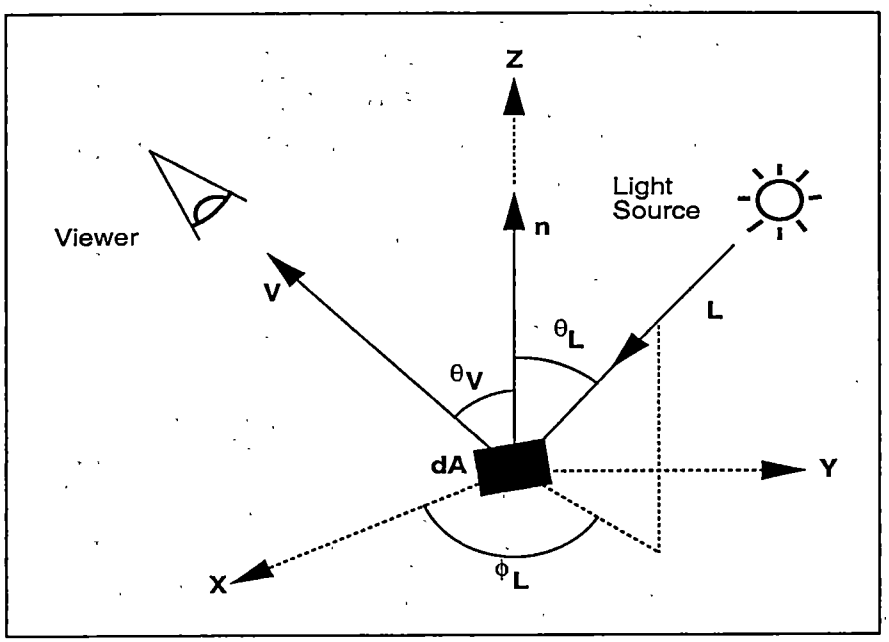


Figure A.1: *Directions of illuminant and viewer with respect to the normal*

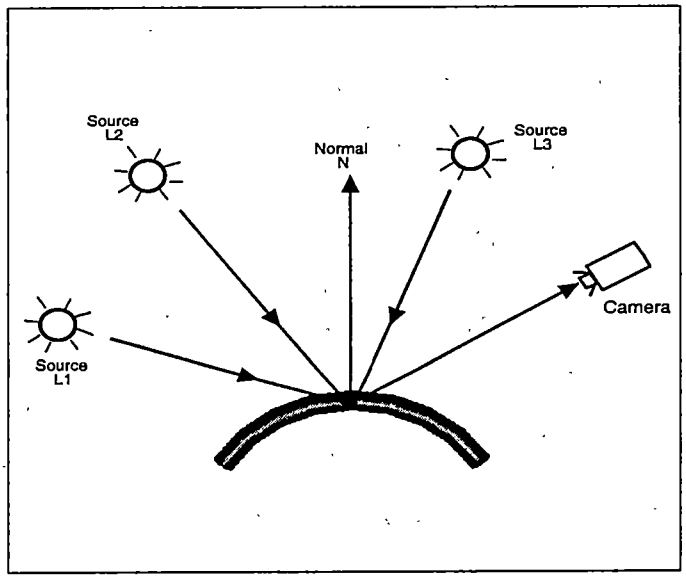


Figure A.2: *The geometry of an image formation using multiple light sources L1, L2 and L3*

said to have a gamma of 2.2. For the intensity to be displayed correctly, the computer monitor must be gamma corrected. Therefore, gamma correction may be defined as an image processing algorithm that compensates for the non-linear effect of signal transfered between electrical and optical devices [59]. The formula for gamma correction for a device with a gamma factor of 2.2 is given by Shao [59]:

$$\begin{aligned} R_{display} &= R_{received}^{2.2} \\ G_{display} &= G_{received}^{2.2} \\ R_{display} &= R_{received}^{2.2} \end{aligned} \tag{A.7}$$

where R, G, B values are normalized to the range of $[0, 1]$. To compensate for the nonlinear relationship between intensity and display, the RGB data must be gamma corrected as follows:

$$\begin{aligned} R_{transmit} &= R_{received}^{p} \\ G_{transmit} &= G_{received}^{p} \\ R_{transmit} &= R_{received}^{p} \end{aligned} \tag{A.8}$$

Where $p = \frac{1}{2.2}$. The R, G, B values are also normalized to the range of $[0, 1]$. The compensation linearizes the displayed signals. Hence, the displayed signal becomes:

$$R_{display} = R_{transmit}^{2.2} \tag{A.9}$$

$$= \left(R_{received}^{p}\right)^{2.2} \tag{A.10}$$

Similar expressions hold for $G_{display}$ and $B_{display}$.

A.4 Euler's Equation

The Euler equation corresponding to the functional (see Courant and Hilbert [60])

$$\int\int_{\Omega} F\left(x, y, \mathbf{n}, \mathbf{n}_x, \mathbf{n}_y\right) dxdy \tag{A.11}$$

is

The Euler equation corresponding to the functional

$$F_{\mathbf{n}} - \frac{\partial}{\partial x} F_{\mathbf{n}_x} - \frac{\partial}{\partial y} F_{\mathbf{n}_y} = 0 \tag{A.12}$$

APPENDIX B

Clustering and Color Image Segmentation

A vast body of knowledge exists in the field of color image segmentation and a complete review of this work is beyond the scope of this proposal. Two color segmentation schemes are used in this work and are: Robust Analysis of Feature Spaces [40], K-Means clustering technique [61], [62], [38], [39], [63].

B.1 Robust Analysis of Feature Spaces

In the robust analysis of feature spaces technique[40], the image is transformed to an isotropic space, the LUV space where dominant features (colors) correspond to high-density regions. The transformation equations are shown in the appendix. The next stage of the analysis of the feature space, Comaniciu et al[40] used a non-parametric method called the Mean Shift Algorithm (MSA)[55] to estimate the density of gradients. In their space, a sphere is used as a search window. The radius of the sphere defines the resolution of the segmentation. The MSA iteratively shifts each data point to the average of data points in its neighborhood until the center of each cluster is found. It may also be referred to as a mode-seeking algorithm

B.2 Mean Shift Algorithm

Let $p(\vec{x})$ be the probability density function of n-dimensional feature vectors, $\vec{x}$ and $S_{\vec{x}}$ a sphere (whose radius r defines the size of the so called search window) centered on $\vec{x}$ and contains feature vectors $\vec{y}$ such that $\|\vec{y} - \vec{x}\| \leq r$. Define a vector $\vec{z} = \vec{y} - \vec{x}$. If its expected value, given $\vec{x}$ and $S_{\vec{x}}$ is μ, then μ is given by

$$\mu = E\left[\vec{z} \mid S_{\vec{x}}\right]$$

$$= \int_{S_{\vec{x}}} (\vec{y} - \vec{x}) p(\vec{y} \mid S_{\vec{x}}) d\vec{y}$$

$$= \int_{S_{\vec{x}}} (\vec{y} - \vec{x}) \frac{p(\vec{y})}{p(\vec{y} \in S_{\vec{x}})} d\vec{y} \tag{B.1}$$

For sufficiently small sphere, $S_{\vec{x}}$, $p(\vec{y} \in S_{\vec{x}}) \approx p(\vec{x}) V_{S_{\vec{x}}}$

where $V_{S_{\vec{x}}}$ = volume of sphere.

Using the first order Taylor series approximation gives

$$p(\vec{y}) = p(\vec{x}) + (\vec{y} - \vec{x})^T \nabla p(\vec{x}) \tag{B.2}$$

where $\nabla p(\vec{x})$ = gradient of probability density function. Substituting eq(B.2) into eq(B.1) gives

$$\mu = \int_{S_{\vec{x}}} \left(\frac{(\vec{y} - \vec{x})(\vec{y} - \vec{x})^T}{V_{S_{\vec{x}}}} \frac{\nabla p(\vec{x})}{p(\vec{x})} \right) d\vec{y}$$

Fukunaga [63] shows that the above integral could be reduced to

$$\mu = \left(\frac{r^2}{n+2} \right) \frac{\nabla p(\vec{x})}{p(\vec{x})} \tag{B.3}$$

Hence,

$$E\left[\vec{x} \mid \vec{x} \in S_{\vec{x}}\right] - \vec{x} = \left(\frac{r^2}{n+2} \right) \frac{\nabla p(\vec{x})}{p(\vec{x})} \tag{B.4}$$

Notice that the mean shift vector is same as the difference between the local mean and the center of the search window. The center of the high density region is deemed to be attained when the mean shift becomes less than or equal to some threshold value. Further details about the derivation can be found in [55], [40] and [63].

The question to be addressed here is having obtained the feature space, how does one initiate the search for the centers of the high density regions. How does one ensure

that the search commences from a point near the high-density region? The size of the search window defines the segmentation resolution. The bigger the size of the window (ie the radius of sphere) the lower the resolution becomes. Comaniciu et al[40] explained how to go about these questions. The initial location of the search window is randomly chosen and to ensure that the search begins from a point close to the high-density region, a number of locations (about 25 pixels at a time) are chosen at random in the image domain. At each location, the mean of 3x3 neighbors is computed and mapped into the feature space. If the neighborhood of these pixels belongs to a large homogeneous region of high probability, then the initial start point is deemed to have been found. The Mean Shift Algorithm (MSA) is then used to locate the closest center of high-density region also referred to as the mode. Ideal convergence is obtained when the mean shift is zero (this is very difficult to realize in practice). The reference [40] regarded a mean shift of less than 0.1 to be adequate for good convergence.

At the mode, the pixels and their 8-connected neighbors yielding the feature vectors inside the search window are deleted from both domains. The above procedure is repeated until the number of feature vectors in the search window becomes less than a preset threshold. Any color deemed significant in the image domain are extracted and then used as initial feature palette for final feature palette. In the final palette, all the pixels that make up the feature vectors inside the search windows are assigned to the color of the window center without regard to the image information. For further information on how this segmentation works, see [40].

B.3 Transformation from RGB to LUV

This equations below are used to transform from image domain displayed in RGB format to a LUV feature space.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \underbrace{\begin{bmatrix} 3.247 & -1.540 & -0.499 \\ -0.972 & 1.875 & 0.042 \\ 0.057 & -0.205 & 1.060 \end{bmatrix}}_{\mathrm{M}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \mathrm{M}^{-1} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Where X, Y and Z are device independent tristimulus called the CIE special sets of mathematical lights. The chromaticity coordinates x, y, and z are related to X, Y and Z by

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = 1 - x - y$$

We can define another chromaticity coordinates u' and v' given by

$$u' = \frac{4X}{X+15Y+3Z}$$

$$v' = \frac{9Y}{X+15Y+3Z}$$

The CIE LUV color space is thus defined as:

$$L^\star = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \quad \text{for} \quad \left(\frac{Y}{Y_n}\right) > 0.008856$$

$$L^\star = 903.3\left(\frac{Y}{Y_n}\right) \quad \text{for} \quad \left(\frac{Y}{Y_n}\right) \leq 0.008856$$

$$u^\star = 13L(u' - u'_n)$$

$$v^\star = 13L(v' - v'_n)$$

where

$$u'_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n}$$

$$v'_n = \frac{9Y_n}{X_n + 15Y_n + 3Z_n}$$

as defined by [64]

The index n denotes the neutral white reference light. Using the CIE daylight standard D_{65} illuminant gives Y_n equal to 100. The other two tristimulus values X_n and Z_n are 95.05 and 108.88 respectively. If Y_n is normalized to 1, X_n and Z_n are then 0.9505 and 1.0888 respectively. For neutral colors, u* = v* = 0. The steps for feature space analysis are summarized in [40].

B.4 K-Means Clustering

This method is also referred to as K-means Partitional Clustering. The spectral content of the pixels, the R-G-B values may be used as the main features for clustering multicolored images. It essentially decomposes the data set into a set of disjoint clusters. It minimizes a criteria function by assigning clusters to the peaks in the probability density function or the global structure [65]. The algorithm is given by the following steps.

1. Choose K different initialization points which are indeed the initial input means.

2. For each pattern, determine which of the K cluster means it is closest to and assign the pattern to that cluster.

3. At the end of each iteration, compute the new mean for each of the K clusters or classes.

4. Stop when no new pattern is being assigned to a cluster. Otherwise, repeat steps 2 and 3 above.

The measure of similarity used in this work is the Euclidean distance. The flow chart that illustrates the algorithm is shown in figure B.1 below.

B.5 Color Segmentation using K-Means Technique

At the end of the iterative clustering procedure, all the pixels in the image must have been classified into groups or classes according to the measure of similarity used which is the Euclidean distance. The mean of each of the classes is computed and all the pixels in each class are assigned to the value of the mean of the class. The number of classes which is of course the number of clusters is defined by the value of K. Higher values of K gives more number of cluster classes. The higher the K, the more and better the features extracted becomes. Hence, K defines the segmentation resolution. The higher the resolution, the better the 3D shape recovery becomes. This is because more of the representative colors in image domain are more properly transformed with correct transformation matrix.

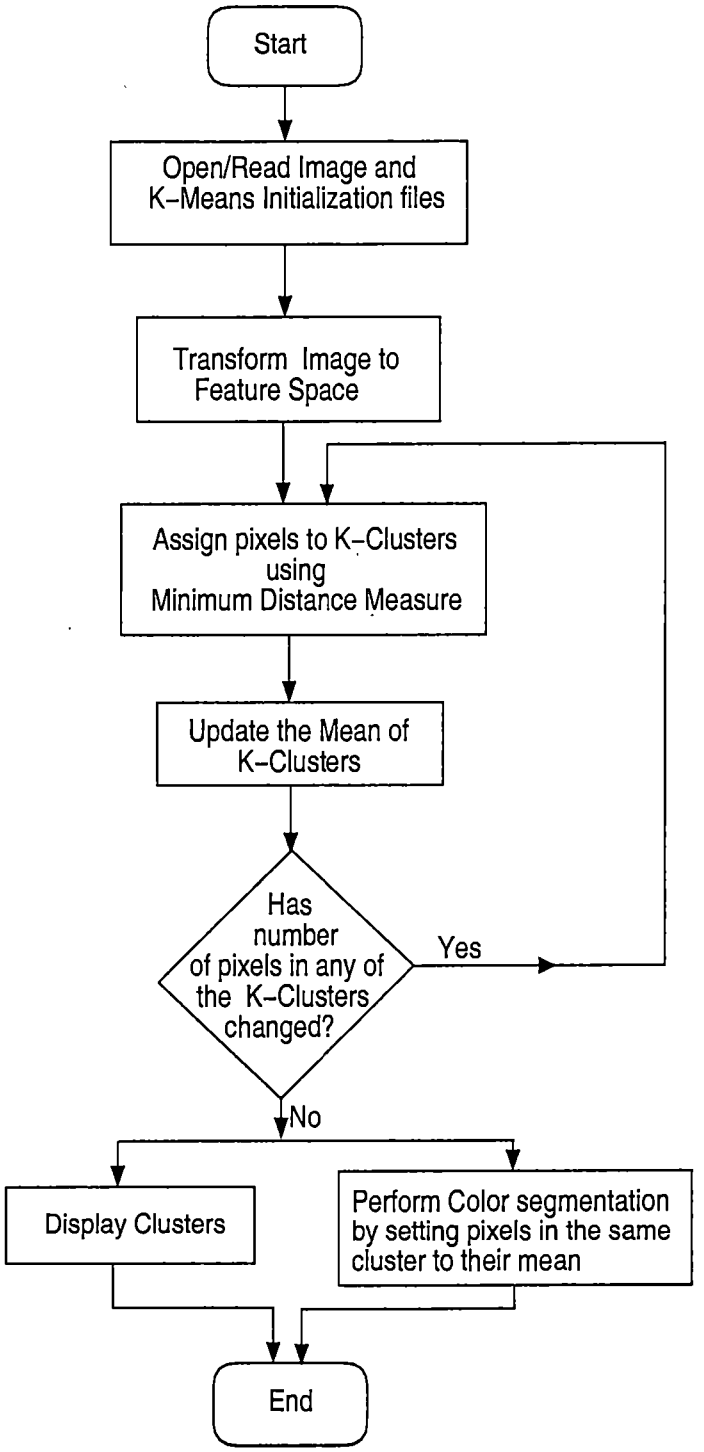


Figure B.1: *K-Means Flow Chart for Color Segmentation*

VITA

Ambrose Ejiofor Ononye was born in Nigeria. He had his primary and high school education at St. Patrick's Primary School and Izzi High School respectively at Abakaliki. In 1981, he received the Bachelor of Science degree (B.Sc.) in Physics from the University of Ibadan. He worked at Bendel State University, Ekpoma (now called Edo State University) where he eventually became a Lecturer. In 1985, he attended the University of Benin and earned a Post Graduate Diploma in Electrical and Electronic Engineering in 1987 and later received a Master of Engineering degree, M.Eng. in Electrical and Electronic Engineering (from same University in 1990) with emphasis on Electronics and Communication. He was admitted at the Physics Department, University of Tennessee where he received an M.S. degree. He later transferred to the Department of Electrical and Computer Engineering and will be presented with a Doctor of Philosophy degree in Electrical Engineering in December 2001. His research interests are computer vision, image processing, pattern recognition and controls.