İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

UYARLANIR DÜZLEME YÖNTEMİYLE GÖRÜNTÜ TONLARINDAN ŞEKİL ELDE ETME

DOKTORA TEZİ

Y. Müh. Atilla GÜLTEKİN

Tezin Enstitüye Verildiği Tarih : 30 Ekim 1996
Tezin Savunulduğu Tarih : 4 Nisan 1997

Tez Danışmanı : Doç. Dr. Muhittin GÖKMEN 15/4/1997
Diğer Jüri Üyeleri : Prof. Dr. Nadir YÜCEL 15/4/1997
Prof. Dr. Erdal PANAYIRCI 15/4/97
Prof. Dr. Bülent SANKUR 15/4/97
Prof. Dr. Aytül ERÇİL 28/4/97

NİSAN 1997

ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

ADAPTIVE SHAPE FROM SHADING

Ph.D. THESIS

Atilla GÜLTEKİN, M.S.

Date of Submission : October 30, 1996
Date of Approval : April 4, 1997

Supervisor : Assoc. Prof. Dr. Muhittin GÖKMEN 15/4/1997

Jury Members : Prof. Dr. Nadir YÜCEL 15/4/1997
Prof. Dr. Erdal PANAYIRCI 15/4/97
Prof. Dr. Bülent SANKUR 15/4/97
Prof. Dr. Aytül ERÇİL 28/4/97

APRIL 1997

ACKNOWLEDGEMENTS

I would first like to express my appreciation to my respected advisor Assoc. Prof. Muhittin Gökmen for his encouragement and guidance. He is the person who taught me how to present a poster session, how to write papers and how to carry out a scientific work. He always motivated me to finish this long run whenever we came together. It was really a wonderful experience for me doing this research under his guidance. I am grateful to him for introducing me to the exciting world of images.

I would like to express my deepest appreciation to my wife for endless supporting and encouraging me to complete this thesis.

I have also my thanks to my parents for their greatest love and support that made all this possible.

I am thankful to Ping-Sing Tsai from University of Central Florida for providing the *Mozart* image and to Ali Sabri Şanal for providing the *Midilli* image.

Atilla Gültekin

April 1997

TABLE OF CONTENTS

List of Symbols

β	:	Intensity gradient constraint parameter
$c(x,y)$	:	Conrtol signal
$(\delta p, \delta q, \delta z)$	:	Increments in (p,q,z) after an iteration
ϵ	:	Space between picture cells
$e(x,y)$	:	Error signal
$E\{\,.\,\}$	:	Expectation operator
$\mathcal{F}(x,y,\lambda_{old})$	:	Function used to calculate $\lambda_{new}(x,y)$ values
$F_1(p,q)$	:	Brightness constraint
$F_2(p,q)$	:	Smoothness constraint
$F_3(p,q,z)$	:	Integrability constraint
$F_4(p,q)$	:	Intensity gradient constraint
$F_1'(p,q)$	:	Modified brightness constraint
$F_2'(p,q)$	:	Spatially controlled regularization function
η	:	Albedo
f,g	:	Stereographic space parameters
f_s, g_s	:	Position of light source in stereographic coordinates
$I(x,y)$	:	Gray level image
$\lambda(x,y)$	:	Spatially controlled regularization parameter
$\lambda_{old}(x,y), \lambda_{new}(x,y)$	:	Old and new values of λ
$\lambda_{min}, \lambda_{max}$	:	Minimum and maximum values of λ

μ	:	Integrability constraint parameter
∇^2	:	Laplacian operator: $\nabla^2 = (\partial^2/\partial x^2) + (\partial^2/\partial y^2)$
$\vec{N}$	:	Unit normal vector
Ω	:	Image domain
p, q	:	x and y partial derivatives of height function $z(x, y)$
p_s, q_s	:	Position of light source in image plane
p^*, q^*	:	Gradients of true solution
$R(p, q)$	:	Reflectance Map
$\vec{S}$	:	Illuminant direction vector
σ, τ	:	Slant and Tilt angles
θ_i	:	Incident angle
x, y	:	Spatial coordinates
$z(x, y)$	:	Surface height function
$z^*(x, y)$	:	Height of true solution
$\bar{z}_{ij}$	:	Local average of z, i.e., $\bar{z}_{ij} = \frac{1}{4}(z_{i,j+1} + z_{i+1,j} + z_{i,j-1} + z_{i-1,j})$

List of Figures

List of Tables

SUMMARY

Shape recovery is a classic problem in computer vision. The goal in shape recovery is to extract surface orientation and surface depth from one or more images. Shape-from-shading (SFS) deals with the recovery of 3-Dimensional shape from a single shaded image. Shape-from-photometric-stereo is another method for shape recovery used in computer vision. In this technique, shape is recovered by using multiple input images of the same scene generated fixed viewing direction and different light source directions.

The shape-from-shading problem is addressed in this study. All the constraints used in SFS are combined in an iterative scheme and the effects of these constraints on the solution are explored. Then, a shape-from-shading algorithm based on the regularization theory is developed, in which the smoothness is controlled spatially over the image space. The spatial control of the smoothness is achieved by employing additional knowledge about the difference between the image which is obtained from the regularized solution and the input image. The adaptive nature of the algorithm eliminates the selection of the optimum value of the smoothness parameter. Using calculus of variations and a linear approximation of the reflectance map, the new adaptive iterative scheme is developed. The new algorithm is robust, data driven and updates both the gradient field and height maps simultaneously. A hierarchical implementation of our adaptive SFS algorithm is also presented. In order to improve the quality of the reconstruction obtained by the SFS algorithm, we integrated our adaptive SFS approach and photometric stereo technique to recover shape by using more than one input image.

ÖZET

UYARLANIR DÜZLEME YÖNTEMİYLE GÖRÜNTÜ TONLARINDAN ŞEKİL ELDE ETME

Bilgisayar görünün (*Computer Vision*) temel amaçlarından biri, objelerin iki boyutlu görüntülerinden, onların üç boyutlu yüzey özelliklerinin elde edilmesidir. Görüntü alma sırasında üç boyutlu objelerin iki boyutlu izdüşümleri elde edilmekte ve bu izdüşüm sırasında objeye ilişkin özelliklerden bir çoğu yitirilmektedir. Üçüncü boyuta ilişkin bilgileri elde etmek amacıyla, genellikle görüntüyü oluşturan benek (*pixel*) değerleri ile obje yüzeyi arasındaki ilişki kullanılır ve bulunabilecek bir çok olası çözümü teke indirmek için, çözüm uzayı üzerinde bir takım sınırlamalar (*constraints*) konur. Görüntüdeki tonlamaların obje şekilleri hakkında taşıdığı bilgiler kullanılarak, obje şeklini bulmak için geliştirilen ve genellikle Tondan Şekil Bulma-TŞB (*Shape from Shading-SFS*) problemi olarak bilinen bu yaklaşımda, önce görüntü ile obje yüzeyleri ve ışık kaynakları arasındaki ilişki elde edilmekte ve daha sonra giriş resmine olabildiğince yakın bir görüntü verecek obje yüzeyi, yani objenin şekli, belirlenmektedir [1, 2].

Tondan-şekil-bulma problemlerinde, obje yüzeyi ve ışık kaynağı ile görüntüyü oluşturan benek değerleri arasındaki ilişki Yansıtma Haritası (*Reflectance Map*) $R(p, q)$ ile belirlenir. Burada, p ve q yüzey parametreleridir ve kurulacak z yüzeyinin kısmi türevlerini göstermek üzere

$$p = \frac{\partial z}{\partial x} \qquad \text{ve} \qquad q = \frac{\partial z}{\partial y} \tag{1}$$

şeklinde ifade edilmektedir.

Işık kaynağından gelen ışınların yüzey tarafından nasıl yansıtılacağı, o yüzeyin yansıtma özelliklerine bağlıdır. Değişik yüzeylerdeki ışık yansımalarını ifade etmek amacıyla çeşitli modeller geliştirilmiştir. Bu modellerin, teorik olarak en basiti ve pratikte olarak en çok kullanılanı Lambertian yansıtma modelidir. Lambertian bir yüzeyin özelliği, üzerine gelen tüm ışığı yansıtması ve

yüzey parlaklığının bakış doğrultularından bağımsız olmasıdır. Lambertian bir yüzey için yansıtma haritası

$$R(p,q) = \eta \frac{\cos\sigma - p\sin\sigma\cos\tau - q\sin\sigma\sin\tau}{\sqrt{1+p^2+q^2}} \tag{2}$$

şeklinde verilir. Burada, η (*albedo*) aydınlanma katsayısını gösteren bir sabit, σ (*slant*) ve τ (*tilt*) ise üç boyutlu resim uzayında ışık kaynağının konumunu belirleyen açılardır.

Giriş resmindeki bir (x,y) noktasında ölçülen aydınlanma değeri $I(x,y)$ olmak üzere, aynı noktadaki yüzey parametreleri ile aydınlanma arasındaki ilişki, görüntü yansıtma denklemi (*image irradiance equation*)

$$I(x,y) = R(p,q) \tag{3}$$

ile belirlenir. Bir yüzeyin parametreleri ve ışık kaynağının yeri verildiği taktirde, bu yüzeye ilişkin aydınlatılmış resim (3) bağıntısından elde edilebilir. Ancak tondan-şekil-bulma probleminde bunun tam tersi sözkonusudur, yani aydınlatılmış resim verilmekte, yüzey parametrelerinin bulunması ve yüzeyin kurulması istenmektedir. Bu durumda sadece $R(p,q)$ ifadesi ve giriş resmi, yüzeyin belirlenmesinde yeterli olmamaktadır. Çünkü, görüntü yansıtma denkleminde iki bağımsız değişken, p ve q, vardır. Bu nedenle, yüzeyin elde edilmesinde yalnız başına (3) bağıntısı yeterli olmamaktadır ve sonsuz sayıda çözüm elde etmek mümkündür. Bilgisayarla görüde karşılaşılan bu tür problemlere kötü konumlu (*ill-posed*) problemler denir. Düzgünleştirme kuramı (*regularization theory*), bu tür kötü konumlu problemleri iyi konumlu (*well-posed*) problemler şekline dönüştürüp çözmeye yarayan yöntemdir ve bilgisayarla görüde yaygın olarak kullanılır. Düzgünleştirme ayrıca çözüm üzerinde düzlük gibi bazı sınırlamalar koymaya olanak sağlar.

TŞB problemlerinde kullanılan sınırlamalar aşağıda kısaca özetlenmiştir:

Aydınlanma sınırlaması (*brightness constraint*), görüntü yansıtma denkleminin bir sınırlama şeklinde ifade edilmesi sonucu elde edilmiştir ve özgün görüntü ile TŞB'den bulunan görüntü arasında

$$\int\int (I(x,y) - R(p,q))^2 \, dx\, dy \tag{4}$$

olarak tanımlanan hatanın enaza indirgenmesi şeklinde ifade edilir. İlave koşullar olmaksızın, sadece aydınlanma sınırlaması kullanılarak, problemin tek çözümünün bulunması olası değildir. Bu nedenle ek sınırlamalara gereksinim vardır.

Düzlük sınırlaması (*smoothness constraint*), ardışıl TŞB yönteminde, problemin tek çözümlü olmasını sağlamak amacıyla, kurulacak yüzeyin düz olması şartını sağlamak için

$$\min_{p,q} \iint (p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy \tag{5}$$

ifadesi kullanılır. Burada p_x, p_y, q_x ve q_y yüzey parametrelerinin kısmi türevlerini göstermektedir.

Integralin alınabilmesi sınırlaması (*integrability constraint*)

$$\min_{p,q} \iint ((z_x - p)^2 + (z_y - q)^2)\, dx\, dy \tag{6}$$

bağıntısı ile tanımlanır. Burada z_x ve z_y, yüzeyin kısmi türevlerini göstermektedir. Bu sınırlama, TBŞ yöntemiyle elde edilen yüzey parametreleri (p, q)'nun fiziksel olarak anlamlı bir yüzeye karşı gelmesini sağlamak amacıyla kullanılmaktadır.

Türev farkları sınırlaması (*intensity gradient constraint*), TŞB'den elde edilen resmin kısmi türevlerinin, giriş resminin kısmi türevlerine yakın olması koşulunu sağlamak amacıyla kullanılır. I_x ve I_y giriş resminin kısmi türevleri ve R_x ve R_y ise algoritmadan elde edilen resmin kısmi türevlerini göstermek üzere, bu sınırlama

$$\min_{p,q} \iint ((R_x - I_x)^2 + (R_y - I_y)^2)\, dx\, dy \tag{7}$$

şeklinde tanımlanır.

Çeşitli kısıtlamaları en iyi sağlayan z yüzeyinin bulunması amacıyla kullanılabilecek bir sayısal yöntemde aşağıdaki adımlar izlenir:

1. Aranan çözüm z ile ideal çözüm arasındaki hatayı ifade etmek için bir F fonksiyonu seçilir ve bu fonksiyonun Ω görüntü uzayındaki integralini,

$$\mathcal{I}(x, y) = \iint_{\Omega} F(x, y, z, ...)\, dx\, dy$$

   enaza indirgeyen z bulunmaya çalışılır.

2. F fonksiyonuna, z'nin Ω üzerinde sağlaması istenen sınırlamalar konulur.

3. Problemin tek çözümü yoksa, $\mathcal{I}$ ifadesine uygun bir düzeltme terimi ilave edilir.

4. z yüzeyinin sağladığı ve $\mathcal{I}$ fonksiyonelini enaza indiren Euler eşitlikleri belirlenir.

5. Tek çözüm elde edebilmek için sınır koşullarının neler olması gerektiği belirlenir.

6. Sınırlı değişimler (*finite difference*) yöntemi kullanılarak Euler eşitlikleri için ayrık domende bir yaklaşıklık bulunur.

7. Euler eşitliklerinin ayrık çözümüne yakınsayan bir sayısal yöntem bulunur.

Tezin ilk bölümünde, bilgisayar görü problemleri genel olarak tanıtılmıştır. Bu problemlerden birisi olan ve giriş resmindeki tonlamalardan yararlanarak resimdeki objenin yüzey özelliklerinin bulunmasını sağlayan tondan-şekil-bulma probleminden genel olarak bahsedilmiştir. Literatürdeki temel TŞB yöntemleri incelenmiş ve bu yöntemler kısaca karşılaştırılmıştır.

İkinci bölümde, yüzeyin fiziksel özellikleri tarafından belirlenen yansıtma modelleri ve Tondan Kaynak Yerini Bulma (*Source from Shading*) tekniği incelenmiştir. Işık kaynağının yerinin belirlenmesi probleminde, giriş resmindeki aydınlanma bilgisi kullanılarak, ışık kaynağının konumunu 3-Boyutlu resim uzayında belirleyen açıların, (σ, τ), bulunmasına çalışılmaktadır. Bu amaçla, Pentland [3], Lee ve Rosenfeld [4] ve Zheng ve Chellappa [5] tarafından geliştirilen üç yöntem incelenmiştir. Bu üç yöntem için, ışık kaynağının yerinin belirlenmesinde kullanılan sonuç bağıntılar, bölüm sonunda tablo halinde verilmiştir.

Üçüncü bölümde, tondan-şekil-bulma problemi tüm ayrıntıları ile incelenmiştir. Burada, önce bilgisayar görü problemlerinde kullanılan ve kötü durumlu problemleri iyi konumlu problemler haline dönüştürerek çözülmesini sağlayan düzgünlük kuramı ele alınmıştır. Yüzey parametrelerinin ifade edilmesi, yansıtma haritası ve görüntü yansıtma eşitliği ile ilgili açıklamalardan sonra, değişimlerin hesaplanması (*calculus of variations*) yaklaşımı kullanılarak TŞB yönteminin ardışıl olarak çözülmesi problemi incelenmiştir. Tondan-şekil-bulma problemlerinde kullanılan sınırlamaların açıklanmasından sonra, literatürdeki temel TŞB metodları özetlenmiştir. Ayrıca bu bölümde, tondan-şekil-bulma problemi ile ilişkili olarak, Fotometrik Stereo'dan Şekil bulma (FSŞB) tekniği de incelenmiştir. TŞB ile FSŞB arasındaki fark, TŞB'de bir tane giriş resmi kullanılmasına karşın, FSŞB'de birden fazla giriş resmi kullanılabilmektedir. Bu giriş resimleri, bakış doğrultuları aynı kalmak üzere, ışık kaynağının konumu değiştirilerek elde edilmektedir. FSŞB yönteminde birden fazla giriş resminin kullanılması sayesinde, resimdeki obje hakkında daha fazla bilgiye sahip olunmaktadır.

Dördüncü bölümde, TŞB problemlerinde kullanılan sınırlamaların çözüm üzerindeki etkileri incelenmektedir. Bu amaçla, sınırlamaların tümünü içeren

$$F(p,q,z) = F_1(p,q) + \lambda F_2(p,q) + \mu F_3(p,q,z) + \beta F_4(p,q) \qquad (8)$$

enerji fonksiyoneli tanımlanmıştır. Buradaki λ, μ ve β parametreleri sabit sayılar olup kendileri ile ilişkili sınırlamaların enerji fonsiyoneli içindeki ağırlıklarını belirlemektedirler. Yukarıda tanımlanan enerji fonksiyonelini enaza

indirme problemi çözülerek, ardışıl bir TŞB yöntemi elde edilmiş ve kullanılmıştır. Her bir sınırlamanın çözüm üzerindeki etkilerini görebilmek amacıyla, algoritma parametreleri λ, μ ve β'nın değişik kombinasyonları kullanılarak TŞB çözümleri elde edilmiştir. Her durumda, toplam enerji (F), ardışıl adımların fonksiyonu olmak üzere grafik olarak çizdirilmiştir. Gerçek yüzey parametreleri (p^*, q^*) ve gerçek yüzey z^* ile, TŞB'den bulunan yüzey parametreleri (p, q) ve yüzey z karşılaştırılmış, ortalama yüzey parametre hataları (p-q hataları) ve yüzey hataları (z hataları) çizdirilerek sonuçların kolaylıkla değerlendirilmesi sağlanmıştır. Bununla birlikte, TŞB'den elde edilen yüzey parametreleri kullanılarak, bu çözümlere ilişkin aydınlatılmış resimler elde edilmiş ve sınırlamaların bu resimler üzerindeki etkileri incelenmiştir. Diğer yandan, elde edilen z yüzeylerinin 1-Boyutlu kesitleri çizdirilerek, sınırlamaların z üzerindeki etkileri görülmeye çalışılmıştır.

Deneysel sonuçlardan, sınırlamaların çözüm üzerindeki etkileri konusunda elde edilen sonuçlar şu şekilde özetlenebilir:

Düzgünlük sınırlaması, TŞB probleminin tek çözümlü olmasını sağlamak amacıyla kullanılmasına rağmen, sayısal çözüm yönteminin yakınsaması açısından kritik öneme sahiptir. Bu terimin çarpanı olan λ parametresinin çok küçük değerleri için algoritmanın ıraksaması olasıdır. Diğer taraftan, λ'nın büyük değerleri, elde edilen yüzeylerin aşırı düzlenmesine neden olmaktadır. Bu düzleme etkisi, TŞB'den bulunan yüzeylerin aydınlatılması ile elde edilen resimlerin bulanık olmasına neden olmaktadır ve bu bulanıklığın miktarı λ'nın sayısal değerine bağlı olarak değişmektedir.

Integralin alınabilmesi sınırlaması, düzgünlük sınırlaması gibi, TŞB probleminde çözümün bulunmasında önemli etkilere sahiptir. Bu sınırlamanın enerji fonksiyoneli içindeki ağırlığını belirleyen parametre μ'dür ve μ'nün çok küçük değerleri, sayısal TŞB yöntemiyle bulunacak çözümlerin hatalı olmasına sebep olabilir. Diğer yandan, μ'nün büyük değerleri için aşırı düzlenmiş çözümler elde edilmektedir. Bu sonuçlardan anlaşılacağı gibi, doğru çözümlere ulaşılabilmesi için μ parametresinin değerinin iyi belirlenmesi gerekmektedir.

Türev farkları sınırlamasının etkisi, giriş resminin türevleri (I_x,I_y) ile TŞB yöntemiyle elde edilen aydınlatılmış resmin türevlerinin (R_x,R_y) karşılaştırılması sonucu belirgin olarak ortaya çıkmaktadır. Bu sınırlamanın katsayısı olan β'nın sıfır olduğu durumda bulunan kısmi türevler, R_x ve R_y, sırasıyla I_x ve I_y ile karşılaştırıldığında, bunların birbirlerine benzemedikleri görülmektedir. Oysa, β'nın sıfırdan farklı olduğu durumda bulunan R_x ve R_y'nin giriş resminin türevlerine oldukça benzediği deneysel sonuçlardan kolaylıkla görülmektedir. Diğer taraftan ardışıl çözüm yönteminde, β'nın büyük değerleri için bulunan çözümlerdeki p-q hatası ve z hatası, başlangıçta düşük olmasına rağmen, iterasyonların devam etmesi durumunda artmaktadır.

Bu sonuçlardan da anlaşılacağı gibi, TŞB probleminde kullanılan sınırla-

maların çözüm üzerinde kritik etkileri vardır. Bu sınırlamaların ağırlıklarını belirleyen algoritma parametreleri λ, μ ve β'nın ancak uygun değerleri için TŞB yönteminden başarılı sonuçlar alınması olasıdır.

Beşinci bölümde, uyarlanır düzleme (*adaptive smoothing*) [6] yönteminin TŞB problemine uygulanması konusunda çalışılmıştır. Dördüncü bölümde, düzleme sınırlamasının, TŞB'den bulunan yüzeylerin oldukça fazla düzlenmesine sebep olduğu sonucu belirlenmişti. Bu istenmeyen etkiyi ortadan kaldırmak için, uyarlanır düzleme yöntemi TŞB problemine uygulanmıştır. Bu amaçla, enerji fonksiyoneli olarak

$$F(p,q,z) = F_1(p,q) + F_2'(p,q) + \mu\, F_3(p,q,z) + \beta\, F_4(p,q) \tag{9}$$

ifadesi tanımlanmıştır. Burada F_1, F_3 ve F_4 terimleri daha önce tanımlandığı gibi, sırasıyla aydınlanma sınırlaması, integralin alınabilmesi sınırlaması ve türev farkları sınırlamasıdır. Düzlük sınırlaması ise

$$F_2'(p,q) = \iint \lambda(x,y)\,(p_x^2 + p_y^2 + q_x^2 + q_y^2)\,dx\,dy \tag{10}$$

şeklinde değiştirilmiştir. Dikkat edilirse, burada $\lambda(x,y)$, uzamsal koordinatların bir fonksiyonudur ve ardışıl TŞB algoritmasında, bu fonksiyonun (x,y) noktasındaki değeri

$$\lambda_{new}(x,y) = \begin{cases} \mathcal{F}(x,y,\lambda_{old}) & \text{if } c(x,y) > 0 \text{ and } \lambda_{old} > \lambda_{min} \\ \lambda_{old}(x,y) & otherwise \end{cases} \tag{11}$$

ile hasaplanır. Burada kullanılan $\mathcal{F}(x,y,\lambda_{old})$ fonksiyonu aşağıdaki gibi tanımlanmıştır:

$$\mathcal{F}(x,y,\lambda_{old}) = (1 - e^{-\frac{c(x,y)}{V_T}})\,\lambda_{min} + (e^{-\frac{c(x,y)}{V_T}})\,\lambda_{old}(x,y). \tag{12}$$

Bu fonksiyondaki $c(x,y)$, kontrol işaretidir ve $abs()$ mutlak değeri göstermek üzere, $c(x,y) = abs(I(x,y) - R(p,q))$'den hesaplanır. V_T, üstel fonksiyonun düşme hızını kontrol eden bir zaman sabitidir. λ_{min} ise, önceden belirlenen ve λ'nın alabileceği en küçük değeri gösteren bir sabittir. λ'nın yeni değerlerinin hesaplanması için kullanılan $\mathcal{F}(x,y,\lambda_{old})$ fonksiyonu, üstel olarak azalan bir fonksiyondur ve

$$\lim_{c(x,y)\to 0} \mathcal{F}(x,y,\lambda_{old}) = \lambda_{old}(x,y) \quad \text{ve} \quad \lim_{c(x,y)\to\infty} \mathcal{F}(x,y,\lambda_{old}) = \lambda_{min} \tag{13}$$

özelliklerine sahiptir. (9)'daki enerji fonksiyonelini enaza indirme problemi çözülerek, uyarlanır TŞB yöntemi için ardışıl çözüm yöntemi elde edilmiş ve deneysel sonuçların bulunmasında kullanılmıştır.

Bu bölümde ayrıca, uyarlanır düzleme yönteminin FSŞB problemine uyarlanması konusunda da çalışılmıştır. Bu amaçla

$$F(p,q,z) = F_1^{'}(p,q) + F_2^{'}(p,q) + \mu\, F_3(p,q,z) + \beta\, F_4(p,q) \qquad (14)$$

enerji fonksiyoneli kullanılmıştır. Burada F_3 ve F_4 terimleri surasıyla integralin alınabilmesi sınırlaması ve türev farkları sınırlamasıdır. $F_2^{'}$ fonksiyonu (10)'da verildiği gibidir ve $F_1^{'}$ ise

$$F_1^{'}(p,q) = \int\int \{(I(x,y) - R(p,q))^2 + (\hat{I}(x,y) - \hat{R}(p,q))^2\}\, dx\, dy \qquad (15)$$

olarak tanımlanmıştır. Buradaki $I(x,y)$ ve $\hat{I}(x,y)$, bakış doğrultuları aynı fakat aydınlanma doğrultuları farklı iki görüntüyü göstermekte, $R(p,q)$ ve $\hat{R}(p,q)$ ise bu iki görüntüye karşı gelen yansıtma haritalarını ifade etmektedir. (14)'deki enerji fonksiyonelini enaza indirme problemi çözülerek, uyarlanır TŞB ve FSŞB yöntemlerinin birleşiminden oluşan ardışıl sayısal çözüm yöntemi geliştirilmiş ve bu yöntem kullanılarak deneysel sonuçlar elde edilmiştir.

Beşinci bölümdeki deneysel sonuçlardan görülmüştür ki, uyarlanır düzleme yönteminin kullanılması, uyarlanır olmayan TŞB yönteminden elde edilen yüzeylerdeki istenmeyen düzlenme etkisini ortadan kaldırmış ve daha başarılı sonuçlar bulunmasını sağlamıştır. Uyarlanır TŞB yöntemi ile fotometrik stereo tekniğinin birleştirilmesi ve elde edilen sayısal TŞB çözüm yönteminin kullanılmasının, sonuçların kalitesini artırmak ve hataları azalmak açısından etkili olduğu deneysel olarak gösterilmiştir.

CHAPTER 1.

INTRODUCTION

The first processing stage in computer vision, also called *early vision*, is to recover 3-Dimensional properties of surfaces from the 2-Dimensional images. Typical 3-D properties are the distance between the surface and the viewer, surface orientation, structure, texture, reflectance and motion parameters obtained from a temporal sequence of images. With the development of computer vision, early vision problems have been formulated rigorously and given them familiar names, such as *shape from shading, structure from stereo, structure from texture, edge detection, visual interpolation, computation of optical flow.* The computational modules that solve them constitute together the core of early vision, and provide spatial and geometrical information about the visible part of 3-D world. The results of this first stage of processing are then used for *higher* level tasks such as navigation in the environment, manipulation of objects, object recognition and also reasoning about objects. Unlike high level vision, early vision is mostly considered as a *bottom-up* set of processes that do not rely upon specific high-level information about the scene to be analysed. These different modules of early vision can be analysed independently at least to a first approximation. Their most natural implementation is in terms of distinct pieces of hardware, whose outputs will be integrated at a later stage.

Although an object in the world is 3-D, its image projected onto the retina is essentially 2-D. The reduction in the image dimension which is a result of projecting of the 3-D world into a 2-D space causes the main lack of information for the recovery problems in early vision. Then, the problem of reconstructing the shapes of 3-D objects from their images which is an inverse problem are regarded as ill-posed; that is, the solution may not exist or it is not unique or

does not continuously depend on data. In order to solve ill-posed problems, a priori information about generic properties of the solution must be used. Regularization theory is a set of techniques that have been developed for this reason.

In early vision, the techniques to recover shape are called shape-from-X techniques. Shape-from-shading (SFS) [1, 7] we consider in this thesis is one of the classic research problems in computer vision. SFS deals with the recovery of surface orientation and surface shape (height) from the gradual variation of shading in the image. It is most often used in conjunction with other computer vision algorithm such as shape from stereo or shape from contour, since it provides information that is complementary to these techniques. Shape-from-shading has also been used widely in the field of *photoclinometry*, where astrogeologists use it for planetary terrain mapping from satellite photographs.

Since a single measurement of image brightness provides only limited information about surface orientation, additional constraints such as that the gradient vary smoothly must be used in order to find a unique solution. On the other hand, using multiple images which are taken with different lighting supplies additional information for robust surface reconstruction. Additional images can also allow us to recover further unknown parameters, such as the albedo of the surface.

Shape-from-photometric-stereo [8] is another method for shape recovery in computer vision. The difference between shape-from-photometric-stereo and SFS is in the number of input images. Shape-from-photometric-stereo recovers shape from *multiple* intensity images of the same scene generated using fixed viewing direction and different light source directions; while SFS provides the shape estimate from a *single* image.

The next section briefly describes the shape-from-shading problem and reviews the recently reported studies on this subject. In the last section of this chapter, we give an outline of our objectives.

Chapter 2 introduces background knowledge related to reflectance models and source from shading which deals with recovery of the light source location

is presented. We explain Pentland's original approach for light source computation, its refinements by Lee and Rosenfeld, and improvements of both methods by Zheng and Chellappa.

In chapter 3, the shape-from-shading problem is discussed in detail. Regularization theory and its application to the SFS problem are considered. After a review of SFS formulation, the constraints used in this problem are presented. A total of eight SFS methods are summarized. Another shape recovery technique, photometric stereo, is introduced.

In chapter 4, effects of the SFS constraints on solution are explored. All constraints used in the shape-from-shading problem are combined into an energy functional. Solving this minimization problem, a general purpose iterative SFS algorithm is derived. Changing the relative weight of each constraint in the energy functional by setting proper values to these parameters, we explore the effects of the SFS constraints on the solution. The experimental results are reported.

Chapter 5 presents the adaptive shape-from-shading approach in which standard regularization is modified to be able to obtain an iterative scheme that improves the results of the present methods. In this adaptive SFS method, smoothness is controlled spatially over the image space to reduce the over-smoothing effect of the smoothness constraint on the solution. In order to improve the quality of the reconstruction, our adaptive SFS method and photometric stereo technique are integrated. A hierarchical implementation of the adaptive SFS scheme is also presented. Typical results obtained for synthetic and real images are given to illustrate the advantages of our adaptive approach.

In chapter 6, the major issues that have been raised and the conclusions of our study are summarized. The major contributions to the literature and the future research are also covered.

1.1 Review of Shape from Shading

The variation in brightness due to changes in surface orientation across a surface, or shading, is an important visual cue for interpreting 3-D scenes. Shading in an image is due to the interaction of four factors: The shape of the surface, the illumination, the reflecting characteristics of the surface material and the image projection. The shading information is used by shape-from-shading techniques to recover 3-D surface properties. In SFS, the basic idea is to model the image brightness as a function of surface shape and then solve the inverse problem-reconstructing the surface from the given intensity image. One of the common features of inverse problems in early vision is that they are often ill-conditioned or ill-posed (An ill-posed problem is the one that does not have any solution, does not have a unique solution, or has a solution that is very sensitive to the given data). So, it is necessary to impose some constraints to get a unique solution.

In computer vision, the recovered shape obtained from shape recovery techniques can be expressed in several ways. One way to do this is to specify a unit vector, $\vec{N}$, perpendicular to the local tangent plane. Another way is to specify the components p and q of the surface gradient. These are the partial derivatives of surface height z above some reference plane perpendicular to the optical axis, that is, $p = (dz/dx)$ and $q = (dz/dy)$. The two notations are associated by the equality

$$\vec{N} = \frac{1}{\sqrt{(1 + p^2 + q^2)}}(-p, -q, 1)^T. \tag{1.1}$$

A third way of specifying surface orientation is to use of surface *slant* (σ) and *tilt* (τ) angles. The relationship of slant and tilt to the unit normal is given by

$$\vec{N} = (\sin\sigma\cos\tau, \sin\sigma\sin\tau, \cos\sigma)^T. \tag{1.2}$$

Each notation for surface orientation has its own advantages and the above three as well as some others have been used in the literature.

Under the assumption that the viewer and the light sources are far away from the objects being imaged, we can introduce the *reflectance map* to specify

the relation between surface orientation and brightness. It basically encodes information about surface reflectance properties and light-source distributions. A graphic representation of the reflectance map, in the so-called *gradient space* is also possible if we plot contours of constant brightness in the pq-plane. The reflectance map can be determined experimentally or derived from the bidirectional reflectance distribution function (BRDF) which specifies the reflectance properties of a surface. If we use surface gradient, (p, q), as a way of specifying surface orientation, we can use the form (p, q) of reflectance map.

If (x, y) is a particular point in the image, we measure the image irradiance as $I(x, y)$ at that point. It is proportional to the radiance at the corresponding point on the image surface. If the surface gradient is (p, q) at the observation point, then the radiance there is $R(p, q)$. Then, we obtain the so-called *image irradiance equation* as

$$I(x, y) = R(p, q), \tag{1.3}$$

where $I(x, y)$ is the image irradiance at the point (x, y), while $R(p, q)$, the reflectance map, is the scene radiance of a surface patch with orientation specified by the partial derivatives $p = dz/dx$ and $q = dz/dy$ of surface height $z(x, y)$ above some reference plane. Actually, image irradiance is not equal to scene radiance, only proportional to it, but the proportionality factor is usually ignored, because it is assumed that some calibration process normalizes one of these quantities so that it is commensurate with the other. The image irradiance equation is fundamental to the methods for solving shape-from-shading problems. The task is to find $z(x, y)$ given the image and the reflectance map $R(p, q)$.

The image irradiance equation is a nonlinear first-order partial differential equation. Then, the shape-from-shading problem is regarded to be mathematically equivalent to a non-linear first-order partial differential equation in surface elevation. The shaded image provides only a single constraint on the surface gradient (p, q) at each point in the image. Additional constraints, such as boundary conditions and singular points, are needed to ensure that there is a unique solution. Hence we can reformulate the problem as one of finding a surface orientation field that minimizes the integral of the brightness error[1].

[1]Brightness error $= I(x, y) - R(p, q)$.

The calculus of variations is then employed to derive the appropriate Euler equations on which iterative schemes can be based. A number of iterative schemes have been proposed to solve shape-from-shading problem based upon the calculus of variations [9, 10, 11, 12].

There are two approaches for finding a solution to the shape-from-shading problem as *global* and *local*. Global approaches further can be divided into global minimization and global propagation approaches. Global minimization approaches obtain the solution by minimizing an energy function. However global propagation methods propagate the shape information from known surface points (e.g., singular points [2]) to the whole image. On the other hand, only a small patch of an image is used to determine the shape of a surface on local approaches.

The shape-from-shading problem was formulated and solved by Horn [7, 13] using the first global propagation technique called characteristic strip. A *characteristic strip* is a line in the image along which the surface depth and orientation can be computed if these quantities are known at the starting point of the line. Horn's method constructs initial surface curves around the neighborhoods of singular points using a spherical approximation. The shape information is propagated simultaneously along the characteristic strips outwards, assuming no crossover of adjacent strips. In order to get a dense shape map, new strips have to be interpolated when neighboring strips separate too much. In characteristic strip method, the partial differential equation in surface elevation fundamental to the problem was converted to an equivalent set of five ordinary differential equations called the characteristic strip equations. Algorithms based on numerical solution of the discrete approximations of these equations are inherently sequential in nature and have difficulty with unavoidable noise in the image data.

One of the earlier global minimization approaches was suggested by Ikeuchi and Horn [11]. Since each surface point has two unknowns for the surface normal, and each pixel in the image provides one gray value, therefore image gray levels alone are not enough to recover the shape. To overcome this, Ikeuchi and Horn introduced two constraints: The brightness constraint and the smoothness

[2]Singular points are the points with maximum intensity in the image.

constraint. The brightness constraint requires that the reconstructed shape will produce the same brightness as the input image at each surface point, while smoothness constraint forces the gradient of the surface to change smoothly. The shape was computed by minimizing an energy functional which consists of the above two constraints. Also using these two constraints, Brooks and Horn [14] minimized the same energy function, in terms of surface normal instead of surface gradient. Frankot and Chellappa [9] enforced the integrability in Brooks and Horn's algorithm in order to recover integrable surfaces (surfaces for which $z_{xy} = z_{yx}$). Surface slope estimates from the iterative scheme were expressed in terms of a linear combination of a finite set of orthogonal Fourier basis functions. The enforcement of integrability was done by projecting the nonintegrable surface slope estimates onto the nearest (in terms of distance) integrable surface slopes. This projection was fulfilled by finding the closest set of coefficients which satisfy integrability in the linear combination. Their results showed improvements in both accuracy and efficiency. Later, Horn [12] also replaced the smoothness constraint in his approach with an integrability constraint. The major problem with Horn's method is its slow convergence. Szeliski [15] speed it up using a hierarchical basis pre-conditioned conjugate gradient descent algorithm. Based on the geometrical interpolation of Brooks and Horn's algorithm, Vega and Yang [16] applied heuristics to the variational approach so that the stability of Brooks and Horn's algorithm was improved.

Instead of smoothness constraint, Zheng and Chellappa [5] introduced an intensity gradient constraint, which specifies that the intensity gradients of the reconstructed image and the input image are close to each other in both x and y directions. Leclerc and Bobick [17] solved directly for depth by using discrete formulation and employing a conjugate gradient technique. The brightness constraint and smoothness constraint were applied to ensure convergence, and a stereo depth map was used an initial estimate. Recently, Lee and Kuo [18] also proposed an approach to recover depth using the brightness and the smoothness constraints. They approximated surfaces by a union of triangular patches. Unlike Leclerc and Bobick's method, this approach did not require the depth from stereo as an initial value.

All of the above approaches deal with a single smooth surface. Malik and Maydan [19] developed the first solution for piecewise smooth surface. They

combined the line drawing and shading constraints in an energy functional, and recovered both surface normal and line labeling through the minimization of the energy function.

Oliensis [20] observed that the smoothness constraint is only needed at the boundaries if we have initial values at the singular points. Based on this idea, Dupuis and Oliensis [21, 22] formulated SFS as an optimal control problem, and solve it using numerical methods. Bichsel and Pentland [23] simplified Dupuis and Oliensis's approach and proposed a minimum downhill approach for SFS which converged in less than ten iterations.

Among the local approaches, two are by Pentland, one by Lee and Rosenfeld, and one bye Tsai and Shah. Pentland [3] recovered shape information from the intensity, and its first and second order derivatives. He used the assumption that the surface is locally spherical at each point. Under the same spherical assumption, Lee and Rosenfeld [4] computed the slant and tilt of the surface in the light source coordinate system through the first derivative of the intensity. A later approach by Pentland [24] used the linear approximation of the reflectance function in terms of the surface gradient, and applied a Fourier transform to the linear approximation to get a closed form solution for the depth at each point. Similar to Pentland's method, Tsai and Shah [25] applied the discrete approximation of the gradient first, then employed the linear approximation of the reflectance function in terms of the depth directly. Their algorithm iteratively recovered the depth at each point without using any global information.

Pentland's linear shape from shading has problems with images of quadratic surface reflectance. Therefore, Pentland [26] proposed photometric motion to solve for shape and reflectance. The images needed in his approach were taken at different time frames while the object was rotated. The quadratic component of the surface reflectance function was factored out by subtracting two of the images. The ratio of one of the images and the difference image was used to cancel out the albedo and obtain the surface shape. Therefore, at least two images were required for the shape recovery. This approach was also extended to three-image photometric motion by considering second derivatives in the discrete form.

None of the above methods deal with interreflection - the mutual illumination between surface facets. Nayar, Ikeuchi and Kanade [27] addressed the shape-from-interreflection problem using photometric stereo. They observed that the erroneous shape extracted by shape-from-photometric-stereo algorithms in the presence of interreflections, was shallower than the real shape. Therefore, they proposed a method to iteratively refine the shape. Their formulation of interreflection was based on Forsyth and Zisserman's result [28].

Some different approaches, recently applied to the shape-from-shading, are explained below.

A global shape-from-shading algorithm which develops a technique to merge local shape from shading results obtained around singular points into the complete shape using the mountaineers theorem was presented by Kimmel [29] and Bruckstein [30]. In their approach, a local shape-from-shading algorithm is used to recover the shape around each singular point. The algorithm inspects the behaviour of iso-height contours around each singular point. The contours are monitored from the singular point "outwards" until another singular point is encountered. An underlying assumption is that the shape to be recovered is a Morse Function. For such functions, according to the mountaineers theorem [31], the number of extrema located within a closed equal height contour of a smooth surface exceeds by one the number of saddle points within that contour. Therefore, when tracking iso-height contours that start as a small circle around an extremum, the first singular point that the extending contours meet must be a saddle point. This saddle point is the singular point whose height is closest to the height of the extremum. When the illumination is in the direction of the camera, the local shape recovered assuming the extremum is a minimum is the reflection of the shape recovered assuming the the extremum is a maximum, therefore the same saddle point will be found. This is not the case when the illumination comes from other directions and the closest saddle point may be different. In both cases the algorithm is run on all possible assignments of extrema as minima or maxima. The global algorithm merges the results of the local algorithm by merging two local surfaces which have the same closest saddle point. After two surfaces have been merged, the local algorithm extends the merged surface until another saddle point is met. This process continues until all local surfaces have been merged together and the global shape has

been recovered.

Shimshoni, Kimmel and Bruckstein [32] enhanced the results given in [29, 30] by providing the completeness and uniqueness obtained by the global shape-from-shading algorithm. Their approach is as follows: The results of the global algorithm can be described as a binary tree where the leaves are the extrema in the image and the inner nodes are the saddle points closest to the sub-surfaces represented by their children. There is a unique such tree for each image. Although the algorithm can merge the subsurfaces in different orders, the resulting tree and the shape recovered, will always be the same because the closest saddle point to an extremum or subsurface is unique by construction. They showed that the algorithm can recover a single surface when the singular points are correctly classified.

A wavelet-based approach to solving shape-from-shading is presented in [33]. The proposed method takes advantage of the nature of wavelet theory, which can be applied to efficiently and accurately represent "things," to develop a faster algorithm for reconstructing better surfaces. To derive the algorithm, the formulation of Horn [12], which combines several constraints into an objective function, is adopted. In order to improve the robustness of the algorithm, two new constraints are introduced into the objective function to strengthen the relation between an estimated surface and its counterpart in the original image. Thus, solving the SFS problem becomes a constrained optimization process. Instead of solving the problem directly by using Euler equation or numerical techniques, the objective function is first converted into the wavelet format. Due to this format, the set of differential operators of different orders which is involved in the whole process can be approximated with connection coefficients of Daubechies bases. In each iteration of the optimization process, an appropriate step size which will result in maximum decrease of the objective function is determined. After finding correct iterative schemes, the solution of the SFS problem will finally be decided.

It is generally agreed that individual visual cues are fallible and often ambiguous. This has generated a lot of interest in design of integrated vision systems which are expected to give a reliable performance in practical situations. The design of such systems is challenging since each vision module

works under a different and possibly conflicting set of assumptions. Therefore, Pankanti and Jain [34] have proposed and implemented a multiresolution system which integrates perceptual organization (grouping), segmentation, stereo, shape from shading and line labeling modules. The output of the integrated system is shown to be insensitive to the constraints imposed by the individual modules. The numerical accuracy of the recovered depth is assessed in case of synthetically generated data. They have qualitatively evaluated their approach by reconstructing geons from the depth data obtained from the integrated system.

Generalized Cylinders are a flexible, loosely-defined class of parametric shapes capable of modeling many real-world objects. Straight Homogeneous Generalized Cylinders (SHGCs) are an important subclass of Generalized Cylinders, whose cross-sections are scaled versions of a reference curve. Although there has been considerable research into recovering the shape of SHGCs from their contour, this work has almost exclusively involved methods that couple contour and heuristic constraints. A rigorous approach to the problem of recovering solid parametric shape from a single intensity view should involve at least two stages: 1) deriving the contour constraints and 2) determining if additional image constraints, e.g., intensity, can be used to uniquely determine the 3-D object shape. This methodology is also important for the recovery of object classes like tubes, where contour and heuristic constraints are shown to be insufficient for shape recovery. Gross and Boult [35] follow the approach just described. First, they prove that SHGC contours generated under orthography have exactly two degrees of freedom. Next, they show that the remaining free parameters can be resolved by using reflectance-based constraints, without knowledge of the number of light sources, their positions, intensities, the amount of ambient light or the surface albedo. Finally, the reflectance-based recovery algorithm is demonstrated on both synthetic and real SHGC images.

1.2 Statement of Objectives

Recovering the surface of an object from its shaded image, known as the shape-from-shading problem, is one of the fundamental problems in computer

vision. There are two main classes of algorithms for computing shape from a shaded image: Global methods and local methods.

In this thesis, we are mainly interested in global shape-from-shading methods. In the global methods, the shape is recovered by minimizing some cost functional involving constraints like smoothness. In these approaches, the variational calculus technique is used to iteratively compute the shape which is globally consistent. The global methods are complex but provide very close solutions to the accurate shape.

A difficulty in shape-from-shading is determining the boundary conditions of these constrained minimization problems. In general, boundary conditions are not known. Our objective here is to derive an iterative algorithm which does not need boundary conditions.

One of the important issue related to robustness is the selection of parameter values. These parameters determine the relative weights of the corresponding constraints in the algorithm. One of our objectives is to explore the effects of these parameter values on shape-from-shading solutions.

A drawback of using the regularization term which is needed to obtained a unique solution, is the smoothing effect on solutions. Our objective is to use adaptive smoothing to reduce oversmoothing along the reconstructed image and to improve the reconstruction.

In the global shape-from-shading problem, many iterations may be needed to recover the true surface shape. Because of that, these methods converge slowly. Our aim is to speed up the convergence of the global SFS algorithm using one of multigrid methods.

Another important issue related to the quality of reconstruction is using more than one input image. Since two or more measurement of the image brightness supply more information about the surface parameters (p, q) and surface height z, this helps to improve the performance of the iterative scheme. Our another objective is to integrate our adaptive shape-from-shading method and photometric stereo technique to improve the quality of reconstructed shapes.

CHAPTER 2.

SOURCE FROM SHADING

2.1 Introduction

Any analysis of image shading must be based on a mathematical model of the process of image formation. Points in a scene, when illuminated, reflect incident light in various directions. Light rays that are reflected to the direction of the sensor cause an image of the scene to be formed. Most machine vision problems involve the analysis of images formed in this manner. The intensity at any given point in the image is closely related to the reflectance properties of the corresponding point in the scene. Therefore, the prediction or the interpretation of image intensities requires modeling the surface reflection. Then, we can recognize an object or measure some of its properties by recovering information about its shape and reflectance from an image. The reflectance models have been used in computer vision is presented in the next section.

2.2 Reflectance Models in Vision

Depending on their physical properties, surfaces can be categorized as Lambertian, specular, hybrid, or more sophisticated surfaces. In this section, we will describe the reflectance models and discuss their properties related to shape from shading.

An ideal Lambertian surface is one that appears equally bright from all viewing directions and reflects all incident light, absorbing none. The brightness of a Lambertian surface is proportional to the energy of the incident light. The amount of light energy falling on a surface element is proportional to the area of the surface element as seen from the light source position (the foreshortened area). The foreshortened area is a cosine function of the angle between the surface normal and the light source direction. Therefore, Lambertian surface can be modeled as the product of the strength of the light source $\mathcal{B}$, the reflectance factor (albedo) of the surface η, and the foreshortened area $\cos\theta_i$ as follows

$$I_L = \mathcal{B}\,\eta\,\cos\theta_i = \mathcal{B}\,\eta\,(\vec{N}\cdot\vec{S}), \tag{2.1}$$

where I_L is the scene radiance and θ_i, the incident angle, is the angle between the surface normal $\vec{N} = (n_x, n_y, n_z)$ and the source direction $\vec{S} = (s_x, s_y, s_z)$.

Specularity only occurs when the incident angle of the light source is equal to the reflected angle. It is formed by two components: The specular spike and the specular lobe. The specular spike is zero in all directions except for a very narrow range around the direction of specular reflection. The specular lobe spreads around the direction of specular reflection.

The simplest method for specular reflection is described by the following delta function

$$I_S = \mathcal{B}\,\delta(\theta_s - 2\theta_r), \tag{2.2}$$

where I_S is the specular brightness, $\mathcal{B}$ is strength of the specular component, θ_s is the angle between the light source direction and the viewing direction, and θ_r is the angle between surface normal and the viewing direction. This model assumes that the lighting caused by specular reflection is only a single point which is not true in real life. Another model was developed by Phong [36]. It represents the specular component of reflection as powers of the cosine of the angle between the perfect specular direction and the viewing direction. This model is capable of predicting specularities which extend beyond a single point; however, the parameters have no physical meaning. A more refined model, the Torrance-Sparrow model [37], assumes that a surface is composed of small, randomly oriented, mirror-like facets. It describes the specular brightness as the

product of four components: Energy of incident light, Frensel coefficient, facet orientation distribution function and geometrical attenuation factor adjusted for foreshortening. On the basis of the Torrance-Sparrow model, Healey and Binford [38] derived a simplified model by using the Gaussian distribution as the facet orientation function and considering the other components as constant. It can be described as

$$I_S = \mathcal{K}\, e^{-(\frac{\alpha}{m})^2}, \tag{2.3}$$

where $\mathcal{K}$ is a constant, α is the angle between the surface normal $\vec{N}$ and the bisector of the viewing direction and source direction, and m indicates the surface roughness.

Most surfaces in the real world are neither purely Lambertian, nor purely specular, they are a combination of both. That is, they are hybrid surfaces. One straightforward equation for a hybrid surface is

$$I = (1-\omega)I_L + \omega I_S, \tag{2.4}$$

where I is the total brightness for the hybrid surface, I_L and I_S are the Lambertian brightness and specular brightness respectively and ω is the weighting factor.

Nayar, Ikeuchi and Kanade [39] proposed a reflectance model which consists of three components: Diffuse lobe, specular lobe and specular spike. The Lambertian model was used to represent the diffuse lobe, and the spike component of the Torrance-Sparrow model was used to describe the specular lobe, and the spike components of the Beckmann-Spizzichino model was used to describe the specular spike. The resulting hybrid model is given as

$$I = K_{dl}\, \cos\theta_i + K_{sl}\, e^{-\frac{\beta^2}{2\sigma^2}} + K_{ss}\, \delta(\theta_i - \theta_r)\delta(\phi_r), \tag{2.5}$$

where K_{dl}, K_{sl} and K_{ss} are the strengths of the three components, β is the angle between the surface normal of a micro-facet on a patch and the mean normal of this surface patch, and σ is its standard derivation. If we consider the surface normal being in the z direction, then (θ_i, ϕ_i) is the direction of incidence light in terms of the slant and tilt in 3-D, (θ_r, ϕ_r) is the direction of reflected light.

Although the Lambertian model is widely used because of its simplicity, it is a poor approximation to the diffuse component of rough surfaces. For a rough surface, the radiance increases as the viewer approaches the source direction. Oren and Nayar [40] derived a reflectance model for rough diffuse surfaces, taking into account complex geometrical effects of masking, shadowing and interreflection. Modeling rough surfaces, they developed a functional approximation of rough surfaces for both isotropic and non-isotropic surfaces, including uni-directional single-slope distribution, isotropic single-slope distribution and Gaussian distribution. A simplified qualitative model was derived by considering the relative significance of the various terms in the functional approximation

$$I = \cos\theta_i (A + B\, Max[0, \cos(\phi_r - \phi_i)] \sin\alpha \tan\beta), \tag{2.6}$$

where $A \approx \eta\,(\frac{1}{\pi} - 0.09\frac{m^2}{m^2+0.4})$, $B \approx \eta\,(0.05\frac{m^2}{m^2+0.18})$, (θ_i, ϕ_i) and (θ_r, ϕ_r) are the same as in the previous section, $\alpha = Max[\theta_i, \theta_r]$, $\beta = Min[\theta_i, \theta_r]$, η is the albedo value and m is the surface roughness. This model reduces to the Lambertian model when $m = 0$.

Clark [41] used perspective rather than orthographic projection in modeling reflectance. In his model, there is no requirement for the light source to be at infinity. Therefore, the reflectance function is

$$I(\vec{x}) = K \frac{R(\vec{S})}{|\, Z\vec{\chi} + \vec{t}\,|^2 |\, Z\vec{\chi}\,|^2}, \tag{2.7}$$

where $\vec{x} = (x, y)$ is the image coordinate vector, K is constant, R is the reflectance map, $\vec{S} = \frac{Z\vec{\chi}+\vec{t}}{|Z\vec{\chi}+\vec{t}|}$, which indicates the direction from the surface point to the light source, $\vec{t}$ is the location of the light source with respect to the coordinate system centered on the focal point of the camera, $\vec{\chi} = (\frac{x}{f}, \frac{y}{f}, -1)^T$, f is the focal length of the camera, and z is the depth. Since the distances from surface to camera and from light source to surface are both considered in this model, it yields a more realistic description of reflectance.

Hougen and Ahuja [42] have observed that the assumption of a single point source overly simplifies the model of the light source distribution. They approximated the light source distribution by a set of m distinct light source vectors,

$\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_m$, where $\vec{S}_k$ is the average value of $\vec{S}$ over a neighborhood angle of $\vec{S}_k$. By writing $\vec{S}_k$ as a product of its magnitude λ_k and unit direction $\vec{S}'_k$, the brightness equation can be expressed by

$$I = \eta \, (\lambda_0 \, R_0 + \sum_{k=1}^{m} \lambda_k \, R(\vec{N}, \vec{S}'_k)), \tag{2.8}$$

where $\eta \lambda_0 R_0$ is due to the contribution of ambient light, and R is the reflectance map which is independent of the magnitude of the light source.

Unlike the classical Lambertian reflectance model, Langer and Zucker [43] introduced the concept of *Shape from Shading in a Cloudy Day.* They claimed that under diffuse lighting, the radiance depends primarily on the amount of the diffuse source visible from each surface element, with the surface normal of secondary importance. Assuming the effect of mutual illumination can be ignored, the brightness at image point $\vec{x} = (x, y)$ is described as

$$I(\vec{x}) = \eta \, I_D \, \frac{1}{\pi} \int_{\nu(\vec{x})} (\vec{N}(\vec{x}) \cdot \vec{S}) \, d\Omega, \tag{2.9}$$

where η is the albedo, I_D is the illuminance from a uniform hemispheric light source, $\nu(\vec{x})$ is the set of unit directions in which the sky is visible from $\vec{x}$ and $d\Omega$ denotes an infinitesimal solid angle.

The above reflectance models attempt to remove one or more of the following constraints used in the simplification of the classical Lambertian model:

- The brightness is independent of the viewing direction,
- The illumination is from an infinite point source,
- The projection of the object onto the image plane is perspective.

Oren and Nayar's and Langer and Zucker's models emphasized the important effect of the viewing direction in the reflectance model. Hougen and Ahuja considered background illumination and more realistic lighting as compared to a single point source illumination widely used. Clark used the perspective projection in the shape recovery and his model does not even depend on any specific reflectance model. In short, all of those models overcome the over-simplification of the traditional Lambertian model which results in the elimination of the modeling error.

2.3 Estimation of Illuminant Direction and Albedo

Most current methods for obtaining shape from shading require known light source direction. Since the light source is assumed to be an infinite distance away from the image being viewed, the light source orientation is constant for all the surface points in the image. Then, one shaded image can provide enough information to estimate the light source direction.

One way of describing a light source direction is to use a three-dimensional vector $\vec{S} = (s_x, s_y, s_z)$. The other way is to use two angles called the slant (σ) and the tilt (τ), where the slant is defined as the angle between the surface normal and the direction toward the viewer and the tilt is defined as the angle between the projection of the surface normal on the image plane and the x-axis,

In shape-from-shading problem, with using the image irradiance equation, the imaging model is specified through a reflectance map [2, 44]

$$I(x, y) = R(p, q), \tag{2.10}$$

where $I(x, y)$ is the scene radiance at the point (x, y) in the image and $R(p, q)$ is the reflectance map function with $p = dz/dx$ and $q = dz/dy$ being the partial derivatives of height $z(x, y)$ with respect to the image coordinates. If we consider the Lambertian model, the reflectance map is

$$R(p, q) = \eta\,(\vec{N} \cdot \vec{S}), \tag{2.11}$$

where η is the albedo, $\vec{N} = (-p, -q, 1)/\sqrt{1 + p^2 + q^2}$ is the surface normal at $(x, y, z(x, y))$ and $\vec{S} = (s_x, s_y, s_z)$ is the direction toward the light source. $\vec{S}$ is constant if we assume the illuminant to be from a point light source far away from the scene.

Under the Lambertian surface, point light source and uniform albedo assumptions, the reflectance map is determined by the parameters $\vec{S}$ and η, i.e., the illuminant direction and surface albedo. The problem of estimating these parameters, and hence estimating the reflectance map, was first attempted by

Pentland [3]. Pentland estimates the illuminant direction from the distribution of image derivatives. By assuming an umbilical surface and isotropic surface normal, a maximum-likelihood analysis was performed to estimate the tilt and the slant angles of the light source. This approach was later extended to the Fourier domain [3, 24, 45]. The basic idea underlying Pentland's method is that, since surface orientation is a random variable over the whole image for most scenes, so both surface normal $\vec{N}$ and the change of the surface normal $d\vec{N}$ are isotropically distributed. Therefore, if we consider any direction (dx_I, dy_I) on the image plane, the z component of the expected value of $d\vec{N}$, $\bar{dn}_z$, is zero. Thus the derivative of the brightness equation

$$E\{dI\} = d\bar{I} = \eta(s_x \bar{dn}_x + s_y \bar{dn}_y), \tag{2.12}$$

where E indicates the expected value.

For a sphere, for example, $z(x,y) = \sqrt{r^2 - x^2 - y^2}$, $z_x = -\frac{x}{z}$, $z_y = -\frac{y}{z}$, hence $\vec{N} = -\frac{1}{r}(x, y, z)$. Consequently, it can be shown that the x and y components of the derivative of normal in any direction θ are given by $dn_x = -\frac{1}{r}\cos\theta$ and $dn_y = -\frac{1}{r}\sin\theta$. Let $(\cos\theta, \sin\theta) = (dx_I, dy_I)$, then $\bar{k}dx_I = \bar{dn}_x$ and $\bar{k}dy_I = \bar{dn}_y$, where $\bar{k} = \frac{1}{r}$ is the mean projected surface curvature, which is the same in all directions using the locally spherical assumption. Repeating the above process in M different directions $(dx_i, dy_i)(i = 1, \ldots, M)$, the regression model can be describe as

$$\begin{pmatrix} d\bar{I}_1 \\ d\bar{I}_2 \\ \vdots \\ d\bar{I}_M \end{pmatrix} = \begin{pmatrix} dx_1 & dy_1 \\ dx_2 & dy_2 \\ \vdots & \vdots \\ dx_M & dy_M \end{pmatrix} \begin{pmatrix} \tilde{s}_x \\ \tilde{s}_y \end{pmatrix}, \tag{2.13}$$

where $d\bar{I}_i$ is the average of the intensity change along the image direction (dx_i, dy_i) and $\tilde{s}_x, \tilde{s}_y$ are the unnormalized x and y components of the τ estimate.

Solving the above we get

$$\begin{pmatrix} \tilde{s}_x \\ \tilde{s}_y \end{pmatrix} = (B^T B)^{-1} B^T \begin{pmatrix} d\bar{I}_1 \\ d\bar{I}_2 \\ \vdots \\ d\bar{I}_M \end{pmatrix}, \tag{2.14}$$

where B is the matrix of directions (dx_i, dy_i) as shown in (2.13). A typical choice for the (dx_i, dy_i) are the eight directions in the image plane: Two in the horizontal, two in the vertical direction and four along the diagonals. Then, $M = 8$ and the direction matrix will be

$$B = \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & -1 & -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & -1 & -\frac{\sqrt{2}}{2} \end{pmatrix}^T. \tag{2.15}$$

The tilt of the illuminant, τ, can be computed by

$$\tau = \arctan\left(\frac{\tilde{s}_y}{\tilde{s}_x}\right), \tag{2.16}$$

and the slant, σ, of the light source direction is computed as

$$\sigma = \arccos\sqrt{1 - s_x^2 - s_y^2}. \tag{2.17}$$

Taking the expected value of the square of intensity derivative $E\{dI^2\}$, and cancelling out of the common terms between $E\{dI^2\}$ and $E\{dI\}^2$ by subtracting one from the other, we have the relation $E\{dI^2\} - E\{dI\}^2 = \eta^2\bar{k}^2$. Since $\tilde{s}_x = \eta\bar{k}s_x$, $\tilde{s}_y = \eta\bar{k}s_y$, by introducing $k = B\bar{k} = \sqrt{E\{dI^2\} - E\{dI\}^2}$, the equation for the slant of the light source can be simplified as

$$\sigma = \arccos\sqrt{1 - \frac{\tilde{s}_x^2 + \tilde{s}_y^2}{k^2}}. \tag{2.18}$$

Instead of taking intensity derivatives along a number of directions, Lee and Rosenfeld [4] considered only the derivatives along the x and y directions. They approximated the surface geometry by a spherical patch in a local region, so their method was also based on an isotropic distribution of the surface orientation. Since the image of a sphere is symmetric about the projection of the light source vector in the image plane, the average direction of the intensity gradient must be parallel to this projection. This gives

$$\frac{E\{I_y\}}{E\{I_x\}} = \frac{s_y}{s_x}. \tag{2.19}$$

So the tilt τ of the illuminant direction can be estimated by

$$\tau = \arctan\left(\frac{E\{I_y\}}{E\{I_x\}}\right), \tag{2.20}$$

where I_x, I_y are the first order partial derivatives of image intensity with respect to the image coordinates and the expectations are taken over the whole image.

Considering the sampling distribution for the slant and expected values of intensity and intensity squared, the slant σ of the illuminant direction satisfies

$$E\{I\} = \frac{4\eta}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{1+\cos\sigma} \tag{2.21}$$

$$E\{I^2\} = \frac{\eta^2}{4}(1+\cos\sigma), \tag{2.22}$$

where the expectations are taken over the whole image.

To compute η and σ from (2.21) and (2.22), we first solve for σ from

$$\frac{E\{I\}}{\sqrt{E\{I^2\}}} = \frac{8}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{(1+\cos\sigma)^{\frac{3}{2}}}. \tag{2.23}$$

After the slant angle σ is estimated, the albedo of the surface can be determined by

$$\eta = \frac{1}{a^2+b^2}\left(aE\{I\}+\sqrt{bE\{I^2\}}\right), \tag{2.24}$$

where

$$a = \frac{4}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{1+\cos\sigma} \quad \text{and} \quad b = \frac{1+\cos\sigma}{4}. \tag{2.25}$$

Zheng and Chellappa [5] modified Lee and Rosenfeld's method by considering not only the area of the illuminated portion in the integral, but also the area of the portion in shadow. Although the shadow does not contribute to the total intensity, it does contribute to the area computation in order to correctly calculate the mean intensity value over the whole image. After the modification, the formulas are

$$E\{I\} = \frac{2\eta}{3\pi}\left[(\pi-\sigma)\cos\sigma+\sin\sigma\right] \tag{2.26}$$

$$E\{I^2\} = \frac{\eta^2}{8}(1+\cos\sigma)^2. \tag{2.27}$$

Thus

$$\frac{E\{I\}}{\sqrt{E\{I^2\}}} = \frac{4\sqrt{2}}{3\pi} \frac{(\pi - \sigma)\cos\sigma + \sin\sigma}{1 + \cos\sigma}. \tag{2.28}$$

Under the assumption that the orientations of the surfaces are uniformly distributed in 3-D space, they also proposed two methods to estimate the tilt of the light source. One of them was the local voting method which assumes that each surface point and its neighbors can be locally approximated by a spherical patch. Zheng and Chellappa considered that, for a small increment along the direction $\vec{d} = (\delta x, \delta y)$, the increment in intensity along that direction will be δI_d. Then, if we consider different directions for $\vec{d}$, $(\tilde{s}_x, \tilde{s}_y)$ can be solved using

$$\begin{pmatrix} \tilde{s}_x \\ \tilde{s}_y \end{pmatrix} = (B^T B)^{-1} B^T \begin{pmatrix} \delta I_1 \\ \delta I_2 \\ \vdots \\ \delta I_M \end{pmatrix}, \tag{2.29}$$

where

$$B = \begin{pmatrix} \delta x_1 & \delta y_1 \\ \delta x_2 & \delta y_2 \\ \vdots & \vdots \\ \delta x_M & \delta y_M \end{pmatrix} \tag{2.30}$$

and M is the number of measured directions of $\vec{d}$. The local estimate of τ is

$$\cos\tau_{(x,y)} = \frac{\tilde{s}_x}{\sqrt{\tilde{s}_x^2 + \tilde{s}_y^2}} \tag{2.31}$$

$$\sin\tau_{(x,y)} = \frac{\tilde{s}_y}{\sqrt{\tilde{s}_x^2 + \tilde{s}_y^2}} \tag{2.32}$$

and the global estimate of the tilt of the illuminant is

$$\cos\tau = E\{\cos\tau_{(x,y)}\} \tag{2.33}$$

$$\sin\tau = E\{\sin\tau_{(x,y)}\} \tag{2.34}$$

$$\tau = \arctan\left(\frac{\sin\tau}{\cos\tau}\right). \tag{2.35}$$

Then, tilt of the source is

$$\tau = \arctan\left(\frac{E\{\frac{\tilde{s}_x}{\sqrt{\tilde{s}_x^2+\tilde{s}_y^2}}\}}{E\{\frac{\tilde{s}_y}{\sqrt{\tilde{s}_x^2+\tilde{s}_y^2}}\}}\right). \tag{2.36}$$

The other was contour-based method which uses shading information along image contours. Under the assumption that the slant of the surface normals along the boundary are constant, the tilt angle α of a boundary pixel, is just the tilt angle of the boundary contour in the image plane, and the summations of $\cos\alpha$ and $\sin\alpha$ over the closed boundary are zero. Then, the tilt angle τ can be computed from

$$\tau = \arctan\left(\frac{x_2}{x_1}\right), \tag{2.37}$$

where

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (C^T C)^{-1} C^T \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_M \end{pmatrix} \tag{2.38}$$

and

$$C = \begin{pmatrix} \cos\alpha_1 & \sin\alpha_1 \\ \cos\alpha_2 & \sin\alpha_2 \\ \vdots & \vdots \\ \cos\alpha_M & \sin\alpha_M \end{pmatrix}. \tag{2.39}$$

For the computation of the slant σ and the albedo η, Zheng and Chellappa defined three functions, $f_1(\sigma)$, $f_2(\sigma)$ and $f_3(\sigma) = f_1(\sigma)/\sqrt{f_2(\sigma)}$, and used them as

$$E\{I\} = \eta f_1(\sigma) \tag{2.40}$$

$$E\{I^2\} = \eta^2 f_2(\sigma). \tag{2.41}$$

Here, $f_1(\sigma)$, $f_2(\sigma)$ and $f_3(\sigma)$ are computed by a numerical method and the results are approximated by seventh-order polynomials of $\cos\sigma$. Thus, σ can be uniquely solved from

$$\frac{E\{I\}}{\sqrt{E\{I^2\}}} = f_3(\sigma) \tag{2.42}$$

and η can be computed by

$$\eta = \frac{1}{f_1^2(\sigma) + f_2(\sigma)}\left[E\{I\} f_1(\sigma) + \sqrt{E\{I^2\} f_2(\sigma)}\right]. \tag{2.43}$$

All of the above three methods are similar. For example, for the estimation of the tilt, if we take the derivatives only in x and y directions in Pentland's regression model, it reduces to Lee and Rosenfeld's method.

In [5], Zheng and Chellappa tested the above methods on a set of three different images. The results showed that for the estimating of the tilt angle, all algorithms work almost perfectly for a sphere without background. However, background and noise will degrade the performance of both Lee and Rosenfeld's and Pentland's methods. Consequently, Zheng and Chellappa's method is more robust to background and noise in most of the cases. For the estimation of the slant angle, Pentland's method is very sensitive to noise but Lee and Rosenfeld's and Zheng and Chellappa's methods are robust to Gaussian noise. If a uniform background is included, results for all three methods are degraded. In terms of the computation time, since M is usually greater than two, Lee and Rosenfeld's method is the most efficient, Pentland's is the second and Zheng and Chellappa's is a bit slower than Pentland's as it needs to solve seventh-order polynomials.

The summary of formulas for estimating illuminant direction and albedo are presented in Table 2.1.

Table 2.1. Summary of Formulas for Illuminant Direction and Albedo.

Pentland's Method	τ	$\tau = \arctan(\frac{\tilde{s}_x}{\tilde{s}_y})$; $\begin{pmatrix}\tilde{s}_x\\ \tilde{s}_y\end{pmatrix} = (B^TB)^{-1}B^T\begin{pmatrix}d\bar{I}_1\\ d\bar{I}_2\\ \vdots\\ d\bar{I}_M\end{pmatrix}$
	σ	$\sigma = \arccos(1 - \frac{\tilde{s}_x^2+\tilde{s}_y^2}{k^2})^{\frac{1}{2}}$; $k = (E\{dI^2\} - E\{dI\}^2)^{\frac{1}{2}}$
Lee and Rosenfeld's Method	τ	$\tau = \arctan(\frac{E\{I_y\}}{E\{I_x\}})$
	σ	Original form: $\frac{E\{I\}}{\sqrt{E\{I^2\}}} = \frac{8}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{(1+\cos\sigma)^{\frac{3}{2}}}$ Modified form: $\frac{E\{I\}}{\sqrt{E\{I^2\}}} = \frac{4\sqrt{2}}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{1+\cos\sigma}$
	η	$\eta = \frac{1}{a^2+b^2}(aE\{I\} + \sqrt{bE\{I^2\}})$ Original form: $a = \frac{4}{3\pi}\frac{(\pi-\sigma)\cos\sigma+\sin\sigma}{1+\cos\sigma}$; $b = \frac{1+\cos\sigma}{4}$ Modified form: $a = \frac{2}{3\pi}[(\pi-\sigma)\cos\sigma + \sin\sigma]$; $b = \frac{(1+\cos\sigma)^2}{8}$
Zheng and Chellappa's Method	τ	Local: $\tau = \arctan(\frac{E\{\tilde{s}_x/\sqrt{\tilde{s}_x^2+\tilde{s}_y^2}\}}{E\{\tilde{s}_y/\sqrt{\tilde{s}_x^2+\tilde{s}_y^2}\}})$; $\begin{pmatrix}\tilde{s}_x\\ \tilde{s}_y\end{pmatrix} = (B^TB)^{-1}B^T\begin{pmatrix}\delta I_1\\ \delta I_2\\ \vdots\\ \delta I_M\end{pmatrix}$ Contour: $\tau = \arctan(\frac{x_2}{x_1})$; $\vec{X} = (C^TC)^{-1}C^T\begin{pmatrix}I_1\\ I_2\\ \vdots\\ I_M\end{pmatrix}$; $C = \begin{pmatrix}\cos\alpha_1 & \sin\alpha_1\\ \cos\alpha_2 & \sin\alpha_2\\ \vdots & \vdots\\ \cos\alpha_M & \sin\alpha_M\end{pmatrix}$
	σ	$\frac{E\{I\}}{\sqrt{E\{I^2\}}} = 0.5577 + 0.6240\cos\sigma + 0.1882\cos^2\sigma - 0.6514\cos^3\sigma$ $-0.5350\cos^4\sigma + 0.9282\cos^4\sigma\cos^5 + 0.3476\cos^6\sigma - 0.4984\cos^7\sigma$
	η	$\eta = [E\{I\}\, f_1(\sigma) + (E\{I^2\}\, f_2(\sigma))^{\frac{1}{2}}] \,/\, (f_1^2(\sigma) + f_2(\sigma))$

CHAPTER 3.

SHAPE FROM SHADING PROBLEM

3.1 Introduction

Extracting surface orientation and surface depth from a shaded image, known as the shape-from-shading problem, is one of the fundamental problems in computer vision. Different parts of the surface are oriented differently and thus will appear with different brightness. This spatial variation of brightness, or *shading*, is used to estimate the orientation of surface patches. Measurement of brightness at a single point in the image, however, only provides one constraint, while surface orientation has two degree of freedom. Without additional information, we cannot recover the orientation of a surface patch from the image irradiance equation

$$I(x,y) = R(z_x(x,y), z_y(x,y)). \tag{3.1}$$

Here, $I(x,y)$ is an image formed by projection of the light onto a plane parallel to the xy-plane, $R(z_x, z_y)$ is the reflectance map relating image brightness to surface orientation, z_x and z_y denote the first partial derivatives of z with respect to x and y respectively. The solution of (3.1) may be expressed mathematically in terms of a first-order partial differential equation. Specifically, a function $z(x,y)$, representing surface depth in the direction of the z-axis, satisfies the image irradiance equation over the image plane. In this formulation, it is implicitly assumed that the light sources are infinitely far away and internal surface reflections are disallowed. On the other hand, the assumption of smoothness of a surface provides a strong constraint. Neighboring patches of

the surface cannot assume arbitrary orientations. They have to fit together to make a continuous smooth surface.

In shape-from-shading algorithms, it is assumed that the reflectance map is given or its form is known. However, it is very important to know whether a solution to the SFS problems exists and whether there is more than one solution. Unfortunately, the existence and uniqueness questions are difficult to decide without additional assumptions. More can be said when specific reflectance maps are chosen such as ones that are linear or those are rotationally symmetric [46]. Since images of most surfaces in the real world can be approximated by Lambertian reflectance, the majority of SFS methods use the Lambertian reflectance model. The important parameters in Lambertian reflectance are albedo and illuminant direction. Commonly, the albedo is assumed to be constant and there are several methods for computing the light source direction.

It has been recently shown that there exist *impossible shaded images*, that is, images that do not correspond to any surface illuminated in the specific way. It may turn out that almost all images with multiple singular points are impossible in this sense [47]. This is an important issue, because it may help explain how our visual system sometimes determines that the surface being viewed cannot possibly be uniform in its reflecting properties. One can easily come up with smoothly shaded images, for example, that do not yield an impression of shape, instead appearing as flat surfaces with spatially varying reflectance or surface *albedo*.

3.2 Ill-Posed Problems and Regularization

3.2.1 Introduction

The well-known rendering problem in the field of computer graphics involves the synthesis of images from explicit representations of 3-D forms. Conversely, early computational vision aims at understanding how explicit geometric representations of the 3-D world may be reconstructed from 2-D images. During the imaging process that maps the 3-D scene to a 2-D image, most

of the information about scene is lost. Then, different scenes may result in the same image. Thus only the pixel gray values in the image do not provide enough information to extract 3-D properties of surfaces due to this ambiguity. Addition to this, the unavoidable noise which is added during the imaging and the sampling process is another factor that makes the inverse problems more difficult.

A theoretical approach which unifies these problems is to consider early vision problems as ill-posed problems [48, 49, 50]. Hadamard [48, 51] defined the conditions which well-posed problems have to satisfy. A problem is said to be well-posed if the following conditions are satisfied:

1. For each data d in a given class of functions Y, there exists a solution u in a prescribed class X (*existence*);
2. The solution u is unique in X (*uniqueness*);
3. The solution u continuously depends on data d; when the error on the data d tends to zero, the included error on the solution u tend also to zero (*continuity*).

Thus a problem is ill-posed in the sense of Hadamard when it fails to satisfy any of the three conditions given above.

3.2.2 Regularization

The general theory and methods for solving ill-posed problems have been developed by Tikhonov [52], Arsenin [52], Ivanov [53] and Morozov [54, 55]. The main idea for solving ill-posed problems is to restrict the class of admissible solutions by introducing a priory knowledge. Thus one of the important issue is to find valid and general constraints which allow us to solve ill-posed problems. In this regard, constraints such as smoothness have been useful expressions of generic, a priori information about possible solutions. The physical justification of the smoothness is that the coherence of the matter tends to give rise to smoothly varying intrinsic scene characteristics relative to the viewing distance.

A priori knowledge can be formulated either in terms of the variational principles that impose constraints on the possible solutions or as statistical properties of the solution space. The general term "regularization" is used for any method which transform an ill-posed problem to a well-posed problem. Variational regularization indicates the regularization methods that reformulate an ill-posed problem in terms of variational principle. The standard regularization methods developed mostly by Tikhonov fall into this class.

Through regularization, a wide range of ill-posed visual reconstruction problems may be reformulated as variational principles. Tikhonov regularization employs a specific class of so-called stabilizing functionals to restrict admissible solutions to spaces of smooth surfaces. Under nonrestrictive conditions, the resulting variational principles can be made well-posed within these spaces; hence, their solutions are effectively computable. Regularization therefore appears to offer a theoretical basis for the smoothness constraints that have been applied to the reconstruction problems of early vision.

Let us consider the regularization of the ill-posed problem of finding z from the data y such that

$$A\,z = y. \tag{3.2}$$

The regularization process requires the choice of suitable norms $\|.\|$ and of a stabilizing functional $\|Pz\|$. In this choice, the analysis of the physical constraints as well as the mathematical considerations play a very important role. In the standard regularization theory, A is a linear operator, the norms are quadratic and P is linear.

There are three methods of standard regularization [50]:

1. Among z that satisfy $\|Pz\| \leq C_1$, where C_1 is a constant, find z that minimizes $\|Az - y\|$,
2. Among z that satisfy $\|Az - y\| \leq C_2$, find z that minimizes $\|Pz\|$,
3. Find z that minimizes $\|Az - y\|^2 + \lambda \|Pz\|^2$, where λ is the regularization parameter.

The first method searches for a function z that satisfies the constraint $\|Pz\| \leq C_1$ and also best approximates the data. The second one consists of finding the function z which is sufficiently close to the data and is most regular. In the third method, the regularization parameter λ controls the compromise between the degree of regularization of the solution and its closeness to the data.

Thus in the standard regularization method, the constraints on the solution are imposed as a variational principle such as a cost functional. The solution minimizing this functional is a good solution since it has to be close to the data and be regular by making the quantity $\|Pz\|$ small. In this formulation, the physical constraints on the solution are formulated as stabilizer P. As mentioned earlier, the smoothness constraint is used widely as an physical constraint on the solution. It can be also justified as the following; the noiseless image has to be smooth in the sense that all its derivatives must exist and be bounded since the image is band limited by the optics. This constraint allows us to eliminate the noise and recover the continuity of the solution on data making the problem well-posed. The smoothness constraint can be imposed in the simplest form by selecting the stabilizer as $P = \frac{\partial^2}{\partial x^2}$ and the norm as $L_2[a, b]$ norm:

$$\|f\| = \int_a^b f(x)^2 dx. \tag{3.3}$$

This stabilizer can be used to compute the derivative of noisy data, which is unstable without regularization. With this choice of P and norm, the regularization for the numerical differentiation is to find the approximating function z which minimizes the functional

$$E(f) = \sum_{i=1}^{N} (f(x_i) - y_i)^2 + \lambda \int_0^1 \left(\frac{\partial^2 f}{\partial x^2}\right)^2 dx, \tag{3.4}$$

where the first term is the measure of the closeness of the solution to data and the second term, stabilizer, is the measure of the smoothness. The compromise between two terms is controlled by the regularization parameter λ.

A more general univariate stabilizer contains a large range of the derivatives as proposed by Tikhonov. In this case, the stabilizer becomes the p-th order weighted Sobolev norm

$$\|\nu\|_p^2 = \sum_{m=o}^{p} \int_{\Re} w_m(x) \left(\frac{d^m \nu(x)}{dx^m} \right)^2 dx, \tag{3.5}$$

where w_m is a nonnegative, prespecified and continuous weighting function. Thus the previous stabilizer is a special case of this stabilizer with $p = 2$, $w_0 = w_1 = 0$ and $w_2 = 1$. In follows from the fact that the Tikhonov regularization imposes the smoothness constraints on the solution by restricting the solution space to the Sobolev space of the smooth functions. There exists also a close relation between the Tikhonov regularization and the approximating by splines. Minimizing the functional $E(\nu)$ is the variation formulation of the smoothing cubic spline problem and Tikhonov regularization in (3.4) is therefore equivalent to fitting a cubic spline to the data [49]. This close relation between splines and the regularization helps to obtain a stabilizer for the multivariate regularization. The generalized splines for two-dimensional case are given by

$$\|\nu\|_m^2 = \int\int_{\Re^2} \sum_{i=0}^{m} \binom{m}{i} \left(\frac{d^m \nu}{dx^i dy^{m-i}} \right)^2 dx\, dy, \tag{3.6}$$

which can be used for surface fitting to scattered data and for visible surface reconstruction in computer vision.

The spline functionals given above have interesting physical interpretations involving equilibria of the elastic bodies with C^{m-1} intrinsic continuity. For $m = 1$, the spline functional reduces to

$$\|\nu\|_1^2 = \int\int_{\Re^2} (\nu_x^2 + \nu_y^2)\, dx\, dy, \tag{3.7}$$

which is proportional to the small deflection energy of a membrane. The associated Euler-Lagrange equation of this functional is the Laplacian equation as

$$-\nabla_u = -(u_{xx} + u_{yy}) = 0. \tag{3.8}$$

For $m = 2$, the functional becomes

$$\|\nu\|_2^2 = \int\int_{\Re^2} (\nu_{xx}^2 + 2\nu_{xy}^2 + \nu_{yy}^2)\, dx\, dy \tag{3.9}$$

and it is proportional to the deflection bending energy of a thin plate. The Euler-Lagrange equation of this functional is the biharmonic equation

$$\nabla_u^2 = (u_{xxxx} + 2u_{xxyy} + u_{yyyy}) = 0. \tag{3.10}$$

Physically the membrane spline represents a surface of C^0 continuity, a continuous surface which need not have continuous first and higher order partial derivatives. The thin plate spline characterizes a C^1 surface, a continuous surface with first partial derivatives which need not have continuous derivatives of degree greater than one.

The regularization parameter λ in the functional

$$\|Az - y\|^2 + \lambda \|Pz\|^2 \tag{3.11}$$

controls the compromise between closeness to the data and the degree of the regularization. One of the important problems in this formulation is the optimal choice of this parameter. Several techniques have been proposed to determine the optimum regularization parameter λ [56, 57, 58, 59, 60]. The so-called min-max principle [60] may be applied to avoid the problems of extremely low and high parameter values. The basic idea is that if lambda is either too high or too low, one of the terms in (3.11) will be inadequately presented in the total cost, and the total cost will be too low. This suggests that we might select a λ which maximizes the minimum energy value. Since each minimum energy level has been already minimized over z, this will avoid problems of λ being excessively low or high.

Another important method proposed for the optimal choice of λ is the generalized cross-validation method [56]. The main idea behind this method is to allow data points themselves to choose the value of the regularization parameter by requiring that a good value of the parameter should predict the missing data points. This method does not require a priori knowledge about the solution or noise. Different approaches on this subject can also be found in references [50, 57, 58, 59].

Visual reconstruction problems tend to be ill-posed in that existence, uniqueness and stability of solutions cannot be guaranteed in the absence of additional constraints.

Regularization techniques are widely used for inverse problems solving in

computer vision such as shape from shading, surface reconstruction, edge detection or optical flow estimation. Formally, regularization transforms an ill-posed problem into a well-posed minimization problem by constraining the solution to belong to a set of smooth functions.

3.3 Review of Problem Formulation

3.3.1 Specifying Surface Orientation

There are several commonly used ways of specifying the orientation of a planar surface patch, including:

- Unit surface normal $\vec{N}$ [10];
- Surface gradient (p, q) [61];
- Slant and tilt angles (σ and τ);
- Stereographic coordinates (f, g) [11].

One way to specify surface orientation is to use a unit vector, $\vec{N} = (n_x, n_y, n_z)$, perpendicular to the local tangent plane. Another way is to specify the components p and q of the surface gradient. These are the partial derivatives of the surface height z above some reference plane perpendicular to the optical axis, that is

$$p = \frac{\partial z}{\partial x} \qquad \text{and} \qquad q = \frac{\partial z}{\partial y}. \tag{3.12}$$

The two notations are connected by the equality

$$\vec{N} = \frac{1}{\sqrt{1 + p^2 + q^2}}(-p, -q, 1)^T. \tag{3.13}$$

A third way of specifying surface orientation is to use slant (σ) and tilt(τ) angles. Slant is defined as the angle between the surface normal and the direction toward the viewer and tilt is defined as the angle between the projection of the surface normal on the image plane and the x-axis. The relationship of slant and tilt to the unit vector is given by

$$\vec{N} = (\sin\sigma\cos\tau, \sin\sigma\sin\tau, \cos\sigma)^T. \tag{3.14}$$

We can also identify surface orientation with points on a unit sphere, named *Gaussian sphere* [62]. Using the stereographic projection, the whole sphere is mapped into a plane called the stereographic plane. Only the south pole ends up at infinity, while the equator is mapped onto a circle of radius two. The coordinates in stereographic space are f and g, and the mapping pq-space to fg-space is specified by the equations

$$f = \frac{2p}{1+\sqrt{1+p^2+q^2}} \quad \text{and} \quad g = \frac{2q}{1+\sqrt{1+p^2+q^2}}. \tag{3.15}$$

Conversely,

$$p = \frac{4f}{4-f^2-g^2} \quad \text{and} \quad q = \frac{4g}{4-f^2-g^2}. \tag{3.16}$$

It is known that the occluding boundary provides important constraints on the solution of the SFS problem [63]. The difficulty with using gradient to specify surface orientation is that at least one of p and q is unbounded on the occluding boundary. This problem can be overcome by specifying surface orientation in stereographic space. Then, the orientation of a point on the occluding boundary corresponds to a point on the circle of radius two in fg-space. Thus the use of the stereographic plane makes it possible to incorporate occluding boundary information. However, one disadvantage of the stereographic parameterization is that it is hard to express the condition of integrability compare with the gradient space parameterization.

The (p, q) components of the surface gradient is the most directly useful for specifying surface orientation. However we can convert between different representations easily. We can use the (p, q) notation to specify the direction to a light source or a small portion of an extended source. We simply give the orientation of a surface element that lies perpendicular to the incident rays. So we can write

$$\vec{S} = \frac{1}{\sqrt{1+p_s^2+q_s^2}}(-p_s, -q_s, 1)^T \tag{3.17}$$

for some p_s and q_s.

3.3.2 Reflectance Map

With the development of computer vision, the reflectance map is introduced to describe the scene radiance as a function of surface orientation. It is a representational tool used in developing methods for recovering surface shape from images. When the light sources and the viewer are far away from the scene being viewed, the use of reflectance map makes the analysis of shape-from-shading algorithms much easier [61, 64]. Several iterative techniques, mostly based on minimization of some functional containing integrals of some error terms, arose later [5, 9, 10, 11, 14, 65, 66]. If we use the unit surface normal $\vec{N}$ as a way of specifying surface orientation, then we can write the brightness as a function of orientation in the form $R(\vec{N})$. If we use p and q instead, we can use the form $R(p, q)$.

The reflectance map may be determined experimentally by mounting a sample of the surface and measuring its brightness under the given illuminating conditions for various orientations. Another way of finding a reflectance map is to use the image of a calibration object such as a sphere for which surface orientation is easily calculated at every point. On the other hand, a reflectance map may be derived in term of the *bidirectional reflectance distribution function* (BRDF) and the given distribution of source radiance [64]. The BRDF was introduced by Nicodemus *et al.* [67] as a unified notation for the specification of reflectance in terms of both incident- and reflected-beam geometry.

An ideal Lambertian surface which reflects all incident light and appears equally bright from all directions, illuminated by a single point source provides a convenient example of a reflectance map. If we let $\vec{N}$ be the unit surface normal and $\vec{S}$ be the unit vector pointing in the direction of the light source, then the cosine of the incident angle is just the dot product of these two vectors,

$$\cos\theta_i = \vec{N} \cdot \vec{S} \tag{3.18}$$

and the scene radiance is given by

$$R(p, q) = \frac{I_0}{\pi} \cos\theta_i, \tag{3.19}$$

where I_0 is the irradiance from the source on a surface oriented perpendicular

to the incident rays. Note that the radiance cannot be negative, so the above formula only applies when $0 \leq \theta_i \leq \pi/2$. The scene radiance is zero for values outside this range. Taking dot products of the corresponding unit vectors, $\cos\theta_i = \vec{N} \cdot \vec{S}$, we obtain

$$R(p,q) = \frac{I_0}{\pi} \frac{1 + pp_s + qq_s}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}} \tag{3.20}$$

as long as the numerator is positive, otherwise $R(p,q) = 0$.

The reflectance map is usually normalized so that its maximum is one for the convenience of computational aspects. For the Lambertian surface which is illuminated by a single distant light source, we can use

$$R(p,q) = \frac{1 + pp_s + qq_s}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}}. \tag{3.21}$$

Thus, aside from a fixed scale factor, the reflectance map gives the dependence of scene radiance on a surface orientation. Using the coordinates of the stereographic plane instead, we can write the reflectance map as

$$R(f,g) = \frac{16(ff_s + gg_s) + (4 - f^2 - g^2)(4 - f_s^2 - g_s^2)}{(4 + f^2 + g^2)(4 + f_s^2 + g_s^2)}. \tag{3.22}$$

The reflectance map can also be expressed in terms of slant and tilt angles. Using the relations between (p_s, q_s) and (σ, τ) given below

$$p_s = -\frac{\cos\tau \sin\sigma}{\cos\sigma} \qquad \text{and} \qquad q_s = -\frac{\sin\tau \sin\sigma}{\cos\sigma}, \tag{3.23}$$

we have

$$R(p,q) = \frac{\cos\sigma - p\sin\sigma\cos\tau - q\sin\sigma\sin\tau}{\sqrt{1+p^2+q^2}}. \tag{3.24}$$

3.3.3 Image Irradiance Equation

At a particular point (x, y) in the image, we can measure the image irradiance $I(x, y)$. It is proportional to the radiance at the corresponding point on the surface imaged, as determined by the projection equation. If the surface

gradient at that point is $(p(x,y), q(x,y))$, then the radiance there is $R(p,q)$. Hence, the basic equation describing the image forming process can be written as

$$I(x,y) = \mathcal{B}\, R(p,q). \tag{3.25}$$

This equation is referred to as the *image irradiance equation* and it is fundamental for recovering surface shape. It is customary to rescale image irradiance so that the proportionality factor, $\mathcal{B}$, may be dropped. Then we can write the image irradiance equation in the simple form

$$I(x,y) = R(p,q). \tag{3.26}$$

This is a first-order partial differential equation that is typically nonlinear, because of the nonlinear dependence of the reflectance map on the gradient.

As an illustration of the shading effect, consider a sphere with a Lambertian surface illuminated by a point source at essentially the same plane as the viewer. In this case $(p_s, q_s) = (0,0)$ and the reflectance map can be written as

$$R(p,q) = \frac{1}{\sqrt{1+p^2+q^2}}.$$

If the sphere is on the optical axis, then for its surface we can write

$$z = z_0 + \sqrt{r^2 - (x^2+y^2)} \qquad \text{for} \qquad x^2 + y^2 \leq r^2,$$

where r is the radius and $-z_0$ is the distance of its center from the lens. For the partial derivatives of the surface height $z(x,y)$

$$p = \frac{-x}{z - z_0} \qquad \text{and} \qquad q = \frac{-y}{z - z_0},$$

so that

$$\frac{1}{\sqrt{1+p^2+q^2}} = \frac{z - z_0}{r}.$$

As a result, we have

$$I(x,y) = R(p,q) = \sqrt{1 - \frac{x^2+y^2}{r^2}}$$

as the image irradiance equation. The brightness falls off smoothly from its maximum at the center of the image of the sphere to zero at the edges.

3.3.4 Variational Formulation

The image irradiance equation (3.26), which is used in image analysis to determine object shape from image irradiance, is a nonlinear first-order partial differential equation. Early solutions to the shape-from-shading problem were based on direct inversion of the image irradiance equation. Unfortunately, an exact solution to the imaging equation (3.26) does not always exist or there may be an infinite number of solutions. In practice, modeling errors such as reflectance map mismatch, imperfect knowledge of the light source, spatial and quantization error, observation noise and albedo variations are inevitable. Further, boundary conditions are generally not completely known and sometimes may not be available at all. These factors all influence the existence and uniqueness of a solution to equation (3.26) and the estimation of a good solution in the case that a unique one does not exist. For these reasons shape-from-shading is a very difficult problem in practice.

In any case, it is desirable to have a shape-from-shading method that neither moves away from correct solutions, nor converges to surfaces that are not solutions. It is more practical to pose shape-from-shading as a constrained minimization problem rather than purely an inversion problem. To derive such a method, we need to impose some constraints which are satisfied by the solution. In order to calculate object shape or equivalently surface orientation, additional assumptions and constraints must be imposed to the iterative scheme used. Suppose, we seek a smooth surface satisfying various constraints over some image domain. It is useful to obtain from the given constraints a non-negative expression that measures the departure of a particular surface from a satisfactory solution. We may then search for a surface that minimizes the expression.

The search for a function that minimizes an integral expression is the major concern of the calculus of variations [68]. In the calculus of variations, we look for extrema of expressions that depend on functions rather than parameters. Such expressions are called *functionals*. Using calculus of variations gives us the valuable result that the extrema of functionals must satisfy an associated *Euler equation*. Then, we can transform the source-recovery problem from one

of minimizing a functional, to one of solving one or more partial differential equations.

The variational approach, first introduced by Ikeuchi and Horn [11], was later used to developed SFS schemes as well as analyze existing ones [10]. This approach to early vision problems soon spread to other areas, such as motion vision and surface interpolation from sparse depth data obtained using binocular stereo [49, 69, 70, 71, 72, 73, 74]. A least-squares variational approach can be used to recover the best-fit surface [2, 10]. Methods based on expansion in terms of basis functions also implicitly provide a best-fit surface [9, 75].

Let us suppose that we obtain from a Euler equation a surface that generates a global minimum of the appropriate functional. The surface that best matches the constraints will generate a global minimum of the functional. This is a very important property of the variational approach. Since vision problems typically involve images that are noisy, this property is very important in computer vision. Exact solutions may not exist in this situation. For example, in the presence of noise, there may not be a smooth surface that satisfies the image irradiance equation $I(x, y) = R(p, q)$ exactly. However, there will be a surface that minimizes the integral of the square of the difference between $I(x, y)$ and $R(p, q)$.

3.3.5 A Procedure for Deriving Iterative Scheme

A procedure for deriving iterative shape-from-shading schemes for recovering surface shape is given in Table 3.1. Here, we seek a surface, $z(x, y)$, that best satisfying various requirements over some image domain Ω.

The approach follows the same pattern if the surface is parameterized in a different way. Also, similar results can be obtained by applying the finite-element method directly to the functional $\mathcal{I}$. Note that it is better to work with a functional that evaluates to zero for perfect solutions. In this way, we may use them to check how close an iterative scheme is to a solution. This is difficult with other functionals, as the minimum value is usually unknown. An additional advantage of functionals evaluating to zero is that there are no

Table 3.1. A procedure for deriving iterative scheme.

i. Select a functional, F, non-negative over Ω, such that

$$\mathcal{I} = \iint_{\Omega} F(x, y, z, ...)\, dx\, dy$$

constitutes a measure of the departure of z from an ideal solution.

ii. Absorb into F any constraint that z should satisfy over Ω, using Lagrangian multipliers if appropriate.

iii. If the problem is not well-posed as it stands, add a suitable regularization term.

iv. Find the Euler equation that must be satisfied by the surface z minimizing the functional $\mathcal{I}$.

v. Determine what boundary conditions are needed to ensure a unique solution. If there are no constraints on the function around the boundary $d\Omega$, determine the appropriate natural boundary conditions.

vi. Develop a discrete approximation of the associated Euler equation, using finite-difference methods.

vii. Design an iterative scheme that converges to the solution of the discrete approximation of the Euler equation.

unwanted surfaces that cause the functional to have a smaller value than that generated by a satisfactory solution.

3.4 Constraints for SFS Problem

The problem of minimizing the integral of the brightness error term is ill-posed, since it has an infinite number of solutions in terms of surface orientation. Then, regularization theory used to overcome this difficulty. Some extra terms, constraints, can be added to the integral to obtain an approximation to a solution. The constraints used in SFS problem are summarized as follows:

- **The brightness constraint** indicates the total brightness error of the reconstructed image compared with the input image and is given by

$$\int\int (I(x,y) - R(p,q))^2 \, dx \, dy, \tag{3.27}$$

where $I(x,y)$ is the measured intensity and $R(p,q)$ is the reflectance map. In any case, it is desirable to have a SFS method that converges to a unique solution. To derive such a method, we need to impose the smoothness constraint.

- **The smoothness constraint** ensures a smooth surface in order to stabilize the convergence to a unique solution and is given by

$$\int\int (p_x^2 + p_y^2 + q_x^2 + q_y^2) \, dx \, dy, \tag{3.28}$$

where p and q are the surface gradients along the x and y directions, respectively. Another version of the smoothness term is less restrictive by requiring constant change of depth only in x and y directions as

$$\int\int (p_x^2 + q_y^2) \, dx \, dy. \tag{3.29}$$

The smoothness constraint can also be described in terms of the surface normal

$$\int\int (\|\vec{N}_x^2 + \vec{N}_y^2\|) \, dx \, dy. \tag{3.30}$$

Although constraints (3.29) and (3.30) look alike, in reality, constraints (3.28) and (3.30) are similar if we consider the relationship between surface normal, $\vec{N}$, and the surface gradient, (p, q). Both (3.28) and (3.30) are most restrictive than (3.29) in terms of the smoothing directions. They are more frequently used, since (3.29) tends to lead to excessively flattened surfaces.

Let us suppose that shape-from-shading method recovers smooth functions $p(x, y)$ and $q(x, y)$ defined over the image. In general, there will be no smooth surface that corresponds to this gradient. This is because the functions p and q must be related in a special way if they are to correspond to a smooth surface [76]. This relation is specified by the constraint of integrability.

- **The integrability constraint** ensures a valid surface, that is, $p_y = q_x$ or $z_{xy} = z_{yx}$. It can be described either

$$\int\int (p_y - q_x)^2 \, dx \, dy \tag{3.31}$$

or

$$\int\int ((z_x - p)^2 + (z_y - q)^2) \, dx \, dy. \tag{3.32}$$

- **The intensity gradient constraint**, finally, requires that the intensity gradient of the reconstructed image be close to the intensity gradient of the input image in both x and y directions and is described as

$$\int\int ((R_x - I_x)^2 + (R_y - I_y)^2) \, dx \, dy. \tag{3.33}$$

3.5 Review of SFS Schemes

In this section, we discuss global minimization, global propagation and local approaches to the shape-from-shading problem.

3.5.1 Global Minimization Approaches

Global minimization approaches compute the solution which minimizes an energy function over the entire image. The function can involve some constraints, such as the brightness constraint, the smoothness constraint, the integrability constraint and the intensity gradient constraint, in order to obtain a unique solution. In this subsection, we discuss three global SFS methods given by Ikeuchi and Horn, Horn, and Zheng and Chellappa.

3.5.1.1 Ikeuchi and Horn

Ikeuchi and Horn [11] were the first to apply the calculus of variations to the shape-from-shading problem and also specified the surface orientation in stereographic space to incorporate occluding boundary information. Ikeuchi and Horn's objective was to find two functions, $f(x,y)$ and $g(x,y)$ with ensuring that the image irradiance equation $I(x,y) = R(f,g)$ was satisfied. Here $R(f,g)$, is the reflectance map expressed in stereographic coordinates. Without additional constraint the problem is ill-posed. In another words, the minimization of the total brightness error alone does not constitute a well-posed problem. Then Ikeuchi and Horn decided to add the measure of *lack of smoothness* given by

$$\int\int (f_x^2 + f_y^2 + g_x^2 + g_y^2)\, dx\, dy. \tag{3.34}$$

Adding this term to the brightness error, Ikeuchi and Horn obtained the functional

$$\int\int \{I(x,y) - R(f,g))^2 + \lambda\,(f_x^2 + f_y^2 + g_x^2 + g_y^2)\}\, dx\, dy, \tag{3.35}$$

that is to be minimized by choosing f and g. In the above formula, f_x, f_y, g_x and g_y are the first partial derivatives of f and g with respect to x and y, and λ is a scalar that assigns a relative weighting to the terms. The corresponding Euler equations of this minimization problem

$$F_f - \frac{\partial}{\partial x}F_{f_x} - \frac{\partial}{\partial y}F_{f_y} = 0 \tag{3.36}$$

$$F_g - \frac{\partial}{\partial x}F_{g_x} - \frac{\partial}{\partial y}F_{g_y} = 0, \tag{3.37}$$

where F_f is the partial derivative of F with respect to f. Applying these formulas we obtain

$$(I - R)R_f + \lambda \nabla^2 f = 0 \tag{3.38}$$

$$(I - R)R_g + \lambda \nabla^2 g = 0, \tag{3.39}$$

where R_f and R_g are the partial derivatives of $R(f, g)$ with respect to f and g, and

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{3.40}$$

is the Laplacian operator. These Euler equations do not have a unique solution without additional constraint. Thus, Ikeuchi and Horn used f and g values on the occluding boundary. Finding the solution of the Euler equations with this particular set of boundary conditions usually constitutes a well-posed problem.

Using the five-point approximation of the Laplacian, we get

$$\{\nabla^2 f\}_{ij} \approx \frac{4}{\epsilon^2}(\bar{f}_{ij} - f_{ij}) \tag{3.41}$$

$$\bar{f}_{ij} = \frac{1}{4}(f_{i,j+1} + f_{i+1,j} + f_{i,j-1} + f_{i-1,j}),$$

where ϵ denotes the spacing between picture cells and $\bar{f}_{ij}$ is the local average of f. The same computation can be done for g. Ikeuchi and Horn used the finite difference approximations in the Euler equations derived above, they obtained the iterative scheme

$$f_{ij}^{k+1} = \bar{f}_{ij}^{\,k} + \frac{\epsilon^2}{4\lambda}\left[I_{ij} - R(f_{ij}^k, g_{ij}^k)\right] R_f(f_{ij}^k, g_{ij}^k) \tag{3.42}$$

$$g_{ij}^{k+1} = \bar{g}_{ij}^{\,k} + \frac{\epsilon^2}{4\lambda}\left[I_{ij} - R(f_{ij}^k, g_{ij}^k)\right] R_g(f_{ij}^k, g_{ij}^k), \tag{3.43}$$

where the superscript denotes the iteration number.

After determining surface orientation in stereographic plane, the next task is to map the solution from fg-space to pq-space and find the surface height. For this, we can find a best-fit surface z to the components of the gradient, p and q, by minimizing the functional

$$\iint ((z_x - p)^2 + (z_y - q)^2)\, dx\, dy \tag{3.44}$$

whose Euler equation is

$$\nabla^2 z = p_x + q_y. \tag{3.45}$$

Using the discrete approximation to the Laplacian employed earlier, we obtain the iterative scheme

$$\begin{aligned} z_{ij}^{k+1} &= \bar{z}_{ij}^{k} - \frac{\epsilon}{4}(h_{ij} + v_{ij}) \\ \bar{z}_{ij} &= \frac{1}{4}(z_{i+1,j} + z_{i-1,j} + z_{i,j+1} + z_{i,j-1}) \\ h_{ij} &= \frac{1}{2}(p_{i+1,j} - p_{ij}) \\ v_{ij} &= \frac{1}{2}(q_{i,j+1} - q_{ij}). \end{aligned} \tag{3.46}$$

Ikeuchi and Horn's method works reasonably well having good stability and convergence properties. However, the solution for surface orientation may not correspond to an underlying smooth surface and that solutions may be distorted by the presence of the regularization term. The degree of distortion depends on the parameter λ. On the other hand, Ikeuchi and Horn's approach requires occluding boundary information in terms of the stereographic coordinated f and g. In order to construct the surface of an object, a mapping step from fg-space to pq-space and then, an integration of the surface orientation is required. Moreover, a well-known problem of this method is the slow convergence rate.

3.5.1.2 Horn

Horn [12] combined the brightness constraint, the smoothness constraint and the integrability constraint, and minimized the following energy function

$$\int\int \{(I - R)^2 + \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2) + \mu((z_x - p)^2 + (z_y - q)^2)\}\, dx\, dy, \tag{3.47}$$

where the first term corresponds to the requirement that the image synthesized from the reconstructed shape be close to the input image. The second term is the regularization constraint used to ensure the convergence of the iterative scheme. The third term comes from the requirement of integrability [9, 10], i.e., the shape-from-shading result should correspond to a physical surface.

The Euler equations lead to the following coupled system of second-order partial differential equations

$$\begin{aligned} \lambda\Delta p &= -(I-R)R_p - \mu(z_x - p) && (3.48)\\ \lambda\Delta q &= -(I-R)R_q - \mu(z_y - q) && (3.49)\\ \Delta z &= p_x + q_y. && (3.50) \end{aligned}$$

A discrete approximation of these equations can be obtained by using the discrete approximation of the Laplacian operator introduced in equation (3.41) and by approximating the reflectance map $R(p,q)$ locally by a linear function of p and q as

$$R(p,q) \approx R(p_0,q_0) + (p-p_0)R_p(p_0,q_0) + (q-q_0)R_q(p_0,q_0), \qquad (3.51)$$

where (p_0, q_0) is the reference gradient, R_p and R_q are the partial derivatives of the reflectance map with respect to p and q. Then, the iterative scheme can be written as follows:

$$\begin{aligned} p^{k+1} &= p^k + \delta p\\ q^{k+1} &= q^k + \delta q\\ z^{k+1} &= \bar{z}^k - \frac{\epsilon}{\kappa}(p_x + q_y), \end{aligned} \qquad (3.52)$$

where $\bar{p}$, $\bar{q}$ and $\bar{z}$ are the local averages, and

$$\begin{aligned} \delta p &= \frac{4}{D}(\lambda'' + R_q^2)A - R_pR_qB\\ \delta q &= \frac{4}{D}(\lambda'' + R_p^2)B - R_qR_pA\\ A &= \kappa\lambda'\delta\bar{p} + \mu\delta z_x + (I-R)R_p\\ B &= \kappa\lambda'\delta\bar{q} + \mu\delta z_y + (I-R)R_q\\ D &= \lambda''(\lambda'' + R_p^2 + R_q^2)\\ \lambda'' &= \kappa\lambda' + \mu\\ \lambda' &= \lambda/\epsilon^2\\ \delta\bar{p} &= \bar{p}^{k+1} - p^k\\ \delta\bar{q} &= \bar{q}^{k+1} - q^k\\ \delta z_x &= {\delta z_x}^{k+1} - p^k\\ \delta z_y &= {\delta z_y}^{k+1} - q^k. \end{aligned}$$

Here, $\kappa = 4$ when the local averages are computed using four edge-adjacent neighbors.

In the previous method, Ikeuchi and Horn, first compute surface orientation and do not ensure that the resulting gradient field is integrable. As a second step, they find a surface whose gradient comes closest to the estimated gradient field in a least-squares sense. These two steps are combined in Horn's method that recovers gradient and height at the same time. Linearization of the reflectance map about the local average surface orientation greatly improves the performance of the method.

The functional (3.47) Horn used contains a penalty term for departure from smoothness, so it may appear that it cannot converge to the exact solution. However, Horn used the regularization term to stabilize the iterative scheme when it is far from the correct solution, and he turned off it as the solution is approached. Thus, he found an exact solutions of given shape-from-shading problem even though a regularization term was included.

3.5.1.3 Zheng and Chellappa

Zheng and Chellappa [5] applied the intensity gradient constraint, instead of the smoothness constraint, and minimized the energy function

$$\int\int \{(I-R)^2+(I_x-R_x)^2+(I_y-R_y)^2+\mu\left((z_x-p)^2+(z_y-q)^2\right)\}\,dx\,dy. \quad (3.53)$$

The Euler equations are simplified by taking the Taylor series of the reflectance map and representing the depth, gradient and their derivatives in discrete form. Then, the iterative scheme, which updates depth and gradients simultaneously, is derived as

$$\begin{aligned} p^{k+1} &= p^k + \delta p \\ q^{k+1} &= q^k + \delta q \\ z^{k+1} &= z^k + \delta z, \end{aligned} \quad (3.54)$$

where

$$
\begin{aligned}
\delta p &= \frac{4}{\Delta}[(C_1 - \frac{1}{4}\mu C_3)(5R_q^2 + \frac{5}{4}\mu) - (C_2 - \frac{1}{4}\mu C_3)(5R_pR_q + \frac{1}{4}\mu)] \\
\delta q &= \frac{4}{\Delta}[(C_2 - \frac{1}{4}\mu C_3)(5R_q^2 + \frac{5}{4}\mu) - (C_1 - \frac{1}{4}\mu C_3)(5R_pR_q + \frac{1}{4}\mu)] \\
\delta z &= \frac{1}{4}(C_3 + \delta p + \delta q)
\end{aligned}
$$

$$
\begin{aligned}
C_1 &= (-R + I + R_p p_{xx} + R_q q_{xx} - I_{xx} + R_p p_{yy} + R_q q_{yy} - I_{yy})R_p - \mu(p - z_x) \\
C_2 &= (-R + I + R_p p_{xx} + R_q q_{xx} - I_{xx} + R_p p_{yy} + R_q q_{yy} - I_{yy})R_q - \mu(q - z_y) \\
C_3 &= -p_x + z_{xx} - q_y + z_{yy} \\
\Delta &= 4\{5\mu[R_p^2 + R_q^2 + \frac{1}{4}(R_p - R_q)^2] + 1.5\mu^2\}.
\end{aligned}
$$

Zheng and Chellappa implemented their algorithm using the hierarchical structure in order to speed up the computation. Their scheme updates the surface slope and height maps simultaneously. There is no special requirement for the initialization of the boundary and the initial values for both depth and gradient can be zero.

A good comparison of Ikeuchi and Horn's method and Zheng and Chellappa's method can be found in our previous work [77].

3.5.2 Global Propagation Approaches

Global propagation approaches start from the surface points where the shape either is known or can be uniquely determined, such as singular points, and propagate the shape information across the whole image. We present three algorithms in this section.

3.5.2.1 Characteristic Strips

The original solution of the general shape from shading problem by Horn uses the method of characteristic strip expansion [7, 13]. The basic idea is quite easy to explain using reflectance map [2, 61]. Suppose that we are at a point on the surface and we wish to extend the solution a small distance in

some direction by taking a step δx in x and δy in y. We need to compute the change in height δz. This we can do if we know the components of the gradient, because

$$\delta z = p\,\delta x + q\,\delta y. \tag{3.55}$$

So, as we explore the surface, we need to keep track of p and q in addition to x, y and z. This means that we also need to be able to compute the changes in p and q when we take the step. This can be done using

$$\delta p = r\,\delta x + s\,\delta y \qquad \text{and} \qquad \delta q = s\,\delta x + t\,\delta y, \tag{3.56}$$

where $r = z_{xx}$, $s = z_{xy} = z_{yx}$ and $t = z_{yy}$ are the second partial derivatives of the height. It seems that we need to now keep track of the second derivatives, and so on.

To avoid this infinite recurrence, we take another tack. To find the brightness gradient we differentiate the image irradiance equation $I(x, y) = R(p, q)$ with respect to x and y, and so obtain

$$I_x = r\,R_p + s\,R_q \qquad \text{and} \qquad I_y = s\,R_p + t\,R_q. \tag{3.57}$$

At this point we exploit the fact that we are free to choose the direction of the step $(\delta x, \delta y)$. Suppose that we pick

$$\delta x = R_p\,\delta\xi \qquad \text{and} \qquad \delta y = R_q\,\delta\xi \tag{3.58}$$

then, from equations (3.56) and (3.57) we have

$$\delta p = I_x\,\delta\xi \qquad \text{and} \qquad \delta q = I_y\,\delta\xi. \tag{3.59}$$

We can summarize the above in the set of ordinary differential equations

$$\begin{aligned} &\dot{p} = I_x, \quad \dot{q} = I_y, \\ &\dot{x} = R_p, \quad \dot{y} = R_q, \quad \dot{z} = pR_p + qR_q, \end{aligned} \tag{3.60}$$

where the dot denotes differentiation with respect to ξ, a parameter that varies along a particular solution curve. Note that we actually have more than a

mere *characteristic curve*, since we also know the orientation of the surface at all points in this curve. This is why particular solution is called *characteristic strip*. The projection of a characteristic curve into the image plane is called a *base characteristic*.

The base characteristics are predetermined straight lines in the image only when the ratio $\dot{x} : \dot{y} = R_p : R_q$ is fixed, that is when the reflectance map is linear in p and q. In general, we cannot integrate along arbitrary curves in the image. Also, an initial curve is needed from which to sprout the characteristic strips.

In turns out that direct numerical implementation of the above equations does not yield particularly good results, since the paths of the characteristics are affected by noise in the image brightness measurements and errors tend to accumulate along their length. In particularly bad cases, the base characteristics may even cross, which does not make any sense in terms of surface shape. It is possible, however, to grow characteristic strips in parallel and use a so-called sharpening process to keep neighboring characteristics consistent [7, 13]. This greatly improves the accuracy of the solution, since the computation of surface orientation is tied more closely to image brightness itself rather than to the brightness gradient. This also makes it possible to interpolate new characteristic strips when existing ones spread too far apart, and to remove some when they approach each other too closely.

3.5.2.2 Dupuis and Oliensis

Most of the shape-from-shading approaches make use of the smoothness constraint. Although this term helps to stabilize the minimization process it pushes the reconstruction toward a smooth surface. Oliensis [20] discovered that the smoothness constraint is often unnecessary since shading determines shape with little ambiguity for a typical image. He also observed that while the occluding boundary does not strongly constrain the surface solution, a singular point does. Therefore the surface should be reconstructed from the interior of the image outward, instead of from the boundary inward. Since both singular points and characteristic strips are independent of the viewing direction and

characteristic strips correspond to steepest ascent curves on the object, the shape of a surface can be constructed strip by strip given the information at the singular points.

Considering the importance of singular points, Dupuis and Oliensis [21, 78] developed an iterative algorithm to recover depth using discretized optimal control and dynamic programming. The proof of equivalence between optimal control representation and shape-from-shading was illustrated. At first, they required a vertical light source and only one singular point, then they removed these restrictions and allow for a general light source and multiple singular points. However their initial algorithm [21] requires priori depth information for all the singular points. A later version [78] can determine this information automatically by assuming twice differentiable depth, isolated singular points and nonzero curvature at singular points.

3.5.2.3 Bichsel and Pentland

Following the main idea of Dupuis and Oliensis's, Bichsel and Pentland [23] developed an efficient minimum downhill approach which directly recovers depth and guarantees a continuous surface. Given initial values at the singular points, the algorithm looks in eight directions in the image and propagates the depth information away from the light source to ensure the proper termination of the process. Since slopes at the surface points in low brightness regions are close to zero for most directions (except the directions which form a very narrow angle with the illumination direction), the image was initially rotated to align the light source direction with one of the eight directions. The inverse rotation was performed on the resulting depth map in order to get the original orientation back.

Bichsel and Pentland used the following form as the reflectance map

$$R(p,q) = \frac{-s_x p - s_y q + s_z}{\sqrt{1+p^2+q^2}}, \tag{3.61}$$

where (s_x, s_y, s_z) is the unit light source vector. Assuming the constraint of parallel slope, the surface gradient , (p,q), was computed by taking the derivative of the equation (3.61) with respect to q in the rotated coordinate system,

setting it to zero and then solving p and q. The solution for p and q were given by

$$p = \frac{-s_x s_z \pm \sqrt{(1-R^2)(R^2-s_y^2)}}{R^2-s_x^2-s_y^2} \tag{3.62}$$

$$q = \frac{p s_y s_x - s_y s_z}{R^2-s_y^2}. \tag{3.63}$$

A drawback to this approach is the requirement for singular points.

3.5.3 Local Approaches

Local approaches use only local intensity information around the current pixel to derive shape. Here, we describe two approaches.

3.5.3.1 Pentland

Pentland's first approach [3] solved for the surface slant and tilt, the radius of curvature, and the light source direction through six equations obtained from the intensity, as well as the first and the second derivatives of the intensity. His approach can classify a surface type as planar, cylindrical, convex, concave or saddle surface. However, the use of the second derivatives makes the algorithm very sensitive to noise.

Pentland's second approach [24] used the linear approximation of the reflectance map in p and q. By taking the Taylor series expansion of the reflectance function (3.61) about $p = p_0$, $q = q_0$ and ignoring the high order terms, we have

$$I(x,y) \approx R(p_0,q_0) + (p-p_0)\frac{\partial R}{\partial p}(p_0,q_0) + (q-q_0)\frac{\partial R}{\partial q}(p_0,q_0). \tag{3.64}$$

For Lambertian reflectance, the above equation at $p_0 = q_0 = 0$, reduces to

$$I(x,y) = \cos\sigma + p\cos\tau\sin\sigma + q\sin\tau\sin\sigma. \tag{3.65}$$

Next, Pentland takes the Fourier transform of both sides of the equation (3.65). Since the first term on the right is a DC term, it can be dropped. Using the identities below

$$\frac{\partial}{\partial x}z(x,y) \longleftrightarrow F_z(w_1,w_2)(-iw_1) \tag{3.66}$$

$$\frac{\partial}{\partial y}z(x,y) \longleftrightarrow F_z(w_1,w_2)(-iw_2), \tag{3.67}$$

where F_z is the Fourier transform of $z(x,y)$ and we get

$$F_I = F_z(w_1,w_2)(-iw_1)\cos\tau\sin\sigma + F_z(w_1,w_2)(-iw_2)\sin\tau\sin\sigma, \tag{3.68}$$

where F_I is the Fourier transform of $I(x,y)$. The depth map $z(x,y)$ can be computed by rearranging the terms in the above equation and then taking the inverse Fourier transform.

This algorithm gives a non-iterative, closed-form solution using Fourier transform. The problem lies in the linear approximation of the reflectance map, which causes trouble when the non-linear terms are large.

3.5.3.2 Tsai and Shah

Tsai and Shah [25] employed the discrete approximations of p and q using finite differences in order to linearize the reflectance map in terms of z. The reflectance function for Lambertian surfaces is the same as equation (3.61). Using the discrete approximations of p and q

$$p = \frac{\partial z}{\partial x} = z(x,y) - z(x-1,y) \tag{3.69}$$

$$q = \frac{\partial z}{\partial y} = z(x,y) - z(x,y-1). \tag{3.70}$$

Then the image irradiance equation $I(x,y) = R(p,q)$ can be written as

$$f(z(x,y)) = I(x,y) - R(z(x,y) - z(x-1,y), z(x,y) - z(x,y-1)). \tag{3.71}$$

By taking the Taylor series expansion of this function about $z(x,y) = z^{n-1}(x,y)$, where $n-1$ denotes the iteration number, we have

$$\begin{aligned} 0 &= f(z(x,y)) \\ &\approx f(z^{n-1}(x,y)) + (z(x,y) - z^{n-1}(x,y)) \frac{df}{dz(x,y)}(z^{n-1}(x,y)). \end{aligned} \tag{3.72}$$

Then for $z(x,y) = z^n(x,y)$, the depth map at n-th iteration can be solved directly from

$$z^n(x,y) = z^{n-1}(x,y) + \frac{-f(z^{n-1}(x,y))}{\frac{d}{dz(x,y)}(z^{n-1}(x,y))}. \tag{3.73}$$

If we assume that the initial estimate $z^0(x,y)$ is set to zero for all image pixels, the depth map can be iteratively refined using equation (3.73). Gaussian smoothing is applied to the final depth map $z^n(x,y)$ to get a smoother result.

3.6 Photometric Stereo

Photometric stereo is another shape recovery technique used in computer vision. Here, there are two or more registered images obtained under different lighting conditions to remove the ambiguity inherent in a single image brightness measurement [8, 79, 80, 81]. If more than two images are available, additional unknown parameters of illuminant or surface reflectance can be recovered [82]. In photometric stereo, since recovery of surface orientation is completely local and very simple, involving little more than table lookup, it does lead to a great practical importance. Calibration of different surface materials and different lighting conditions is straightforward, requiring only observation of an object of known shape, such as a sphere, and construction of the lookup table.

Traditional *binocular stereo* techniques determine range by relating images of an object viewed from two different directions. If the correspondence between picture elements is known, then distance to the object can be calculated by triangulation. The idea of photometric stereo is to instead vary the direction of the incident illumination between successive images, while holding the viewing direction constant. Since the imaging geometry is not changed, the correspondence between image points is known a *priori*. The technique is photometric because it uses the radiance values recorded at a single image location in successive views, rather than the relative positions of displaced features.

Suppose two images $I_1(x,y)$ and $I_2(x,y)$ are obtained by varying the direction of incident illumination. Since there has been no change in the imaging geometry, each picture element (x,y) in the two images corresponds to the same object point and hence to the same gradient (p,q). The effect of varying the direction of incident illumination is to change the reflectance map $R(p,q)$ that characterizes the imaging situation. Then, let the reflectance maps corresponding to $I_1(x,y)$ and $I_2(x,y)$ be $R_1(p,q)$ and $R_2(p,q)$ respectively. The two views are characterized by two independent equations as

$$I_1(x,y) = R_1(p,q) \tag{3.74}$$

$$I_2(x,y) = R_2(p,q). \tag{3.75}$$

If these equations are linear and independent, there will be a unique solution for p and q. In general, however, equations (3.74) and (3.75) are nonlinear and we can have either no solutions or several solutions. Often it is better to use a third image as

$$I_3(x,y) = R_3(p,q) \tag{3.76}$$

rather than two different illuminating conditions. In some cases this makes the equations linear. More important, it can improve accuracy and increase the range of surface orientations over which a solutions can be obtained. The third image can also allow us to recover the surface albedo.

If there are n images, we can formulate the photometric stereo problem in terms of the minimization of brightness error as

$$\sum_{i=1}^{n} \int\int (I_i(x,y) - R_i(p,q))^2 \, dx \, dy, \tag{3.77}$$

where I_i is the brightness measured in the i-th image and R_i is the corresponding reflectance map. Note that other shape-from-shading constraints should be added to the functional (3.77) in order to obtain a well-posed problem.

Multiple images required for photometric stereo can be obtained by explicitly moving a single light source, by using multiple light sources calibrated with respect to each other, or by rotating the object surface and imagining hardware together to simulate the effect of moving a single light source. The equivalent

of photometric stereo can also be achieved in a single view, by using multiple illumination sources that can be separated by color.

3.7 Conclusions

Shape-from-shading techniques recover 3-D description of an object from a single view of the object. In section 3.5, we analyzed a total eight existing SFS algorithms and grouped them into three different categories: Global minimization techniques, global propagation techniques and local techniques. These groupings are based on the conceptual differences among the algorithms.

Unlike local approaches which provided a close form solution, most global approaches use iterative schemes. A few schemes, such as Ikeuchi and Horns's, Horn's and Zheng and Chellappa's approaches which are explained in section 3.5, use variational calculus to derive a nonlinear iterative scheme. There is no guarantee of correct convergence because of the nonlinear dependence of the reflectance map $R(p, q)$, the discrete formulation in these approaches, the unavailability of the suitable boundary conditions and the noise in real images. In some cases, an algorithm can give the correct solution by converging the global minimum; in others, it can get stuck at a local minimum. Sometimes, an algorithm can diverge and walk away from the correct solution, therefore, the initial value will greatly affect the speed of convergence and the solution to which an algorithm converges. In order to enforce proper convergence, the shape information at occluding boundaries are used as initial values in Ikeuchi and Horn's method.

Oliensis [20] proved that even the complete shape information at an occluding boundary does not well-determine the surface reconstruction. However, the surface orientation is uniquely determined at singular points. On the basis of Oliensis's idea, Bichsel and Pentland's algorithm starts at singular points and propagates the depth through a minimum downhill technique. The proper termination of the algorithm is guaranteed by propagating the depth away from the light source. The correctness of the solution depends on the initialization at the singular points.

The uniqueness of SFS can be proven under the condition that the light source direction is equal to, or symmetric around, the viewing direction [20]. With an initial known curve, the method of characteristic strips yields a unique solution, if the first derivative of surface depth is continuous. For other cases, the uniqueness is unknown.

If we consider local uniqueness instead of global uniqueness over the entire image, the uniqueness of a solution can be easily determined at singular points and occluding boundaries. These are the points at which we can determine the surface orientation directly from the image brightness. The brightness pattern in any arbitrary region could arise from an infinite number of different surfaces. However, the information at singular points and at occluding boundaries, can be used to constrain the possible solutions.

In the cases where there is a unique solution, existence is clear. However, there are cases where no solution exists. Horn, Szeliski and Yullie [83] presented some impossibly shaded images that could not have originated from a smooth continuous surface with uniform albedo and illumination. These cases can be detected through the examination of conditions for impossibly shaded images or by checking for the existence of singularities in the solution to shape from shading.

Implementations, comparsions and performance analysis of various shape-from-shading techniques are presented in "Analysis of Shape from Shading Techniques" [84].

CHAPTER 4.

INTEGRATING SFS CONSTRAINTS

4.1 Introduction

In general, image irradiance depends on surface geometry, scene illuminance, surface reflectance and imaging geometry. The shape-from-shading problem is to recover shape of surfaces from image irradiance. By assuming that illuminance, reflectance and imaging geometry are constant and known, image irradiance can be related directly to surface orientation. If $z(x,y)$ is a surface patch with constant albedo defined over a bounded planar image region, the relationship between the surface orientation at an image point (x,y) and the image irradiance there $I(x,y)$ is denoted by $R(p,q)$, where $p = z_x$ and $q = z_y$ are the first partial derivatives of the surface function at (x,y). Then, the shape-from-shading problem can be posed as a nonlinear, first-order partial differential equation in two unknowns, called the image irradiance equation $I(x,y) - R(p,q) = 0$. Surface orientation cannot be computed strictly locally because image irradiance provides a single measurement, while surface orientation has two independent components. Thus, some additional constraints should be imposed to the iterative scheme. Using the calculus of variations technique, the shape is iteratively recovered by minimizing a cost functional involving some constraints. The shape-from-shading problem, then, can be regarded as a constrained minimization problem mathematically. The SFS constraints used in the iterative scheme were presented in section 3.4 and exploring the effects of these constraints on the solution is addressed in this chapter.

4.2 Integration of Constraints

The shape-from-shading constraints presented in section 3.4 have been used in various combinations in several algorithms. Here, to explore the effects of these constraints, we combined all four SFS constraints in a single energy functional in order to obtain a general purpose algorithm. Thus, we construct the following energy functional:

$$F(p,q,z) = F_1(p,q) + \lambda F_2(p,q) + \mu F_3(p,q,z) + \beta F_4(p,q), \tag{4.1}$$

where

$$F_1(p,q) = \iint (I(x,y) - R(p,q))^2 \, dx \, dy \tag{4.2}$$

$$F_2(p,q) = \iint (p_x^2 + p_y^2 + q_x^2 + q_y^2) \, dx \, dy \tag{4.3}$$

$$F_3(p,q,z) = \iint ((z_x - p)^2 + (z_y - q)^2) \, dx \, dy \tag{4.4}$$

$$F_4(p,q) = \iint ((R_x - I_x)^2 + (R_y - I_y)^2) \, dx \, dy. \tag{4.5}$$

Here, F_1 is the brightness constraint which is based on the image irradiance equation. F_2 is the smoothness constraint and used to ensure the convergence of the iterative scheme. The third term F_3, represents the integrability constraint that enforces the resulting gradient be integrable. The last term, F_4 is called the intensity gradient constraint which comes from the requirement that the gradients of the reconstructed image to be close to the gradients of the input image. The parameters λ, μ and β are the scalars and each one determines the relative weight of the corresponding term in the energy functional.

In the next section, we will consider the problem of minimizing (4.1) and derive an iterative scheme by solving this constrained minimization problem.

4.3 Formulation

Under Lambertian surface, point light source and uniform albedo assumptions, the reflectance map can be written as

$$R(p,q) = \eta \frac{\cos\sigma - p\sin\sigma\cos\tau - q\sin\sigma\sin\tau}{\sqrt{1+p^2+q^2}}, \tag{4.6}$$

where η is the albedo, σ and τ are the slant and tilt angles of the illuminant respectively.

Using the calculus of variations, minimizing (4.1) is equivalent to solving the following Euler equations:

$$\begin{aligned} F_p - \frac{\partial}{\partial x}F_{p_x} - \frac{\partial}{\partial y}F_{p_y} &= 0 \\ F_q - \frac{\partial}{\partial x}F_{q_x} - \frac{\partial}{\partial y}F_{q_y} &= 0 \\ F_z - \frac{\partial}{\partial x}F_{z_x} - \frac{\partial}{\partial y}F_{z_y} &= 0, \end{aligned} \tag{4.7}$$

where subscripts denote the partial derivatives.

Let variables with primes ($'$) represent the values after updating and the variables without primes stand for the values before updating. Then, the iterative scheme can be written simply as

$$p' = p + \delta p, \qquad q' = q + \delta q \quad \text{and} \quad z' = z + \delta z. \tag{4.8}$$

By introducing a linear approximation of the reflectance map with respect to (p,q), we have

$$R(p',q') \approx R(p,q) + (p'-p)R_p(p,q) + (q'-q)R_q(p,q). \tag{4.9}$$

Then, using (4.8) and (4.9), F_p in (4.7) can be written as

$$\frac{1}{2}F_p = (\mu + R_p^2)\delta p + R_pR_q\delta q + \mu\delta z + (R-I)R_p + \mu(p - z_x). \tag{4.10}$$

The other terms in (4.7) can be derived similarly and the results are listed as follows:

$$\begin{aligned} \frac{1}{2}\frac{\partial}{\partial x}F_{p_x} &= -2(\lambda + \beta R_p^2)\delta p - 2\beta R_pR_q\delta q \\ &\quad +\lambda p_{xx} + \beta R_p(p_{xx}R_p + q_{xx}R_q - I_{xx}) \end{aligned} \tag{4.11}$$

$$\frac{1}{2}\frac{\partial}{\partial y}F_{p_y} = -2(\lambda+\beta R_p^2)\delta p - 2\beta R_p R_q \delta q + \lambda p_{yy} + \beta R_p(p_{yy}R_p + q_{yy}R_q - I_{yy}) \quad (4.12)$$

$$\frac{1}{2}F_q = R_p R_q \delta p + (\mu + R_q^2)\delta q + \mu\delta z + (R-I)R_q + \mu(q - z_y) \quad (4.13)$$

$$\frac{1}{2}\frac{\partial}{\partial x}F_{q_x} = -2\beta R_p R_q \delta p - 2(\lambda+\beta R_q^2)\delta q + \lambda q_{xx} + \beta R_q(p_{xx}R_p + q_{xx}R_q - I_{xx}) \quad (4.14)$$

$$\frac{1}{2}\frac{\partial}{\partial y}F_{q_y} = -2\beta R_p R_q \delta p - 2(\lambda+\beta R_q^2)\delta q + \lambda q_{yy} + \beta R_q(p_{yy}R_p + q_{yy}R_q - I_{yy}) \quad (4.15)$$

$$F_z = 0 \quad (4.16)$$

$$\frac{1}{2}\frac{\partial}{\partial x}F_{z_x} = \mu(\delta p - 2\delta z + z_{xx} - p_x) \quad (4.17)$$

$$\frac{1}{2}\frac{\partial}{\partial y}F_{z_y} = \mu(\delta q - 2\delta z + z_{yy} - q_y). \quad (4.18)$$

After substituting (4.10)-(4.18) into (4.7) and writing it in a matrix form, i.e., $Ax = B$, we obtain

$$\begin{pmatrix} 4(\lambda+\beta R_p^2)+\mu & R_pR_q(1+4\beta) & \mu \\ R_pR_q(1+4\beta) & 4(\lambda+\beta R_q^2)+\mu & \mu \\ 1 & 1 & -4 \end{pmatrix} \begin{pmatrix} \delta p \\ \delta q \\ \delta z \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}, \quad (4.19)$$

where

$$\begin{aligned} B_1 &= \lambda(p_{xx}+p_{yy}) + \mu(z_x - p) + R_p\{I - R \\ &\quad +\beta[(p_{xx}+p_{yy})R_p + (q_{xx}+q_{yy})R_q - I_{xx} - I_{yy}]\} \\ B_2 &= \lambda(q_{xx}+q_{yy}) + \mu(z_y - q) + R_q\{I - R \\ &\quad +\beta[(p_{xx}+p_{yy})R_p + (q_{xx}+q_{yy})R_q - I_{xx} - I_{yy}]\} \\ B_3 &= p_x + q_y - z_{xx} - z_{yy}. \end{aligned}$$

and $(\delta p, \delta q, \delta z)$ are the increments in (p, q, z). Solving (4.19) for $(\delta p, \delta q, \delta z)$, we obtain the updating scheme

$$\begin{aligned}
\delta p &= \frac{1}{\Delta}[A_{22}(B_1 + \frac{1}{4}\mu B_3) - A_{12}(B_2 + \frac{1}{4}\mu B_3)] \\
\delta q &= \frac{1}{\Delta}[A_{11}(B_2 + \frac{1}{4}\mu B_3) - A_{21}(B_1 + \frac{1}{4}\mu B_3)] \\
\delta z &= \frac{1}{4}(\delta p + \delta q - B_3),
\end{aligned} \tag{4.20}$$

where

$$\begin{aligned}
A_{11} &= 4\lambda + \frac{5}{4}\mu + R_p^2(1 + 4\beta) \\
A_{12} &= A_{21} = \frac{1}{4}\mu + R_p R_q(1 + 4\beta) \\
A_{22} &= 4\lambda + \frac{5}{4}\mu + R_q^2(1 + 4\beta) \\
\Delta &= A_{11}A_{22} - A_{12}^2.
\end{aligned}$$

Then, the updating scheme is written as

$$\begin{aligned}
p^{k+1} &= p^k + \delta p \\
q^{k+1} &= q^k + \delta q \\
z^{k+1} &= z^k + \delta z.
\end{aligned} \tag{4.21}$$

4.4 Effects of Constraints on Solution

To evaluate the effects of the constraints, we need choose some error measures. Horn [12] discusses a number of possible quality measures and methods for displaying the algorithm's progress. Here, we chose to study three different measures:

- The energy (or cost) $F^{(k)}$ associated with the current estimate and defined as

  $$F^{(k)} = F_1^{(k)} + \lambda F_2^{(k)} + \mu F_3^{(k)} + \beta F_4^{(k)},$$

  where superscripts indicate the iteration index. This measure actually is the quantity we are trying to minimize in SFS problem.

- The norm of the gradient error

  $$| p^{(k)} - p^* | + | q^{(k)} - q^* |,$$

where (p^*, q^*) is the gradient of the true solution.

- The norm of the height error

$$| z^{(k)} - z^* |,$$

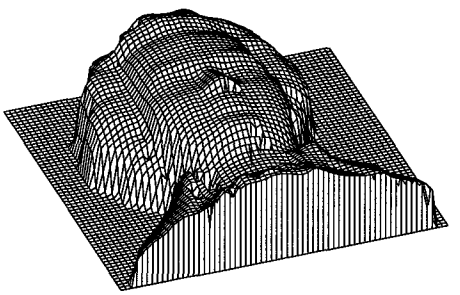
where z^* is the height of the true solution.

To see the effects of the constraints on the SFS solution, the iterative scheme we derived in the previous section is used. The synthetically generated Mozart image with size 128x128 shown in Figure 4.1 is used as the input image to the algorithm. Figure 4.1(a) shows a 3-D plot of the true height map whose size is reduced to 64x64. Figure 4.1(b) is the true height map obtained by a laser range finder used to generate the synthetic input image. Figure 4.1(c) shows the input image which is synthesized from the true height map with parameter $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$. Figure 4.1(d) shows a profile of the true height map. In this plot, the horizontal axis corresponds to the 64-th line in the true height map. That is, it forms a line parallel to the x-axis running from left to right through the center of the image.

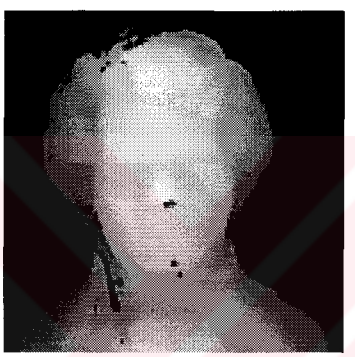
The first term in the energy functional (4.1)

$$\int\int (I(x,y) - R(p,q))^2 \, dx \, dy \tag{4.22}$$

measures the error in brightness, that is the difference between the observed image brightness and that predicted from the computed shape. This mean-squared error term allows for modeling errors and noise, and used as a part of an energy functional which is minimized with respect to (p, q) to find the solution of the SFS problem. Each parameter λ, μ or β in the energy functional (4.1), establishes a tradeoff between the brightness error and the corresponding constraint term. In the following subsections, we use various combinations of these three parameters and observe the effects of each constraint by changing its parameter value.

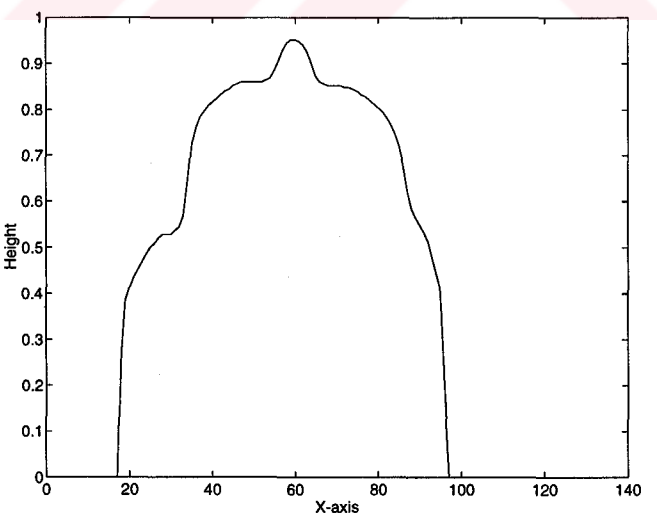


(a)



(b)

(c)



(d)

Figure 4.1. Mozart test image: (a) A 3-D plot of the true height map, (b) the true height map used to generate the synthetic image, (c) the input image which is synthesized from (a) with parameters $\eta = 250$, $\tau = 45°$ and $\sigma = 45°$, (d) a profile of the true height map.

4.4.1 Effects of the Smoothness Constraint

The smoothness constraint is defined as

$$\lambda \iint (p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy, \tag{4.23}$$

p_x, p_y, q_x and q_y are the partial derivatives of the surface gradient and λ is a scalar which determines the weight of the smoothness term in the energy functional (4.1). This constraint which is also called regularization term is used to stabilize the iterative shape-from-shading scheme and to ensure that it has a unique minimum.

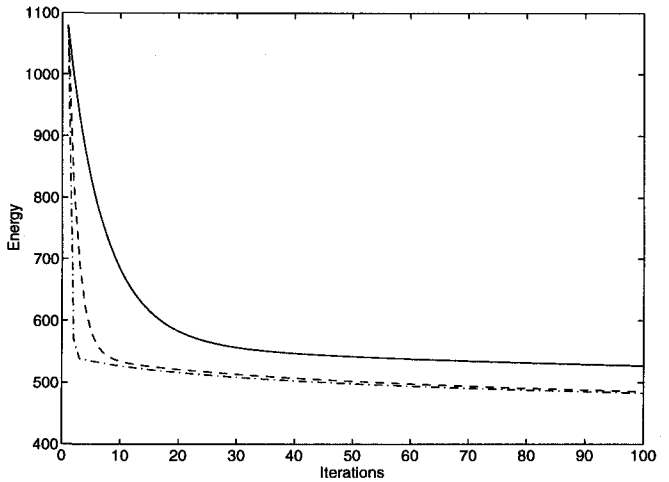
The synthetically generated Mozart image shown in Figure 4.1(c) is used as an input image to the iterative algorithm. For the algorithm parameters, we set $\mu = 1$ and $\beta = 0$, and three different values to the regularization parameter λ as 5, 1 and 0.01 respectively. The obtained results are shown in Figure 4.2. The plots shown here are the three error measures in the first 100 iterations. The first plot shows the energy F as a function of iteration number. From this plot, we see that decreasing the value of the regularization parameter also decreases the total energy in the minimization process. The second plot shows the average gradient error. It can be seen from Figure 4.2(b) that for $\lambda = 5$, this error is small and decreasing λ causes more error. We also note that, for $\lambda = 1$ and $\lambda = 0.01$, the p-q error is increasing after an initial improvement. The third plot in Figure 4.2 shows the average height error obtained from the iterative algorithm. In this plot, we have continuously decreasing height error values for $\lambda = 5$ while the iterations are proceeded. For $\lambda = 1$, although the height error is small for the first 80 iterations, this error is started to increase for the further iterations. But the worst case is obtained for $\lambda = 0.01$ and the height error decreases for the first 25 iterations and then starts to going up. These results shows that the optimum choice of the smoothness constraint parameter value is critical and too small values may cause instabilities in the iterative algorithm even they caused smaller error.

In Figure 4.3, the shaded images synthesized from the (p, q) maps obtained by the iterative SFS algorithm using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$ for three different λ values are shown. Looking at these shaded images,

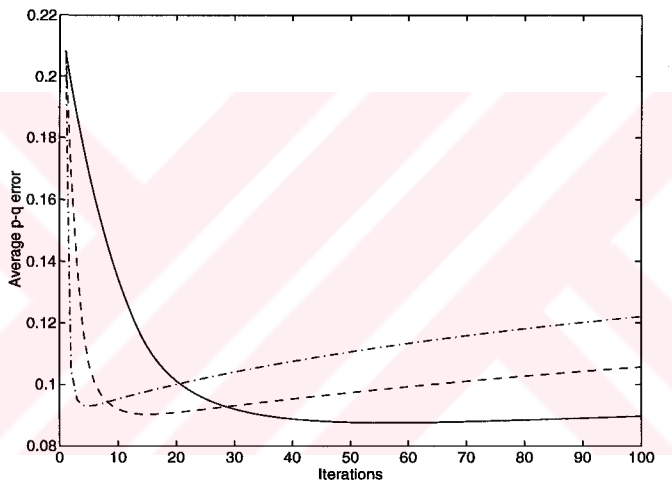
we see that the solution being computed is too smooth for $\lambda = 5$. On the other hand, decreasing the regularization parameter value improves the results as shown in Figure 4.3(b) and Figure 4.3(c).

The one dimensional slices compared in Figure 4.4 better illustrate how the reconstructed surface changes according to the regularization parameter value. In these plots, the horizontal axis corresponds to the 64-th line in the reconstructed height maps. The smoothing effect of a large value of λ can also be seen from these representative 1-D cuts through the surfaces obtained by the algorithm.

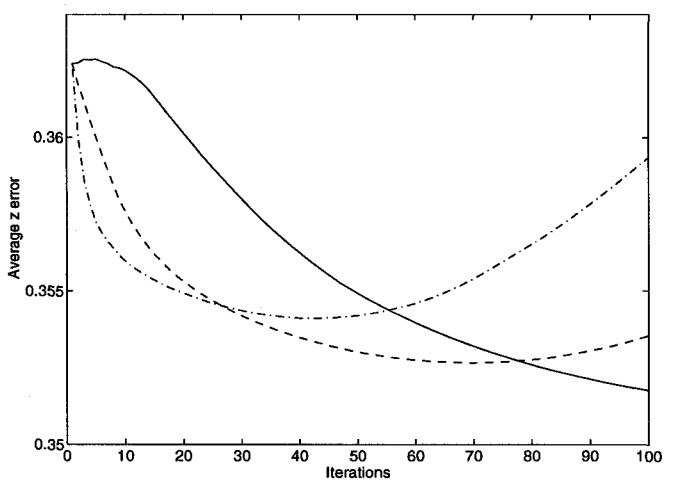
As a result, using of the regularization term in the iterative SFS algorithm produces solutions that are too smooth, with the amount of distortion depending on the choice of the parameter λ.



(a)



(b)



(c)

Figure 4.2. Error terms: (a) Energy, (b) p-q error and (c) z error for $\lambda = 5$ (solid), $\lambda = 1$ (dashed) and $\lambda = 0.01$ (dash-dotted) when $\mu = 1$ and $\beta = 0$.

(a) $\lambda = 5$ (b) $\lambda = 1$ (c) $\lambda = 0.01$

Figure 4.3. SFS results for different λ values when $\mu = 1$ and $\beta = 0$.

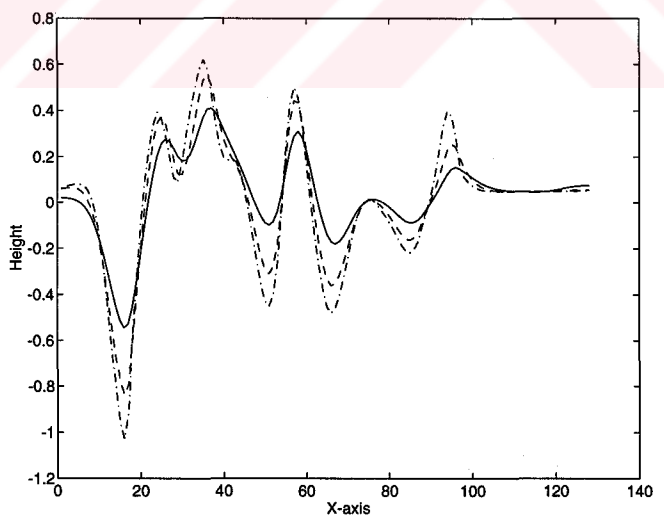


Figure 4.4. One-dimensional slices showing the effects of the smoothness constraint for $\lambda = 5$ (solid), $\lambda = 1$ (dashed) and $\lambda = 0.01$ (dash-dotted) when $\mu = 1$ and $\beta = 0$.

4.4.2 Effects of the Integrability Constraint

The integrability constraint is

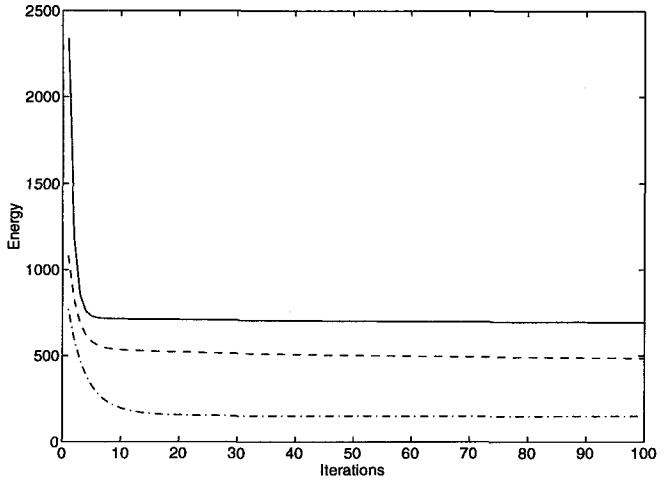
$$\mu \iint ((z_x - p)^2 + (z_y - q)^2)\, dx\, dy, \tag{4.24}$$

where z_x and z_y are the partial derivatives of the height field z, and μ is a weighting factor. This constraint is used to ensure that the gradient field (p, q) computed by the shape-from-shading algorithm corresponds to a physical surface.

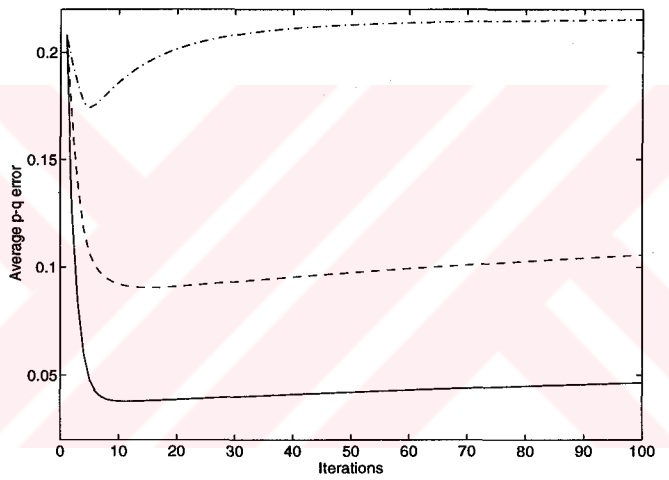
To explore the effects of the integrability constraint, we used our iterative algorithm by setting $\lambda = 1$ and $\beta = 0$. For the integrability constraint parameter μ, we set three values as 5, 1 and 0.01 respectively and obtained the error measures shown in Figure 4.5. As we are expecting, small values of μ cause small energy values as shown in Figure 4.5(a). However, by looking at the average gradient error in Figure 4.5(b), we can say that, this error increases while μ is decreasing. The largest error values are obtained for $\mu = 0.01$. Reasonable solutions are found for $\mu = 5$ and $\mu = 1$. If we look at the plots in Figure 4.5(c), one can observe that the height error is small for $\mu = 0.01$ up to the first 30 iterations but increases rapidly while the further iterations are proceeded. The reasonable results, on the other hand, are obtained for $\mu = 5$ and $\mu = 1$. The results in Figure 4.5 show that, although a big value of μ, i.e., $\mu = 5$, produces small p-q error, it causes more distortion in depth map. Conversely, we have rather small z error for $\mu = 0.01$ up to the first 30 iterations while p-q error is high. However, a reasonable solution for this example is obtained for $\mu = 1$.

The shaded images synthesized from the obtained gradient maps by the SFS algorithm for $\lambda = 1$, $\beta = 0$ and three different values of μ are shown in Figure 4.6. By looking at these images, we see that the two solutions obtained for $\mu = 5$ and $\mu = 0.01$ as shown in Figure 4.6(a) and Figure 4.6(c) are smoother than the one obtained for $\mu = 1$ as shown in Figure 4.6(b).

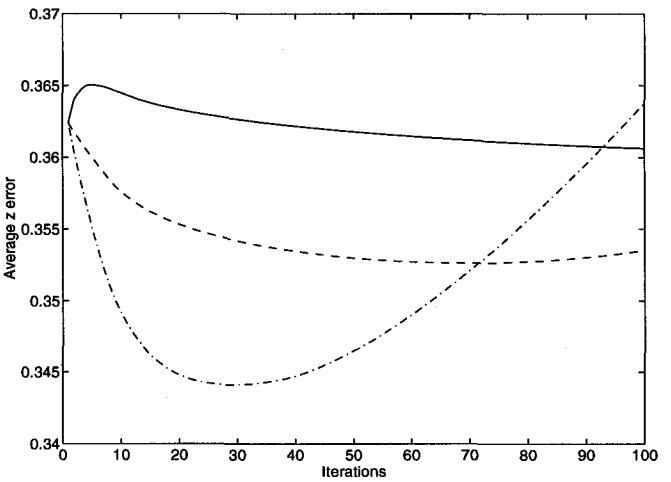
The one-dimensional slices comparing three different values of μ are shown in Figure 4.7. The smoothing effect of large μ values can be clearly seen from these plots.



(a)



(b)



(c)

Figure 4.5. Error terms: (a) Energy, (b) p-q error and (c) z error for $\mu = 5$ (solid), $\mu = 1$ (dashed) and $\mu = 0.01$ (dash-dotted) when $\lambda = 1$ and $\beta = 0$.

(a) $\mu = 5$ (b) $\mu = 1$ (c) $\mu = 0.01$

Figure 4.6. SFS results for different μ values when $\lambda = 1$ and $\beta = 0$.

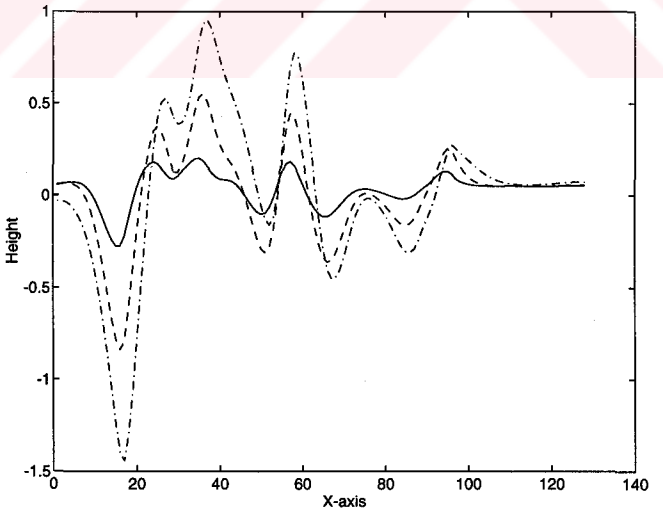


Figure 4.7. One-dimensional slices showing the effects of the integrability constraint for $\mu = 5$ (solid), $\mu = 1$ (dashed) and $\mu = 0.01$ (dash-dotted) when $\lambda = 1$ and $\beta = 0$.

4.4.3 Effects of the Intensity Gradient Constraint

The intensity gradient constraint is described as

$$\beta \iint ((R_x - I_x)^2 + (R_y - I_y)^2)\, dx\, dy, \tag{4.25}$$

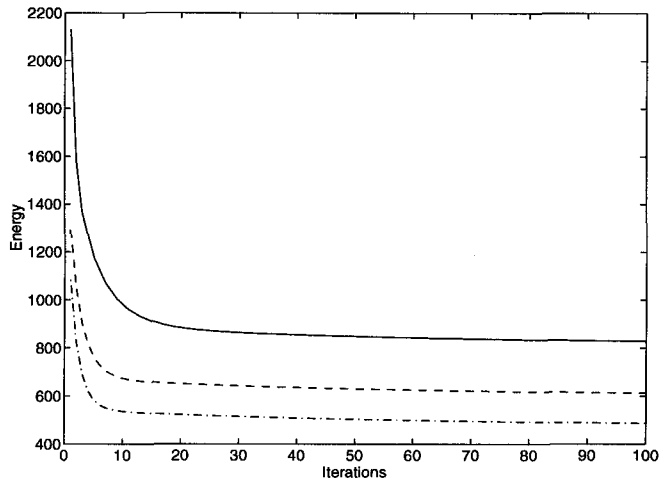
where I_x and I_y are the partial derivatives of the input image, while R_x and R_y are the partial derivatives of the reflectance map and β is a scalar. This constraint is used to require that the gradients of the reconstructed image to be close to the gradients of the input image in both x and y directions.

Here, we used the iterative SFS algorithm by setting $\lambda = 1$, $\mu = 1$ and three different values to β. The plots in Figure 4.8 show the convergence of the iterative algorithm for values 5, 1 and 0.01 of β. The total energy is low for small values of β in Figure 4.8(a) because of the less contribution of the intensity gradient constraint in the energy functional. It can be seen from Figure 4.8(b) and Figure 4.8(c) that both the p-q error and the z error are high for a large value of the intensity gradient coefficient, i.e., $\beta = 5$. Moreover, the algorithm converges to a reasonable solution for both the surface slopes and the height map by using small values of β. Note that, for $\beta = 1$ and $\beta = 0.01$, there are no significant changes in the corresponding gradient error and the height error and reasonable solutions are found using these parameter values.

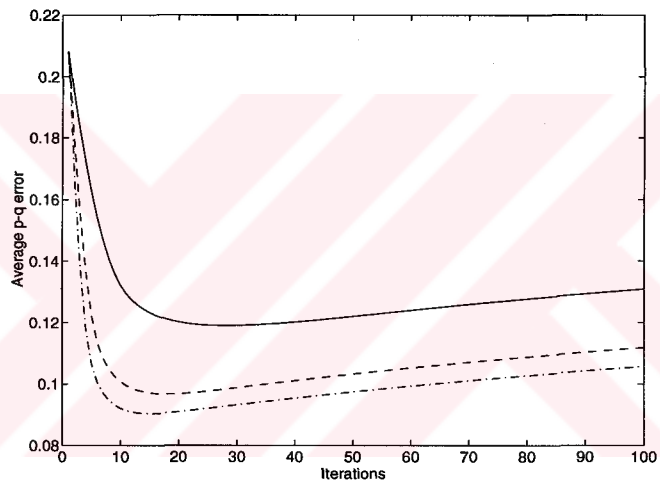
The shaded images synthesized from the (p, q) maps obtained by the SFS algorithm using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$ are shown in Figure 4.9. The image shown in Figure 4.9(a) is quite similar to the input image shown in Figure 4.1(c) but while β is decreasing, the shaded images become more smoother as shown in Figure 4.9(b) and Figure 4.9(c) for $\beta = 1$ and $\beta = 0.01$ respectively.

If we compare the one-dimensional slices obtained by the SFS algorithm for three different β values shown in Figure 4.10, one can easily see that changing the intensity gradient parameter value does not cause significant changes in the height map. Only the peaks are suppressed while β is decreasing.

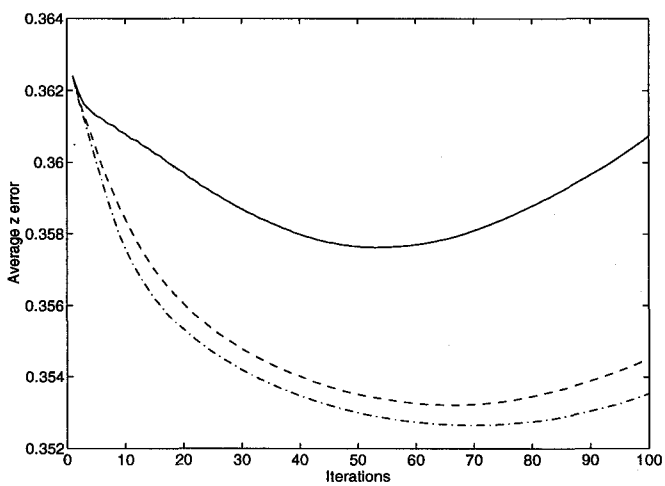
To see how the intensity gradient constraint effects the solution, we can look at the partial derivatives of the input image, I_x and I_y, and the partial



(a)



(b)



(c)

Figure 4.8. Error terms: (a) Energy, (b) p-q error and (c) z error for $\beta = 5$ (solid), $\beta = 1$ (dashed) and $\beta = 0.01$ (dash-dotted) when $\lambda = 1$ and $\mu = 1$.

(a) $\beta = 5$ (b) $\beta = 1$ (c) $\beta = 0.01$

Figure 4.9. SFS results for different β values when $\lambda = 1$ and $\mu = 1$.

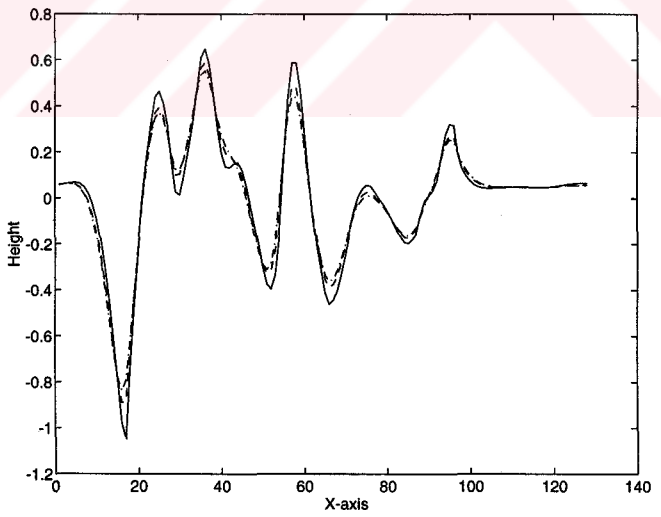


Figure 4.10. One-dimensional slices showing the effects of the intensity gradient constraint for $\beta = 5$ (solid), $\beta = 1$ (dashed) and $\beta = 0.01$ (dash-dotted) when $\lambda = 1$ and $\mu = 1$.

derivatives of the reflectance map, R_x and R_y. The derivatives of the input image are shown in Figure 4.11. The derivative images obtained by the algorithm for $\lambda = 1$, $\mu = 1$ and $\beta = 0$ are shown in Figure 4.12. If we compare the images in Figure 4.11 and Figure 4.12, we see that the partial derivatives of the input image and the partial derivatives of the reconstructed image do not look similar when the intensity gradient constraint is not in the energy functional, i.e., $\beta = 0$. In order to see the effect of the intensity gradient term on the derivative images, we run the iterative algorithm with parameters $\lambda = 1$, $\mu = 1$ and $\beta = 5$, and the resulting derivative images are shown in Figure 4.13. If we compare I_x and I_y in Figure 4.11 to R_x and R_y in Figure 4.13, the similarities between the partial derivative images can be seen easily. These results show how this constraint effects the solution.

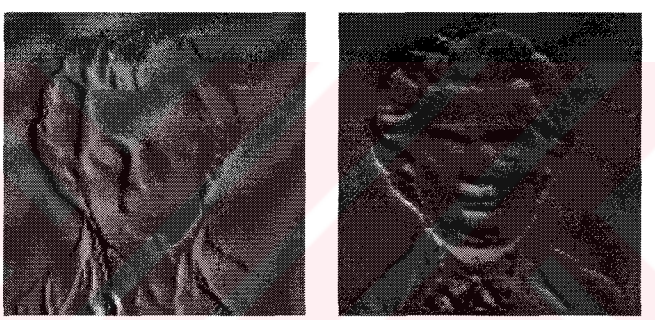
Now, we want to see the effects of the intensity gradient constraint when the smoothness constraint is dropped. This time, the algorithm parameters are $\lambda = 0$, $\mu = 1$. For β, we set three different values and run the SFS algorithm to see how the absence of the smoothness constraint effects the results while β is changing. The error measures obtained in this situation are shown in Figure 4.14. If we compare these results to those shown in Figure 4.8 which are obtained when $\lambda = 1$, we can see that, the iterative algorithm may become unstable for the large values of β when the smoothness constraint is not present in the iterative process. However, there are no significant changes on p-q error and z error while β is small even if $\lambda = 0$.

The shaded images synthesized from the reconstructed (p, q) maps obtained by the algorithm are shown in Figure 4.15. These images are less smoother than the images obtained when $\lambda = 1$ which are shown in Figure 4.9. This is because of the absence of the smoothness constraint. However, the smoothing effect of the intensity gradient constraint, while β is decreasing, can also be seen from the shaded images.

The one-dimensional slices comparing three different values of β when $\lambda = 0$ are shown in Figure 4.16. These plots look quite similar to ones obtained when $\lambda = 1$ which are shown in Figure 4.9.

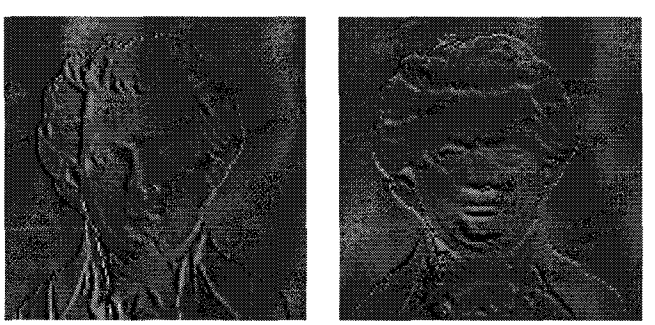
(a) I_x (b) I_y

Figure 4.11. Partial derivatives of the input image.



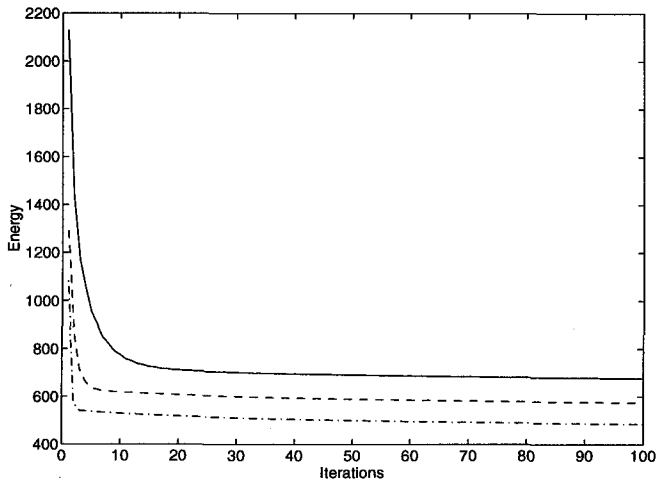
(a) R_x (b) R_y

Figure 4.12. Partial derivatives of the reconstructed image for $\lambda = 1$, $\mu = 1$ and $\beta = 0$.

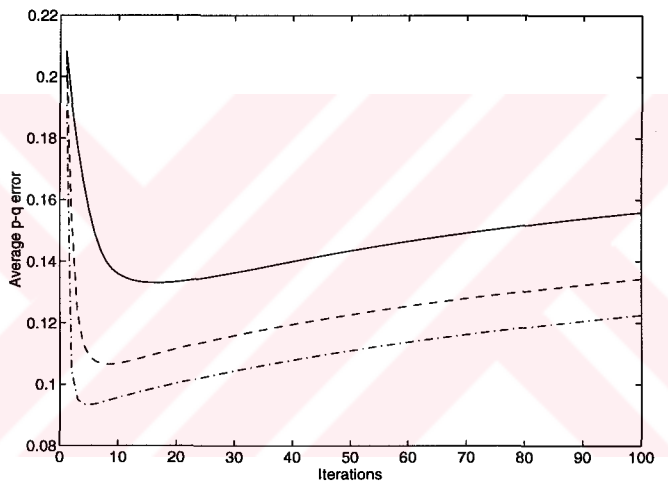


(a) R_x (b) R_y

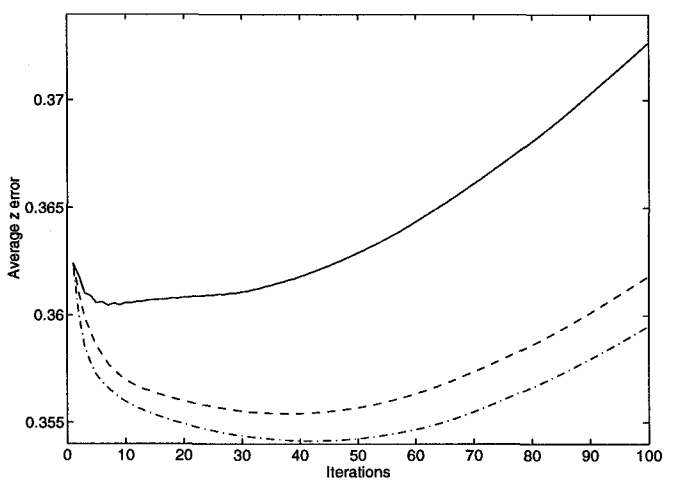
Figure 4.13. Partial derivatives of the reconstructed image for $\lambda = 1$, $\mu = 1$ and $\beta = 5$.



(a)



(b)



(c)

Figure 4.14. Error terms: (a) Energy, (b) p-q error and (c) z error for $\beta = 5$ (solid), $\beta = 1$ (dashed) and $\beta = 0.01$ (dash-dotted) when $\lambda = 0$ and $\mu = 1$.

(a) $\beta = 5$ (b) $\beta = 1$ (c) $\beta = 0.01$

Figure 4.15. SFS results for different β values when $\lambda = 0$ and $\mu = 1$.

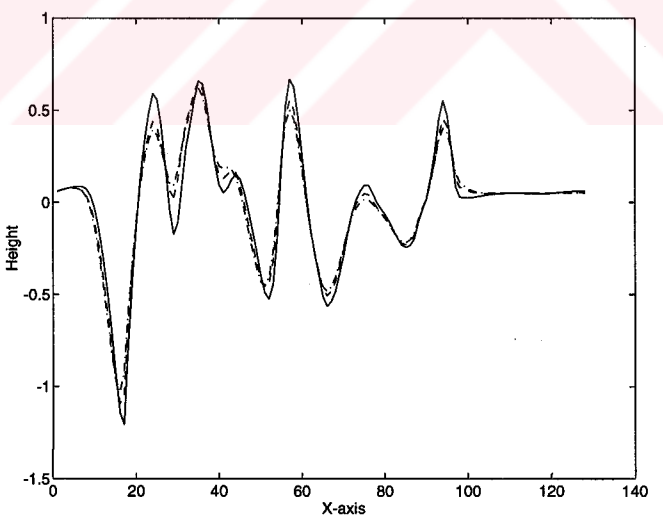


Figure 4.16. One-dimensional slices showing the effects of the intensity gradient constraint with $\beta = 5$ (solid), $\beta = 1$ (dashed) and $\beta = 0.01$ (dash-dotted) when $\lambda = 0$ and $\mu = 1$.

The experimental results presented in this chapter are obtained by using various combinations of the three algorithm parameter values. When we are observing individual effects of each constraint, the corresponding parameter value of this constraint is changing, while other two parameters are fixed to constant values. For evaluating the effects of the smoothness constraint, for example, we set three different values to the regularization parameter λ as 5, 1 and 0.01, while we set constant values for other two parameters as $\mu = 1$ and $\beta = 0$. Additionally, we also use various constant values for the fixed parameters in order to see if they affect the solution or not. The experimental results we obtained show that changing the values of the fixed parameters does not effect the solution significantly. Therefore, results obtained by changing the fixed parameter values are not presented here.

On the other hand, we also obtained the same error measures, shaded images and one-dimensional profiles by using four more synthetic input images. When we compare the results obtained from different input images, we see that the algorithm parameters affect the SFS solutions in the same way. For this reason, we here only presented the results which are found by using the Mozart image.

Based on these experiments, although algorithm parameters μ and β affect SFS solution, the regularization parameter λ is the most effective one for the iterative scheme. The SFS results will closely depend on the value of λ used. The choice of λ is critical, as it should be large enough to avoid instability and small enough to obtain a meaningful solution. From the experimental results, the edges in the reconstructed images become sharper when the smoothness constraint is weakened (small λ's) but too small λ values may cause instability in the iterative SFS algorithm. On the other hand, the solution being computed is too smooth for large values of λ. In order to overcome the difficulty in determining the regularization parameter value, an automatic method for adjusting the value of λ needs to be developed.

In the next chapter, an adaptive shape-from-shading method will be developed and implemented by controlling the smoothness spatially over the image space.

CHAPTER 5.

ADAPTIVE SHAPE FROM SHADING

5.1 Introduction

In the previous chapter, we experimentally observed that, although the regularization term helps to stabilize the minimization process, it flattens the SFS reconstruction and causes distortions along image discontinuities. The regularization parameter, λ, determines the tradeoff between the smoothness constraint and the other constraints in an energy functional and the choice of its optimum value is critical. It should be large enough to avoid instability and small enough to produce a meaningful solution. On the other hand, the smoothness term which is used in the energy functional to enforce the condition that the reconstructed surface should be smoothly connected is quadratic and suppresses changes in intensities. In areas where image intensity changes rapidly, the corresponding surface patch may not be smooth. The solution of these problems is to use an adaptive smoothing function to reduce the over-smoothing along the reconstructed image.

In this chapter, we introduce a new adaptive shape-from-shading approach in which the smoothness is controlled spatially over the image space [6, 85, 86]. The adaptive nature of the algorithm eliminates the selection of the optimum value of the smoothness parameter and protects the iterative scheme "to walk away" from the correct solution. The spatial control of smoothness is achieved by employing additional knowledge about the difference between the image which is obtained from the regularized solution and the input image. Using

calculus of variations and a linear approximation of the reflectance map, a new iterative scheme which updates the slope and the height maps simultaneously is developed.

5.2 Formulation

In our implementation, we combine all four SFS constraints which are presented in section 3.4 in a single energy functional and use a spatially controlled regularization parameter instead of a fixed regularization parameter in order to adaptively control the smoothness over the image space. To achieve this, we construct the following energy functional

$$F(p,q,z) = F_1(p,q) + F_2'(p,q) + \mu\, F_3(p,q,z) + \beta\, F_4(p,q), \tag{5.1}$$

where

$$F_1(p,q) = \int\!\!\int (I(x,y) - R(p,q))^2\, dx\, dy \tag{5.2}$$

$$F_2'(p,q) = \int\!\!\int \lambda(x,y)\,(p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy \tag{5.3}$$

$$F_3(p,q,z) = \int\!\!\int ((z_x - p)^2 + (z_y - q)^2)\, dx\, dy \tag{5.4}$$

$$F_4(p,q) = \int\!\!\int ((R_x - I_x)^2 + (R_y - I_y)^2)\, dx\, dy. \tag{5.5}$$

Here, F_1 is the brightness constraint, F_2' is the spatially controlled constraint, where

$$\int\!\!\int (p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy \tag{5.6}$$

represents the smoothness constraint, F_3 is the integrability constraint and finally, F_4 is the intensity gradient constraint. Note that, in the energy functional (5.1), the regularization parameter $\lambda(x,y)$ is a function of spatial coordinates while μ and β are scalars.

In our adaptive SFS approach, the smoothness of the solution is spatially determined by the value $\lambda(x,y)$ at each image point. That is, a large $\lambda(x,y)$ pushes the reconstruction toward a smooth surface, while a small $\lambda(x,y)$ forces the solution to be close to the input image data $I(x,y)$ by weakening the

connection between successive pixels. An important task in our approach is to determine what to use in controlling the regularization parameter and how to use it. As indicated before, we would like to use a somewhat small $\lambda(x,y)$ in the vicinity of a discontinuity for the accurate shape recovery. For this reason, our choice for the control of $\lambda(x,y)$ should provide some information about discontinuities. The error or residual between the input image intensity values and the regularized solution, i.e., $I(x,y)-R(p,q)$, exhibits some useful features for this purpose. Thus, the error signal is defined as

$$e = I(x,y) - R(p,q) \tag{5.7}$$

and indicates the existence of discontinuity at the corresponding image point.

In the proposed algorithm, the control signal is extracted from the error signal defined above in order to control the smoothness of the solution by adjusting the regularization parameter $\lambda(x,y)$. Let c be the control signal, then

$$c = abs(e), \tag{5.8}$$

where $abs(e)$ is the absolute value of the error signal $e = I - R$. Taking the absolute value of the error signal obtains a signal which has positive values on both sides of the discontinuity. The positive values in this signal are then used to control the parameter $\lambda(x,y)$. Since the control signal has large positive values in the vicinity of discontinuities, this signal can be used to reduce the amount of smoothing in these regions by decreasing the value of $\lambda(x,y)$ at these points. To determine the new value of the regularization parameter at location (x,y), we use the following decreasing function:

$$\lambda_{new}(x,y) = \begin{cases} \mathcal{F}(x,y,\lambda_{old}) & \text{if } c(x,y) > 0 \text{ and } \lambda_{old} > \lambda_{min} \\ \lambda_{old}(x,y) & otherwise \end{cases} \tag{5.9}$$

where

$$\mathcal{F}(x,y,\lambda_{old}) = (1 - e^{-\frac{c(x,y)}{V_T}})\,\lambda_{min} + (e^{-\frac{c(x,y)}{V_T}})\,\lambda_{old}(x,y), \tag{5.10}$$

$c(x,y)$ is the control signal as described in (5.8), V_T is a time-constant that controls the rate of exponential decrease and λ_{min} is a preselected minimum value that $\lambda(x,y)$ may have. This is an exponentially decreasing function with the following properties:

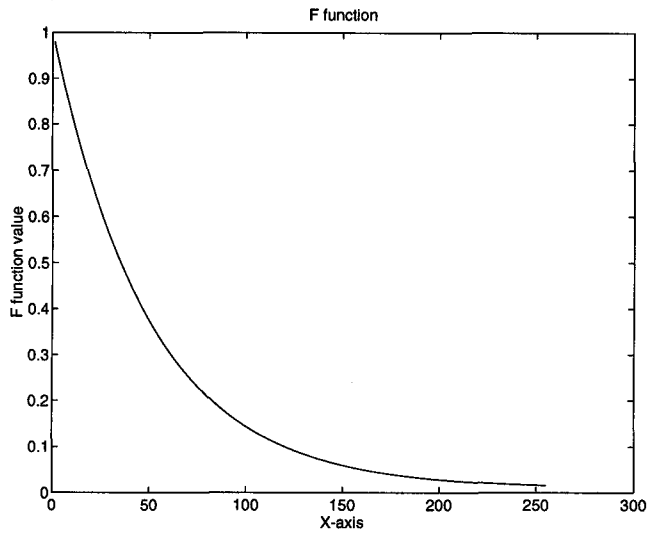


Figure 5.1. A profile of the $\mathcal{F}$ function which is used to calculate the regularization parameter values.

$$\lim_{c(x,y)\to 0} \mathcal{F}(x,y,\lambda_{old}) = \lambda_{old}(x,y) \quad \text{and} \quad \lim_{c(x,y)\to \infty} \mathcal{F}(x,y,\lambda_{old}) = \lambda_{min}. \quad (5.11)$$

Thus the regularization parameter is gradually decreased by this control scheme. A profile of the function $\mathcal{F}(x,y,\lambda_{old})$ obtained by setting $\lambda_{min} = 0.01$, $V_T = 50$, $\lambda_{old}(x) = \lambda_{old} = 1$, $1 \leq x \leq 255$ and $0 \leq c(x) \leq 255$ is shown in Figure 5.1.

Figure 5.2 shows the experimental results obtained by the SFS algorithm developed in chapter 4. The Mozart image in Figure 4.1 is used as the input image and initial settings of the algorithm parameters are $\lambda = 5$, $\mu = 1$ and $\beta = 1$. The input image and the obtained regularized solution profiles are shown in Figure 5.2(a). In these plots, the regularized solution is too smooth because of using the constant λ value. The error signal and extracted control signal for this example are shown in Figure 5.2(b) and Figure 5.2(c) respectively. Figure 5.2(d) shows the calculated regularization parameter values.

Figure 5.3 demonstrates the iterative refinement process on the Mozart test image. Figure 5.3(a) shows the regularized solution obtained with a large regularization parameter $\lambda(x,y) = 5$, by which the discontinuity and the continuous regions as well as smoothed out. Figure 5.3(b) shows the refined solution after 3 iterations, where the gap between the input image and the regularized solution is reduced around discontinuities. Figure 5.3(c) shows the iterative refinement

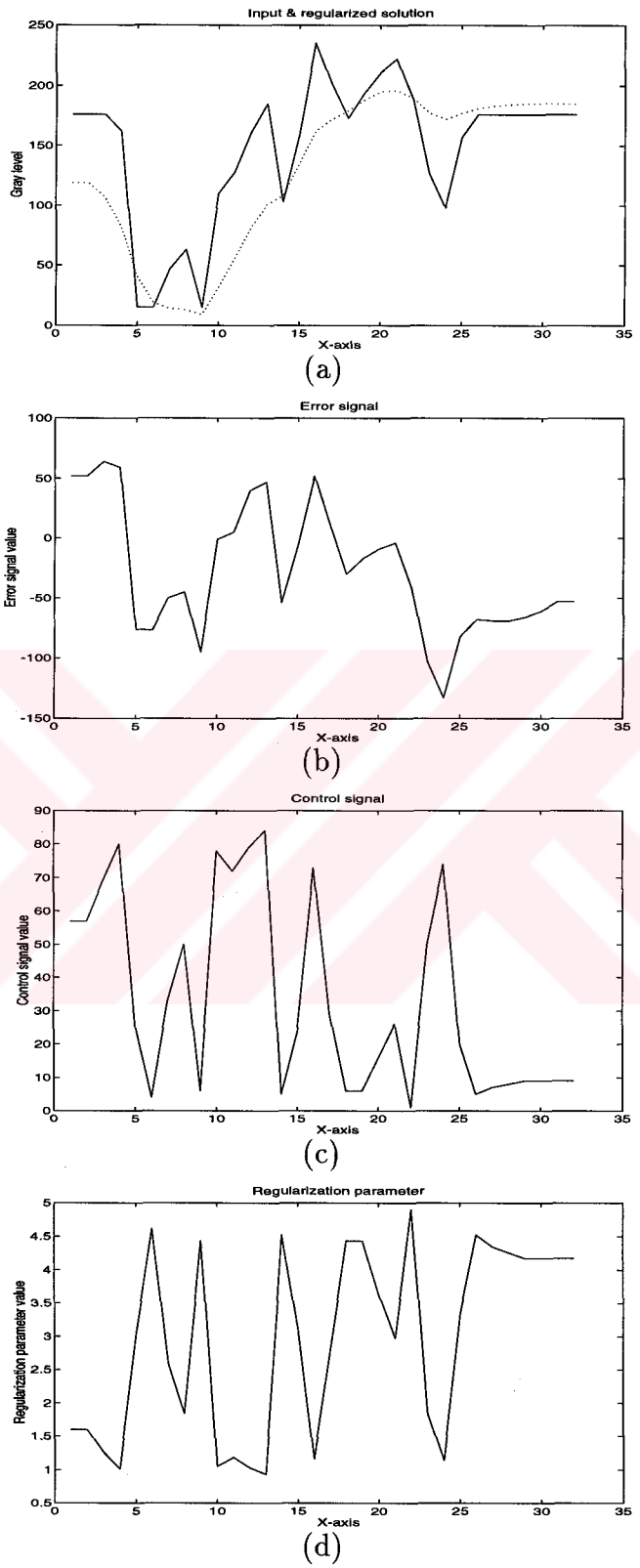
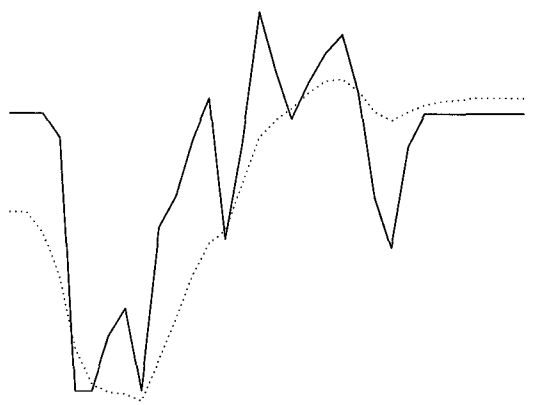
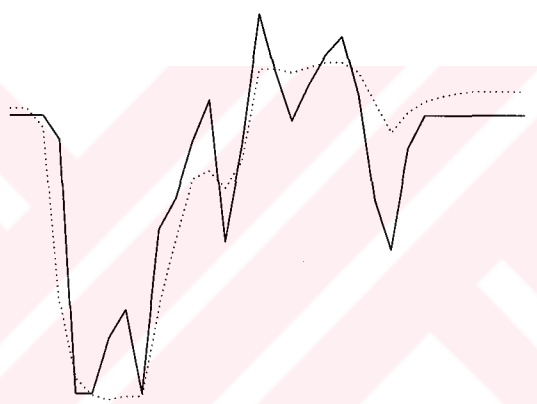


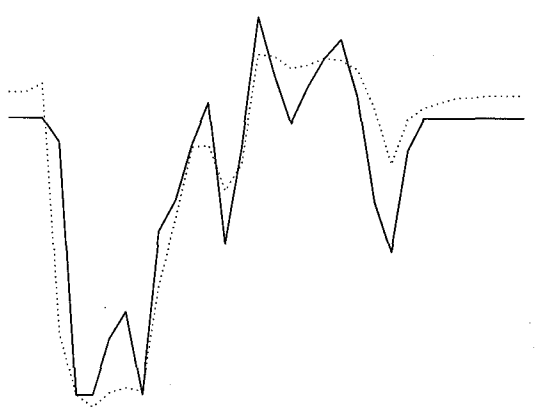
Figure 5.2. (a) The input (solid) and the regularized solution (dotted) profiles, (b) the error signal, (c) the control signal and (d) the regularization parameter $\lambda(x, y)$ values.



(a)



(b)



(c)

Figure 5.3. Evolution of the Mozart test image: The input image (solid) and the regularized solution (dotted) profiles, (a) standard regularization, (b) iterative refinement after 3 iterations, (c) iterative refinement after 5 iterations.

after 5 iterations. More iterations can be done to improve the reconstruction.

Under Lambertian surface, point light source and uniform albedo assumptions, the reflectance map can be written as

$$R(p,q) = \eta \frac{\cos\sigma - p\sin\sigma\cos\tau - q\sin\sigma\sin\tau}{\sqrt{1+p^2+q^2}}, \tag{5.12}$$

where η is the albedo, σ and τ are the slant and tilt angles of the illuminant respectively. Using the calculus of variations, minimizing (5.1) is equivalent to solving the following Euler equations:

$$\begin{aligned} F_p - \frac{\partial}{\partial x}F_{p_x} - \frac{\partial}{\partial y}F_{p_y} &= 0 \\ F_q - \frac{\partial}{\partial x}F_{q_x} - \frac{\partial}{\partial y}F_{q_y} &= 0 \\ F_z - \frac{\partial}{\partial x}F_{z_x} - \frac{\partial}{\partial y}F_{z_y} &= 0, \end{aligned} \tag{5.13}$$

where subscripts denote the partial derivatives.

If the variables with primes ($'$) represent the values after updating and the variables without primes stand for the values before updating, then, the iterative scheme simply can be written as

$$p' = p + \delta p, \qquad q' = q + \delta q \quad \text{and} \quad z' = z + \delta z. \tag{5.14}$$

The corresponding increments in the partial derivatives of (p,q,z) after updating are

$$\begin{array}{lll} p'_x = p_x - \delta p & q'_x = q_x - \delta q & z'_x = z_x - \delta z \\ p'_y = p_y - \delta p & q'_y = q_y - \delta q & z'_y = z_y - \delta z \\ p'_{xx} = p_{xx} - 2\delta p & q'_{xx} = q_{xx} - 2\delta q & z'_{xx} = z_{xx} - 2\delta z \\ p'_{yy} = p_{yy} - 2\delta p & q'_{yy} = q_{yy} - 2\delta q & z'_{yy} = z_{yy} - 2\delta z. \end{array} \tag{5.15}$$

Further, by introducing a linear approximation of the reflectance map with respect to (p,q), we have

$$R(p',q') \approx R(p,q) + (p'-p)R_p(p,q) + (q'-q)R_q(p,q). \tag{5.16}$$

Then using (5.14), (5.15) and (5.16), F_p in (5.13) can be written as

$$\frac{1}{2}F_p = (\mu + R_p^2)\delta p + R_p R_q \delta q + \mu \delta z + (R - I)R_p + \mu(p - z_x). \tag{5.17}$$

The other terms in (5.13) can be derived similarly and the results are listed as follows:

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial x}F_{p_x} &= -(2\lambda + \lambda_x + 2\beta R_p^2)\delta p - 2\beta R_p R_q \delta q \\ &\quad +\lambda p_{xx} + \lambda_x p_x + \beta R_p(p_{xx}R_p + q_{xx}R_q - I_{xx})\end{aligned} \tag{5.18}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial y}F_{p_y} &= -(2\lambda + \lambda_y + 2\beta R_p^2)\delta p - 2\beta R_p R_q \delta q \\ &\quad +\lambda p_{yy} + \lambda_y p_y + \beta R_p(p_{yy}R_p + q_{yy}R_q - I_{yy})\end{aligned} \tag{5.19}$$

$$\begin{aligned}\frac{1}{2}F_q &= R_p R_q \delta p + (\mu + R_q^2)\delta q + \mu \delta z \\ &\quad +(R - I)R_q + \mu(q - z_y)\end{aligned} \tag{5.20}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial x}F_{q_x} &= -2\beta R_p R_q \delta p - (2\lambda + \lambda_x + 2\beta R_q^2)\delta q \\ &\quad +\lambda q_{xx} + \lambda_x q_x + \beta R_q(p_{xx}R_p + q_{xx}R_q - I_{xx})\end{aligned} \tag{5.21}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial y}F_{q_y} &= -2\beta R_p R_q \delta p - (2\lambda + \lambda_y + 2\beta R_q^2)\delta q \\ &\quad +\lambda q_{yy} + \lambda_y q_y + \beta R_q(p_{yy}R_p + q_{yy}R_q - I_{yy})\end{aligned} \tag{5.22}$$

$$F_z = 0 \tag{5.23}$$

$$\frac{1}{2}\frac{\partial}{\partial x}F_{z_x} = \mu(\delta p - 2\delta z + z_{xx} - p_x) \tag{5.24}$$

$$\frac{1}{2}\frac{\partial}{\partial y}F_{z_y} = \mu(\delta q - 2\delta z + z_{yy} - q_y). \tag{5.25}$$

After substituting (5.17)-(5.25) into (5.13) and writing it in a matrix form, we obtain

$$\begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ 1 & 1 & -4 \end{pmatrix} \begin{pmatrix} \delta p \\ \delta q \\ \delta z \end{pmatrix} = \begin{pmatrix} B_1 + \frac{1}{4}\mu B_3 \\ B_2 + \frac{1}{4}\mu B_3 \\ B_3 \end{pmatrix}, \tag{5.26}$$

where

$$
\begin{aligned}
A_{11} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + R_p^2(1+4\beta) \\
A_{12} &= A_{21} = \frac{1}{4}\mu + R_p R_q(1+4\beta) \\
A_{22} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + R_q^2(1+4\beta) \\
B_1 &= \lambda(p_{xx}+p_{yy}) + \lambda_x p_x + \lambda_y p_y + \mu(z_x - p) \\
&\quad + R_p\{I - R + \beta[(p_{xx}+p_{yy})R_p + (q_{xx}+q_{yy})R_q - I_{xx} - I_{yy}]\} \\
B_2 &= \lambda(q_{xx}+q_{yy}) + \lambda_x q_x + \lambda_y q_y + \mu(z_y - q) \\
&\quad + R_q\{I - R + \beta[(p_{xx}+p_{yy})R_p + (q_{xx}+q_{yy})R_q - I_{xx} - I_{yy}]\} \\
B_3 &= p_x + q_y - z_{xx} - z_{yy}
\end{aligned}
$$

and $(\delta p, \delta q, \delta z)$ are the increments in (p, q, z) after an iteration. Solving (5.26) for $(\delta p, \delta q, \delta z)$, we obtain the updating scheme

$$
\begin{aligned}
\delta p &= \tfrac{1}{\Delta}[A_{22}(B_1 + \tfrac{1}{4}\mu B_3) - A_{12}(B_2 + \tfrac{1}{4}\mu B_3)] \\
\delta q &= \tfrac{1}{\Delta}[A_{11}(B_2 + \tfrac{1}{4}\mu B_3) - A_{12}(B_1 + \tfrac{1}{4}\mu B_3)] \\
\delta z &= \tfrac{1}{4}(\delta p + \delta q - B_3),
\end{aligned}
\tag{5.27}
$$

where

$$
\Delta = A_{11}A_{22} - A_{12}^2. \tag{5.28}
$$

5.3 Hierarchical Structure

The solution to the shape-from-shading problem requires long-range error to be reduced through strictly local interactions. This situation arises mainly from the local nature of the representation used for gradient and height fields, and causes very large number of iterations in the iterative scheme. Multigrid methods are often used in computer vision to speed up computationally intensive tasks such as shape-from-shading. In these methods, an approximate solution computed for a reduced image is used to guide the algorithm toward the complete solution on a larger image.

Multigrid implementation can be simplified by having a 2 : 1 decrease in grid nodes during a transition from a finer grid to an adjacent coarser grid.

Thus, the image resolution is reduced by a factor of 2 between adjacent resolution layers. We use the simplest pyramid method for image sampling and interpolation. The grid size for the lowest layer is 32x32. The input images for various resolution layers are derived from the given highest resolution image by averaging the pixels that belong to the same cell in the low-resolution layer. Let the variables with a tilde ($\tilde{}$) stand for the shape descriptors of the higher resolution layer while the variables without tilde denote those of the lower resolution layer. Then, the communication of the results from one layer to another are specified by the following rules:

Rule 1: The illuminant direction and albedo are the same

$$(\tilde{\sigma}, \tilde{\tau}, \tilde{\eta}) = (\sigma, \tau, \eta).$$

This assumption is made based on experiments reported in [87] that the estimated illuminant direction and surface albedo almost insensitive to changes in resolution. However, the illuminant direction and albedo can also be estimated for each layer independently.

Rule 2: The surface descriptions of a higher resolution layer are interpolated from the descriptions of the adjacent lower resolution layer. Let $\tilde{N}$ be the image size of the higher resolution layer. For $i, j \in \{2, \ldots, \tilde{N}\}$, the shape descriptions for the higher resolution layer are

$$(\tilde{p}, \tilde{q}, \tilde{z})_{i,j} = \begin{cases} (p, q, 2z)_{\frac{i}{2}, \frac{j}{2}}, & \text{if } i \text{ and } j \text{ are evens} \\ \frac{1}{2}[(p, q, 2z)_{\frac{i+1}{2}, \frac{j}{2}} + (p, q, 2z)_{\frac{i-1}{2}, \frac{j}{2}}], & \text{if } i \text{ is odd and } j \text{ is even} \\ \frac{1}{2}[(p, q, 2z)_{\frac{i}{2}, \frac{j+1}{2}} + (p, q, 2z)_{\frac{i}{2}, \frac{j-1}{2}}], & \text{if } i \text{ is even and } j \text{ is odd} \\ \frac{1}{4}[(p, q, 2z)_{\frac{i+1}{2}, \frac{j+1}{2}} + (p, q, 2z)_{\frac{i+1}{2}, \frac{j-1}{2}} & \\ +(p, q, 2z)_{\frac{i-1}{2}, \frac{j+1}{2}} + (p, q, 2z)_{\frac{i-1}{2}, \frac{j-1}{2}}] & \text{if } i \text{ and } j \text{ are odds.} \end{cases}$$

Rule 3: The natural boundary condition is used for the interpolation of boundary pixels. For example, for the boundaries of $i = 1$ and $j = 1$,

$$\begin{array}{lll} (\tilde{p}, \tilde{q})_{1,j} = 2(\tilde{p}, \tilde{q})_{2,j} - (\tilde{p}, \tilde{q})_{3,j} & \tilde{z}_{1,j} = \tilde{z}_{2,j} - \tilde{p}_{1,j}, & \text{for } j \geq 2 \\ (\tilde{p}, \tilde{q})_{i,1} = 2(\tilde{p}, \tilde{q})_{i,2} - (\tilde{p}, \tilde{q})_{i,3} & \tilde{z}_{i,1} = \tilde{z}_{i,2} - \tilde{q}_{i,1}, & \text{for } i \geq 2 \\ (\tilde{p}, \tilde{q})_{1,1} = 2(\tilde{p}, \tilde{q})_{2,2} - (\tilde{p}, \tilde{q})_{3,3} & \tilde{z}_{1,1} = \tilde{z}_{2,2} - \tilde{p}_{1,1} + \tilde{q}_{1,1}. & \end{array}$$

5.4 Iterative Scheme

Combining the hierarchical implementation and our new adaptive approach, we obtain the following shape-from-shading scheme:

1. Estimate the reflectance map parameters (σ, τ, η).

2. Normalize the input image:

$$\bar{I}(x,y) = I(x,y) \,/\, \eta.$$

3. Reduce the input image size to that of the lowest resolution layer.

4. a) Input the initial status p^0, q^0 and z^0.
   b) Initially set the regularization parameters at each pixel to the maximum value, i.e., $\lambda(x,y) = \lambda_{max}$ and set λ_{min}.
   c) Set the values for μ and β.

5. Update the current shape reconstruction. For each pixel, the partial derivatives are approximated by

$$\begin{aligned}
p_x &= p^k_{(x+1,y)} - p^k_{(x,y)} & p_y &= p^k_{(x,y+1)} - p^k_{(x,y)} \\
q_x &= q^k_{(x+1,y)} - q^k_{(x,y)} & q_y &= q^k_{(x,y+1)} - q^k_{(x,y)} \\
z_x &= z^k_{(x+1,y)} - z^k_{(x,y)} & z_y &= z^k_{(x,y+1)} - z^k_{(x,y)} \\
\lambda_x &= \lambda_{(x+1,y)} - \lambda_{(x,y)} & \lambda_y &= \lambda_{(x,y+1)} - \lambda_{(x,y)} \\
\bar{I}_x &= \bar{I}_{(x+1,y)} - \bar{I}_{(x,y)} & \bar{I}_y &= \bar{I}_{(x,y+1)} - \bar{I}_{(x,y)} \\
p_{xx} &= p^k_{(x+1,y)} + p^k_{(x-1,y)} - 2p^k_{(x,y)} & p_{yy} &= p^k_{(x,y+1)} + p^k_{(x,y-1)} - 2p^k_{(x,y)} \\
q_{xx} &= q^k_{(x+1,y)} + q^k_{(x-1,y)} - 2q^k_{(x,y)} & q_{yy} &= q^k_{(x,y+1)} + q^k_{(x,y-1)} - 2q^k_{(x,y)} \\
z_{xx} &= z^k_{(x+1,y)} + z^k_{(x-1,y)} - 2z^k_{(x,y)} & z_{yy} &= z^k_{(x,y+1)} + z^k_{(x,y-1)} - 2z^k_{(x,y)} \\
\bar{I}_{xx} &= \bar{I}_{(x+1,y)} + \bar{I}_{(x-1,y)} - 2\bar{I}_{(x,y)} & \bar{I}_{yy} &= \bar{I}_{(x,y+1)} + \bar{I}_{(x,y-1)} - 2\bar{I}_{(x,y)}
\end{aligned}$$

   where for values outside the image frame, we use natural boundary conditions similar to Rule 3 given in section 5.3. The reconstructed shape is updated by

$$\begin{aligned}
p^{k+1} &= p^k + \delta p \\
q^{k+1} &= q^k + \delta q \\
z^{k+1} &= z^k + \delta z
\end{aligned}$$

and

$$\begin{aligned}
\delta p &= \frac{1}{\Delta}[A_{22}(B_1 + \frac{1}{4}\mu B_3) - A_{12}(B_2 + \frac{1}{4}\mu B_3)] \\
\delta q &= \frac{1}{\Delta}[A_{11}(B_2 + \frac{1}{4}\mu B_3) - A_{12}(B_1 + \frac{1}{4}\mu B_3)] \\
\delta z &= \frac{1}{4}(\delta p + \delta q - B_3)
\end{aligned}$$

where

$$\begin{aligned}
A_{11} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + R_p^2(1 + 4\beta) \\
A_{12} &= \frac{1}{4}\mu + R_p R_q(1 + 4\beta) \\
A_{22} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + R_q^2(1 + 4\beta) \\
B_1 &= \lambda(p_{xx} + p_{yy}) + \lambda_x p_x + \lambda_y p_y + \mu(z_x - p) \\
&\quad + R_p\{I - R + \beta[(p_{xx} + p_{yy})R_p + (q_{xx} + q_{yy})R_q - I_{xx} - I_{yy}]\} \\
B_2 &= \lambda(q_{xx} + q_{yy}) + \lambda_x q_x + \lambda_y q_y + \mu(z_y - q) \\
&\quad + R_q\{I - R + \beta[(p_{xx} + p_{yy})R_p + (q_{xx} + q_{yy})R_q - I_{xx} - I_{yy}]\} \\
B_3 &= p_x + q_y - z_{xx} - z_{yy} \\
R_p &= [R(p^k_{(x,y)} + \delta, q^k_{(x,y)}) - R(p^k_{(x,y)}, q^k_{(x,y)})] \,/\, \delta \\
R_q &= [R(p^k_{(x,y)}, q^k_{(x,y)} + \delta) - R(p^k_{(x,y)}, q^k_{(x,y)})] \,/\, \delta \\
\Delta &= A_{11}A_{22} - A_{12}^2.
\end{aligned}$$

6. Go to the next step if solution is stable or iterations has reached N_{max} of current layer; otherwise repeat Step 5.

7. Compute the control signal $c(x, y) = abs(I(x, y) - R(p^k_{(x,y)}, q^k_{(x,y)}))$.

8. Update $\lambda(x, y)$ according to the following rule:

$$\lambda_{new}(x, y) = \begin{cases} \mathcal{F}(x, y, \lambda_{old}) & \text{if } c(x, y) > 0 \text{ and } \lambda_{old} > \lambda_{min} \\ \lambda_{old}(x, y) & otherwise \end{cases} \tag{5.29}$$

where

$$\mathcal{F}(x, y, \lambda_{old}) = (1 - e^{-\frac{c(x,y)}{V_T}})\,\lambda_{min} + (e^{-\frac{c(x,y)}{V_T}})\,\lambda_{old}(x, y).$$

9. If no change in $\lambda(x, y)$ values continue to Step 10; otherwise go back to Step 5 and repeat the process with new $\lambda(x, y)$ values.

10. If current image is in the highest resolution stop; otherwise
    a) Increase the image size and expand the shape reconstruction to the adjacent higher resolution layer,
    b) Expand $\lambda(x, y)$ using the rules given in Section 5.3,
    c) Reduce the normalized input image to the current resolution,
    d) Go to Step 5.

5.5 Relationship to Existing Techniques

In our adaptive shape-from-shading implementation, we combined all SFS constraints in a single energy functional. Thus, we have a general purpose algorithm which encompasses most of the existing methods. These methods are special instances of our generalized method and can be obtained by appropriately selecting parameter values.

- In our method, we made λ as a function of spatial coordinates, i.e., $\lambda(x, y)$, in order to control the degree of smoothness of the reconstructed surface. If we set a constant value to λ and drop the intensity gradient constraint in the integrand ($\beta = 0$), we obtain the scheme developed by Horn [12].

- If we make λ constant and drop both the integrability and the intensity gradient constraints ($\mu = 0$ and $\beta = 0$), we derive the algorithm presented by Ikeuchi and Horn [11].

- If we instead, remove both the departure from smoothness and the intensity gradient terms ($\lambda = 0$ and $\beta = 0$), we obtained something reminiscent of the iterative scheme of Strat [1], although Strat dealt with the integrability in a different way.

- If we drop the brightness error term and the intensity gradient constraint ($\beta = 0$), we obtain the scheme of Harris [88, 89] for interpolating from depth and slope. This method solves the biharmonic equation for z, by solving a coupled set of second-order partial differential equations.

- We find Zheng and Chellappa's scheme [5] if we drop the departure from smoothness term ($\lambda = 0$) in our method.

5.6 Integrating Adaptive SFS and Photometric Stereo

The reflectance map is used in computer graphics where an image is created from description of the shape of an object. Conversely, given an image, we would like to be able to recover the shape. There is a unique mapping from surface orientation, specified by p and q, to radiance, given by the reflectance map $R(p,q)$. However, the inverse mapping is not unique. An infinite number of surface orientations give rise to the same brightness. In general, the mapping from brightness to surface orientation cannot be unique, since brightness has one degree of freedom, while orientation has two.

To recover surface orientation locally, we must introduce additional information. To determine two unknowns, p and q, we need two equations. Two images, taken with different lighting, will yield two equations for each image point as

$$I = R(p,q) \qquad \text{and} \qquad \hat{I} = \hat{R}(p,q). \tag{5.30}$$

We also applied our adaptive regularization approach to the photometric stereo technique in order to improve the quality of the reconstructed surface obtained by the iterative scheme. Suppose we have two input images, $I(x,y)$ and $\hat{I}(x,y)$, illuminated by two light sources placed different locations in the image space as (σ,τ) and $(\hat{\sigma},\hat{\tau})$ respectively. Let the reflectance maps corresponding to these illuminant directions be

$$R(p,q) = \eta \frac{\cos\sigma - p\sin\sigma\cos\tau - q\sin\sigma\sin\tau}{\sqrt{1+p^2+q^2}} \tag{5.31}$$

$$\hat{R}(p,q) = \hat{\eta} \frac{\cos\hat{\sigma} - p\sin\hat{\sigma}\cos\hat{\tau} - q\sin\hat{\sigma}\sin\hat{\tau}}{\sqrt{1+p^2+q^2}}. \tag{5.32}$$

Then, for the photometric stereo application, we construct

$$F(p,q,z) = F_1'(p,q) + F_2'(p,q) + \mu\, F_3(p,q,z) + \beta\, F_4(p,q), \tag{5.33}$$

where

$$F_1' = \iint \{(I(x,y) - R(p,q))^2 + (\hat{I}(x,y) - \hat{R}(p,q))^2\}\, dx\, dy \tag{5.34}$$

$$F_2' = \iint \lambda(x,y)\,(p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy \tag{5.35}$$

$$F_3 = \iint ((z_x - p)^2 + (z_y - q)^2)\, dx\, dy \tag{5.36}$$

$$F_4 = \iint ((R_x - I_x)^2 + (R_y - I_y)^2)\, dx\, dy. \tag{5.37}$$

Using calculus of variations, minimizing $F(p,q,z)$ in (5.33) is equivalent to solving the Euler equations given in (5.13). Approximating the reflectance map in (5.33) around (p,q) by Taylor series expansion of up to first-order terms and replacing in (5.13), we obtain

$$\begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ 1 & 1 & -4 \end{pmatrix} \begin{pmatrix} \delta p \\ \delta q \\ \delta z \end{pmatrix} = \begin{pmatrix} B_1 + \frac{1}{4}\mu B_3 \\ B_2 + \frac{1}{4}\mu B_3 \\ B_3 \end{pmatrix}, \tag{5.38}$$

where

$$\begin{aligned}
A_{11} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + (R_p^2 + \hat{R}_p^2)(1 + 4\beta) \\
A_{12} &= A_{21} = \frac{1}{4}\mu + (R_p R_q + \hat{R}_p \hat{R}_q)(1 + 4\beta) \\
A_{22} &= 4\lambda + \lambda_x + \lambda_y + \frac{5}{4}\mu + (R_q^2 + \hat{R}_q^2)(1 + 4\beta) \\
B_1 &= \lambda(p_{xx} + p_{yy}) + \lambda_x p_x + \lambda_y p_y + \mu(z_x - p) \\
&\quad + R_p\{(I - R) + \beta[(p_{xx} + p_{yy})R_p + (q_{xx} + q_{yy})R_q - I_{xx} - I_{yy}]\} \\
&\quad + \hat{R}_p\{(\hat{I} - \hat{R}) + \beta[(p_{xx} + p_{yy})\hat{R}_p + (q_{xx} + q_{yy})\hat{R}_q - \hat{I}_{xx} - \hat{I}_{yy}]\} \\
B_2 &= \lambda(q_{xx} + q_{yy}) + \lambda_x q_x + \lambda_y q_y + \mu(z_y - q) \\
&\quad + R_q\{(I - R) + \beta[(p_{xx} + p_{yy})R_p + (q_{xx} + q_{yy})R_q - I_{xx} - I_{yy}]\} \\
&\quad + \hat{R}_q\{(\hat{I} - \hat{R}) + \beta[(p_{xx} + p_{yy})\hat{R}_p + (q_{xx} + q_{yy})\hat{R}_q - \hat{I}_{xx} - \hat{I}_{yy}]\} \\
B_3 &= p_x + q_y - z_{xx} - z_{yy}.
\end{aligned}$$

Solving (5.38) for δp, δq and δz, we obtain the updating scheme

$$\delta p = \frac{1}{\Delta}[A_{22}(B_1 + \frac{1}{4}\mu B_3) - A_{12}(B_2 + \frac{1}{4}\mu B_3)]$$

$$\begin{aligned} \delta q &= \frac{1}{\Delta}[A_{11}(B_2 + \frac{1}{4}\mu B_3) - A_{12}(B_1 + \frac{1}{4}\mu B_3)] \\ \delta z &= \frac{1}{4}(\delta p + \delta q - B_3), \end{aligned} \tag{5.39}$$

where

$$\Delta = A_{11}A_{22} - A_{12}^2. \tag{5.40}$$

5.7 Experimental Results and Discussions

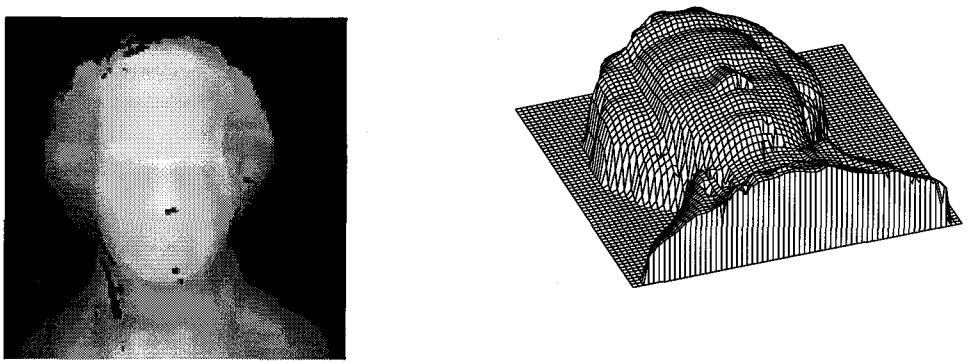
We tested our adaptive SFS algorithm on synthetic and real images. In all experiments, we set $\lambda = 1$, $\mu = 1$, $\beta = 1$, $\lambda_{min} = 0.01$ and N_{max} for the highest resolution layer be 500. The iterations start from the zero gradient field and the zero height map, i.e., $(p, q, z) = (0, 0, 0)$. All the images presented in our examples are 128x128. The 3-D plots of the height maps are given in three resolution layers as 32x32, 64x64 and 128x128. The Mozart test image is generated from a laser range finder while the Midilli image is simulated using a digital terrain model (DTM) of the Midilli Island in Aegaen Sea.

Figure 5.4(a) shows the true height map of the Mozart image obtained by a laser range finder used to generate the synthetic input image to the algorithm and Figure 5.4(b) is a 3-D plot of the true height map.

Figure 5.5 shows the synthesized Mozart images with three resolution layers applied to the multigrid adaptive SFS algorithm. These images are generated from the true height map shown in Figure 5.4(a) using the Lambertian reflectance map with parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$.

If we convert our adaptive SFS algorithm to the non-adaptive algorithm by setting a constant value to the regularization parameter, i.e., $\lambda(x, y) = \lambda = 1$, we obtain the images with three resolution layers shown in Figure 5.6. These images are synthesized from the (p, q) maps using the reflectance map parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$.

Figure 5.7 shows the images with different resolutions synthesized from the (p, q) maps obtained by the adaptive SFS algorithm using parameters $\tau = 45°$,



(a) (b)

Figure 5.4. Mozart test image: (a) The true height map. (b) A 3-D plot of the true height map.

$\sigma = 45°$ and $\eta = 250$. The similarity between the input images in Figure 5.5 and the reconstructed images in Figure 5.7 illustrates that the irradiance requirement in the energy functional is satisfied by the SFS solution.

If we compare the results in Figure 5.6 obtained from the non-adaptive SFS algorithm and Figure 5.7 obtained from the adaptive SFS algorithm, we can see that non-adaptive SFS algorithm gives smoother surfaces because of setting a constant value to λ which flattens the reconstructed height maps. On the other hand, the adaptive SFS scheme can automatically determine the value of λ at every image point.

The 3-D plots of the reconstructed surfaces obtained by the non-adaptive and the adaptive algorithms are shown in Figure 5.8 and Figure 5.9 respectively. The smoothing effect of using a constant λ can be easily seen if we compare these plots. It can also be seen that the adaptive control of the regularization parameter spatially improves the solution around discontinuities.

Figure 5.10(a) and (b) show the p and q maps of the ground truth, and (c) and (d) show the p and q maps obtained by the adaptive SFS algorithm.

A comparison of the reconstructed (p, q, z) with the ground truth for the Mozart image shows that there are errors in the background region which is

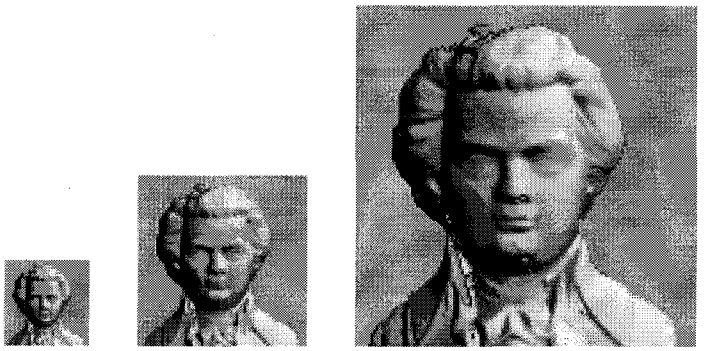


Figure 5.5. The input images synthesized from the true height map with parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$ with sizes 32x32, 64x64 and 128x128 from left to right.

Figure 5.6. The images synthesized from the (p, q) maps obtained by the non-adaptive SFS algorithm using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$ with sizes 32x32, 64x64 and 128x128 from left to right.

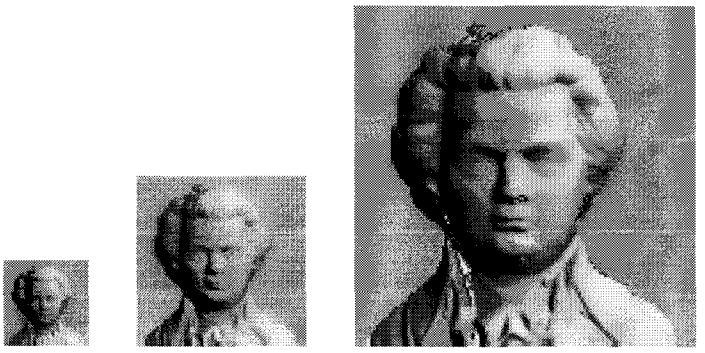
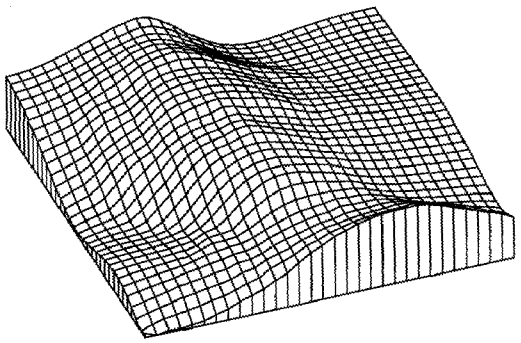
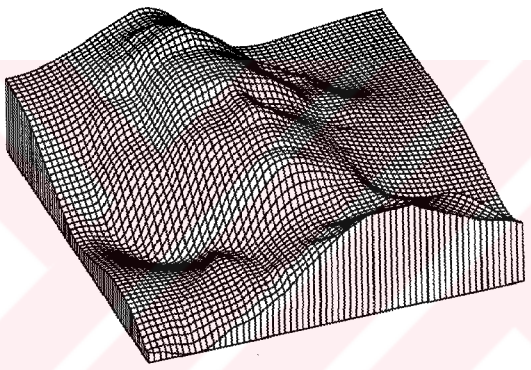


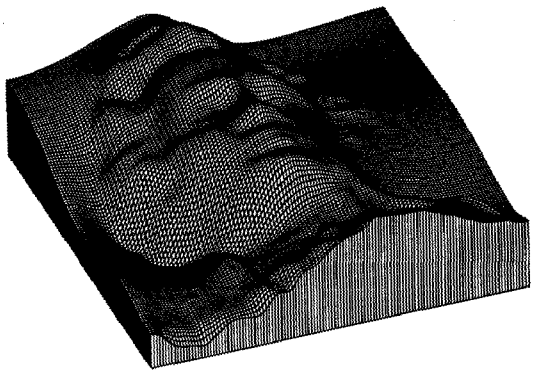
Figure 5.7. The images synthesized from the (p, q) maps obtained by the adaptive SFS algorithm using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$ with sizes 32x32, 64x64 and 128x128 from left to right.



(a)

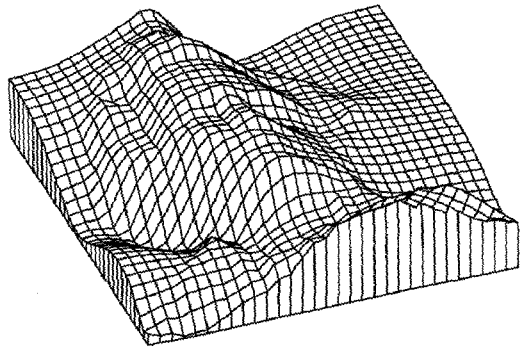


(b)

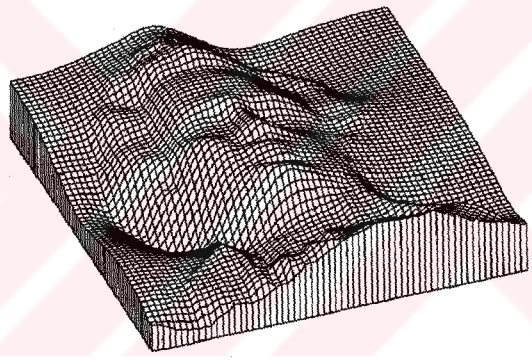


(c)

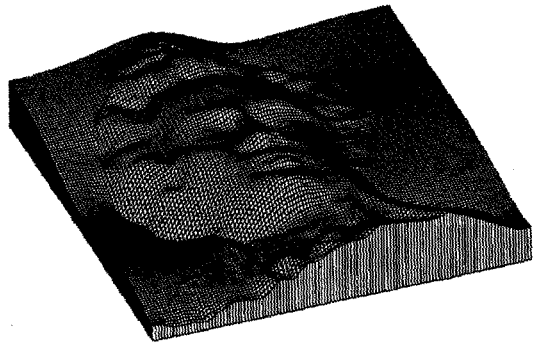
Figure 5.8. 3-D plots of the reconstructed height maps for the Mozart image obtained by the non-adaptive SFS algorithm with sizes (a) 32x32, (b) 64x64 and (c) 128x128.



(a)

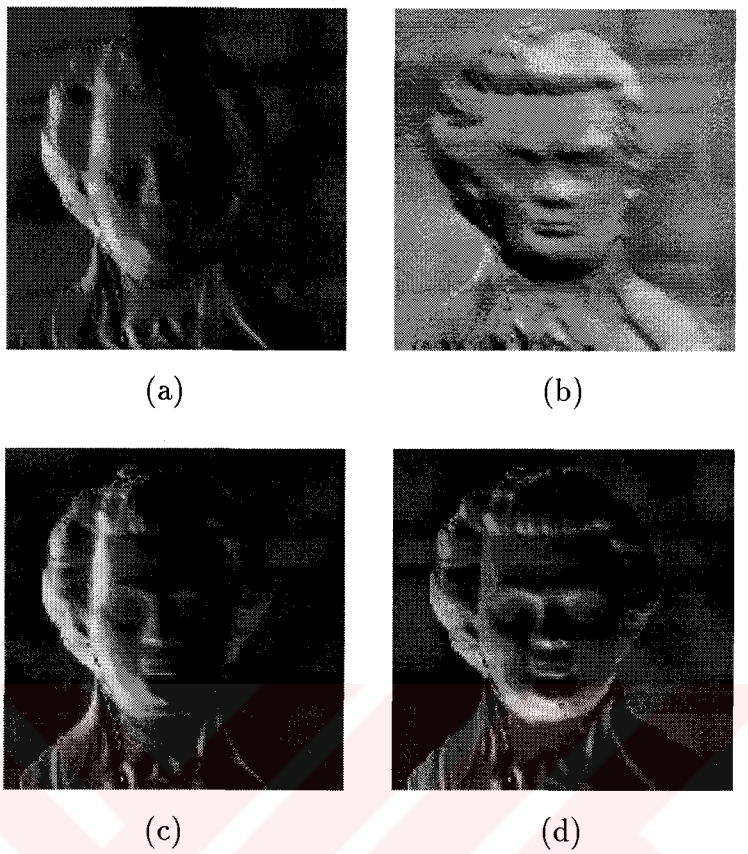


(b)



(c)

Figure 5.9. 3-D plots of the reconstructed height maps for the Mozart image obtained by the adaptive SFS algorithm with sizes (a) 32x32, (b) 64x64 and (c) 128x128.



(a) (b)

(c) (d)

Figure 5.10. Partial derivatives: (a) The true z_x map. (b) The true z_y map. (c) The p map reconstructed by the adaptive SFS. (d) The q map reconstructed by the adaptive SFS.

probably due to a lack of proper boundary conditions. Apart from this, the reconstruction is successful and the shapes of hair, forehead, eyebrows, nose, cheeks, etc. are correct. This conclusion is supported by the similarities between the images generated from the ground truth and the SFS results using different illuminant directions.

Figure 5.11(a) and (d) compare the images synthesized from the ground truth and the adaptive SFS result, both using the parameters $\tau = 135°$, $\sigma = 45°$ and $\eta = 250$. In this case, the illuminant direction is orthogonal to the illuminant in the input image. Figure 5.11(b) and (e) give an image pair, the images synthesized from the ground truth and the SFS result using parameters $\tau = 225°$, $\sigma = 45°$ and $\eta = 250$. In this case, the illuminant direction is opposite to the illuminant in the input image. Figure 5.11(c) and (f) give another image pair, the images synthesized from the ground truth and the SFS result, both

using the parameters $\tau = 315°$, $\sigma = 45°$ and $\eta = 250$.

Figure 5.12 shows two input images generated with the parameters (a) $\tau = 45°$, $\sigma = 45°$, $\eta = 250$ and (b) $\tau = 135°$, $\sigma = 45°$, $\eta = 250$, and used in the integrated adaptive photometric stereo algorithm as input images. Figure 5.13 shows the images synthesized from the (p, q) maps obtained by integrating the photometric stereo and adaptive SFS algorithm. In Figure 5.14, a 3-D plot of the reconstructed height map from this integrated approach is shown. By comparing the results in Figure 5.9 and Figure 5.14, one can observe that this integration improves the reconstructed surface by providing more symmetrical results with the help of information gathered from two images.

Figure 5.15 shows the Midilli image which is simulated using a digital terrain model (DTM) of the Midilli island. Figure 5.15(a) shows a 3-D plot of the true DTM and Figure 5.15(b) is the true height map used to generate the input image. Figure 5.15(c) shows the input image to the adaptive SFS algorithm which is synthesized from Figure 5.15(b) with using the reflectance map parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$.

In Figure 5.16(a), the height map recovered by the adaptive SFS algorithm is shown. Figure 5.16(b) shows the Midilli image synthesized from the (p, q) maps obtained by the SFS algorithm using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$. Figure 5.17 shows the 3-D plots of the reconstructed images with three resolution layers. Comparison of the 3-D plot in Figure 5.15(a) with those in 5.16 show that the general shape of the surface has been reconstructed successfully.

Figure 5.18(a) and (b) show the p and q maps of the true DTM, and (c) and (d) show the p and q maps obtained by the SFS algorithm. Figure 5.19(a)-(f) compare images synthesized from the true DTM and the SFS results both using the parameters $\sigma = 45°$ and $\eta = 250$. In these figures, three different tilt angles are used as 135°, 225° and 315°. It can be seen from these images that, using different illuminant directions are consistent.

Figure 5.20 shows the SFS results for the Lenna image. Figure 5.20(a) shows the input Lenna image. The reflectance map parameters are estimated

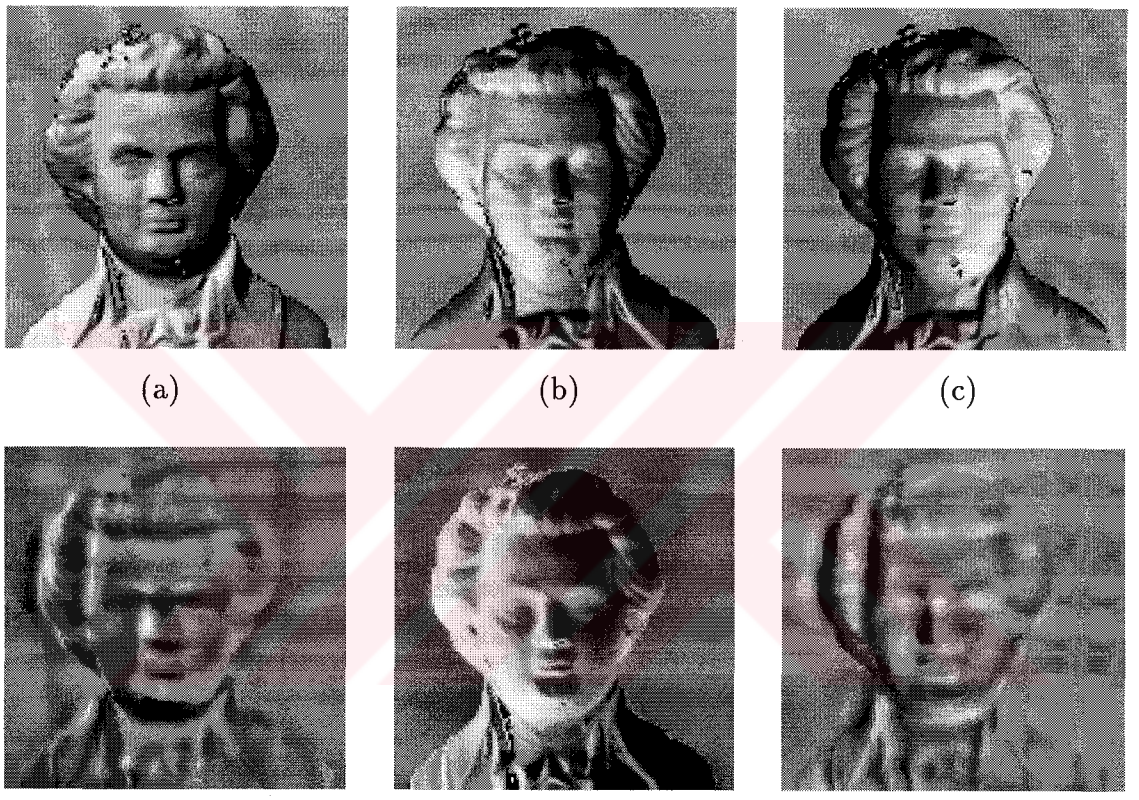


Figure 5.11. The images synthesized from the true height with parameters $\sigma = 45°$, $\eta = 250$, and (a) $\tau = 135°$, (b) $\tau = 225°$ and (c) $\tau = 315°$. The images synthesized from the (p, q) maps obtained by the adaptive SFS algorithm using parameters $\sigma = 45°$, $\eta = 250$, and (d) $\tau = 135°$, (e) $\tau = 225°$ and (f) $\tau = 315°$.

Figure 5.12. The input images for the adaptive photometric stereo algorithm, which are synthesized from the true height map using the parameters (a) $\tau = 45°$, $\sigma = 45°$, $\eta = 250$ and (b) $\tau = 135°$, $\sigma = 45°$, $\eta = 250$.

Figure 5.13. The images synthesized from the (p, q) maps obtained by the adaptive photometric stereo algorithm using the parameters (a) $\tau = 45°$, $\sigma = 45°$, $\eta = 250$ and (b) $\tau = 135°$, $\sigma = 45°$, $\eta = 250$.

by using Zheng and Chellappa's source from shading method as $\tau = 9.9°$, $\sigma = 57.8°$ and $\eta = 191.7$. Figure 5.20(b) is the height map obtained by the adaptive SFS algorithm. Figure 5.20(c) and (d) are the (p, q) maps of the SFS results. Figure 5.20(e) shows the image synthesized from the reconstructed (p, q) maps using the estimated reflectance map parameters. Figure 5.20(f)-(h) show images synthesized from the reconstructed (p, q) maps using parameters $\sigma = 57.8°$ and $\eta = 191.7$, and τ equals to 99.9°, 189.9° and 279.9° respectively, corresponding to whether illuminant from directions opposite or orthogonal to the estimated direction for the input image. It can be seen that the images synthesized from the SFS result using different illuminant directions are consistent.

In Figure 5.21, the 3-D plots of the reconstructed surfaces are shown with

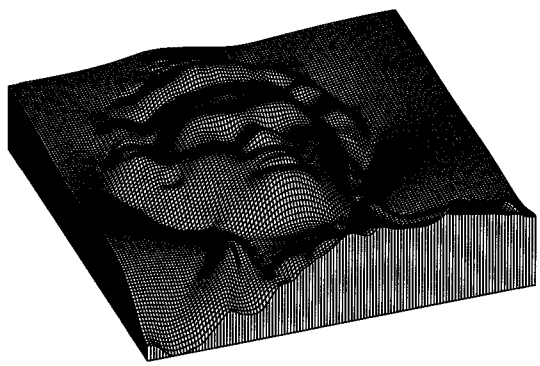
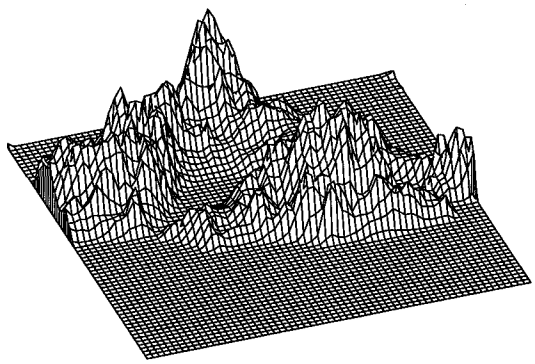


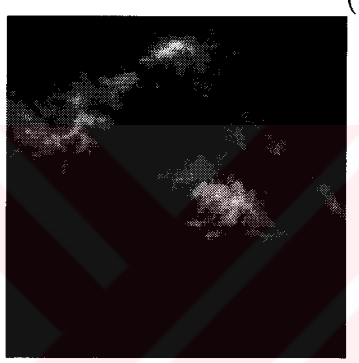
Figure 5.14. A 3-D plot of the reconstructed height map from the adaptive photometric stereo algorithm.

three resolution layers. It can be seen from these plots that the shapes of the face and shoulder are recovered correctly and that features such as the nose, lips, cheeks, chin, etc. are easily identified.

Figure 5.22 shows the SFS results for the pepper image. Figure 5.22(a) is the input image. The estimated reflectance map parameters by using Zheng and Chellappa's method are $\tau = 33.9°$, $\sigma = 45.2°$ and $\eta = 221.7$. Figure 5.22(b) is the z map obtained by our adaptive SFS algorithm. The p and q maps of the SFS result are shown in Figure 5.22(c) and 5.22(d) respectively. Figure 5.22(e) shows the image synthesized from the reconstructed (p, q) maps using the estimated reflectance map parameters. Figure 5.22(f)-(h) show images synthesized from the reconstructed (p, q) maps using parameters $\sigma = 45.2°$ and $\eta = 250$, and τ equals to 123.9°, 213.9° and 303.9° respectively, corresponding to whether illuminant from directions opposite or orthogonal to the estimated direction for the input image. It can be seen that the images synthesized from the SFS result using different illuminant directions are consistent. The 3-D plots of the reconstructed surfaces are shown with three resolutions in Figure 5.23. It can be seen from these plots that the basic shape for the pepper are recovered with enough details.



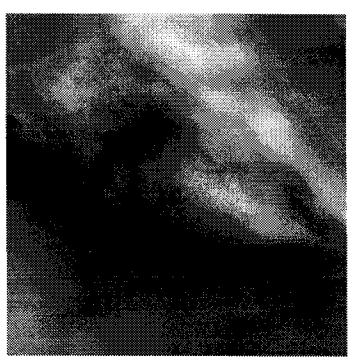
(a)



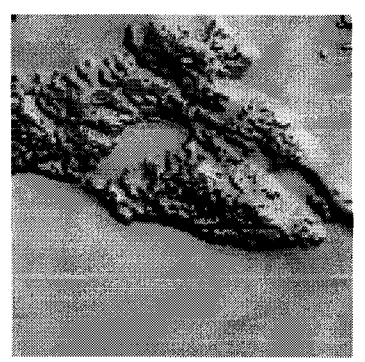
(b)

(c)

Figure 5.15. Midilli DTM image: (a) A 3-D plot of the true height map. (b) The true height map. (c) The input image which is synthesized from (b) with parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$.

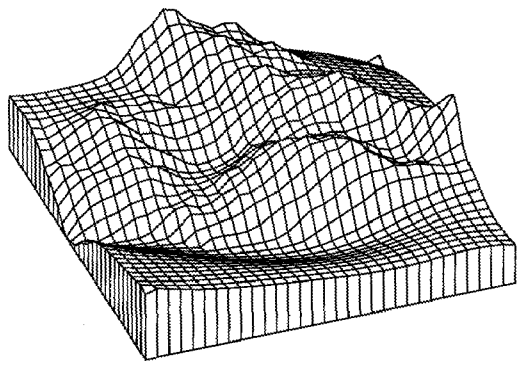


(a)

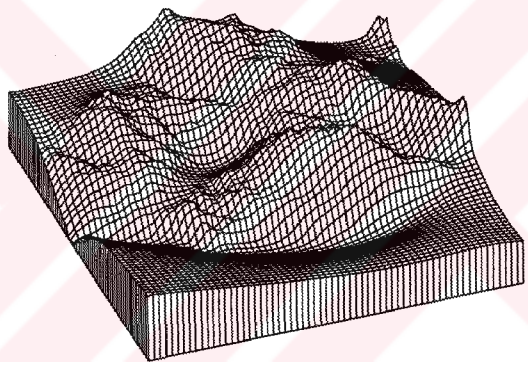


(b)

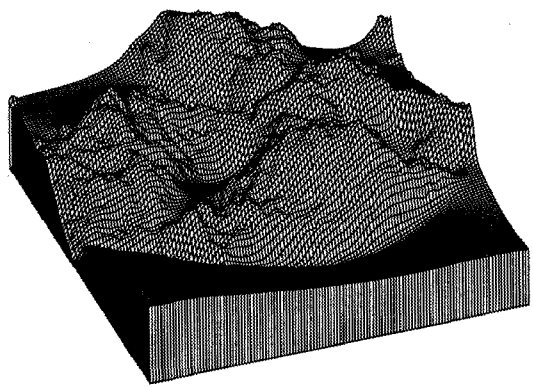
Figure 5.16. (a) The height map reconstructed by the adaptive SFS algorithm. (b) The image generated from the SFS result using the parameters $\tau = 45°$, $\sigma = 45°$ and $\eta = 250$.



(a)

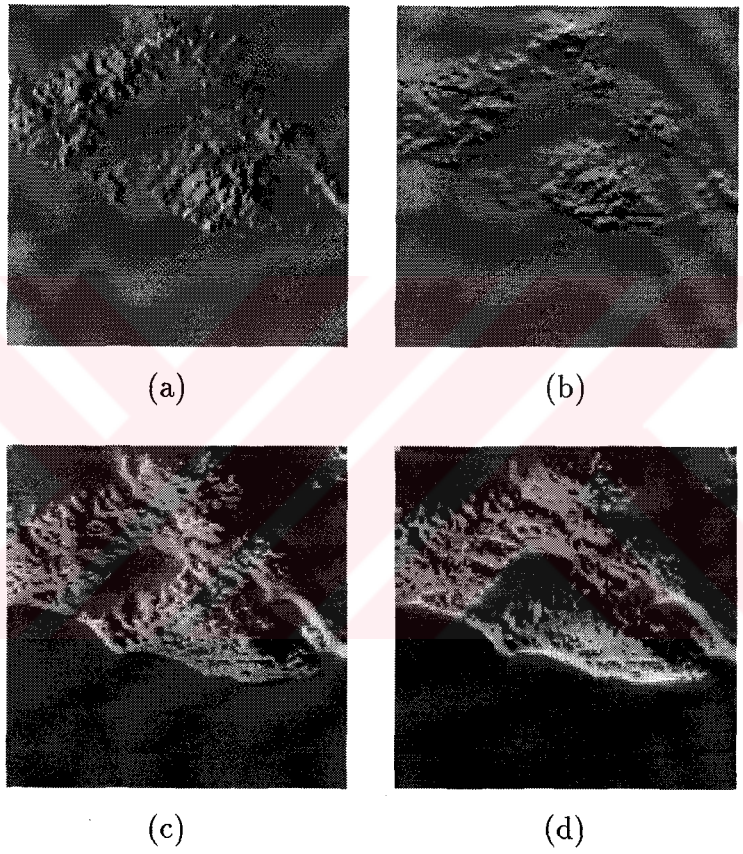


(b)



(c)

Figure 5.17. 3-D plots of the reconstructed height maps for the Midilli image obtained by the adaptive SFS algorithm with sizes (a) 32x32, (b) 64x64 and (c) 128x128.



(a) (b)

(c) (d)

Figure 5.18. Partial derivatives: (a) The true z_x map. (b) The true z_y map. (c) The SFS reconstructed p map. (b) The SFS reconstructed q map.

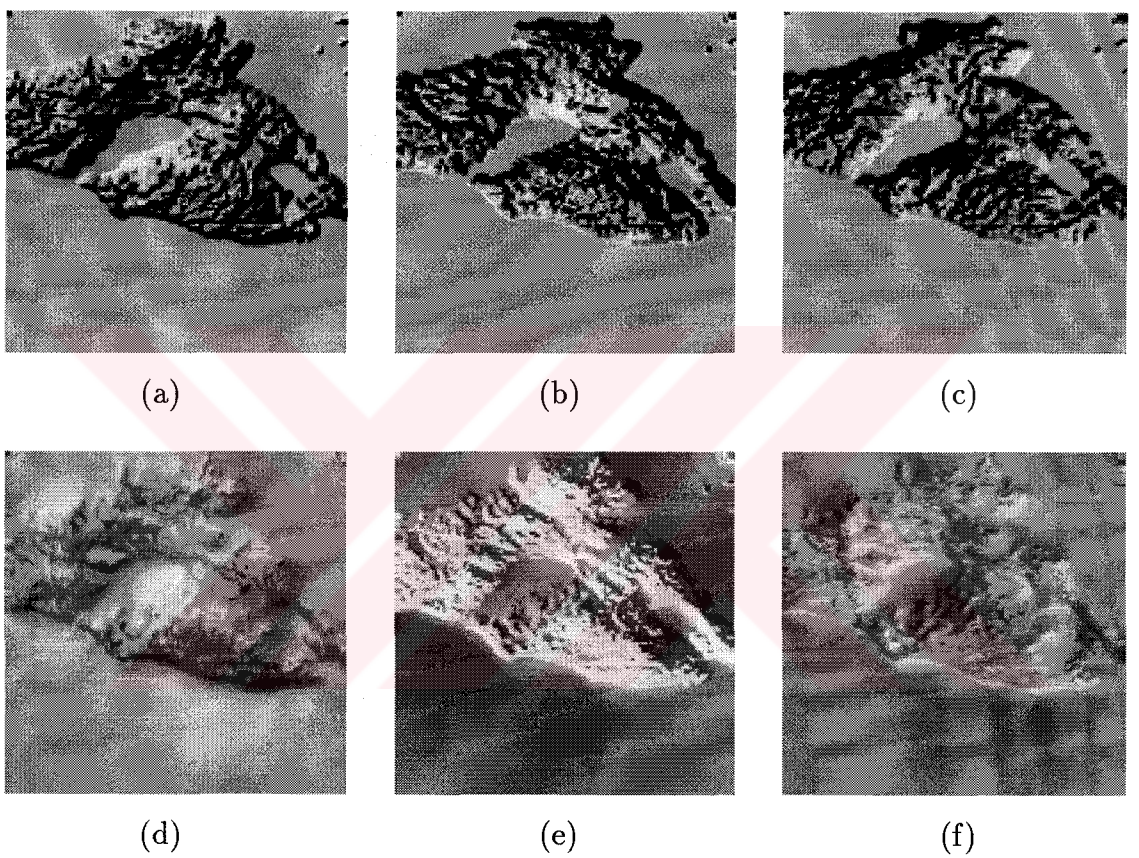


Figure 5.19. The images synthesized from the true height using parameters $\sigma = 45°$, $\eta = 250$, and (a) $\tau = 135°$, (b) $\tau = 225°$ and (c) $\tau = 315°$. The images synthesized from the (p, q) maps obtained by the adaptive SFS algorithm using parameters $\sigma = 45°$, $\eta = 250$, and (d) $\tau = 135°$, (e) $\tau = 225°$ and (f) $\tau = 315°$.

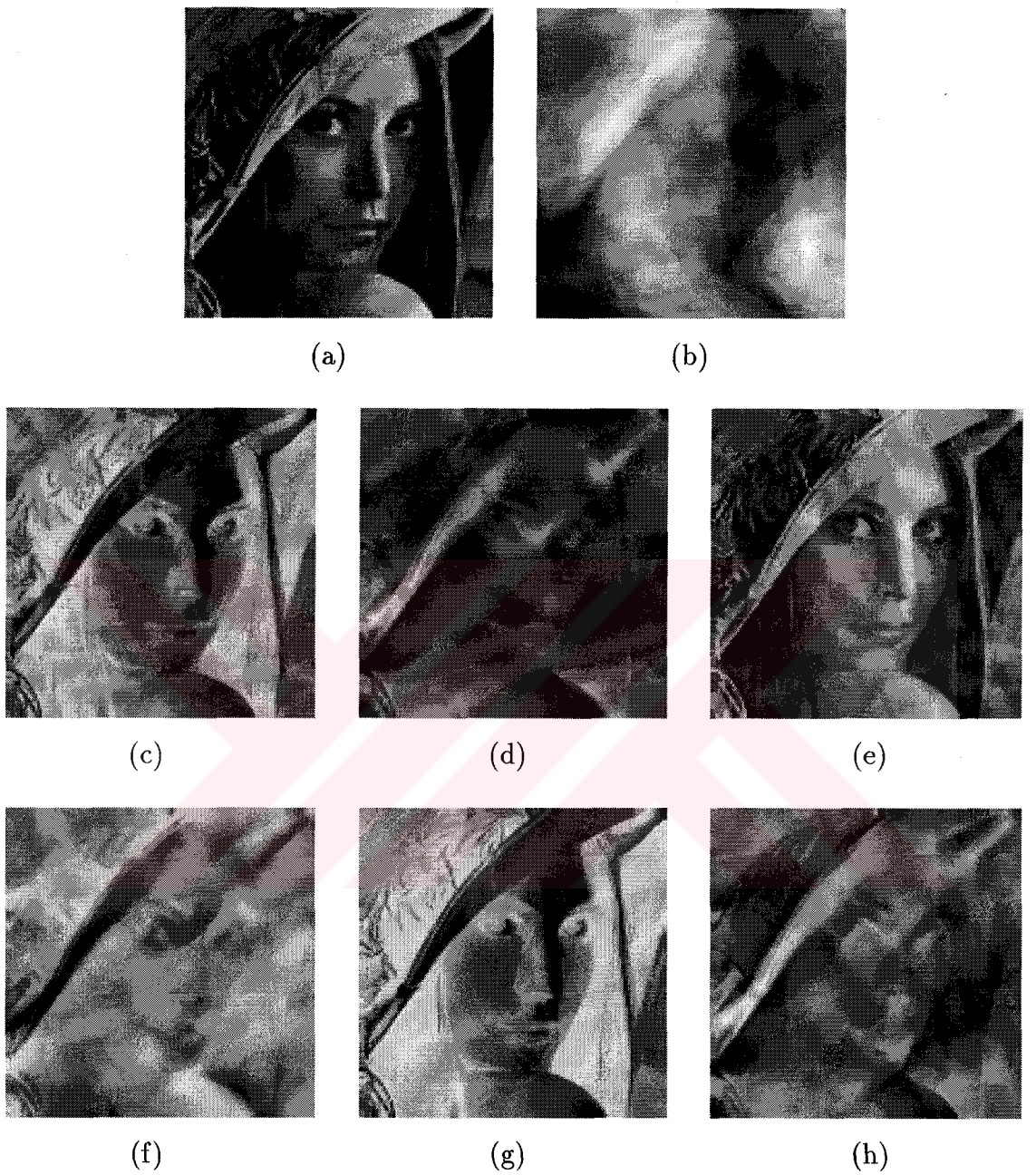
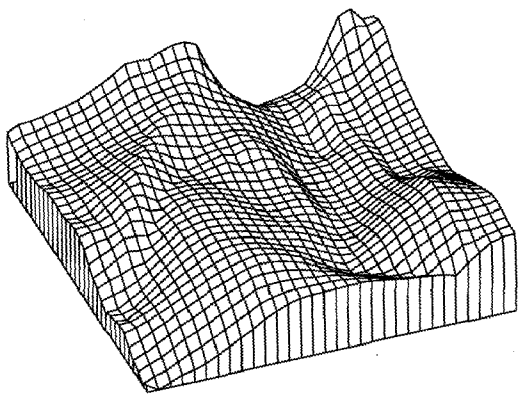
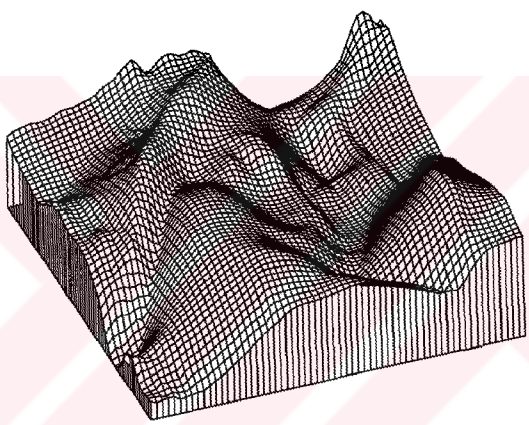


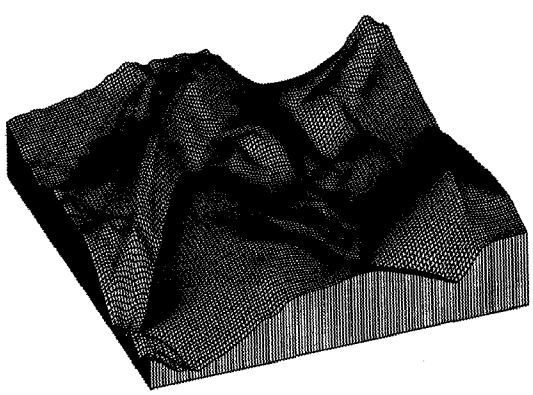
Figure 5.20. SFS results for Lenna image: (a) The input image. (b) The height map obtained by the adaptive SFS algorithm. (c) and (d) are the p and q maps of the SFS results. (e) The image synthesized from the reconstructed (p, q) maps using the estimated reflectance map parameters. (f), (g) and (h) are images synthesized from the reconstructed (p, q) maps using τ equals to 99.9°, 189.9° and 279.9° respectively, corresponding to illumination from directions opposite or orthogonal to the estimated directions for the input image.



(a)



(b)



(c)

Figure 5.21. 3-D plots of the reconstructed height maps for the Lenna image obtained by the adaptive SFS algorithm with sizes (a) 32x32, (b) 64x64 and (c) 128x128.

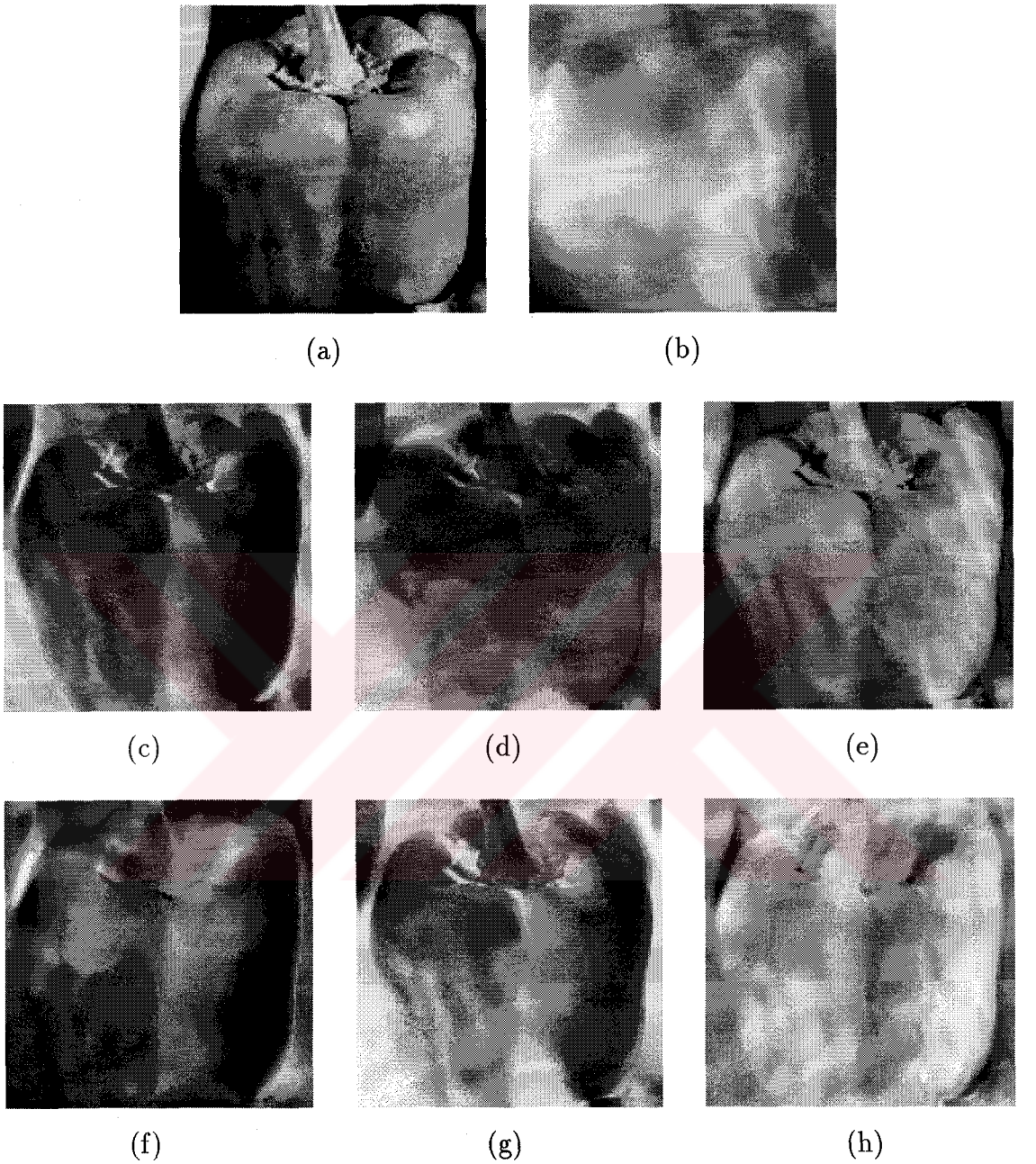
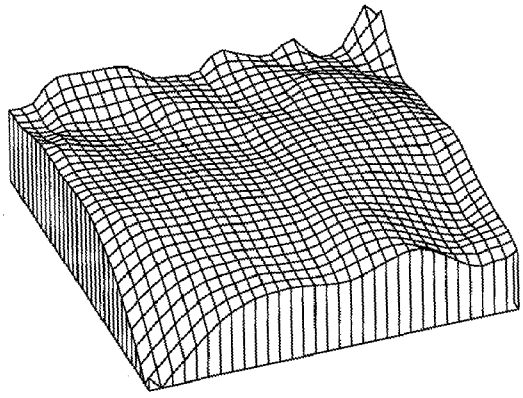
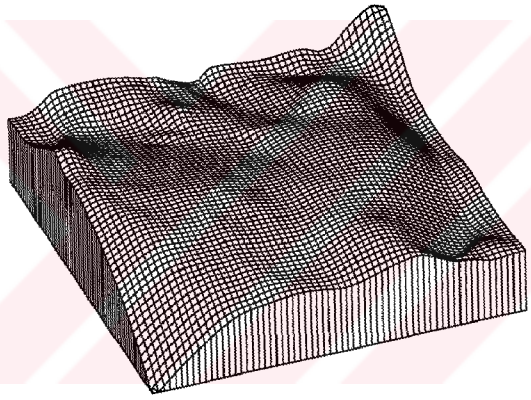


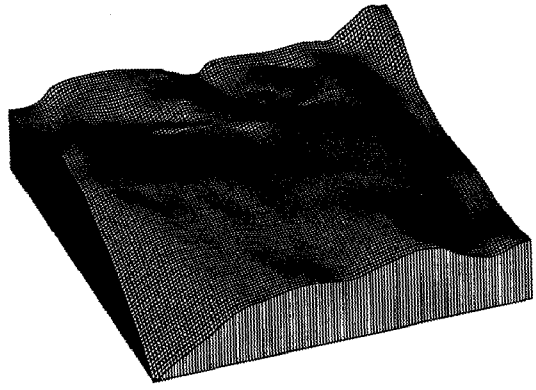
Figure 5.22. SFS results for the pepper image: (a) The input image. (b) The height map obtained by the adaptive SFS algorithm. (c) and (d) are the p and q maps of the SFS results. (e) The image synthesized from the reconstructed (p, q) maps using the estimated reflectance map parameters. (f), (g) and (h) are images synthesized from the reconstructed (p, q) maps using τ equals to 123.9°, 213.9° and 303.9° respectively, corresponding to illumination from directions opposite or orthogonal to the estimated directions for the input image.



(a)



(b)



(c)

Figure 5.23. 3-D plots of the reconstructed height maps for the pepper image obtained by the adaptive SFS algorithm with sizes (a) 32x32, (b) 64x64 and (c) 128x128.

CHAPTER 6.

CONCLUSIONS

In this final chapter, we summarize the results and contributions of this thesis and suggest problems for future research.

6.1 Summary of Results

The shape-from-shading problem is studied in this thesis. The review of the algorithms reveals that the major difficulties in this problem are determining the boundary conditions of the minimization problem, selecting the algorithm parameter values, slow convergence of the global iterative SFS algorithm and having poor information about the scene by using a single measurement of the image brightness.

Chapter 2 represents a background knowledge related to the reflectance models and source from shading technique. Depending on their physical properties, surfaces can be categorized as Lambertian, specular, hybrid and more sophisticated surfaces. In this chapter, we describe the reflectance models and discuss their properties related to shape-from-shading. Source-from-shading technique which deals with the recovery of the light source location and albedo from a shaded image is presented. We considered three source-from-shading methods. We explained Pentland's original approach for light source computation, its refinements by Lee and Rosenfeld and improvements of the both methods by Zheng and Chellappa. The resulting formulas of these methods for illuminant direction and albedo estimation are summarized in the table given at the end of this chapter.

In chapter 3, the shape-from-shading problem is discussed in detail. At the beginning of this chapter, we present the regularization theory which enables us to transform the ill-posed problems in the early vision to the well-posed problems by imposing the physical constraints on the solution in the variational form. After that, we review the shape-from-shading formulation. We present a number of ways for specifying surface orientation. They are unit surface normal ($\vec{N}$), surface gradient (p, q), slant and tilt angles (σ, τ) and stereographic coordinates (f, g). Then, the reflectance map is introduced and the image irradiance equation which is fundamental for recovering surface shape is described. The variational solution of the image irradiance equation is discussed and a procedure for deriving a SFS iterative scheme is presented. After the introduction of the constraints used in the shape-from-shading problem, a total of eight SFS methods are summarized. At the end of this chapter, another shape recovery technique used in computer vision, known as shape-from-photometric-stereo, is introduced. Photometric stereo deals with the recovery of 3-D shape using multiple input images of the same scene generated fixed viewing direction and different light source directions.

Chapter 4 discusses the effects of the SFS constraints, which are the brightness constraint, the smoothness constraint, the integrability constraint and the intensity gradient constraint, on the solution. Here, an energy functional which contains all the constraints are constructed and minimization of this constrained minimization problem is considered. An iterative SFS scheme is derived by using the calculus of variations and a linear approximation of the reflectance map. Changing the relative weight of each constraint in the energy functional by setting the proper values to the algorithm parameters λ, μ and β, we explored the individual effects of the SFS constraints on the solution and observed the following results:

The brightness constraint measures the error in brightness , that is the difference between the observed image brightness and the predicted from the computed shape. This constraint which is related directly to the image irradiance equation allows for modeling errors and noise, and used as a part of an energy functional.

The regularization term also called smoothness constraint is used to sta-

bilize the iterative SFS scheme and to ensure that it has a unique minimum. Although the smoothness term is used to ensure the convergence of the iterative process, it flattens the SFS reconstruction, causing distortions along the image discontinuities. On the other hand, determining the optimum value of the regularization term is critical as it should be large enough to avoid instability and small enough to produce a meaningful solution. As a result, using of the smoothness constraint in the shape-from-shading problem produces solutions that are too smooth with the amount of distortion depending on the choice of regularization parameter value.

The integrability constraint is used to ensure that the gradient field computed by the SFS algorithm corresponds to a physical surface. Experimental results show that the choice of the value of the integrability constraint parameter μ is also critical and too small values may cause instabilities in the iterative scheme. However, big values for μ cause smooth surfaces and this effect can be seen on the shaded images which are synthesized from the (p, q) maps obtained by the SFS algorithm.

The intensity gradient constraint is used to require that the gradients of the reconstructed image to be close to the gradients of the input image in both x and y directions. To see how this constraint effects the solution, we calculated the partial derivatives of the input image (I_x, I_y) and the partial derivatives of the SFS generated images (R_x, R_y) and compared them. Although these partial derivatives are not similar when the intensity gradient constraint is not present in the iterative scheme, they look quite similar when we add this constraint to the algorithm. Changing the value of the intensity gradient parameter β does not cause significant changes on the recovered height map but too small values of β causes smoother shaded images.

In chapter 5, a new adaptive shape-from-shading approach which uses adaptive smoothing function to reduce the oversmoothing along the reconstructed image is introduced. In order to control the smoothness over the image space, the fixed regularization parameter of the standard regularization is replaced by a space varying regularization parameter. The space varying regularization parameter is updated in an iterative process utilizing the knowledge about the difference between the image which is obtained from the regularized solution

and the input image. The adaptive SFS algorithm starts with an oversmoothed regularization solution and iteratively refines this solution around discontinuities. At each iteration, a control signal is obtained from the error signal which has large positive values around discontinuities. The control signal is used to update the space varying regularization parameter. In order to speed up the iterative algorithm, one of the multigrid methods is applied to the iterative scheme and a hierarchical implementation of our adaptive SFS approach is presented. Integration of the adaptive SFS approach and the photometric stereo technique improves the reconstruction since we have two input images of the same scene generated using fixed viewing direction and different light source directions. The adaptive SFS algorithm were tested on various synthetic and real images and encouraging results were obtained. It can be seen from these results that the spatial control of the smoothness improves the reconstructed gradient field and height map obtained by the adaptive algorithm.

6.2 Summary of Contributions

A number of research contributions were presented in this thesis. These contributions can be summarized as follows:

- All the SFS constraints are combined in an energy functional and a general purpose SFS algorithm is derived by solving this minimization problem.

- The effects of the SFS constraints on the solution were explored using this general purpose SFS scheme.

- An adaptive SFS scheme was derived, which controls the smoothness over the image space by spatially varying regularization parameters determined from the error signal in an iterative process.

- The adaptive SFS algorithm has been shown by experiments to reduce the smoothing effect of the regularization term.

- The integration of the adaptive shape-from-shading approach and the photometric stereo technique has been achieved in order to improve the quality of the reconstruction.

- The multigrid method was applied to the adaptive SFS algorithm to speed up the iterative process.

6.3 Suggestions for Future Research

There are several possible directions for future research. Reflectance models used in SFS methods are too simplistic; recently more sophisticated models have been proposed (see Section 2.2). This is not only includes more accurate models for Lambertian, specular and hybrid reflectance but also includes replacing the assumption of the orthographic projection with perspective projection, which is a more realistic model of cameras in the real world. The traditional simplification of lighting conditions, assuming an infinite point light source, can also be eliminated by either assuming a non-infinite light source or simulating lighting conditions using a set of point sources. SFS methods employing more sophisticated models should be developed to provide more accurate and realistic results.

Another direction is the combination of the computational tasks used in early vision. One can use the results of the other tasks in order to improve the results of shape-from-shading.

In shape-from-shading problem, the shape is recovered by minimizing an energy functional involving some constraints. The relative weights of these constraints are determined by the corresponding parameter value of each constraint. Experimental results show that, choosing these parameter values are critical in order to obtain a reasonable solution. An iterative SFS scheme which determines algorithm parameter values automatically should be investigated.

Multiple images can also be employed by moving the viewer in order to successively refine the shape. The successive refinement can improve the quality

of estimates by combining estimates between image frames and reduce the computational time.

One problem with SFS is that the shape information in the shadow areas is not recovered since these areas do not provide enough intensity information. This can be solved if we make use of the information available from shape-from-shadow (shape-from-darkness) and combine it with the results from SFS. The depth values on the shadow boundaries from SFS can be used either as the initial values for shape-from-shadow or as constraints for the shape-from-shadow algorithm. In the case of multiple image frames, the information recovered from shadow in the previous frame can also be used for SFS in the next frame.

The adaptive shape-from-shading technique developed in this thesis can be adapted to synthetic aperture radar (SAR) imagery by incorporating radar reflectance models and by solving for surface slopes in a rotated system of coordinates which represents surface height relative to a plane parallel to the line-of-sight.

Finally, the application of the adaptive shape-from-shading to other early vision problems which have been solved using the regularization theory should be considered.

BIBLIOGRAPHY

[1] B. K. P. Horn and M. J. Brooks. *Shape from Shading.* The MIT Press, Cambridge, MA, 1989.

[2] B. K. P. Horn. *Robot Vision.* McGraw-Hill, New York, NY, 1986.

[3] A. P. Pentland. Local Shading Alaysis. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 6(2):170–187, March 1984.

[4] C. H. Lee and A. Rosenfeld. Improved Methods of Estimating Shape from Shading using the Light Source Coordinate System. *Artificial Intelligence*, 26(2):125–143, 1985.

[5] Q. Zheng and R. Chellappa. Estimation of Illuminant Direction, Albedo and Shape from Shading. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 13(7):680–702, July 1991.

[6] M. Gökmen and C. C. Li. Edge Detection and Surface Reconstruction using Refined Regularization. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 15(5):492–499, 1993.

[7] B. K. P. Horn. Obtaining Shape from Shading Information. In *The Psychology of Computer Vision*, pages 115–155, McGraw-Hill, New York, NY, 1975.

[8] R. J. Woodham. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering*, 19(1):139–144, January-February 1980.

[9] R. T. Frankot and R. Chellappa. A Method for Enforcing Integrability in Shape from Shading Algorithms. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 10(4):439–451, July 1988.

[10] B. K. P. Horn and M. J. Brooks. The Variational Approach to Shape from Shading. *Computer Vision, Graphics and Image Processing*, 33(2):174–208, February 1986.

[11] K. Ikeuchi and B. K. P. Horn. Numerical Shape from Shading and Occluding Boundaries. *Artifical Intelligence*, 17(1-3):141–184, August 1981.

[12] B. K. P. Horn. Height and Gradient from Shading. *International Journal of Computer Vision*, pages 37–75, 1990.

[13] B. K. P. Horn. *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View.* PhD thesis, Department of Electrical Engineering, MIT, 1970.

[14] M. J. Brooks and B. K. P. Horn. Shape and Source from Shading. In *Proc. of International Joint Conf. Artificial Intelligence*, pages 932–936, Los Angeles, August 1985.

[15] R. Szeliski. Fast Shape from Shading. *Computer Vision, Graphics and Image Processing*, 53:129–153, 1991.

[16] O. E. Vega and Y. H. Yang. Shading Logic: A Heuristic Approach to Recover Shape from Shading. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 15(6):592–597, 1993.

[17] Y. G. Leclerc and A. F. Bobick. The Direct Computation of Height from Shading. *Proc. of Computer Vision and Pattern Recognition*, pages 552–558, 1991.

[18] K. M. Lee and C. C. J. Kuo. Shape from Shading with a Linear Triangular Element Surface Model. Tech. Report No:172, Signal and Image Processing Institute, University of Southern California, 1991.

[19] J. Malik and D. Maydan. Recovering Three Dimensional Shape from a Single Image of Curved Objects. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 11(6), 1988.

[20] J. Oliensis. Shape from Shading as a Partially Well-Constrained Problem. *Computer Vision, Graphics and Image Processing*, 54:163–183, 1991.

[21] P. Dupuis and J. Oliensis. Direct Method for Reconstructing Shape from Shading. *Proc. of Computer Vision and Pattern Recognition*, pages 453–458, 1992.

[22] J. Oliensis and P. Dupuis. A Global Algorithm for Shape from Shading. *Proc. of International Conference on Computer Vision*, pages 692–701, 1993.

[23] M. Bichsel and A. P. Pentland. A Simple Algorithm for Shape from Shading. *Proc. of Computer Vision and Pattern Recognition*, pages 459–465, 1992.

[24] A. P. Pentland. Shape Information from Shading: A Theory about Human Perception. *Proc. of International Conference on Computer Vision*, pages 404–413, 1988.

[25] P. S. Tsai and M. Shah. A Simple Shape from Shading Algorithm. *Proc. of Computer Vision and Pattern Recognition*, pages 734–736, 1992.

[26] A. P. Pentland. Photometric Motion. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 6:879–890, 1991.

[27] S. K. Nayar, K. Ikeuchi and T. Kanade. Shape from Interreflections. *Proc. of International Conference on Computer Vision*, pages 1–11, 1990.

[28] D. Forsyth and A. Zisserman. Mutual Illumination. *Proc. of Computer Vision and Pattern Recognition*, pages 466–473, 1989.

[29] R. Kimmel. *Curve Evolution on Surfaces.* PhD thesis, Computer Science Department, Israel Institute of Technology, May 1995.

[30] R. Kimmel and A. M. Bruckstein. Global Shape from Shading. *Computer Vision and Image Understanding*, 62(3):360–369, 1995.

[31] H. B. Griffiths. *Surfaces.* Cambridge University Press, Cambridge, 1981.

[32] I. Shimshoni, R. Kimmel and A. M. Bruckstein. Global Shape from Shading. *Computer Vision and Image Understanding*, 64(1):188–189, July 1996.

[33] J. W. Hsieh, H. Y. M. Liao, M. T. Ko and K. C. Fan. Wavelet-Based Shape from Shading. Tech. Report, TR-IIS-95-001, Academia Sinica Institute of Information Science, 1995.

[34] S. Pankanti and A. K. Jain. Integrating Vision Modules: Stereo, Shading and Line Labeling. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 17(9), September 1995.

[35] A. D. Gross and T. E. Boult. Recovery of SHGCs from a Single View. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 18(2), February 1996.

[36] B. T. Phong. Illumination for Computer Generated Pictures. *Comm. of the Association for Computing Machinery*, 18(6):311–317, June 1975.

[37] K. E. Torrance and E. M. Sparrow. Theory for Off-Specular Reflection from Roughened Surfaces. *Journal of the Optical Society of America*, 57(9):1105–1114, September 1967.

[38] G. Healey and T. O. Binford. Local Shape from Specularity. *Computer Vision, Graphics and Image Processing*, 42(1):62–86, April 1988.

[39] S. K. Nayar, K. Ikeuchi and T. Kanade. Surface Reflection: Physical and Geometrical Perspectives. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 13(7):611–634, 1991.

[40] M. Oren and S. K. Nayar. Diffuse Reflectance from Rough Surfaces. *Proc. of Computer Vision and Pattern Recognition*, pages 763–764, 1993.

[41] J. J. Clark. Active Photometric Stereo. *Proc. of Computer Vision and Pattern Recognition*, pages 29–34, 1992.

[42] D. R. Hougen and N. Ahuja. Estimation of the Light Source Distribution and its use in Integrated Shape Recovery from Stereo and Shading. *Proc. IEEE International Conf. on Computer Vision*, pages 29–34, 1993.

[43] M. S. Langer and S. W. Zucker. Diffuse Shading, Visibility Fields and the Geometry of Ambient Light. *Proc. of International Conference on Computer Vision*, pages 29–34, 1993.

[44] B. K. P. Horn. Hill Shading and Reflectance Map. *Proceedings of the IEEE*, 69(1):14–47, January 1981.

[45] A. P. Pentland. Linear Shape from Shading. *International Journal of Computer Vision*, 4:153–162, 1990.

[46] A. R. Bruss. The Eikonal Equation: Some Results Applicable to Computer Vision. *Journal of Mathematical Physics*, 23(5):890–896, May 1982.

[47] B. V. H. Saxberg. *A Modern Differential Geometric Approach to Shape from Shading.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1988.

[48] T. Poggio, V. Torre and C. Koch. Computational Vision and Regularization Theory. *Nature*, 317(26):314–319, September 1985.

[49] D. Terzopoulos. Image Analysis using Multigrid Relaxation Methods. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 8(2):129–139, March 1986.

[50] M. Bertero, T. Poggio and V. Torre. Ill-Posed Problems in Early Vision. Tech. Report, Memo 924, MIT-AI, 1987.

[51] J. Hadamard. Lectures on The Cauchy Problem in Linear Partial Differential Equations. Yale University Press, New Heaven CT, 1923.

[52] A. N. Tikhonov and V. Y. Arsenin. *Solution of Ill-Posed Problems.* Winston and Sons, Washington D.C., 1977.

[53] V. K. Ivanov. On Linear Problems which are not Well-Posed. *Soviet Math. Dokl.*, (3):981–983, 1962.

[54] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems.* Springer-Verlag, New York, NY, 1984.

[55] V. A. Morozov. On the Solution of Functional Equations by the Method of Regularization. *Soviet Math. Dokl.*, (7):414–417, 1966.

[56] G. Wahba. Practical Approximate Solutions to Linear Operator Equations when the Data are Noisy. *SIAM J. Numer. Anal.*, 14:651–667, 1977.

[57] J. T. Marti. An Algorithm for Computing Minimum Norm Solutions of Fredholm Integral Equations of the First Kind. *SIAM J. Numer. Anal.*, 15:1071–1076, 1978.

[58] D. Geiger and T. Paggio. An Optimal Scale for Edge Detection. Tech. Report, Memo 1078, MIT-AI, 1988.

[59] H. W. Engle and A. Neubauer. An Improved Version of Martis Method for Solving Ill-Posed Linear Integral Equations. *Math. Comp.*, 45(172):405–416, 1985.

[60] M. A. Gennert and A. L. Yuille. Determining the Optimal Weights in Multiple Objective Function Optimization. *Proc. IEEE International Conf. on Computer Vision*, pages 87–89, 1988.

[61] B. K. P. Horn. Understanding Image Intensities. *Artifical Intelligence*, 8(2):201–231, April 1977.

[62] D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination.* Chelsea Publishing Company, New York, NY, 1952.

[63] A. R. Bruss. Is What You See What You Get? In *Proc. of International Joint Conf. Artificial Intelligence*, pages 1053–1056, Karlsruhe, West Germany, August 1983.

[64] B. K. P. Horn and R. W. Sjoberg. Calculating the Reflectance Map. *Applied Optics*, 18(11):1770–1779, June 1979.

[65] R. J. Woodham. A Cooperative Algorithm for Determining Surface Orientation from a Single View. In *Proc. of International Joint Conf. Artificial Intelligence*, pages 635–641, Cambridge, MA, August 1977.

[66] T. Strat. A Numerical Method for Shape from Shading for a Single Image. S.M. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1979.

[67] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*, NBS, Monograph 160, National Bureau of Standarts, U.S. Department of Commerce, Washington, D.C., October 1977.

[68] R. Court and D. Hilbert. *Methods of Mathematical Physics.* Volume I, Wiley, New York, NY, 1962.

[69] W. E. L. Grimson. *From Images to Surfaces: A Computational Study of the Human Early Visual Systems.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1981.

[70] W. E. L. Grimson. An Implementation of a Compuational Theory of Visual Surface Interpolation. *Computer Vision, Graphics and Image Processing*, 22(1):39–69, 1983.

[71] D. Terzopoulos. Multilevel Reconstruction of Visual Surfaces: Variational Principles and Finite Element Representations. Tech. Report, Memo 671, MIT-AI, April 1982.

[72] D. Terzopoulos. Multilevel Computational Process for Visual Surface Reconstruction. *Computer Vision, Graphics and Image Processing*, 24(1):52–96, 1983.

[73] D. Terzopoulos. Multigrid Relaxation Methods and the Analysis of Lightness, Shading and Flow. Tech. Report, Memo 803, MIT-AI, October 1984.

[74] J. M. Brady and B. K. P. Horn. Rotationally Symmetric Operators for Surface Interpolation. *Computer Vision, Graphics and Image Processing*, 22(1):70–94, April 1983.

[75] R. T. Frankot and R. Chellappa. Application of a Shape from Shading Tecnique to Synthetic Aperture Radar Imaginery. In *International Geoscience and Remote Sensing Symposium*, pages 1323–1329, Ann Arbor, MI, May 1987.

[76] M. J. Brooks. *Shape from Shading Discretely.* PhD thesis, Department of Computer Science, Essex University, Colchester, England, 1982.

[77] A. Gültekin and M. Gökmen. Görüntü Tonlarından Şekil Elde Etme. In *2. Sinyal İşleme ve Uygulamaları Kurultayı Bildiriler Kitabı*, pages 71–76, Gökova, Muğla, Türkiye, April 1994.

[78] P. Dupuis and J. Oliensis. A Global Algorithm for Shape from Sahing. *Proc. of International Conference on Computer Vision*, pages 692–701, 1993.

[79] B. K. P. Horn, R. J. Woodham and W. M. Silver. Determining Shape and Reflectance using Multiple Images. Tech. Report, Memo 490, MIT-AI, August 1978.

[80] R. J. Woodham. Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from a Single View. *Image Understanding Systems and Industrial Applications*, SPIE, 155:136–143, 1978.

[81] R. J. Woodham. *Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1978.

[82] W. Silver. Determining Shape and Reclectance using Multiple Images. S.M. Thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, June 1980.

[83] B. K. P. Horn, R. Szeliski and A. L. Yullie. Impossible Shaded Images. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 15(2):166–170, 1993.

[84] R. Zhang, P. S. Tsai, J. E. Crayer and M. Shah. Analysis of Shape from Shading Techniques. Tech. Report No:23, Department of Computer Science, University of Central, Florida, Orlando, 1992.

[85] A. Gültekin and M. Gökmen. Adaptive Shape from Shading. In *Proc. of the Eleventh International Symposium on Computer and Information Sciences (ISCIS-XI)*, pages 83–92, Antalya, Türkiye, November 1996.

[86] A. Gültekin and M. Gökmen. Adaptive Shape from Shading. *ELEKTRIK - Turkish Journal of Electrical Engineering and Computer Science*, to appear, 1997.

[87] Q. Zheng and R. Chellappa. A Robust Algorithm for Inferring Shape from Shading. USC-SIPI Report No:159, Univ. of Southern California, Los Angeles, 1990.

[88] J. G. Harris. The Coupled Depth/Slope Approach to Syrface Reconstruction. S.M. Thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1986.

[89] J. G. Harris. A New Approach to Surface Reconstruction: The Coupled Depth/Slope Model. *Proc. of International Conference on Computer Vision*, pages 277–283, June 1987.

RESUME

Atilla Gültekin was born in Samsun on March 21, 1965. He graduated from Maçka Teknik Lisesi in 1983. He received his B.S. degree in Electrical and Electronics Engineering from Istanbul Technical University in 1987.

He earned the M.S. degree in 1989 in Electrical and Electronics Engineering from Istanbul Technical University. He worked as a research assistant in Istanbul Technical University between 1987-1988. He worked for Siemens-Nixdorf Bilgisayar Sistemleri A.Ş. as a system engineer from 1988 to 1995. He also, at the same time, started to study for Ph.D. degree in Control and Computer Engineering at Istanbul Technical University in 1989.

He has been working in Bilpa A.Ş. since 1995 as a Project Manager.