

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Engenharia de Computação e
Automação Industrial

Algoritmos de Shape from Shading e
Aplicação para Casos com Posições de Fonte de Luz e
Observador Coincidentes

por

Carlos Henrique Quartucci Forster

sob orientação de

Clésio Luis Tozzi

Tese apresentada à Faculdade de Engenharia
Elétrica e de Computação, FEEC-UNICAMP,
como requisito parcial para obtenção do título
de MESTRE EM ENGENHARIA ELÉTRICA.

Novembro - 1999

Campinas - SP

Este exemplar corresponde a redação final da tese
defendida por Carlos Henrique Q. Forster
e aprovada pela Comissão
Julgada em 22/11/1999.
Orientador [Assinatura]



20 0002571

UNIDADE	Be
N.º CHAMADA:	F/UNICAMP
	F774a
V.	Er.
TOMBO BC/	40488
PROC.	278100
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	311,00
DATA	14/03/00
N.º CPD	

CM-00135073-9

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

F774a Forster, Carlos Henrique Quartucci
Algoritmos de shape from shading e aplicações para
casos com posições de fonte de luz e observador
coincidentes / Carlos Henrique Quartucci Forster.--
Campinas, SP: [s.n.], 1999.

Orientador: Clésio Luis Tozzi
Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Visão por computador. 2. Endoscopia. 3. Cor. 4.
Recuperação de dados (Computação). I. Tozzi, Clésio
Luis. II. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação.
III. Título.

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Engenharia de Computação e Automação Industrial

**Algoritmos de Shape from Shading e Aplicações
para Casos com Posições de Fonte de Luz e
Observador Coincidentes**

por

Carlos Henrique Quartucci Forster

Tese apresentada à Faculdade de Engenharia
Elétrica e de Computação, FEEC-UNICAMP,
como requisito parcial para obtenção do título
de MESTRE EM ENGENHARIA ELÉTRICA

Aprovada em 22/11/1999 por

Clésio Luis Tozzi (orientador)

Fernando José Von Zuben (FEEC-UNICAMP)

Alexandre Xavier Falcão (IC-UNICAMP)

AGRADECIMENTOS

Agradeço, em primeiro lugar, ao Prof. Dr. Clésio Luis Tozzi por ter me convidado para realizar este trabalho e pela confiança em minha competência, tendo me dado bastante liberdade para atuar nesse projeto.

Agradeço ao Gastrocentro da Unicamp, principalmente ao Dr. Fábio Guerrazzi e ao Dr. José Geraldo Ferraz pela colaboração e os esclarecimentos sobre a questão da endoscopia. Gostaria de agradecer também o falecido Dr. Mauro Dantas Neder, que nem cheguei a conhecer, mas que foi responsável pela iniciativa da colaboração entre o Gastrocentro e o DCA.

Agradeço aos professores do departamento e aos demais professores desta universidade por minha formação, que é um dos bens mais preciosos para mim. Agradeço a todos meus amigos e parentes próximos pelo apoio e conforto.

Agradeço à FAPESP pelo apoio financeiro. Admiro a seriedade dessa instituição.

Aos meus pais.

RESUMO

Este trabalho trata o problema de reconstrução tridimensional a partir de um sistema monocular de captura de imagens. Especificamente, o problema de Shape from Shading, reconstrução tridimensional a partir de imagens de sombreamento, é tratado na forma de uma revisão bibliográfica e de construção de algoritmos para se resolver um caso específico onde as posições da fonte de luz e da câmera coincidem e estão próximas do objeto. Dois métodos de Shape from Shading são implementados, um deles é baseado em otimização e o outro é baseado em expansão de curvas. A fim de aplicar a técnica em imagens de endoscopia, alguns problemas práticos são solucionados, como a correção da distorção da imagem e a remoção do componente especular de reflexão. A remoção da reflexão especular se faz através de um filtro baseado no modelo dicromático de reflexão. Considerando-se a recuperação da forma, resultados satisfatórios são obtidos, tanto para imagens sintéticas quanto para imagens de objetos reais.

ABSTRACT

The subject of this work is the solution of the 3d reconstruction problem from a single shading image captured by a monocular system. More specifically, the Shape from Shading problem is addressed. A bibliographic review and the implementation of algorithms to solve the specific case in which the camera and light source positions are coincident and close to the object are presented. Two of the Shape from Shading methods described were implemented, one based on optimization and the other based on curve expansion. In order to test the application of the technique in endoscope images, some practical problems were addressed: the correction of image distortion and the removal of the specular component of reflection, using the dichromatic reflection model. Considering the shape reconstruction, the obtained results for synthetic real objects images were all satisfactory.

ÍNDICE

AGRADECIMENTOS	2
RESUMO	4
ABSTRACT	5
ÍNDICE	6
TABELA DE SÍMBOLOS	9
CAPÍTULO 1 - INTRODUÇÃO	13
CAPÍTULO 2 - FUNDAMENTOS DA FORMAÇÃO DE IMAGENS DE SOMBREAMENTO	16
2.1 MODELOS DE ÓPTICA GEOMÉTRICA PARA FORMAÇÃO DE IMAGENS	17
2.2 A LEI DE REFLEXÃO DE LAMBERT	22
2.3 MAPA DE REFLETÂNCIA	22
2.4 A EQUAÇÃO DA IRRADIÂNCIA DA IMAGEM	26
2.4.1 <i>Ambigüidade</i>	27
CAPÍTULO 3 - ALGORITMOS DE SHAPE FROM SHADING	29
3.1 SHAPE FROM SHADING POR OTIMIZAÇÃO	30
3.1.1 <i>O Método Variacional de Horn</i>	32
3.1.2 <i>Método de Frankot e Chellappa</i>	39
3.1.3 <i>O Método de Zheng e Chellappa</i>	40
3.1.4 <i>Métodos de Gradiente</i>	41

3.1.5 <i>Outras Variações</i>	42
3.2 SHAPE FROM SHADING POR EXPANSÃO DE CURVAS	43
3.2.1 <i>O Método de Kimmel e Bruckstein</i>	43
3.2.2 <i>Shape from Shading pela Transformada da Distância Ponderada</i>	48
3.2.3 <i>Representação como Problema de Controle Ótimo</i>	49
3.2.4 <i>Algoritmos de Oliensis e Dupuis</i>	50
3.2.5 <i>Equivalência com o Método de Kimmel e Bruckstein</i>	51
3.2.6 <i>Condições de Borda ou Pontos Singulares?</i>	51
3.2.7 <i>O Problema de Shape from Shading Global</i>	51
3.3 SHAPE FROM SHADING POR ANÁLISE LOCAL	55
3.4 DISCUSSÃO SOBRE AS ALTERNATIVAS APRESENTADAS	56
CAPÍTULO 4 - ESTUDO DO ENDOSCÓPIO PARA CAPTURA DE IMAGENS DE SOMBREAMENTO	57
4.1 O MODELO DO ENDOSCÓPIO	57
4.1.1 <i>Aquisição de Imagens de Endoscopia, um Modelo</i>	57
4.1.2 <i>Problemas Relacionados à Endoscopia</i>	58
4.2 IMAGENS COM DISTORÇÃO RADIAL	59
4.2.1 <i>Warping</i>	60
4.2.2 <i>Parametrização da Superfície</i>	61
4.2.3 <i>Calibração da Distorção da Câmera</i>	61
4.3 IMAGENS COLORIDAS	62
4.3.1 <i>Espaços de Cor</i>	62
4.3.2 <i>O Modelo de Reflexão de Phong</i>	66
4.3.3 <i>O Modelo Dicromático</i>	70
4.3.4 <i>Utilizando o Modelo Dicromático para a Remoção de Reflexão Especular</i>	72
CAPÍTULO 5 - IMPLEMENTAÇÃO DOS ALGORITMOS PARA RECONSTRUÇÃO TRIDIMENSIONAL EM IMAGENS DE ENDOSCOPIA	75
5.1 ADAPTAÇÃO DO ALGORITMO DE OKATANI E DEGUCHI	75
5.1.1 <i>O Algoritmo Original</i>	75
5.1.2 <i>Adaptação para uma Câmera com Distorção</i>	77
5.2 ALGORITMO DE RECONSTRUÇÃO POR OTIMIZAÇÃO	79
5.2.1 <i>Outras Alternativas para Resolver o Sistema de Equações</i>	82
CAPÍTULO 6 - RESULTADOS	84
6.1 CORREÇÃO DA DISTORÇÃO	84
6.2 REMOÇÃO DA REFLEXÃO ESPECULAR	87
6.3 RECONSTRUÇÃO DE OBJETOS	89

6.3.1 Hemisfério	90
6.3.2 Diferença de Gaussianas	93
6.3.3 Monkey Saddle	96
6.3.4 Chapéu (diferença de gaussianas)	98
6.3.5 Imagens de Objetos Reais	100
6.4 RESULTADOS EM IMAGENS DE ENDOSCOPIA	103
CAPÍTULO 7 - CONCLUSÃO	104
APÊNDICE A - OUTROS ALGORITMOS DE SHAPE FROM SHADING	106
A.1 SHAPE FROM SHADING POR ANÁLISE LOCAL	106
A.1.1 Estimadores da Direção da Fonte de Luz	106
A.1.2 Estimadores da Normal à Superfície utilizando Estimador de Albedo	107
A.2 O MÉTODO CLÁSSICO DE EXPANSÃO DE CURVAS CARACTERÍSTICAS DE HORN	108
A.2.1 Formulação do Esquema	109
A.2.2 Tratamento de Pontos Especiais	110
A.2.3 Tratamento de Bordas	111
A.2.4 Problemas Topológicos	112
REFERÊNCIAS	113
ÍNDICE REMISSIVO POR AUTORES	119

TABELA DE SÍMBOLOS

$\ \bullet\ $	norma euclidiana
\bar{x}_{ij}	média dos vizinhos de x_{ij}
z_x	derivada parcial de z em relação a x
$[]^T$	transposta da matriz
$\nabla r(u, v)$	campo vetorial normal à superfície definida por $r(u, v)$
α	ângulo em relação ao eixo óptico
α	ângulo formado pelo centro da circunferência, o centro de projeção e um ponto da circunferência para o experimento do item 4.2.3
$\delta(x)$	função impulso ou função delta de Dirak
ΔI	pequena área da imagem
ΔO	pequena área da superfície
Δt	passo de integração
ε	distância entre pixels vizinhos
(θ_i, ϕ_i)	ângulos que determinam a direção do raio incidente
(θ_e, ϕ_e)	ângulos que determinam a direção do observador
λ	coeficiente de regularização
λ	comprimento de onda ou índice que identifica uma função de base espectral

μ	coeficiente que pondera a importância do termo de integrabilidade na função objetivo
ρ	albedo da superfície
σ	albedo, na formulação de Okatani e Deguchi [1997]
σ, θ	ângulo polar (<i>slant</i>)
τ, ϕ	azimute (<i>tilt</i>)
$\phi(x, y, t)$	função associada à representação implícita de uma curva de profundidade constante e igual a t
$C(t), C(s, t)$	representação paramétrica de uma curva de profundidade constante e igual a t
$c_B(\theta)$	coeficiente de reflexão de corpo associado à geometria da superfície
$c_I(\theta)$	coeficiente de reflexão de interface associado à geometria da superfície
D	distância que uma curva se expande na direção normal
d	diâmetro da lente
E	irradiância
$E(\lambda)$	irradiância espectral na superfície
e	função objetivo
e_r	erro em intensidade da imagem ou erro de reconstrução
e_s	termo de suavidade ou regularização
F	velocidade normal à curva para direção da fonte de luz equivalente à direção do observador
f	distância focal da lente
f	parâmetro de distorção radial
$f(\theta_i, \phi_i; \theta_e, \phi_e)$	função de distribuição de refletância bidirecional ou BRDF
(f, g)	coordenadas estereográficas para o espaço gradiente
$G(\bullet)$	distorção do mapa de refletância lambertiano na formulação de Okatani e Deguchi [1997]
$G(x, y, \hat{n})$	velocidade normal à curva para direção de fonte de luz genérica
$h(x, y)$	mapa de profundidade

I_0	intensidade da fonte de luz
I_a	intensidade da iluminação ambiente
I_b	intensidade azul do pixel
I_g	intensidade verde do pixel
I_r	intensidade vermelha do pixel
k_a	coeficiente de iluminação ambiente
k_d	coeficiente de reflexão difusa
k_s	coeficiente de reflexão especular
\vec{l}	direção da fonte de luz
L	radiância
n	expoente de reflexão especular
\vec{n}	vetor normal à superfície na forma $(p, q, -1)$
\hat{n}	vetor unitário normal à curva
\vec{N}	vetor normal à superfície na forma (n_1, n_2, n_3)
(p, q)	coordenadas gnômicas para o espaço gradiente
(p_l, q_l)	direção da fonte de luz em coordenadas gnômicas
(p_s, q_s)	direção da fonte de luz em coordenadas gnômicas
$R(p, q)$	mapa de refletância
$Rs(f, g)$	mapa de refletância
$r(u, v)$	distância do centro de projeção ao ponto da superfície correspondente ao pixel de coordenadas (u, v)
$S_B(\lambda)$	coeficiente de reflexão de corpo dependente do comprimento de onda
$S_I(\lambda)$	coeficiente de reflexão de interface dependente do comprimento de onda
sr	esterorradiano
t	tempo e também está relacionado com a profundidade nos métodos de expansão de curvas de profundidade constante
(u, v)	coordenadas do pixel na imagem

$v(x, y, t)$	velocidade normal à curva
(x, y)	coordenadas do ponto da superfície e, no caso da projeção ortográfica, eqüivalem às coordenadas do pixel na imagem
$Y(\theta, \lambda)$	intensidade de luz refletida em função da geometria e do comprimento de onda ou função de base espectral
z	distância de um ponto ao plano perpendicular ao eixo óptico que contem o centro de projeção
z_{pi}	distância do plano imagem ao centro de projeção
$z(x, y)$	mapa de profundidade
$\begin{cases} z(u, v) \\ x(u, v) \\ y(u, v) \end{cases}$	coordenadas do ponto da superfície correspondente ao pixel de coordenadas (u, v) na imagem

Capítulo 1 - INTRODUÇÃO

Os modelos de formação de imagem permitem que, a partir de modelos geométricos tridimensionais que descrevem superfícies, câmeras e fontes de luz, obtenha-se a imagem bidimensional que se formaria no ato da fotografia. Esses mesmos modelos de formação de imagem podem ser utilizados para resolver o problema inverso: obter um modelo geométrico de objetos do mundo real a partir de imagens bidimensionais capturadas por câmeras. O problema de reconstrução tridimensional definido dessa maneira é dificultado dada a perda de informação de profundidade decorrente da aquisição da imagem.

Informações adicionais ou restrições quanto à natureza dos objetos e ao modo de aquisição de imagens são necessárias para se obter uma solução, compensando o que foi perdido nessa projeção. Por exemplo, no caso da Visão Estéreo em que se trabalha com pares de imagem, deve-se ser capaz de relacionar os pontos de uma imagem com os pontos da outra. Outro exemplo é a restrição de objeto rígido, isto é, os movimentos do objeto podem ser utilizados para se estimar a forma a partir de uma seqüência de imagens porque as distâncias, no espaço euclidiano da cena, não se alteram. Em geral, procura-se utilizar informações que a própria visão humana utiliza, como movimento, silhueta, textura e sombreamento. Essas características são utilizadas na prática em algoritmos de Visão Computacional.

A intensidade da imagem é produto do albedo pelo sombreamento. O sombreamento é a grandeza que está relacionada com a geometria do objeto e com a iluminação e é independente do material da superfície. O albedo é a grandeza que define as características

fotométricas do material. O sombreamento pode ser obtido diretamente do modelo geométrico. No caso inverso, é possível se obter o sombreamento a partir da imagem se o albedo for conhecido. Assim, uma vez que se tenha o sombreamento, quer-se determinar a geometria do objeto. Esse problema é conhecido como Shape from Shading (SFS) e foi introduzido por Horn [1975] na forma de um sistema de equações diferenciais parciais (PDE).

De modo geral, o problema de Shape from Shading é tratado sob as seguintes condições para a formação de imagem: projeção ortográfica, fonte de luz direcional e reflexão lambertiana. Mesmo sob essas condições, não é um problema simples. Além disso, há casos em que a câmera fica muito próxima ao objeto, não sendo possível assumir a hipótese de projeção ortográfica. Da mesma forma, a fonte de luz pode ficar muito próxima ao objeto, impedindo o emprego de um modelo de luz direcional. O algoritmo de Shape from Shading, neste caso, deve ser adaptado para um sistema com fonte de luz pontual e projeção perspectiva.

O caso mais estudado de Shape from Shading é o caso particular de reflexão lambertiana. Em geral, aproximar a realidade utilizando o modelo de reflexão lambertiana é aceitável [Pentland 1984]. Porém, sabe-se que a reflexão especular é um componente importante da reflexão em superfícies. Além disso, existem pontos em que a reflexão especular é intensa e satura o sensor da câmera. É desejável que, ao processar essas imagens, essas regiões de reflexão especular intensa possam ser removidas. Um resultado melhor ainda seria eliminar o componente de reflexão especular, obtendo uma imagem de reflexão difusa pura (reflexão lambertiana). Para superfícies para as quais o modelo dicromático (e o modelo de Phong) é válido, a disponibilidade de imagens coloridas permite não só segmentar os pontos especulares, mas também, isolar os componentes de reflexão especular e difusa. Isolando o componente de reflexão difusa e conhecendo a solução da reconstrução para o caso lambertiano, é possível estender a solução para o caso dicromático.

Muitos algoritmos de Shape from Shading, por se tratarem de métodos de solução de equações diferenciais, necessitam de valores de contorno ou condições iniciais. Essas condições podem ser fornecidas na forma da normal à superfície na borda de oclusão, de uma curva inicial de profundidade constante, ou, ainda, de pontos singulares.

Este trabalho aborda a questão da reconstrução tridimensional baseada em sistemas monoculares de captura de imagens, posicionados próximos ao objeto e com uma única fonte de luz pontual, também próxima ao objeto, e tem como objetivo a avaliação de algoritmos para reconstrução tridimensional de imagens a partir da informação de sombreamento. Como aplicação e a fim de testar os resultados, foram utilizadas imagens de endoscopia, cujos sistemas de captura de imagens se encaixam no modelo de perspectiva e fonte de luz próxima. Adicionalmente, foi tratada a distorção da câmera do endoscópio, utilizando um modelo de projeção esférica. As imagens de intensidade a serem utilizadas com o algoritmos de reconstrução tridimensional deste trabalho são restritas a objetos que, por partes, tenham albedo uniforme. Esta restrição do albedo pode ser relaxada dado o conhecimento do albedo em todo ponto da imagem.

Este texto está estruturado da seguinte forma. No capítulo 2, uma revisão de conceitos básicos é apresentada a fim de definir a função de distribuição de refletância bidirecional e o mapa de refletância formulando a equação da irradiância da imagem, que é a equação central para o problema de Shape from Shading. No capítulo 3, os algoritmos de Shape from Shading mais referenciados na literatura são apresentados na forma de uma taxonomia simples e aqueles de interesse do projeto são descritos mais detalhadamente. Aproveita-se para fazer considerações quanto à natureza do problema e à natureza das soluções. Os aspectos considerados são a integrabilidade da solução, o uso de regularização, a necessidade de condições de borda, a análise dos pontos singulares e solução de problemas inversos.

No capítulo 4, são descritos os modelos geométricos e radiométricos que podem ser utilizados no tratamento de imagens. Os itens discutidos são imagens coloridas, modelos de representação de cor e o modelo de câmera e fonte de luz.

Os algoritmos implementados, um baseado em otimização e outro baseado em expansão de curvas, são descritos no capítulo 5 e, em seguida, no capítulo 6 são apresentados os resultados experimentais obtidos.

No apêndice, são apresentados exemplos de algoritmos de análise local de sombreamento e o método clássico de Shape from Shading por expansão de curvas.

Capítulo 2 - FUNDAMENTOS DA FORMAÇÃO DE IMAGENS DE SOMBREAMENTO

Neste capítulo, é apresentado um modelo de formação de imagem baseado em Óptica Geométrica, a fim de se definir o problema de Shape from Shading e de se obter uma formulação para este problema. Esse modelo depende de informação precisa sobre o material da superfície, expressa na forma de uma função de quatro variáveis, a BRDF (Função de Distribuição da Refletância Bidirecional).

O problema de Shape from Shading consiste em obter a partir de uma única imagem de sombreamento, um mapa de profundidade, isto é, a função que relaciona cada pixel da imagem com a distância do ponto correspondente da cena até o plano imagem. A imagem de sombreamento é determinada pelo conhecimento do albedo da superfície e pela imagem de intensidade, que pode, por exemplo, ser obtida por uma câmera comum (CCD, Vidicon ou até filme fotográfico). A relação entre o sombreamento e a geometria da superfície é dada pela equação da irradiância da imagem e pelo mapa de refletância, que serão apresentados neste capítulo.

Este capítulo trata do mecanismo de formação da imagem através de modelos de Óptica Geométrica, cuja compreensão não é totalmente necessária para entendimento dos demais capítulos do texto, mas esclarece quanto a origem da formulação utilizada. Recomenda-se que sejam lidos, pelo menos, os itens 2.3 e 2.4 para que não haja problemas na compreensão dos capítulos posteriores.

2.1 Modelos de Óptica Geométrica para Formação de Imagens

Alguns conceitos básicos de Óptica Geométrica serão revistos nesta seção. Com base nestes conceitos, será mostrado como utilizar a função de distribuição de refletância bidirecional (BRDF) para calcular a intensidade de luz recebida por uma região da imagem a partir de uma fonte de luz que ilumina uma superfície. Tendo-se a BRDF tabelada com precisão, esse modelo é muito útil em problemas de Visão. O modelo de formação de imagens apresentado neste capítulo é bastante abrangente quanto às características da superfície por utilizar a BRDF. No item 4.3, outros modelos de formação de imagens são considerados com a finalidade de extrair informação de imagens coloridas.

Esta seção foi baseada em [Horn 1990] e [Okatani e Deguchi 1997].

♦ Ângulo Sólido

O ângulo sólido é uma medida sobre cones. Podemos entender cone como o sólido formado pela varredura de toda semi-reta que passa por um ponto determinado e tem interseção não nula com uma determinada região fechada de um plano. Criando-se uma casca esférica unitária (de raio unitário) com centro na ponta do cone, o ângulo sólido, medido em esterorradianos corresponde à área da superfície formada pela interseção do cone com a casca esférica. O ângulo sólido máximo que se pode obter com essa definição é o de um hemisfério de direções e tem valor de $2\pi sr$.

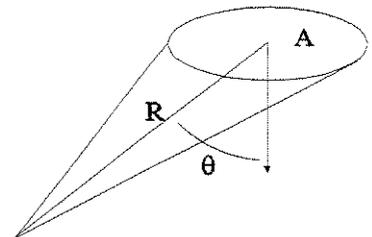


Figura 2-1 Ângulo sólido

O ângulo sólido que compreende uma região plana pequena de área A e distância R da origem do cone é dado por

$$\Omega = \frac{A \cos \theta}{R^2}$$

onde θ é o ângulo entre a reta que conecta a origem do cone com um ponto da região plana e a reta normal à região plana. O efeito causado por esse ângulo θ é conhecido como *foreshortening* (redução da área aparente).

- **Radiância L e Irradiância E**

Define-se irradiância como a quantidade de luz que uma superfície recebe. Essa quantidade pode ser medida em W/m^2 (Watt por metro quadrado), ou seja, potência por área ou intensidade. Essa grandeza será representada pelo símbolo E neste texto.

Define-se radiância como a quantidade de luz emitida por uma superfície para um ângulo sólido específico. Ela é medida levando em conta o ângulo sólido que está sendo avaliado e, portanto, a unidade utilizada é W/m^2sr (Watt por metro quadrado por esterorradiano), ou seja, intensidade por ângulo sólido. A notação que será utilizada para radiância é L .

- **Formação da Imagem**

Consideram-se uma região do objeto de área ΔO e a área correspondente da imagem ΔI que é dada pela projeção perspectiva de ΔO sobre o plano imagem dado o centro de projeção da câmera. A radiância da fonte de luz compreendida no cone formado pela fonte de luz e a região de área ΔO será chamada L_f . A irradiância da superfície do objeto na região de área ΔO será E_s e a radiância da mesma região será L_s . A irradiância da região considerada da imagem chamar-se-á E_i .

O esquema da figura 2-2 mostra a relação entre essas variáveis. A fim de se relacionar L_f com E_i , estabelecem-se as relações entre essas duas variáveis e as variáveis intermediárias L_s e E_s .

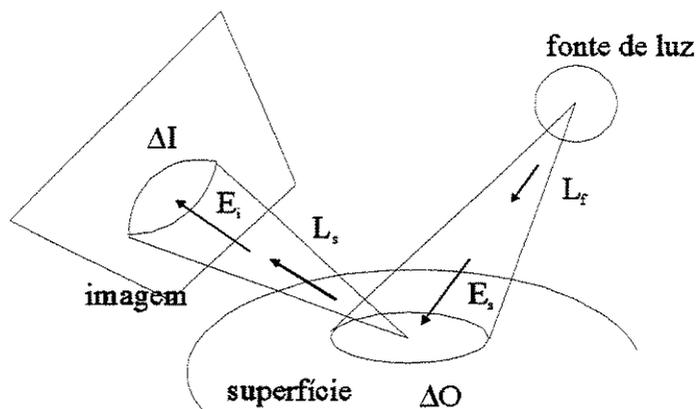


Figura 2-2 Relação entre variáveis para formação da imagem

• **Relação entre E_i (Irradiância na Imagem) e L_s (Radiância na Superfície)**

Para uma câmera cuja lente tenha diâmetro d e distância f do plano imagem será calculada a relação $\frac{E_i}{L_s}$.

Define-se o eixo óptico da lente como perpendicular ao plano imagem, com origem no centro de projeção e crescente no sentido do plano imagem ao centro de projeção.

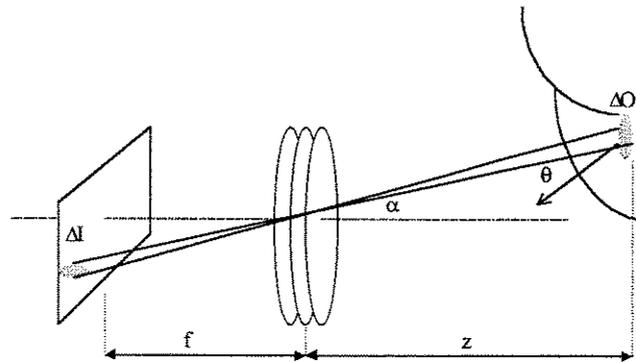


Figura 2-3 Cálculo da irradiância da imagem

O ângulo sólido correspondente à área ΔI em relação ao centro de projeção da câmera é dado por

$$\frac{\Delta I \cos \alpha}{(f / \cos \alpha)^2},$$

onde α é o ângulo entre o eixo óptico da lente, perpendicular ao plano imagem, e a reta que liga o centro de projeção ao centro da região de área ΔI . Note-se que $f / \cos \alpha$ é a distância do centro de projeção ao centro da região de área ΔI .

O ângulo sólido gerado pela área ΔO em relação ao centro de projeção é, portanto, $\frac{\Delta O \cos \theta}{(z / \cos \alpha)^2}$ onde z é a projeção do centro da região de área ΔO sobre o eixo óptico e θ é o ângulo entre a reta normal à superfície no centro de ΔO e a reta que liga o centro de projeção ao centro de ΔO .

Como os ângulos sólidos calculados acima são iguais porque são opostos pelo vértice, tem-se

$$\frac{\Delta O}{\Delta I} = \frac{\cos \alpha}{\cos \theta} \left(\frac{z}{f} \right)^2.$$

O ângulo sólido gerado pela lente em relação ao centro de ΔO é

$$\Omega = \frac{\pi d^2 \cos \alpha}{4 (z / \cos \alpha)^2} = \frac{\pi \left(\frac{d}{z} \right)^2}{4} \cos^3 \alpha.$$

A potência, que atravessa a lente, originada na região ΔO do objeto em consideração é dada por

$$\Delta P = L_s \cdot \Delta O \cdot \Omega \cdot \cos \theta = L_s \cdot \Delta O \cdot \frac{\pi \left(\frac{d}{z} \right)^2}{4} \cos^3 \alpha \cdot \cos \theta.$$

Essa potência corresponde exatamente à potência que atinge a região de área ΔI , logo

$$E_i = \frac{\Delta P}{\Delta I} = L_s \cdot \frac{\Delta O}{\Delta I} \cdot \frac{\pi \left(\frac{d}{z} \right)^2}{4} \cos^3 \alpha \cdot \cos \theta.$$

Substituindo, obtém-se

$$\frac{E_i}{L_s} = \frac{\pi \left(\frac{d}{f} \right)^2}{4} \cos^4 \alpha.$$

Na prática, o termo $\cos^4 \alpha$ pode ser substituído por uma constante dado que os sistemas de lentes normalmente compensam este termo.

♦ **Relação entre L_s e E_s , Radiância e Irradiância na Superfície, respectivamente**

A relação $\frac{L_s}{E_s}$ depende da direção do raio incidente, da direção do observador e do tipo de material e é tabelada para esses parâmetros. As direções são representadas por pares de números determinados pelos ângulos *ângulo polar* (*slant*) e *azimute* (*tilt*). O ângulo polar, representado por θ , é normalmente associado ao ângulo formado entre a reta (que representa a direção) e o eixo z . Neste texto, o eixo z será normal à superfície no ponto em questão. O ângulo polar é o complemento da *elevação*. O azimute ϕ é o ângulo formado pela projeção da reta (que representa a direção)

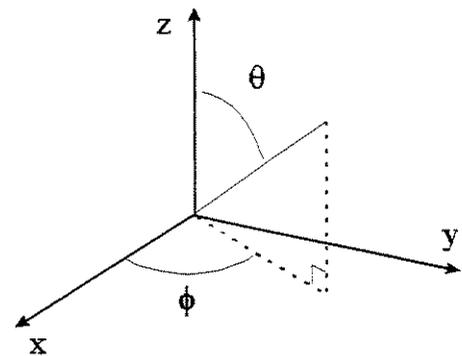


Figura 2-4 ângulo polar e azimute

sobre o plano xy em relação ao eixo x . O eixo x será determinado pela microtextura da superfície. Se não houver microtextura orientada (como pequenas escamas) na superfície, então pode-se fixar o eixo x sobre a projeção do raio incidente no plano xy .

Sejam (θ_i, ϕ_i) a direção do raio incidente e (θ_e, ϕ_e) a direção do observador. Define-se a função de distribuição de refletância bidirecional (BRDF) como

$$f(\theta_i, \phi_i; \theta_e, \phi_e) = \frac{\Delta L(\theta_e, \phi_e)}{\Delta E(\theta_i, \phi_i)}$$

Uma propriedade importante para esta função é a condição de reciprocidade de Helmholtz:

$$f(\theta_i, \phi_i; \theta_e, \phi_e) = f(\theta_e, \phi_e; \theta_i, \phi_i).$$

É devido a essa propriedade que se utiliza o termo bidirecional na denominação desta função.

Na ausência de microtextura orientada, a BRDF pode ser simplificada na forma de uma função de 3 variáveis escrita como $f(\theta_i, \theta_e, \phi_e - \phi_i)$.

A radiância da superfície é relacionada com a irradiância e a função de reflexão através da integral definida sobre um ângulo sólido:

$$L_s(\theta_e, \phi_e) = \int_{\pi}^{\pi} \int_0^{\pi/2} f(\theta_e, \phi_e; \theta_i, \phi_i) E_s(\theta_i, \phi_i) \sin \theta_i \cos \theta_i d\theta_i d\phi_i,$$

onde o termo $\cos \theta_i$ é devido ao *foreshortening*. Esta integral determina a radiância numa direção computando a irradiância em todas as direções de incidência.

Suponha que E_s , em um ponto da superfície, seja devido a um raio de luz apenas (fonte pontual). A fim de simplificar o resultado, define-se $E_s^{L_f}$ da forma

$$E_s^{L_f}(\theta_i, \phi_i) = E_s^{TOTAL} \frac{\delta(\theta_i - \theta_s) \delta(\phi_i - \phi_s)}{\sin \theta_s},$$

onde δ é o impulso ou função delta de Dirak. E_s^{TOTAL} , ou simplesmente E_s , é a irradiância total no ponto da superfície, resultado da integração da equação acima no hemisfério. $E_s^{L_f}$ é o componente da irradiância na direção considerada e é igual a E_s na direção de incidência e zero em todas as outras direções. O uso da função delta de Dirak permite eliminar a integral da BRDF, gerando o resultado:

$$\frac{L_s}{E_s} = f(\theta_s, \phi_s; \theta_e, \phi_e) \cos \theta_s$$

♦ **Relação do Modelo de Fonte de Luz com a Irradiância E_s na Superfície do Objeto**

Uma fonte de luz pontual isotrópica de intensidade I_0 gera uma irradiância $E_s = \frac{I_0 \cos \theta_s}{4\pi r^2}$ num ponto do objeto. O termo $\cos \theta_s$, devido ao *foreshortening*, já foi considerado no equacionamento da superfície no item anterior e o termo $\frac{1}{4\pi r^2}$ é devido à área da casca esférica de raio r , onde r é a distância da fonte de luz ao ponto da superfície.

2.2 A Lei de Reflexão de Lambert

A reflexão lambertiana ideal ocorre quando a radiância é isotrópica e toda luz incidente é refletida, sem absorção. Dessa forma, a BRDF é constante e pode ser demonstrado que essa constante vale $\frac{1}{\pi}$.

Calculando-se a integral do produto da BRDF pela irradiância devido à fonte de luz pontual na direção (θ_s, ϕ_s) obtém-se

$$L_s = \frac{1}{\pi} E_s \cos \theta_s.$$

Esta é a lei de reflexão de Lambert.

A reflexão lambertiana é equivalente à reflexão difusa do modelo de Phong, que será vista no item 4.3.2.

2.3 Mapa de Refletância

O mapa de refletância é um mecanismo mais direto para relacionar a direção normal à superfície com a intensidade do pixel correspondente na imagem, facilitando o tratamento dos problemas de reconstrução. O espaço gradiente é o domínio para o caso mais comum de mapa de refletância, pois constitui-se numa representação mais adequada para a direção normal à superfície.

♦ Espaço Gradiente

Uma maneira de se representar a direção normal à superfície é representá-la no domínio definido pela esfera gaussiana (esfera com centro na origem do sistema de referência, raio igual a 1), isto é, utilizando um versor (vetor unitário) ao qual se denominará \hat{n} . Dado que o versor é unitário, existe portanto um grau de liberdade a menos na definição do vetor. Portanto, desconsiderando-se ambigüidades do tipo +/-, é possível representar a direção utilizando duas variáveis. Por exemplo, pode-se utilizar ângulo polar (simbolizado por θ ou σ) e azimute (simbolizado por ϕ ou τ) para se representar uma direção ao invés dos três valores do vetor normal.

Considere uma superfície S definida pelo mapa de profundidade contínuo $h(x, y)$. A equação da superfície pode ser escrita na forma implícita como $h(x, y) - z = 0$. A direção da normal à superfície no ponto $(x_0, y_0, h(x_0, y_0))$ é paralela ao gradiente da função $h(x, y) - z$, que é representado por $\left(\frac{\partial h(x_0, y_0)}{\partial x}, \frac{\partial h(x_0, y_0)}{\partial y}, -1 \right)$. Para demonstrar isso basta determinar o plano tangente.

O plano tangente à superfície S no ponto $(x_0, y_0, h(x_0, y_0))$ é representado por

$$z = px + qy + k, \text{ onde}$$

$$p = \frac{\partial h}{\partial x}(x_0, y_0) \text{ e}$$

$$q = \frac{\partial h}{\partial y}(x_0, y_0).$$

Considerando a superfície e o plano tangente na forma implícita:

$$F(x, y, z) = h(x, y) - z = 0 \text{ (superfície) e}$$

$$T(x, y, z) = px + qy - z + k = 0 \text{ (plano tangente),}$$

o gradiente das duas funções é igual no ponto de tangência, portanto afirma-se que

$$\vec{n} = \vec{\nabla}F = \vec{\nabla}T = (p, q, -1).$$

Assim, a direção normal à superfície pode ser representada por um par de valores (p, q) .

A fim de simplificar a notação, será utilizado $z(x, y)$ ao invés de $h(x, y)$ para se representar o mapa de profundidade, visto que $z = h(x, y)$ somente em todo ponto da

superfície. Além disso, apenas o ponto visível de cada reta paralela ao eixo z é considerado, que pode-se convencionar como aquele de menor profundidade. Assim a notação utilizada será

$$p = z_x = \frac{\partial z}{\partial x}(x_0, y_0)$$

$$q = z_y = \frac{\partial z}{\partial y}(x_0, y_0)$$

Nota-se que o gradiente de $z(x, y)$, representado pelo vetor bidimensional igual a (p, q) tem uma interpretação diferente do gradiente da função $F(x, y, z) = h(x, y) - z$. Este último vetor gradiente é tridimensional e é dado por $(p, q, -1)$. O vetor (p, q) será perpendicular a toda curva de nível $z(x, y) = c$. O vetor $(p, q, -1)$ será normal a toda superfície de nível $F(x, y, z) = c$, dentre as quais está S para $c = 0$. O espaço gradiente é definido por todo (p, q) e representa a direção de $(p, q, -1)$.

Observa-se que o espaço gradiente, assim definido, é capaz de representar apenas um hemisfério de direções. Para ângulo polar $\theta = 90^\circ$, o vetor (p, q) tende ao infinito. Este sistema de coordenadas, conhecido como projeção gnômica, é muito comum em algoritmos de Shape from Shading. Porém, quando se faz necessário representar $\theta = 90^\circ$ para, por exemplo, definir as condições de borda, uma outra representação, conhecida como projeção estereográfica, pode ser empregada. Essa projeção será abordada junto com o algoritmo variacional de Horn, no item 3.1.1.

♦ Definição de Mapa de Refletância

Mapa de refletância é uma função definida no domínio do espaço gradiente e seu valor correspondente é o valor da irradiância da imagem decorrente da direção da reta normal à superfície, para uma fonte de luz de direção fixa. Em geral, o contra-domínio é definido por um intervalo fechado de 0 a 1.

Para superfícies lambertianas, o mapa de refletância é dado por

$$R(\vec{N}) = \frac{\vec{N} \cdot \vec{L}}{\|\vec{N}\| \cdot \|\vec{L}\|}, \quad (2.1)$$

onde $R(\vec{N})$, o mapa de refletância, é a intensidade de luz refletida na direção do observador

se a normal à superfície for $\vec{N} = (n_1, n_2, n_3)$ e \vec{L} for a direção da fonte de luz. Supõe-se a fonte de luz direcional e, portanto, \vec{L} constante em todo ponto da superfície.

Sejam \vec{n} e \vec{l} , respectivamente, as projeções gnômicas das direções de \vec{N} e \vec{L} , dados pelas equações

$$\vec{n} = (p, q, -1)$$

$$\vec{l} = (p_s, q_s, -1)$$

$$\vec{n} \cdot \vec{l} = pp_s + qq_s + 1$$

$$\|\vec{n}\| = \sqrt{1 + p^2 + q^2}$$

$$\|\vec{l}\| = \sqrt{1 + p_s^2 + q_s^2}$$

e, portanto,

$$R(p, q) = \frac{1 + pp_s + qq_s}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}} \quad \text{no caso de fonte de luz direcional, superfície}$$

lambertiana e projeção ortográfica (centro de projeção no infinito). A figura 2-5 é um gráfico de curvas de nível de uma mapa de refletância de acordo com essa equação.

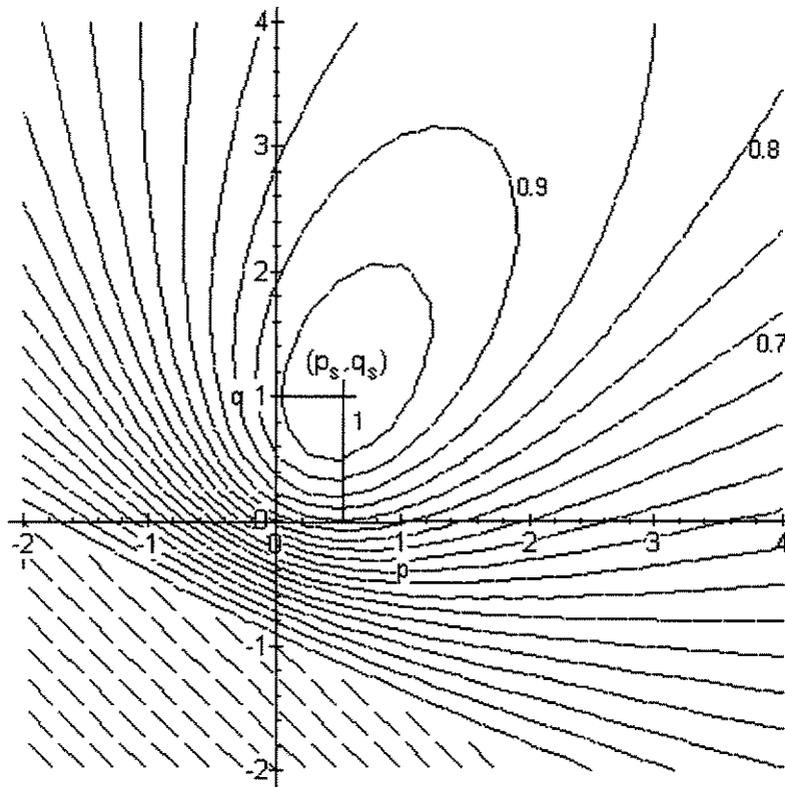


Figura 2-5 Exemplo de mapa de refletância

Quando a direção da fonte de luz coincide com a direção do observador, $p_s = q_s = 0$, o mapa de refletância é escrito como $R(p, q) = \frac{1}{\sqrt{1 + p^2 + q^2}}$, que é rotacionalmente simétrico (as curvas de nível são circunferências concêntricas). Se a normal à superfície for representada no sistema de coordenadas da fonte de luz, esta fórmula rotacionalmente simétrica também poderá ser utilizada para qualquer direção da fonte de luz. Veja [Lee e Rosenfeld 1985] e [Kimmel e Bruckstein 1995b].

2.4 A Equação da Irradiância da Imagem

Tendo-se o mapa de profundidade $z(x, y)$, pode-se calcular o vetor normal à superfície em todo ponto. Conhecido o mapa de refletância $R(p, q)$, pode-se aplicá-lo para obter a intensidade de luz que chega a uma ponto da imagem. Considerando projeção ortográfica, pode-se definir a imagem no mesmo domínio do mapa de profundidade. Define-se, então,

$E(x, y)$ com valores no intervalo fechado de 0 a 1, supondo superfície com albedo homogêneo e que a irradiância máxima é 1.

A equação da irradiância da imagem, que descreve como a imagem é formada, é escrita como

$$E(x, y) = R(p, q), \quad (2.2)$$

ou, também,

$$E(x, y) = R\left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right). \quad (2.3)$$

Quando se conhece $E(x, y)$ e não se conhece $z(x, y)$, tem-se um problema de reconstrução tridimensional, que pode ser visto como um problema de inversão do mapa de refletância ou um sistema de equações diferenciais parciais geralmente não-linear. É essa a formulação matemática para o problema de Shape from Shading.

2.4.1 Ambigüidade

A ambigüidade encontrada nos problemas de Shape from Shading é a incapacidade de se determinar se uma superfície é côncava ou convexa. No caso de projeção ortográfica e luz distante, a imagem gerada por um cone e por um disco com a mesma inclinação será a mesma. Um hemisfério côncavo e um hemisfério convexo produzem a mesma imagem. Observe que a inclusão do valor da normal nas bordas resolve a ambigüidade. Porém, basta a informação de que o ponto singular do hemisfério é côncavo (ou convexo) para se resolver a ambigüidade. Assim, tanto a informação da borda quanto a informação do ponto singular são suficientes para resolver a ambigüidade no Shape from Shading. Alguns exemplos estão ilustrados na figura 2-6. Veja [Pentland 1984] para um exemplo interessante em que não é possível distinguir numa imagem se se trata de uma cratera ou de uma montanha.

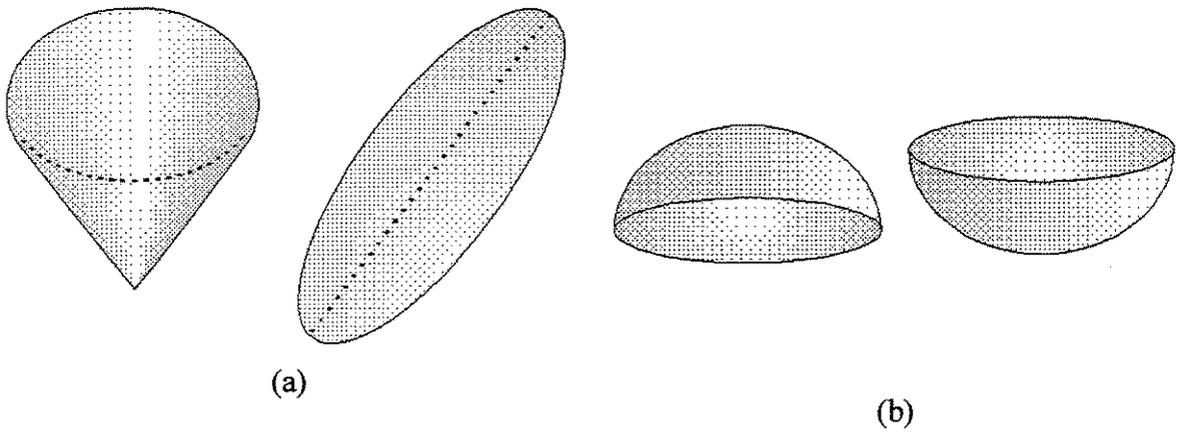


Figura 2-6 Ambigüidade: pares de objetos que resultam na mesma imagem para projeção ortográfica e fonte de luz direcional. (a) Cone e disco cujo ângulo entre a normal à superfície e a direção da fonte de luz é constante. (b) Hemisfério côncavo e convexo.

Capítulo 3 - ALGORITMOS DE SHAPE FROM SHADING

O objetivo desse capítulo é introduzir os principais algoritmos de Shape from Shading encontrados na literatura. Algumas das soluções apresentadas são consideradas para o desenvolvimento dos algoritmos utilizados neste trabalho. Referências a outras soluções serão feitas tendo em vista informar o leitor. Alguns métodos menos importantes estão descritos no apêndice.

Os algoritmos de Shape from Shading serão apresentados seguindo uma taxonomia simples onde o critério de classificação é o tipo de solução empregada. Os métodos mais referenciados podem ser classificados em métodos baseados em otimização, em expansão de curvas e métodos de análise local.

métodos por otimização	método variacional de Horn método de Frankot e Chellappa método de Zheng e Chellappa métodos de gradiente
métodos por expansão de curvas	método clássico de Horn método de Kimmel e Bruckstein
métodos por análise local	método de Pentland método de Lee e Rosenfeld

Tabela 3-1 Métodos de Shape From Shading

Entre os métodos baseados em otimização, serão discutidos o método variacional de Horn [1990] e variações desse método, a saber, o método de Frankot e Chellappa [1988], que trata a questão da integrabilidade, o método de Zheng e Chellappa [1991], mais estável e eficiente e métodos baseados em gradiente, estes últimos mais próximos da abordagem utilizada para desenvolver o algoritmo utilizado neste projeto. Outras variações baseadas em otimização serão comentadas, incluindo o método de Lee e Kuo [1994] e o método de Leclerc e Bobbick [1992].

Os principais métodos baseados em expansão de curvas são o método clássico de Horn [1975] e o método de Kimmel e Bruckstein [1995b], em que se baseia o algoritmo de Okatani e Deguchi [1997] para imagens de endoscópio. O algoritmo de Okatani e Deguchi será apresentado no item 5.1.1.

Além desses métodos, será tratada a formulação de Oliensis e Dupuis [1994], que obtiveram uma solução analítica para o problema de Shape from Shading baseada em soluções de viscosidade para equações hiperbólicas. A solução tem a forma de uma transformada da distância ponderada. O método de Kimmel e Bruckstein [1995b], por exemplo, pode ser derivado dessa formulação. Será apresentada, também, uma extensão para o caso global proposta por Kimmel e Bruckstein [1995a].

Alguns dos métodos de análise local de sombreamento estão descritos no apêndice como informação adicional. São métodos rápidos, aproximados e robustos em relação à imprecisão do modelo, tendo, assim, objetivos distintos dos métodos de Shape from Shading por otimização e por expansão de curvas.

3.1 Shape from Shading por Otimização

Nesta seção, será apresentado o algoritmo variacional de Horn [1990] e variações deste desenvolvidas para tratar o problema de Shape from Shading através da minimização de uma função energia. Serão discutidos aspectos específicos do Shape from Shading, como integrabilidade e problemas inversos.

Para se resolver o problema de Shape from Shading utilizando otimização, é necessário definir uma função de energia (função objetivo) a ser minimizada. Essa função de energia, com domínio definido no espaço do estado do sistema (espaço de busca), deve ser escolhida de forma que o estado do sistema procurado (solução) fique no mínimo global. O

estado do sistema procurado é a solução do sistema de equações diferenciais parciais dado pela equação da irradiância da imagem.

Esses métodos apresentam uma série de desvantagens. As principais são a convergência lenta e a possível convergência para mínimos locais ao invés do mínimo global correspondente à solução procurada. Por outro lado, apresentam como vantagem, o fato de que quando convergem para a solução, o erro é muito pequeno. Além disso, tratam a imagem como um todo (método global) apresentando imunidade ao ruído. A implementação é bastante cômoda e permite incluir outros parâmetros no espaço de busca, como, por exemplo, a determinação automática da direção da fonte de luz [Brooks e Horn 1985] e de parâmetros de calibração de câmera [Hasegawa e Tozzi 1996]. Restrições à solução também são facilmente acopláveis ao algoritmo, somando-se uma parcela a mais à função objetivo, ou empregando-se um multiplicador de Lagrange.

Seguindo as definições dadas por [Tarantolla 1987], considere um espaço de modelo M e um espaço de observações D , que, em geral, são *manifolds*, isto é, espaços topológicos conexos e localmente euclidianos. Seja $T : M \rightarrow D$ um mapeamento do espaço M ao espaço D . Para $d \in D; m \in M$, considere a equação $d = T(m)$. O conhecimento de m e T e desconhecimento de d é um problema de realização do sistema T . O conhecimento de m e d , e desconhecimento de T é o que se chama de identificação de sistemas. O problema inverso é todo problema em que se conheça d e T , e seja necessário conhecer m . No caso geral, podem existir infinitas soluções ou nenhuma solução para o problema inverso. Um método genérico de se resolver esse tipo de problema consiste do seguinte: considere \hat{A} o conjunto de todos os subconjuntos de um conjunto A . Tomando-se $A = M \times D$, pode-se definir uma medida sobre \hat{A} que seja máxima (por exemplo, quadrados mínimos) quando $d = T(m)$ e aplicar algum algoritmo de otimização (por exemplo, método do gradiente) a fim de encontrar a solução do problema inverso.

Observa-se que o Shape from Shading se enquadra na classe de problemas inversos. A obtenção da imagem a partir da geometria e do mapa de refletância é o problema de realização do modelo, que é bem mais fácil de ser resolvido. A obtenção do mapa de refletância de um objeto de geometria conhecida a partir de sua imagem é o problema de identificação de sistemas. A solução do Shape from Shading com o uso de algoritmos de otimização é uma instância da solução geral de problemas inversos.

3.1.1 O Método Variacional de Horn

Os métodos de Horn apresentados aqui podem ser considerados métodos clássicos, porque foi através deles que a técnica de Shape from Shading foi introduzida. Devido à dificuldade de resolver o sistema de equações diferenciais parciais, foram procurados métodos alternativos de solução, baseados em expansão de curvas e Cálculo de Variações. O termo *ill-posed* é normalmente utilizado para qualificar o problema de Shape from Shading, devido à instabilidade dos esquemas elaborados para resolver o problema. Daum e Dudek [1998] chamam de *ill-posed* os problemas em que pequenas alterações na entrada tem grande repercussão na solução. Descrições detalhadas do método variacional de Horn são apresentadas em [Horn 1990], [Ikeuchi e Horn 1981] e [Brooks e Horn 1986].

♦ Projeção Estereográfica

Ikeuchi e Horn [1981] buscaram a solução do problema de Shape from Shading utilizando o mapa de refletância em função do espaço gradiente. Porém, o sistema de coordenadas (p, q) não se mostrou adequado quando as condições de contorno definidas na borda de oclusão do objeto passaram a ser utilizadas porque essas coordenadas podem vir a assumir valor infinito. Assim, Ikeuchi e Horn propuseram utilizar o sistema de coordenadas estereográficas.

Como mostrado na figura 3-1, o par (p, q) pode ser interpretado como a intersecção da reta que contém o ponto definido pela normal unitária (n_1, n_2, n_3) e a origem do espaço com o plano $z = -1$. Esta projeção é conhecida como projeção gnômica. A projeção estereográfica consiste da intersecção da reta que contém o ponto definido pela normal unitária (n_1, n_2, n_3) e o ponto $(0, 0, 1)$ com o plano $z = -1$. No caso da projeção gnômica, se $n_3 = 0$, isto é, o ângulo polar $\theta = 90^\circ$, a projeção só será possível se (p, q) tender ao infinito. No caso da projeção estereográfica, a linha do equador é projetada numa circunferência de raio 2. Assim, é possível se representar o equador evitando problemas numéricos. Normalmente, se representa a projeção estereográfica por (f, g) .

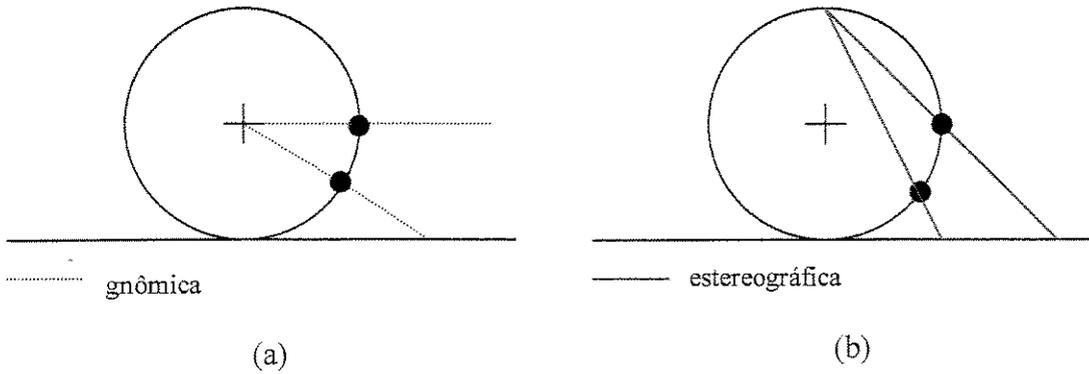


Figura 3-1 Projeção gnômica (a) e projeção estereográfica (b). Para um caso normal e para o caso em que $\theta \rightarrow 90^\circ$, onde $(p, q) \rightarrow \infty$ e $f^2 + g^2 \rightarrow 4$.

A conversão de coordenadas gnômicas para estereográficas é dada pelo par de equações (3.1):

$$f = \frac{2p}{1 + \sqrt{1 + p^2 + q^2}} \quad g = \frac{2q}{1 + \sqrt{1 + p^2 + q^2}} \quad (3.1)$$

A conversão inversa (estereográficas para gnômicas) é dada pelas equações (3.2):

$$p = \frac{4f}{4 - f^2 - g^2} \quad q = \frac{4g}{4 - f^2 - g^2} \quad (3.2)$$

Nota-se que o uso de coordenadas estereográficas não traz, necessariamente, vantagens. A maior parte das vezes, é mais aconselhável o uso de coordenadas gnômicas ou a própria direção do vetor em \mathbb{R}^3 .

♦ **Escolha da Função Objetivo**

Partindo-se da equação da irradiância da imagem em coordenadas estereográficas

$$E(x, y) = Rs(f, g),$$

onde R_s é o mapa de refletância calculado para parâmetros (f, g) , define-se o erro em intensidade da imagem como

$$e_r = \iint_{\Omega} (E(x, y) - Rs(f(x, y), g(x, y)))^2 dx dy, \quad (3.3)$$

onde Ω é a região de suporte da imagem.

Aparentemente, basta minimizar o erro e_r e a equação da irradiância da imagem estará resolvida. Porém, observa-se que a solução não é única. Existem infinitas funções f e g para as quais a aplicação do mapa de refletância resultará na imagem E . Isto acontece se for permitido escolher funções f e g descontínuas. Essas soluções são chamadas soluções fracas. Dentre as soluções fracas é necessário escolher aquela que é mais plausível.

A opção de Horn [1990] foi utilizar um termo de erro que considerasse a suavidade da superfície (e conseqüentemente a continuidade). Esse termo é

$$e_s = \iint_{\Omega} (f_x^2 + f_y^2 + g_x^2 + g_y^2) dx dy, \quad (3.4)$$

onde f_x é a derivada parcial de f em relação a x .

A idéia é utilizar o termo de suavidade como uma parcela da função objetivo. A técnica de adicionar um termo de suavidade a uma função objetivo é conhecida como regularização. O objetivo da regularização é escolher a solução fraca mais próxima da solução desejada, em geral, considerando-se a suavidade da solução. Com a inclusão do novo termo, o problema de otimização passa a ter múltiplos objetivos e portanto, um fator λ ponderar a relevância da parcela de regularização. Esse fator λ pode ser visto como o inverso da confiança na medida da imagem. A função objetivo total é escrita como

$$e = e_r + \lambda e_s. \quad (3.5)$$

Nota-se, porém, que o termo de suavidade contribuirá para uma elevação do erro de reconstrução e_r , no caso em que o valor de λ não for adequadamente arbitrado. A regularização também auxilia a convergência de esquema iterativos, neste caso costuma-se chamar o termo adicional de estabilizador. A técnica de regularização é muito utilizada em problemas inversos e aproximação de funções, podendo-se encontrar muitas referências em textos a respeito dessas teorias. Veja [Poggio *et al.* 1985] para a aplicação de regularização em diversos problemas de Visão Computacional.

Além da restrição de suavidade, pode-se utilizar uma restrição de integrabilidade para garantir que existirá uma função potencial ao integrar-se f e g . Esta é uma alternativa para escolher uma solução fraca mais adequada.

♦ **Minimização**

A minimização do erro definido na equação (3.5) corresponde ao problema de minimizar o funcional

$$\iint F(f, g, f_x, f_y, g_x, g_y) dx dy.$$

A solução em que esse funcional é mínimo é dada pelas equações de Euler-Lagrange:

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

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

onde F_f é a derivada parcial de F em relação a f , F_{f_x} é a derivada parcial de F em relação a f_x e assim por diante.

Como, neste caso,

$$F = \lambda(f_x^2 + f_y^2 + g_x^2 + g_y^2) + (E(x, y) - Rs(f, g))^2,$$

as equações de Euler-Lagrange para resolver o problema são

$$\begin{aligned} (E - Rs)Rs_f + \lambda \nabla^2 f &= 0 \\ (E - Rs)Rs_g + \lambda \nabla^2 g &= 0 \end{aligned} \tag{3.6}$$

♦ **Esquema Iterativo**

A fim de se formar um esquema iterativo, é necessário discretizar o sistema de equações (3.6). A aproximação por diferenças finitas do laplaciano, considerando ε a distância entre pixels vizinhos é dada por

$$\{\nabla^2 f\}_{ij} \approx \frac{4}{\varepsilon^2}(\bar{f}_{ij} - f_{ij}), \text{ onde}$$

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

Esta aproximação é obtida aplicando a diferença central dos vizinhos, da seguinte forma

$$\frac{\Delta f}{\Delta x} \approx \frac{f\left(x + \frac{\varepsilon}{2}, y\right) - f\left(x - \frac{\varepsilon}{2}, y\right)}{\varepsilon}$$

$$\frac{\Delta^2 f}{\Delta x^2} \approx \frac{\frac{f(x+\varepsilon, y) - f(x, y)}{\varepsilon} - \frac{f(x, y) - f(x-\varepsilon, y)}{\varepsilon}}{\varepsilon} = \frac{f(x+\varepsilon, y) + f(x-\varepsilon, y) - 2f(x, y)}{\varepsilon^2}$$

$$\frac{\Delta^2 f}{\Delta x^2} + \frac{\Delta^2 f}{\Delta y^2} \approx \frac{f(x+\varepsilon, y) + f(x-\varepsilon, y) + f(x, y+\varepsilon) + f(x, y-\varepsilon) - 4f(x, y)}{\varepsilon^2}.$$

Substituindo-se esta aproximação por diferenças finitas na equação (3.6) o seguinte esquema iterativo é obtido

$$\begin{aligned} f_{ij}^{k+1} &= \bar{f}_{ij}^k + \frac{\varepsilon^2}{4\lambda} (E_{ij} - Rs(f_{ij}^k, g_{ij}^k)) Rs_f(f_{ij}^k, g_{ij}^k) \\ g_{ij}^{k+1} &= \bar{g}_{ij}^k + \frac{\varepsilon^2}{4\lambda} (E_{ij} - Rs(f_{ij}^k, g_{ij}^k)) Rs_g(f_{ij}^k, g_{ij}^k) \end{aligned} \quad (3.7)$$

Ikeuchi e Horn [1981] fazem as seguintes sugestões quanto à discretização do laplaciano e a convergência do esquema iterativo:

1. Utilizar $\bar{f}_{ij} = \frac{1}{5}(f_{ij+1} + f_{i+1j} + f_{ij-1} + f_{i-1j}) + \frac{1}{20}(f_{i+1j+1} + f_{i+1j-1} + f_{i-1j-1} + f_{i-1j+1})$, que é mais precisa do que a aproximação acima;
2. Empregar o esquema:

$$\begin{aligned} f_{ij}^{k+1} &= \bar{f}_{ij}^k + \frac{\varepsilon^2}{4\lambda} (E_{ij} - Rs(\bar{f}_{ij}^k, \bar{g}_{ij}^k)) Rs_f(\bar{f}_{ij}^k, \bar{g}_{ij}^k) \\ g_{ij}^{k+1} &= \bar{g}_{ij}^k + \frac{\varepsilon^2}{4\lambda} (E_{ij} - Rs(\bar{f}_{ij}^k, \bar{g}_{ij}^k)) Rs_g(\bar{f}_{ij}^k, \bar{g}_{ij}^k) \end{aligned} \quad (3.8)$$

que pode resolver alguns problemas de instabilidade numérica na forma de *checkerboard*, isto é, o erro alterna para mais e para menos em pixels adjacentes.

♦ Aplicação em Estéreo Fotometria

Horn [1990] generalizou seu algoritmo para o caso de várias imagens da mesma cena em condições de iluminação (mapas de refletância) diferentes. Criou-se, assim, um algoritmo genérico de Estéreo Fotometria [Woodham 1980] para um número qualquer de imagens.

Considerando-se n imagens E_i e os n mapas de refletância R_i respectivos, a função objetivo a ser minimizada é

$$e = \iint_{\Omega} (f_x^2 + f_y^2 + g_x^2 + g_y^2) dx dy + \sum_{i=1}^n \alpha_i \iint_{\Omega} (E_i(x, y) - R_i(f(x, y), g(x, y)))^2 dx dy, \text{ onde}$$

α_i são parâmetros que determinam a confiança que se tem na medida de intensidade da imagem E_i . Quanto menos ruidosa for a imagem, maior poderá ser α_i .

♦ Integrabilidade

O resultado do algoritmo global de Shape from Shading por otimização de Horn é um campo vetorial (função que tem como contra-domínio um espaço vetorial) que representa o gradiente do mapa de profundidade. Não é todo campo vetorial (vetores bidimensionais, neste caso) que pode ser integrado e se obter o que se chama de função potencial do campo vetorial. A função potencial só vai existir se a integral de linha em todo caminho fechado for nula (como no trabalho de uma força conservativa, não há perda de energia). Para cada região Γ da imagem, tem-se

$$\oint_{\Gamma} (p dx + q dy) = 0, \text{ se o campo vetorial for Riemann-integrável.}$$

De acordo com o teorema de Green, uma condição equivalente é

$$z_{xy} = z_{yx},$$

ou ainda,

$$p_y = q_x. \tag{3.9}$$

♦ Integração por Otimização

A solução de Ikeuchi e Horn [1981] para integrar o campo vetorial obtido para o gradiente do mapa de profundidade foi utilizar a mesma estratégia empregada na solução do Shape from Shading: transformar o problema usando o Cálculo Variacional e quadrados mínimos para fazer a integração. Definiu-se um funcional a ser minimizado em função da estimativa de $z(x, y)$. A minimização da função objetivo $\iint_{\Omega} (z_x - p)^2 + (z_y - q)^2 dx dy$,

resulta na equação de Euler-Lagrange

$$\nabla^2 z = p_x + q_y.$$

Um esquema iterativo para resolver essa equação é

$$z_{ij}^{k+1} = \bar{z}_{ij}^k - \frac{\varepsilon}{4} \left(\frac{p_{i+1j} - p_{i-1j}}{2} + \frac{q_{ij+1} - q_{ij-1}}{2} \right). \quad (3.10)$$

Por se tratar de um sistema bem comportado, com solução única, não se faz necessário utilizar o termo de regularização.

♦ Restrição de Integrabilidade na Função Objetivo

A incorporação da restrição de integrabilidade no Shape from Shading por otimização permitirá a escolha de uma solução integrável dentre as soluções fracas. Espera-se da solução integrável que se possa integrar por acumulação, pois por qualquer caminho que se faça a integração, deve-se ter o mesmo resultado.

A formulação de Horn e Brooks [1989] para se incluir o termo de integrabilidade no algoritmo de Shape from Shading é a utilização do erro de integração:

$$e_{INT} = \iint_{\Omega} \left[(E(x, y) - R(p(x, y), q(x, y)))^2 + \mu(p_y - q_x)^2 \right] dx dy. \quad (3.11)$$

O coeficiente μ pode ser um multiplicador de Lagrange ou uma constante, a ser arbitrada, que multiplica o termo de penalidade. O termo de suavidade pode ser empregado, mas com menor importância. Segundo Frankot e Chellappa [1988], há dificuldade para se obter um esquema que apresente convergência, baseado na formulação de Horn e Brooks. Veja em [Frankot e Chellappa 1988], ou no item 3.1.2, uma solução para a questão da integrabilidade em que se projeta a cada iteração a solução estimada no espaço das soluções integráveis.

♦ Algumas Considerações quanto ao Método de Horn

Para perfeito funcionamento do método de Horn, são necessárias condições de contorno devido, por exemplo, à questão da ambigüidade da equação da irradiância da imagem, abordada no item 2.4.1. Essas condições devem ser apresentadas na forma de uma curva em que se conheça a normal à superfície, em geral, se utiliza a borda de oclusão do objeto onde a normal à superfície é perpendicular ao eixo óptico. Lee [1988] demonstrou a existência de um algoritmo, semelhante ao algoritmo de Horn, que converge dadas condições de contorno, porém sabe-se que sua solução não é integrável [Frankot e Chellappa 1988]. Por

outro lado, Oliensis e Dupuis [1994] demonstraram que a informação de borda é menos importante para a solução do problema de Shape from Shading que os pontos singulares.

Os problemas com o método de Horn são decorrentes da necessidade de se resolver o problema de otimização de forma global. A convergência para a solução correta não é garantida, podendo convergir para um mínimo local além de necessitar de muitas iterações para convergir. Sem as condições de borda, a convergência se torna ainda mais difícil. O termo de regularização, se for escolhido de forma não apropriada, estabiliza o sistema ao mesmo tempo que afasta o resultado da solução correta. Além disso, como já foi dito, é necessária uma etapa de integração após o término do algoritmo, pois o que se recuperaram foram as orientações em cada ponto da imagem. A etapa de integração permite obter, também, a profundidade.

3.1.2 Método de Frankot e Chellappa

Frankot e Chellappa [1988] resolveram a questão da integrabilidade projetando a solução, a cada iteração do algoritmo de Horn, no espaço das soluções integráveis. A técnica utilizada é chamada projeção em espaços convexos, ou POCS. O método foi intitulado “Shape from Shading com Integrabilidade Reforçada”.

Considere z_x a derivada da profundidade na direção x , \hat{z}_x o valor estimado e \tilde{z}_x o valor estimado cujo campo vetorial é integrável.

Deseja-se obter \tilde{z}_x, \tilde{z}_y que minimizem

$$d\left\{\left(\hat{z}_x, \hat{z}_y\right), \left(\tilde{z}_x, \tilde{z}_y\right)\right\} = \iint \left(\tilde{z}_x - \hat{z}_x\right)^2 + \left(\tilde{z}_y - \hat{z}_y\right)^2 dx dy$$

Sujeitos a $\frac{\partial}{\partial y} \tilde{z}_x = \frac{\partial}{\partial x} \tilde{z}_y$, que é a condição de integrabilidade.

Seja uma base ortogonal $\{\phi(x, y, \omega)\}$, onde ω é um índice.

Escreve-se $z(x, y)$ na forma

$$z(x, y) = \sum_{\omega \in \Omega} C(\omega) \phi(x, y, \omega)$$

$$z_x(x, y) = \sum_{\omega \in \Omega} C(\omega) \phi_x(x, y, \omega)$$

$$z_y(x, y) = \sum_{\omega \in \Omega} C(\omega) \phi_y(x, y, \omega)$$

onde $C(\omega)$ é o conjunto de parâmetros.

Em primeiro lugar, aplica-se uma iteração do algoritmo de Horn:

$$\begin{bmatrix} \hat{z}_x \\ \hat{z}_y \end{bmatrix}_{k+1} = \begin{bmatrix} \tilde{z}_x \\ \tilde{z}_y \end{bmatrix}_k + \lambda_1 (E - R) \begin{bmatrix} R_{z_x} \\ R_{z_y} \end{bmatrix}$$

Obtêm-se novos \hat{z}_x e \hat{z}_y , porém o campo vetorial definido por eles pode não ser integrável. O que se deseja é \tilde{z}_x e \tilde{z}_y , que formam um campo vetorial integrável. Para isso, calculam-se \hat{C}_1 e \hat{C}_2 , tais que

$$\hat{z}_x(x, y) = \sum_{\omega \in \Omega} \hat{C}_1(\omega) \phi_x$$

$$\hat{z}_y(x, y) = \sum_{\omega \in \Omega} \hat{C}_2(\omega) \phi_y$$

Minimizando-se o funcional $d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\}$, deseja-se encontrar o conjunto de coeficientes $\tilde{C}(\omega)$. Está demonstrado em [Frankot e Chellappa 1988] que o conjunto de coeficientes desejado é dado por

$$\tilde{C}(\omega) = \frac{P_x(\omega) \hat{C}_1(\omega) + P_y(\omega) \hat{C}_2(\omega)}{P_x(\omega) + P_y(\omega)}, \quad \text{onde}$$

$$P_x(\omega) = \iint (\phi_x(x, y, \omega))^2 dx dy \quad P_y(\omega) = \iint (\phi_y(x, y, \omega))^2 dx dy$$

Finalmente, calculam-se \tilde{z}_x, \tilde{z}_y utilizando \tilde{C} e executa-se nova iteração. Para a base ortogonal, pode-se, por exemplo, utilizar funções de Fourier com implementação através da Transformada Rápida de Fourier, obtendo-se uma implementação bastante eficiente.

3.1.3 O Método de Zheng e Chellappa

Zheng e Chellappa [1991] propuseram um método para diminuir o efeito negativo do termo de suavização da função objetivo. Além de buscar uma solução que satisfaça

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

a solução deve satisfazer também

$$E_x(x, y) = R_p(p, q) p_x + R_q(p, q) q_x \text{ e}$$

$$E_y(x, y) = R_p(p, q) p_y + R_q(p, q) q_y. \quad (3.12)$$

Assim, o funcional a ser minimizado passa a ser:

$$e = (R(p, q) - E(x, y))^2 + (R_p(p, q)p_x + R_q(p, q)q_x - E_x(x, y))^2 + (R_p(p, q)p_y + R_q(p, q)q_y - E_y(x, y))^2 + \mu((p - z_x)^2 + (q - z_y)^2) \quad (3.13)$$

A última parcela da equação (3.13) é um termo de integrabilidade e as variáveis do espaço de busca são p , q e z . Zheng e Chellappa ainda aplicam linearização do mapa de refletância e demonstram a estabilidade do esquema montado. Além dessas técnicas para garantir a robustez da solução, a implementação proposta utiliza múltipla escala, que é uma técnica que estabiliza e acelera a convergência e pode ser aplicada a variáveis definidas sobre uma grade.

3.1.4 Métodos de Gradiente

Um esquema iterativo equivalente ao método de Horn pode ser derivado através de métodos de gradiente. Uma breve descrição é apresentada na seqüência.

O método do gradiente baseia-se em variar o estado atual ao longo do tempo, sempre na direção que minimize um funcional de energia. Seja $F : \mathcal{R}^n \rightarrow \mathcal{R}$ um funcional real a ser minimizado, isto é, deseja-se encontrar o valor de $\mathbf{x} = (x_1, \dots, x_n)$ em que $F(\mathbf{x})$ seja mínima. O método do gradiente consiste em tomar passos no espaço de busca na direção em que $F(\mathbf{x})$ decai mais rápido, isto é, na direção oposta ao gradiente de F em relação a \mathbf{x} . O algoritmo é escrito na forma de uma equação diferencial no tempo, e o valor procurado é o estado do sistema em regime. Essa equação tem a forma

$$\frac{\partial \mathbf{x}}{\partial t} = - \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n} \right).$$

Um esquema iterativo para solução dessa equação é o seguinte

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \Delta t \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n} \right).$$

A convergência depende do valor adotado para o parâmetro de passo de integração Δt . Métodos de ordem mais elevada (Runge-Kutta) podem ser utilizados na abordagem de otimização por gradiente, assim como o método do gradiente conjugado. Veja [Tarantola 1987], [Press *et al.* 1992] e [Strikwerda 1989] para métodos de gradiente.

Szeliski [1991] mostra como o esquema de Horn pode ser obtido através de um método de gradiente, após discretização da função objetivo. Horn obtém primeiro as equações de

Euler-Lagrange e depois aplica discretização. O artigo de Szeliski [1991] contém também bons resumos dos métodos de gradiente e relaxação de Jacobi.

Lai e Vemuri [1997] utilizam o método do gradiente conjugado pré-condicionado para acelerar a convergência do método de Horn. O método é aplicado também a outros problemas de Visão.

O método desenvolvido neste projeto, apresentado no item 5.2, é baseado no método de descida na direção do gradiente.

3.1.5 Outras Variações

♦ Rotulação Densa e Shape from Shading

Malik e Maydan [1989] propuseram o primeiro algoritmo de Shape from Shading por partes, utilizando o processo de otimização para escolher uma rotulação apropriada para as bordas presentes na imagem. Mais sobre rotulação de bordas pode ser visto em [Horn 1990].

♦ Elementos Finitos – Esquema de Lee e Kuo

Lee e Kuo [1994, 1997] resolveram o problema de Shape from Shading através de uma malha (*mesh*) triangular aplicando o método de elementos finitos e linearizando o mapa de refletância.

♦ Height from Shading

Leclerc e Bobbick [1992] utilizaram Shape from Shading para melhorar os resultados obtidos num processo de Shape from Stereo. Segundo [Cryer, Tsai e Shah 1995], Shape from Stereo (método de Barnard) fornece informação de alta frequência porque utiliza características da imagem para fazer o *matching* (correspondência entre os pontos de duas imagens). Por outro lado, o Shape from Shading (método de Pentland baseado em linearização do mapa de refletância) é mais preciso para superfícies suaves e, portanto, a informação de baixa frequência é mais relevante. E é dessa forma que os dois métodos podem ser complementados.

O método de Leclerc e Bobbick [1992] parte da profundidade calculada por Estéreo e minimiza o erro da equação da irradiância da imagem utilizando como espaço de parâmetros o próprio mapa de profundidade, ao invés de suas derivadas.

3.2 Shape from Shading por Expansão de Curvas

Expansão de curvas é um método alternativo de solução de sistemas de equações diferenciais. Partindo-se de uma curva inicial conhecida, propaga-se a informação dessa curva, calculando-se uma nova curva. Uma vez que se conheça curvas que cubram toda a região de suporte, pode se dar como terminado o algoritmo.

É difícil estabelecer um critério para se escolher um método baseado em expansão de curvas porque esse tipo de método pode admitir vários níveis de sofisticação. Horn empregou um método de expansão de curvas na solução do Shape from Shading, mas trata-se de um método local e, por isso, problemático quanto ao ruído na imagem. Já, o esquema elaborado por Osher e Sethian [1988] para solução de expansão de curvas dependente da curvatura trata a imagem como um todo e é empregado no método de Kimmel e Bruckstein [1995b].

O método clássico de Horn de expansão de curvas será deixado para o apêndice, pois não é relevante ao projeto e foi incluído no texto para conhecimento do leitor. Porém, é preciso ressaltar a característica local deste método e a implementação das curvas através de pontos de marca que se movimentam conforme as equações que definem o sistema de expansão de curvas. A curva é reamostrada a cada iteração seguindo uma heurística. Às vezes é necessário o rompimento da curva, que era inicialmente fechada, devido a problemas da topologia. Assim, o método fica precário quando se afasta muito da curva inicial.

3.2.1 O Método de Kimmel e Bruckstein

A maioria dos problemas encontrados na solução por curvas características se resolve através do método de Kimmel e Bruckstein [1995b]. A equivalência desta formulação com a solução pela transformada da distância ponderada (veja item 3.2.2) garante a unicidade da solução e a convergência para a solução correta. Além disso, este método apresenta uma interpretação bastante intuitiva. Veja também [Kimmel 1995].

♦ Propriedades de Curvas de Isoprofundidade

Será tratado o caso de superfície lambertiana, projeção ortográfica e fonte de luz direcional na direção do eixo óptico. A formulação de Kimmel e Bruckstein [1995b] utiliza curvas de isoprofundidade. Nota-se que o azimute da direção da normal à superfície, se forem desconsideradas ambigüidades, é conhecido ao longo dessas curvas, porque correspondem à direção do vetor normal à curva, que é paralelo à projeção da normal à superfície no plano da curva. Observa-se, também, que se conhece o ângulo polar por toda a superfície porque este pode ser obtido diretamente da imagem através da Lei de Lambert, colocada da forma $E = \rho \cos \sigma$, onde E é a intensidade do pixel, σ é o ângulo polar da normal à superfície e ρ é o albedo da superfície, que será considerado constante e igual a 1, sem perda de generalidade.

Uma curva em movimento apresenta em cada ponto velocidade tangencial e velocidade normal à curva. A velocidade tangencial vai depender da parametrização da curva. Porém, a velocidade tangencial pode ser ignorada porque é possível definir uma parametrização tal que esta se anule (veja [Kimmel 1995]). O movimento de uma curva pode ser definido com base na velocidade normal. A notação utilizada para representar a curva é $C(t) = (x(s), y(s))$ para $z = t$, onde s é um parâmetro com valor pertencente a um intervalo de números reais e é mapeado para um ponto da curva.

Para cada ponto da curva, pode-se associar um ponto da nova curva de isoprofundidade, seguindo a direção \hat{n} perpendicular à curva atual. Essa direção é a direção da projeção da normal à superfície \vec{N} no plano da curva. A distância no plano do novo ponto ao ponto atual é determinada pela inclinação do plano tangente e, portanto, do ângulo polar σ da normal. No caso canônico (superfície lambertiana, projeção ortográfica e fonte de luz direcional coincidente com o observador), o ângulo polar σ é obtido diretamente da imagem, usando $E = \cos \sigma$. Suponha que se queira descer Δz , isto é, encontrar uma curva de profundidade $z + \Delta z$, onde z é a profundidade da curva

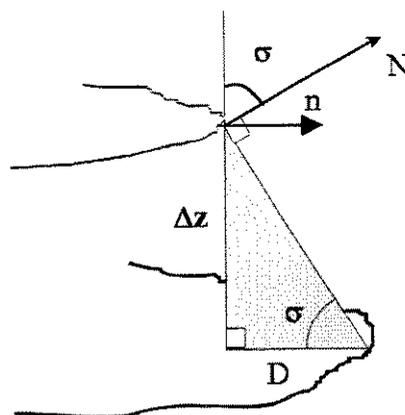


Figura 3-2 Expansão de curvas de isoprofundidade

atual. Utilizando relações trigonométricas num triângulo retângulo, é fácil ver na figura 3-2 que $D = \Delta z \cot \sigma$ é a distância no plano entre o ponto da curva atual e o ponto da próxima curva. Calculando-se diretamente a partir da imagem, o que se tem é

$$D = \Delta z \cot \sigma = \Delta z \frac{\cos \sigma}{\sin \sigma} = \Delta z \frac{\cos \sigma}{\sqrt{1 - \cos^2 \sigma}} = \Delta z \frac{E}{\sqrt{1 - E^2}} \quad (3.14)$$

onde E é a intensidade do pixel pertencente à curva.

• Expansão de Curvas de Isoprofundidade

Será apresentada inicialmente a formulação lagrangeana do método de expansão de curvas. Implementar diretamente um método numérico utilizando essa formulação resultará em muitos problemas de topologia e estabilidade numérica. A sugestão é utilizar a formulação euleriana de Osher e Sethian [1988] que será apresentada no item intitulado “A Forma Euleriana”.

Considere um contorno de igual profundidade $\{x(s), y(s)\}, s \in [0, s]$ conhecido. O parâmetro s varia ao longo da curva

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} = 0 \Rightarrow px_s + qy_s = 0,$$

onde o s subscrito representa a derivada em relação a s .

O vetor unitário normal à curva é dado por

$$\hat{n}(s) = \frac{1}{\sqrt{x_s^2(s) + y_s^2(s)}} [y_s(s), -x_s(s)].$$

Para o caso $\hat{l} = \hat{z} = (0, 0, 1)$, isto é, a direção de iluminação é paralela ao eixo z , $\cos \sigma = \hat{l} \cdot \vec{N}$, onde \vec{N} é o vetor unitário normal à superfície.

Para se avançar dz , deve-se afastar D na direção de \hat{n} tal que

$$D = dz \cot \sigma$$

como calculado anteriormente,

$$D = dz \frac{E}{\sqrt{1 - E^2}}.$$

A propagação do contorno é descrita como

$$\{x(s, dz), y(s, dz)\} = \{x(s, 0), y(s, 0)\} + D(s)\hat{n}(s)$$

$$\begin{cases} x_t(s,t) = \frac{y_s}{\sqrt{x_s^2 + y_s^2}} \frac{E}{\sqrt{1-E^2}} \\ y_t(s,t) = \frac{-x_s}{\sqrt{x_s^2 + y_s^2}} \frac{E}{\sqrt{1-E^2}} \end{cases},$$

onde s e t subscritos indicam diferenciação.

Seja $F = \frac{E}{\sqrt{1-E^2}}$ a velocidade normal à curva, abrevia-se a expansão da curva como

$$C_t = F \cdot \hat{n} \tag{3.15}$$

onde $C_t = \frac{\partial C}{\partial t}$ e $C(t)$ é o contorno de isoprofundidade no instante t . A figura 3-3 ilustra a reconstrução da superfície por expansão das curvas de isoprofundidade $C(t)$.

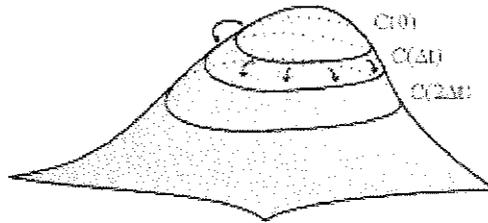


Figura 3-3 Ilustração do mecanismo de expansão de curvas.

♦ **Tratamento da Direção da Fonte de Luz**

Kimmel e Bruckstein [1995b] utilizam o sistema de coordenadas da fonte de luz para tratar fontes de luz não paralelas ao eixo óptico. As curvas de isoprofundidade estão em planos perpendiculares à direção da fonte de luz, que é diferente da direção do observador. Deve-se, portanto, calcular a velocidade normal da curva para este caso específico. Neste caso o sistema será escrito assim:

$$C_t = G(x, y, \hat{n})\hat{n} \tag{3.16}$$

onde

$$G(x, y, \hat{n}) = \frac{1}{\sqrt{1+p_l^2+q_l^2}} \left[F \sqrt{n_1^2(1+q_l^2)+n_2^2(1+p_l^2)-n_1n_2 2p_lq_l} + (p_l n_1 + q_l n_2) \right],$$

$$F = \frac{E}{\sqrt{1-E^2}} \text{ e } \hat{n} = (n_1, n_2).$$

- ♦ **A Forma Euleriana**

A forma euleriana corresponde ao algoritmo numérico de expansão de curvas de Osher e Sethian [1988]. Esta solução é numericamente estável e resolve problemas topológicos. A forma euleriana do algoritmo de Kimmel e Bruckstein [1995b] será mostrada na seqüência.

- ♦ **Discretização da Função Implícita no Plano**

Define-se a função $\phi(x, y, t)$ tal que $\phi(x, y, 0) = 0$ nos pontos (x, y) pertencentes à curva $C(0)$. É a forma implícita da curva. Deseja-se propagar $\phi(x, y, 0) = 0$ para que $\phi(x, y, t)$ corresponda à curva $C(t)$ a cada iteração.

Pela regra da cadeia,

$$\phi_t = \phi_x x_t + \phi_y y_t = (\phi_x, \phi_y) \cdot (x_t, y_t)$$

$$\phi_t = \nabla\phi \cdot C_t(t)$$

Fazendo $\hat{n} = \frac{\nabla\phi}{\|\nabla\phi\|}$, e substituindo $C_t = G \cdot \hat{n}$

$$\phi_t = \nabla\phi \cdot G \cdot \hat{n} = G \cdot \nabla\phi \cdot \frac{\nabla\phi}{\|\nabla\phi\|}.$$

Logo,

$$\phi_t = G \cdot \|\nabla\phi\|,$$

onde $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$ é o vetor gradiente em relação a (x, y) e G é equivalente à velocidade normal à curva.

No caso lambertiano com direção de iluminação $(p_t, q_t, -1)$

$$\phi_t = \frac{1}{\sqrt{1 + p_t^2 + q_t^2}} \left[F(x, y) \sqrt{\phi_x^2 (1 + q_t^2) + \phi_y^2 (1 + p_t^2) - \phi_x \phi_y 2 p_t q_t - (p_t \phi_x + q_t \phi_y)} \right] \quad (3.17)$$

- ♦ **Implementação da Expansão da Função Implícita**

Um ponto fundamental para o correto funcionamento do algoritmo de Osher e Sethian [1988] é a escolha do esquema de discretização. O esquema empregado em [Kimmel e Bruckstein 1995b] é o esquema de Godunov, que escolhe aproximar a derivada pela diferença à esquerda ou à direita. A discretização utilizada é *upwind* e tem como finalidade

preservar a condição de entropia: as novas curvas não podem ocupar o espaço das curvas anteriores. Veja a explicação em [Osher e Sethian 1988]. As expressões abaixo mostram a maneira correta de se realizar a discretização para gerar um esquema iterativo.

$$\begin{aligned}\phi_{ij}^n &= \phi(i\Delta x, j\Delta y, n\Delta t) \\ \phi_t &\approx \frac{\phi^{n+1} - \phi^n}{\Delta t} \\ \phi_x^2 &\approx \max\{D_+^x\phi_{ij}, -D_-^x\phi_{ij}, 0\} \\ D_+^x\phi_{ij} &= \frac{\phi_i - \phi_{i-1}}{\Delta x} \\ D_-^x\phi_{ij} &= \frac{\phi_{i+1} - \phi_i}{\Delta x} \\ \phi_x &\approx \min \text{mod}(D_-^x\phi_{ij}, D_+^x\phi_{ij}) \\ \min \text{mod}(a, b) &= \begin{cases} \text{sgn}(a) \cdot \min(|a|, |b|), & ab > 0 \\ 0 & \text{caso contrário} \end{cases}\end{aligned}$$

A inicialização pode ser feita aplicando *threshold* T para procurar pontos singulares como mostrado na equação

$$\phi(x, y, 0) = E(x, y) - T.$$

Uma outra opção é utilizar um cone gerado por uma pequena circunferência centrada no ponto singular.

A cada iteração, quando $\phi_{ij}^n \cdot \phi_{ij}^{n-1} < 0$, ou seja, a curva atingiu aquele ponto (O ponto passou da região externa para a região interna definidas pela curva, nesta iteração), então

$$\text{height}_{ij} = \left[\Delta t \left(n - \frac{\phi_{ij}^n}{\phi_{ij}^n - \phi_{ij}^{n-1}} \right) \right] \cdot \sqrt{1 + p_l^2 + q_l^2} + i\Delta x p_l + j\Delta y q_l. \quad (3.18)$$

Assim, são atribuídos os valores de profundidade para cada pixel.

3.2.2 Shape from Shading pela Transformada da Distância Ponderada

Recentemente, alguns problemas de Visão vem sendo tratados através de uma abordagem baseada em Controle Ótimo. Veja, por exemplo, [Kimia *et al.* 1994]. Oliensis e Dupuis [1994] demonstraram, utilizando uma abordagem de Controle Ótimo, que o

problema de Shape from Shading é equivalente ao cálculo da Transformada da Distância Ponderada (WDT). Na verdade, esta denominação foi empregada em [Kimmel e Bruckstein 1995a] e [Kimmel 1995], interpretando a expressão encontrada para a solução do Shape from Shading. Considere uma função $G(x, y)$ que relacione um custo para cada ponto da imagem. Considere um conjunto S de pontos da imagem e defina para cada um destes pontos a função de custo terminal $V(x, y)$. O custo de uma trajetória sobre a imagem é definido como a integração da função custo G ao longo da trajetória. O valor da Transformada da Distância Ponderada de um ponto P em relação ao conjunto de pontos S consiste no menor custo total para conectar P a S , através de uma trajetória (curva), considerando o custo total como a soma do custo da trajetória e do custo terminal do ponto de S que é conectado a P . Para isso, é necessário cobrir todas as trajetórias possíveis de P a S , procurando o menor custo.

A solução do problema de Shape from Shading, utilizando esta abordagem é dada pela equação (3.19).

$$z(x, y) = \inf_{l(s) \in L} \left\{ \int_0^{|l|} G(l(s)) ds : \left| \frac{\partial l}{\partial s} \right| = 1, \quad l(0) = (x, y), \quad l(|l|) = (x_0, y_0) \right\}, \quad (3.19)$$

onde L é o conjunto de todas as trajetórias $l(s)$, G é a função custo, $z(x, y)$ é o custo terminal do ponto (x, y) e corresponde ao mapa de profundidade. O ponto (x_0, y_0) é um ponto singular em que se adotou $z(x_0, y_0) = 0$.

Para o caso de Shape from Shading para superfície lambertiana, com projeção ortográfica e fonte de luz direcional perpendicular ao plano imagem, a função custo é dada por

$$G(x, y) = \frac{\sqrt{1 - E^2}}{E}$$

3.2.3 Representação como Problema de Controle Ótimo

Segundo Kimmel [1995], as equações de Hamilton-Jacobi, encontradas em problemas de controle ótimo, têm a seguinte forma, para variável de estado em \mathbb{R}^d :

$$\begin{aligned} \phi_t + H(\phi_{x_1}, \dots, \phi_{x_d}) &= 0 \\ \phi(x, 0) &= \phi_0(x) \end{aligned}$$

Este tipo de equação admite soluções fracas, devendo-se procurar uma solução que seja a mais fisicamente plausível. A solução fisicamente correta que é procurada é chamada solução de viscosidade e é obtida considerando-se as condições de entropia.

A equação da irradiância da imagem pode ser escrita da forma

$$H(x, y, \nabla z) = E(x, y) \sqrt{1 + \|\nabla z\|^2} - 1 = 0. \quad (3.20)$$

onde $H(x, y, \alpha)$ é o hamiltoniano. Para o cálculo da transformada da distância ponderada, é necessário obter o custo $G(x, y)$ para cada pixel. Essa função de custo está relacionada com o lagrangeano, que é obtido a partir do hamiltoniano e vice-versa através da transformação de Legendre. Veja detalhes em [Oliensis e Dupuis 1994].

A existência e a unicidade da solução para o problema de Shape from Shading, como demonstrado por Oliensis e Dupuis [1994] é consequência da unicidade e existência da solução de viscosidade, ou do custo da trajetória de menor custo para cada ponto da imagem.

3.2.4 Algoritmos de Oliensis e Dupuis

Uma primeira aproximação grosseira, para exemplificar o cálculo da transformada da distância ponderada, seria buscar pelas trajetórias compostas de trechos horizontais, verticais e diagonais a 45 graus. Pode-se utilizar um esquema iterativo, inicializando todos os pixels com profundidade máxima e o ponto de máximo com profundidade zero e, a cada iteração, a profundidade de um pixel recebe o valor da menor profundidade dentre a profundidade atual e a profundidade de seus oito vizinhos somadas ao custo do pixel. Se o custo estimado do pixel analisado já tiver atingido o custo ótimo, este valor se estabilizará. Cabe citar aqui o princípio da programação dinâmica, "se uma trajetória é ótima então todo pedaço desta trajetória também é uma trajetória ótima".

Veja em [Oliensis e Dupuis 1994] um esquema iterativo aproximando a trajetória real, através de uma trajetória quantizada utilizando probabilidades.

As condições iniciais para os algoritmos descritos neste item são os pontos singulares, que devem ter profundidade conhecida. Porém, qualquer ponto de profundidade conhecida pode ser utilizado como condição adicional. Para o algoritmo de Kimmel e Bruckstein [1995b], que resolve o problema de Shape from Shading localmente, se existirem mais de

um ponto singular, o problema passa a ser um problema de Shape from Shading Global, que será tratado em 3.2.7.

3.2.5 Equivalência com o Método de Kimmel e Bruckstein

Considerando o custo terminal de um pixel equivalente à profundidade, o método de expansão de curvas de profundidade constante de Kimmel e Bruckstein [1995b] é também um método de expansão de curvas de distância ou custo terminal constante. O custo em um pixel corresponde ao inverso da velocidade normal à curva no método de Kimmel e Bruckstein [1995b].

3.2.6 Condições de Borda ou Pontos Singulares ?

Esses resultados criam uma questão quanto à necessidade de se conhecer a orientação nas bordas do objeto para poder reconstruir sua forma a partir da imagem. Antes, se pensava que essa condição era necessária, porque o algoritmo de Horn [1990] funcionava adequadamente só quando a condição de borda era conhecida. O resultado de Oliensis e Dupuis [1994], por exemplo, desmistifica essa questão, mostrando que a condição que estabelece a existência e unicidade do Shape from Shading é a profundidade nos pontos singulares e na borda do domínio.

3.2.7 O Problema de Shape from Shading Global

Os algoritmos de Kimmel e Bruckstein [1995b] e de Oliensis e Dupuis [1994] resolvem o problema de Shape from Shading nas redondezas de um ponto singular ou para vários pontos singulares de profundidade conhecida. Há a necessidade de se estender esses algoritmos para o caso global, em que existem vários pontos singulares e não se tem nenhuma informação a priori sobre a profundidade destes pontos. A idéia é resolver localmente para cada ponto singular e combinar as soluções.

O que pode tornar difícil a combinação das soluções locais é a classificação dos pontos singulares. Não é possível se determinar se um extremo é um máximo ou um mínimo. Porém pontos singulares podem ser classificados através de, por exemplo, o teorema dos montanheiros. Outro problema que pode acontecer é a ausência no domínio da imagem de um ponto singular essencial para encontrar a solução. Veja em [Okatani e Deguchi 1998]

uma nova restrição para classificar pontos singulares baseada em isófotos (linhas de mesma intensidade).

♦ **Função de Morse e o Teorema dos Montanheiros**

Uma função de Morse é uma função localmente quadrática. Esta restrição garante que haverá apenas três tipos de pontos singulares: mínimos, máximos e pontos de sela. O teorema dos montanheiros diz que para esse tipo de função, o interior de um contorno de altura constante contém um número de extremos igual ao número de pontos de sela mais um.

De acordo com esse teorema, se for expandida uma curva de isoprofundidade a partir de um extremo, o primeiro ponto singular que será encontrado será um ponto de sela.

♦ **Algoritmo de Shape from Shading Global**

O primeiro passo do algoritmo de Shape from Shading global é calcular a transformada da distância ponderada por toda a imagem para cada ponto singular encontrado. Deve-se relacionar a cada ponto singular, o primeiro ponto singular encontrado na expansão de curvas de isoprofundidade ou aquele de menor distância. Define-se então um procedimento para combinar os resultados obtidos para dois pontos singulares (supostos extremos) que encontram, em primeiro lugar, o mesmo ponto singular (suposto ponto de sela). O resultado de cada ponto singular a ser mesclado deve ser normalizado para que a profundidade do suposto ponto de sela seja igual em ambos. O restante dos pontos do resultado da combinação é calculado como o menor valor entre os dois mapas de profundidade que estão sendo combinados. Considera-se então este resultado como se fosse gerado por um único extremo e procura-se o ponto singular de menor distância. Se os extremos combinados não forem extremos, a mesclagem deve ser desfeita e deve-se considerar outra combinação.

♦ **Prova de Unicidade da Solução Global**

Shimshoni, Kimmel e Bruckstein [1996] provaram que é única a classificação correta de pontos extremos e pontos de sela, utilizando o teorema dos montanheiros.

Suponha que os extremos sejam classificados no conjunto E' e os pontos de sela classificados no conjunto S' . Suponha que a classificação correta seja E e S , diferentes

do obtido. O teorema dos montanheiros diz que $|E'| = |E| = n + 1$ se $|S'| = |S| = n$. Como $|E'| - |S| = 1$, há pelo menos um extremo e' tal que $e' \in (E \cap E')$. O ponto singular de menor distância a e' é s' e é claro que $s' \in (S \cap S')$. Assim, $|E' - \{e'\}| - |S - \{s'\}| = 1$ e, como conseqüência, este argumento pode ser aplicado até que se esgotem os pontos classificados e se conclua que $E' = E$ e $S' = S$ contrariando a hipótese de não unicidade.

♦ Implementação por Backtracking

A demonstração de unicidade apresentada segue o mesmo paradigma de uma implementação por backtracking. Faz-se uma busca pelas classificações possíveis, retornando e desfazendo a combinação de mapas de profundidade quando se perceber que a classificação é incorreta. Fatalmente a classificação correta será obtida. Veja o algoritmo GSFS apresentado a seguir.

Algoritmo principal GSFS

- Detecte os pontos singulares.
- Calcule o mapa de profundidade para cada um deles.
- Enquanto não houver um mapa de profundidade que inclua todos os extremos e nem todas as classificações possíveis foram testadas, execute o algoritmo GSFS-1 com e' = o mapa de profundidade de um dos supostos extremos ou uma combinação de mapas de profundidade. Se falhou execute GSFS-1 com outro suposto extremo ou combinação.

Algoritmo GSFS-1

- Se houver apenas um mapa de profundidade, que combine todos os mapas de profundidade dos extremos, este é a solução. Termine o algoritmo GSFS.
- Classifique como ponto de sela o ponto s' de menor distância ao extremo e' (entrada), segundo seu mapa de profundidade. Encontre outro ponto singular ou combinação tal que s' tenha a menor distância e chame-o e'' . Se não encontrou retorne. Tendo encontrado, aplique o algoritmo GSFS-Merge com e' , s' e e'' .
- Enquanto todas as classificações possíveis em que e' e e'' sejam classificados extremos ou combinações e s' , ponto de sela, não foram testadas, execute GSFS-1 com e' = mapa de profundidade de um dos supostos extremos ou uma combinação de mapas de profundidade.

Algoritmo GSFS-Merge

- Translade os mapas de profundidade para que s' tenha a mesma profundidade nos dois mapas.
- A profundidade de cada pixel é dada pelo mínimo das profundidades dos dois mapas, assumindo que todo extremo é um mínimo.

3.3 Shape from Shading por Análise Local

A proposta dos algoritmos de análise local é determinar que tipo de informação pode ser obtida localmente, de forma a tornar o menos computacionalmente custoso o cálculo de Shape from Shading e de outras estimativas, como por exemplo, raio de curvatura, direção da fonte de luz e albedo da superfície. Tendo-se uma estimação local pode-se obter a orientação em um ponto da imagem sem ter que calcular a orientação de todos os outros pontos, além disso, procura-se obter um método robusto que aproxime a solução de Shape from Shading com pouco conhecimento sobre as condições da cena. É desejável que os algoritmos sejam robustos quanto a diversas formas de iluminação, de superfícies e variações do albedo. Porém, são necessárias restrições sobre a superfície que são bastante limitadoras, por exemplo, a suposição de que todo ponto da superfície é esférico. Outro problema com esses algoritmos é que a análise local é muito sensível ao erro de alta frequência.

No apêndice, serão apresentados os algoritmos mais conhecidos de análise local de sombreamento. Os métodos de Pentland [1984] relacionam os parâmetros de intensidade da imagem e suas derivadas até segunda ordem $(E, E_x, E_y, E_{xx}, E_{xy}, E_{yy})$ com τ (azimute da normal), σ (ângulo polar da normal), τ_l (azimute da fonte de luz), σ_l (ângulo polar da fonte de luz), R (raio de curvatura) e ρ (albedo da superfície). Os métodos de Lee e Rosenfeld [1985] utilizam derivadas até a primeira ordem. Ambos os conjuntos de estimadores baseiam-se na hipótese de que todo ponto da superfície é esférico, ou seja, as curvaturas principais são iguais. É uma hipótese bastante limitadora. Nos casos em que $E_{xx}E_{yy} - E_{xy}^2 < 0$, sabe-se de antemão que esta aproximação não funcionará e se deve utilizar uma aproximação por ponto de sela.

O modelo de iluminação utilizado é a lei de Lambert, enunciada como $E = \rho(\vec{n} \cdot \vec{l})$, onde \vec{n} é a normal à superfície e \vec{l} é a direção da fonte de luz, que se assume constante.

Mais recentemente, Zheng e Chellappa [1991] desenvolveram estimadores locais mais acurados e Šára [1995] desenvolveu um método baseado em isófotos (linhas de mesma intensidade). Na literatura, Shape from Shading por análise local é tratado à parte, como um outro problema, recebendo o nome de Local Shading Analysis.

3.4 Discussão sobre as Alternativas Apresentadas

A escolha de um método de Shape from Shading deve ponderar a velocidade, a robustez e a precisão. Precisando-se de um método rápido e robusto e não podendo determinar com segurança o modelo de iluminação, a sugestão é utilizar um método local. Veja [Pentland 1984] e [Šára 1995]. Se a necessidade for precisão ao invés de rapidez, para um modelo de iluminação e aquisição de imagem coerente com a realidade, deve-se utilizar um método de otimização ou um método de expansão de curvas. Para o método de expansão de curvas, se o esquema desenvolvido por Osher e Sethian [1988] puder ser utilizado, um algoritmo bastante robusto pode ser obtido. O método por otimização é questionável quanto à robustez porque sempre pode terminar em um mínimo local, não atingindo a solução. Por outro lado, é menos sensível ao ruído de alta frequência, devido a sua natureza global.

Em se tratando do método de Kimmel e Bruckstein [1995b] para o Shape from Shading, as soluções obtidas são locais para cada ponto singular identificado. É necessário agregar as soluções como mostrado no item 3.2.7, onde o problema de Global Shape from Shading é tratado. Pode haver a necessidade de aplicar o algoritmo de Kimmel e Bruckstein para várias curvas iniciais em torno de cada ponto singular (normal paralela à direção da fonte de luz) presente na imagem.

Para este trabalho, escolheu-se utilizar um método baseado em otimização, devido à flexibilidade que permite adaptá-lo a várias situações. Além disso, o funcionamento do algoritmo pode ser testado com uma implementação de um método básico e depois pode-se sofisticar o método. Um método de expansão de curvas baseado na solução de Okatani e Deguchi [1997] para endoscopia também foi implementado para fins de comparação.

Um levantamento de métodos de Shape from Shading e avaliação da eficiência destes pode ser encontrado em [Zhang, Tsai, Cryer e Shah 1994].

Capítulo 4 - ESTUDO DO ENDOSCÓPIO PARA CAPTURA DE IMAGENS DE SOMBREAMENTO

Neste capítulo, descreve-se o modelo de endoscópio segundo Okatani e Deguchi [1997]. Tendo em vista que este modelo não considera a distorção radial das lentes do endoscópio, um modelo de distorção radial baseado em projeção esférica é descrito. Utilizando este modelo, uma solução para o problema da distorção radial baseada em calibração prévia do endoscópio é apresentada neste capítulo. O modelo do endoscópio e da distorção radial é incorporada aos algoritmos de Shape from Shading no capítulo 5.

Um estudo sobre a refletância baseada em cor se faz para complementar o que já se tinha para imagens monocromáticas. O objetivo do uso da cor é isolar os componentes de reflexão especular e de reflexão difusa, para que se possa aplicar os algoritmos baseados na lei de Lambert. Esse estudo se baseia no modelo dicromático e no modelo de Phong. Em primeiro lugar, são apresentados os mecanismos de representação de cor. Em seguida, discutem-se o modelo de Phong e o modelo dicromático.

4.1 O Modelo do Endoscópio

4.1.1 Aquisição de Imagens de Endoscopia, um Modelo

Okatani e Deguchi [1997] utilizaram um modelo de formação de imagens, baseado no fato de que a fonte de luz e a câmera do endoscópio estão muito próximos. Nestas

condições não vai existir sombra projetada, todo ponto em que a reflexão difusa é máxima é também um ponto em que a reflexão especular é máxima.

Sejam θ e ϕ , o ângulo polar e o azimute, respectivamente, da normal à superfície em relação à direção do endoscópio (direção do observador, que é igual à direção da fonte de luz) e r , a distância entre o centro de projeção e o ponto da cena. Utilizando o modelo de fonte de luz pontual próxima à superfície, segundo Okatani e Deguchi [1997] modela-se a formação da imagem no endoscópio como

$$E(x, y) = \sigma \frac{G(\cos\theta)}{r^2}.$$

Veja que a refletância, neste caso, devido à fonte de luz próxima, depende da profundidade. Há também o problema do observador próximo, que faz com que as coordenadas (X, Y) da cena tridimensional não sejam equivalentes às coordenadas (x, y) da imagem. A conversão de coordenadas dependerá da profundidade. Por isso, é interessante utilizar as coordenadas (x, y) e definir o mapa de profundidade utilizando a distância $r(x, y)$.

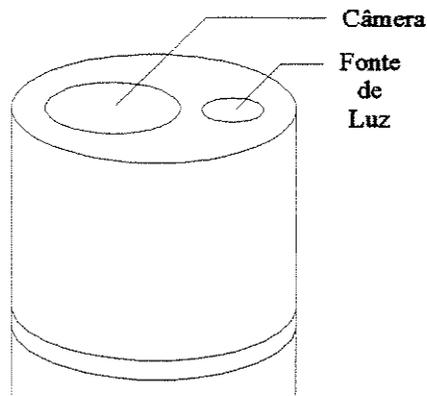


Figura 4-1 Cabeça do endoscópio – câmera e fonte de luz próximos.

4.1.2 Problemas Relacionados à Endoscopia

Há uma série de dificuldades ao se trabalhar com imagens reais de endoscopia. Muitas delas, como o problema da interreflexão foram desprezadas. Outras, como a presença de reflexão especular, foram tratadas. A seguir, citam-se as dificuldades mais relevantes para

processamento de imagens de endoscopia, identificadas ao dialogar com especialistas na área de endoscopia do Gastrocentro da Unicamp.

As paredes do estômago se movimentam muito e o movimento não é rígido, isto é, as distâncias não são preservadas. A não ser que se trate de peça cirúrgica, retirada de um paciente. Mas nesse caso, pode haver iluminação externa, destruindo a aplicabilidade dos modelos de iluminação. Também, não é muito fácil se fixar o endoscópio numa posição determinada. As superfícies da parede do estômago não são lambertianas. Há um componente especular de reflexão que vai depender da umidade da superfície (muco). Além disso, por se tratar de uma estrutura tubular, a iluminação devido a interreflexões ficará relevante para se obter um resultado mais preciso.

Cada paciente terá tonalidades diferente da mucosa estomacal. As tonalidades das paredes do estômago variam no mesmo paciente de região para região e também devido ao que se chama de idade gástrica. Além disso, quando há uma anormalidade, mais sangue é bombeado para a região problemática, tornando os vasos mais evidentes. Outra alteração da tonalidade ocorre em casos de sangramento ou presença de líquidos digestivos, como a bile.

4.2 Imagens com Distorção Radial

Adicionalmente ao modelo apresentado por Okatani e Deguchi [1997] (fonte de luz pontual que coincide com o centro de projeção), será considerada a projeção equidistante, que ocorre nas câmeras tipo olho de peixe (fisheye) ou grande angulares (wideangle) com distorção radial, visto que os endoscópios disponíveis para o trabalho apresentam essa configuração. O modelo que será utilizado aqui corresponde a uma projeção esférica seguida de uma projeção ortográfica.

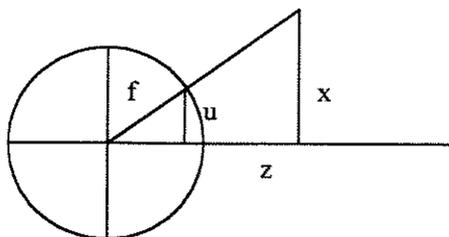


Figura 4-2 Modelo de distorção de câmera

Considera-se uma superfície esférica de raio f centrada na origem do sistema de referência que é também o centro de projeção. Projeta-se o ponto do espaço tridimensional sobre essa superfície esférica. Projeta-se então, o ponto da superfície esférica sobre o plano imagem, bastando para isso anular a coordenada z . O espaço tridimensional é mapeado, dessa maneira, sobre um círculo de raio f . A circunferência de raio f , que é o limite da imagem, corresponde à abertura de 180 graus. Sejam (u, v) as coordenadas da imagem distorcida, (x, y, z) as coordenadas no espaço do objeto e r a distância da origem até o ponto (x, y, z) . Lembra-se que a fonte de luz pontual é considerada coincidente com o centro de projeção portanto está na origem. Observe que a origem do plano de (u, v) é o centro da imagem (onde passa o eixo óptico ou eixo z).

Pela geometria do problema, usando semelhança de triângulos, obtém-se as relações

$$r(u, v) = \sqrt{x^2 + y^2 + z^2} \quad (4.1)$$

$$z(u, v) = \frac{\sqrt{f^2 - u^2 - v^2}}{f} r(u, v) \quad (4.2)$$

$$x(u, v) = \frac{z(u, v) \cdot u}{\sqrt{f^2 - u^2 - v^2}} \quad (4.3)$$

$$y(u, v) = \frac{z(u, v) \cdot v}{\sqrt{f^2 - u^2 - v^2}} \quad (4.4)$$

É importante notar que, uma vez determinada a função $z(u, v)$ ou a função $r(u, v)$, as funções $x(u, v)$ e $y(u, v)$ já estão determinadas. Portanto, a solução do problema pode ser simplesmente expressa na forma $z(u, v)$. Desejando-se obter uma malha de polígonos no espaço tridimensional, basta calcular as funções x e y a partir de $z(u, v)$.

4.2.1 Warping

A equação de warping para transformar uma imagem capturada por uma câmera com distorção radial em uma imagem normal de perspectiva pode ser obtida calculando a projeção dos pontos do plano imagem da projeção perspectiva sobre a imagem distorcida.

Assim, a profundidade z é constante $z = z_{pi}$, focal da projeção perspectiva (plano imagem).

Além disso, x e y são determinados uma vez que se conhece z .

$$x = \frac{z_{pi} \cdot u}{\sqrt{f^2 - u^2 - v^2}} \quad (4.5)$$

$$y = \frac{z_{pi} \cdot v}{\sqrt{f^2 - u^2 - v^2}}$$

Calculando a inversa,

$$u = \frac{f \cdot x}{\sqrt{z_{pi}^2 + x^2 + y^2}} \quad (4.6)$$

$$v = \frac{f \cdot y}{\sqrt{z_{pi}^2 + x^2 + y^2}}$$

Esta inversão é necessária para se implementar a transformação entre imagens. Para cada pixel da imagem em perspectiva normal que se quiser obter, deve-se consultar, na imagem original distorcida, o pixel cujas coordenadas são obtidas pelas expressões acima.

4.2.2 Parametrização da Superfície

O resultado do problema de Shape from Shading pode ser fornecido na forma $z(u, v)$, pois esta é uma parametrização da superfície visível e, portanto, contém toda informação que se pode obter através do Shape from Shading. Consultando as expressões acima, é possível obter os pontos no espaço euclidiano e interpolar como uma malha de triângulos, por exemplo.

4.2.3 Calibração da Distorção da Câmera

A distorção radial da imagem, segundo o modelo apresentado, pode ser representada pela equação

$$\sqrt{u^2 + v^2} = f \cdot \text{sen } \alpha \quad (4.7)$$

onde α é o deslocamento angular em relação ao eixo óptico. Um experimento que pode ser feito é gerar uma imagem impressa de círculos concêntricos correspondentes a diversos valores de α e filmá-la com um endoscópio corretamente posicionado. Com os valores de (u, v) e α , é possível se determinar f .

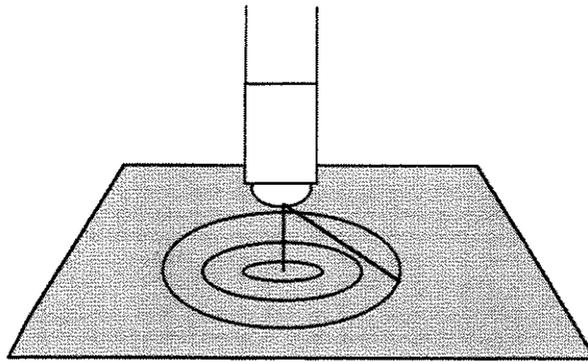


Figura 5-1 Correção da distorção da câmera

Por exemplo, posicionando-se o endoscópio perpendicularmente distando 5 cm do centro da folha de papel, um círculo de 5 cm de raio corresponderia a um ângulo de 45 graus. As coordenadas (u, v) são as coordenadas dos pixels contidos no círculo da imagem capturada, considerando a origem no centro dos círculos. Assim, podem-se tabelar os valores nominais de α com os valores experimentais de (u, v) , determinando, assim, um valor médio para f . O valor estimado é dado pela equação

$$\bar{f} = \frac{1}{n} \sum_{i=1}^n \frac{\sqrt{u_i^2 + v_i^2}}{\text{sen } \alpha_i}, \quad (4.8)$$

onde α_i é o ângulo em relação ao eixo óptico e (u_i, v_i) são coordenadas dos pontos da circunferência, assumindo a origem dos eixos no ponto central das circunferências.

4.3 Imagens Coloridas

A motivação de se incluir um estudo sobre cor nesse trabalho é que as imagens de endoscopia são sempre coloridas. A informação de cor é utilizada para remover a reflexão especular da imagem, para melhorar a aproximação ao se assumir que a superfície é lambertiana. Imagens coloridas podem ser utilizadas também para segmentação, por exemplo, em [Tian e Tsui 1997], utiliza-se segmentação de imagens coloridas para auxiliar o processo de reconstrução tridimensional.

4.3.1 Espaços de Cor

Este item do texto é baseado principalmente em [Foley *et al.* 1990].

A cor é uma função do comprimento de onda no intervalo do espectro visível, mas, normalmente, o que se tem é um conjunto de três valores que representa a cor. Isto é explicado pela teoria do triestímulo. Esta teoria assume a hipótese de que existem três tipos de sensores de cor na retina. Os componentes de cor são equacionadas como funções base.

♦ O Espectro Cromático

A luz é uma onda eletromagnética de comprimento de onda entre 400 nm (violeta) e 700 nm (vermelho). Apesar de a luz visível conter uma mistura dessas ondas nos seus vários comprimentos possíveis, a sensação de cor no olho humano é limitada a três dimensões. A resposta dos sensores da retina para cada comprimento de onda é a chamada função de resposta espectral.

Se for desejado criar uma sensação luminosa de uma cor específica, é possível, por exemplo, utilizar pontos de fósforo vermelho, azul e verde num tubo de raios catódicos. A soma ponderada do espectro da luz emitida por estas três fontes corresponde ao espectro que produzirá a sensação de cor. Utilizando-se as distribuições espectrais dos três tipos de fósforo, determina-se funções base para representar a cor visível equivalente. Note-se, entretanto, que nem toda cor visível pode ser representada por misturas de R, G e B, pois o valor de R deve ser negativo para algumas cores.

♦ RGB e CMY

O sistema RGB utiliza funções de base resultante da distribuição espectral dos três tipos de fósforo colorido do tubo de raios catódicos (vermelho, verde e azul). As variáveis R, G e B ponderam a soma das funções base. Essas funções são normalizadas para que $R=G=B$ correspondam a níveis de cinza. Assume-se 1 como o valor de ganho máximo e (1,1,1) a cor branca. O sistema RGB tem a interpretação intuitiva de cores primárias aditivas. Isto é, sabe-se intuitivamente que a mistura de uma luz vermelha (1,0,0) com uma luz verde (0,1,0) resulta numa luz de cor amarela (1,1,0). O sistema RGB pode ser representado em coordenadas cartesianas na forma de um cubo. O cubo é o fecho convexo das combinações baricêntricas de R, G, B, branco, preto, ciano, magenta e amarelo.

O inconveniente do sistema RGB é que nem toda cor visível pode ser representada. A não ser que se utilizem valores negativos.

O sistema CMY é equivalente ao RGB, mais especificamente, é formado pelo complemento do RGB. CMY representa as cores primárias subtrativas (ciano, magenta e amarelo). Novamente, há a noção intuitiva de mistura de cores. Se for misturado um pigmento ciano com um amarelo, o resultado será um pigmento verde.

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

O sistema CMYK é utilizado por impressoras coloridas que possuem tinta preta. A mistura das primitivas ciano, magenta e amarelo, na mesma proporção, resultam em preto. Assim, para se economizar tinta colorida e se evitar excesso de tinta no papel, utiliza-se um volume proporcional a K de tinta preta.

$$\begin{pmatrix} C_{CMYK} \\ M_{CMYK} \\ Y_{CMYK} \\ K \end{pmatrix} = \begin{pmatrix} C - K \\ M - K \\ Y - K \\ \min\{C, M, Y\} \end{pmatrix}$$

♦ CIE XYZ e CIE LUV

A *Commission Internationale de l'Éclairage* (CIE) padronizou um sistema de representação de cor capaz de representar toda cor visível através de um sistema de coordenadas com valores no intervalo de 0 a 1. Definiu-se três funções base ponderadas pelos parâmetros X, Y e Z. Esse sistema preserva a propriedade de combinação baricêntrica. A mistura de cores estará dentro do fecho convexo. O diagrama formado pela região de cores visíveis desenhada no plano XY é o diagrama de cromaticidade do CIE. Normalmente, a luz branca (luz do dia) é ressaltada. A parte curva da borda da região de cor visível é marcada com a frequência dominante. A parte reta corresponde a cores acromáticas. São cores que não tem equivalente a um impulso no espectro. Considera-se o comprimento de onda do complemento da cor neste caso. O complemento de uma cor é determinado pela reta que liga o branco a esta, de forma que a mistura de uma cor e seu complemento resulta em branco (tons de cinza).

A propriedade de combinação baricêntrica permite que se desenhe sobre o diagrama de cromaticidade as regiões de cores que um sistema pode reproduzir. Por exemplo, uma impressora que tenha três tipos de tinta colorida pode gerar as cores presentes no fecho

convexo das combinações baricêntricas das três cores, isto é, um triângulo no diagrama de cromaticidade.

O sistema CIE LUV tem a propriedade de que a distância euclidiana de duas cores corresponde à distância percebida pelo olho humano. Esse sistema, porém, não apresenta a propriedade de combinação baricêntrica.

♦ YIQ

O sistema YIQ é utilizado nas televisões comerciais norte-americanas, projetado para eficiência de transmissão e compatibilidade com o sistema de televisão preto e branco. O componente Y corresponde ao componente Y do CIE e é o componente visível na TV preta e branca. Esse componente pode ser obtido do sinal de vídeo colorido através de uma filtragem passa-baixa, como está ilustrado na figura 4-3. O sinal preto e branco é visível na TV colorida, porque a ausência do sinal de frequência mais alta corresponde à ausência de crominância, sendo apresentado em níveis de cinza. Por outro lado, a TV preta e branca ignora o sinal de mais alta frequência, recebendo apenas o componente Y.

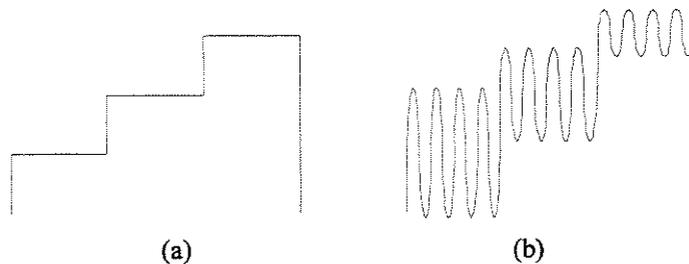


Figura 4-3 (a) Sinal de vídeo monocromático e (b) sinal de vídeo colorido.

♦ HSV

A maneira mais intuitiva de se caracterizar uma cor é utilizando o sistema HSV. Por não se tratar de funções de base, este sistema não apresenta propriedades como a de combinações baricêntricas, mas, para se interpretar uma única cor, é muito mais útil. H representa *hue* ou matiz. A interpretação é de comprimento de onda dominante, mas inclui cores acromáticas. O valor de H é representado como um ângulo de 0 a 360 graus, isto é, tem a topologia de uma curva fechada.

A saturação S corresponde à pureza da cor. Quanto mais seu espectro em comprimento de onda for próximo de um impulso, maior será a saturação. A saturação é mínima quando não se distingue o matiz, ou seja, nos tons de cinza, branco e preto. Define-se a saturação como $S = \frac{\max\{R, G, B\} - \min\{R, G, B\}}{\max\{R, G, B\}}$.

O valor V é equivalente à luminância, energia total do espectro. Neste sistema, $V = \max\{R, G, B\}$.

4.3.2 O Modelo de Reflexão de Phong

O modelo de Phong ganhou muita popularidade na síntese de imagens realistas devido à sua simplicidade e ao realismo visual que proporciona. Esse modelo é baseado em observações de fundo empírico e foi projetado para valorizar a continuidade da reflexão pela superfície. A principal aplicação para o modelo de Phong é o traçado de raios. Veja [Foley *et al.* 1990].

Nesse modelo, a reflexão é dividida em vários componentes que somados representam a luz refletida. Os principais componentes são a luz ambiente, a reflexão difusa e a reflexão especular.

♦ Iluminação Ambiente

Com o propósito de modelar fontes de luz de pequena importância e interreflexões, propõe-se supor a presença de uma fonte de luz constante e isotrópica (igual em todas as direções). A intensidade dessa luz ambiente é ponderada por uma constante fixa para cada material. A equação da iluminação ambiente no modelo é $I = I_a k_a$, onde I é a intensidade resultante, I_a é a intensidade da iluminação ambiente e k_a é uma constante que pondera a reflexão devido à luz ambiente. Observe o efeito da iluminação ambiente na figura 4-4.



Figura 4-4 Várias intensidades de iluminação ambiente.

♦ Reflexão Difusa

A reflexão em superfícies rugosas é modelada de forma a ser linearmente proporcional ao cosseno do ângulo entre o raio incidente e a reta normal à superfície. É uma consequência da Lei de Reflexão de Lambert, já referida neste texto.

Esta é a expressão do modelo para a reflexão difusa:

$$I = I_0 k_d \cos\theta,$$

onde I é a intensidade resultante, I_0 é a intensidade devido à fonte de luz, k_d é o coeficiente de reflexão difusa e θ é o ângulo entre a normal à superfície \vec{N} e a direção do raio incidente \vec{L} . Note que a intensidade não depende da posição do observador. Veja na figura 4-5 os ângulos e vetores envolvidos na determinação da reflexão difusa.

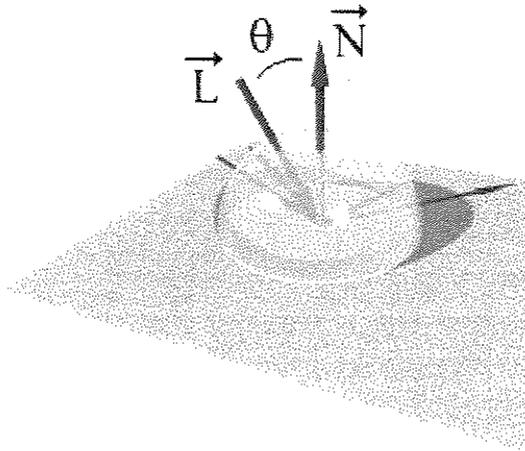


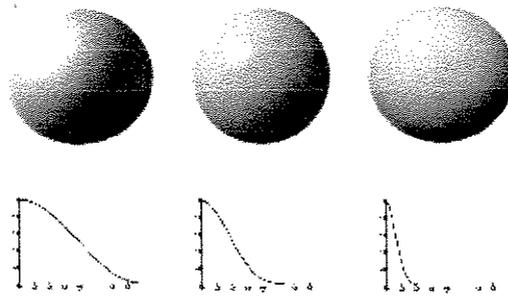
Figura 4-5 Vetores para determinação da reflexão difusa

♦ Reflexão Especular

A reflexão em superfícies polidas foi modelada por Phong como uma função de dois ângulos, como mostrado na figura 4-7. A reflexão especular depende do ângulo θ entre o raio de luz incidente \vec{L} e a reta normal à superfície \vec{N} e do ângulo α entre o raio de luz refletido \vec{R} e a direção do observador \vec{V} . A expressão utilizada no modelo de reflexão especular é:

$$I = I_0 W(\theta) \cos^n \alpha .$$

A função $W(\theta)$, na maioria dos casos, é implementada como uma constante k_s , o coeficiente de reflexão especular. O expoente de reflexão especular n está relacionado com a especularidade da superfície, de forma que, quanto mais elevado, mais a superfície se aproxima de uma superfície especular ideal (espelho). Observe, na figura 4-6, que a função $\cos^n \alpha$ decresce rapidamente, principalmente para, n elevado. O tamanho do lóbulo especular, portanto, é pequeno para n grande.



(a) (b) (c)

Figura 4-6 $\cos^n \alpha$ e imagens correspondentes, para diversos valores do expoente de reflexão especular
(a) $n=4$ (b) $n=16$ (c) $n=64$

O modelo de Phong completo é a soma dos três componentes descritos e é dado pela equação

$$I = I_a k_a + I_0 (k_d \cos \theta + k_s \cos^n \alpha).$$

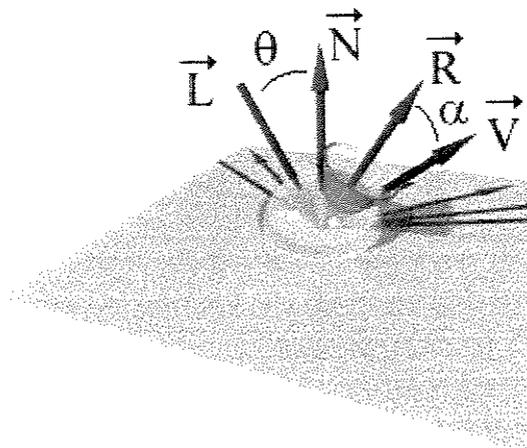


Figura 4-7 Vetores e ângulos para se determinar a reflexão segundo o modelo de Phong.

♦ O Modelo de Phong para Imagens Coloridas

Em imagens coloridas, o modelo de Phong é aplicado separadamente para cada comprimento de onda ou função base, como no caso de RGB. Porém deve-se ter o cuidado de variar os parâmetros que especificam a reflexão difusa e a reflexão especular para cada

comprimento de onda ou função base. Em geral, a constante de reflexão difusa é um vetor paralelo ao vetor que representa a cor da superfície e a constante de reflexão especular é paralela à cor branca. O modelo dicromático generaliza o modelo de Phong e justifica essa afirmação.

4.3.3 O Modelo Dicromático

O modelo dicromático é baseado em fatos experimentais e consiste da observação da reflexão de vários tipos de superfície que possuem uma camada gordurosa ou cerosa. Pode ser visto como uma generalização do modelo de Phong para imagens coloridas. No modelo de Phong, essa camada de interface será responsável pela reflexão especular, enquanto que a camada interna será suposta lambertiana. Precisamente, o modelo dicromático não faz considerações quanto à BRDF (veja ítem 2.1), mas simplesmente à existência de duas respostas diferentes ao espectro da luz incidente. Ao invés de se utilizar o espectro da luz visível, pode-se utilizar funções base, o que facilita o uso do modelo com imagens coloridas na forma RGB. Este estudo sobre o modelo dicromático é baseado em [Tominaga 1991], [Tominaga 1998] e [Tian e Tsui 1997].

♦ Descrição do Modelo

Seja λ um comprimento de onda ou um índice que seleciona uma função base. Seja $E(\lambda)$ a luz incidente, $S_I(\lambda)$ a refletância especular da camada de interface e $S_B(\lambda)$ a refletância de corpo pelo objeto. Seja θ uma variável que englobe todas as dependências da geometria do objeto. A luz refletida $Y(\theta, \lambda)$ é dada por

$$Y(\theta, \lambda) = c_I(\theta)S_I(\lambda)E(\lambda) + c_B(\theta)S_B(\lambda)E(\lambda) \quad (4.9)$$

onde $c_I(\theta)$ e $c_B(\theta)$ dependem da BRDF de cada camada da superfície.

[Tominaga 1991] utiliza o modelo dicromático para classificar e identificar superfícies. Variando a iluminação e fixando a geometria, ou fixando a iluminação e alterando-se a geometria do objeto, isto é, percorrendo a imagem, faz-se medidas de S_I e S_B . Considerando S_I constante em uma cena e S_B dependente do objeto, é obtida uma assinatura do material da superfície na forma de parâmetros que ponderam a combinação linear das funções base.

Entre os materiais que podem ser descritos pelo modelo dicromático, podem ser citados: plásticos, tintas, cerâmicas, vinil, azulejos, frutas, folhas e madeiras. Por outro lado, os seguintes materiais não podem ser descritos pelo modelo dicromático na forma apresentada aqui: metais, panos e papéis.

♦ **Interpretação de Diagramas através do Modelo**

O diagrama de dispersão de cor¹ é obtido plotando-se em um gráfico (bidimensional ou tridimensional) pontos cuja posição corresponde à cor de cada pixel. A frequência de ocorrência de uma cor é visualizada pela concentração da nuvem de pontos. O tipo de diagrama que será utilizado aqui é uma medida de $R \times G$, porém outros eixos podem ser utilizados. Veja [Tominaga 1998] para um estudo das partes dos diagramas de dispersão (chamados de histogramas no artigo) obtidos para o modelo dicromático e o modelo de Phong.

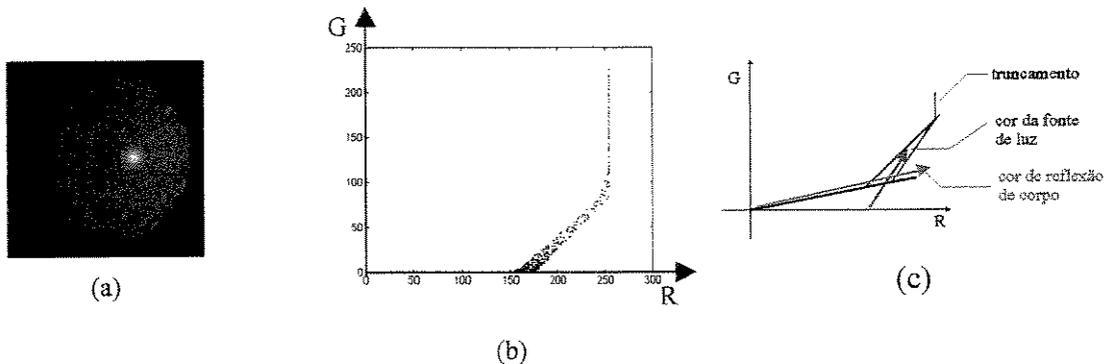


Figura 4-8 (a) Imagem sintética gerada utilizando o modelo de Phong colorido. (b) Diagrama de dispersão de cor para o modelo dicromático. Eixo horizontal é a intensidade de vermelho e eixo vertical é a intensidade de verde. (c) Como interpretar o diagrama segundo [Tominaga 1998].

Para ilustrar o modelo dicromático, foi gerada a imagem da esfera da figura 4-8, utilizando o modelo de Phong e o traçador de raios POV-Ray. Gerou-se também o diagrama de dispersão de cor, correspondente. Observa-se nesse diagrama como identificar

¹ Preferiu-se utilizar o termo diagrama de dispersão de cor ao invés de histograma, porque o diagrama não apresenta a frequência em que cada cor ocorre, mas sim a ocorrência ou não de cada cor.

a cor da superfície e a cor da iluminação (vetores). Observa-se, também, o efeito de saturação em intensidade¹ devido à iluminação ultrapassar o fundo de escala da câmera.

No caso em que o observador e a fonte de luz são coincidentes ou próximos, perde-se a capacidade de se determinar o valor de máxima reflexão difusa, porque esta se mistura com a reflexão especular. No caso em que o máximo de reflexão difusa está fora do lóbulo especular, determina-se este máximo pela cor de máxima saturação. A cor do ponto em que é máxima a reflexão especular é aquela cor que mais se aproxima da cor do iluminante, quando não houver saturação em intensidade.

4.3.4 Utilizando o Modelo Dicromático para a Remoção de Reflexão Especular

A fim de aplicar a técnica de Shape from Shading no caso em que os objetos apresentam reflexão difusa e especular e dispõe-se de imagens coloridas, o modelo dicromático pode ser empregado na separação da reflexão em componentes difusa e especular. No modelo de câmera e fonte de luz que está sendo empregado (sistema do endoscópio), não é possível se determinar o valor máximo de reflexão difusa porque os pontos em que este componente da reflexão é máxima pertencem à região do lóbulo especular. Porém, ainda assim, é possível extrair da imagem colorida uma imagem proporcional à imagem que contém apenas o componente de reflexão difusa.

Deseja-se obter apenas o componente da reflexão difusa devido ao corpo da superfície, que será assumida lambertiana. Isto é fácil quando se conhece a cor para a qual a reflexão difusa é máxima e a cor da fonte de luz. De fato, a cor de um pixel, obtida de uma superfície, é gerada por uma combinação baricêntrica (mistura de cores) do preto, da cor da fonte de luz e da cor da reflexão de corpo da superfície, mas sempre supondo que S_l , coeficiente que pondera a reflexão especular para cada comprimento de onda, é constante para todo comprimento de onda. Veja o item 4.3.3.

¹ Não confundir com saturação de cor.

Seja $[I_r, I_g, I_b]^T$ o vetor dos três componentes RGB da cor de um pixel, $[kd_r, kd_g, kd_b]^T$ o vetor dos coeficientes da reflexão de corpo e $[ks_r, ks_g, ks_b]^T$ o vetor dos coeficientes da reflexão da camada de interface. Supõe-se luz branca e, portanto,

$$E = [1,1,1]^T.$$

Considere $c_B(\theta)$ e $c_I(\theta)$ os coeficientes que correspondem à geometria para a camada de corpo e de interface, respectivamente. Se a camada de corpo for lambertiana, obtém-se $c_B(\theta) = \cos \sigma$, onde σ é o ângulo entre a normal e a direção de iluminação. É importante lembrar que θ é um vetor que contém, para um dado ponto da superfície, toda a informação geométrica relevante para o cálculo da reflexão.

Pelo modelo dicromático, sabe-se que

$$Y = \begin{bmatrix} I_r \\ I_g \\ I_b \end{bmatrix} = \begin{bmatrix} kd_r & ks_r \\ kd_g & ks_g \\ kd_b & ks_b \end{bmatrix} \begin{bmatrix} c_B \\ c_I \end{bmatrix}.$$

Assim, é necessário buscar um vetor $[a_1, a_2, a_3]^T$ perpendicular a $[ks_r, ks_g, ks_b]^T$, de forma que $a_1 ks_r + a_2 ks_g + a_3 ks_b = 0$.

$$\begin{bmatrix} a_1 I_r \\ a_2 I_g \\ a_3 I_b \end{bmatrix} = \begin{bmatrix} a_1 kd_r & a_1 ks_r \\ a_2 kd_g & a_2 ks_g \\ a_3 kd_b & a_3 ks_b \end{bmatrix} \begin{bmatrix} c_B \\ c_I \end{bmatrix}$$

Dessa forma,

$$a_1 I_r + a_2 I_g + a_3 I_b = (a_1 kd_r + a_2 kd_g + a_3 kd_b) c_B + (a_1 ks_r + a_2 ks_g + a_3 ks_b) c_I$$

$$a_1 I_r + a_2 I_g + a_3 I_b = (a_1 kd_r + a_2 kd_g + a_3 kd_b) c_B.$$

Portanto $c_B(\theta)$ é proporcional a $a_1 I_r + a_2 I_g + a_3 I_b$ para $[a_1, a_2, a_3]^T$ perpendicular a $[ks_r, ks_g, ks_b]^T$. Se for possível encontrar $[kd_r, kd_g, kd_b]^T$, pode-se determinar exatamente $c_B(\theta)$. Como a determinação desse parâmetro é difícil, no caso de observador e fonte de luz próximos, será escolhido um vetor adequado para $[a_1, a_2, a_3]^T$ dentre aqueles perpendiculares a $[ks_r, ks_g, ks_b]^T$ e, então, deve-se fazer uma estimativa da norma de $[kd_r, kd_g, kd_b]^T$.

Supondo-se $[ks_r, ks_g, ks_b] = [1 \ 1 \ 1]$, ou seja, que a reflexão de interface é branca, pode-se adotar $[a_1, a_2, a_3] = [1 \ -1 \ 0]$, por exemplo, perpendicular ao primeiro. Portanto uma imagem lambertiana pode ser obtida por $I_r - I_g$ ou qualquer outra combinação com $a_1 + a_2 + a_3 = 0$.

♦ Observações para aplicação em endoscopia

Para o caso de imagens de endoscopia gástrica, a reflexão na mucosa do estômago deve-se a 3 componentes. Há reflexão no muco que reveste as paredes do estômago, reflexão especular na camada de interface entre o tecido da mucosa e o meio e reflexão difusa de corpo do tecido da mucosa.

A reflexão devido ao muco tem expoente de reflexão especular n elevado, o que faz com que estes reflexos de alta intensidade ocupem uma área pequena. Esses pontos dão uma idéia precisa de que a normal à superfície aponta para o observador, considerando-se o modelo do endoscópio onde a fonte de luz e o centro de projeção da câmera coincidem. Porém, estes pontos devem ser isolados e não se deve utilizá-los no algoritmo de Shape from Shading devido ao erro de saturação em intensidade. Por outro lado, há a reflexão especular devido à camada de interface citada acima. Esta reflexão deve ser eliminada de forma a se obter uma imagem próxima ao que se espera de uma superfície lambertiana. A técnica apresentada pode ser utilizada com este intuito e será avaliada no item 6.2.

Capítulo 5 - IMPLEMENTAÇÃO DOS ALGORITMOS PARA RECONSTRUÇÃO TRIDIMENSIONAL EM IMAGENS DE ENDOSCOPIA

São apresentados, neste capítulo, os algoritmos de reconstrução tridimensional implementados. O algoritmo de Okatani e Deguchi [1997] é adaptado para englobar a distorção da câmera. Em seguida, é proposto um algoritmo baseado em otimização, semelhante aos algoritmos de Szeliski [1991] e de Ikeuchi e Horn [1981] ou Horn [1990], mas tendo como objetivo a reconstrução de imagens de perspectiva, com centro de projeção e fonte de luz coincidentes.

5.1 Adaptação do Algoritmo de Okatani e Deguchi

O algoritmo de Okatani e Deguchi [1997] é a extensão para o problema do endoscópio do algoritmo de Kimmel e Bruckstein [1995b], discutido no item 3.2.2. A adaptação feita neste projeto corresponde a incorporar a distorção radial da câmera ao algoritmo de Okatani e Deguchi. Isso permite evitar que se tenha que realizar uma etapa de warping com interpolação, que adicionaria erro ao processo.

5.1.1 O Algoritmo Original

Em seu algoritmo, Okatani e Deguchi [1997] representam o mapa de profundidade pela distância $r(x, y) = \sqrt{x^2 + y^2 + z^2}$ por ser mais adequada que $z(x, y)$ ao sistema de aquisição de imagem do endoscópio (centro de projeção e fonte de luz pontual coincidentes

e próximos ao objeto). Utilizam, assim, curvas para as quais a distância à origem é constante ao invés de curvas de profundidade constante. Essas curvas de distância à origem constante são a interseção da superfície com uma esfera de raio r e centro no centro de projeção da câmera. Lembra-se que, no modelo do endoscópio, é esperado que a fonte de luz coincida com o centro de projeção da câmera e é nesse ponto que se convencionou a origem do sistema de referência.

No método de Kimmel e Bruckstein [1995b] as curvas são propagadas sobre a imagem na direção normal à curva, na proporção dada pela velocidade v , obtida da imagem. A expansão pode ser escrita na forma:

$$\frac{\partial C}{\partial t} = v\hat{n}, \quad (5.1)$$

onde $C(x, y; t)$ representa a curva e \hat{n} a direção normal.

No caso de imagens de endoscopia, a velocidade normal v à curva (responsável pela expansão das curvas) deve ser calculada levando em conta o modelo do endoscópio, isto é, projeção perspectiva e fonte de luz próxima do objeto. Os cálculos de Okatani e Deguchi [1997] levam à expressão:

$$v(x, y, t) = \frac{\sqrt{(x^2 + y^2 + z_{pi}^2)((x^2 + z_{pi}^2)n_1^2 + 2n_1n_2xy + (y^2 + z_{pi}^2)n_2^2)}}{z_{pi}t(n_1^2 + n_2^2) \tan \sigma}, \quad (5.2)$$

onde z_{pi} é a distância focal do endoscópio, (n_1, n_2) é o vetor normal à curva de nível, σ é o ângulo entre a normal à superfície e o raio de luz incidente e t representa o tempo e é equivalente à distância de qualquer ponto da curva ao centro de projeção.

Determina-se, também, que

$$\tan \sigma = \frac{\sqrt{1 - (G^{-1}(t^2 E(x, y) / \rho))^2}}{G^{-1}(t^2 E(x, y) / \rho)}, \quad (5.3)$$

onde ρ é o albedo e $G(\cos \sigma)$ é uma função que corrige o mapa de refletância. Para as implementações desse trabalho, E deve ser normalizada de forma que $\rho = 1$. Vale lembrar que a qualidade da reconstrução é muito dependente desse fator, ou da correta normalização da imagem [Okatani e Deguchi 1997].

O esquema iterativo que resolve o problema é implementado na forma euleriana apresentada no item 3.2.1. Deve-se inicializar a função ϕ com uma curva inicial, que pode

ser uma pequena circunferência em torno de um ponto singular. O esquema iterativo obtido é dado por

$$\phi_t = -\frac{G^{-1}(t^2 E(x, y) / \rho)}{z_{pi} t \sqrt{1 - (G^{-1}(t^2 E(x, y) / \rho))^2}} \sqrt{(x^2 + y^2 + z_{pi}^2) \left((x^2 + z_{pi}^2) \phi_x^2 + 2xy \phi_x \phi_y + (y^2 + z_{pi}^2) \phi_y^2 \right)}. \quad (5.4)$$

5.1.2 Adaptação para uma Câmera com Distorção

Na seqüência, o algoritmo de Okatani e Deguchi [1997] é adaptado a fim de incorporar a correção da distorção radial da câmera.

A partir das equações do warping, calcula-se o jacobiano, que tem a finalidade de efetuar a conversão das derivadas de ϕ do sistema de coordenadas da imagem com distorção para o sistema da imagem sem distorção correspondente. O jacobiano é a matriz 2×2 na expressão abaixo, que faz a conversão do gradiente de ϕ .

$$\nabla \phi = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} \phi_u \\ \phi_v \end{bmatrix}$$

Assim,

$$\begin{aligned} \phi_x &= u_x \phi_u + v_x \phi_v \\ \phi_y &= u_y \phi_u + v_y \phi_v \end{aligned} \quad (5.5)$$

Onde cada derivada, obtida a partir das equações (4.6), é dada por

$$\begin{aligned} u_x &= \frac{f}{\sqrt{z_{pi}^2 + x^2 + y^2}} - x^2 f \cdot (z_{pi}^2 + x^2 + y^2)^{-\frac{3}{2}}, \\ u_y &= v_x = -xyf \cdot (z_{pi}^2 + x^2 + y^2)^{-\frac{3}{2}}, \\ v_y &= \frac{f}{\sqrt{z_{pi}^2 + x^2 + y^2}} - y^2 f \cdot (z_{pi}^2 + x^2 + y^2)^{-\frac{3}{2}}, \end{aligned} \quad (5.6)$$

onde f é o raio da superfície esférica da projeção como descrito no item 4.1.3.

Substituindo-se, na equação (5.4), as equações (5.5) e (5.6), obtém-se

$$\frac{\partial \phi}{\partial t} = \frac{-tE(u, v)}{\sqrt{1 - t^4 E^2}} \sqrt{(x^2 + y^2 + 1) \left[(x^2 + 1) \phi_x^2 + 2xy \phi_x \phi_y + (y^2 + 1) \phi_y^2 \right]}, \quad (5.7)$$

onde a reflexão foi considerada lambertiana e $z_{pi} = 1$.

O algoritmo consiste em iterar a equação (5.7) discretizada no tempo. A cada iteração corresponde um valor de distância ao centro de projeção. Quando um pixel (u, v) sofrer troca de sinal do valor de $\phi(u, v)$, este pixel passou da região exterior à curva em expansão para a região interior no instante correspondente à iteração. Dessa forma, atribui-se a $r(u, v)$ o valor de distância associado à iteração.

Se faz necessário adicionar a restrição

$$1 - t^4 E^2 > 0 \Rightarrow t < \sqrt{\frac{1}{E}}. \quad (5.8)$$

Esta restrição é consequência do limite superior do coseno do ângulo entre a normal à superfície e a direção da fonte de luz. Este limite ocorre nos pontos singulares da função $r(u, v)$, onde a superfície do objeto é tangente a uma esfera cujo centro coincide com o centro de projeção .

Assim, iniciando-se com o valor mínimo de t , não se deve utilizar $t > \sqrt{\frac{1}{E}}$ em qualquer pixel. Devido a essa restrição foi assumida a seguinte postura quando $t > \sqrt{\frac{1}{E}}$. Se o pixel tratado já tem sua profundidade $r(u, v)$ determinada, substitui-se t por $r(u, v)$; caso contrário, o pixel deve corresponder a um novo ponto singular, assim, associa-se a ele a profundidade $\sqrt{\frac{1}{E}}$, que é a profundidade correta para qualquer ponto singular. Essa postura adiciona um caráter global ao algoritmo, pois resolve o problema localmente e permite o tratamento de pontos singulares.

Nota-se que há necessidade de utilizar derivadas *upwind* para ϕ_u e ϕ_v , a fim de manter as condições de entropia [Osher e Sethian 1988]. Veja no item 3.2.1 como construir o esquema iterativo.

5.2 Algoritmo de Reconstrução por Otimização

O algoritmo descrito neste item foi desenvolvido neste projeto. Trata-se de um método de Shape from Shading baseado em otimização e é próximo da solução proposta por Szeliski [1991] (veja item 3.1.4) ou da solução proposta por Horn [1990] (veja item 3.1.1).

Para resolver o problema utilizando otimização, é necessário definir uma função objetivo que seja mínima na solução procurada. Uma maneira de se fazer isso é através do método de quadrados mínimos. Define-se a função objetivo como a soma dos quadrados das diferenças entre os valores medidos e os valores estimados. Para mais detalhes sobre o método de quadrados mínimos veja [Tarantola 1987].

Trabalhando-se com superfícies lambertianas, uma imagem $E(u, v)$ terá a intensidade de seus pixels proporcional ao cosseno ($\cos \sigma$) entre a normal à superfície e a direção da fonte de luz e inversamente proporcional ao quadrado da distância r do ponto da superfície à fonte de luz, representado pela equação

$$E = \frac{\cos \sigma}{r^2}. \quad (5.9)$$

Nesta equação, assume-se que o albedo é unitário. Os valores dos pixels da imagem devem ser divididos pelo valor do albedo a fim de cumprir essa condição. Observa-se, na equação (5.9), que o conhecimento do valor de r para um ponto singular, onde $\cos \sigma = 1$, é o suficiente para se determinar o valor de escala a ser aplicado na intensidade da imagem.

♦ Escolha da função objetivo

Para aplicação de um método de otimização, é possível escolher, além da função a ser otimizada (função objetivo), o domínio a que as variáveis dessa função pertencem (espaço de busca). O resultado do método de otimização é um ponto pertencente a esse domínio.

Para o problema de Shape from Shading com o sistema do endoscópio, escolheu-se a função objetivo tendo a variável $\cos \sigma(u, v)$ como espaço de busca. A condição de visibilidade define os limites do ângulo entre a direção da luz e a normal à superfície como zero e 90 graus. Assim, o cosseno varia de um a zero. Essa variável, portanto, tem extremos bem definidos dentro do intervalo $[0;1]$, motivando seu uso, ao invés de $r(u, v)$, como espaço de busca. Na prática, o valor um é raro e corresponde a pontos singulares ou a pontos de uma superfície esférica com centro no centro de projeção e o valor zero

corresponde à inclinação de 90 graus da normal de forma que a área que contém valor zero é nula.

A função objetivo pode ser definida de forma procedural, como apresentado na seqüência. A partir de uma estimativa inicial de $\cos\sigma(u, v)$, pode-se obter o mapa $r(u, v)$ da distância de cada pixel ao centro de projeção, utilizando a equação (5.9), pois a imagem $E(u, v)$ é conhecida. Como descrito no item “Cálculo da normal à superfície”, a partir de $r(u, v)$, é possível calcular o ângulo entre a direção da normal à superfície e a direção da fonte de luz, obtendo uma função que será denominada $(\diamond r)(u, v)$. Observa-se que \diamond pode ser visto como um operador não-linear. No caso em que $(\diamond r)(u, v) = \cos\sigma(u, v)$, o problema está resolvido e a solução pode ser dada na forma de $r(u, v)$. Assim, define-se o erro a ser minimizado através da equação (5.10), já na forma discretizada.

$$e = \sum_{ij \in \Omega} (\cos\sigma_{ij} - (\diamond r)_{ij})^2 \quad (5.10)$$

onde Ω é o domínio das coordenadas da imagem.

É possível escrever uma expressão para o operador \diamond na forma

$$(\diamond r)(u, v) = F\left(r, \frac{\partial r}{\partial u}, \frac{\partial r}{\partial v}, u, v\right), \quad (5.11)$$

mas é muito mais fácil utilizar um algoritmo para esse cálculo.

• Cálculo da normal à superfície

A superfície é parametrizada através das coordenadas da imagem. Isso permite limitar a região parametrizável da superfície à região visível da cena. O procedimento para calcular o vetor normal à superfície, necessário para avaliar a função objetivo, é apresentado a seguir.

Seja $\vec{P} = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}$ um ponto do objeto.

$\vec{n} = \frac{\vec{P}_u \times \vec{P}_v}{\|\vec{P}_u\| \cdot \|\vec{P}_v\|}$ é o vetor normal obtido por produto vetorial. As variáveis \vec{P}_u e \vec{P}_v são,

respectivamente, a derivada do vetor \vec{P} em relação a u e a derivada do vetor \vec{P} em relação a v .

$$\begin{bmatrix} n1 \\ n2 \\ n3 \end{bmatrix} = \vec{P}_u \times \vec{P}_v = \begin{vmatrix} i & j & k \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} = \begin{bmatrix} y_u z_v - z_u y_v \\ -x_u z_v + z_u x_v \\ x_u y_v - y_u x_v \end{bmatrix}$$

$$\vec{n} = \frac{(n_1, n_2, n_3)}{\sqrt{n_1^2 + n_2^2 + n_3^2}}$$

O valor de $\cos\sigma$ ou de $(\hat{r})(u, v)$ é obtido pelo produto escalar do vetor normal à superfície e a direção da fonte de luz. Para o problema do endoscópio, a direção da fonte de luz é a direção do vetor (x, y, z) .

A discretização pode ser feita utilizando diferenças finitas das derivadas em u e v , que são as coordenadas da imagem e parâmetros da superfície visível.

• Método baseado no gradiente

A fim de resolver o problema de otimização, um método simples de descida na direção do gradiente foi utilizado. Veja no item 3.1.4 a descrição do método de gradiente de forma genérica.

Para o caso modelado pela equação (5.9), este método é implementado para encontrar a função $z(u, v)$ ou $r(u, v)$ para a qual a função objetivo seja mínima. Lembrando que a variável $\cos\sigma(u, v)$ foi escolhida como variável de busca, observa-se que basta conhecer $\cos\sigma(u, v)$, para se obter $r(u, v)$ e depois $z(u, v)$.

Assim, o modelo empregado é, na forma de equação diferencial,

$$\frac{\partial \cos\sigma_{ij}}{\partial t} = -\frac{\partial e}{\partial \cos\sigma_{ij}}. \quad (5.12)$$

Como a equação é definida proceduralmente e a derivada seria bastante trabalhosa de se computar, um método quasi-gradiente pode ser empregado, substituindo a derivada por uma diferença finita

$$\frac{\partial e}{\partial \cos \sigma_{ij}} \approx \frac{e(\dots, \cos \sigma_{ij} + \Delta c, \dots) - e(\dots, \cos \sigma_{ij}, \dots)}{\Delta c}. \quad (5.13)$$

O método de otimização sempre pode ser utilizado quando não é certa a existência de uma solução para o sistema de equações. No caso de existirem diversas soluções, ou diversos mínimos locais, a técnica de regularização, apresentada no item 3.1.1, pode ser empregada para escolher a solução mais suave. Porém, com o uso de regularização, a função objetivo é alterada e o método pode se afastar da solução verdadeira caso o coeficiente de regularização seja definido de modo inadequado.

O método do gradiente converge para uma solução do sistema de equações $\nabla e = 0$. Na verdade, o que se deseja é a solução para $e = 0$, mas não são conhecidas as condições para que essa solução exista. Como e foi definida, em (5.10), utilizando quadrados mínimos, se $e = 0$, então $\nabla e = 0$. No entanto, $\nabla e = 0$ pode ocorrer para $e \neq 0$, caracterizando um mínimo local.

A implementação que gerou os resultados que serão apresentados no capítulo 6 utiliza o método do gradiente e o termo de regularização. O esquema iterativo é dado por

$$\Delta \cos \sigma_{ij} = \Delta t \frac{e(\dots, \cos \sigma_{ij} + \Delta c, \dots) - e(\dots, \cos \sigma_{ij}, \dots)}{\Delta c} - \lambda(4 \cos \sigma_{ij} - \cos \sigma_{i-1j} - \cos \sigma_{i+1j} - \cos \sigma_{ij-1} - \cos \sigma_{ij+1}), \quad (5.14)$$

onde λ pondera o termo de regularização.

Para fins de eficiência, observa-se que $e(\dots, \cos \sigma_{ij} + \Delta c, \dots) - e(\dots, \cos \sigma_{ij}, \dots)$ pode ser calculado localmente, pois se trata de uma diferença entre somatórios por toda a imagem, em que no máximo 5 parcelas variam (o pixel ij e sua vizinhança).

5.2.1 Outras Alternativas para Resolver o Sistema de Equações

São muitas as alternativas para o método proposto. É possível empregar outros métodos de otimização e métodos de solução de sistema de equações. Além disso, pode-se utilizar a técnica de múltipla escala para acelerar a convergência.

O método do gradiente pode ser melhorado através do método do gradiente conjugado ou através do método de Newton. No caso do método de Newton, é necessário se resolver um sistema linear a cada iteração. A matriz que representa a hessiana da função objetivo, a ser invertida, é esparsa e tem a forma chamada de tridiagonal com franjas. Este tipo de

esparsidade acontece quando se utiliza vizinhança de 4 pixels na conversão do caso contínuo para o discreto através de diferenças finitas. Porém a inversa não é esparsa. Deve-se, então, resolver o sistema através de relaxação de Jacobi ou através de decomposição LU, visto que as matrizes da decomposição são também esparsas. Veja em [Press *et al.* 1992] a explicação detalhada do método de Newton.

O método de relaxação se apoia no fato de que a variação deve ser nula quando a solução for encontrada. O sistema é realimentado com o próprio erro, ou seja, a diferença entre o dado observado e o dado estimado. Assim, um esquema para o problema objeto deste trabalho pode ser, por exemplo,

$$\Delta r_{ij} = \gamma \cdot (\text{image}(r)_{ij} - E_{ij}), \text{ onde } \text{image}(r)_{ij} = \frac{(\langle r \rangle)_{ij}}{r_{ij}^2}$$

$$\text{ou } \Delta \cos \sigma_{ij} = \gamma \cdot (\langle r_{ij} \rangle - \cos \sigma_{ij}).$$

Trata-se de um método de convergência extremamente lenta. O parâmetro ε pondera a lentidão e a estabilidade. Métodos, como o de sobre-relaxação sucessiva (SOR), que calculam γ a cada iteração corrigem esta desvantagem. Veja [Press *et al.* 1992] para maiores detalhes.

Uma outra forma de se estabelecer um esquema iterativo visando a otimização de uma função é o simulated annealing, que possui propriedades de otimização global. É um método lento, porém, permite implementação paralela. Essa técnica foi utilizada em [Torreão 1992] para problemas de Visão Computacional. Pode ser utilizada a fim de melhorar o método do gradiente adicionando a propriedade de convergência global para a solução. A explicação detalhada da técnica de simulated annealing também pode ser encontrada em [Press *et al.* 1992].

A técnica de múltipla resolução ou múltipla escala consiste em armazenar várias cópias da imagem com resolução sucessivamente reduzida. Por exemplo, se a imagem original for 64 por 64, pode-se utilizar imagens reduzidas 32 por 32, 16 por 16 e 8 por 8, simulando uma equação diferencial para o nível mais reduzido e transferindo as alterações para o nível seguinte de maior resolução. Veja, por exemplo, [Dold e Eckmann 1981] ou [Haar Romeny 1994].

Capítulo 6 - RESULTADOS

São apresentados, neste capítulo, resultados dos algoritmos implementados, a fim de avaliar sua aplicabilidade. Inicia-se com a determinação da distorção da câmera. Em seguida, avalia-se o filtro para remoção da reflexão especular. E, por fim, resultados para os dois algoritmos de reconstrução a partir de sombreamento são apresentados.

6.1 Correção da Distorção

Para se determinar o parâmetro da distorção da câmera f considerado na equação do modelo de câmera (4.7), utilizou-se uma imagem calibradora, como a da figura 6-1, fotografada por um endoscópio. O raio da circunferência indicada como 45 graus é 5 cm, portanto o centro de projeção deve ser posicionado 5 cm acima do centro das circunferências, de forma que o ângulo formado entre o eixo óptico e a reta que passa pelo centro de projeção e qualquer ponto da circunferência seja de 45 graus.

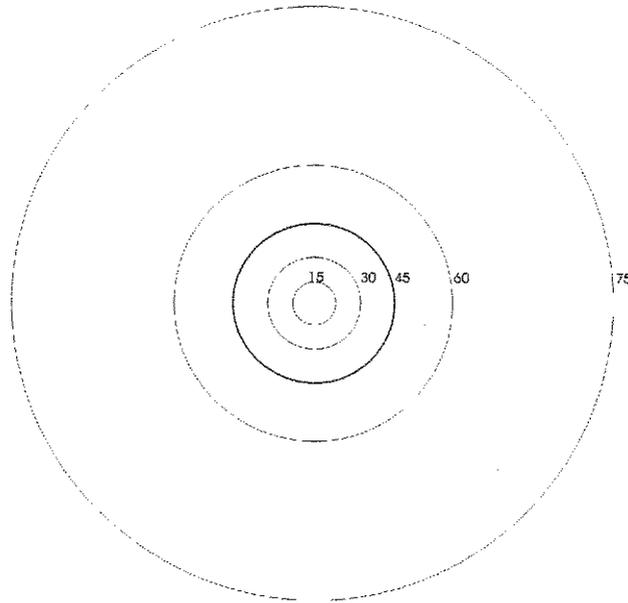


Figura 6-1 Imagem para calibração do parâmetro de distorção da câmera

Há uma circunferência a cada 15 graus de 15 a 75 graus. Porém a abertura da câmera permite ver apenas 100 graus, isto é, um ângulo de 50 graus para cada lado do eixo óptico. Assim, a circunferência correspondente ao ângulo de 60 graus com o eixo óptico mal aparece na imagem fotografada na figura 6-2.

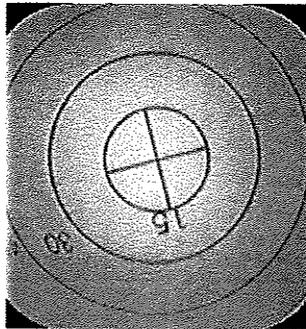


Figura 6-2 Imagem capturada da imagem calibradora

Utilizando-se a imagem da figura 6-2 é possível estimar o parâmetro de distorção f da câmera medida em pixels como foi mostrado no item 4.2.3. Com um editor de imagens, selecionam-se pixels das circunferências, medem-se suas distâncias em pixels em relação ao centro (alvo) e substituem-se os valores de ângulo e raio na equação (4.7). A estimativa do parâmetro de distorção pode ser obtida, por exemplo, pela média destas medidas. O valor estimado é dado pela equação (4.8).

Através desse método obteve-se um valor próximo de 286 para a distorção f . Entretanto, as condições em que a imagem foi capturada não foram ideais. Utilizou-se uma régua para manter o endoscópio à distância estabelecida e não se dispunha de equipamento para manter o endoscópio na posição correta. Assim, esta medida não é confiável. Um outro procedimento baseado na obtenção de imagens de uma grade regular foi empregado. Transformações de correção de distorção (*warping*) foram feitas para vários valores atribuídos ao parâmetro da distorção. Analisando-se os resultados do *warping* aplicado sobre imagens da grade regular, notou-se que resultados são melhores para o parâmetro de distorção em torno de 255.

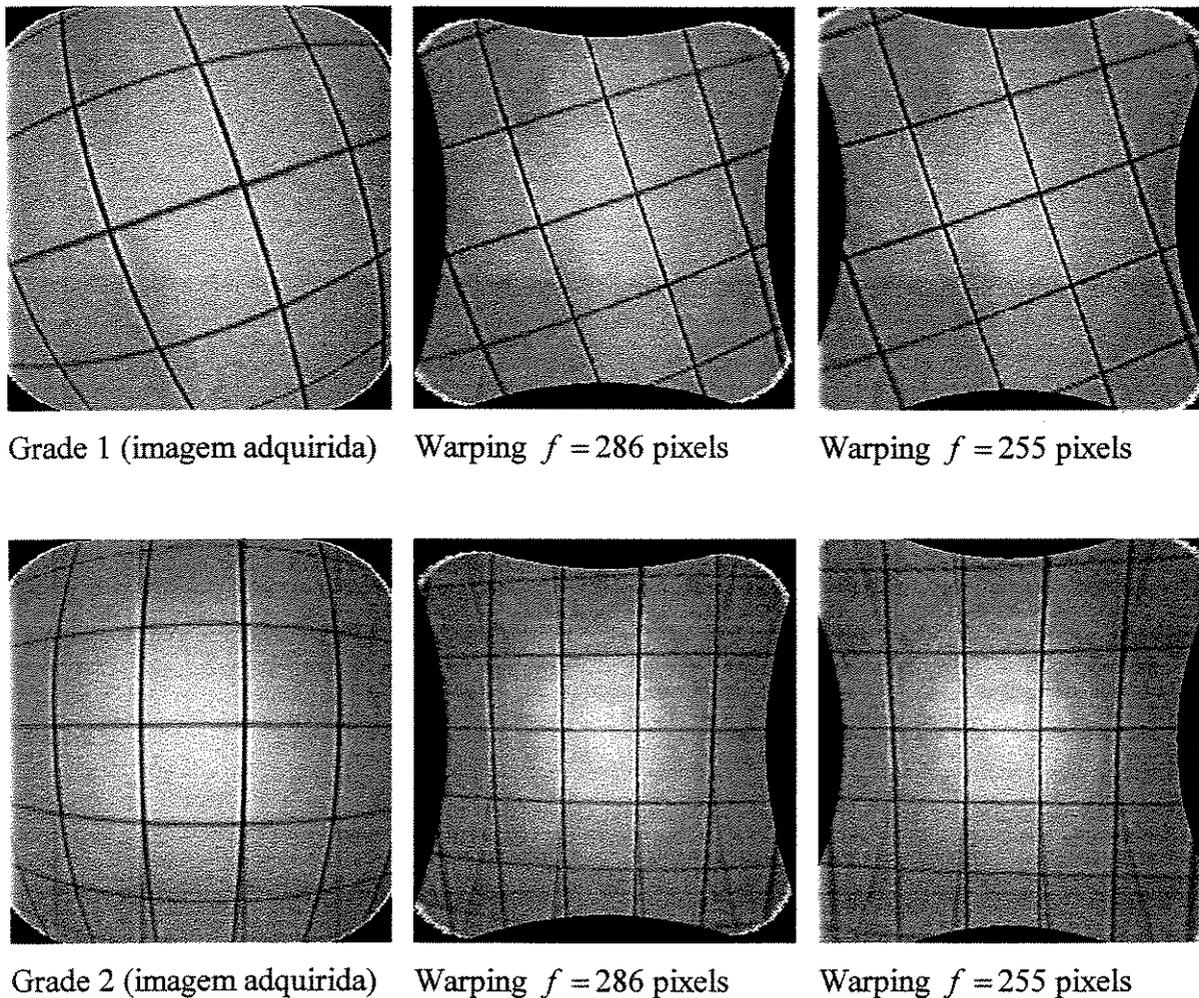
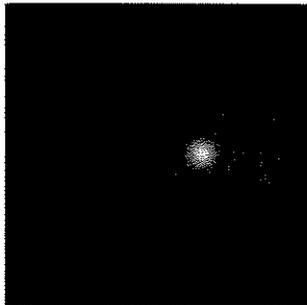


Figura 6-3 Teste para determinação do parâmetro de distorção do endoscópio

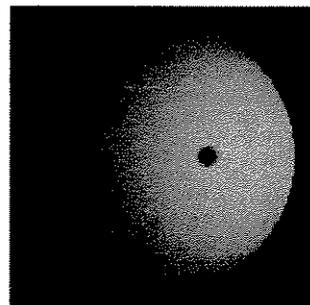
6.2 Remoção da Reflexão Especular

O filtro para remoção do componente de reflexão especular utilizando o modelo dicromático, apresentado no item 4.4, foi aplicado a diversas imagens que são mostradas a seguir. Devido à falta de uma técnica para quantificar a intensidade da reflexão difusa, a avaliação se limitará a observações do resultado da aplicação do filtro.

Para a imagem da esfera da figura 6-4 (a), gerada no POV-Ray, a imagem lambertiana procurada, vista na figura 6-4 (b), foi obtida subtraindo o canal verde do canal vermelho.



(a) original



(b) resultado

Figura 6-4 Remoção da reflexão especular numa imagem sintética

O ponto preto na figura 6-4 (b) é decorrente da saturação de intensidade na imagem original. A região na qual há saturação do valor de intensidade em 255 (máximo valor do byte), não pode ser considerada. Basta, para remover a região problemática, isolar a região em que qualquer dos valores da tripla RGB apresente valor igual a 255.

Observe o resultado do filtro para a imagem de endoscopia na figura 6-5 (a). O filtro utilizado foi a diferença entre o canal vermelho e o canal azul. Nota-se que a mancha de bile, região amarelada na porção superior central da imagem, foi ressaltada (se tornou mais brilhante). Nota-se também que, se as imagens da figura 6-5 forem comparadas visualmente, concluí-se que parte da informação de profundidade foi perdida (veja a região apontada pela seta).

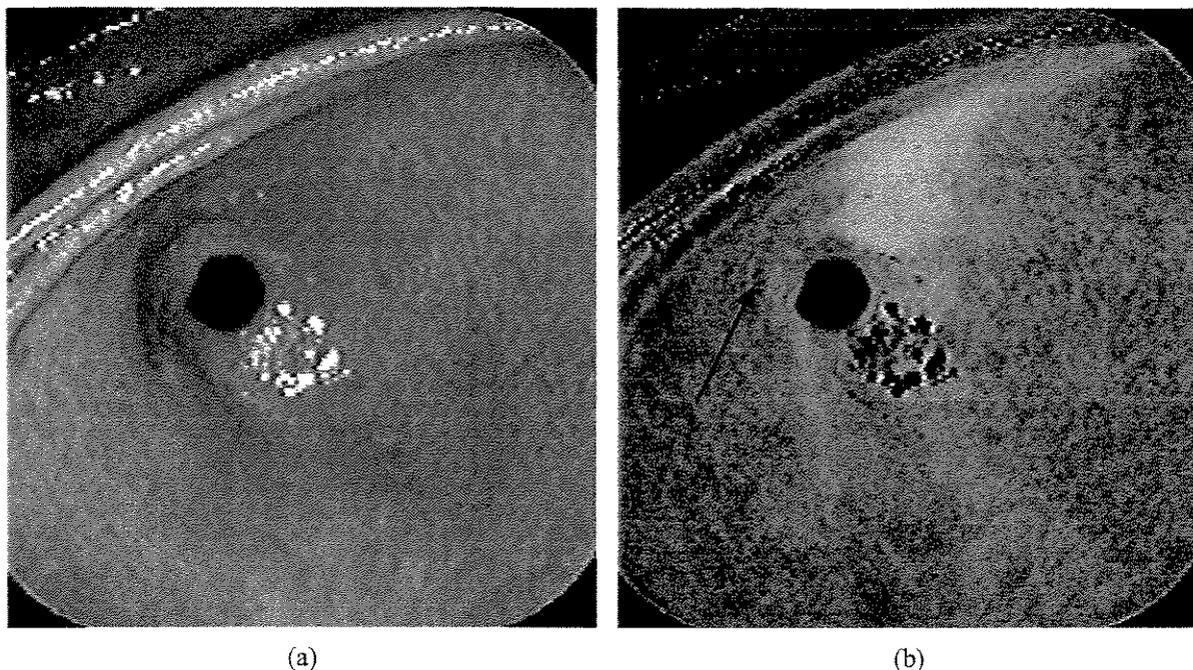


Figura 6-5 Remoção da reflexão especular numa imagem de endoscopia

A imagem do boneco de resina, na figura 6-6, foi adquirida utilizando um scanner de mesa. Por ser feito de um material cuja reflexão segue o modelo dicromático (resina), o resultado obtido está bem próximo do esperado dado que, visualmente, a imagem resultante parece proceder de um material que apresente apenas reflexão difusa.

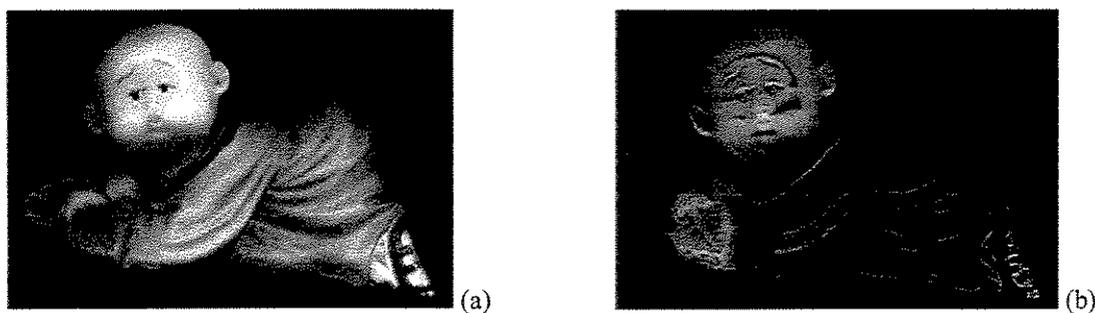


Figura 6-6 Remoção da reflexão especular utilizando o modelo dicromático de reflexão

A filtragem para remoção da reflexão especular mostrou bons resultados para imagens geradas artificialmente utilizando o modelo de Phong. Para o bule de cerâmica da figura 6-8 (b), cuja imagem filtrada pode ser vista na figura 6-30, a observação do resultado indica sucesso, apesar de ser possível apenas julgar o resultado através de sua visualização.

É preciso lembrar que fatores como a iluminação ambiente ou interreflexões prejudicam o resultado. Há também a necessidade de se comprovar experimentalmente que a reflexão na superfície do estômago pode ser representada por um modelo dicromático. Uma lista de materiais que podem ser modelados com o modelo dicromático foi apresentada no item 4.3.3.

6.3 Reconstrução de Objetos

Utilizou-se para testes imagens sintéticas de tamanho 64x64 pixels e 256 níveis de cinza geradas por programa e imagens reais de objetos obtidas por um endoscópio. Abaixo, se encontram as imagens sintéticas.

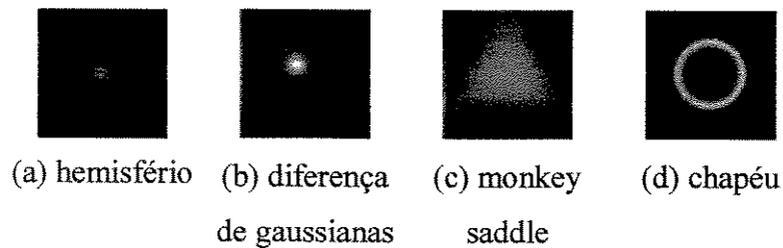
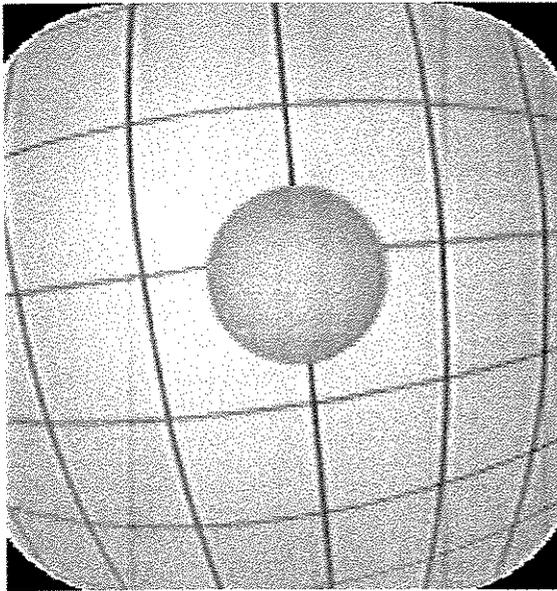


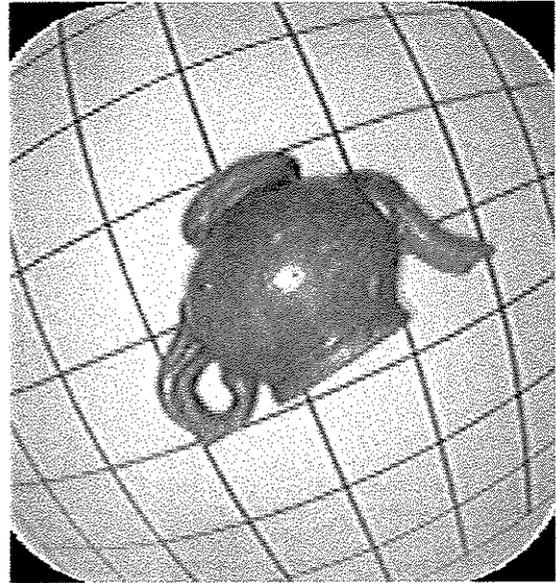
Figura 6-7 Imagens sintéticas utilizadas para teste

As imagens (a) a (d) da figura 6-7 foram geradas utilizando as equações (6.1) a (6.4), respectivamente, que serão apresentadas na seqüência deste capítulo, para gerar a forma geométrica e a equação (5.9) para obter a imagem de intensidade correspondente a uma superfície lambertiana.

Na figura 6-8, encontram-se as imagens de objetos reais. Essas imagens foram obtidas com um endoscópio a distância aproximada de 5 cm do ponto mais alto do objeto. Percebe-se pela grade desenhada como pano de fundo a ocorrência de distorção radial. O material do bule, cerâmica, segue o modelo dicromático, como pode ser visto em [Tominaga 1991]. Portanto, para o caso do bule, será utilizado o filtro para remoção da reflexão especular. No caso da bola de borracha cinza, a superfície foi suposta lambertiana. O modelo dicromático não poderia ser aplicado porque o vetor que representa a cor da fonte de luz é paralelo, ou muito próximo disso, ao vetor que representa a cor da bola em RGB. A dependência linear desses vetores impede o uso do método de eliminação da reflexão especular.



(a) bola de borracha



(b) bule de cerâmica

Figura 6-8 Imagens de objetos reais utilizadas nos testes

Nos próximos itens serão apresentados os casos de teste.

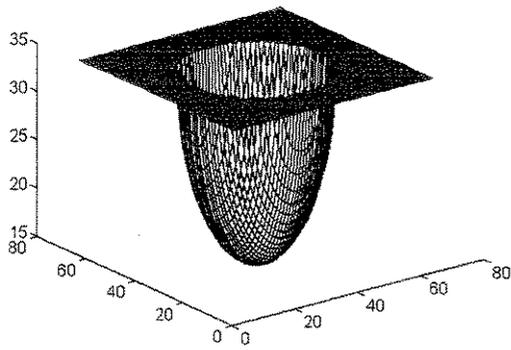
6.3.1 Hemisfério

A figura do hemisfério foi gerada pela equação

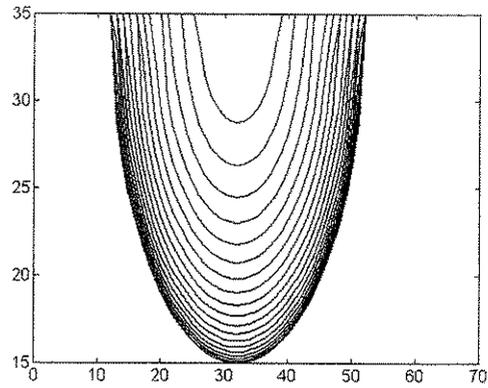
$$z = 35 - \sqrt{20 - u^2 - v^2} . \quad (6.1)$$

onde (u, v) são coordenadas da imagem.

Corresponde a um hemisfério no espaço tridimensional definido pelas coordenadas reais (u, v, z) . Os gráficos da figura 6-9 mostram a profundidade em cada ponto do objeto original de teste.



(a)

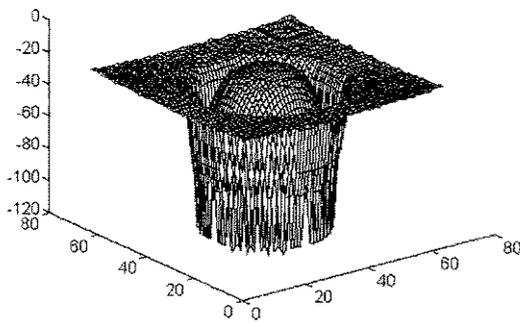


(b)

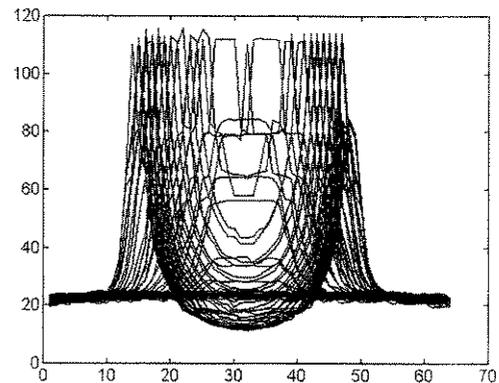
Figura 6-9 Forma geométrica do hemisfério. (a) Objeto em três dimensões. (b) Projeção no plano $v=0$.

A projeção no plano $v=0$ permite a visualização da qualidade da medida de profundidade.

Aplicando-se o algoritmo de reconstrução por otimização proposto na tese, obtiveram-se os gráficos da figura 6-10.



(a)



(b)

Figura 6-10 Hemisfério reconstruído por otimização. (a) Objeto em três dimensões, porém com eixo z invertido para melhorar a visualização. (b) Projeção no plano $v=0$.

Para o gráfico (a), o eixo z foi invertido a fim de permitir melhor visualização. Nota-se o efeito da descontinuidade na borda do hemisfério sobre a solução, elevando o erro. Os gráficos da figura 6-11 mostram a evolução da função objetivo e dos erros em imagem e profundidade.

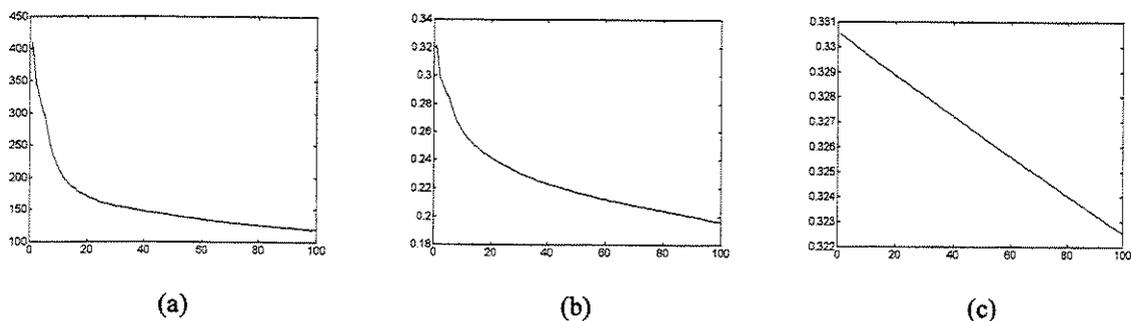


Figura 6-11 Evolução (a) da função objetivo, (b) do erro em imagem e (c) do erro em profundidade para o método baseado em otimização apresentado no item 5.2.

O primeiro gráfico (a) corresponde à função energia dada por $e = \sum_{ij \in \Omega} (\cos \sigma_{ij} - (\diamond r)_{ij})^2$. O eixo das abcissas corresponde ao número da iteração. O segundo gráfico (b) representa o erro em imagem, isto é, o erro quadrático médio entre a imagem original e a imagem gerada a partir do objeto reconstruído. O último gráfico (c) representa o erro em profundidade. O algoritmo foi interrompido após 100 iterações visto que a evolução do erro e da função objetivo estava sendo monitorada. Deve-se estipular um número máximo de iterações e um valor limite para que a função objetivo seja satisfatória como condição de término do algoritmo.

Utilizando, agora, o método de Okatani-Deguchi modificado (expansão de curvas), obtém-se os gráficos da figura 6-12.

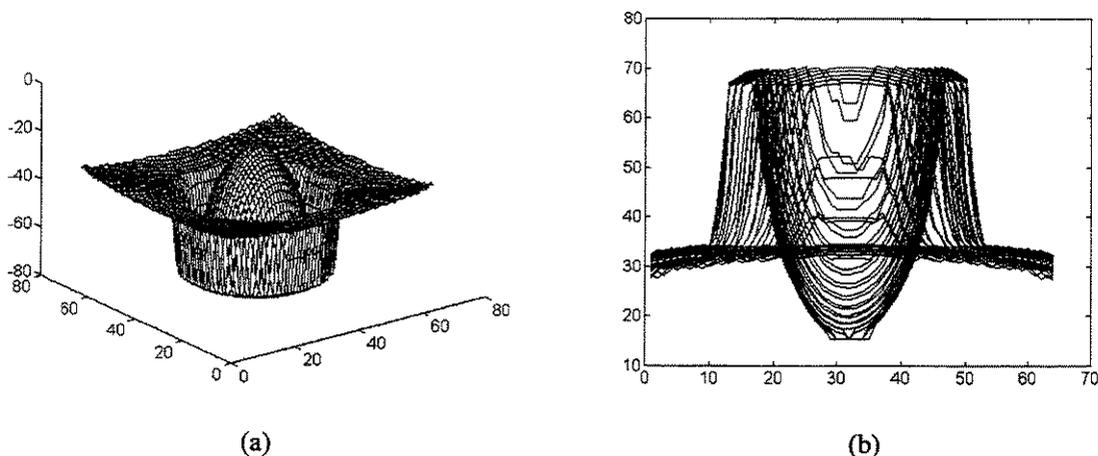


Figura 6-12 Hemisfério reconstruído por expansão de curvas. (a) Objeto em três dimensões. (b) Projeção no plano $v=0$.

A fim de comparar os algoritmos, mediram-se os erros quadráticos médios em profundidade. O método por otimização para a última iteração calculada (100) apresentou erro igual a 21 e o método por expansão de curvas apresentou erro igual a 18. comparando-se o valor dos erros com relação às proporções dos gráficos das figuras 6-9, 6-10 e 6-12, conclui-se que estes valores de erro são altos para se realizar medidas de profundidade. Porém, se o objetivo for simplesmente resgatar a forma dos objetos, estes erros são toleráveis. Uma análise do comportamento do erro em profundidade através de histogramas dos desvios dão uma idéia melhor dos resultados e são apresentados na figura 6-13 para o método baseado em otimização e para o método baseado em expansão de curvas.

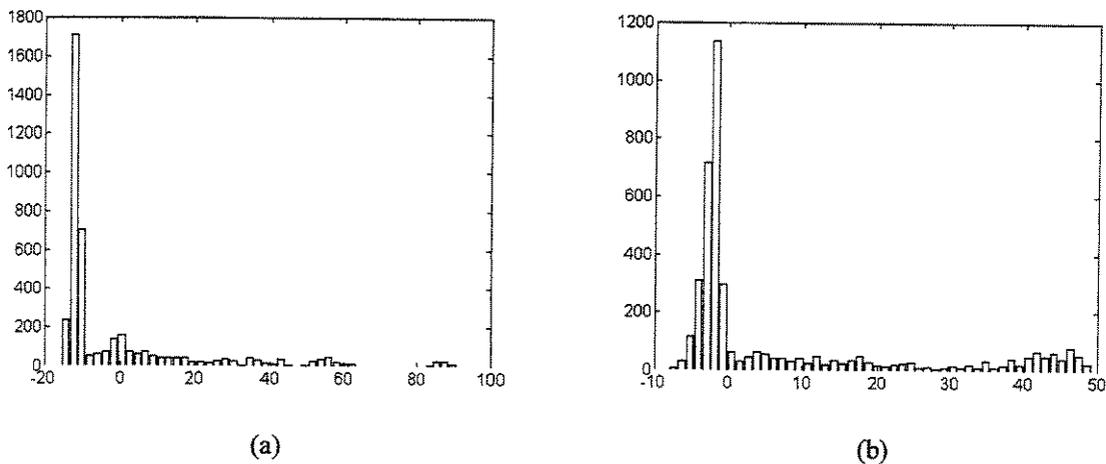


Figura 6-13 Histograma do desvio do mapa de profundidade para (a) o método baseado em otimização e para (b) o método de expansão de curvas.

Observa-se pelos histogramas, que os desvios do mapa de profundidade apresentam uma tendência a valores negativos, isto é, os valores estimados tendem a ser maiores do que os valores nominais. Essa tendência do desvio à uma direção determinada indica que a forma foi preservada.

6.3.2 Diferença de Gaussianas

O modelo da diferença de gaussianas foi obtido através da equação

$$z = 10 \exp\left(\frac{-(u-36)^2 - (v-36)^2}{720}\right) - 10 \exp\left(\frac{-(u-30)^2 - (v-30)^2}{216}\right) + 12. \quad (6.2)$$

A figura 6-14 mostra os gráficos do objeto criado dessa maneira:

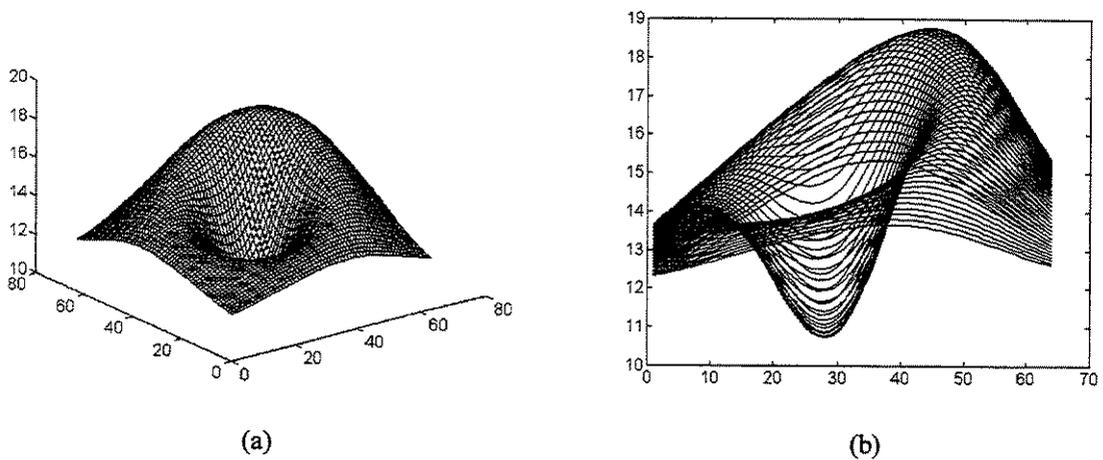


Figura 6-14 Forma geométrica da diferença de gaussianas. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

A reconstrução por otimização a partir da imagem da diferença de gaussianas (figura 6-7-b) resulta no mapa de profundidade representado pelos gráficos da figura 6-15.

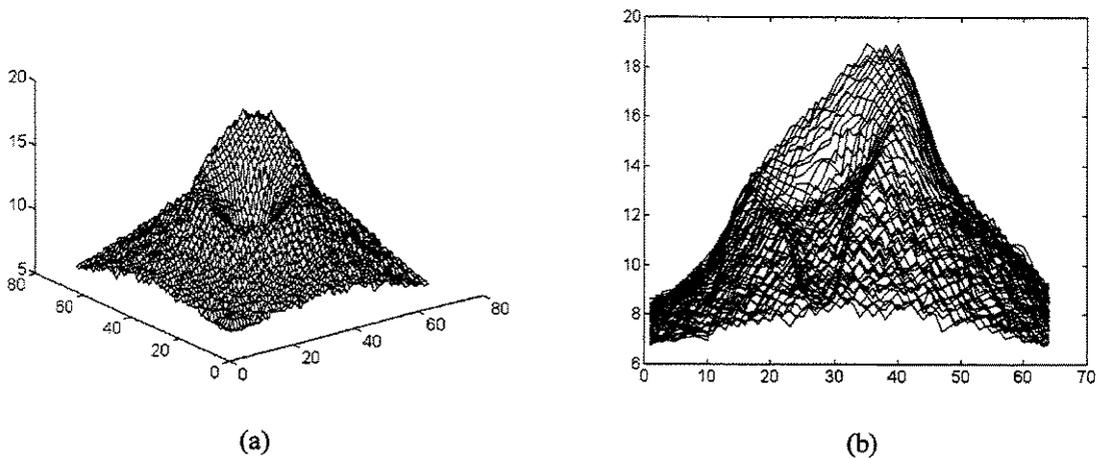


Figura 6-15 Diferença de gaussianas reconstruída por otimização. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

Na figura 6-16, apresentam-se os gráficos obtidos do algoritmo de expansão de curvas (Okatani-Deguchi modificado).

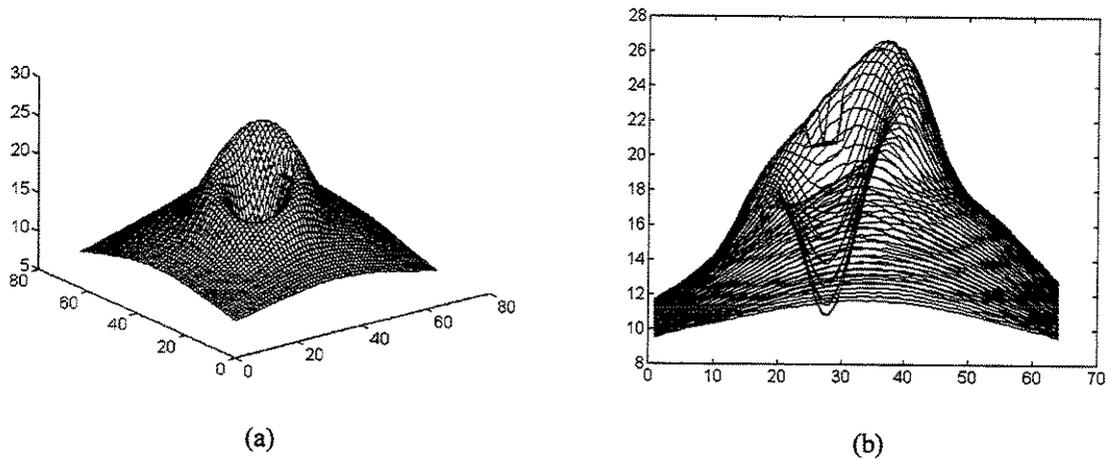


Figura 6-16 Diferença de gaussianas reconstruída por expansão de curvas. (a) Objeto em três dimensões. (b) Projeção no plano $v=0$.

A evolução da função de energia, do erro em imagem e do erro em profundidade para o método baseado em otimização são apresentados na figura 6-17.

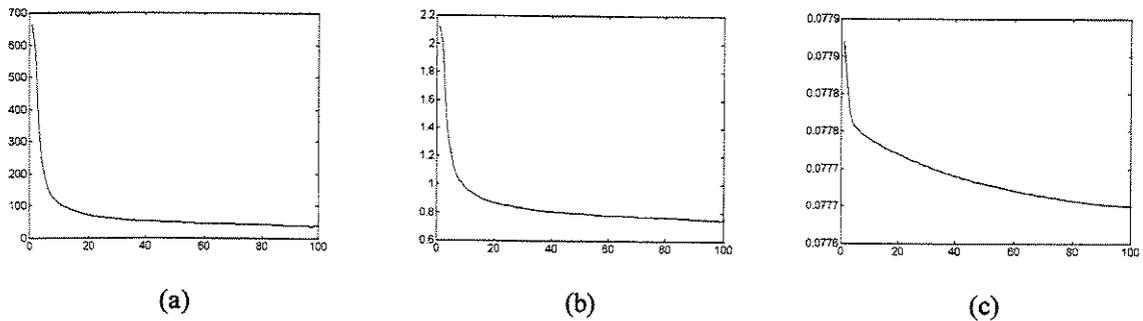
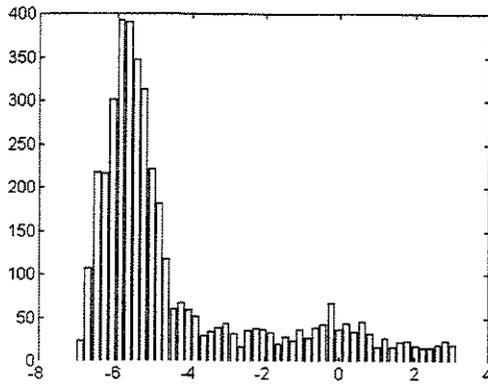


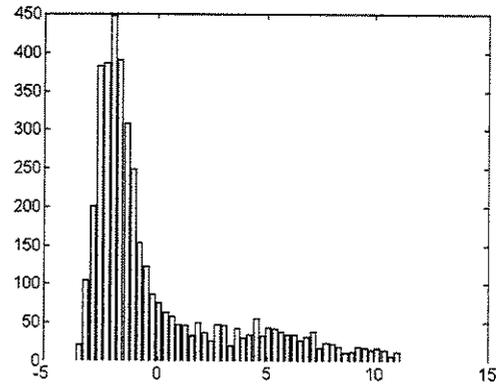
Figura 6-17 Evolução (a) da função objetivo, (b) do erro em imagem e (c) do erro em profundidade para o método baseado em otimização apresentado no item 5.2.

Comparando-se os métodos, obtém-se erro quadrático médio em profundidade de 5,0 para o método por otimização e 3,3 para o método por expansão de curvas.

Na figura 6-18, encontram-se os histogramas dos erros em profundidade para o método por otimização e para o método por expansão de curvas. Observa-se a mesma tendência do erro que ocorreu no caso anterior.



(a)



(b)

Figura 6-18 Histograma do desvio do mapa de profundidade para (a) o método baseado em otimização e para (b) o método de expansão de curvas.

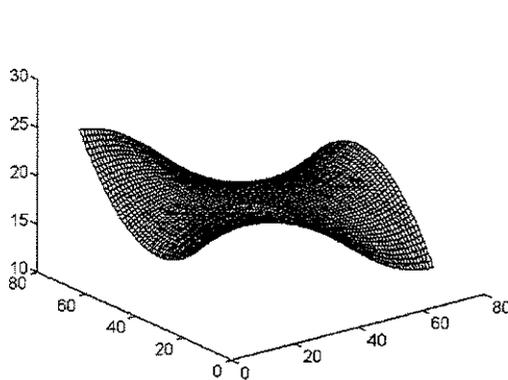
6.3.3 Monkey Saddle

Este modelo foi gerado pela equação

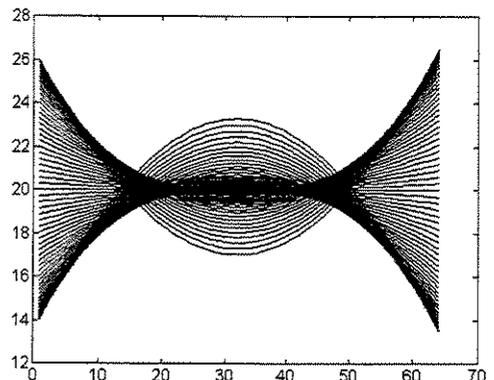
$$z = (u^3 - 3uv^2)/10000 + 20 \tag{6.3}$$

e contém um ponto de sela especial conhecido como *monkey saddle*.

Os gráficos que representam o objeto gerado pela equação (6.3) são mostrados na figura 6-19.



(a)



(b)

Figura 6-19 Forma geométrica da Monkey saddle. (a) Objeto em três dimensões.

(b) Projeção no plano $v=0$.

O objeto reconstruído pelo método por otimização é representado pelos gráficos da figura 6-20.

comportamento, incluem-se a convergência para um mínimo local e a arbitragem inadequada do coeficiente de regularização λ .

Na figura 6-22, a evolução da função energia, do erro em imagem e do erro em profundidade para o método por otimização são apresentadas.

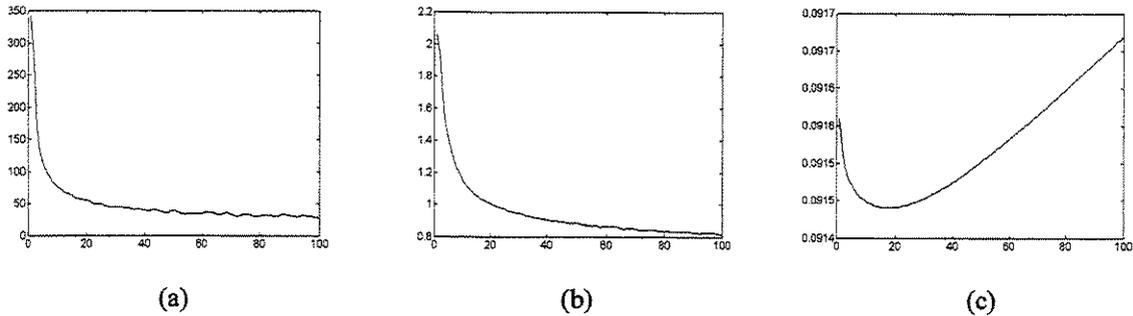


Figura 6-22 Evolução (a) da função objetivo, (b) do erro em imagem e (c) do erro em profundidade para o método baseado em otimização apresentado no item 5.2.

Os histogramas do erro em profundidade são dados na figura 6-23.

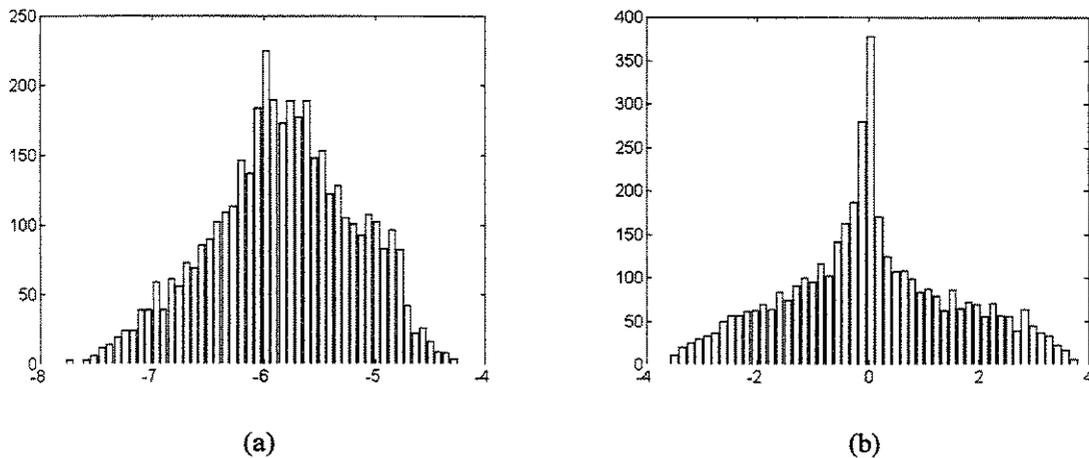


Figura 6-23 Histograma do desvio do mapa de profundidade para (a) o método baseado em otimização e para (b) o método de expansão de curvas.

6.3.4 Chapéu (diferença de gaussianas)

O objeto de teste foi gerado pela equação

$$z = 200 \exp\left(\frac{-(u-32)^2 - (v-32)^2}{100}\right) - 100 \exp\left(\frac{-(u-32)^2 - (v-32)^2}{400}\right) + 100, \quad (6.4)$$

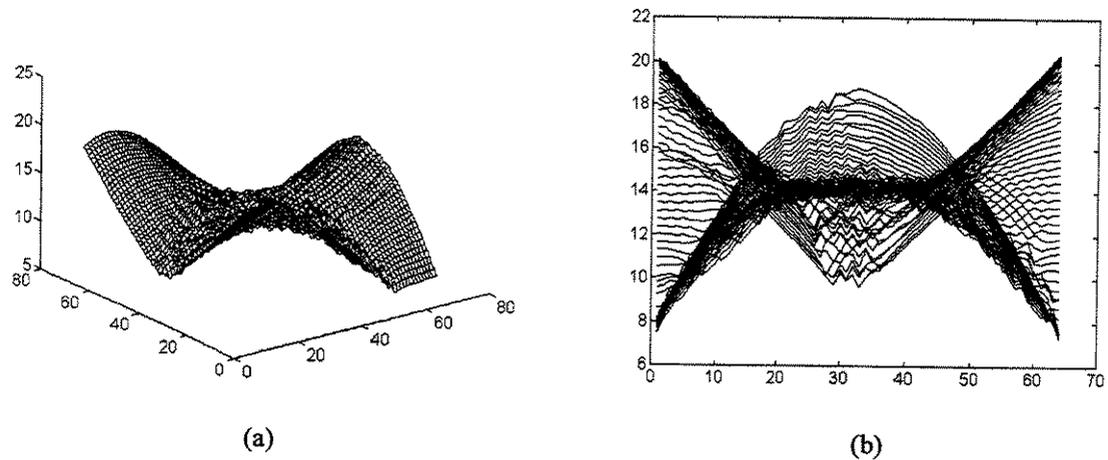


Figura 6-20 Monkey saddle reconstruída por otimização. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

O objeto reconstruído por expansão de curvas é mostrado na figura 6-21.

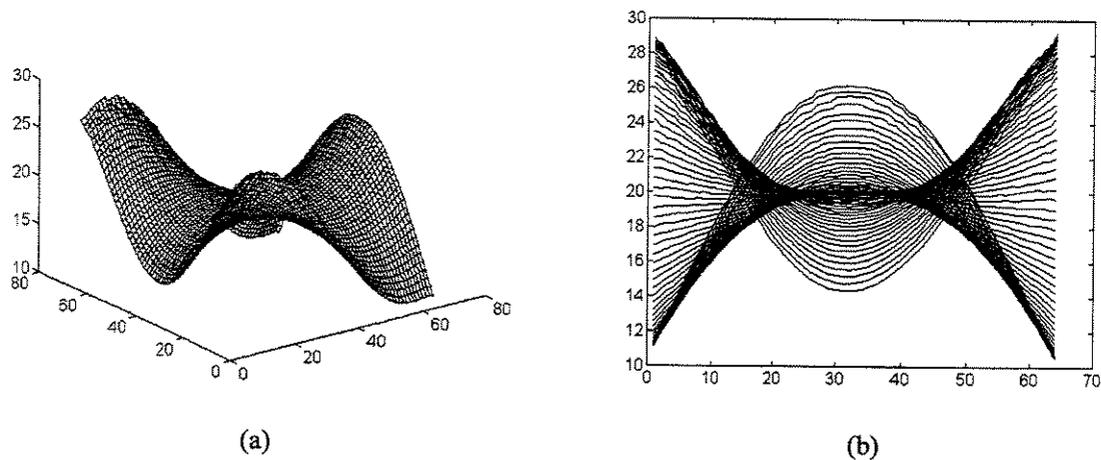


Figura 6-21 Monkey saddle reconstruída por expansão de curvas. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

Para o método por otimização, o erro quadrático médio obtido foi igual a 5,9. Para o método por expansão de curvas, obteve-se 1,5.

Observa-se da figura 6-20 que apesar de diminuir o erro em imagem a cada iteração, o método por otimização eleva o erro em profundidade. O algoritmo não tem controle do erro em profundidade, mas sim do erro em imagem. Entre os possíveis motivos para tal

que é uma diferença de gaussianas.

Os gráficos do mapa de profundidade do objeto gerado por essa equação são mostrados na figura 6-24.

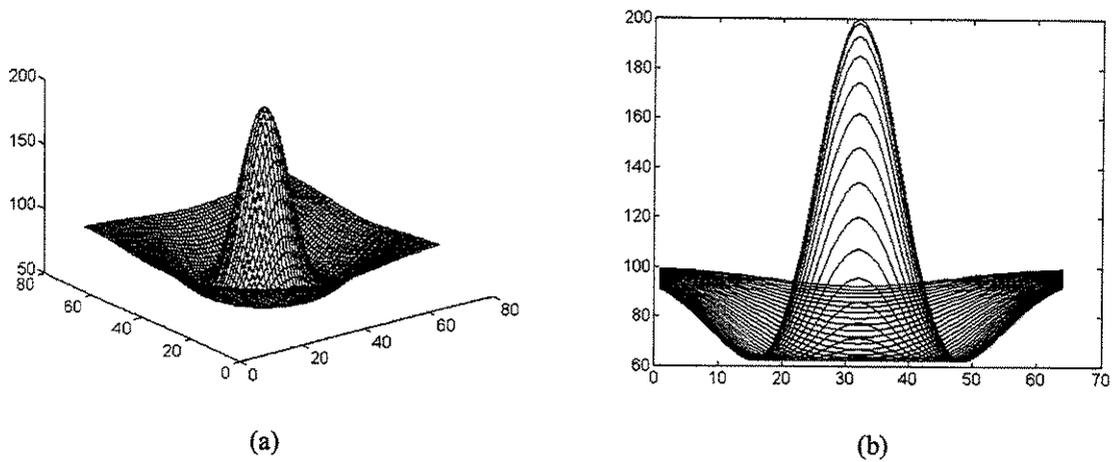


Figura 6-24 Forma geométrica do chapéu. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

Na figura 6-25 são mostrados os gráficos do objeto reconstruído pelo método por otimização. Na figura, 6-26, são mostrados os gráficos para a reconstrução pelo método baseado em expansão de curvas.

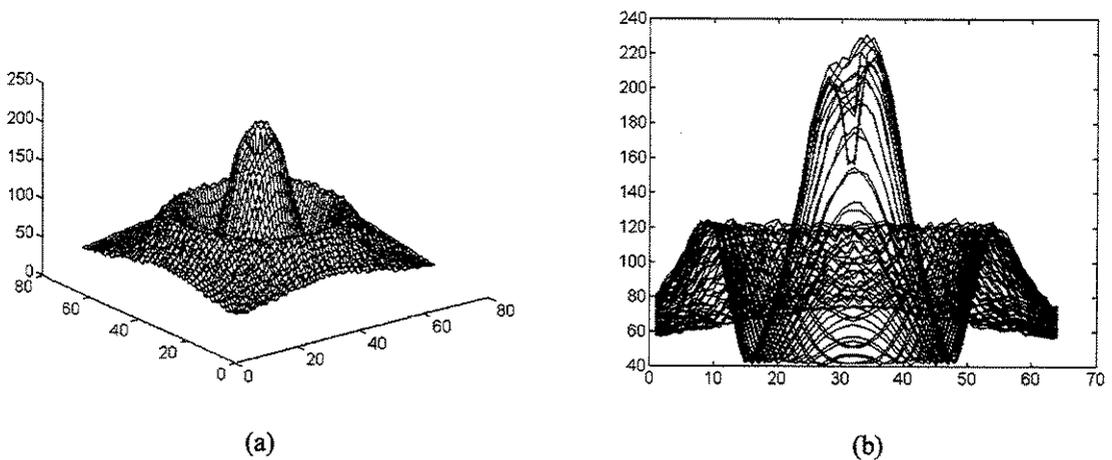


Figura 6-25 Chapéu reconstruído por otimização. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

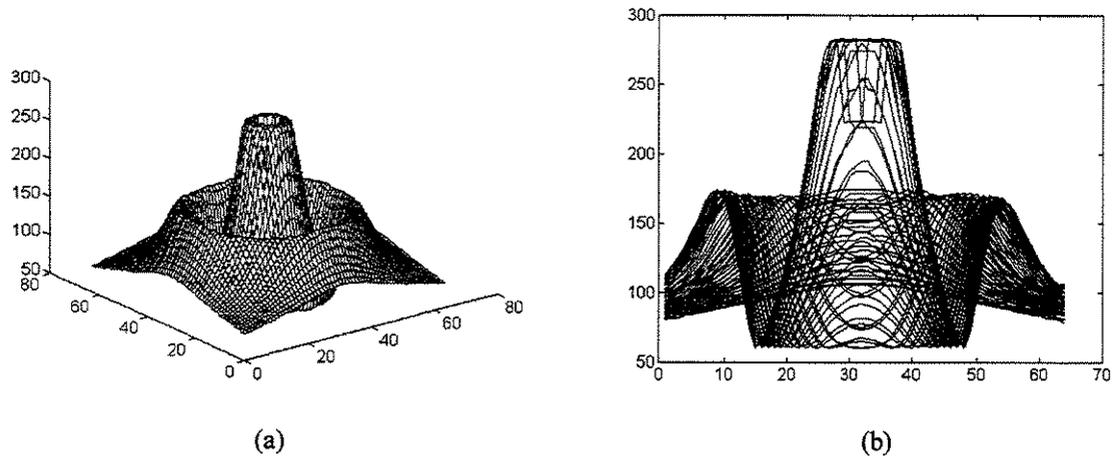


Figura 6-26 Chapéu reconstruído por expansão de curvas. (a) Objeto em três dimensões.
(b) Projeção no plano $v=0$.

Para o método por otimização, o erro quadrático médio da profundidade é 26. O erro para o método de expansão de curvas foi calculado como 54. Os histogramas dos erros são apresentados na figura 6-27.

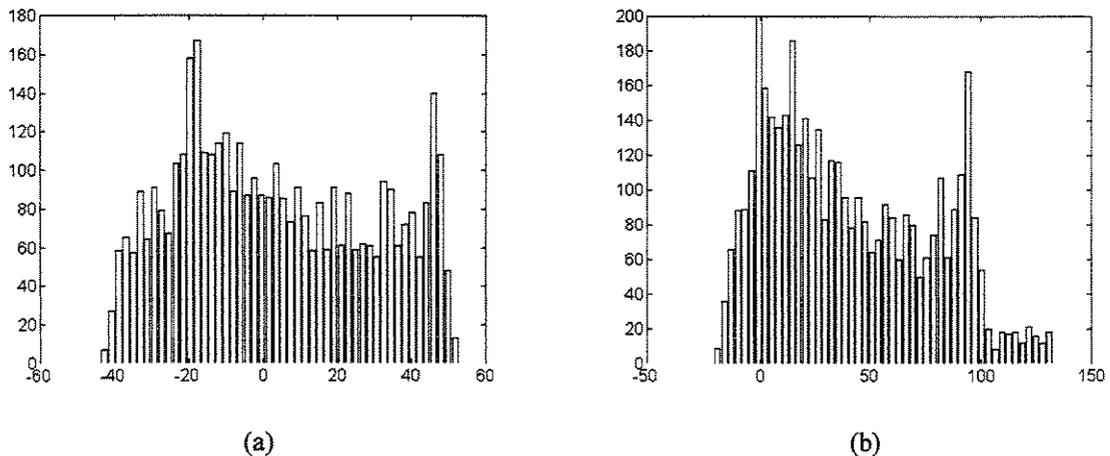


Figura 6-27 Histograma do desvio do mapa de profundidade para (a) o método baseado em otimização e para (b) o método de expansão de curvas.

6.3.5 Imagens de Objetos Reais

- ♦ **Bola de Borracha**

As imagens da bola de borracha e do bule de cerâmica foram obtidas com um endoscópio. A segmentação foi feita manualmente no caso da bola. Foi assumido que a

reflexão na bola é lambertiana. A imagem da bola segmentada manualmente é apresentada na figura 6-28.

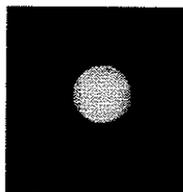


Figura 6-28 Imagem da bola de borracha

O resultado da reconstrução tridimensional para a bola de borracha utilizando o método baseado em otimização pode ser visto na figura 6-29.

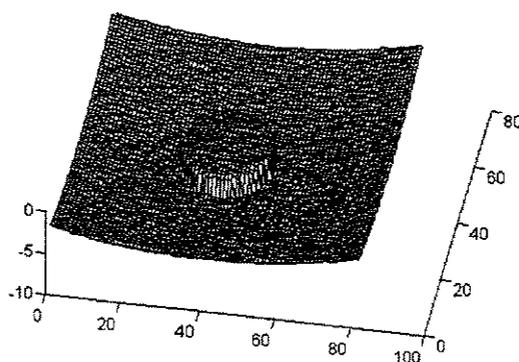


Figura 6-29 Bola de borracha reconstruída

A imagem da bola foi adquirida sob condições não muito apropriadas, em particular, sob presença de iluminação ambiente. Assim, é explicável o fato da reconstrução da bola aproximar-se da forma de uma superfície plana.

♦ Bule de Cerâmica

Para o bule da figura 6-8 (b), aplicou-se o filtro para remoção da reflexão especular. Em seguida, o valor 2 de intensidade foi atribuído aos pixels cuja intensidade era inferior a 2 para eliminar valores de intensidade próximos de zero. No caso, a intensidade do pixel pode variar de 0 a 255. Finalmente, aplicou-se o algoritmo de reconstrução por otimização.

O filtro para o modelo dicromático aplicado ao bule consiste da diferença entre o canal azul e o vermelho. De fato, como explicado no item 4.3.4, qualquer combinação linear $a_1 I_r + a_2 I_g + a_3 I_b$ tal que $a_1 + a_2 + a_3 = 0$ pode ser utilizada, desde que o valor de albedo seja corrigido. O resultado da filtragem pode ser vista na figura 6-30.

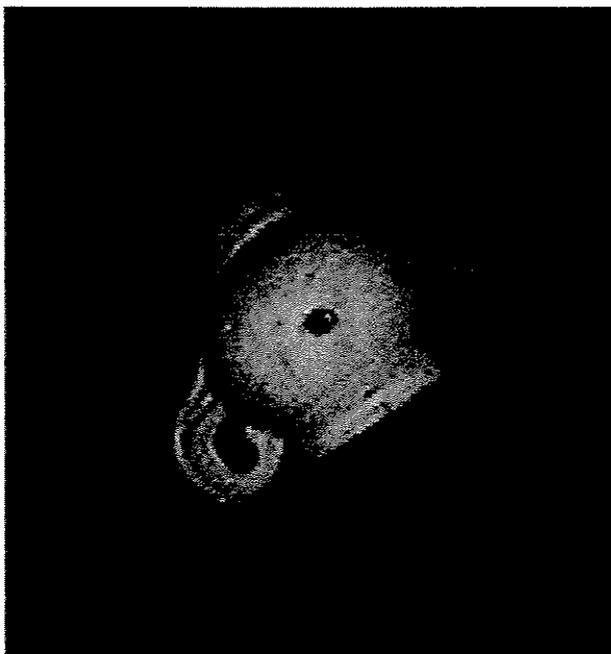


Figura 6-30 Imagem do bule com reflexão especular removida

O resultado da reconstrução pode ser visualizado na figura 6-31.

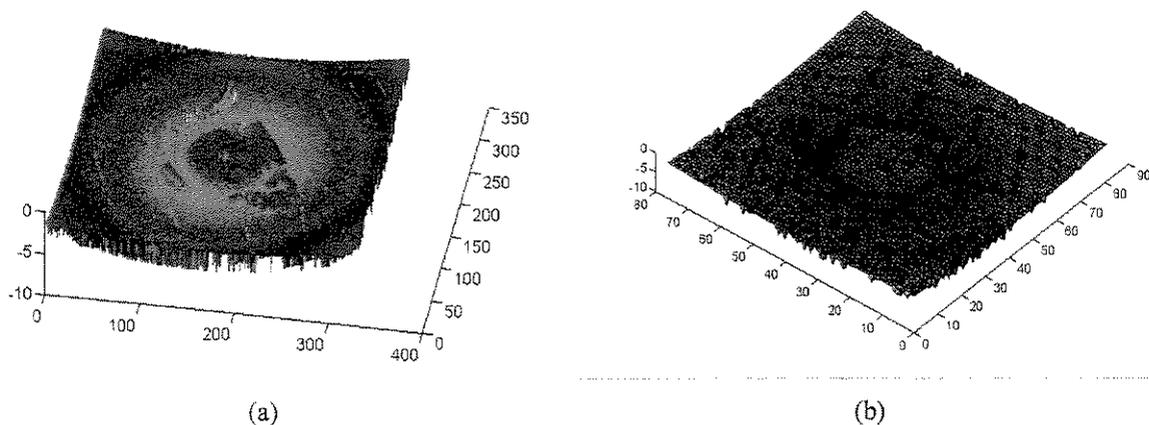


Figura 6-31 Reconstrução tridimensional do bule (a) representado utilizando uma cor para cada nível de profundidade e (b) representado através de uma malha poligonal.

6.4 Resultados em Imagens de Endoscopia

A imagem da figura 6-32 (a) foi obtida durante uma endoscopia e é uma imagem de estômago normal. O canal vermelho foi utilizado imaginando-se que a imagem seria lambertiana. A imagem reconstruída e apresentada na figura 6-32 (b) foi renderizada no POV-Ray a partir dos triângulos fornecidos com as coordenadas da reconstrução.

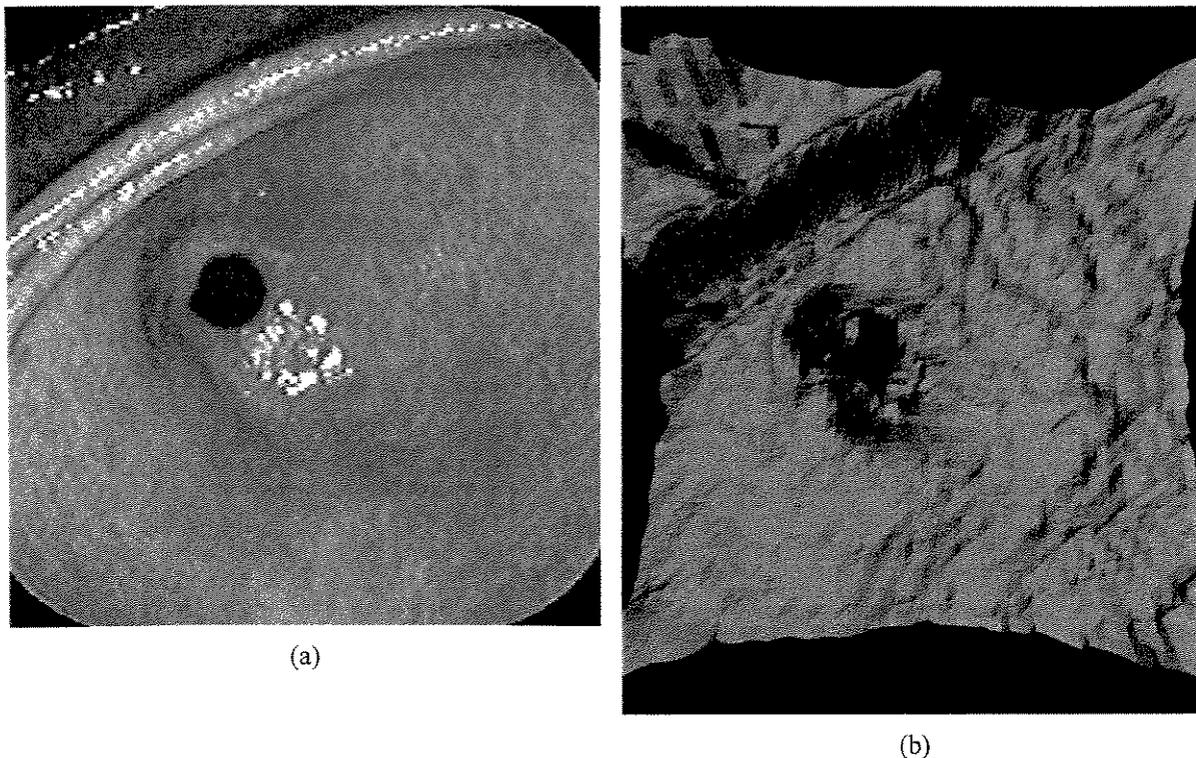


Figura 6-32 Reconstrução de uma imagem de endoscopia. (a) Imagem original. (b) Objeto reconstruído.

Observa-se que o algoritmo tratou as pregas na região superior da imagem como superfície côncava ao invés de convexa. Por outro lado, a reconstrução da região mais ao centro da imagem se mostrou bastante adequada, pois há preservação da forma que é percebida na imagem original.

Capítulo 7 - CONCLUSÃO

As contribuições deste trabalho consistem de:

1. Foi feita uma revisão dos métodos de Shape from Shading mais referenciados e suas variações a fim de encontrar um método adequado para se tratar imagens de endoscopia. Esta revisão pode ser útil na escolha de um método para uma situação específica, em que se aplique a técnica de Shape from Shading.
2. Abordagem mais prática do Shape from Shading. Na literatura, é difícil se encontrar resultados associados a uma abordagem prática, no sentido de que não é comum de se ver uma implementação de algoritmos de Shape from Shading para se resolver um problema prático.
3. Estudo de um modelo de câmera e iluminação para o Shape from Shading que pode ser utilizado para o caso da endoscopia e outros casos com câmera e fonte de luz próximas do objeto e um algoritmo que é consequência desse modelo.
4. Remoção de reflexão especular de imagens. Essa técnica permite aproveitar a informação de cor para se estimar o componente de reflexão difusa.

O problema de Shape from Shading é um problema complexo. O tratamento do problema resulta num sistema de equações diferenciais parciais não-lineares sujeito a problemas de instabilidade. É muito comum se referir ao problema de Shape from Shading como *ill-posed*. Métodos alternativos para resolver sistemas de equações diferenciais aplicados a esse problema são relatados na literatura, tais como otimização e expansão de

curvas. É comum, também, encontrar abordagens que buscam soluções aproximadas para o problema.

A abordagem deste projeto se direcionou à busca de uma solução que dispensasse a entrada de condições iniciais e que resolvesse o problema para câmera e fonte de luz próximos do objeto a se reconstruir. É muito conveniente dispensar a determinação de condições iniciais, porque há casos em que é muito complicado determiná-las. O método por otimização desenvolvido e o método de Okatani e Deguchi modificado atendem essas características para as imagens de teste. Apesar de apresentarem, em alguns casos, alto valor de erro em profundidade, o desvio é tendencioso a uma direção, preservando a forma da superfície. O alto valor de erro inviabiliza o uso em aplicações para medida de profundidade em que seja necessária uma precisão melhor, todavia não impede a determinação da forma da superfície. A causa do erro, para o método por otimização, é a necessidade de se arbitrar um valor para o termo de regularização, sem o qual não haveria convergência para uma solução adequadamente suave. A comparação dos resultados obtidos pelos dois métodos revela que o comportamento do erro é muito semelhante. Na maioria dos casos, a versão modificada do método de Okatani e Deguchi [1997] apresentou melhor erro e superfície reconstruída mais suave, em relação ao método baseado em otimização.

Entre as melhorias que se pode planejar para o método baseado em otimização, estão a busca de outras soluções mais eficiente como simulated annealing e gradiente conjugado ou técnicas para métodos iterativos como múltipla resolução [Dold e Eckmann, 1981].

Quanto à filtragem da reflexão especular em imagens coloridas, os resultado obtidos mostram que o funcionamento é adequado para imagens sintéticas, satisfatório para imagens de objetos de materiais cuja reflexão obedeça o modelo dicromático, porém é duvidosa a sua aplicabilidade para imagens reais de endoscopia, sendo necessário determinar experimentalmente as características de refletância da superfície do estômago. Em particular, não funcionou bem com as imagens de endoscopia utilizadas, onde se notou perda da informação de profundidade. Há várias hipóteses para o fracasso desse método, entre elas: o material do objeto pode não se conformar ao modelo dicromático, pode haver presença de interreflexões e a cor da fonte de luz pode estar incorretamente estimada (foi suposta branca).

Apêndice A - OUTROS ALGORITMOS DE SHAPE FROM SHADING

A.1 Shape from Shading por Análise Local

Este item visa ilustrar o que seriam os métodos locais de Shape from Shading, uma vez que foram comentados no texto, mas não foram utilizados.

A.1.1 Estimadores da Direção da Fonte de Luz

Na abordagem de Pentland [1984], chama-se \dot{E}_i a derivada direcional média (por toda região da imagem) de E na direção indexada por i , direção dada por (\dot{x}_i, \dot{y}_i) . Fazendo-se n observações para as diversas direções, monta-se o sistema

$$\begin{pmatrix} \dot{E}_1 \\ \vdots \\ \dot{E}_n \end{pmatrix} = \begin{pmatrix} \dot{x}_1 & \dot{y}_1 \\ \vdots & \vdots \\ \dot{x}_n & \dot{y}_n \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}$$

\hat{x} e \hat{y} podem ser obtidos através de uma regressão linear, utilizando a matriz pseudoinvertida.

O estimador de azimute da fonte de luz é dado por:

$$\tau_i = \arctan\left(\frac{\hat{y}}{\hat{x}}\right).$$

Considerando isotrópica a distribuição da normal à superfície, o estimador de ângulo polar da fonte de luz é dado por

$$\sigma = \arccos \sqrt{1 - \frac{\hat{x}^2 + \hat{y}^2}{|E\{\hat{E}^2\} - E\{\hat{E}\}^2|}}, \text{ onde } E\{\bullet\} \text{ é o operador esperança matemática.}$$

Lee e Rosenfeld [1985] obtiveram um método mais simples e acurado. O estimador de azimute da fonte de luz, segundo eles, é:

$$\tau_l = \arctan\left(\frac{E\{E_y\}}{E\{E_x\}}\right), \text{ calculando-se a esperança na região de interesse da imagem.}$$

O estimador de ângulo polar da fonte de luz considera *foreshortening* no cálculo da distribuição isotrópica da normal, e pode ser calculado numericamente através da expressão:

$$\frac{E\{E\}}{\sqrt{E\{E^2\}}} = 8 \frac{(\pi - \sigma_l) \cos \sigma_l + \sin \sigma_l}{3\pi(1 + \cos \sigma_l)^{3/2}},$$

cujos segundo membro é uma função monotonicamente decrescente. Observa-se que, na abordagem de Lee e Rosenfeld, utilizam-se apenas as derivadas de primeira ordem.

A.1.2 Estimadores da Normal à Superfície utilizando Estimador de Albedo

[Lee e Rosenfeld 1985] utilizaram o sistema de coordenadas da fonte de luz para definir os estimadores de azimute e ângulo polar da normal da superfície. Os eixos desse sistema são

$$\begin{aligned} \hat{x}' &= (\cos \sigma_l \cos \tau_l; \cos \sigma_l \sin \tau_l; -\sin \sigma_l)^T \\ \hat{y}' &= (-\sin \tau_l; \cos \tau_l; 0)^T \\ \hat{z}' &= (\sin \sigma_l \cos \tau_l; \sin \sigma_l \sin \tau_l; \cos \sigma_l)^T \end{aligned}$$

No sistema da fonte de luz, o estimador de azimute é dado por

$$\tau = \arctan\left(\frac{\bar{E}_y}{\bar{E}_x}\right), \text{ onde } \bar{\bullet} \text{ é considerado no sistema da fonte de luz.}$$

Nota-se que se deve, inicialmente, estimar a direção da fonte de luz, se esta não for conhecida.

Se o albedo da superfície for conhecido, basta utilizar

$$\sigma = \arccos \frac{E}{\rho}$$

Entretanto, existe um estimador de albedo que pode ser utilizado.

Considere um pixel P, um pixel Q definido na direção do azimute em P e um pixel R, tal que P fique entre R e Q. Aplicando o modelo de iluminação, tem-se

$$E_P = \rho(\vec{n}_P \cdot \vec{l}) = \rho \cos \theta$$

$$E_Q = \rho(\vec{n}_Q \cdot \vec{l}) = \rho \cos(\theta + \Delta\theta_1)$$

$$E_R = \rho(\vec{n}_R \cdot \vec{l}) = \rho \cos(\theta - \Delta\theta_2).$$

Aproximando $\Delta\theta_1 \approx \Delta\theta_2 = \delta$,

$$E_Q = \rho \cos(\theta + \delta) = \rho(\cos \theta \cos \delta - \sin \theta \sin \delta)$$

$$E_R = \rho \cos(\theta - \delta) = \rho(\cos \theta \cos \delta + \sin \theta \sin \delta).$$

Assim,

$$E_R + E_Q = 2\rho \cos \theta \cos \delta = 2E_P \cos \delta$$

$$E_R - E_Q = 2\rho \sin \theta \sin \delta$$

Com isso,

$$\cos \delta = \frac{E_R + E_Q}{2E_P}.$$

Logo,

$$\rho^2 = (\rho \sin \theta)^2 + (\rho \cos \theta)^2 = \left(\frac{E_R - E_Q}{2 \sin \delta} \right)^2 + E_P^2$$

e, portanto,

$$\rho^2 = \frac{E_P^2 - E_R E_Q}{E_P^2 - ((E_R + E_Q)/2)^2} E_P^2.$$

A.2 O Método Clássico de Expansão de Curvas Características de Horn

A formulação do problema de Shape from Shading através de um sistema de equações diferenciais parciais permite escolher métodos desenvolvidos por matemáticos para a solução dessas equações e empregá-los na solução de Shape from Shading. Um método

reconhecido é a expansão de curvas características. Se forem conhecidos todos os parâmetros ao longo de uma curva, pode-se utilizar a equação diferencial parcial para aproximar numericamente a próxima curva e os parâmetros ao longo desta.

Uma série de inconvenientes torna o método de curvas características pouco popular. Talvez, o inconveniente mais relevante seja que a discretização das curvas não ocorre da mesma forma que a imagem foi discretizada. É por esse motivo que o distanciamento das curvas características acaba deixando regiões da imagem inexploradas. E como consequência, quanto mais distante da curva inicial, maior o erro. Osher e Sethian [1988] demonstraram que a expansão de curvas discretizadas por pontos é numericamente instável se não for feita uma reamostragem a cada iteração. Outra adversidade é a topologia da superfície que não é considerada pelo método.

A.2.1 Formulação do Esquema

Considere um parâmetro arbitrário ξ . Derivando z em relação a ξ e aplicando a regra da cadeia, tem-se

$$\frac{\partial z}{\partial \xi} = p \frac{\partial x}{\partial \xi} + q \frac{\partial y}{\partial \xi}$$

Calculando, agora, as derivadas de p e q ,

$$\frac{\partial p}{\partial \xi} = r \frac{\partial x}{\partial \xi} + s \frac{\partial y}{\partial \xi}$$

$$\frac{\partial q}{\partial \xi} = s \frac{\partial x}{\partial \xi} + t \frac{\partial y}{\partial \xi}$$

onde r , s e t são as segundas derivadas de z . Precisamente,

$$r = \frac{\partial^2}{\partial x^2} z$$

$$s = \frac{\partial^2}{\partial x \partial y} z$$

$$t = \frac{\partial^2}{\partial y^2} z$$

Diferenciando $E = R$ em relação a x e y e aplicando a regra da cadeia, obtém-se

$$E_x = rR_p + sR_q$$

$$E_y = sR_p + tR_q$$

que pode ser escrita como

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix} \cdot \begin{pmatrix} R_p \\ R_q \end{pmatrix}.$$

Utilizando a matriz hessiana $H = \begin{pmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{pmatrix}$, pode-se escrever

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = H \begin{pmatrix} R_p \\ R_q \end{pmatrix} \quad \text{e}$$

$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = H \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}$$

Imagine uma curva parametrizada por ξ . Para um valor pequeno de $\delta\xi$, escolhe-se a variação de x e y como sendo

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} R_p \\ R_q \end{pmatrix} \delta\xi$$

e, portanto, multiplicando-se H em ambos os lados,

$$\begin{pmatrix} \delta p \\ \delta q \end{pmatrix} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} \delta\xi.$$

Sabe-se também que

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

O problema foi transformado em 5 equações diferenciais ordinárias, cujas soluções são curvas características que devem ser buscadas uma a uma a partir de condições de contorno conhecidas (silhueta e pontos singulares).

$$\begin{aligned} \dot{x} &= R_p & \dot{y} &= R_q & \dot{z} &= pR_p + qR_q \\ \dot{p} &= E_x & \dot{q} &= E_y \end{aligned}$$

A.2.2 Tratamento de Pontos Especiais

Pontos singulares e pontos especulares são utilizados como pontos de partida para a busca da solução. Pontos singulares são pontos em que a intensidade da reflexão difusa é máxima e portanto a normal à superfície aponta para a fonte de luz. Uma curva inicial pode ser obtida utilizando-se uma pequena circunferência em torno do ponto singular. Isto pode

ocorrer em máximos locais, mínimos locais e pontos de sela da superfície, no sistema de coordenadas da fonte de luz. Para pontos de sela, a construção da circunferência não se aplica. Pontos especulares são pontos em que a normal à superfície é paralela à bissetriz da direção do observador e da direção da fonte de luz. Nestes pontos, a intensidade é máxima e podem ser reconhecidos por terem cor próxima da cor da fonte de luz em superfícies que obedecem o modelo dicromático. São pontos em que se conhece a direção da normal e, portanto, podem ser úteis em alguns métodos.

A.2.3 Tratamento de Bordas

[Horn 1975] reconheceu seis tipos de borda presentes nas imagens de sombreamento:

1. Bordas de oclusão;
2. Bordas de sombra projetada;
3. Bordas de sombra própria;
4. Bordas de descontinuidade;
5. Bordas de descontinuidade da normal;
6. Bordas de ambigüidade.

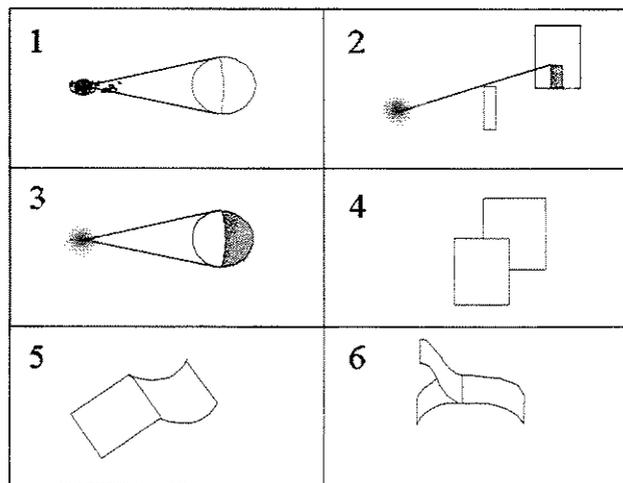


Figura A-1 Tipos de borda identificados por Horn.

Um exemplo de borda de ambigüidade ocorre com as funções:

$$z = z_0 + x^3$$

$$z = z_0 + |x|^3$$

Um outro tipo de borda que se pode considerar é similar a borda de sombra projetada, que é borda de variação de albedo. A profundidade e seu gradiente são contínuos, porém a refletância se altera.

Em [Horn 1975] encontram-se alguns detalhes de implementação como o uso de algoritmos de Runge-Kutta, preditor, modificador e corretor e tratamento de problemas na expansão das características, a saber:

- Espalhamento, deixando regiões da imagem sem explorar.
- Cruzamento das características base (projeção das curvas na imagem).
- Solução usando curvas de mesmo s (comprimento de arco)
- Curva inicial (ponto singular).

A.2.4 Problemas Topológicos

[Horn 1975] utilizou uma série de heurísticas para resolver os problemas topológicos na evolução das curvas.

- Se a distância de dois pontos vizinhos da curva ficar maior que 1,5 vezes o passo, criar novo ponto com média de x , y , z , p , q .
- Se a distância for menor que 0,7 vezes o passo, remover um ponto.
- Quebrar a curva se houver cruzamento de características base.

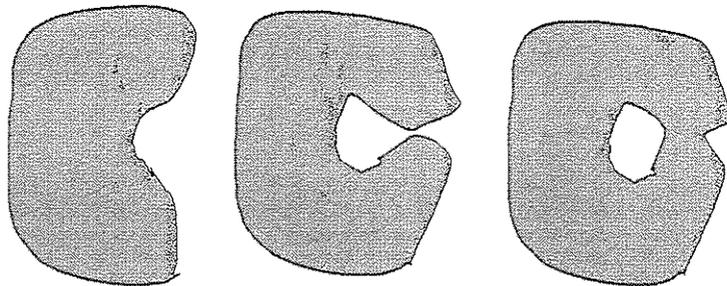


Figura A-2 Evolução de uma curva, mostrando alteração de topologia.

Esta heurísticas correspondem ao mecanismo de discretização das curvas características que é aplicado a cada nova curva determinada.

REFERÊNCIAS

1. Brooks, M. J. e Horn, B. K. P., 1985, Shape and Source from Shading, *Proceedings of the 1985 International Joint Conference on Artificial Intelligence*, Los Angeles, CA, August, Morgan Kaufmann Publishers, pp 932-936.
2. Brooks, M. J. e Horn, B. K. P., 1986, The Variational Approach to Shape from Shading, *Computer Vision, Graphics & Image Processing*, vol 33, no 2, pp 174-208.
3. Cryer, J. E., Tsai, P. S. e Shah, M., 1995, Integration of Shape from Shading and Stereo, *Pattern Recognition*, vol 28, no 7, pp 1033-1043.
4. Daum, M., Dudek, G., 1998, Out of the Dark: Using Shadows to Reconstruct 3D Surfaces, *Lecture Notes in Computer Science* , vol 1351, pp 72-79.
5. Dold, A., Eckmann B. eds., 1981, *Multigrid Methods*. Proceedings, Köln-Porz, *Lecture Notes in Mathematics*, vol 960, Springer-Verlag.
6. Foley, J., van Dam, A., Feiner, S., Hughes, J., 1990, *Computer Graphics - Principles and Practice*, Addison Wesley, Reading, Massachusetts.

7. Frankot, R. T. e Chellappa, R., 1988, A Method for Enforcing Integrability in Shape from Shading Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 10, no 4, pp 439-451.
8. Haar Romeny, B. M. ter ed., 1994, *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, Kluwer Academic Publishers.
9. Hasegawa, J. K. e Tozzi, C. L., 1996, Shape from Shading with Perspective Projection and Camera Calibration, *Computers & Graphics*, vol 20, no 3, pp 351-364.
10. Horn, B. K. P., 1975, Obtaining Shape from Shading Information, In *Shape from Shading*, pp 121-171.
11. Horn, B. K. P., 1990, *Robot Vision*, The MIT Press, Cambridge, Massachusetts.
12. Horn, B. K. P. e Brooks, M. J. editors, 1989, *Shape from Shading*, The MIT Press, Cambridge, Massachusetts.
13. Ikeuchi, K. e Horn, B. K. P., 1981, Numerical Shape from Shading and Occluding Boundaries, *Artificial Intelligence*, vol 17, no 1-3, pp 141-184.
14. Kimia, B., Tannenbaum, A., Zucker, S., 1994, On Optimal Control Methods in Computer Vision and Image processing, In Haar Romeny, B. M. ter ed., *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, Kluwer Academic Publishers, pp 307-338.
15. Kimmel, R., 1995, *Curve Evolution on Surfaces*, Sc. D. thesis, Technion Israel Institute of Technology.
16. Kimmel, R. e Bruckstein, A. M., 1995a, Global Shape From Shading, *Computer Vision and Image Understanding*, vol 62, no 3, pp 360-369.

17. Kimmel, R. e Bruckstein, A. M., 1995b, Tracking Level Sets by Level Sets: A Method for Solving the Shape from Shading Problem, *Computer Vision and Image Understanding*, vol 62, no 2, pp 47-58.
18. Lai, S. H. e Vemuri, B. C., 1997, Physically Based Adaptive preconditioning for Early Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 19, no 6, pp 594-607.
19. Leclerc, Y. G., Bobick, A. F., 1992, The Direct Computation of Height from Shading, *Proceedings of the Image Understanding Workshop*, San Diego, CA, January.
20. Lee, C. e Rosenfeld A., 1985, Improved Methods of Estimating Shape from Shading Using the Light Source Coordinate System, In *Shape from Shading*.
21. Lee, D., 1988, A Provably Convergent Algorithm for Shape from Shading, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 1988, Ann Arbor, MI, pp 478-485.
22. Lee K. M. e Kuo, C. C. J., 1994, Shape from Shading with Perspective Projection, *CVGIP: Image Understanding*, vol 59, no 2, pp 202-212.
23. Lee K. M. e Kuo C. C. J., 1997, Shape from Shading with a Generalized Reflectance Map Model, *Computer Vision and Image Understanding*, vol 67, no 2, pp 143-160.
24. Malik, J., Maydan, D., 1989, Recovering Three Dimensional Shape from a Single Image of Curved Objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 11, no 6.

25. Okatani, T. e Deguchi, K., 1997, Shape Reconstruction from an Endoscope Image by Shape from Shading Technique for a Point Light Source at the Projection Center, *Computer Vision and Image Understanding*, vol 66, no 2, pp 119-131.
26. Okatani, T. e Deguchi, K., 1998, On the Classification of Singular Points for the Global Shape from Shading Problem: A Study of the Constraints Imposed by Isophotes. *Lecture Notes in Computer Science*, vol 1351, pp 48-55.
27. Oliensis J. e Dupuis P., 1994, Shape from Shading: Provably Convergent Algorithms and Uniqueness Results, *Lecture Notes in Computer Science*, vol 801, Jan-Olof Eklundn editor, ECCV'94 vol II, Stockholm, Springer, pp 259-268.
28. Osher S. e Sethian J., 1988, Fronts Propagating with Curve-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations, *Journal of Computational Physics*, vol 79, pp 12-49.
29. Pentland, A., 1984, Local Shading Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 6 no 2, pp 170-187.
30. Poggio, T., Torre, V. e Koch, C., 1985, Computational Vision and Regularization Theory, *Nature*, vol 317, no 6035, pp 314-319.
31. Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992, *Numerical Recipes in C, The Art of Scientific Computing*, Second Edition, Cambridge University Press.
32. Šára, R., 1995, Isophotes: the Key to Tractable Local Shading Analysis, *Lecture Notes in Computer Science*, vol 970, pp 416-423.
33. Shimshoni, I., Kimmel, R. e Bruckstein A. M., 1996, Global Shape from Shading, *Computer Vision and Image Understanding*, vol 64 no 1, pp 188-189.

34. Strikwerda, J. C., 1989, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Brooks/Cole Advanced Books & Software.
35. Szeliski, R., 1991, Fast Shape from Shading. *CVGIP: Image Understanding*, vol 53, no 2, pp 129-153.
36. Tarantola, A., 1987, *Inverse Problem Theory, Methods for Data Fitting and Model Parameter Estimation*, Elsevier Science Publishers.
37. Tian, Y. e Tsui, H. T., 1997, Shape Recovery from a Color Image for Non-Lambertian Surfaces, *Journal of the optical Society of America A*, vol 14, no 2, pp 397-404.
38. Tominaga, S., 1998, Estimation of Reflection Parameters from a Color Image, *Lecture Notes in Computer Science*, vol 1351.
39. Tominaga, S., 1991, Surface Identification Using the Dichromatic Reflection Model, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 13, no 7, pp 658-670.
40. Torreão, J. R A., 1992, *Métodos Estocásticos em Computação Visual*, VIII Escola de Computação, Gramado RS.
41. Woodham, R. J., 1980, Photometric Method for Determining Surface Orientation from Multiple Images, *Optical Engineering*, vol 19, no 1, pp 41-50.
42. Zhang, R., Tsai, P. S., Cryer, J. E., Shah, M., 1994, Analysis of Shape from Shading Techniques, *IEEE CVPR-94*, Seattle, Washington.

43. Zheng, Q. e Chellappa, R., 1991, Estimation of Illuminant Direction, Albedo, and Shape from Shading, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 13, no 7, pp 680-702.

ÍNDICE REMISSIVO POR AUTORES

- Brooks e Horn 1985, 31
Brooks e Horn 1986, 32
Cryer *et al.* 1995, 42
Daum e Dudek 1998, 32
Dold e Eckmann 1981, 83, 105
Foley *et al.* 1990, 62, 66
Frankot e Chellappa 1988, 30, 38, 39, 40
Haar Romeny 1994, 83
Hasegawa e Tozzi 1996, 31
Horn 1975, 14, 30, 111, 112
Horn 1990, 17, 30, 32, 34, 36, 42, 51, 75, 79
Horn e Brooks 1989, 38
Ikeuchi e Horn 1981, 32, 36, 37, 75
Kimia *et al.* 1994, 48
Kimmel 1995, 43, 44, 49
Kimmel e Bruckstein 1995a, 30, 49
Kimmel e Bruckstein 1995b, 26, 30, 43, 44, 46, 47, 50, 51, 56, 75, 76
Lai e Vemuri 1997, 42
Leclerc e Bobick 1992, 30, 43
Lee e Rosenfeld 1985, 26, 55, 107

Lee 1988, 38
Lee e Kuo 1994, 30, 42
Lee e Kuo 1997, 42
Malik e Maydan 1989, 42
Okatani e Deguchi 1997, 17, 30, 56, 57, 58, 59, 75, 76, 77, 105
Okatani e Deguchi 1998, 51
Oliensis e Dupuis 1994, 30, 39, 48, 50, 51
Osher e Sethian 1988, 43, 45, 47, 48, 56, 78, 109
Pentland 1984, 14, 27, 55, 106
Poggio *et al.* 1985, 34
Press *et al.* 1992, 41, 83
Šara 1995, 55
Shimshoni *et al.* 1996, 52
Strikwerda 1989, 41
Szeliski 1991, 41, 42, 75, 79
Tarantola 1987, 31, 41, 79
Tian e Tsui 1997, 62, 70
Tominaga 1991, 70, 89
Tominaga 1998, 70, 71
Torreão 1992, 83
Woodham 1980, 36
Zhang *et al.* 1994, 56
Zheng e Chellappa 1991, 30, 40, 55