

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO E AUTOMAÇÃO INDUSTRIAL

Shape from Shading com Projeção Perspectiva e Calibração de Câmera

Autor: **Júlio Kiyoshi Hasegawa**

Orientador: **Prof. Dr. Clésio Luis Tozzi**

Tese submetida à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, para preenchimento dos pré-requisitos parciais para obtenção do Título de Doutor em Engenharia Elétrica.

agosto 1997

Este exemplar corresponde a redação final da tese
defendida por Júlio Kiyoshi Hasegawa
Julgada em 19/08/97 aprovada pela Comissão
Orientador CLT

H271s

31794/BC

UNICAMP
BIBLIOTECA CENTRAL

UNIDADE	BC
N.º CHAMADA:	
	T/Unicamp
	H271s
V	
Tombo BC/	31794
PELO	282/97
C	
D	☒
PREÇO	R\$ 11,00
DATA	17/10/97
N.º CPD	

CM-00101533-6

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

H271s

Hasegawa, Júlio Kiyoshi

Shape from shading com projeção perspectiva e
calibração de câmara / Júlio Kiyoshi Hasegawa.--
Campinas, SP: [s.n.], 1997.

Orientador: Clésio Luis Tozzi.

Tese (doutorado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Radiação. 2. Recuperação de dados (Computação).
3. Calibração. 4. Sistemas lineares. I. Tozzi, Clésio Luis.
II. Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. III. Título.

Resumo

A maioria dos algoritmos encontrados na literatura, para a reconstrução de uma superfície 3D a partir de imagens de intensidade, usando a técnica de *shape from shading* (SFS), utiliza o sistema de projeção ortográfico. Este fato, no caso de imagens reais, causa deformações na superfície reconstruída. Neste trabalho desenvolve-se um procedimento simultâneo de reconstrução da superfície e calibração da câmara, que utiliza as técnicas de SFS e técnicas fotogramétricas, considerando a projeção perspectiva. São utilizados modelos matemáticos explícitos (equação de *shading* e colinearidade) e é aplicado um processo iterativo para recuperação da informação 3D da superfície e dos parâmetros de orientação da câmara. Devido ao uso de equações não lineares no modelo a linearização por série de Taylor foi utilizada. O Método dos Mínimos Quadrados (MMQ) foi aplicado, para obter um valor único a partir de observações superabundantes. Estes procedimentos se diferenciam em relação aos citados na literatura pelo fato de se calcular simultaneamente as coordenadas tridimensionais dos pontos da superfície e os parâmetros de orientação da câmara (calibração), não sendo necessárias restrições sobre a superfície. Imagens de profundidade e de intensidade foram geradas (simuladas), onde o processo de reconstrução foi aplicado na imagem de intensidade, cujo resultado foi verificado, comparando-a com a imagem de profundidade com a finalidade de analisar a exatidão do método.

Abstract

The majority of algorithms making use of the Shape from Shading technique to reconstruct a 3 D surface from its image employs the orthographic projection system which, in the case of real images, always causes deformation on the reconstructed surface. To remove this constraint a simultaneous procedure (Hybrid Model) is proposed which uses the Shape from Shading technique (SFS) combined with the photogrammetric technique and the perspective projection to reconstruct the surface and determine the camera calibration parameters. The procedure is based on an explicit mathematical model and an iterative process is adopted to solve the resulting system of equations. The Least Square Method (LSM) is adopted and the Taylor's Series Linearization technique is applied due to the non-linearity of the model. The proposed procedure differs from those found in the literature in that it carries out the simultaneous computation of the tridimensional coordinates of both the surface points and the camera calibration parameters while employing perspective projection. In addition to that, solving the resulting system of equations does not demand any restriction on the surface, as it happens in other related procedures. The intensity image and its respective depth image were generated and the proposed model is applied on the intensity image. The results obtained are compared with the depth image to analyse the accuracy of the proposed method.

Agradecimentos

- . ao Prof. Dr. Clésio Luis Tozzi pela orientação, pelas proveitosas discussões e pelo apoio no decorrer do trabalho;
- . aos meus colegas Antônio M. G. Tommaselli e João F. Galera Monico pela leitura do trabalho e pelas preciosas sugestões;
- . aos meus colegas Maurício Galo e Maria de Lourdes B. T. Galo pela amizade, apoio e preciosas sugestões no desenvolvimento do trabalho;
- . aos meus colegas João F. C. da Silva e Arlete A. C. Meneguete pela leitura e sugestões;
- . aos demais professores, funcionários e alunos do DCA, que de uma forma ou de outra contribuíram para que este trabalho fosse concluído;
- . à Universidade Estadual Paulista, UNESP/FCT, pela concessão de afastamento integral durante o período de realização dos créditos e de desenvolvimento da tese;
- . aos companheiros do departamento de Cartografia, da UNESP/FCT, que tornaram possível este trabalho;
- . à coordenação de Aperfeiçoamento do Pessoal de Ensino Superior, CAPES, pela concessão de bolsa.

À

Mitiko,

Aline, Juliana e Toshi.

.....

Sumário

RESUMO	i
ABSTRACT	ii
AGRADECIMENTOS	iii
LISTA DE TABELAS	iv
LISTA DE FIGURAS	vi
1 Introdução	1
1.1 Considerações iniciais	1
1.2 Problemas da reconstrução	3
1.3 Solução Proposta	3
1.4 Objetivos do Trabalho	4
1.5 Tópicos do Trabalho	5
2 Fundamentação Teórica	7
2.1 Técnicas de " <i>Shape from X</i> "	7
2.1.1 A Técnica de SFS	7
2.1.2 A Técnica de <i>Shape from Texture</i>	9
2.1.3 A Técnica de <i>Shape from Contour</i>	10
2.1.4 A Técnica de <i>Shape from Motion</i>	12
2.1.5 A Técnica de <i>Shape from Stereo</i>	12

2.2	Formação Radiométrica da Imagem	15
2.2.1	Hemisfério de Iluminação e Ângulo Sólido	15
2.2.2	Modelo de Iluminação	17
2.3	Formação Geométrica da Imagem	20
2.4	Considerações Sobre a Superfície	22
2.5	Mapa de Refletância	24
2.5.1	Mapa de refletância com a fonte de luz próxima do observador	27
2.5.2	Mapa de refletância com a fonte de luz afastada do observador	29
2.5.3	Mapa de refletância para Superfícies Não Lambertianas	30
2.6	Considerações Sobre o Mapa de Refletância	31
2.7	Influência da Componente Especular	32
2.7.1	Modelo Matemático de Iluminação Especular	33
2.7.2	Modelo de Phong	34
2.8	Soluções Analisadas da Técnica de SFS	36
2.8.1	SFS com Projeção Ortogonal	36
2.8.2	SFS com Projeção Perspectiva	40
2.9	Técnica utilizada por Lee e Kuo	41
2.9.1	Desenvolvimento Teórico	41
2.9.2	Formulação Matemática	42
3	Modelo Proposto	45
3.1	As equações de colinearidade	45
3.2	Equação de <i>Shading</i>	49
3.3	Modelo Matemático de Reconstrução Proposto	51
3.4	Considerações sobre as Bordas	56
3.4.1	Bordas de Sombreamento	56
3.4.2	Bordas de Oclusão	56
3.4.3	Caracterização das Bordas	57
3.4.4	Detecção de Bordas	58

3.5	Método dos Mínimos Quadrados (MMQ)	61
3.5.1	Matriz Variância-Covariância	66
3.6	Injunções para a Calibração da Câmara	67
3.7	Injunções Relativas às Bordas de Oclusão e Sombreamento	70
3.8	Solução Implementada	71
3.8.1	Vetorização da matriz	75
4	Experimentos e Resultados	77
4.1	Imagens Simuladas	77
4.1.1	Experimentos com Imagens de uma superfície Plana	78
4.1.2	Experimentos com imagens simuladas de uma Esfera	81
4.1.3	Experimentos com Imagens Simuladas de um Cilindro	93
4.1.4	Experimentos com ruídos nas Imagens de intensidade	100
4.2	Análise dos resultados	104
4.2.1	Análise dos resultados da superfície plana	104
4.2.2	Análise dos resultados para as superfícies esféricas e cilíndricas	105
5	Conclusões	107
5.1	Síntese do Trabalho	107
5.2	Conclusões	108
5.3	Recomendações	110
A	Derivadas Parciais	111
A.1	Derivadas Parciais das Equações de Colineridade	111
A.2	Derivadas Parciais da Equação de <i>Shading</i>	113
	Referências Bibliográficas	122

Lista de Tabelas

4.1	Média e Desvio Padrão dos erros para as coordenadas nos testes de 1 a 6, obtida de Hasegawa e Tozzi [16]	80
4.2	Tabela com os valores dos parâmetros de calibração da câmara	81
4.3	Parâmetros utilizados nas imagens geradas variando a posição da luz, com distância de projeção de 1800 mm	85
4.4	Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 5400 mm	85
4.5	Parâmetros utilizados nas imagens geradas de esferas com variação da direção da luz (Z_v e $Z_l = 1800\text{mm}$)	88
4.6	Parâmetros utilizados nas imagens geradas de esferas com variação da direção da luz (Z_v e $Z_l = 5400\text{ mm}$)	88
4.7	Médias e Desvios-Padrão dos erros das coordenadas nos testes (1 a 7) realizados	89
4.8	Médias e Desvios-Padrão dos erros das coordenadas nos testes (8 a 14) realizados	91
4.9	Médias e Desvios-Padrão dos erros das coordenadas nos testes (15 a 28) realizados	92
4.10	Médias e Desvios-Padrão dos erros das coordenadas nos testes (29 a 42) realizados	93
4.11	Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 1800 mm	95
4.12	Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 5400 mm	95
4.13	Parâmetros utilizados na geração de imagens variando-se a direção da fonte de luz (Z_v e $Z_l = 1800\text{ mm}$)	97
4.14	Parâmetros utilizados na geração de imagens com variação da direção da luz (Z_v e $Z_l = 5400\text{ mm}$)	98
4.15	Médias e Desvios-Padrão dos erros das coordenadas nos testes (1 a 7) realizados	98

4.16 Médias e Desvios-Padrão dos erros das coordenadas nos testes (8 a 15) realizados 99

4.17 Médias e Desvios-Padrão dos erros das coordenadas nos testes (16 a 29) realizados101

4.18 Médias e Desvios-Padrão para as coordenadas nos testes (30 a 43) realizados 101

4.19 Médias e Desvios-Padrão para as coordenadas nos testes (1 a 14 - da superfície
esférica) realizadas em imagens com ruídos 102

4.20 Médias e Desvios-Padrão para as coordenadas nos testes (1 a 14 - superfície
cilíndrica) realizadas em imagens com ruídos 103

Lista de Figuras

2.1	Esquema de um par estereoscópico nivelado em relação ao sistema de coordenadas referencial	13
2.2	Mapa de refletância com três contornos de intensidade constante	15
2.3	Projeção de uma fonte de luz circular sobre o hemisfério de iluminação, adaptado de Hall [13]	16
2.4	Sistema de formação da imagem, adaptado de Horn e Sjoberg [21]	19
2.5	Sistema de coordenadas uv da tela	21
2.6	(a) Superfície rugosa - (b) Superfície não homogênea	23
2.7	Caminhos da luz e ângulos de fase - Shading	24
2.8	Sistema de coordenadas esféricas com coordenada Z passante pelo pólo.	26
2.9	Geometria do sistema de iluminação Lambertiana - fonte e observador na mesma direção	28
2.10	Mapa de refletância com a posição do observador e da fonte de luz coincidentes com $\rho = 1$	29
2.11	Mapa de refletância $R(p,q)$ - Com a fonte de luz ($p_l = 0.7, q_l = 0.3$ e $\rho = 1.0$), obtido de Horn [17].	30
2.12	Mapa de refletância para superfície Lunar de intensidade constante.	31
2.13	(a) ambigüidade provocada quando a fonte de luz está próxima do observador; (b) ambigüidade provocada quando a fonte de luz afastada do observador, adaptado de Kimmel e Bruckstein [27].	32
2.14	(a) Superfície especular maior intensidade de reflexão; (b) Superfície especular menor intensidade de reflexão.	33
2.15	Mapa de refletância para superfície especular, obtido de Horn e Sjoberg [21]	34
2.16	Sistema de iluminação de Phong	35

3.1	Sistema de coordenadas imagem e objeto paralelos, com origem no centro perspectivo	46
3.2	Sistemas de coordenadas da imagem e do objeto transladado e rotacionado .	47
3.3	Modelo de iluminação considerando-se a superfície Lambertiana, fonte de luz pontual e projeção perspectiva.	50
3.4	Esquema que define a correspondência entre o <i>patch</i> no espaço objeto e no espaço imagem.	52
3.5	Vetores que definem a normal à superfície	53
3.6	(a) Representação de uma estrutura 3x3 pontos, resultando em 8 <i>patches</i> tri- angulares; (b) Representação esquemática de uma estrutura de nxn pontos. .	56
3.7	Silhueta gerada pela borda com observador e fonte de luz na mesma posição.	57
3.8	Silhueta gerada por uma borda com observador e fonte de luz em posições diferentes	58
3.9	Operadores de Sobel; (a) máscara para calcular gradiente em x; (b) máscara para calcular gradiente em y;	59
3.10	Estrutura da matriz das equações normais N somente com as equações de colinearidade.	74
3.11	Estrutura da matriz das equações normais N em relação às equações de coli- nearidade e de <i>shading</i>	74
4.1	Imagem de um plano gerada com projeção perspectiva	78
4.2	Imagem de uma esfera gerada com projeção perspectiva	82
4.3	Representação <i>wire frame</i> construída a partir da imagem de profundidade, gerada da superfície esférica	86
4.4	Representação <i>wire frame</i> construída a partir da imagem de profundidade ger- ada, da superfície esférica, com o raio de 80% do tamanho do quadro da imagem	87
4.5	Representação <i>wire frame</i> construída a partir da imagem de profundidade ger- ada, da superfície esférica, com o raio de 90% do tamanho do quadro da imagem	87
4.6	Curva gerada a partir dos desvios-padrão dos erros para imagens esféricas com distância de 1800 mm de projeção.	89
4.7	Curva gerada a partir dos desvios-padrão dos erros para imagens esféricas com distância de 5400 mm de projeção.	91
4.8	Imagem de um cilindro gerada com projeção perspectiva	94

4.9 Representação *wire frame* construída a partir da imagem de profundidade, gerada da superfície cilíndrica, com o raio com 80% do quadro da imagem . . 96

4.10 Representação *wire frame* construída a partir da imagem de profundidade gerada, da superfície cilíndrica, com o raio com 90% do quadro da imagem . . . 97

4.11 Curva gerada a partir dos desvios-padrão dos erros para imagens cilíndricas com distância de 1800 mm de projeção. 99

4.12 Curva gerada a partir dos desvios-padrão dos erros para imagens cilíndricas com distância de 5400 mm de projeção. 100

Capítulo 1

Introdução

1.1 Considerações iniciais

A determinação da forma de um objeto usando observações da luz refletida por sua superfície é considerada um problema de reconstrução, e a principal dificuldade na determinação da superfície do objeto por meio desta técnica é a diferença na dimensionalidade dos espaços envolvidos. Dado que a imagem de um objeto é representada num espaço bidimensional, a reconstrução do objeto através de sua imagem requer uma transformação do espaço bidimensional (2D) para o espaço tridimensional (3D).

Assim, o paradigma 2D/3D pode ser entendido como uma questão de relacionamento de uma superfície no espaço 3D com sua imagem no espaço 2D. Neste contexto, pode-se considerar a superfície física como sendo composta por duas informações de naturezas distintas: parcela bidimensional (posição planimétrica) e profundidade. Por outro lado, na imagem, pode-se dispor de dois tipos de informações: a posição do *pixel* na imagem e a intensidade dos pontos. Partindo deste ponto de vista, a questão da reconstrução, pode ser resolvida concebendo duas parcelas distintas na imagem: a geométrica e a radiométrica. A porção geométrica da imagem define a posição planimétrica de cada ponto (pixel) no plano em que ela é formada e a sua porção radiométrica proporciona, através da sua luminosidade, a identificação destes pontos na imagem. Assim, utilizando-se destas concepções pode-se aplicar um modelo matemático de reconstrução, relacionando as informações 3D do espaço objeto com as informações 2D do espaço imagem.

Vários procedimentos de reconstrução têm sido desenvolvidos, dependendo do tipo de informação ou condição que é imposto na solução do problema. Desta forma, com o conhecimento de algumas características óticas do objeto imageado e o valor de brilho (tons

de cinza) de cada pixel, pode-se recuperar a orientação ou reconstruir os *patches* da superfície. Técnicas que utilizam a intensidade de luz refletida pelo objeto e captada pelo sistema de imageamento como um dos parâmetros para a determinação da superfície do objeto, podem ser denominadas, genericamente como, fotométricas¹ ou *Shape from Shading* (SFS).

No contexto da exploração do *shading* pode-se dizer que a geração de imagens (intensidade) a partir de modelos de objetos conhecidos é relativamente fácil de ser resolvida, geralmente utilizada em computação gráfica (*rendering* e *ray tracing*). Entretanto, a obtenção da forma (reconstrução 3D) a partir da imagem de intensidade (2D) é de difícil exploração, visto que a transformação 2D-3D é indeterminada (não inversível). Conseqüentemente, no processo de reconstrução a partir de imagens de intensidade, algumas informações não podem ser recuperadas.

Por seu interesse, principalmente na área de visualização e reconstrução de superfícies, o problema de SFS tem recebido grande atenção e inúmeras publicações com propostas metodológicas para a sua solução têm sido encontradas na literatura. Uma das primeiras soluções para o problema de SFS foi apresentada por Horn, segundo Ballard e Brown [2], e outros ([6], [18], [20]).

A partir da formulação básica de Horn, várias outras propostas foram apresentadas aplicando as mais variadas ferramentas para resolver a questão, tais como: técnica do cálculo variacional ([5], [18], [21]); métodos estocásticos ([29], [32], [41]); e recuperação da altura de objetos junto com o gradiente ([39]). Todas estas abordagens, conforme Sugihara [38], buscam a solução com base na redução dos graus de liberdade da superfície (3 graus para o caso da superfície plana).

Pode-se verificar que quase todas as soluções mostradas na literatura impõem restrições, quer pela necessidade de conhecimento prévio da cena, de modo a prover uma solução matemática única para a orientação da superfície, quer pela utilização de projeção ortográfica ao invés da projeção perspectiva. Poucos trabalhos sobre SFS considerando projeção perspectiva são encontrados na literatura ([34], [33], [30], [17]).

No caso de imagens reais, erros significativos podem ser provocados pelo fato de se adotar o sistema de projeção ortográfico ao invés de perspectivo, visto que, este último é a base de formação de imagens na maioria dos dispositivos de aquisição. Esses erros podem ser tanto mais acentuados quanto mais próximo estiver o dispositivo de aquisição do objeto, ou quando existirem grandes variações nas dimensões do objeto ao longo do eixo ótico (profundidade).

¹Segundo Bernd [24], a fotometria refere-se à resposta do sistema visual humano à radiação refletida pelo ponto.

1.2 Problemas da reconstrução

Como já mencionado, o processo de reconstrução de superfícies usando imagens digitais é problemático, mais ainda quando se utiliza de uma única imagem. Neste caso, verifica-se que, tanto na técnica de SFS quanto na fotogramétrica, há inconsistências matemáticas no processo de solução com uma única imagem.

O grande problema dos modelos que determinam a forma do objeto está na necessidade de uma definição prévia da superfície, ou seja, devem ser feitas várias considerações e restrições para se resolver a questão. Matematicamente, pode-se verificar que há uma única equação para duas incógnitas (p e q), problema este que só pode ser solucionado com mais uma equação de condição. Por outro lado, os modelos de determinação da superfície (coordenadas X , Y e Z) estão ainda, em alguns casos, vinculados a algum conhecimento prévio da superfície, ou como no caso de Lee and Kuo [30], na impossibilidade de se trabalhar com um sistema de referencial global e a desconsideração de uma fonte de luz pontual.

O sistema de iluminação é de fundamental importância no processo de reconstrução por SFS pois, na maioria dos ambientes, utiliza-se um sistema de iluminação artificial: lâmpada incandescente, *abajur*, *spots* e outros. O padrão de iluminação gerado no meio ambiente é de difícil modelagem mas, genericamente, pode-se representar os dispositivos artificiais e naturais por um modelo matemático simplificado ou aproximado da luz real.

Com base no desenvolvimento realizado por Hasegawa e Tozzi [17], foi observado que a qualidade da forma inicial da superfície pode ser enriquecida com a inclusão de informações adicionais, como por exemplo, pontos singulares, pontos especulares e silhuetas (mais especificamente as bordas de oclusão e sombreamento). Tal procedimento permite a redução da ambigüidade na reconstrução da superfície, a aceleração da convergência e a diminuição do erro de reconstrução.

Desse modo, a fim de solucionar os problemas, tanto em SFS, como na fotogrametria, além de algumas considerações (a nível de definições dos modelos matemáticos), devem ser impostas algumas condições iniciais.

1.3 Solução Proposta

Geralmente, o sistema de iluminação utilizado nos procedimentos analisados considera a fonte de luz afastada do objeto e a incidência de raios paralelos sobre uma superfície

Lambertiana. Nesse trabalho considera-se uma superfície Lambertiana iluminada por uma fonte de luz pontual. Seu posicionamento próximo da superfície produz uma condição mais realística da iluminação do objeto.

Assim, na busca de um modelo completo para a solução do problema de SFS, e que considere a projeção perspectiva para a imagem, propõe-se a utilização de um modelo híbrido baseado nas soluções convencionais de SFS e de técnicas fotogramétricas. Neste último caso, mais especificamente, é utilizada a equação de colinearidade, que relaciona, matematicamente, a projeção perspectiva do ponto no espaço imagem e no espaço objeto, considerando que o ponto no espaço imagem, o centro perspectivo e o ponto no espaço objeto são colineares.

Além disso, como resultado do uso da técnica fotogramétrica, propõe-se a obtenção dos parâmetros que determinam o posicionamento e a orientação da câmara, e a descrição da superfície, a partir de um referencial global.

Portanto, neste trabalho, desenvolve-se um procedimento simultâneo de reconstrução da superfície e da calibração da câmara, que utiliza as técnicas de SFS e técnicas fotogramétricas, considerando a projeção perspectiva e a utilização de informações de silhuetas para reduzir os erros. É utilizado um modelo matemático explícito, sendo aplicado um processo iterativo para recuperação da informação 3D da superfície e dos parâmetros de orientação da câmara. Devido ao uso de equações não lineares no modelo, a linearização por série de Taylor foi utilizada, e o Método dos Mínimos Quadrados (MMQ) foi aplicado em razão da superabundância de observações. Esta solução se diferencia daquelas citadas na literatura pelo fato de se calcular simultaneamente as coordenadas tridimensionais dos pontos da superfície e os parâmetros de orientação da câmara (calibração²), não sendo necessário, para isto, nenhuma restrição sobre a superfície.

1.4 Objetivos do Trabalho

Detectados os problemas relacionados com a reconstrução de uma superfície através de uma imagem de intensidade e da não consideração da projeção perspectiva, o objetivo principal do presente trabalho pode ser definido como:

- desenvolver e implementar um procedimento de recuperação da superfície através de uma imagem de intensidade, considerando a projeção perspectiva como geradora desta imagem.

²Determinação somente dos parâmetros extrínsecos ou de orientação exterior da câmara

Além disso, como resultado do uso dessa abordagem, pode-se delinear como objetivos secundários:

- obter parâmetros que determinem, simultaneamente, o posicionamento e a orientação da câmara (calibração da câmara) bem como a descrição da superfície a partir de um referencial global;
- verificar a eficiência do procedimento em função dos experimentos com dados sintéticos, determinando assim a exatidão e a configuração mais adequada entre a posição da luz e o observador.

1.5 Tópicos do Trabalho

No presente capítulo (Introdução) foram feitas as considerações iniciais relacionadas à técnica de SFS, bem como uma revisão bibliográfica das técnicas de reconstrução existentes, e estabelecidos os objetivos do presente trabalho.

No segundo capítulo são apresentadas as técnicas que exploram informações radiométricas (*shape from X*) e feitas algumas considerações sobre a formação da imagem, questões sobre a superfície, mapa de reflectância, especularidade da superfície e sobre a técnica de reconstrução utilizada por Lee e Kuo [30].

O terceiro capítulo descreve a base teórica do procedimento proposto neste trabalho, apresentando-se, inicialmente, a equação de colinearidade modificada. Neste procedimento, a normal à superfície é calculada utilizando-se as coordenadas do sistema do ponto objeto, criando *patches* triangulares planares. A equação de *shading* utilizada considera que a superfície é Lambertiana. Devido a existência de um número maior de equações do que o mínimo necessário para solucionar o problema, na determinação das incógnitas faz-se uso do Método dos Mínimos Quadrados (MMQ). Em seguida, são apresentados os algoritmos e as otimizações implementadas.

O quarto capítulo refere-se aos testes realizados com o modelo desenvolvido, usando vários tipos de dados simulados, bem como dados reais. Para ambos os casos, os resultados obtidos são apresentados e analisados.

No quinto capítulo são apresentados comentários finais sobre o trabalho, bem como as conclusões e algumas recomendações.

No apêndice A é apresentado o desenvolvimento matemático para a obtenção das derivadas parciais das equações que compõem os elementos da matriz dos coeficientes das derivadas parciais.

Capítulo 2

Fundamentação Teórica

Neste capítulo é apresentada uma revisão de algumas técnicas de reconstrução 3D, bem como a fundamentação teórica das etapas envolvidas desde a formação da imagem no processo de reconstrução tridimensional da imagem.

2.1 Técnicas de "*Shape from X*"

A descrição da superfície física visível pode ser reconstruída através das informações contidas na imagem. Desta forma, existem diversas maneiras de determinar a forma de uma superfície visível através de uma ou mais imagens. Estas técnicas, denominadas de *shape from X* por Aloimonos [1], são: *shading*, *texture*, *contour*, *motion*, *stereo* e outros. Todas estas técnicas procuram extrair as propriedades tridimensionais da superfície a partir das características bidimensionais das imagens.

Nesta seção serão revisadas algumas das técnicas de *shape from X*. A técnica de SFS, que é objeto de pesquisa deste trabalho, será enfatizada com maiores detalhes na seção 2.8.

2.1.1 A Técnica de SFS

A técnica de SFS consiste em explorar informações contidas na imagem, posição do ponto (x , y - coordenadas no espaço imagem) e intensidade obtidas de objetos que possuem informações tridimensionais (3D).

Um dos primeiros trabalhos considerando a informação de *shading* foi desenvolvido na década de 60, segundo Horn e Brooks [20]. Esse trabalho mapeava a superfície lunar, em

preparação para a exploração humana. Desde então Horn começou a pesquisar o assunto e, em 1970 publicou a formulação básica da técnica de SFS. Segundo Horn e Brooks [20], o desenvolvimento desta técnica de recuperação da superfície tem evoluído lentamente devido à grande complexidade da análise matemática do sistema de iluminação e da diversidade de superfícies quanto a sua reflectância (*albedo*).

O procedimento de SFS baseia-se na informação contida na imagem (brilho) para determinar a orientação da superfície, ou seja, faz da intensidade refletida pelo objeto e captada pelo sistema de imageamento, um dos parâmetros para a determinação da superfície. Numa imagem, as variações de intensidade podem indicar a variação da orientação da superfície (inclinação da superfície em relação à fonte de luz) exceto pela influência da distância e de alguns efeitos da especularidade.

Um dos problemas implícitos na técnica de SFS é a deficiência de característica devido à impossibilidade de projetar 2D para 3D (não inversibilidade da projeção 2D para 3D), ou seja, deve-se resolver uma equação com duas incógnitas, o que é impossível se não há consideração de injunções (restrições) ou imposições iniciais. Além destas injunções, várias outras considerações ou simplificações são realizadas para adequar o modelo à situação real. A principal simplificação é a de que a fonte de luz e o observador estão afastados do objeto, significando que o brilho de uma porção de superfície é independente da sua posição espacial (o brilho depende somente da orientação), implicando também na adoção da projeção ortogonal, segundo Horn e Brooks [20]. Desta forma, estando o observador bastante afastado do objeto, pode-se considerar que o sistema de projeção (formação da imagem) seja ortogonal, sem maiores problemas de erro de posicionamento do ponto na imagem.

De uma forma geral, o problema de SFS pode ser resolvido adotando-se um modelo de iluminação para a superfície, geralmente superfície difusa ideal ou superfície *matte*¹, e um modelo matemático que descreva a superfície da melhor forma possível.

Segundo Bruckstein [6], a imagem de intensidade de uma superfície descrita por $H(x,y)$ ² é definida por um mapa de intensidade luminosa $A(x,y)$, sendo que $A(x,y)$ depende das propriedades da superfície, da sua orientação em "x" e "y" e da iluminação. Assim sendo, o problema de SFS é a recuperação da função $H(x,y)$ sobre uma dada região, através da imagem $A(x,y)$ sobre a região definida, e alguma informação adicional possível.

A imagem gerada por uma superfície *matte* (Lambertiana) pode ser descrita matema-

¹Superfície Lambertiana, que reflete os raios de luz incidente com brilhos de mesma intensidade em todas as direções.

² $H(x,y)$ é uma função contínua (segunda ordem ou superior) que descreve uma superfície tridimensional em função de duas variáveis (x,y) .

ticamente a partir do produto escalar entre dois vetores: vetor normal e o vetor direção da fonte de luz. Desta forma, a reconstrução da superfície *patch* a *patch* significa definir a orientação deste *patch* da superfície, ou seja, determinar o vetor normal unitário ($\bar{n}$) à superfície.

2.1.2 A Técnica de *Shape from Texture*

A textura de um objeto pode ser muito útil para o olho humano, ajudando-o com uma informação adicional na percepção da profundidade.

A reconstrução 3D de uma superfície texturizada é obtida através da sua imagem projetada, sob a consideração de que a textura é homogeneamente distribuída. Ainda, a recuperação 3D da superfície através da textura só é possível se há um conhecimento *a priori* sobre a verdadeira textura, segundo Kanatani e Chou [26].

Segundo Haralick e Shapiro [15], a técnica de *Shape from texture* refere-se à capacidade de determinação das características da forma 3D de uma superfície homogeneamente texturizada, através da variação da densidade de textura, manifestada pela superfície na imagem, sob a projeção perspectiva ou ortográfica.

Na recuperação da forma, o efeito da distorção da projeção pode ser determinado através das propriedades da textura sobre a qual a distorção age. Isto requer a introdução de algumas considerações sobre a textura.

Em geral, as distorções introduzidas pela projeção perspectiva podem ser consideradas como oriundas dos seguintes efeitos:

- **efeito distância**, quando objetos do mesmo tamanho parecem maiores quando estão mais próximos do plano da imagem;
- **efeito posição**, quando a distorção do padrão é dependente do ângulo formado entre a linha de visada e o plano da imagem, isto é, o padrão é dependente da sua posição na imagem; e
- **efeito redução**, quando a distorção do padrão é dependente do ângulo formado entre a linha de visada e a normal à superfície.

Destes efeitos, somente o último, efeito redução, atua quando do uso da projeção ortogonal, sendo que os outros dois são ignorados.

Gonzalez [13] considera que uma técnica importante para a descrição de regiões é a quantificação do conteúdo da textura. Isto pode ser obtido intuitivamente tomando como medida as propriedades físicas, tais como: suavidade, rugosidade e regularidade.

As três principais técnicas usadas em processamento de imagens para descrever a textura de uma região são, segundo Gonzalez [13]:

- técnicas estatísticas, que caracterizam a textura como sendo suaves, rugosas, granulares, etc.;
- técnicas estruturais, que descrevem a textura através de composição de primitivas; e
- técnicas espectrais, que descrevem a textura baseada nas propriedades do espectro de frequências.

As variações nas feições no plano imagem, tais como: tamanho, forma, densidade e relação de aparência das texturas primitivas, podem ser exploradas para recuperar a orientação da superfície a partir da imagem, segundo Jain et al [25].

Uma das maiores dificuldades da técnica de *shape from texture* é a necessidade de se reconhecer, primeiramente, a estrutura (análise estrutural) da verdadeira textura, através da regularidade, periodicidade, colinearidade, paralelismo, ortogonalidade, simetria e outros, visto que a textura observada não exibe a estrutura esperada. Segundo Kanatani e Chou [26], vários trabalhos apresentam, a partir de muitas considerações e aproximações, modelos fundamentados estatisticamente. No trabalho apresentado por Kanatani e Chou [26] desenvolve-se um modelo baseado em geometria diferencial e teoria das distribuições, relacionando a textura discretizada corretamente com a descrição de superfícies suaves em termos diferenciais.

2.1.3 A Técnica de *Shape from Contour*

A técnica de *Shape from Contour* refere-se à capacidade de inferir a forma 3D de um objeto a partir de uma imagem (2D) sob a projeção perspectiva, observada a partir de um conjunto de contornos regularmente marcados sobre a superfície do objeto (Haralick e Shapiro [15]).

Esta técnica pode ser entendida como um processo de reconstrução de superfície, usando bordas de uma imagem segmentada que melhor se ajuste aos contornos já conhecidos.

Yuen [43] considera a existência de duas classes de pesquisa sobre *Shape from Contour*, *top-down* e *bottom-up*:

- as classes definidas como *top-down* são baseadas no melhor resultado da comparação (*matching*) entre o contorno obtido da imagem e vários modelos parametrizados representando um objeto, cuja projeção resulta numa linha;
- as classes definidas por *bottom-up*³ consideram o processo de *shape from contour* como sendo um processo de pesquisa para encontrar uma forma métrica, ou seja, neste método, a medida de alguma propriedade do objeto é associada à silhueta.

Sanz [37] descreve que há vários métodos para determinar a forma a partir de contornos, dentre os quais, os já citados por Yuen [43] e o método que emprega o descritor de Fourier normalizado. O descritor de Fourier pode ser utilizado como uma aproximação do contorno, já que a posição ao longo do contorno fechado pode ser descrito por uma função periódica.

Yuen [43] apresenta um método misto para contornar as dificuldades dos dois procedimentos (*bottom-up* e *top-down*). No caso do método *bottom-up*, que utilizam das primitivas de baixo nível, não capturam as propriedades essenciais do objeto. Por outro lado, no caso do método *top down*, que usam modelos parametrizados tem alto custo computacional. Assim, para contornar estas dificuldades, Yuen propõe um método que identifica as propriedades invariantes (que tem implicações tridimensionais) nas projeções dos contornos, aplicando uma projeção inversa que melhor preserve estas invariáveis.

Ulupinar e Nevatia [40] apresentam uma metodologia que possibilita a inferência da forma 3D através do contorno de objetos cilíndricos, sob a condição de que eles exibam certa propriedade simétrica, procedimento este que está relacionado com o método *top-down*.

As bordas são de fundamental importância neste procedimento, pois elas podem representar, na imagem, algum tipo de descontinuidade provocada pela variação de refletância, profundidade ou orientação. Pode-se ter outro tipo de borda que não apresente a descontinuidade mencionada anteriormente, e que são chamadas de *silhuetas*, quando se referem a imagem e, *limb*⁴ ou *occluding boundary*⁵, quando se referem à cena.

³Definidas como *método dos extremos* por Brady [4] e Aloimonos [1].

⁴A *limb* de uma superfície é uma curva dependente do ponto de vista, tal que para cada ponto na curva a normal à superfície é ortogonal à direção do ponto de vista, segundo Ulupinar e Nevatia [40]

⁵Segundo Ikeuchi e Horn [23]

As bordas de oclusão (descritas detalhadamente no item 3.4), que produzem contornos ou silhuetas nas imagens, permitem inferir diretamente a orientação da superfície ao longo do contorno, podendo eliminar a ambigüidade nestes pontos.

2.1.4 A Técnica de *Shape from Motion*

O procedimento de *shape from motion* é caracterizado pela consideração de duas ou mais imagens seqüenciais, onde cada imagem gravada representa um instante da cena num certo tempo. As mudanças nas cenas podem ser causadas pelo movimento da câmara em relação ao objeto, pelo movimento do objeto, pela troca de iluminação, pela troca na estrutura, e pela mudança no tamanho ou forma de um objeto. Geralmente, supõe-se que a mudança na cena ocorre em função do movimento da câmara ou do objeto, segundo Jain et al [25].

Recuperar a geometria de uma cena e o movimento da câmara a partir de uma seqüência de imagens é uma importante tarefa numa variedade de aplicações, tais como: inspeção e automação industrial, navegação, engenharia biomédica, sensoriamento remoto, robótica e fotogrametria, segundo Poelman e Kanade [35] e Haralick e Shapiro [15].

A técnica de *shape from motion*, conhecida como *structure from motion*, consiste basicamente de duas etapas:

- processar estruturas observáveis na imagem, isto é, pontos, linhas, contornos, etc;
- relacionar estas estruturas com eventos no espaço.

Wechsler [16] coloca a importância da técnica de *shape from motion* na navegação, onde um observador ativo vê novas informações para explorar o ambiente. A exploração, geralmente, envolve movimentos, deslocamentos de objetos de interesse, planejamento de uma rota segura para evitar colisão e integração de informações para uma navegação melhor e mais eficiente.

2.1.5 A Técnica de *Shape from Stereo*

Esta técnica tem causado muita indefinição na classificação dos métodos, pois verifica-se que Aloimonos [1] e Brady [4] definem *shape from stereo* como sendo estéreo visão, ou seja, a técnica de reconstruir superfícies utilizando duas imagens tomadas de dois pontos de vista

distintos. Assim, a mesma cena seria gravada em duas imagens, com alguma disparidade, provocada pelo deslocamento do ponto de vista.

Desta forma, reconstruir a superfície a partir destas imagens, supondo-as calibradas, seria explorar as propriedades geométricas contidas na câmara e nas imagens. Supondo um par de imagens tomadas sem inclinações e orientadas em relação ao referencial utilizado, (ver Figura 2.1), pode-se calcular as coordenadas tridimensionais dos pontos no espaço objeto, usando as equações 2.1.

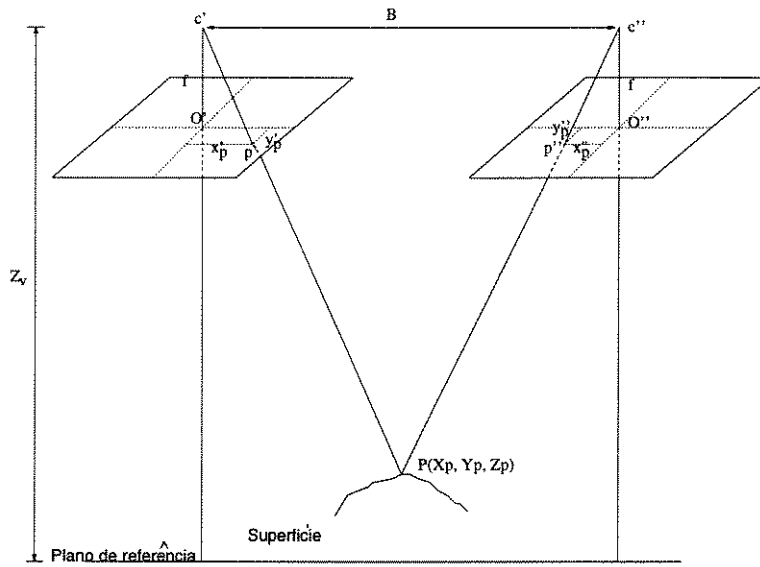


Figura 2.1: Esquema de um par estereoscópico nivelado em relação ao sistema de coordenadas referencial

$$\begin{aligned}
 X_p &= x'_p \frac{Z_v - Z_p}{f} \\
 Y_p &= y'_p \frac{Z_v - Z_p}{f} \\
 Z_p &= Z_v - B \frac{f}{px}
 \end{aligned} \tag{2.1}$$

$$\text{e } px = x'_p - x''_p$$

onde:

X_p , Y_p e Z_p são as coordenadas tridimensionais do ponto no espaço objeto;

B é a base, distância entre as câmaras;

f é a distância focal;

Z_v é a distância de projeção;

x'_p, x''_p, y'_p, y''_p são as coordenadas do ponto imagem, no sistema de coordenadas com origem no ponto principal⁶ de cada foto O' e O'' , onde $(')$ e $('')$ representam a câmara da esquerda e da direita, respectivamente.

A condição definida pela figura 2.1 pode ser obtida (na prática) por estéreo-câmaras, que são dispositivos que contém duas câmaras semelhantes, montadas nos extremos de uma base (B) e orientadas de forma a não permitir movimentos angulares entre elas.

Do ponto de vista da visão computacional, a técnica de *Shape from Stereo* para reconstrução de superfícies, supondo as imagens retificadas⁷, seria realizar o *stereo matching*.

O *stereo matching*⁸ é um procedimento de complexa realização, mas neste caso, está fora da concepção dos procedimentos adotados nas resoluções dos métodos que usam da fotometria.

Portanto, no contexto do procedimento adotado neste trabalho, considera-se que *shape from stereo* é o procedimento conhecido na literatura consultada como *photometric stereo* ([37], [9], [11]). Desta forma, pode-se sub-dividir a técnica de *shape from stereo* em dois procedimentos: estéreo-visão e estéreo-fotometria.

A técnica de estéreo fotometria é aquela que aumenta a informação disponível através da variação da refletância da superfície. Basicamente, são tomadas várias imagens do mesmo ponto de vista, mas com diferentes posicionamentos da fonte de luz. Desta forma, esta técnica produz vários mapas de refletância para uma mesma superfície, eliminando-se, assim, a ambigüidade provocada pelo uso de um mapa.

Na Figura 2.2 verifica-se superposição das três curvas colocadas num único mapa de refletância (considerando que foram obtidas a partir de três imagens distintas); a interseção das três isolinhas de mesma intensidade determina um ponto neste mapa resultante. A posição deste ponto (interseção) determina um valor que é o gradiente da superfície.

Neste procedimento, mesmo tendo mais de uma imagem, não há necessidade de se realizar *matching*, pois existe uma correspondência direta entre todos os pontos de todas as imagens. Isto se deve à manutenção do mesmo ponto de vista nas tomadas.

⁶É o ponto definido no plano focal, pela interseção da linha perpendicular ao plano focal que passa pelo ponto nodal da lente da câmara, segundo Wolf [42].

⁷Imagens retificadas são imagens corrigidas das rotações da câmara em relação ao referencial utilizado (nivelamento da câmara em relação ao sistema de coordenadas).

⁸*Stereo matching* é um procedimento de determinação de pontos ou feições correspondentes, em duas ou mais imagens.

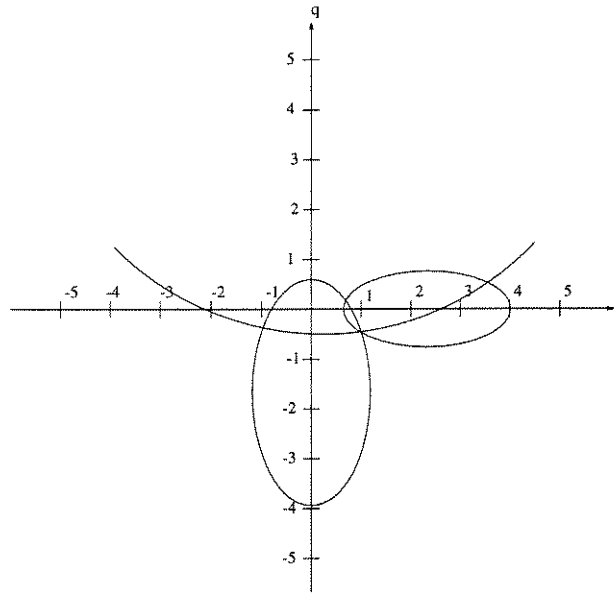


Figura 2.2: Mapa de refletância com três contornos de intensidade constante

2.2 Formação Radiométrica da Imagem

Nesta seção desenvolve-se os conceitos de hemisfério de iluminação, definindo e conceituando ângulo sólido, definições estas utilizadas no desenvolvimento da equação geral de iluminação.

2.2.1 Hemisfério de Iluminação e Ângulo Sólido

Segundo Hall [14], o hemisfério de iluminação é uma notação conveniente para descrever um evento de iluminação acima ou abaixo da superfície. Estes eventos, tais como fontes de luz e outras superfícies refletoras, são projetados sobre este hemisfério de raio unitário.

Um ângulo sólido, ω , descreve a quantidade do hemisfério de iluminação que é coberto pela projeção de um evento de iluminação sobre o hemisfério de iluminação.

Considerando uma fonte de luz circular (Fig. 2.3), o ângulo sólido desta fonte é determinado projetando-a sobre o hemisfério de iluminação e integrando a área de projeção, conforme a equação 2.2.

$$\omega = \int_0^{2\pi} \int_0^{\arctan(\frac{r_s}{r_0})} \sin(\theta) d\theta d\Phi = 2\pi(1 - \cos(\arctan(\frac{r_s}{r_0}))) \quad (2.2)$$

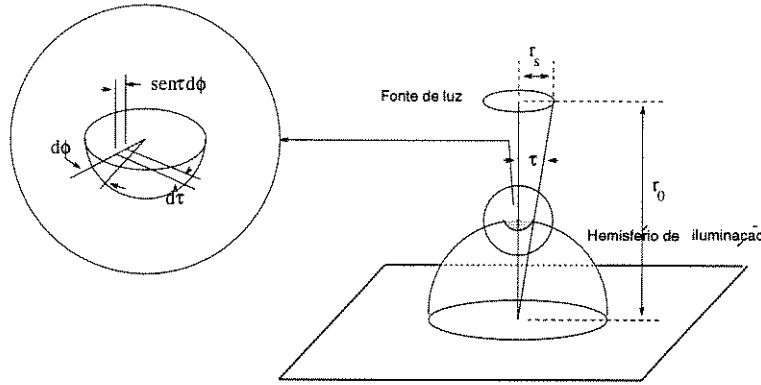


Figura 2.3: Projeção de uma fonte de luz circular sobre o hemisfério de iluminação, adaptado de Hall [13]

O cálculo do ângulo sólido é computacionalmente difícil de ser obtido, mas, considerando o caso de $r_s \ll r_0$, quando o ângulo sólido é muito pequeno, vale a aproximação:

$$\theta = \arctan\left(\frac{r_s}{r_0}\right) \cong \text{sen}\left(\frac{r_s}{r_0}\right) \cong \frac{r_s}{r_0}$$

Desta forma um ângulo sólido pequeno é aproximado por:

$$\omega = \int^{2\pi} \int_0^{\frac{r_s}{r_0}} \theta d\theta d\Phi = \pi \frac{r_s^2}{r_0^2} \quad (2.3)$$

Na figura 2.3 verifica-se que a área da fonte de luz (no plano) é πr_s^2 . Denominando-a de A_p , a equação 2.3 pode ser escrita como:

$$d\omega = \frac{dA_p}{r_0^2}$$

que é o ângulo sólido, quando o observador e a fonte de luz estão próximos.

Quando o observador e a fonte de luz estão afastados, há uma redução da área (dA_p) pelo valor do co-seno do ângulo formado entre o vetor normal e a direção da fonte de luz, conforme a equação 2.4, que é o ângulo sólido projetado.

$$d\Omega = \frac{dA_p \cos(\alpha)}{r_0^2} \quad (2.4)$$

2.2.2 Modelo de Iluminação

Na maioria dos casos em Visão Computacional, considera-se o meio ambiente como sendo composto por corpos opacos imersos num meio transparente. Assim, raios de luz que passam nesse meio (transparente) não sofrem influências na trajetória e não penetram no objeto. Isto significa que pode-se ignorar o meio, e que somente a superfície do objeto é de interesse.

Esta consideração não seria válida se o meio ambiente considerado fosse translúcido e o objeto semi-opaco. Nestas condições, as imagens (bidimensionais) seriam difíceis de ser interpretadas. Tal condição pode ser exemplificada como no caso de imagens microscópicas de espécimes biológicos e imagens de objetos nos meios aquosos (imagens sub-aquáticas).

Desta forma, pode-se considerar a Intensidade Luminosa como sendo uma parcela da luz refletida por uma superfície, a qual depende da sua micro estrutura e da distribuição da luz incidente. A imagem obtida por um dispositivo de *imageamento* é uma cena bidimensional, que fornece informações de uma complexa estrutura do espaço tridimensional. Assim, pode-se explorar estas informações sob dois aspectos, a fim de realizar a reconstrução: um problema é ligado a geometria da projeção (perspectiva) que gerencia a formação da imagem enquanto que o outro é relacionado com a intensidade da luz gravada na imagem.

Para descrever matematicamente o complexo processo de formação da imagem, é necessário realizar algumas considerações. Desta forma, considerando que exista uma perfeita focalização do dispositivo de aquisição da imagem, isto é, um ponto no espaço objeto passa através da lente e é refratado como um ponto no plano imagem. Esta consideração possibilita a utilização das equações fundamentais da ótica geométrica no processo de formação da imagem.

Assim, matematicamente, raios refletidos por uma área infinitesimal dA_0 do espaço objeto serão projetados sobre o plano imagem com uma área dA_p , e nenhum outro raio atingirá esta porção da imagem. Assim, assume-se que não há aberrações e/ou distorções (que ocorrem nos sistemas de lentes), ou seja, a lente usada é ideal.

A exposição de um filme numa câmara é proporcional à irradiância⁹ E_p da imagem e os níveis de cinza, num sistema digital de imageamento, são medidas quantificadas da radiância¹⁰ da imagem, segundo Horn [22]. Assim, para calcular a radiância da imagem deve-

⁹Irradiância é a densidade de fluxo radiante incidente sobre uma superfície plana, segundo [Bernd [24], Haralick e Shapiro [15]]

¹⁰Radiância é o fluxo radiante que deixa determinada superfície por unidade de ângulo sólido e unidade de área projetada na direção da medida, segundo [Bernd [24], Haralick e Shapiro [15]].

se, primeiro, determinar o fluxo de energia $d\Phi_L$ que passa através da abertura (diafragma) vindo da área dA_0 do objeto.

Desta forma, pode-se escrever:

$$d\Phi_L = dA_0 \int_{\Omega_r} L_r d\Omega_r \quad (2.5)$$

onde:

Ω_r é o ângulo sólido projetado.

A irradiância da imagem E_p é o fluxo de energia por unidade de área, e é dada por:

$$E_p = \frac{d\Phi_L}{dA_p} \quad (2.6)$$

Conforme a Figura 2.4, θ'_r é o ângulo formado entre a normal à superfície no ponto e a linha deste ponto com o ponto nodal da lente, e α é o ângulo entre esta linha e o eixo óptico. Desta forma, pela equação do ângulo sólido projetado:

$$\frac{dA_0 \cos(\theta'_r)}{f_0^2} = \frac{dA_p \cos(\alpha)}{f_p^2} \quad (2.7)$$

Realizando algumas operações matemáticas, obtém-se:

$$E_p = \left(\frac{f_0}{f_p}\right)^2 \cos(\alpha) \int_{\omega_r} L_r \frac{\cos(\theta_r)}{\cos(\theta'_r)} d\omega_r \quad (2.8)$$

A integral da equação 2.8 é definida sobre o ângulo sólido ω_r . Assumindo que a lente é relativamente pequena em relação a sua distância do objeto, pode-se considerar que θ_r é, aproximadamente, igual a θ'_r . Desta forma, a radiância refletida L_r tenderá a ser constante, podendo assim ser retirada da integral. Então, a equação 2.8 torna-se:

$$E_p = \left(\frac{f_0}{f_p}\right)^2 \cos(\alpha) \omega \quad (2.9)$$

O ângulo sólido ocupado pela lente, quando visto da superfície, é aproximadamente igual a um fator de redução, dado pela equação 2.10:

$$\omega = \left(\frac{\pi}{4}\right) \frac{d^2 \cos(\alpha)^3}{f_0^2} \quad (2.10)$$

Substituindo ω da equação 2.10 na equação 2.9 obtém-se a equação 2.11, conhecida como equação da irradiância da imagem.

$$E_p = L_r \frac{\pi}{4} \left(\frac{d}{f_p} \right)^2 \cos(\alpha)^4 \quad (2.11)$$

onde:

d é a abertura do diafragma da lente da câmara; e

f_p é a distância focal da câmara.

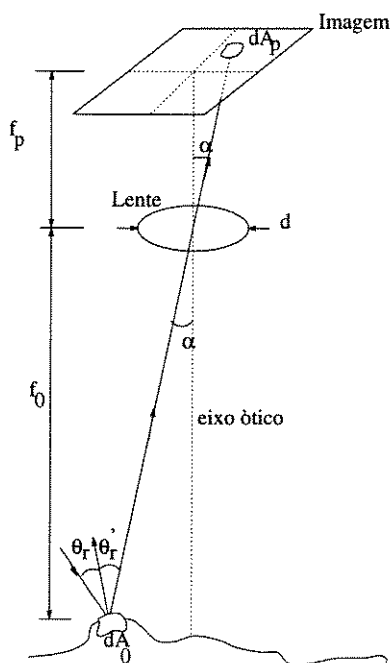


Figura 2.4: Sistema de formação da imagem, adaptado de Horn e Sjoberg [21]

Observando a equação 2.11, pode-se notar que a sensibilidade em uma imagem não é uniforme, mas é constante para um ponto particular na imagem. Desta forma, mesmo não considerando as aberrações e as distorções na dedução, o modelo matemático (Eq. 2.11) produz um efeito na lente conhecido como *vignetting*¹¹. Portanto, um sistema de imageamento ideal deverá ser montado para que este efeito, nas bordas da imagem, seja reduzido com o objetivo de melhorar a qualidade da imagem.

Se não forem considerados os efeitos das aberrações e das distorções, pode-se verificar que a irradiância da imagem E_p no ponto (x, y) é relacionada à radiância da cena L_r no

¹¹ *Vignetting*, traduzido como vinhetamento, é um efeito provocado pelo sistema ótico de uma câmara, no qual o diafragma bloqueia o feixe de luz incidente na lente, provocando uma redução na iluminação dos pontos na imagem próximos das bordas.

correspondente ponto no espaço objeto.

Na Equação 2.11 verifica-se que a irradiância da imagem é proporcional à radiância da cena, onde este fator é inversamente proporcional ao $n^o f$ ($n^o f = \frac{f_p}{d}$).

Desta forma, pode-se concluir que a irradiância da imagem é diretamente proporcional à radiância da cena, a qual depende dos seguintes fatores:

- distribuição e intensidade das fontes de luz;
- microestrutura da superfície; e
- orientação da superfície.

2.3 Formação Geométrica da Imagem

Praticamente, todos os instrumentos que reproduzem imagens (câmaras digitais) dependem, de alguma maneira dos elementos óticos para seu funcionamento. O número e o tipo de elementos óticos usados variam conforme o tipo de equipamento utilizado. Desta forma, pode-se dizer que a imagem formada (gerada) por uma câmara pode ser modelada pelas técnicas da ciência ótica (física e geométrica).

A ciência ótica é definida por duas principais áreas: a ótica física e a geométrica. Na ótica física considera-se que o deslocamento da luz num meio transparente se dá por uma série de ondas eletromagnéticas emanadas a partir da fonte.

Na ótica geométrica, o deslocamento da luz num meio transparente, a partir da fonte, é considerado em linhas retas, chamado de raios de luz. Estes raios produzem infinitos feixes de luz que se propagam em todas as direções a partir da fonte.

Geralmente, as equações fundamentais utilizadas na formação da imagem estão referenciadas às linhas retas (caminho) que o raio de luz descreve no meio ambiente, caracterizando assim, o comportamento da luz na área da ótica geométrica.

A mais simples e primitiva câmara que representa com maior fidelidade as considerações da ótica geométrica é a câmara de *pinhole*¹²(que é uma caixa hermeticamente fechada com um pequeno orifício). Este dispositivo, teoricamente, deixa passar os raios de luz provenientes dos pontos objeto, projetando a cena invertida no plano oposto ao do orifício, produzindo uma imagem sem aberrações e sem distorções.

¹²Também conhecida como câmara de orifício ou buraco de agulha.

Entretanto, esta câmara deixa passar pouca luminosidade pelo pequeno orifício, tornando-a extremamente lenta, incapacitando-a para a maioria dos trabalhos fotográficos. Com a finalidade de eliminar este problema, usam-se lentes para aumentar a quantidade de luz que passa pelo orifício. Desta forma, as lentes utilizadas numa câmara fotográfica têm por finalidade reunir os raios de luz que vêm do objeto e focalizá-los em alguma posição do lado oposto à lente.

Contudo, o uso de lentes provoca imperfeições na imagem, degradando a nitidez e produzindo aberrações (aberração esférica, astigmatismo e curvatura de campo, coma e aberração cromática). Para diminuir estas aberrações usam-se combinações de lentes, procedimento este que provoca as distorções.

Geralmente, nas deduções das equações, adota-se a condição ideal para descrever o evento matematicamente, ficando, para cada situação, a modelagem ou não dos erros inerentes às imperfeições das lentes. Esta modelagem é conhecida na área de Visão Computacional e Fotogrametria como calibração dos parâmetros intrínsecos da câmara, determinados juntamente com os parâmetros extrínsecos de calibração durante o processamento computacional.

Assim, o processo de formação da imagem pode ser descrito por uma transformação (Eq. 2.12) 2D para 3D, usando como base a câmara de *pinhole* e o sistema de coordenadas homogêneas, assumindo uma posição arbitrária para a posição da lente.

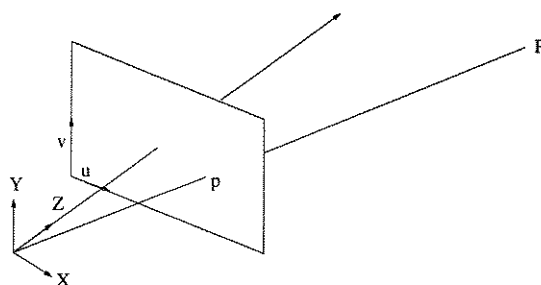


Figura 2.5: Sistema de coordenadas uv da tela

A transformação perspectiva pode ser representada com coordenadas homogêneas, modelo matemático que foi desenvolvido como uma ferramenta em aplicações na geometria projetiva analítica, utilizada em computação gráfica e automação industrial, segundo Barnard [3].

Desta forma, a representação de um ponto com coordenadas homogêneas $[x, y, t]$, cujas coordenadas cartesianas bidimensionais são $[u, v]$, onde

$$u = \frac{x}{t} \quad e \quad v = \frac{y}{t} \quad (2.12)$$

e a sua relação com as coordenadas cartesianas tridimensionais ordinárias se dá, segundo Jahne [24]:

$$\begin{bmatrix} x \\ y \\ t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -f \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.13)$$

ou

$$\begin{bmatrix} x \\ y \\ t \end{bmatrix} = \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ \frac{Z - Z_0}{f} \end{bmatrix} \quad (2.14)$$

onde:

X , Y e Z são as coordenadas de um ponto no espaço 3D;

X_0 , Y_0 e Z_0 são as coordenadas do centro perspectivo da lente;

f é a distância focal da câmara; e

u e v são as coordenadas da projeção perspectiva de (X, Y, Z) sobre o plano da tela.

2.4 Considerações Sobre a Superfície

Quando um raio de luz atinge a superfície de um objeto, ele pode ser absorvido, transmitido ou refletido. Se a superfície é plana e constituída basicamente de material homogêneo, o raio refletido estará no plano formado pelo raio incidente e a normal à superfície, e formará um ângulo com a normal local, igual ao ângulo formado entre o raio incidente e a normal local. Esta condição é conhecida como reflexão especular, metálica ou dielétrica.

Se a superfície de um objeto é suave (liso), a luz refletida e transmitida é muito pouco espalhada. Da mesma forma, se a superfície do objeto é muito rugosa, a luz é espalhada em todas as direções.

Muitas superfícies não são perfeitamente lisas quando vistas mais detalhadamente (ampliadas), espalhando, assim, os raios incidentes em várias direções (ver Figura 2.6 (a)).

Se este espalhamento for pequeno em relação à normal média do local (haverá várias normais em torno do ponto), a maioria dos raios refletidos estarão próximos da direção da reflexão especular ideal, contribuindo para a formação de uma superfície prateada ou espelhada.

Outras superfícies¹³ que não são homogêneas espalham a luz incidente, que penetram na superfície por refração e reflexão. Os raios espalhados podem reemergir próximos do ponto de incidência, em várias direções, e contribuem na formação da reflexão difusa, plana ou *matte*, ver Figura 2.6 (b).

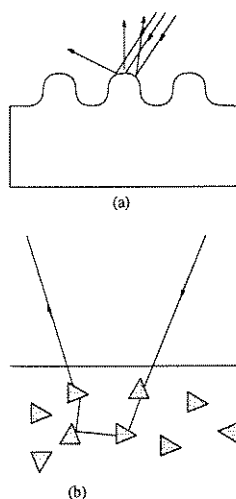


Figura 2.6: (a) Superfície rugosa - (b) Superfície não homogênea

O processo de reflexão de muitas superfícies pode ser convenientemente modelado como uma combinação das componentes difusa ou *matte* e especular. Um caso de superfícies puramente especulares, que é uma superfície perfeitamente plana ou lisa, é a espelhada. Por outro lado, o modelo de reflexão da luz, que define a superfície mais simples, é o modelo ideal difuso, chamada de superfície Lambertiana.

A superfície Lambertiana é a ideal, pois tem a função refletividade proporcional ao ângulo incidente. Esta superfície tem a propriedade de manter o brilho uniforme sob uma iluminação uniforme ou colimada, ou seja, a intensidade de um elemento da superfície é constante com relação à posição do observador, segundo Ballard e Brown [2].

A elaboração de uma imagem de intensidade de uma superfície, através de sua orientação, é única e pode ser determinada por diferentes tipos de materiais e condições de iluminação. Desta forma, o uso de *shading* em computação gráfica é um problema simples de se resolver, podendo-se dizer que é um problema "direto". Na técnica de SFS o problema

¹³Superfícies constituídas de diferentes índices de reflexão.

é "inverso", o que torna mais complexa sua resolução.

2.5 Mapa de Refletância

A quantidade de luz refletida por um *patch* de superfície depende, basicamente, da sua orientação, da distribuição das fontes de luz e da natureza da superfície.

Para uma determinada superfície e distribuição das fontes de luz existe um valor fixo da refletância para cada orientação da normal à superfície. Assim, o mapa de refletância pode fornecer a radiância da cena, como uma função do gradiente da superfície ($\mathbf{p}$ e $\mathbf{q}$) num sistema de coordenadas centrado no observador.

O mapa de refletância, geralmente, é representado como uma série de contornos constantes da radiância da cena (Fig. 2.10). Ela pode ser medida, diretamente, através de goniômetro ou, indiretamente, a partir da imagem de um objeto com a forma conhecida. Por outro lado, o mapa de refletância pode ser calculado se as propriedades do material da superfície e a distribuição da fonte de luz forem conhecidas.

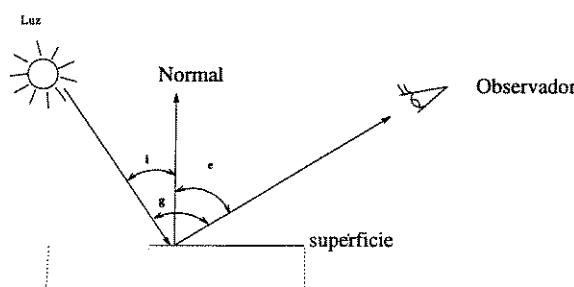


Figura 2.7: Caminhos da luz e ângulos de fase - Shading

O modelo de iluminação, ilustrado na figura 2.7, é composto por quatro elementos fundamentais: a superfície (com as suas características), a normal à superfície no ponto, a fonte ou fontes de luz e o observador (posição no espaço).

Através destes elementos definem-se três variáveis:

i , denominado ângulo de incidência, que é formado entre os vetores normal e fonte de luz;
 e , denominado ângulo emergente, que é formado entre os vetores normal e observador; e
 g , denominado ângulo de fase, que é formado entre os vetores observador e fonte de luz.

Desta forma, a refletância característica de uma superfície pode ser representada como uma função $\beta(i, e, g)$ dos três ângulos i , e e g , definido na figura 2.7. A função $\beta(i, e, g)$

determina a relação de radiância da superfície com a irradiância da superfície.

Pode-se definir uma superfície, explicitamente por $H(x,y)$ e as quantidades (p, q) ¹⁴ chamada de gradiente de $H(x,y)$. O espaço gradiente é um espaço bidimensional de todos os pontos (p, q) . Este espaço produz uma forma eficiente de representar a orientação da superfície.

Um dispositivo de imageamento ideal deveria produzir irradiâncias na imagem proporcionais às radiâncias da cena. As quantidades medidas (irradiâncias na imagem) são geralmente denominadas de níveis de cinza.

Com a consideração da projeção ortogonal, a direção do observador e, portanto, o ângulo de fase g , é constante para todos os pontos da imagem. Desta forma, para uma posição fixa da fonte de luz e do observador, a relação da radiância da cena com a irradiância da imagem, depende somente de componentes do gradiente (p, q) .

Graficamente, pode-se representar a irradiância da imagem através do mapa de refletância, que é definido em função dos gradientes p e q . Um mapa de refletância é específico para o tipo de material imageado, para a fonte de luz específica e para a geometria do observador.

Definem-se algumas funções trigonométricas a fim de calcular as componentes e, assim, construir o mapa de refletância. Deste modo, $\cos(i)$, $\cos(e)$ e $\cos(g)$ podem ser determinados através dos produtos escalares dos vetores normais a superfície $(p, q, -1)^T$, do vetor direção ponto-fonte de luz $(p_l, q_l, -1)^T$ e do vetor que define a posição do observador $(0, 0, -1)^T$.

$$\cos(i) = \frac{1 + p_l p + q_l q}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_l^2 + q_l^2}} \quad (2.15)$$

$$\cos(e) = \frac{1}{\sqrt{1 + p^2 + q^2}} \quad (2.16)$$

$$\cos(g) = \frac{1}{\sqrt{1 + p_l^2 + q_l^2}} \quad (2.17)$$

as variáveis p_l e q_l são as componentes do gradiente na direção x e y do vetor luz, respectivamente.

¹⁴Definidas na equação 2.32

Pode-se identificar a orientação da superfície através da esfera unitária Gaussiana, quando um plano é colocado tangente ao pólo (Norte), e o centro de projeção no pólo oposto (Sul). Define-se, desta maneira, o sistema de projeção estereográfico.

O plano estereográfico (tangente ao pólo norte) possui um sistema de eixos cartesianos denominados de f e g , que podem ser relacionados aos gradientes através das equações 2.18, segundo Ikeuchi e Horn [23],

$$f = \frac{2p}{1 + \sqrt{1 + p^2 + q^2}} \text{ e } g = \frac{2q}{1 + \sqrt{1 + p^2 + q^2}} \quad (2.18)$$

e a equação 2.19 fornecendo a relação inversa,

$$p = \frac{4f}{4 - f^2 - g^2} \text{ e } q = \frac{4g}{4 - f^2 - g^2} \quad (2.19)$$

A projeção estereográfica é uma projeção conforme, isto é, preserva os ângulos entre as linhas e, portanto, a forma das figuras. Geralmente, esta projeção é utilizada quando há necessidade de explorar os limites de uma superfície, mais especificamente, quando a normal no contorno é zero e seus pontos, quando projetados no plano estereográfico, definem a linha projetada do equador da esfera Gaussiana.

Uma outra forma de descrever a superfície é especificando as inclinações desta através dos ângulos σ (inclinação) e τ (direção da inclinação), conhecidas como co-latitute e longitude, conforme a figura 2.8.

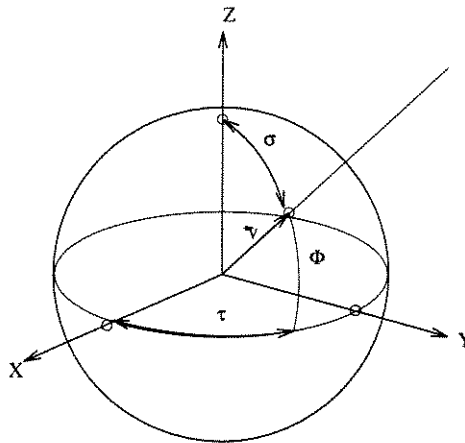


Figura 2.8: Sistema de coordenadas esféricas com coordenada Z passando pelo pólo.

Na figura 2.8, um ponto é definido pela sua orientação em termos de co-latitute (σ) e longitude (τ) sobre a esfera. O ponto pode ser expressado em coordenadas esféricas

(co-latidade e longitude), que descreve direção do vetor direção do ponto, e é descrita pela equação 2.20,

$$\bar{p} = [\text{sen}(\sigma) \cos(\tau), \text{sen}(\sigma) \text{sen}(\tau), \cos(\sigma)] \quad (2.20)$$

Geralmente, utiliza-se do mapa de refletância para determinar a superfície imageada. Esta representação serve como um meio de especificar a dependência do brilho em função da orientação da superfície. Assim, o brilho pode ser representado como função da orientação na forma $R(\bar{n})$, se utilizado o vetor normal unitário, e $R(p,q)$ quando se utiliza dos componentes gradientes. Ainda, uma representação gráfica do mapa de refletância pode ser plotada através dos contornos constantes do brilho no plano pq (ver Figura 2.10), chamado de espaço gradiente, segundo Horn [18].

2.5.1 Mapa de refletância com a fonte de luz próxima do observador

Conforme a definição de superfície Lambertiana, a intensidade luminosa de um ponto pode ser dada por:

$$R(p, q) = \rho \cos(\mathbf{i}) \quad (2.21)$$

ou na forma expandida:

$$R(p, q) = \rho \frac{V_n \cdot V_l}{||V_n|| ||V_l||} \quad (2.22)$$

onde:

V_n é a normal à superfície,

V_l é o vetor direção da fonte de luz e

ρ é o albedo da superfície,

Dentre os vários modelos de iluminação, o mais simples é o que considera uma superfície Lambertiana com a fonte de luz (paralela) próxima (coincidente) do observador. Esta consideração resulta em $\mathbf{g} = 0$ e, conseqüentemente, em $\mathbf{e} = \mathbf{i}$, conforme a Figura 2.9.

Observa-se, na Figura 2.9, que os feixes de luz estão na mesma direção do observador (câmara); então, a partir da equação 2.22, o vetor luz pode ser reduzido a:

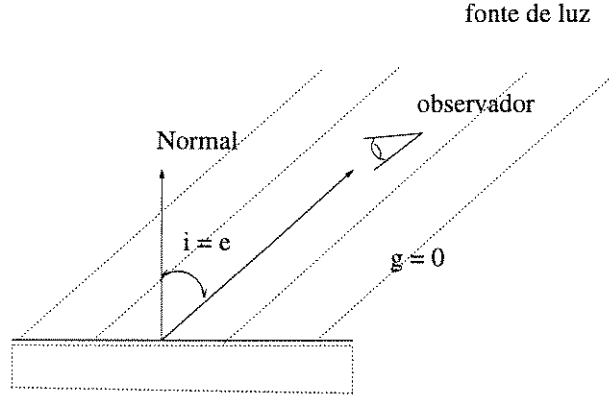


Figura 2.9: Geometria do sistema de iluminação Lambertiana - fonte e observador na mesma direção

$$V_l = (0, 0, -Z_l)$$

resultando em $(0, 0, -1)$, utilizando o vetor unitário orientado na direção do eixo z. Assim;

$$\cos(\mathbf{i}) = \frac{V_n}{||V_n||} \quad (2.23)$$

Esta condição, $\mathbf{i} = \mathbf{e}$ e $\mathbf{g} = 0$, só ocorrerá se a fonte de luz estiver no infinito (pode ser o caso do Sol), projetando os feixes paralelos, e se o sistema de projeção utilizado for a ortográfica. Esta condição gera um mapa de refletância com as iso-linhas de intensidade circulares e concêntricas na origem do sistema (ver Figura 2.10).

O mapa de refletância é construído a partir da equação 2.15, realizando a seguinte consideração:

$$p_l = 0 \text{ e } q_l = 0$$

resultando em

$$\cos(\mathbf{i}) = \frac{1}{\sqrt{1 + p^2 + q^2}} \quad (2.24)$$

que é igual a equação 2.16, resultando em:

$$R(p, q) = \rho \cos(\mathbf{e}) = \rho \cos(\mathbf{i}) = \frac{1}{\sqrt{1 + p^2 + q^2}} \quad (2.25)$$

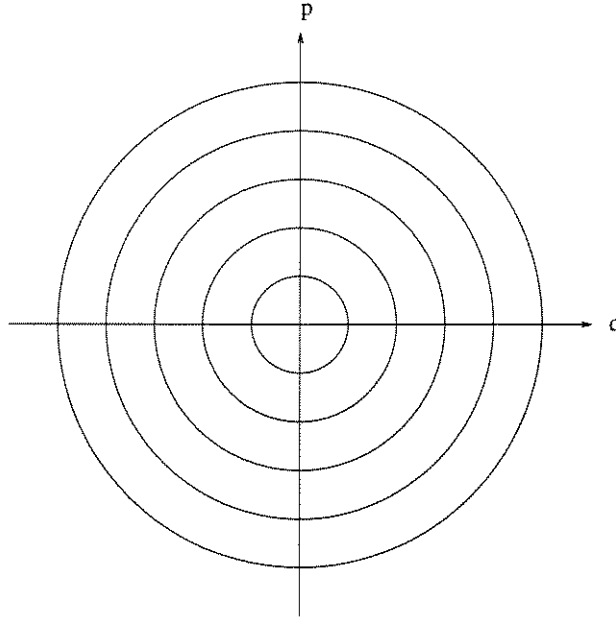


Figura 2.10: Mapa de refletância com a posição do observador e da fonte de luz coincidentes com $\rho = 1$

2.5.2 Mapa de refletância com a fonte de luz afastada do observador

No caso da superfície Lambertiana com a fonte de luz afastada do observador, a equação 2.22 não se altera, ficando:

$$V_l = (X_l, Y_l, Z_l)$$

onde X_l , Y_l e Z_l são constantes para toda a imagem.

A representação em forma de gradiente é dada pela equação 2.21, resultando na equação 2.26, que define o mapa de refletância para este caso.

$$R(p, q) = \rho \frac{(-p, -q, 1)}{\sqrt{p^2 + q^2 + 1}} \frac{(-p_l, -q_l, 1)}{\sqrt{p_l^2 + q_l^2 + 1}} \quad (2.26)$$

Verifica-se, na Figura 2.11, que as iso-linhas têm forma elíptica em algumas regiões e, no caso extremo (valor = 0), degenerando para uma reta.

Da mesma forma que no caso anterior, estas condições seguirão as mesmas considerações, ou seja, os feixes de luz e a projeção são determinadas pela projeção ortográfica.

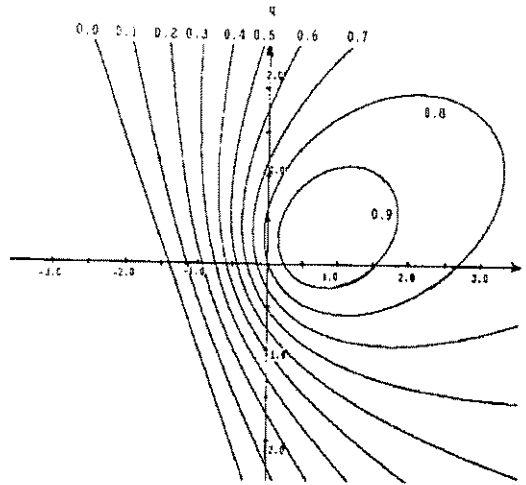


Figura 2.11: Mapa de refletância $R(p,q)$ - Com a fonte de luz ($p_l = 0.7$, $q_l = 0.3$ e $\rho = 1.0$), obtido de Horn [17].

2.5.3 Mapa de refletância para Superfícies Não Lambertianas

A superfície Lunar (ou algum outro satélite) pode ser considerada como sendo um terceiro tipo de superfície, segundo Horn [18], pois ela mantém uma característica particular em relação à superfície *matte* (Lambertiana) especular.

Neste caso, pode-se verificar que o ângulo de fase (g) é zero, pois a fonte de luz (sol) sempre estará na mesma direção do observador; na realidade, sempre estará atrás, pois somente nesta condição a Lua será visível da superfície terrestre.

Devido à grande distância de observação, a microestrutura não aparece da mesma forma que quando observada de perto. Desta forma, a função refletividade foi obtida com uma relação entre o co-seno do ângulo de incidência e a emissão.

$$I_{(i,j)} = \frac{I}{E}$$

onde, $I = \cos(i)$ e $E = \cos(e)$.

Este fato pode ser notado quando se observa a Lua (cheia). Verifica-se que o brilho da superfície Lunar é, aproximadamente, constante. Portanto, ao observar a Lua, nota-se que ela tem a forma de um disco plano e não a forma esférica que deveria ter.

Assim, o mapa de refletância da superfície Lunar (ver Figura 2.12) será constituído de linhas, a partir da linha de sombra (extremo). Estas linhas de contorno serão perpendiculares

à linha que define a direção da posição da fonte de luz.

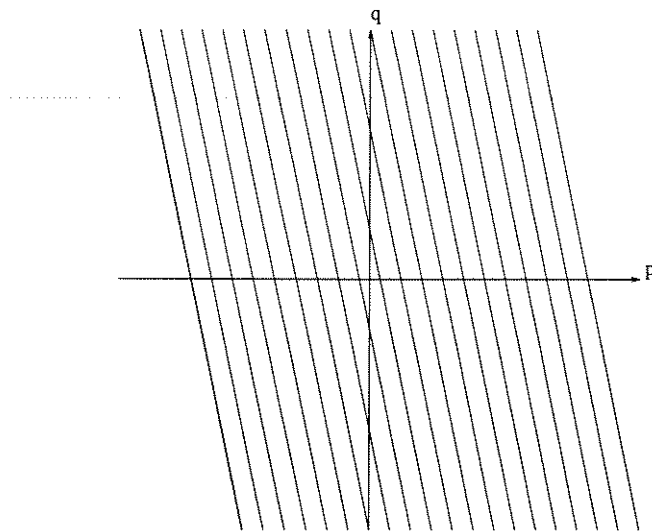


Figura 2.12: Mapa de refletância para superfície Lunar de intensidade constante.

2.6 Considerações Sobre o Mapa de Refletância

Pode-se verificar, por meio das equações de *shading* que resolver o problema de reconstrução através do mapa de refletância¹⁵ é impossível, a não ser com algumas considerações, pois matematicamente existem duas incógnitas (p , q) para se definir a direção da normal.

Esta ambigüidade pode ser verificada na Figura 2.13 (a) e (b). O vetor normal pode estar apontado para qualquer direção em torno da direção da fonte de luz, pois na determinação do vetor normal (ou da função co-seno) calcula-se o seu valor absoluto com a direção da fonte de luz. Na realidade, forma-se um cone de "confusão" que, no mapa de refletância, pode ser qualquer valor de p e q para uma dada iso-linha (intensidade).

Desta forma, para se resolver esta questão, imposições em relação à superfície devem ser realizadas, particularizando a solução. O uso do mapa de refletância, como um meio de especificar a dependência do brilho sobre a orientação da superfície imageada, só pode ser adotado devido às considerações (simplificações) impostas, isto é, o observador e a fonte de luz estão afastados do objeto.

¹⁵A técnica de SFS usa como ferramenta básica o mapa de refletância.

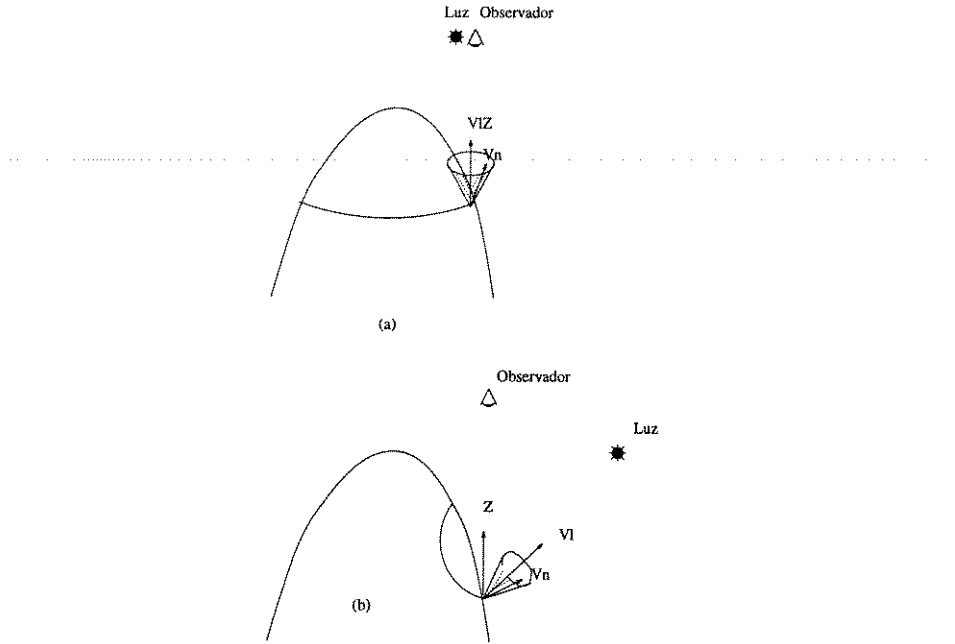


Figura 2.13: (a) ambigüidade provocada quando a fonte de luz está próxima do observador; (b) ambigüidade provocada quando a fonte de luz afastada do observador, adaptado de Kimmel e Bruckstein [27].

2.7 Influência da Componente Especular

A componente especular é de fundamental importância, no sentido de permitir o processamento de imagens mais realistas. Quando a luz incide sobre uma superfície, uma parte dela é refletida. Considerando uma superfície perfeitamente lisa (espelhada), a luz é refletida somente em uma direção, tal que o ângulo de incidência é igual ao ângulo de reflexão. No sentido de processar imagens reais, há a necessidade da consideração da especularidade, pois com isto pode-se simular (processar), além das superfícies matte, também as superfícies brilhantes (espelhadas).

A radiância de um objeto opaco que não emite sua própria radiação é determinada pela irradiância incidente num pequeno *patch* da sua superfície e suas propriedades óticas com respeito à reflexão. Geralmente, esta relação pode ser expressa como a razão entre a radiância emitida sob os ângulos polar e azimutal (θ_e e ϕ_e) e pela irradiância recebida sob os ângulos θ_i e ϕ_i . Esta razão é chamada de função da distribuição da reflectância bidirecional (*bidirectional reflectance distribution function*- BRDF), e pode ser expressa, segundo Jähne [24], por:

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{\partial L(\theta_e, \phi_e)}{\partial L(\theta_i, \phi_i)}$$

Uma definição apropriada da BRDF é de fundamental importância no sentido de proporcionar imagens mais realistas em computação gráfica (Foley et alli [10] e Hall [14]).

Numa superfície especular, a fração de luz que incide sobre a superfície tem uma parcela refletida com maior ou menor intensidade, dependendo do tipo de superfície, conforme a Figura 2.14. No caso de uma superfície ideal, ela formaria uma perfeita imagem virtual da fonte de luz. Um dos modelos utilizados para representar este fenômeno é o de Phong, segundo Rogers [36] e Hall [14].

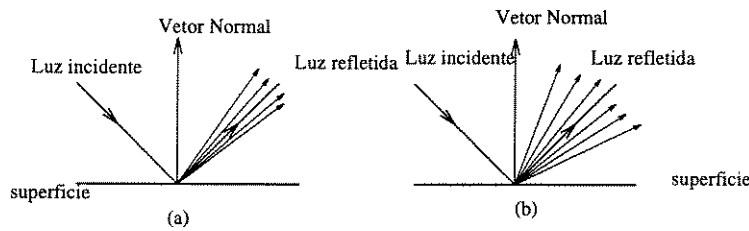


Figura 2.14: (a) Superfície especular maior intensidade de reflexão; (b) Superfície especular menor intensidade de reflexão.

2.7.1 Modelo Matemático de Iluminação Especular

Muitas superfícies têm alguma componente especular ou reflexão espelhada, ou seja, elas não são completamente *matte*. Numa perfeita reflexão especular $\mathbf{i} = \mathbf{e}$, os vetores incidentes, emitidos e a normal estão no mesmo plano ($\mathbf{i} + \mathbf{e} = \mathbf{g}$).

Mas, na prática, poucas superfícies são perfeitamente especulares. Assim, uma boa aproximação dada por Horn em [18] é obtida combinando-se a componente *matte* com a especular através da equação 2.27:

$$I_{i,j} = (1 - s)I + \frac{1}{2}s(n + 1)(2IE - G)^n \quad (2.27)$$

onde:

s varia entre zero (0) e um (1), determinando a fração da luz incidente que é refletida;

n determina o grau de especularidade da superfície (pico especular); e

$G = \cos(g)$.

Este modelo matemático produz o mapa de refletância mostrado na figura 2.15, onde se verifica uma ligeira mudança em relação ao mapa ilustrado na figura 2.11, produzido pela inserção da componente especular e pronunciado pelo parâmetro n .

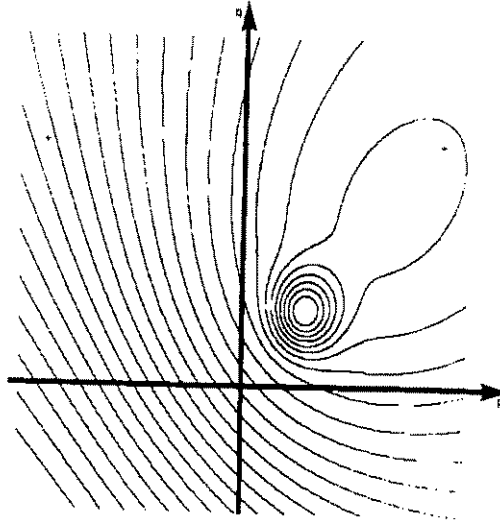


Figura 2.15: Mapa de refletância para superfície especular, obtido de Horn e Sjöberg [21]

2.7.2 Modelo de Phong

O modelo de iluminação de Phong (*Shading*) foi desenvolvido em 1975 na área da computação gráfica. Este modelo considera os efeitos das reflexões difusa e especular, sem considerar a questão física do fenômeno (modelo empírico). O modelo matemático é descrito por:

$$I_{i,j} = I_a K_a + \frac{I}{d + K} [K_d (V_n \cdot V_l) + K_s (V_r \cdot V_o)^n] \quad (2.28)$$

a equação 2.28 foi obtida de Rogers [36].

onde,

I_a é a intensidade de iluminação da luz ambiente;

I é a intensidade de iluminação da fonte de luz;

K_a é a constante de reflexão difusa da luz ambiente;

K_d é a constante de reflexão difusa da superfície;

K_s é a constante de reflexão especular;

d é a distância da fonte ao ponto;

K é uma constante arbitrária que evita a divisão por zero quando a fonte está adjacente a superfície;

n é a potência que aproxima a distribuição da luz refletida especularmente;

V_n é a direção do vetor normal à superfície;

V_l é a direção do vetor iluminação;

V_o é a direção do vetor observação;

V_r é a direção do vetor reflexão.

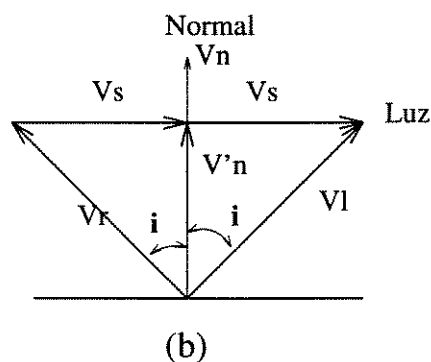
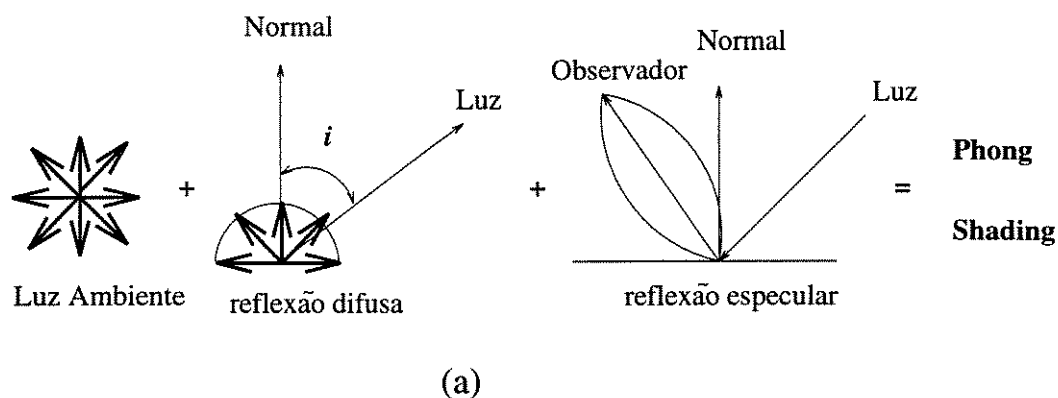


Figura 2.16: Sistema de iluminação de Phong

A figura (2.16 b) nos fornece informações sobre o cálculo do vetor reflexão V_r . Segundo Whitted (1980) apud Hall [14], calculou o vetor refletido usando álgebra vetorial e semelhança de triângulos, considerando nesta aplicação, a condição imposta pela lei de reflexão da ótica.

Multiplicando o vetor normal pelo produto escalar entre o vetor normal e o vetor incidente, cujo resultado é um vetor V'_n com o comprimento do cosseno do ângulo incidente:

$$V'_n = V_n(V_n \cdot V_l) \quad (2.29)$$

Subtraindo o vetor luz incidente do vetor V'_n resulta um vetor seno:

$$V_s = V'_n - V_l \quad (2.30)$$

da figura 2.16 b, verifica-se que o vetor seno e cosseno são ortogonais. Assim, por semelhança de triângulos, o vetor seno é adicionado ao vetor cosseno resultando no vetor reflexão V_r :

$$V_r = V'_n + V_s = 2V_n(V_n \cdot V_l) - V_l \quad (2.31)$$

2.8 Soluções Analisadas da Técnica de SFS

Analisando a bibliografia existente na área pode-se classificar os trabalhos de SFS basicamente em dois grupos: o primeiro grupo ([6], [5], [28], [39]) que determina a forma da superfície calculando o gradiente e, conseqüentemente, a normal à superfície, procedimentos estes que são baseados na técnica de SFS convencional; o segundo grupo ([34], [33] e [30]) que determina as coordenadas cartesianas dos pontos, fornecendo, assim, a reconstrução da superfície.

2.8.1 SFS com Projeção Ortogonal

Os métodos de reconstrução da superfície através de SFS com uma única imagem apresentam algumas insuficiências de informações. Para suprir estas insuficiências, várias técnicas para resolver este problema de SFS foram desenvolvidas, considerando sempre a projeção ortogonal na formação da imagem.

Geralmente, como já observado no início deste capítulo, a reconstrução da superfície é dada pela determinação dos vetores normais às superfícies. Mas, em SFS, uma outra forma é comumente utilizada, que é a especificação das componentes $\mathbf{p}$ e $\mathbf{q}$, denominada de gradiente da superfície. Estas grandezas são as derivadas parciais da função altura $Z = H(x,y)$, sobre algum plano de referência perpendicular ao eixo ótico.

As grandezas $\mathbf{p}$ e $\mathbf{q}$ são as derivadas parciais da (função) superfície:

$$p = \frac{\partial Z}{\partial x} \text{ e } q = \frac{\partial Z}{\partial y} \quad (2.32)$$

Como os valores de $\mathbf{p}$ e $\mathbf{q}$ representam as componentes do gradientes da superfície (derivadas direcionais) no ponto (x, y) , os vetores $[1 \ 0 \ p]^t$ e $[0 \ 1 \ q]^t$ definem um plano. Realizando o produto vetorial obtém-se o vetor $[-p \ -q \ 1]^t$, que é um vetor normal ao plano formado por aqueles vetores. Isto quer dizer que o vetor definido pelo produto vetorial é a normal à um *patch* de superfície no ponto (x,y) . Assim, as duas notações podem ser relacionadas pela seguinte igualdade:

$$\bar{n} = \frac{1}{\sqrt{1 + p^2 + q^2}}(-p, -q, 1)^T \quad (2.33)$$

A solução do problema de SFS, geralmente, gira em torno da equação (2.34) de **irradiância da imagem**, relacionando a irradiância da imagem à radiância da cena,

$$E(i, j) = R(\bar{n}(x_{i,j}, y_{i,j})) \text{ ou } E(i, j) = R(p_{i,j}, q_{i,j}), \quad (2.34)$$

onde $E(i,j)$ é a intensidade da irradiância na imagem do ponto $(x_{i,j}, y_{i,j})$ e R é a radiância do *patch* de superfície.

A equação da irradiância da imagem é uma equação diferencial parcial não linear de primeira ordem, que pode ser expressa em termos de $\mathbf{p}$ e $\mathbf{q}$. Um método que pode ser utilizado na resolução do problema de SFS é o das faixas características, que reduz a equação diferencial parcial a um conjunto de cinco equações diferenciais ordinárias correlacionadas, para x, y, z, p e q (Brady[4] appud Horn):

$$\begin{aligned} \frac{dx}{ds} &= R_p \\ \frac{dy}{ds} &= R_q \\ \frac{dz}{ds} &= pR_p + qR_q \\ E_x &= \frac{dp}{ds} \\ E_y &= \frac{dq}{ds} \end{aligned} \quad (2.35)$$

onde: R_p é a derivada parcial de $R(p,q)$ em relação a p ; R_q é a derivada parcial de $R(p,q)$ em relação a q ; E_x é a derivada parcial de $E(i,j)$ na direção x ; E_y é a derivada parcial de $E(i,j)$ na direção y e ds é o deslocamento do ponto ao longo de uma dada direção.

A formulação básica utilizada por Horn é fundamentada no processo iterativo. Basicamente, o processo se desenvolve através do conhecimento prévio de que o ponto imagem (x_n, y_n) corresponde a um ponto na superfície, no qual os componentes do gradiente são (p_n, q_n) . Deslocando-se de ds ao longo da faixa característica, o ponto passa de (x_n, y_n) para (x_{n+1}, y_{n+1}) e, similarmente, de (p_n, q_n) para (p_{n+1}, q_{n+1}) . Segundo Horn [19], a idéia básica deste procedimento parte de um pequeno deslocamento dx e dy provocado no ponto inicial sobre a superfície. Este deslocamento provoca uma alteração na coordenada Z de dz , que pode ser calculada a partir das equações 2.35.

Uma solução particular destas equações gera as chamadas curvas características sobre a superfície. A projeção desta curva sobre a imagem é chamada de base característica. Este procedimento, basicamente, calcula, a partir de um dado ponto (singular) conhecido e as cinco equações (Eqs. 2.35) diferenciais, um novo valor para os parâmetros $(x, y, z, p$ e $q)$ deslocado de ds .

Segundo Szeliski [39], Horn desenvolveu um procedimento que calcula simultaneamente a altura Z e as componentes do gradiente p e q , combinando nesta formulação três condições (brilho, suavidade do gradiente e integrabilidade) numa única função que é então minimizada, ou seja, usa-se do cálculo variacional para determinar os parâmetros da equação de Euler ($Z(x,y) = 0$).

A técnica de Szeliski [39] é baseada na mesma formulação de Horn. Entretanto, ao invés de usar o cálculo variacional, ele utiliza as diferenças finitas no ambiente discretizado e minimiza a nova função diretamente. Vários problemas com estas técnicas podem ser citados, tais como: o processamento de um grande volume de informações simultaneamente, tornando o processo crítico computacionalmente; a convergência do processo que se torna lenta sem auxílio de outras técnicas, segundo Szeliski [39] e a necessidade de valores aproximados aos parâmetros.

Brooks e Horn [5] utilizam-se de um processo iterativo para calcular o vetor normal unitário da superfície juntamente com a posição da fonte de luz. Basicamente, este procedimento calcula uma nova normal tomando uma média local e ajustando-a na direção da fonte de luz, conforme a equação de *Shading*. Da mesma forma que a anterior, o procedimento pode se tornar de lenta convergência. Pode-se verificar, ainda, a dificuldade na atribuição dos valores iniciais para a normal, para a posição da fonte de luz e informações de bordas.

Vários outros procedimentos, usando basicamente a mesma concepção (cálculo variacional no modo discretizado) e variando apenas na parametrização, são citados em Brooks e Horn [5]: método de Strat, método de Ikeuchi e Horn e o método de Smith. Todos os procedimentos citados são iterativos, partindo-se assim de um conhecimento inicial dos valores dos parâmetros ((p,q), (f,g) ou (l,m)), havendo, ainda, uma dependência do conhecimento do contorno do objeto.

Segundo Vega e Yang [41], um dos grandes problemas destes algoritmos fundamentados no cálculo das variações usada por Horn e outros, é que há um conflito entre a necessidade do uso da condição de intensidade e suavidade. A intensidade é uma grandeza observada, enquanto que a suavidade é idealizada (imposta). Nestes casos, somente com esta imposição o problema pode ser resolvido. Assim sendo, Vega e Yang [41] propuseram um algoritmo, também iterativo, em que o cálculo das normais numa primeira etapa é realizada usando-se uma ponderação dos "n" vizinhos mais próximos, onde o vizinho com maior aproximação (menor erro) da equação de *Shading* com a intensidade teria um peso maior. O segundo cálculo é realizado aumentando ou diminuindo o valor do vetor normal (no eixo Z), no caso de 10 %, escolhendo, assim, a que resultar menor erro. Finalmente, na terceira etapa calcula-se a normal complementar e, da mesma forma que no caso anterior, escolhe-se a que resultar menor erro.

Kimmel e Bruckstein [28] exploram o contorno de igual intensidade existente na imagem para eliminar a ambigüidade provocada na técnica de SFS. Kimmel e Bruckstein [27] apresentam um algoritmo de reconstrução global de SFS, baseado em propriedades topológicas de superfícies suaves, classificando pontos singulares na imagem como pontos de máxima, mínima ou pontos de sela (*saddle*) que servem como parâmetros (conhecidos) na solução do problema.

Cryer et al [8] exploram um novo procedimento no qual se utilizam de imagens reconstruídas por dois procedimentos: técnicas de SFS e estéreo-visão. Nesta técnica eles associam a imagem reconstruída pela estéreo-visão ao filtro passa-baixa e a obtida por SFS com o filtro passa-alta. Assim, o processamento destas duas imagens reconstruídas são processadas no domínio da frequência, produzindo uma imagem de melhor qualidade.

A maioria destes procedimentos tem pontos críticos de fundamental importância no desenvolvimento do processo. Estes pontos críticos são, escolha de um ponto singular para iniciar o processo recursivo e a definição do tipo de superfície (côncava ou convexa) a ser desenvolvida a partir deste ponto.

Nota-se que a maioria das técnicas realiza o cálculo da orientação da superfície pela

normal $\bar{n}$ ou pelos componentes do gradiente $\mathbf{p}$ e $\mathbf{q}$ numa primeira etapa para, posteriormente, recuperar a superfície. Nesta segunda etapa há mais um ponto crítico, ou seja, a necessidade de um conhecimento prévio da superfície ou alguma condição de contorno.

Nos procedimentos citados anteriormente somente o gradiente da superfície é determinado. Com isto, somente a forma da superfície é definida, necessitando-se, assim, de mais uma etapa de processamento para reconstruir a superfície. Assim, devido a necessidade de informações à respeito do objeto, o método é limitado, tornando-se muito restritivo.

Portanto, pode-se verificar que, na maioria destas técnicas, o processamento se dá a um número limitado de superfícies. Ainda, conforme as simplificações iniciais (observador e fonte de luz afastados), utiliza-se a projeção ortogonal mas, em alguns casos, como em fotogrametria, em *Motion-Vision* e na robótica (em casos de imagens reais), deve-se considerar a projeção perspectiva, pois o deslocamento devido à variação de profundidade do objeto, causa erros no posicionamento do ponto.

Assim, a fim de eliminar esta deficiência, pode-se utilizar técnicas fotogramétricas (projeção perspectiva), mais especificamente, as equações de colinearidade. Estas equações relacionam matematicamente a projeção perspectiva da imagem no instante da tomada, considerando que o ponto no espaço imagem, o centro perspectivo e o ponto no espaço objeto estejam no mesmo alinhamento, isto é, sejam colineares. Quando se trabalha com uma única imagem esta técnica também apresenta uma deficiência que, neste caso, é resolver um sistema onde um ponto medido fornece duas equações com três incógnitas.

2.8.2 SFS com Projeção Perspectiva

Poucos trabalhos sobre SFS considerando projeção perspectiva são encontrados na literatura. Penna [34] propõe um algoritmo local aplicado a imagens de objetos opacos e definidos por superfícies suaves. Utiliza, para a solução do problema, um sistema de equações que relaciona a intensidade da imagem e suas derivadas com a geometria (planos osculadores) do objeto no ponto.

Penna [33] propõe um algoritmo de SFS local no qual, para os vértices do poliedro em análise, são determinadas as profundidades, ou seja, as distâncias do centro ótico (origem do sistema de referência) ao vértice no espaço objeto. A dificuldade deste procedimento é a informação inicial necessária para solução que inclui a rotulação das arestas de modo a definir se o poliedro é côncavo ou convexo.

Lee e Kuo [30] apresentam um algoritmo iterativo de SFS com projeção perspectiva baseado numa minimização global do erro, o que torna a solução não susceptível aos inevitáveis ruídos existentes em imagens reais. Segundo Lee e Kuo [30], as soluções obtidas por algoritmos de SFS com minimização do erro local, em geral, são afetadas pelos ruídos.

A idéia básica de Lee e Kuo [30] é dividir a imagem em *patches* triangulares com brilhos homogêneos em cada triângulo. Assim, cada triângulo é representado por três pontos no espaço objeto. Usando-se as equações de colinearidade e de intensidade da imagem, pode-se estabelecer um novo modelo matemático em função somente das coordenadas altimétricas. Assim, para determinar as coordenadas altimétricas, faz-se necessário a minimização da função custo global dada pela soma ao quadrado do erro do brilho sobre cada *patch* triangular projetado no domínio da imagem.

2.9 Técnica utilizada por Lee e Kuo

Esta técnica foi desenvolvida dividindo-se a superfície em *patches* triangulares e relacionando-a com um mapa de refletância linearizado. Assim, a imagem a ser reconstituída é dividida também em *patches* triangulares com brilho homogêneo, onde cada pixel corresponde a um *patch*.

2.9.1 Desenvolvimento Teórico

A equação de irradiância (*Shading*) da imagem é dada por:

$$E(x,y) = E(X,Y,Z) = R(p(X,Y), q(X,Y))$$

e o mapa de refletância pode ser escrito como:

$$R(p, q) = \begin{cases} \rho \frac{k}{\sqrt{1+p^2+q^2}} & k \geq 0 \\ 0 & k < 0 \end{cases}$$

onde:

$$k = -p \cos(\tau) \sin(\sigma) - q \sin(\tau) \sin(\sigma) + \cos(\sigma)$$

e ρ é o albedo da superfície, τ e σ são, respectivamente, os ângulos de inclinação e orientação da direção da fonte de luz.

A relação matemática entre o espaço objeto e o espaço imagem é dada pela equação de perspectiva (3.2) que reproduz matematicamente o processo de formação da imagem, fazendo a ligação entre os pontos no espaço objeto 3D e seus correspondentes pixels imageados no espaço 2D.

2.9.2 Formulação Matemática

Para uma superfície suave, existe uma correspondência biunívoca dos *patches* triangulares do espaço objeto com o plano imagem. Eles podem ser denotados por S_i e T_i , $i = 1, 2, \dots, M_t$, onde M_t é o número total de *patches* triangulares. Similarmente, há a correspondência entre os pontos projetados com a imagem denotados por P_i e p_i $i=1, 2, \dots, M_n$, onde M_n é o número total de pontos.

Para um conjunto de três vértices ($V_k = (i, j, l)$) que definem um *patch* T_k ou S_k terá o valor da intensidade da imagem E_k sobre T_k determinado por,

$$E_k = R_k(p_k, q_k) = \rho \frac{\cos(\sigma) - p_k \cos(\tau) \text{sen}(\sigma) - q_k \text{sen}(\tau) \text{sen}(\sigma)}{\sqrt{1 + p_k^2 + q_k^2}} \quad (2.36)$$

onde (p_k, q_k) é o gradiente de S_k

A normal à superfície V_{n_k} de um *patch* triangular S_k pode ser determinada como o produto vetorial dos três pontos P_i, P_j e P_l :

$$V_{n_k} = \frac{(P_j - P_i) \times (P_l - P_i)}{|(P_j - P_i) \times (P_l - P_i)|} \quad (2.37)$$

$$V_{n_k} = \frac{(DX_{ji}, DY_{ji}, DZ_{ji}) \times (DX_{li}, DY_{li}, DZ_{li})}{|(DX_{ji}, DY_{ji}, DZ_{ji}) \times (DX_{li}, DY_{li}, DZ_{li})|} \quad (2.38)$$

onde $DC_{ki} = C_k - C_i$ para $C = (X, Y, Z)$ e $k = j, l$.

Substituindo a equação 3.2 na 2.38 obtém-se:

$$V_{n_k} = \frac{(dX_{ji}, dY_{ji}, DZ_{ji}) \times (dX_{li}, dY_{li}, DZ_{li})}{|(dX_{ji}, dY_{ji}, DZ_{ji}) \times (dX_{li}, dY_{li}, DZ_{li})|} \quad (2.39)$$

onde $dC_{ki} = \frac{1}{f}(C_i Z_i - C_k Z_k)$ para $C = (x, y)$ e $k = j, l$.

A normal também pode ser relacionada ao gradiente via:

$$V_{n_k} = \frac{(-p_k, -q_k, 1)^T}{\sqrt{1 + p_k^2 + q_k^2}} \quad (2.40)$$

Assim, igualando as equações 2.39 e 2.40 obtém-se:

$$p_k = \frac{\alpha_k}{\gamma_k} \text{ e } q_k = \frac{\beta_k}{\gamma_k} \quad (2.41)$$

onde:

$$\alpha_k = f [(y_i - y_j)Z_iZ_j + (y_l - y_i)Z_iZ_l + (y_j - y_l)Z_jZ_i]$$

$$\beta_k = f [(x_j - x_i)Z_iZ_j + (x_l - x_i)Z_iZ_l + (x_l - x_j)Z_jZ_i] \quad (2.42)$$

$$\gamma_k = (x_iy_j - x_jy_i)Z_iZ_j + (x_ly_i - x_iy_l)Z_iZ_l + (x_jy_l - x_ly_j)Z_jZ_i.$$

Desta forma, observando-se as equações 2.42, verifica-se que as coordenadas dos pontos imagens (x_i, y_i) , (x_j, y_j) e (x_l, y_l) são conhecidas e α_k , β_k e γ_k são funções somente de uma das coordenadas dos três pontos Z_i , Z_j e Z_l . Reescrevendo as equações obtém-se uma função não linear do mapa de refletância com variáveis Z_i , Z_j e Z_l :

$$R_k(Z_i, Z_j, Z_l) = \rho \frac{\gamma_k \cos(\sigma) - \alpha_k \cos(\tau) \sin(\sigma) - \beta_k \sin(\tau) \sin(\sigma)}{\sqrt{\alpha_k^2 + \beta_k^2 + \gamma_k^2}} \quad (2.43)$$

A fim de resolver a equação, é necessário uma linearização do mapa de refletância, que pode ser realizada pela expansão da série de Taylor de $R_k(Z_i, Z_j, Z_l)$ sobre um valor inicial dos parâmetros (Z_i^0, Z_j^0, Z_l^0) . Desprezando-se os termos de segunda ordem e superiores, obtém-se:

$$\begin{aligned} R_k(Z_i, Z_j, Z_l) &= R_k(Z_i^0, Z_j^0, Z_l^0) \\ &+ (Z_i - Z_i^0) \frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_i} \Big|_{(Z_i^0, Z_j^0, Z_l^0)} \\ &+ (Z_j - Z_j^0) \frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_j} \Big|_{(Z_i^0, Z_j^0, Z_l^0)} \\ &+ (Z_l - Z_l^0) \frac{\partial R_k(Z_i, Z_j, Z_l)}{\partial Z_l} \Big|_{(Z_i^0, Z_j^0, Z_l^0)} \end{aligned} \quad (2.44)$$

Assim, pode-se expressar o brilho E_k sobre T_k como uma combinação linear das coordenadas que representam a profundidade dos pontos no espaço objeto Z_i , $i=1,2,...,M_n$.

Desenvolvendo o procedimento de somatória proposto por Lee e Kuo, na forma matricial, obtém-se a seguinte equação:

$$W^T W z = W^T L \quad (2.45)$$

onde:

$$W = \sum_{i=1}^{M_t} \sum_{j=1}^{M_n} w_{ij}$$

com:

$$w_{ij} = \begin{cases} \frac{\partial R_i(Z_k, Z_l, Z_m)}{\partial Z_j} \big|_{(Z_k^0, Z_l^0, Z_m^0)} & \text{se } j \in V_k = [k, l, m] \text{ de } T_k \\ 0 & \text{para outros valores} \end{cases} \quad (2.46)$$

$$z = [Z_1, Z_2, \dots, Z_{M_n}]$$

$$L = E_i - \xi_i$$

$$\xi_i = R_i(Z_k^0, Z_l^0, Z_m^0) - \sum_{j=1}^{M_n} w_{ij} Z_j^0$$

Este procedimento tem vantagens em relação aos outros já citados: não há necessidade do conhecimento prévio da superfície, possibilitando a aplicação deste em qualquer tipo de superfície Lambertiana; consideração da projeção perspectiva no processo de reconstrução; e o modo de processamento global, eliminando, em parte, os possíveis ruídos da imagem.

Algumas limitações do procedimento adotado por Lee e Kuo [30] podem ser verificadas: fonte de luz situada em posição distante, ou seja, os raios de luz com a mesma direção em todos os pontos da imagem (luz no infinito); e sistema de coordenadas centrado e orientado no observador, possibilitando o uso da equação perspectiva.

Capítulo 3

Modelo Proposto

Neste item são apresentados conceitos necessários à formulação do modelo proposto, incluindo o tratamento dos aspectos relacionados às equações de colinearidade, de *shading* e o modelo de iluminação.

3.1 As equações de colinearidade

As equações de colinearidade reproduzem matematicamente o processo de formação da imagem, fazendo a ligação entre as coordenadas dos pontos no espaço objeto (3D) e suas correspondentes coordenadas no espaço imagem (2D). A equação de colinearidade supõe que os três pontos (centro perspectivo, ponto imagem e ponto objeto) pertençam à mesma reta. Esta condição pode ser aplicada a todos os raios que passam pelo centro de projeção (ou ponto de vista ou, ainda, ponto perspectivo), produzindo, analiticamente, a imagem.

Define-se um sistema de coordenadas cartesianas com origem no centro de projeção (ver figura 3.1), denominado referencial da imagem. Pode-se relacionar o ponto $P = (X, Y, Z)^T$ com o ponto imagem $p = (x, y, -f)^T$, descritos no mesmo referencial através da equação da perspectiva.

Desta forma, conforme a figura 3.1, pode-se deduzir a equação de uma reta no espaço passante pela origem do sistema referencial cartesiano,

$$\frac{x_p}{x_p} = \frac{y_p}{y_p} = \frac{z_p}{z_p} \quad (3.1)$$

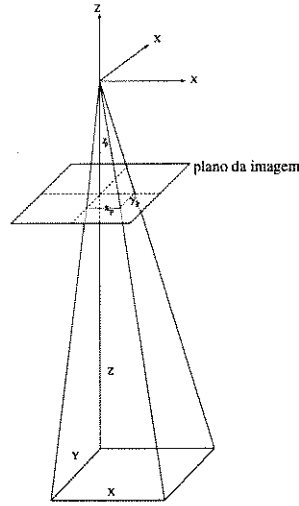


Figura 3.1: Sistema de coordenadas imagem e objeto paralelos, com origem no centro perspectivo

Reescrevendo a equação 3.1 em duas possíveis equações linearmente independentes, têm-se:

$$x_p = z_p \frac{X}{Z} \quad y_p = z_p \frac{Y}{Z} \quad (3.2)$$

Considerando, agora, o sistema de coordenadas da imagem posicionado e orientado no espaço objeto (ver figura 3.2), o ponto $P(X_P, Y_P, Z_P)$ no espaço objeto pode ser relacionado com o ponto no espaço imagem $p(x_p, y_p, z_p)$ através da equação 3.3:

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \lambda M \begin{bmatrix} X_p - X^c \\ Y_p - Y^c \\ Z_p - Z^c \end{bmatrix} \quad (3.3)$$

que é uma transformação de similaridade (ou isogonal) que realiza matematicamente a transformação entre os dois sistemas de coordenadas,

onde:

$$M = R_z(\kappa)R_y(\phi)R_x(\omega)$$

é uma matriz de rotação definida por:

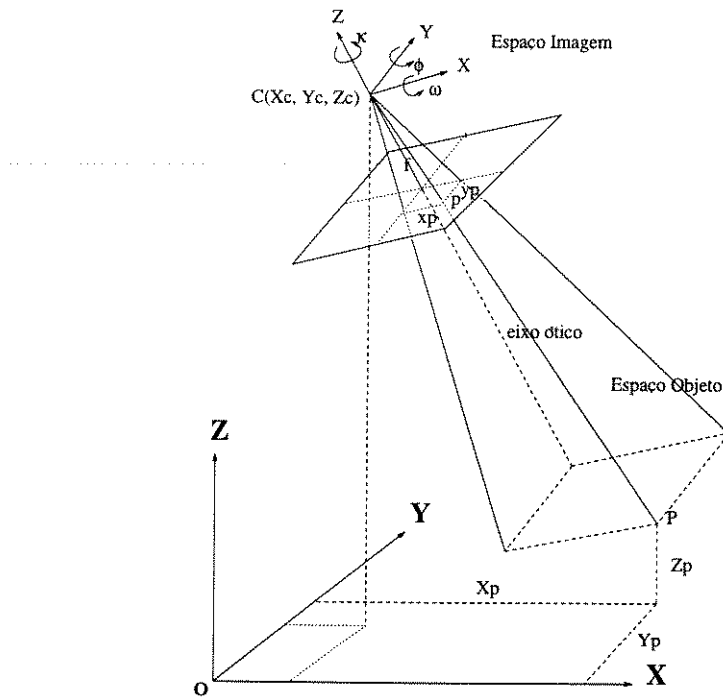


Figura 3.2: Sistemas de coordenadas da imagem e do objeto transladado e rotacionado

$$R_z(\kappa) = \begin{bmatrix} \cos(\kappa) & \sin(\kappa) & 0 \\ -\sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\phi) = \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix}$$

$$R_x(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & \sin(\omega) \\ 0 & -\sin(\omega) & \cos(\omega) \end{bmatrix}$$

assim:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (3.4)$$

com:

$$\begin{aligned}
 m_{11} &= \cos(\phi) \cos(\kappa) \\
 m_{12} &= \cos(\omega) \sin(\kappa) + \sin(\omega) \sin(\phi) \cos(\kappa) \\
 m_{13} &= \sin(\omega) \sin(\kappa) - \cos(\omega) \sin(\phi) \cos(\kappa) \\
 m_{21} &= -\cos(\phi) \sin(\kappa) \\
 m_{22} &= \cos(\omega) \cos(\kappa) - \sin(\omega) \sin(\phi) \sin(\kappa) \\
 m_{23} &= \sin(\omega) \cos(\kappa) + \cos(\omega) \sin(\phi) \sin(\kappa) \\
 m_{31} &= \sin(\phi) \\
 m_{32} &= -\sin(\omega) \cos(\phi) \\
 m_{33} &= \cos(\omega) \cos(\phi)
 \end{aligned}$$

e:

x_p , y_p e z_p são as coordenadas cartesianas de P em relação ao sistema de coordenadas da imagem (referencial do espaço imagem);

X_p , Y_p e Z_p são as coordenadas cartesianas de P em relação ao sistema de coordenadas do objeto (referencial do espaço objeto);

X^c , Y^c e Z^c são as coordenadas do centro perspectivo em relação ao sistema de coordenadas do espaço objeto;

λ é o fator de escala entre os dois sistemas de referência;

M é a matriz de rotação, determinada em função das três rotações em torno dos eixos cartesianos e;

κ , ϕ e ω são os ângulos de rotação em torno dos eixos Z, Y e X, respectivamente.

Assim, desenvolvendo-se as equações (3.3), têm-se:

$$\begin{aligned}
 x_p &= \lambda[m_{11}(X_p - X^c) + m_{12}(Y_p - Y^c) + m_{13}(Z_p - Z^c)] \\
 y_p &= \lambda[m_{21}(X_p - X^c) + m_{22}(Y_p - Y^c) + m_{23}(Z_p - Z^c)] \\
 z_p &= \lambda[m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)]
 \end{aligned} \tag{3.5}$$

Substituindo x_p , y_p e z_p , das equações (3.5) nas equações (3.2) obtém-se:

$$x_p = z_p \frac{m_{11}(X_p - X^c) + m_{12}(Y_p - Y^c) + m_{13}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)}$$

$$y_p = z_p \frac{m_{21}(X_p - X^c) + m_{22}(Y_p - Y^c) + m_{23}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)}$$

Fazendo z_p igual a distância focal (-f) (o sinal negativo é devido a compatibilização de sistemas imagem/objeto)) tem-se:

$$x_p = -f \frac{m_{11}(X_p - X^c) + m_{12}(Y_p - Y^c) + m_{13}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)} \quad (3.6)$$

$$y_p = -f \frac{m_{21}(X_p - X^c) + m_{22}(Y_p - Y^c) + m_{23}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)} \quad (3.7)$$

Com base nas equações (3.6) e (3.7), conhecidas como equações de colinearidade, pode-se formalizar os problemas de calibração da câmara e reconstrução da superfície. No caso de calibração, estas equações devem ser solucionadas para a obtenção de X^c , Y^c e Z^c e as rotações que definem a matriz M (ver equação (3.4)). No caso de reconstrução, estas equações devem ser solucionadas para a obtenção de X_p , Y_p e Z_p para todos os pontos da imagem.

3.2 Equação de *Shading*

Seja uma imagem de uma superfície Lambertiana iluminada por uma fonte de luz pontual, e obtida por um sistema no qual a formação da imagem é gerenciada pela projeção perspectiva (conforme a Figura 3.3). Para cada ponto (i,j) da imagem correspondente ao ponto objeto P(i, j), definido pelas coordenadas $X_{i,j}$, $Y_{i,j}$ e $Z_{i,j}$, determina-se um valor de intensidade luminosa pela equação (3.8), denominada equação de *shading*.

$$I_{i,j} = K_d \cos(\mathbf{i}) \quad \text{ou} \quad I_{i,j} = K_d \frac{V_n \cdot V_l}{|V_n \cdot V_l|} \quad (3.8)$$

onde:

K_d é o índice de reflexão difusa da superfície;

V_n é o vetor normal a superfície no ponto P(i, j);

V_l é o vetor direção do raio de luz que ilumina o ponto P(i,j) na superfície e;

$\mathbf{i}$ é o ângulo formado entre os vetores V_n e V_l .

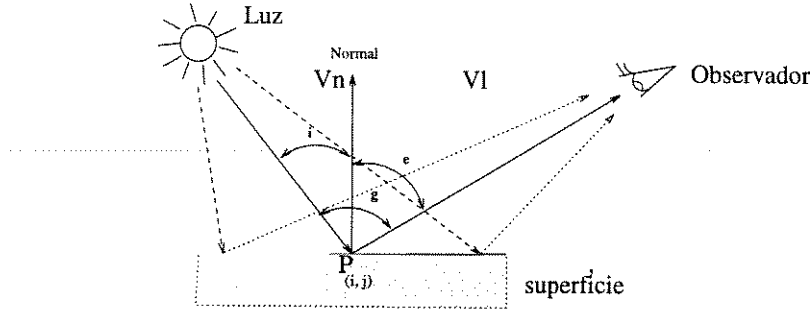


Figura 3.3: Modelo de iluminação considerando-se a superfície Lambertiana, fonte de luz pontual e projeção perspectiva.

No caso de uma fonte de luz pontual, o vetor V_l pode ser desenvolvido em função das coordenadas do ponto $P(i,j)$ da superfície e da posição da fonte de luz P_l (conforme Fig. 3.3), e pode ser escrito como:

$$V_l = P_{(i,j)} - P_l \quad (3.9)$$

ou:

$$V_l = \frac{[X_{(i,j)} - X_l]}{D_l} \vec{i} + \frac{[Y_{(i,j)} - X_l]}{D_l} \vec{j} + \frac{[Z_{(i,j)} - Z_l]}{D_l} \vec{k} \quad (3.10)$$

onde:

$$D_l = \sqrt{(X_{(i,j)} - X_l)^2 + (Y_{(i,j)} - X_l)^2 + (Z_{(i,j)} - Z_l)^2}$$

e:

$X_{i,j}$, $Y_{i,j}$ e $Z_{i,j}$ são as coordenadas do ponto $P(i,j)$ no sistema de coordenadas do espaço objeto; X_l , Y_l e Z_l são as coordenadas da posição da fonte de luz no sistema de coordenadas do espaço objeto.

Substituindo a equação 3.10 na equação 3.8, verifica-se que a intensidade da luz refletida é função de sete variáveis,

$$I_{i,j} = R_p(X_{i,j}, Y_{i,j}, Z_{i,j}, X_l, Y_l, Z_l, V_n) \quad (3.11)$$

A equação 3.11, que é uma representação simbólica da equação de *shading*, mostra que a intensidade de cada ponto é dada em função da posição da fonte de luz, do vetor normal

e das coordenadas tridimensionais do ponto $P_{i,j}$. O procedimento adotado considera que a posição da fonte de luz é conhecida, restando as determinações das coordenadas do ponto e do vetor normal.

Desta forma, pode-se admitir o conhecimento das coordenadas e do vetor normal a fim de elaborar o mapa de reflectância analiticamente, condição esta que possibilita a reconstrução da superfície.

3.3 Modelo Matemático de Reconstrução Proposto

Como pode ser verificado pelas revisões da literatura, o problema de SFS é resolvido, na maioria dos casos, utilizando-se o conhecimento do mapa de reflectância. Neste caso, o mapa de reflectância gráfico não é disponível e, no mapa de reflectância analítico, verifica-se a existência de vários parâmetros $(X_{i,j}, Y_{i,j}, Z_{i,j}, V_n)$ que não são conhecidos.

Ainda, o vetor normal V_n é dado em função de várias incógnitas:

$$(X_{i,j}, Y_{i,j}, Z_{i,j}, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1})$$

(conforme a equação 3.13), aumentando o número de incógnitas da equação 3.11. Sabe-se que o problema de reconstrução é resolvido a partir do conhecimento do mapa de reflectância. Assim, resolver o mapa de reflectância é determinar os parâmetros incógnitos da equação 3.11.

Desta forma, a inexistência do mapa de reflectância (explícito) deixaria esta situação indeterminada mas, com algumas considerações na forma de se resolver e na superfície utilizada, pode-se determiná-la sem problemas:

- considerando que a superfície a ser reconstruída é suave e contínua;
- é uma superfície Lambertiana;
- existe somente uma fonte de luz pontual e conhecida;
- a superfície é dividida em *patches* triangulares (Fig. 3.4) tendo a sua correspondência tanto no espaço imagem quanto no espaço objeto; e
- os valores aproximados das incógnitas são conhecidos *a priori*.

O modelo híbrido proposto para reconstrução da superfície e calibração da câmara será desenvolvido a partir das equações de colinearidade (Eqs. (3.6) e (3.7)) e da equação de *Shading* (Eq. (3.8), reescritas aqui para maior compreensão,

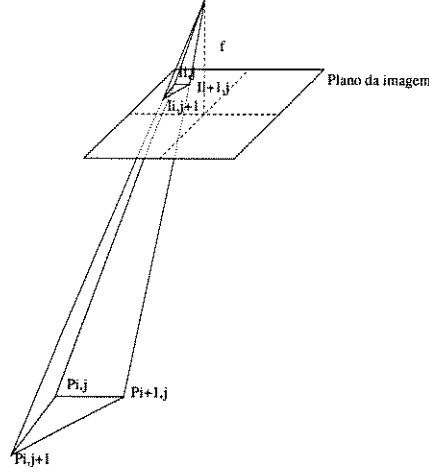


Figura 3.4: Esquema que define a correspondência entre o *patch* no espaço objeto e no espaço imagem.

$$x_p = -f \frac{m_{11}(X_p - X^c) + m_{12}(Y_p - Y^c) + m_{13}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)} \quad (3.6)$$

$$y_p = -f \frac{m_{21}(X_p - X^c) + m_{22}(Y_p - Y^c) + m_{23}(Z_p - Z^c)}{m_{31}(X_p - X^c) + m_{32}(Y_p - Y^c) + m_{33}(Z_p - Z^c)} \quad (3.7)$$

$$I_{i,j} = K_d \frac{V_n \cdot V_l}{|V_n \cdot V_l|} \quad (3.8)$$

O vetor V_n , normal à superfície no ponto $P(i,j)$, pode ser expresso, genericamente, pelo produto vetorial dos vetores definidos pelos pontos $[P(i,j), P(i+1,j)]$ e $[P(i,j), P(i,j+1)]$, o que corresponde a uma aproximação do vetor V_n pelo vetor normal relativa a um *patch* planar definido pelos três pontos considerados.

A figura 3.5 mostra um dos *patches* com os pontos envolvidos no cálculo da normal. Os índices i e j referem-se a posição do pixel na imagem, enquanto que $X_{(i,j)}, Y_{(i,j)}$ e $Z_{(i,j)}$ correspondem às coordenadas do ponto no sistema de coordenadas do objeto.

Assim, o vetor normal ao ponto $P(i,j)$ pode ser escrito como:

$$V_n = (P_{i,j} - P_{i+1,j}) \times (P_{i,j} - P_{i,j+1})$$

que na forma expandida resulta em:

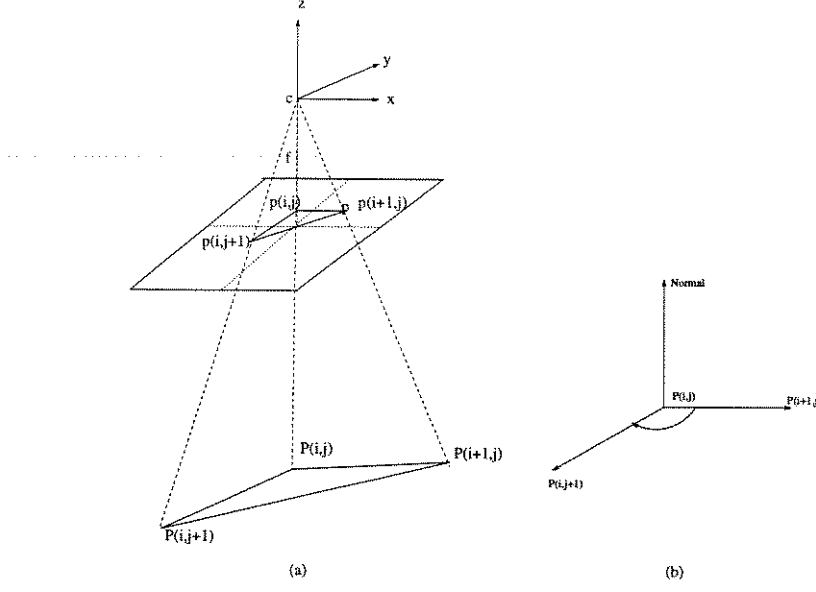


Figura 3.5: Vetores que definem a normal à superfície

$$V_n = [(Y_{(i,j)} - Y_{(i+1,j)})(Z_{(i,j)} - Z_{(i,j+1)}) - (Y_{(i,j)} - Y_{(i,j+1)})(Z_{(i,j)} - Z_{(i+1,j)})] \vec{i} +$$

$$[(X_{(i,j)} - X_{(i,j+1)})(Z_{(i,j)} - Z_{(i+1,j)}) - (X_{(i,j)} - X_{(i+1,j)})(Z_{(i,j)} - Z_{(i,j+1)})] \vec{j} + \quad (3.12)$$

$$[(X_{(i,j)} - X_{(i+1,j)})(Y_{(i,j)} - Y_{(i,j+1)}) - (X_{(i,j)} - X_{(i,j+1)})(Y_{(i,j)} - Y_{(i+1,j)})] \vec{k}$$

Substituindo V_l da equação (3.10) e V_n da equação (3.12) na equação (3.8), pode-se observar que a intensidade $I_{(i,j)}$ da imagem é função das coordenadas dos pontos da superfície. Desta forma, a equação (3.8) pode ser reescrita como:

$$I_{i,j} = F_{I_{i,j}}(X_{i,j}, Y_{i,j}, Z_{i,j}, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1}) \quad (3.13)$$

onde $F_{I_{i,j}}$ é uma representação simbólica das equações (3.8), (3.6) e (3.7).

De forma idêntica, tomando como pivôs (ponto base do cálculo da normal) os pontos $P(i+1,j)$ e $P(i,j+1)$, obtém-se, respectivamente:

$$I_{i+1,j} = F_{I_{i+1,j}}(X_{i,j}, Y_{i,j}, Z_{i,j}, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1}) \quad (3.14)$$

e

$$I_{i,j+1} = F_{I_{i,j+1}}(X_{i,j}, Y_{i,j}, Z_{i,j}, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1}). \quad (3.15)$$

Considerando o mesmo ponto $P(i,j)$, utilizado no cálculo da normal, e aplicando os índices nas equações de colinearidade (Eqs. (3.6), (3.7)), obtém-se as seguintes equações:

$$x_{i,j} = -f \frac{m_{11}(X_{i,j} - X^c) + m_{12}(Y_{i,j} - Y^c) + m_{13}(Z_{i,j} - Z^c)}{m_{31}(X_{i,j} - X^c) + m_{32}(Y_{i,j} - Y^c) + m_{33}(Z_{i,j} - Z^c)} \quad (3.16)$$

$$y_{i,j} = -f \frac{m_{21}(X_{i,j} - X^c) + m_{22}(Y_{i,j} - Y^c) + m_{23}(Z_{i,j} - Z^c)}{m_{31}(X_{i,j} - X^c) + m_{32}(Y_{i,j} - Y^c) + m_{33}(Z_{i,j} - Z^c)} \quad (3.17)$$

que podem ser representadas como uma função dos 9 parâmetros envolvidos:

$$x_{i,j} = F_{x_{i,j}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i,j}, Y_{i,j}, Z_{i,j}) \quad (3.18)$$

$$y_{i,j} = F_{y_{i,j}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i,j}, Y_{i,j}, Z_{i,j}) \quad (3.19)$$

Similarmente, para os pontos $P(i+1,j)$ e $P(i,j+1)$, as seguintes equações são obtidas, respectivamente:

$$x_{i+1,j} = F_{x_{i+1,j}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}) \quad (3.20)$$

$$y_{i+1,j} = F_{y_{i+1,j}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i+1,j}, Y_{i+1,j}, Z_{i+1,j}) \quad (3.21)$$

e

$$x_{i,j+1} = F_{x_{i,j+1}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1}) \quad (3.22)$$

$$y_{i,j+1} = F_{y_{i,j+1}}(\kappa, \phi, \omega, X_c, Y_c, Z_c, X_{i,j+1}, Y_{i,j+1}, Z_{i,j+1}) \quad (3.23)$$

Observa-se nas equações (3.13) a (3.15) e (3.18) a (3.23) que existem 6 incógnitas relativas à calibração da câmara ($\kappa, \phi, \omega, X_c, Y_c, Z_c$) que devem ser determinadas, e mais

9 incógnitas referentes às coordenadas dos pontos considerados. Na solução deste sistema, verifica-se que há uma inconsistência do modelo matemático, pois existem 15 incógnitas a serem determinadas e apenas 9 equações, condição esta que proporcionam infinitas soluções.

Para resolver esta indefinição matemática, um maior número de pontos deve ser considerado. Desta forma, uma estrutura mínima de 3 x 3 pontos (ver Figura 3.6 a) deve ser utilizada. Para uma estrutura deste tipo, verifica-se a existência de 27 incógnitas correspondentes às coordenadas 3D dos 9 pontos e mais 6 incógnitas referentes aos parâmetros de calibração, perfazendo 33 incógnitas. Por outro lado, são disponíveis 18 equações correspondentes às equações de colinearidade e mais 24 equações correspondentes às equações de *shading*, resultando em 42 equações.

Entretanto, a utilização de uma configuração mínima não permite obter uma solução confiável, podendo o resultado convergir para um valor não adequado. De uma forma geral, pode-se dizer que quanto maior o número de pontos considerados melhor o resultado obtido, pois o grau de liberdade (número de equações - número de incógnitas) é o que indica a confiabilidade do resultado.

Desta forma, a solução mais adequada para o problema é obtida pela utilização de todos os pontos da imagem, como mostrado na Figura 3.6b. Uma imagem de $n \times n$ pixels resulta num sistema com u incógnitas e v equações, com u e v definidos pelas equações (3.24) e (3.25). Assim, este método caracteriza-se por uma superabundância de equações (para imagens maiores que 3x3 pixels) e permite a utilização de um método com minimização dos erros para a solução do sistema e com uma maior probabilidade de convergência da solução para os valores corretos.

$$u = 3n^2 + 6 \quad (3.24)$$

$$v = 2(n - 1)(n - 1)3 + 2n^2 \quad (3.25)$$

Observa-se, aqui, o fato de que cada *patch* de superfície pode ser gerada a partir de três pontos distintos, gerando três equações de observação. Rigorosamente, pode-se dizer que existe somente uma observação para cada *patch*. Desta forma, o procedimento adotado neste trabalho, utiliza-se de observações adicionais¹ para resolver o sistema de equações. Esta condição leva a gerar uma superfície suavizada.

¹Duas por *patch*, que são ortogonais a primeira utilizada.

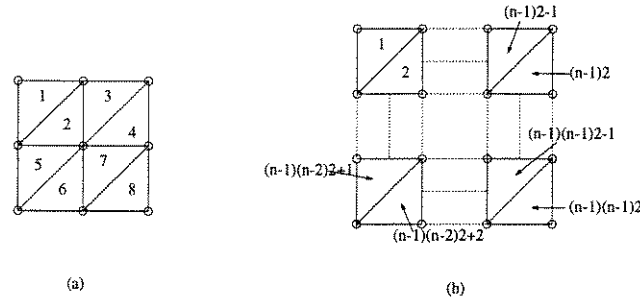


Figura 3.6: (a) Representação de uma estrutura 3x3 pontos, resultando em 8 *patches* triangulares; (b) Representação esquemática de uma estrutura de $n \times n$ pontos.

3.4 Considerações sobre as Bordas

Na solução apresentada originalmente em Hasegawa e Tozzi [17], as informações de borda não eram consideradas. Entretanto, como já observado na introdução deste trabalho (item 2.1.3), as bordas de oclusão e de sombreamento contêm informações importantes que podem ser exploradas diretamente pela imposição de injunções (condição de contorno) na solução do problema de SFS.

Segundo Ikeuchi and Horn [23] o contorno da projeção de um objeto no plano da imagem é chamado de silhueta, a qual pode ser originada pela posição do observador (borda de oclusão) ou pela posição da fonte de luz (borda de sombreamento).

3.4.1 Bordas de Sombreamento

Esta situação ocorre quando o plano tangente ao ponto na superfície contém a fonte de luz (ver Fig. 3.7). Neste caso, as bordas de sombreamento têm a sua orientação conhecida, ou seja, a direção do vetor normal à superfície e a direção da fonte de luz formam um ângulo reto (90°). Tal situação é caracterizada na imagem por regiões da silhueta com intensidade nula e valores próximos de zero nos pontos vizinhos, na direção interna à superfície imageada.

Assim, uma vez detectados na imagem os pontos que formam as bordas de sombreamento, fica estabelecida a orientação do ponto correspondente na superfície.

3.4.2 Bordas de Oclusão

Esta situação ocorre quando o plano tangente ao ponto na superfície contém o observador (ver Fig. 3.8), e é caracterizada na imagem por regiões da silhueta com valores de

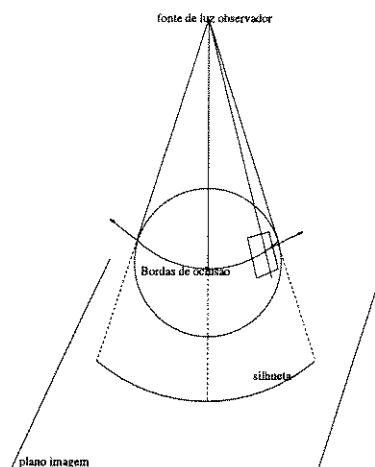


Figura 3.7: Silhueta gerada pela borda com observador e fonte de luz na mesma posição.

intensidade não nula. Contrariamente ao observado para a borda de sombreamento, neste caso nada se pode afirmar sobre a orientação do ponto correspondente na superfície.

3.4.3 Caracterização das Bordas

Uma vez detectados na imagem os pontos que formam as silhuetas, fica pendente o estabelecimento de sua classificação em bordas de oclusão ou sombreamento. A definição de limiares adequados para separação das bordas é dependente da imagem e sua determinação não é imediata.

Para superar a dificuldade na classificação dos pontos da silhueta é utilizada a característica do procedimento aplicado na solução do sistema de equações. Este procedimento permite ponderar a qualidade das medidas que, neste caso, são representadas pelos valores iniciais atribuídos. Para as bordas são estipulados valores de ponderação inversamente proporcionais aos valores da intensidade. As bordas de sombreamento são ponderadas com valores altos, pois possuem valores próximos de zero, enquanto que as bordas de oclusão são ponderadas com valores menores, pois as intensidade dos pixels são diferentes de zero.

No caso de silhuetas produzidas quando o observador e a fonte de luz não estão coincidentes, produz-se dois tipos de silhuetas definidas pelas bordas de sombreamento e oclusão. Os pontos relacionados com a borda de oclusão são menos ponderados e, à medida que se aproximam das bordas de sombreamento, os valores de intensidade se aproximam de zero e são mais ponderadas. Quando não existe grande variação entre a posição do

observador e da fonte de luz, as bordas de oclusão e sombreamento tornam-se coincidentes e são ponderadas com valores altos, pois suas intensidades são próximas de zero.

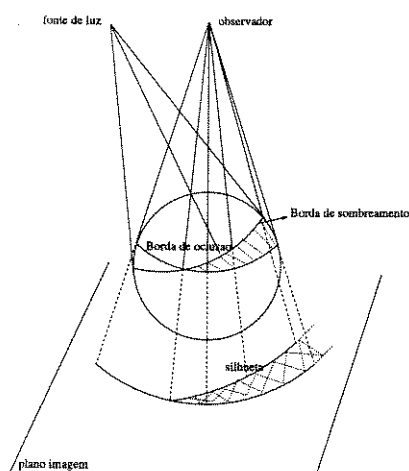


Figura 3.8: Silhueta gerada por uma borda com observador e fonte de luz em posições diferentes

3.4.4 Detecção de Bordas

A operação de detecção de bordas está inserida na segmentação da imagem e faz parte de um processo maior, que é o de obter a descrição da imagem. A descrição da imagem é proporcionada pela identificação de objetos que a compõem, que podem ser caracterizados por linhas ou regiões. Assim, esta imagem será dividida em várias regiões não sobrepostas.

A detecção das regiões numa imagem pode ser feita através de agrupamento de pontos vizinhos com características semelhantes (usando o critério de similaridade) ou através da determinação da aresta (critério de descontinuidade) da região.

A operação de detecção de bordas consiste em agrupar *elementos* de descontinuidade numa entidade contínua. A idéia básica de uma técnica de detecção de bordas é aplicar uma pequena máscara de convolução espacial a todos os *pixels* da imagem, determinando um valor que será comparado a um limiar (*thresholding*). Segundo Castelman [7], *thresholding* é uma técnica vantajosa de detecção de regiões para cenas que contêm objetos sólidos sobre um fundo contrastante, e é uma técnica computacionalmente simples, que trabalha bem quando o interior do objeto de interesse tem intensidade (tons de cinza) uniforme (característica de superfícies *matte*). Assim, aplicando-se a máscara, os valores obtidos que estiverem acima do

limiar estabelecido são considerados como pontos de borda (aresta).

Basicamente, a técnica de detecção de bordas consiste no cálculo de um operador de derivada local (máscara). Neste trabalho foi escolhido o detector de bordas baseado nos operadores de Sobel.

Assim, pode-se detectar uma borda usando as equações 3.26:

$$\delta_x = (I_{(i+1,j-1)} + 2I_{(i+1,j)} + I_{(i+1,j+1)}) \quad (3.26)$$

$$- (I_{(i-1,j-1)} + 2I_{(i-1,j)} + I_{(i-1,j+1)})$$

$$\delta_y = (I_{(i-1,j+1)} + 2I_{(i,j+1)} + I_{(i+1,j+1)}) \quad (3.27)$$

$$- (I_{(i-1,j-1)} + 2I_{(i,j-1)} + I_{(i+1,j-1)})$$

onde $I_{(i,j)}$ é a intensidade do pixel (i,j) .

As equações 3.26 podem ser caracterizadas pelas máscaras definidas pelas figuras 3.9 a e b.

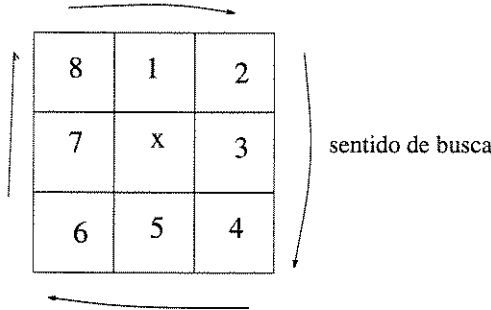


Figura 3.9: Operadores de Sobel; (a) máscara para calcular gradiente em x; (b) máscara para calcular gradiente em y;

A operação para detectar bordas consiste em aplicar os operadores a cada pixel na imagem e sua vizinhança, calculando a magnitude do gradiente pela equação 3.28.

$$G_{(i,j)} = \sqrt{\delta_x^2 + \delta_y^2} \quad (3.28)$$

onde, $G_{(i,j)}$ é a magnitude do gradiente resultante da aplicação dos operadores de Sobel ao pixel i,j .

Aplicando este modelo matemático por toda imagem, obtém-se uma imagem de gradientes que indicam as descontinuidades nos níveis de cinza. Para definir as bordas através destes gradientes, faz-se necessário estipular um valor limite (*thresholding*), que será utilizado como critério.

Existem várias formas de calcular o valor do *threshold*. Neste caso, será utilizado o processo estatístico, cujas operações são:

- cálculo da média aritmética (m) da imagem de gradientes, equação 3.29;
- cálculo do desvio padrão da imagem de gradientes, conforme a equação 3.30; e
- cálculo do limiar (*threshold*) na imagem de gradientes, definido como a soma do valor médio e o desvio-padrão, equação 3.31.

$$m = \frac{\sum_{i=0}^{i=n} \sum_{j=0}^{j=n} G_{(i,j)}}{n} \quad (3.29)$$

$$\sigma = \sqrt{\frac{\sum_{i=0}^{i=n} \sum_{j=0}^{j=n} (G_{(i,j)} - m)^2}{n^2}} \quad (3.30)$$

$$T = m + \sigma. \quad (3.31)$$

Desta forma, aplicando-se as equações (Eqs. 3.29, 3.30 e 3.31) na imagem de intensidade, a imagem resultante é uma imagem binária, onde os valores unitários correspondem aos limites das regiões, conforme a igualdade mostrada na equação 3.32:

$$B_{(i,j)} = \begin{cases} 1 & G_{(i,j)} \geq T \\ 0 & G_{(i,j)} < T \end{cases} \quad (3.32)$$

Este procedimento de determinar o *threshold* tem a praticidade de poder ser implementado facilmente e possibilitar o processamento automático. Entretanto, uma das deficiências deste método é a detecção falsa de bordas nas imagens, principalmente nas reais (com ruídos), pois, devido ao procedimento estatístico, os ruídos e as variações de intensidade nas imagens são consideradas, dependendo da amplitude, como descontinuidades.

3.5 Método dos Mínimos Quadrados (MMQ)

Os aspectos matemáticos utilizados na resolução do sistema de equações resultante da aplicação do modelo proposto serão discutidos nesta seção.

O sistema de equações resultante do modelo proposto envolve equações não lineares e, para sua linearização foi aplicada a técnica de Série de Taylor, desprezando-se os termos de ordem igual e superior à segunda.

As derivadas parciais das equações (3.6), (3.7) e (3.8), que dão origem ao sistema de equações resultante da aplicação do modelo proposto, são apresentadas no apêndice A.

Os aspectos concernentes à aplicação do Método dos Mínimos Quadrados (MMQ) são discutidos a seguir.

Observa-se, nas equações (3.6), (3.7) e (3.8) que são funções explícitas dos parâmetros incógnitos. Assim, todas as v funções consideradas no modelo proposto podem ser representadas, genericamente, da seguinte forma:

$$F(X_a) = L_a \quad (3.33)$$

onde, para a imagem $n \times n$:

L_a é um vetor ($vx1$) e corresponde aos valores observados ajustados (coordenadas x e y da imagem e intensidade I do pixel (i,j));

$$L_a = [x_{11}, y_{11}, x_{12}, y_{12}, \dots, x_{1n}, y_{1n}, x_{21}, y_{21}, \dots, x_{nn}, y_{nn}, P_1, P_2, \dots, P_{(n-1)2}, \dots, P_{(n-1)(n-2)2+2}, \dots, P_{(n-1)(n-1)2}]^T$$

onde:

P_w contém as observações de *shading* dos 3 pontos que definem o *patch* w , ou seja,

$P_w = (I_{(i,j)}, I_{(i,j+1)}, I_{(i+n,j)})$, e

$I_{(i,j)}$ representa a intensidade do pixel (i,j) .

X_a é um vetor ($ux1$) correspondente às variáveis incógnitas (coordenadas dos pontos no espaço objeto e parâmetros de calibração da câmara):

$$X_a = [\omega, \phi, \kappa, X_c, Y_c, Z_c, X_{11}, Y_{11}, Z_{11}, X_{12}, Y_{12}, Z_{12}, \dots, X_{1n}, Y_{1n}, Z_{1n}, X_{21}, Y_{21}, Z_{21}, \dots, X_{2n}, Y_{2n}, Z_{2n}, \dots, X_{n1}, Y_{n1}, Z_{n1}, \dots, X_{nn}, Y_{nn}, Z_{nn}]^T$$

Considerando a existência de erros nas observações, há uma diferença entre valores observados e ajustados, podendo ser reescrita da seguinte forma:

$$L_a = L_b + V \quad (3.34)$$

onde:

L_b é o vetor dos valores observados; e

V é o vetor dos resíduos.

Devido a não linearidade do sistema de equações, atribui-se valores iniciais aproximados para as incógnitas do vetor X_a , representados pelo vetor X_0 e, iterativamente, determina as correções a estes valores iniciais (equação 3.35),

$$X_a = X_0 + X \quad (3.35)$$

onde:

X_0 é um vetor (ux1) dos valores aproximados das incógnitas e

X é um vetor (ux1) das correções aos valores iniciais.

Substituindo as equações (3.34) e (3.35) na equação (3.33), tem-se:

$$L_b + V = F(X_0 + X) \quad (3.36)$$

Aplicando a técnica de linearização por série de Taylor na equação (3.36) e desconsiderando todos os elementos de segunda ordem e superiores, a equação (3.36) pode ser escrita:

$$L_b + V = F(X_0) + \frac{\partial F}{\partial X_a} \bigg|_{X_a=X_0} X \quad (3.37)$$

Definindo:

$$L_0 = F(X_0) \quad (3.38)$$

e

$$A = \frac{\partial F}{\partial X_a} \Big|_{X_a=X_0} \quad (3.39)$$

A equação (3.37) pode ser reescrita como:

$$L_b + V = L_0 + AX, \text{ e portanto : } V = AX + L_0 - L_b \quad (3.40)$$

ou:

$$V = A \cdot X + L \quad (3.41)$$

onde:

A é uma matriz (vxu) formada pelas derivadas parciais das equações envolvidas (anexo A)

e

$$L = L_0 - L_b \quad (3.42)$$

A equação (3.41) representa o modelo linearizado do método de ajustamento denominado paramétrico (MMQ) (Gemael [12]).

Para a solução do problema em questão é necessário que se determine a matriz A. Cada linha da matriz A é definida pelas derivadas parciais de uma observação que, no presente caso, corresponde às equações de colinearidade e *shading*. O modelo matemático que calcula o número de linhas e colunas na matriz A é dado pelas equações (3.25) e (3.24), respectivamente.

Desta forma, são 9 (nove) os elementos não nulos de uma linha da matriz A. Estes elementos correspondem, para a equação de colinearidade, aos 6 parâmetros de calibração e às 3 coordenadas do ponto. Para as linhas correspondentes às equações de *shading*, os 9 elementos que assumem valores diferentes de zero referem-se às coordenadas dos 3 pontos envolvidos.

Considerando a redundância (n° total de observações - n° mínimo necessário) de equações, utilizou-se o Método dos Mínimos Quadrados para a determinação de uma solução única para as incógnitas, bem como a estimativa de precisão (matriz variância-covariância).

O critério utilizado no MMQ é representado pela forma quadrática expressada pela equação (3.43), que pode ser escrita como:

$$\min_X \|V(X)\|_p^2 \quad (3.43)$$

onde $V(X) = AX + L$ e

P é a matriz peso das observações, dada por:

$$P = \sigma_0^2 \sum_{b=1}^{-1} L_b \quad (3.44)$$

com:

σ_0^2 é a variância da observação de peso unitário (variância) *a priori* e

$\sum_{L_b}^{-1}$ é uma matriz de ordem (vxv) que representa a matriz variância-covariância dos valores observados.

Supondo que as observações não são correlacionadas, o que é válido para esta aplicação, a matriz $\sum_{L_b}^{-1}$ será reduzida a uma matriz diagonal,

Substituindo o vetor V da equação (3.41) na função (3.43) e aplicando o problema dos mínimos quadrados com norma (mínima) ponderada, obtém-se:

$$\min_X (V(X)^t P V(X)) = \min_X (AX + L)^t P (AX + L) \quad (3.45)$$

$$\min_X (X^t A^t P A X + X^t A^t P L + L^t P A X + L^t P L) = \min_X f(X)$$

aplicando a condição de mínimo: $\frac{\partial f(X)}{\partial X} = 0$

$$\frac{\partial f(X)}{\partial X} = 2A^t P A X + 2A^t P L = 0 \quad (3.46)$$

$$A^t P A X + A^t P L = 0 \quad (3.47)$$

resultando em,

$$X = -(A^t P A)^{-1} (A^t P L) \quad (3.48)$$

Fazendo:

$$N = (A^t P A) \quad (3.49)$$

e

$$U = (A^t P L) \quad (3.50)$$

resulta em:

$$NX + U = 0 \quad (3.51)$$

Relembrando que a solução é obtida iterativamente, isto é, atribui-se valores iniciais (aproximados) às incógnitas e determina-se as correções a estes valores a cada passo. As iterações são realizadas, usando as equações (3.35) e (3.52), até que a convergência (para o critério estabelecido) seja atingida. De acordo com a (3.52),

$$X_0^i = X^{i-1} + X_0^{i-1} \quad (3.52)$$

a convergência pode ser definida como $X^{i-1} < \varepsilon$, onde ε é o critério.

Segundo Mikhail e Ackermann [31], a princípio, nem todos os processos iterativos convergem. Alguns podem divergir, outros podem oscilar ou repetir os resultados e ainda, alguns podem convergir, mas não no valor esperado. Assim, o critério de convergência adotado deve ser um indicador representativo para esta tarefa. Desta forma, eles citam vários critérios que podem ser adotados como convergência dos métodos, dentre eles o que foi utilizado neste trabalho: a estabilização da forma quadrática $V^t P V$ que implicitamente está ligada à convergência dos parâmetros ajustados.

Com o uso do MMQ, pressupondo que nenhum erro (ruído, modelos matemáticos e parâmetros aproximados impróprios) afete os resultados, pode-se adotar como critério de convergência a condição de que as correções são menores a cada iteração e que, no final do ajustamento, elas serão nulas. Portanto, pode-se dizer que, quando as correções dos parâmetros se tornarem "insignificantes", segundo um valor pré-definido, a solução do sistema será determinada.

3.5.1 Matriz Variância-Covariância

No MMQ foi utilizada a matriz peso P (Eq. 3.44), que é composta a partir do modelo estocástico das medidas (observações) efetuadas. Assim, após a estimação dos parâmetros, pode-se estimar a variância da observação de peso unitário (variância a posteriori) e a matriz variância-covariância destes parâmetros.

Matriz Variância-Covariância das Incógnitas X e X_a

Substituindo o valor de U da equação 3.50 na equação 3.49 e fazendo $L = L_0 - L_b$, obtém-se:

$$X = -N^{-1}A^tPL_0 + N^{-1}A^tPL_b \quad (3.53)$$

Aplicando a lei da propagação das covariâncias, (Gemael [12])

$$\sum X = G \sum L_b G^t \quad (3.54)$$

onde,

$\sum X$ representa a matriz variância-covariância das incógnitas X ; e $\sum L_b$ representa a matriz variância-covariância das observações.

Fazendo,

$$G = N^{-1}A^tP \quad (3.55)$$

e

$$G^t = PAN^{-1} \quad (3.56)$$

devido a P e N^{-1} serem matrizes simétricas.

Substituindo G e G^t das equações 3.55 e 3.56 na equação 3.54 obtém-se, após algumas operações:

$$\sum X = \bar{\sigma}_0^2 N^{-1} \quad (3.57)$$

A matriz variância-covariância dos parâmetros estimados ajustados é obtida aplicando-se a lei da propagação das covariâncias na equação 3.35. Observa-se, entretanto, que o vetor X_0 na equação 3.35, por ser um valor aproximado, não apresenta características estocáticas. Logo,

$$\sum X_a = \sum X = \bar{\sigma}_0^2 N^{-1} \quad (3.58)$$

Onde $\bar{\sigma}_0^2$ é a variância da observação de peso unitário a *posteriori*, que é determinada por:

$$\bar{\sigma}_0^2 = \frac{V^t P V}{u - v} \quad (3.59)$$

3.6 Injunções para a Calibração da Câmara

A seqüência apresentada na Seção 3.5 não leva a resolver o problema completamente, visto que há uma singularidade² nas equações (3.48). A solução para o problema pode ser obtida somente com o prévio conhecimento de alguns pontos (coordenadas) da superfície que será reconstruída e que serão introduzidos como injunções, no caso, como injunções posicionais.

Mikhail e Ackermann [31] definem injunções (*constraint* - condição de contorno) como sendo uma equação que relaciona somente os parâmetros entre si. E a sua presença no modelo matemático funcional implica em que os parâmetros sejam dependentes funcionalmente.

Na prática, ocorre quando alguns ou todos os parâmetros do ajustamento devem satisfazer relações estabelecidas a partir das características físicas ou geométricas do modelo. Como exemplos de injunções pode-se citar o conhecimento da distância entre dois pontos que aparecem numa imagem, pontos estes situados sobre uma reta entre outros.

Genericamente, a injunção ($G(X_a)$) pode ser representada em forma matricial pela seguinte equação:

$$G(X_a) = R_{X_a} \quad (3.60)$$

²A singularidade, neste caso, é devida à indefinição de um sistema de referência, pois as equações arroladas no modelo estão relacionadas a um sistema de coordenadas global.

onde:

R_{X_a} é um vetor (ux1) contendo os valores dos parâmetros de injunção observados ajustados.

Realizando as mesmas considerações no desenvolvimento da equação 3.33, a equação 3.60 pode ser escrita como:

$$G(X_a) = R_{X_a} = R_{X_b} + \bar{V} \quad (3.61)$$

onde:

R_{X_b} é um vetor (ux1) dos valores de injunção observados e

$\bar{V}$ é o vetor (ux1) dos resíduos das injunções.

Linearizando a equação (3.61), tem-se:

$$R_{X_b} + \bar{V} = G(X_0) + \frac{\partial G}{\partial X_a} \bigg|_{X_a=X_0} X \quad (3.62)$$

Fazendo as mesmas considerações do caso anterior, tem-se:

$$\bar{V} = C \cdot X + R \quad (3.63)$$

onde:

$R = R_0 - R_{X_b}$, que representa a diferença entre a injunção calculada em função dos parâmetros estimado e o valor fixado pela injunção;

C é uma matriz formada pelas derivadas parciais das equações de injunção em relação aos parâmetros.

Aplicando o MMQ nas equações (3.41) e (3.63), agora considerando as injunções, resulta a equação (3.64):

$$\min_X (V^t P V + \bar{V}^t \bar{P} \bar{V}) \quad (3.64)$$

Minimizando a forma quadrática (3.64), obtém-se a equação 3.65:

$$X = -(A^t P A + C^t \bar{P} C)^{-1} (A^t P L + C^t \bar{P} R) \quad (3.65)$$

Assim, a equação 3.65 é a equação normal geral (com injunções) do método paramétrico para se realizar a estimação com MMQ.

Considerando o caso em que as injunções são definidas por três coordenadas de um simples ponto $P(i,j)$, as equações de injunções $G(X_a)$, para estas coordenadas, podem ser expressas da seguinte forma:

$$\begin{aligned} X_{i,j} &= \overline{X_{i,j}} \\ Y_{i,j} &= \overline{Y_{i,j}} \\ Z_{i,j} &= \overline{Z_{i,j}} \end{aligned} \quad (3.66)$$

onde:

$\overline{X_{i,j}}, \overline{Y_{i,j}}, \overline{Z_{i,j}}$ são as coordenadas de injunções do ponto $P(i,j)$ no sistema de coordenadas do espaço objeto; e

$X_{i,j}, Y_{i,j}, Z_{i,j}$ são as variáveis incógnitas correspondentes as coordenadas ajustadas do ponto $P(i,j)$ no sistema de coordenadas do espaço objeto;

e a matriz peso ($\overline{P}$) terá a seguinte forma:

$$\overline{P} = \begin{bmatrix} P_{X_{i,j}} & 0 & 0 \\ 0 & P_{Y_{i,j}} & 0 \\ 0 & 0 & P_{Z_{i,j}} \end{bmatrix} \quad (3.67)$$

e:

$$P_{X_{i,j}} = \frac{1}{\sigma_{X_{i,j}}^2}; P_{Y_{i,j}} = \frac{1}{\sigma_{Y_{i,j}}^2}; P_{Z_{i,j}} = \frac{1}{\sigma_{Z_{i,j}}^2} \quad (3.68)$$

onde:

$\sigma_{X_{i,j}}^2, \sigma_{Y_{i,j}}^2, \sigma_{Z_{i,j}}^2$ são as variâncias das coordenadas dos pontos de injunção.

Os valores das variâncias indicam a precisão da determinação dos pontos envolvidos. Quando a variância é nula, o peso ($P \rightarrow \infty$) tende para o infinito. Esta condição pode ser entendida como injunção absoluta, que expressa muita confiança na determinação dos pontos de controle. Em caso contrário, é definida como injunção relativa e as variâncias assumem valores diferentes de zero (0); quando elas assumirem valores muito altos, tendendo para o infinito ($\sigma^2 \rightarrow \infty$), os pesos tendem a zero (0), indicando, aqui, a inexistência de injunções.

A matriz C , que é formada pelas derivadas parciais da função injunção em relação aos parâmetros, as mesmas consideradas na formação da matriz A , terá o mesmo número de colunas da matriz A e terá tantas linhas quantas forem as injunções. No problema em questão, dadas as equações 3.66, a matriz C terá elementos com valor igual a 1, na posição correspondente à injunção efetuada, e zero, em outras posições.

$$\text{posição na matriz} \Rightarrow \begin{matrix} 1 & 2 & . & . & i_j & i+1_j & i+2_j & . & . & . & v \end{matrix}$$

$$C = \begin{bmatrix} 0 & 0 & . & . & 1 & 0 & 0 & . & . & 0 & 0 \\ 0 & 0 & . & . & 0 & 1 & 0 & . & . & 0 & 0 \\ 0 & 0 & . & . & 0 & 0 & 1 & . & . & 0 & 0 \end{bmatrix} \quad (3.69)$$

Assim, na matriz $\bar{N} = C^t \bar{P} C$ os elementos não nulos estarão sobre a diagonal, em posição correspondente à posição dos parâmetros fixados, e estes valores serão o peso da injunção. Da mesma forma, o vetor $\bar{U} = C^t \bar{P} \bar{L}$ terá também elementos não nulos nas posições correspondentes aos parâmetros fixos, valor este que será o produto do peso pela diferença entre os valores ajustados e fixados.

3.7 Injunções Relativas às Bordas de Oclusão e Sombreamento

No caso das informações adicionais consideradas neste trabalho, tendo sido determinadas as bordas de oclusão ou de sombreamento, os pontos que as constituem serão utilizados como pontos de injunção. O modelo matemático funcional para as injunções será determinado pelo produto interno do vetor direção da luz e do vetor normal à superfície no ponto.

O modelo matemático desta injunção será definido por:

$$\frac{V_n \cdot V_l}{|V_n||V_l|} = \cos(90^\circ) = 0 \quad (3.70)$$

com:

$$V_l = [(X_{(i,j)} - X_l), (Y_{(i,j)} - Y_l), (Z_{(i,j)} - Z_l)]$$

$$V_n = [n_x, n_y, n_z].$$

Neste caso, o vetor normal V_n será calculado para todos os *patches* que incluem o ponto considerado, e o valor escolhido para a normal ao ponto será o que produzir o maior ângulo entre o vetor normal e vetor direção da fonte de luz.

A equação (3.70) pode ser reescrita na forma funcional

$$H(X_a) = S_{X_a} \quad (3.71)$$

onde:

S_{X_a} é um vetor (ux1) contendo os valores dos parâmetros de injeção observados; e $H(X_a)$ representa uma função (injeção) que descreve uma condição a ser imposta.

Fazendo as mesmas considerações apresentadas no item 3.6, tem-se, para a equação de injeção:

$$B = \frac{\partial H}{\partial X_a} \Big|_{X_a=X_0}. \quad (3.72)$$

A matriz B é formada pelas derivadas parciais da equação (3.70) em relação aos parâmetros variáveis $(X_{(i,j)}, Y_{(i,j)}, Z_{(i,j)})$. Cada linha da matriz B é definida por uma equação, que identifica o ponto pertencente a silhueta que descreve a borda. O número de colunas da matriz B é igual ao da Matriz A, pois os parâmetros envolvidos nesta injeção são os mesmos.

Assim, desenvolvendo da mesma forma que no item 3.6, obtém-se a equação (3.73):

$$X = -(A^t P A + C^t \bar{P} C + B^t P_b B)^{-1} (A^t P L + C^t \bar{P} R + B^t P_b S) \quad (3.73)$$

que é o modelo geral (com as injeções) para o processamento com MMQ do modelo proposto, considerando também as informações de borda.

3.8 Solução Implementada

Nesta seção, os algoritmos e as otimizações implementadas serão discutidos e analisados. O procedimento adotado para resolver o sistema de equações, aplicando o MMQ, consiste das 13 etapas descritas a seguir:

1. obter o vetor dos parâmetros aproximados X_0 ;
2. calcular $L_0 = F(X_0)$;

3. calcular $L = L_0 - L_b$;
4. montar a matriz P (peso);
5. montar a matriz normal $N = A^t P A$;
6. montar o vetor $U = A^t P L$;
7. inserir a injeção na forma $\bar{N} = C^t \bar{P} C$;
8. inserir o vetor $\bar{U} = C^t \bar{P} L$;
9. inserir a injeção na forma $N_b = B^t P_b B$;
10. inserir o vetor $U_B = B^t P_b S$;
11. resolver o sistema de equações $X = (N + \bar{N} + N_b)^{-1}(U + \bar{U} + U_b)$;
12. calcular o vetor dos parâmetros ajustados $X_a = X_0 + X$;
13. verificar a convergência $X_i \leq \text{critério}$.

Caso a convergência não seja atingida, faz-se $X_0 = X_a$ e repete-se o processamento a partir da segunda etapa. Caso a convergência seja atingida, a solução será dada pelo último X_a .

A determinação dos valores iniciais X_0 (etapa 1) é feita a partir de uma transformação geométrica com doze parâmetros de transformação, como nas equações 3.74, onde os valores dos pontos conhecidos (pontos de calibração) são usados na determinação dos parâmetros desta transformação. Um mínimo de 4 pontos conhecidos, nos dois sistemas (imagem e objeto), são necessários para determinar os parâmetros de transformação.

$$\begin{bmatrix} x_{i,j} \\ y_{i,j} \\ I_{i,j} \end{bmatrix} = \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \\ \mathbf{d} & \mathbf{e} & \mathbf{f} \\ \mathbf{g} & \mathbf{h} & \mathbf{i} \end{bmatrix} \times \begin{bmatrix} X_{i,j} \\ Y_{i,j} \\ Z_{i,j} \end{bmatrix} + \begin{bmatrix} j \\ k \\ l \end{bmatrix} \quad (3.74)$$

onde:

$X_{i,j}$, $Y_{i,j}$, and $Z_{i,j}$ são as coordenadas do ponto conhecido no sistema de coordenadas do espaço objeto, usado na calibração da câmara;

$x_{i,j}$ and $y_{i,j}$ são as coordenadas do ponto conhecido no sistema de coordenadas do espaço imagem, usado na calibração da câmara;

$I_{i,j}$ é a intensidade do pixel considerado;

$\mathbf{a}$, $\mathbf{b}$, ..., $\mathbf{k}$ and $\mathbf{l}$ são as variáveis incógnitas da transformação.

Desta forma, conhecendo-se os parâmetros da transformação (Eq. (3.74)), aplica-se a transformação inversa (Eq. (3.75)) em todos os pontos da imagem, obtendo-se os valores iniciais para os pontos no espaço objeto.

$$\begin{bmatrix} X'_{i,j} \\ Y'_{i,j} \\ Z'_{i,j} \end{bmatrix} = \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \\ \mathbf{d} & \mathbf{e} & \mathbf{f} \\ \mathbf{g} & \mathbf{h} & \mathbf{i} \end{bmatrix}^{-1} \times \begin{bmatrix} x_{i,j} - j \\ y_{i,j} - k \\ I_{i,j} - l \end{bmatrix} \quad (3.75)$$

onde:

$X'_{i,j}$, $Y'_{i,j}$, and $Z'_{i,j}$, são as coordenadas iniciais aproximadas do ponto $P(i,j)$.

O modelo matemático 3.74 utilizado realiza a transformação geométrica separadamente nas três coordenadas, mesmo com o processamento simultâneo, condição esta que é necessária, pois as duas primeiras informações do ponto são geométricas e a terceira é radiométrica.

Os valores iniciais dos parâmetros de calibração são atribuídos manualmente no arquivo de dados. Os dois parâmetros angulares, ϕ e ω , que representam a inclinação da câmara em relação ao sistema de coordenadas global recebem valores nulos. O terceiro parâmetro κ é calculado a partir das relações definidas entre as coordenadas da imagem e do objeto. Este ângulo, desvio da direção, é formado pelo eixo x do sistema de coordenadas da imagem com o eixo X do sistema de coordenadas do objeto. Os três elementos de translação são considerados conhecidos aproximadamente, cujos valores são atribuídos manualmente.

As etapas 2 e 3 correspondem a implementação para o cálculo dos valores das observações. As equações de colinearidade (Eqs. (3.6) e (3.7)) e *shading* (Eq. (3.8)) são utilizadas nesta etapa.

Com o objetivo de reduzir o custo computacional, as etapas 4, 5, 6 e 7 foram implementadas a partir do conhecimento prévio dos padrões das matrizes A, N e U. Assim, a matriz dos coeficientes normais $N = A^t P A$ de ordem (uxu) e o vetor $U = A^t P L$ de ordem (ux1) são estabelecidos sem a necessidade de montar a matriz A usada no produto $A^t P A$ e $A^t P L$.

Considerando as equações de colinearidade, a matriz N resulta na estrutura mostrada na figura 3.10. A sub-matriz N_{11} é formada pelas derivadas dos elementos de orientação da câmara, sendo simétrica e de ordem 6. A sub-matriz N_{12} é a transposta de N_{21} , e é formada pelos produtos mistos das derivadas dos elementos de orientação e das coordenadas. A sub-matriz N_{22} é formada pelas derivadas das equações de colinearidade em relação as coordenadas

dos pontos (incógnitas). Ela contém sub-matrizes de ordem 3 ao longo das diagonais e zeros nas demais posições.

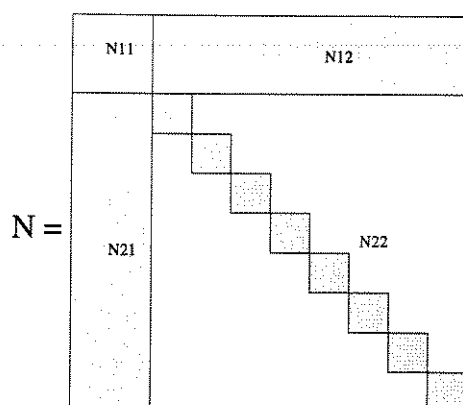


Figura 3.10: Estrutura da matriz das equações normais N somente com as equações de colinearidade.

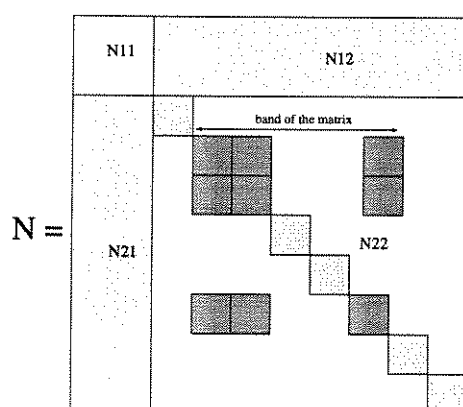


Figura 3.11: Estrutura da matriz das equações normais N em relação às equações de colinearidade e de *shading*

Como se verifica na figura 3.11, as equações de *shading* somente contribuirão na formação da sub-matriz N_{22} , onde ela provocará um alargamento dos elementos não nulos a partir da diagonal, produzindo uma banda (faixa com elementos significativos) de 3 vezes o número de colunas da imagem.

Conhecendo-se o padrão da matriz N e os padrões das matrizes $\bar{N}$ e N_b , que são diagonais, pode-se otimizar a solução do sistema (Etapa 8).

Desta forma uma das otimizações implementadas foi a exploração do padrão da matriz A , que não foi dimensionada no processo de montagem da matriz N , sendo necessário somente a alocação da memória de dois vetores, que corresponderão à duas linhas da matriz

A. Assim, foi utilizado o processo de montagem cumulativa da matriz N a partir dos dois vetores auxiliares da matriz A .

Ainda, foi explorada a simetria da matriz N gerando-se somente a porção triangular superior e a diagonal da matriz, armazenando-a num vetor onde a posição correspondente na matriz é dada pelo algoritmo 3.77.

3.8.1 Vetorização da matriz

Na implementação realizada, o sistema de equações (que é simétrica e positiva definida) foi solucionado usando-se o método de Cholesky, a condição de simetria da matriz N e a diagonalidade das matrizes $\overline{N}$ e N_b para vetorizar a porção triangular superior da matriz $(N + \overline{N} + N_b)$.

A vetorização foi realizada percorrendo-se a matriz simétrica da primeira linha até a última, sempre a partir da diagonal principal da linha correspondente. Cada elemento da porção triangular superior da matriz foi armazenado num vetor, conforme segue:

$$\begin{bmatrix} a_{(0)} \\ \dots \\ a_{(u)} \\ a_{(u+1)} \\ \dots \\ a_{([u+u*u]/2)} \end{bmatrix} = T \left(\begin{bmatrix} n_{(0,0)} & \dots & \dots & n_{(0,u)} \\ & n_{(1,1)} & \dots & n_{(1,u)} \\ & & \dots & \dots \\ & & & \dots & n_{(u,u)} \end{bmatrix} \right) \quad (3.76)$$

onde T é uma função de transformação, descrito pelo algoritmo 3.77,

$$k = (i + i * i) / 2$$

$$l = j * u + i - k \quad (3.77)$$

$$a[l] = n[i, j]$$

onde:

$a[l]$ é o l -ésimo elemento do vetor gerado; e

$n[i, j]$ é o elemento da linha i e da coluna j da matriz $(N + \overline{N})$. Observa-se que o valor de

j sempre será maior ou igual a i , condição esta que impõe que somente a porção triangular superior da matriz será mapeada para o vetor.

Capítulo 4

Experimentos e Resultados

Nesta seção são apresentados e discutidos os experimentos utilizados para avaliar o modelo proposto. Os resultados foram analisados estatisticamente através do desvio padrão das coordenadas, determinadas com base no erro entre as coordenadas obtidas e as coordenadas reais.

4.1 Imagens Simuladas

Foram geradas imagens de superfícies planas, esféricas e cilíndricas, realizando as seguintes variações:

- aproximando e afastando a fonte de luz;
- alterando a geometria dos raios da luz (pontual ou paralela);
- inclinando-se o feixe de raios da luz; e
- mudando a posição da câmara (variação da escala).

Devido à especificidade de cada objeto utilizado na composição da imagem, foram elaborados três (3) programas para gerar imagens de profundidade e de intensidade, considerando a projeção perspectiva. Assim, para cada objeto (plano, esfera e cilindro), foi gerada uma imagem de profundidade, dada pela distância do ponto ao plano de referência no espaço objeto, e uma de intensidade.

4.1.1 Experimentos com Imagens de uma superfície Plana

Simulação da imagem

A superfície plana foi o modelo utilizado no trabalho apresentado por Hasegawa e Tozzi [17], pois ela tem a geometria conhecida e de fácil manipulação. A superfície plana foi gerada pixel a pixel, conhecendo-se a normal correspondente ao pixel no espaço objeto, conforme a figura 4.1, a posição da fonte de luz e da câmara.

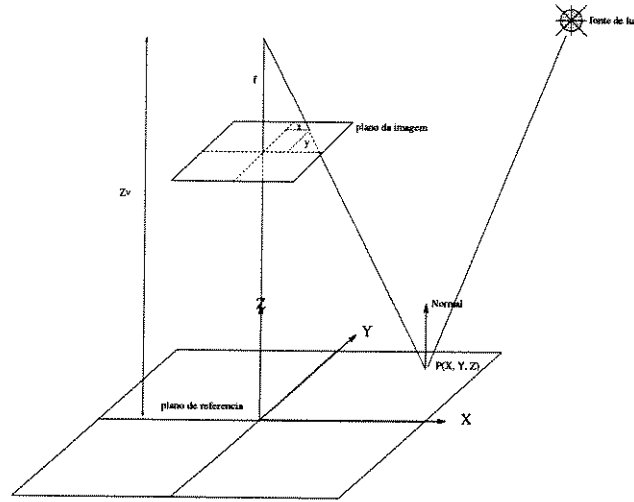


Figura 4.1: Imagem de um plano gerada com projeção perspectiva

Considerou-se que a câmara estava orientada em relação ao sistema de coordenadas utilizado. O sistema de coordenadas do objeto tem como origem o centro geométrico da figura, cuja imagem é centralizada. O sistema de coordenadas tem a orientação paralela ao plano de referência e a posição da fonte de luz \$(X_l, Y_l, Z_l)\$ pré-definida (ver figura 4.1). Estas considerações resultaram nas seguintes equações (Eqs. 4.1) para a imagem de intensidade:

$$\begin{aligned} X_{(i,j)} &= i \frac{Z_v}{f}; \\ Y_{(i,j)} &= j \frac{Z_v}{f}; \\ Z_{(i,j)} &= 0; \end{aligned} \quad (4.1)$$

Como \$Z_{(i,j)}\$ é constante, pode-se utilizá-lo para montar a imagem de profundidade e o vetor normal terá sempre o mesmo valor:

$$V_n = (0, 0, 1);$$

assim, a intensidade de um ponto $P_{(i,j)}$ será:

$$I(i,j) = \frac{(Z_l - Z_{(i,j)})}{\sqrt{(X_l - X_{(i,j)})^2 + (Y_l - Y_{(i,j)})^2 + (Z_l - Z_{(i,j)})^2}}.$$

Experimentos

Com a finalidade de verificar o desempenho do método proposto, foram sintetizadas imagens de intensidade e profundidade. Para cada cena foram geradas imagens de intensidade e suas respectivas imagens de profundidade, conforme Hasegawa e Tozzi [17].

As seguintes imagens foram geradas para avaliar o modelo proposto:

- Imagem 1: Superfície plana Lambertiana com coeficiente de reflexão difusa ($k_d = 0,35$), iluminada por uma fonte de luz pontual, numa posição coincidente com o centro perspectivo da câmara;
- Imagem 2: Superfície plana considerando o coeficiente de reflexão especular ($k_s = 0,30$) e com coeficiente de reflexão difusa ($k_d = 0,35$), iluminada por uma fonte de luz pontual numa posição coincidente com o centro perspectivo da câmara;
- Imagem 3: Superfície plana Lambertiana com coeficiente de reflexão difusa ($k_d = 0,35$), iluminada por uma fonte de luz pontual no canto direito da imagem.

Com estas imagens foram realizadas as seguintes experiências:

1. Calibração da câmara e reconstrução da superfície usando a Imagem 1, usando 5 pontos de injunções (4 nos cantos e um no centro). Neste experimento, o modelo de iluminação da imagem gerada e do método proposto são equivalentes;
2. Calibração da câmara e reconstrução da superfície usando a Imagem 2, com a mesma configuração descrita na experiência 1. Neste experimento, o modelo de iluminação da imagem gerada e do método proposto são diferentes. O objetivo deste experimento foi a determinação dos erros provocados ao desconsiderar o coeficiente de reflexão difusa da luz no modelo proposto;
3. Reconstrução da superfície usando a Imagem 1. O posicionamento da câmara foi considerado conhecido e os seus parâmetros considerados como injunções. Determinação do efeito da injunção nos parâmetros de calibração da câmara e da reconstrução da superfície simultaneamente;

4. Calibração da câmara e reconstrução da superfície como descrito na experiência 1, mas usando um modelo de reconstrução na qual a cena é iluminada por raios paralelos. O objetivo deste experimento foi a determinação do erro provocado quando se considera, no modelo, que a fonte de luz emita raios paralelos, equivalente a fonte de luz no infinito, ao invés de raios divergentes, com direção dependente da posição da fonte de luz e do ponto iluminado na superfície;
5. Calibração da câmara e reconstrução da superfície usando a Imagem 3, considerando as mesmas injunções do experimento 1. O objetivo deste experimento foi a determinação da influência da posição da fonte de luz;
6. Calibração da câmara e reconstrução da superfície usando a Imagem 1, com 9 pontos de injunções. O objetivo deste experimento foi a verificação do efeito causado pela quantidade dos pontos de injunções na reconstrução da superfície e na calibração da câmara.

Resultados

A partir dos resultados obtidos, foram determinadas as diferenças, ponto a ponto, entre as coordenadas dadas pela imagem de profundidade sintetizada e a obtida pelo método proposto, a fim de verificar a precisão do método. A tabela 4.1 resume, para cada experimento, os valores médios e os desvios-padrão dos erros para as três coordenadas.

Tabela 4.1: Média e Desvio Padrão dos erros para as coordenadas nos testes de 1 a 6, obtida de Hasegawa e Tozzi [16]

teste	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
1	0.0001	-0.0205	-0.7208	0.1929	0.1973	1.3966
2	-0.0002	-0.0210	-0.7147	0.1906	0.1938	1.3843
3	-0.0123	-0.0351	-0.5253	1.2215	1.2249	3.7774
4	0.2976	0.2963	-0.1051	0.2547	0.2552	1.3205
5	0.0738	0.0647	-0.1354	0.0471	0.0431	0.2516
6	0.0002	-0.0226	-0.7358	0.1846	0.1874	1.3277

Observando a tabela 4.1, verifica-se que as coordenadas referentes às profundidades (coordenadas Z), consideradas neste trabalho como as altimétricas, do objeto são as que produzem os piores resultados, compatíveis com o procedimento de reconstrução utilizado.

Tabela 4.2: Tabela com os valores dos parâmetros de calibração da câmara

teste	ω (rad)	κ (rad)	ϕ (rad)	X_c (mm)	Y_c (mm)	Z_c (mm)
1	0.00000	0.00000	0.00000	0.00008	-0.00026	1000.00512
6	0.00000	0.00000	0.00000	0.00009	-0.00054	999.99487

Nas reconstruções de objetos a partir de imagens (uma ou mais), as coordenadas altimétricas são determinadas por interseção de raios (duas ou mais fotos) ou por alguma condição imposta no caminho do raio (um equação de *shading*), provocando uma degradação mais acentuada nas coordenadas altimétricas.

O teste 05 é o que produz melhor resultado, analisada a coordenada Z, condição esta que pode ter sido provocada pela não coincidência do observador com a fonte de luz, resultando numa configuração favorável para este tipo de superfície.

Da tabela 4.2, pode-se verificar que os parâmetros de calibração da câmara são determinados com precisão, independentemente da quantidade de pontos utilizados, desde que se utilize de, no mínimo, cinco (5) pontos de injeção (controle). Sabe-se que três (3) pontos de injeção é a quantidade mínima para se determinar os parâmetros de calibração extrínsecos de uma câmara, mas na prática, e para aplicar o MMQ, utiliza-se quatro (4) ou mais pontos.

Os pontos de injeção (controle) devem estar distribuídos uniformemente na imagem; no caso da utilização dos cinco pontos, quatro devem estar situados nos cantos e o quinto na posição central.

4.1.2 Experimentos com imagens simuladas de uma Esfera

Simulação da imagem

A fim de gerar as imagens de profundidade e de intensidade, foram realizadas as seguintes considerações:

- adotou-se um plano de referência paralelo ao plano da imagem, transladado da distância de projeção (Z_v);
- a esfera foi dividida em dois hemisférios pelo plano de referência, provocando a coincidência do plano com o centro da esfera;

- a origem do sistema de coordenadas global (espaço objeto) foi definida no centro da esfera; e
- o sistema de coordenadas global foi definido como paralela ao plano de referência e orientado na mesma direção do sistema de coordenadas da imagem.

Definido e orientado o sistema de coordenadas, a fonte de luz (considerada conhecida) é posicionada pelas coordenadas $(X_l, Y_l \text{ e } Z_l)$, em cada imagem gerada, conforme a Figura 4.2.

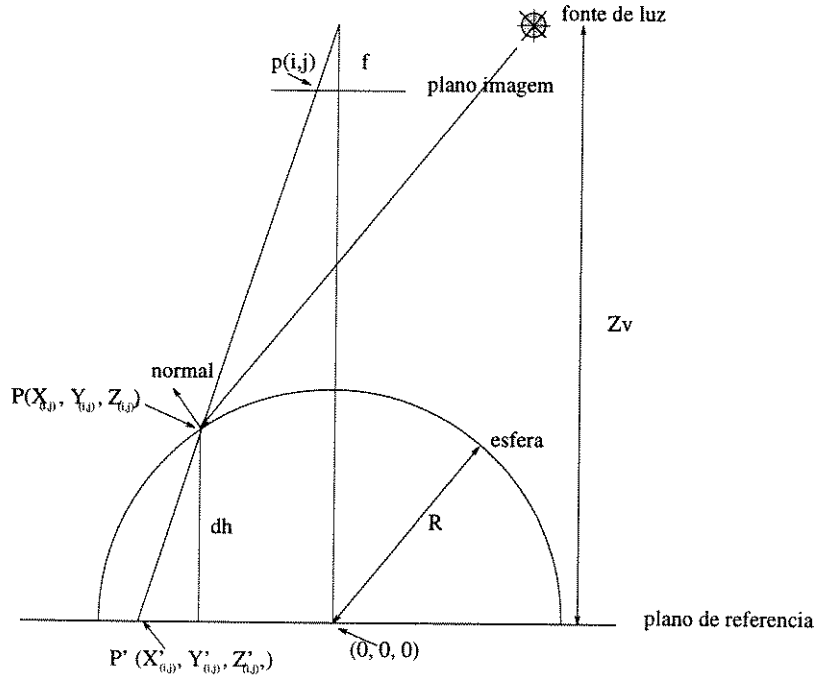


Figura 4.2: Imagem de uma esfera gerada com projeção perspectiva

Pela Figura 4.2, pode-se verificar que, somente a partir das coordenadas da imagem de um ponto $p(x,y)$, não é possível calcular as coordenadas tridimensionais do ponto P no espaço objeto. Entretanto, pode-se calcular as coordenadas do ponto P' através das equações 4.1.

Desta forma, por semelhança de triângulos (ver figura 4.2), pode-se obter as seguintes equações:

$$\frac{Z_v - dh}{A} = \frac{Z_v}{D} \quad (4.2)$$

$$A^2 = R^2 - dh^2 \quad (4.3)$$

$$D = \sqrt{X^2 + Y^2} \quad (4.4)$$

Substituindo as variáveis A e D das equações 4.3 e 4.4 na equação 4.2, obtém-se a equação 4.5 do 2º grau, com a variável dh como incógnita:

$$(D^2 + Z_v)dh^2 - (Z_v D^2 dh + Z_v^2 D^2 - R^2 Z_v^2) = 0 \quad (4.5)$$

Assim, resolvendo a equação acima, o valor de $Z_{(i,j)}$ (utilizado para montar a imagem range) do ponto P será o valor de dh positivo, pois o outro valor será o que interceptar a esfera na porção inferior desta, que neste caso não será considerado.

O cálculo das coordenadas $X_{(i,j)}$ e $Y_{(i,j)}$ é efetuado a partir das equações 4.6:

$$\begin{aligned} X_{(i,j)} &= \frac{(Z_v - Z_{(i,j)})}{f} \cdot i \\ Y_{(i,j)} &= \frac{(Z_v - Z_{(i,j)})}{f} \cdot j \end{aligned} \quad (4.6)$$

Na geração da imagem de intensidade faz-se necessário o uso do vetor normal:

$$V_n = (n_x, n_y, n_z).$$

Assim, a equação 4.7 é uma representação algébrica da equação da esfera no espaço tridimensional:

$$Z = \sqrt{R^2 - X^2 - Y^2} \quad (4.7)$$

cujos componentes do vetor normal são calculados a partir das derivadas parciais nas direções X e Y (gradientes) da equação 4.7, resultando nos seguintes valores,

$$n_x = p = \frac{\partial Z_{(i,j)}}{\partial X_{(i,j)}} = \frac{X_{(i,j)}}{Z_{(i,j)}^2}$$

$$n_y = q = \frac{\partial Z_{(i,j)}}{\partial Y_{(i,j)}} = \frac{Y_{(i,j)}}{Z_{(i,j)}^2} \quad (4.8)$$

$$n_z = 1$$

Desta forma, a intensidade de um ponto $P_{(i,j)}$ na imagem será dada pela equação 4.9:

$$I(i,j) = \frac{V_n \cdot V_l}{|V_n| |V_l|}; \quad (4.9)$$

com $V_l = [(X_l - X_{(i,j)}) (Y_l - Y_{(i,j)}) (Z_l - Z_{(i,j)})]$.

Experimentos

A partir do modelo proposto, foram geradas várias imagens para avaliar, analisar e/ou verificar alguns parâmetros específicos e o comportamento do sistema quanto à configuração utilizada. Assim, foram geradas imagens simulando uma câmara com distância focal de 60 mm, 15x15 pixels, distâncias de projeção médias de 1800 mm e 5400 mm e a fonte de luz posicionada em várias posições, conforme as tabelas 4.3 e 4.4.

As tabelas 4.3 e 4.4 apresentam um conjunto de imagens geradas de superfícies esféricas, variando-se (aproximando ou afastando da superfície) a posição da fonte de luz. Nestas tabelas (4.3 e 4.4) :

K_d representa o coeficiente de reflexão difusa da superfície;

K_s representa o coeficiente de reflexão especular da superfície;

I indica a inclinação da direção da fonte de luz em relação ao sistema de coordenadas adotado;

Z_v indica a distância do centro perspectivo ao plano de referência adotado;

Z_l indica a distância da posição da fonte de luz ao plano de referência adotado;

As imagens geradas nas tabelas 4.3 e 4.4 têm como objetivo verificar a influência da posição da fonte de luz em relação à superfície.

A figura 4.3 é uma representação *wire frame* da superfície range (no caso com distância de projeção de 1800 mm) utilizada nos experimentos da tabela 4.3.

Tabela 4.3: Parâmetros utilizados nas imagens geradas variando a posição da luz, com distância de projeção de 1800 mm

Imagem	K_d	K_s	I	$Z_v(mm)$	Z_l
1	0.5	0.0	0 ⁰	1800.0	360.0
2	0.5	0.0	0 ⁰	1800.0	450.0
3	0.5	0.0	0 ⁰	1800.0	900.0
4	0.5	0.0	0 ⁰	1800.0	1350.0
5	0.5	0.0	0 ⁰	1800.0	1800.0
6	0.5	0.0	0 ⁰	1800.0	2700.0
7	0.5	0.0	0 ⁰	1800.0	3600.0

Tabela 4.4: Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 5400 mm

Imagem	K_d	K_s	I	$Z_v(mm)$	$Z_l(mm)$
8	0.5	0.0	0 ⁰	5400.0	1080.0
9	0.5	0.0	0 ⁰	5400.0	1350.0
10	0.5	0.0	0 ⁰	5400.0	2700.0
11	0.5	0.0	0 ⁰	5400.0	4050.0
12	0.5	0.0	0 ⁰	5400.0	5400.0
13	0.5	0.0	0 ⁰	5400.0	8100.0
14	0.5	0.0	0 ⁰	5400.0	10800.0

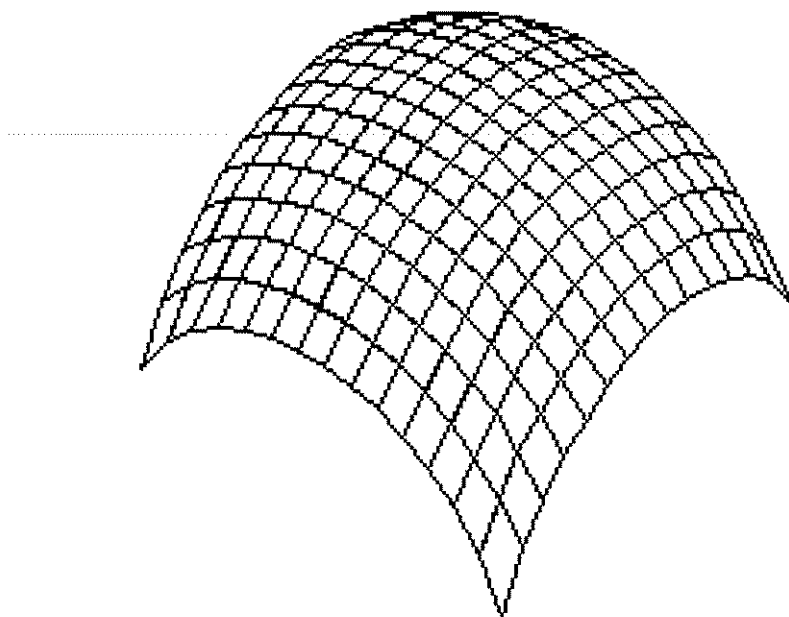


Figura 4.3: Representação *wire frame* construída a partir da imagem de profundidade, gerada da superfície esférica

Experimentos Explorando Bordas

Para avaliar o modelo proposto, foram geradas imagens com arestas provocadas pelas bordas da superfície esférica. Desta forma, foram geradas imagens a partir da superfície esférica com raio correspondente a 80% e 90% (Imagens 4.4 e 4.5) do tamanho do quadro da imagem, criando aresta fechada na imagem.

As tabelas 4.5 e 4.6 apresentam um conjunto de imagens geradas de superfícies esféricas, variando-se (inclinando-se a direção da fonte de luz em relação ao eixo Z) a posição da fonte de luz e o tamanho do objeto contido na imagem. Nestas tabelas (4.5 e 4.6) :

- **tam** representa o tamanho da esfera na imagem;
- **A** representa que o raio da esfera é de 80% do tamanho do quadro;
- **B** representa que o raio da esfera é de 90% do tamanho do quadro; e
- **IB** representa se o processamento foi realizado com ou sem informações de bordas.

As imagens geradas com os parâmetros apresentados nas tabelas 4.5 e 4.6 têm como objetivo verificar a influência das bordas e a posição da fonte de luz no processo de reconstrução proposto.

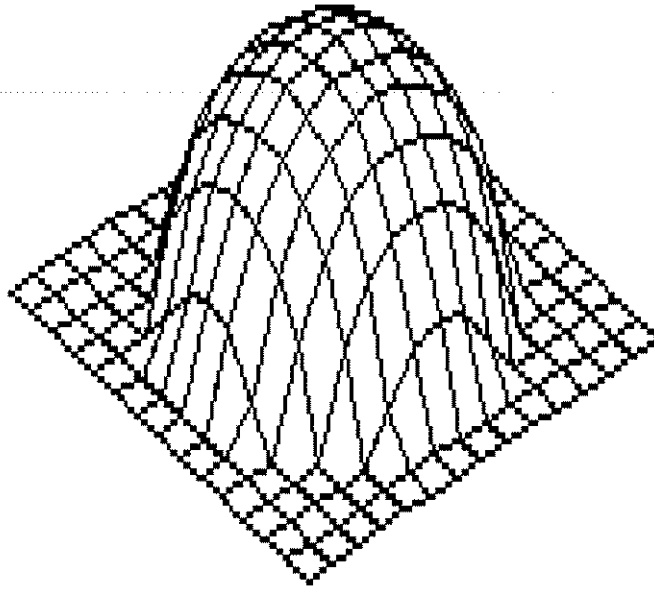


Figura 4.4: Representação *wire frame* construída a partir da imagem de profundidade gerada, da superfície esférica, com o raio de 80% do tamanho do quadro da imagem

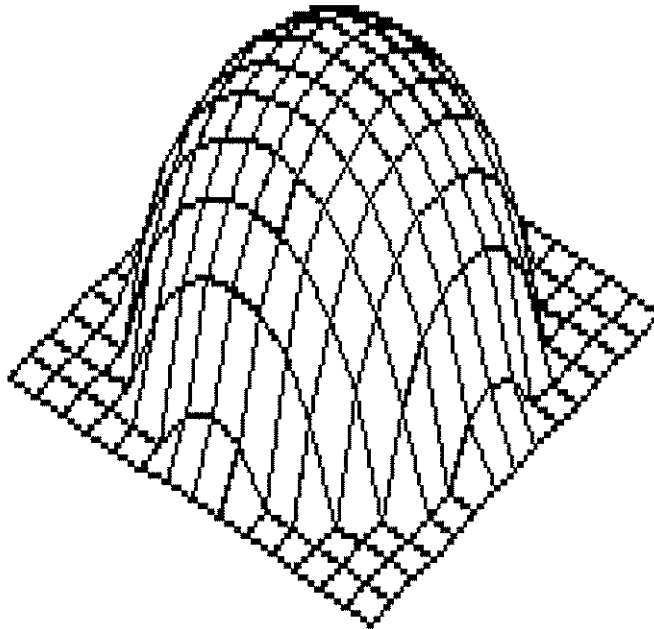


Figura 4.5: Representação *wire frame* construída a partir da imagem de profundidade gerada, da superfície esférica, com o raio de 90% do tamanho do quadro da imagem

Tabela 4.5: Parâmetros utilizados nas imagens geradas de esferas com variação da direção da luz (Z_v e $Z_l = 1800\text{mm}$)

Imagem	K_d	K_s	I	tam	IB
15	0.5	0.0	0^0	A	não
16	0.5	0.0	0^0	A	sim
17	0.5	0.0	1^0	A	sim
18	0.5	0.0	3^0	A	sim
19	0.5	0.0	5^0	A	sim
20	0.5	0.0	7^0	A	sim
21	0.5	0.0	9^0	A	sim
22	0.5	0.0	0^0	B	não
23	0.5	0.0	0^0	B	sim
24	0.5	0.0	1^0	B	sim
25	0.5	0.0	3^0	B	sim
26	0.5	0.0	5^0	B	sim
27	0.5	0.0	7^0	B	sim
28	0.5	0.0	9^0	B	sim

Tabela 4.6: Parâmetros utilizados nas imagens geradas de esferas com variação da direção da luz (Z_v e $Z_l = 5400\text{ mm}$)

Imagem	K_d	K_s	I	tam	IB
29	0.5	0.0	0^0	A	não
30	0.5	0.0	0^0	A	sim
31	0.5	0.0	1^0	A	sim
32	0.5	0.0	3^0	A	sim
33	0.5	0.0	5^0	A	sim
34	0.5	0.0	7^0	A	sim
35	0.5	0.0	9^0	A	sim
36	0.5	0.0	0^0	B	não
37	0.5	0.0	0^0	B	sim
38	0.5	0.0	1^0	B	sim
39	0.5	0.0	3^0	B	sim
40	0.5	0.0	5^0	B	sim
41	0.5	0.0	7^0	B	sim
42	0.5	0.0	9^0	B	sim

Resultados

Para avaliação da precisão da reconstrução foram determinadas as diferenças, ponto a ponto, entre as coordenadas definidas na imagem de profundidade sintetizada e a obtida pelo método de reconstrução proposto. As Tabelas 4.7 e 4.8 apresentam, para cada experimento, os desvios-padrão dos erros determinados para as três coordenadas.

Tabela 4.7: Médias e Desvios-Padrão dos erros das coordenadas nos testes (1 a 7) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
1	0.0000	0.0000	6.7045	0.5465	0.5465	1.4842
2	0.0000	0.0000	6.9685	0.5393	0.5393	1.3978
3	0.0000	0.0000	6.6046	0.5323	0.5323	1.5961
4	0.0000	0.0000	6.1871	0.5207	0.5207	1.6798
5	0.0000	0.0000	6.0229	0.5174	0.5174	1.7773
6	0.0000	0.0000	5.8360	0.5127	0.5127	1.8820
7	0.0000	0.0000	5.7677	0.5102	0.5101	1.9130

Elaborando-se um gráfico onde a ordenada é representada pelo desvio-padrão, e tendo como abscissa a relação das distâncias da fonte de luz Z_l com a distância de projeção Z_v , conforme a figura 4.6, pode-se traçar uma curva de interpolação.

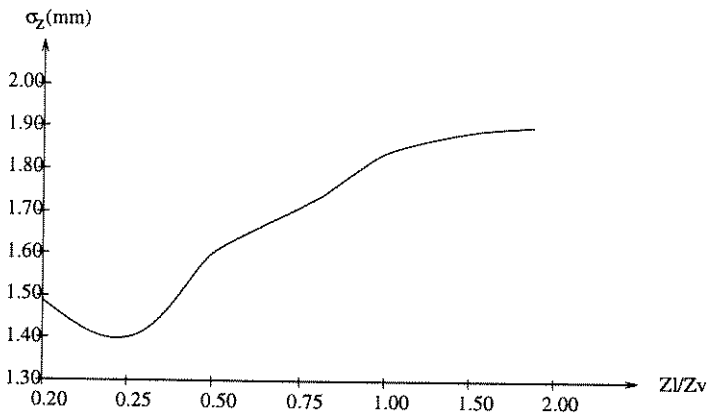


Figura 4.6: Curva gerada a partir dos desvios-padrão dos erros para imagens esféricas com distância de 1800 mm de projeção.

Da tabela 4.7, pode-se observar que os desvios-padrão dos erros das coordenadas Z (que se relacionam com a profundidade) são maiores que os das coordenadas X e Y. Pode-se ainda, verificar que as médias de X e Y são zero, o que não ocorre com as coordenadas Z,

indicando a presença de erros sistemáticos.

Observando-se a figura 4.6, verifica-se que há um ponto de mínimo, próximo da ordenada 0,25. Ainda, há uma tendência de estabilização da curva quando a distância da fonte de luz é aproximadamente o dobro da distância de projeção, ou seja, quando a fonte de luz tende para o infinito.

Tabela 4.8: Médias e Desvios-Padrão dos erros das coordenadas nos testes (8 a 14) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
8	0.0005	-0.0006	17.7682	1.5003	1.4897	6.6401
9	0.0005	-0.0007	18.5603	1.5231	1.4866	5.7800
10	0.0004	-0.0005	18.3712	1.4915	1.4915	4.8802
11	0.0003	-0.0003	17.7356	1.4786	1.4786	5.3152
12	0.0003	-0.0003	17.4278	1.4038	1.4037	5.5063
13	0.0002	-0.0002	17.1287	1.4677	1.4677	5.6992
14	0.0001	-0.0002	16.9942	1.4661	1.4661	5.7826

Da mesma forma, verificando-se a tabela 4.8, pode-se elaborar um gráfico considerando as mesmas variáveis utilizadas na montagem da tabela 4.6.

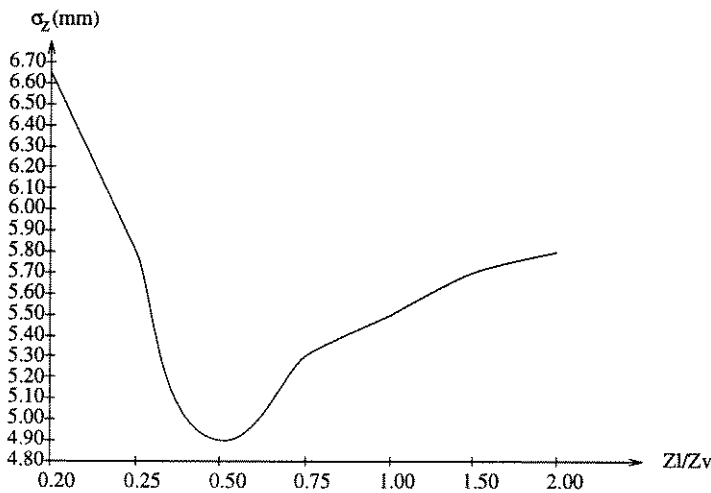


Figura 4.7: Curva gerada a partir dos desvios-padrão dos erros para imagens esféricas com distância de 5400 mm de projeção.

Analisando-se a tabela 4.8, como no caso da tabela 4.7, os desvios-padrão dos erros das coordenadas tem comportamentos semelhantes, isto é, as coordenadas Z têm erros maiores que as coordenadas X e Y. Da mesma forma, verifica-se uma possível tendência a erros sistemáticos nas coordenadas Z.

Observando-se a figura 4.7, da mesma forma que no caso da figura 4.6, há um ponto de mínimo próximo da ordenada 0,50 e uma tendência de estabilização no valor próximo de 2,0.

Tabela 4.9: Médias e Desvios-Padrão dos erros das coordenadas nos testes (15 a 28) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
15	0.0000	0.0000	-0.9575	0.2100	0.2100	3.5250
16	0.0021	0.0021	-0.6273	0.1383	0.1383	2.2917
17	-0.0015	-0.0015	-0.6277	0.1384	0.1384	2.2947
18	-0.0026	-0.0026	-0.6290	0.1401	0.1401	2.3227
19	-0.0066	-0.0066	-0.6260	0.1413	0.1413	2.3468
20	-0.0105	-0.0105	-0.6228	0.1431	0.1431	2.3809
21	-0.0144	-0.0144	-0.6162	0.1450	0.1450	2.4179
22	0.0000	-0.0000	-0.3153	0.2025	0.2004	3.0557
23	-0.0022	-0.0022	0.0463	0.1286	0.1286	1.9824
24	-0.0029	-0.0029	0.0461	0.1289	0.1289	1.9874
25	-0.0132	-0.0132	0.0574	0.1320	0.1320	2.0342
26	-0.0396	-0.0396	0.1320	0.1335	0.1335	2.0380
27	-0.0469	-0.0469	0.1858	0.1401	0.1401	2.1339
28	-0.0539	-0.0539	0.2415	0.1479	0.1479	2.2527

A partir da análise dos resultados apresentados nas Tabelas 4.9 e 4.10, observa-se uma redução no erro médio da reconstrução quando as informações de borda são utilizadas, condição esta que era esperada, pois as silhuetas provocadas pelas bordas de sombreamento representam uma região de mudança brusca de orientação.

Das tabelas 4.9 e 4.10, observa-se que, a tendência a erros sistemáticos (em Z), pode ser verificada comparando-se as médias das discrepâncias nas coordenadas X, Y e Z. Isto pode ser verificado no teste realizado sem a utilização de informações de bordas, comparado com os testes utilizando-as. Entretanto, quando se usa informações de bordas no processamento, isto provoca uma tendência a erros sistemáticos nas coordenadas X e Y, o que não ocorre no outro processamento.

Tabela 4.10: Médias e Desvios-Padrão dos erros das coordenadas nos testes (29 a 42) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
29	0.0000	-0.0000	-2.1476	0.3877	0.3877	6.2001
30	-0.0056	-0.0056	-1.8850	0.3322	0.3322	5.3933
31	0.0061	0.0061	-1.8851	0.3323	0.3322	5.3950
32	-0.0110	-0.0109	-1.8853	0.3317	0.3317	5.3952
33	0.0082	0.0082	-1.8832	0.3330	0.3331	5.4000
34	-0.0040	-0.0040	-1.8844	0.3325	0.3325	5.4028
35	-0.0243	-0.0243	-1.8903	0.3373	0.3373	5.4140
36	-0.0000	0.0000	0.7152	0.2872	0.2872	4.5658
37	-0.0001	-0.0001	1.3186	0.2713	0.2713	4.3323
38	-0.0006	-0.0006	1.3156	0.2715	0.2715	4.3237
39	-0.0018	-0.0019	1.3088	0.2721	0.2721	4.3432
40	-0.0033	-0.0033	1.3117	0.2726	0.2726	4.3496
41	-0.0061	-0.0060	1.3034	0.2722	0.2722	4.3497
42	-0.0069	-0.0069	1.2984	0.2723	0.2723	4.3502

4.1.3 Experimentos com Imagens Simuladas de um Cilindro

Simulação da Imagem

A imagem do cilindro foi gerada fazendo-se as mesmas considerações da esfera, sendo que, o eixo principal do cilindro, está posicionado na direção Y do sistema de coordenadas adotado, conforme a Figura 4.8.

O processamento das coordenadas pode ser realizado utilizando-se o procedimento iterativo, determinando-se a equação algébrica do cilindro 4.10.

$$Z = \sqrt{R^2 - X^2} \quad (4.10)$$

A equação 4.10 descreve um cilindro centrado no eixo Y do sistema de coordenadas cartesiano tridimensional, que tem como base o círculo definido no plano Z-X.

Da figura 4.8 pode-se obter as seguintes relações matemáticas, para se realizar o processo iterativo:

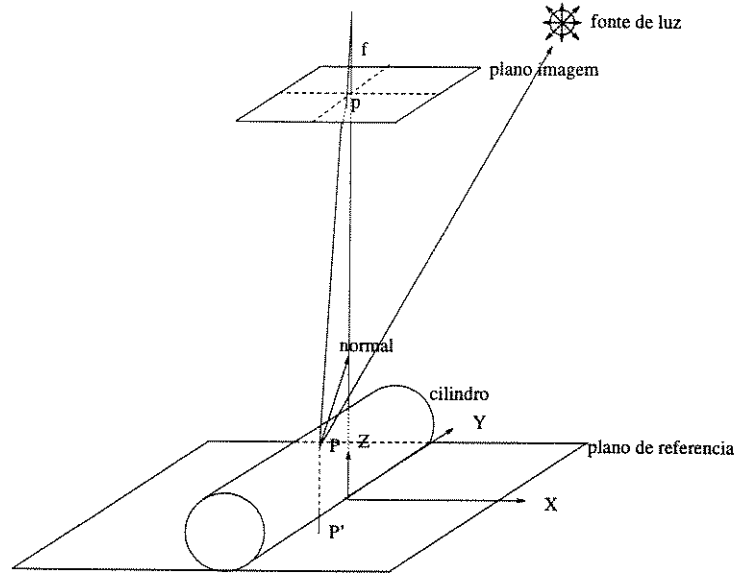


Figura 4.8: Imagem de um cilindro gerada com projeção perspectiva

1. atribui-se um valor inicial para a coordenada Z^i (zero);
2. calcula-se as coordenadas X^i e Y^i , usando as equações 4.6, com o valor de Z^i ;
3. determina-se o valor de $Z^{i+1} = \sqrt{R^2 - X^i X^i}$;
4. calcula-se a diferença dos valores de Z ($|Z^i - Z^{i+1}|$).
5. se $|Z^i - Z^{i+1}| > \xi$, atualiza-se a coordenada Z e retorna para a etapa 2;
6. se $|Z^i - Z^{i+1}| \leq \xi$, finaliza-se o processo.

Caso a convergência seja atingida, aplica-se a equação 4.6, determinando-se os valores atualizados das coordenadas dos pontos.

Na determinação da intensidade do ponto também se utiliza a normal à superfície no ponto, cujos componentes são calculados a partir dos gradientes da função (equação do cilindro), resultando nos seguintes valores:

$$n_x = p = \frac{\partial Z_{(i,j)}}{\partial X_{(i,j)}} = \frac{X_{(i,j)}}{Z_{(i,j)}^2}$$

$$n_y = q = \frac{\partial Z_{(i,j)}}{\partial Y_{(i,j)}} = 0 \quad (4.11)$$

$$n_z = 1$$

A intensidade será calculada usando-se a equação do caso da esfera (Eq. 4.9).

Experimentos

Foram geradas imagens a partir dos parâmetros descritos nas tabelas 4.11 e 4.12, no mesmo padrão e configuração usado no caso das imagens esféricas do item 4.1.2, considerando, agora, o raio do cilindro como padrão.

Tabela 4.11: Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 1800 mm

Imagem	K_d	K_s	I	Z_v (mm)	Z_l (mm)
1	0.5	0.0	0 ⁰	1800.0	360.0
2	0.5	0.0	0 ⁰	1800.0	450.0
3	0.5	0.0	0 ⁰	1800.0	900.0
4	0.5	0.0	0 ⁰	1800.0	1350.0
5	0.5	0.0	0 ⁰	1800.0	1800.0
6	0.5	0.0	0 ⁰	1800.0	2700.0
7	0.5	0.0	0 ⁰	1800.0	3600.0

Tabela 4.12: Parâmetros utilizados nas imagens geradas variando a posição da luz com distância de projeção de 5400 mm

Imagem	K_d	K_s	I	Z_v (mm)	Z_l (mm)
8	0.5	0.0	0 ⁰	5400.0	810.0
9	0.5	0.0	0 ⁰	5400.0	1080.0
10	0.5	0.0	0 ⁰	5400.0	1350.0
11	0.5	0.0	0 ⁰	5400.0	2700.0
12	0.5	0.0	0 ⁰	5400.0	4050.0
13	0.5	0.0	0 ⁰	5400.0	5400.0
14	0.5	0.0	0 ⁰	5400.0	8100.0
15	0.5	0.0	0 ⁰	5400.0	10800.0

Da mesma forma que no caso da superfícies esféricas, estes experimentos, descritos nas tabelas 4.11 e 4.12, têm como objetivo verificar a influência da posição da fonte de luz em relação à superfície.

Experimentos Explorando Bordas

De forma semelhante ao caso da superfície esférica, foram geradas imagens com arestas provocadas pelas bordas da superfície cilíndrica. Assim, foram geradas imagens cilíndricas com raio correspondentes a 80% e 90% (imagens 4.9 e 4.10) do quadro da imagem, provocando arestas em diferentes posições da imagem.

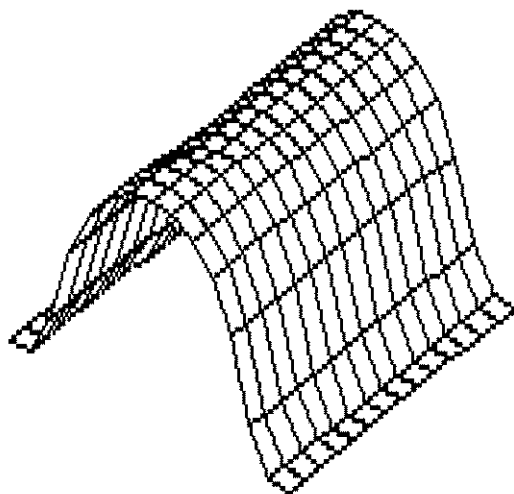


Figura 4.9: Representação *wire frame* construída a partir da imagem de profundidade, gerada da superfície cilíndrica, com o raio com 80% do quadro da imagem

As tabelas 4.13 e 4.14 apresentam os parâmetros utilizados para a geração de um conjunto de imagens de superfícies cilíndricas, variando-se (inclinando-se a direção da fonte de luz em relação ao eixo Z) a posição da fonte de luz e o tamanho desta superfície contida na imagem.

Resultados

A partir dos resultados obtidos, foram determinadas as diferenças das coordenadas, ponto a ponto, como realizado no item 4.1.1.

Da mesma forma que no caso das superfícies esféricas, verificando-se a tabela 4.15, pode-se elaborar um gráfico considerando as mesmas variáveis utilizadas na montagem das tabelas 4.6 e 4.7.

Da tabela 4.15, verifica-se que, os desvios-padrão dos erros das coordenadas, têm o mesmo comportamento que no caso das superfícies esféricas.

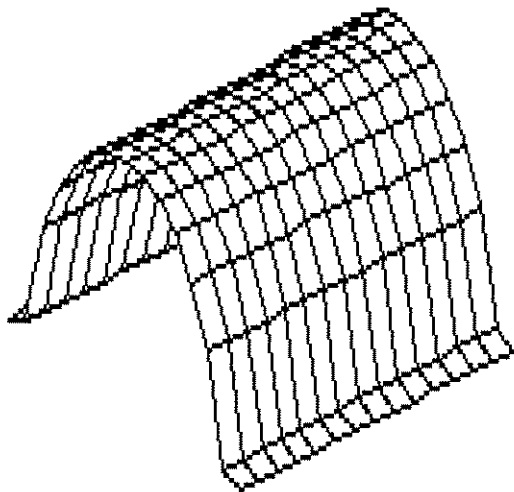


Figura 4.10: Representação *wire frame* construída a partir da imagem de profundidade gerada, da superfície cilíndrica, com o raio com 90% do quadro da imagem

Tabela 4.13: Parâmetros utilizados na geração de imagens variando-se a direção da fonte de luz (Z_v e $Z_l = 1800$ mm)

Imagem	K_d	K_s	I	tam	IB
16	0.5	0.0	0^0	A	não
17	0.5	0.0	0^0	A	sim
18	0.5	0.0	1^0	A	sim
19	0.5	0.0	3^0	A	sim
20	0.5	0.0	5^0	A	sim
21	0.5	0.0	7^0	A	sim
22	0.5	0.0	9^0	A	sim
23	0.5	0.0	0^0	B	não
24	0.5	0.0	0^0	B	sim
25	0.5	0.0	1^0	B	sim
26	0.5	0.0	3^0	B	sim
27	0.5	0.0	5^0	B	sim
28	0.5	0.0	7^0	B	sim
29	0.5	0.0	9^0	B	sim

Tabela 4.14: Parâmetros utilizados na geração de imagens com variação da direção da luz (Z_v e $Z_l = 5400$ mm)

Imagem	K_d	K_s	I	tam	IB
30	0.5	0.0	0^0	A	não
31	0.5	0.0	0^0	A	sim
32	0.5	0.0	1^0	A	sim
33	0.5	0.0	3^0	A	sim
34	0.5	0.0	5^0	A	sim
35	0.5	0.0	7^0	A	sim
36	0.5	0.0	9^0	A	sim
37	0.5	0.0	0^0	B	não
38	0.5	0.0	0^0	B	sim
39	0.5	0.0	1^0	B	sim
40	0.5	0.0	3^0	B	sim
41	0.5	0.0	5^0	B	sim
42	0.5	0.0	7^0	B	sim
43	0.5	0.0	9^0	B	sim

Tabela 4.15: Médias e Desvios-Padrão dos erros das coordenadas nos testes (1 a 7) realizados

Média (mm)				Desvio Padrão (mm)		
Imagem	X	Y	Z	X	Y	Z
1	-0.0003	0.0002	7.2453	0.7942	0.7846	2.3361
2	-0.0000	0.0000	5.0411	0.7835	0.6950	2.0941
3	-0.0000	0.0000	5.6277	0.7635	0.7227	1.5385
4	-0.0000	0.0000	5.1786	0.7554	0.7015	1.6947
5	-0.0000	0.0000	4.9677	0.7514	0.6918	1.7977
6	0.0000	0.0000	4.7686	0.7476	0.6830	1.9076
7	0.0000	0.0000	4.6756	0.7454	0.6787	1.9600

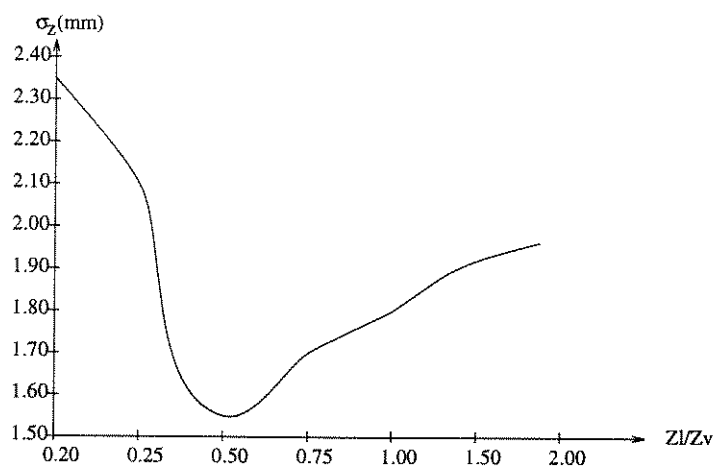


Figura 4.11: Curva gerada a partir dos desvios-padrão dos erros para imagens cilíndricas com distância de 1800 mm de projeção.

Observando-se a figura 4.11, da mesma forma que no caso das superfícies esféricas, há um ponto de mínimo no eixo X, cujo valor é da ordem de 0,50 e uma tendência de estabilização a partir de 2,0.

Tabela 4.16: Médias e Desvios-Padrão dos erros das coordenadas nos testes (8 a 15) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
8	0.0001	0.0003	8.4304	0.7694	0.6627	3.2478
9	0.0002	0.0002	9.0491	0.7715	0.6915	2.6242
10	0.0001	0.0000	8.0981	0.7597	0.6605	3.2983
11	0.0000	-0.0000	7.6580	0.7522	0.6487	3.7084
12	0.0001	-0.0000	7.5755	0.7426	0.6470	3.5977
13	0.0000	-0.0000	7.5116	0.7501	0.6450	3.8325
14	-0.0000	-0.0000	7.3767	0.7483	0.6419	3.9374
15	0.0000	-0.0000	7.3472	0.7480	0.6412	3.9586

Semelhantemente aos casos anteriores, pode-se elaborar um gráfico baseado nos desvios-padrão contidos na tabela 4.16:

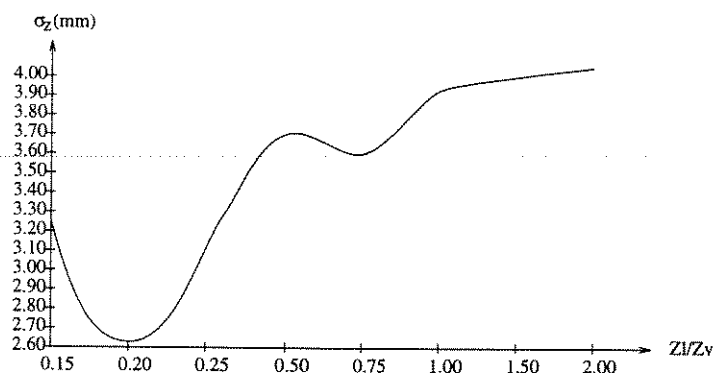


Figura 4.12: Curva gerada a partir dos desvios-padrão dos erros para imagens cilíndricas com distância de 5400 mm de projeção.

Da análise dos resultados apresentados nas Tabelas 4.17 e 4.18, da mesma forma que no caso das superfícies esféricas, há uma redução nos desvios-padrão dos erros das coordenadas quando as informações de borda são utilizadas.

Ainda, nas tabelas 4.9 e 4.10, observa-se uma tendência a erros sistemáticos semelhante ao caso das superfícies esféricas. E da mesma forma, quando as informações de bordas são introduzidas no processamento, ocorre uma tendência a erros sistemáticos nas coordenadas X e Y.

4.1.4 Experimentos com ruídos nas Imagens de intensidade

A fim de analisar a influência dos ruídos, foram introduzidos ruídos gaussianos de 2% do valor da intensidade máxima de uma imagem, cujo valor é de 5 unidades em tons de cinza.

Assim, foram realizados testes nas imagens de superfícies esféricas, correspondentes às imagens 1 a 7 e 8 a 14 das tabelas 4.3 e 4.4, respectivamente.

Tabela 4.17: Médias e Desvios-Padrão dos erros das coordenadas nos testes (16 a 29) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
16	0.3834	-0.3840	-1.0475	0.4296	0.3852	5.2275
17	0.3826	-0.3829	-1.2074	0.3418	0.3059	4.3141
18	0.3914	-0.3673	-1.1813	0.3481	0.2990	4.2964
19	1.0091	-0.3800	-1.2087	0.3656	0.3030	4.3176
20	1.0091	-0.3800	-1.2087	0.3656	0.3030	4.3176
21	1.0088	-0.3802	-1.2094	0.3656	0.3031	4.3176
22	1.0086	-0.3803	-1.2104	0.3657	0.3032	4.3198
23	0.0000	0.0000	-1.8147	0.5278	0.4025	5.1730
24	0.0004	0.0000	-1.2513	0.3060	0.3147	3.7285
25	0.0035	-0.0153	-1.2489	0.3029	0.3143	3.6948
26	-0.0056	0.0002	-1.2617	0.3057	0.3149	3.6948
27	-0.5085	-0.0001	-1.2774	0.3349	0.2957	3.7396
28	-0.5122	-0.0005	-1.2801	0.3366	0.2954	3.7335
29	-0.6607	-0.0493	-1.3009	0.3282	0.2992	3.7880

Tabela 4.18: Médias e Desvios-Padrão para as coordenadas nos testes (30 a 43) realizados

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
30	-0.0000	0.0091	-3.1517	0.6593	0.5921	7.6419
31	-0.0002	0.0000	-2.1262	0.3778	0.4656	5.5869
32	-0.0307	-0.0115	-2.1223	0.3779	0.4646	5.5835
33	0.0049	-0.0001	-2.1199	0.3781	0.4656	5.5914
34	0.0081	-0.0003	-2.1216	0.3788	0.4660	5.6014
35	0.1517	-0.0151	-2.1290	0.3814	0.4720	5.6214
36	1.5841	-0.0221	-2.1711	0.4502	0.4655	5.6627
37	0.0000	-0.0001	-0.4505	0.5833	0.4760	6.4509
38	-0.0001	-0.0000	-0.2903	0.5409	0.5543	5.5288
39	-0.3786	0.0631	-0.2931	0.5549	0.5379	5.5339
40	-1.4639	0.0018	-0.3010	0.6185	0.4943	5.5538
41	-1.4680	0.0019	-0.3005	0.6194	0.4944	5.5555
42	-1.4723	0.0019	-0.3014	0.6200	0.4945	5.5563
43	-1.4767	0.0018	-0.3022	0.6206	0.4946	5.5573

Tabela 4.19: Médias e Desvios-Padrão para as coordenadas nos testes (1 a 14 - da superfície esférica) realizadas em imagens com ruídos

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
1	-0.0008	0.0003	5.9872	0.5243	0.5237	1.9463
2	-0.0000	-0.0000	6.7523	0.5477	0.5475	1.4724
3	-0.0001	-0.0000	7.0130	0.5411	0.5408	1.4100
4	-0.0003	0.0000	6.5207	0.5347	0.5341	1.5870
5	-0.0006	0.0001	6.1356	0.5265	0.5259	1.8135
6	-0.0010	0.0004	5.8585	0.5213	0.5207	2.0559
7	-0.0012	0.0005	5.7953	0.5195	0.5189	2.1042
8	0.0018	0.0011	17.4218	1.4732	1.4729	5.5694
9	0.0004	-0.0014	17.7895	1.5003	1.4901	6.6309
10	0.0037	-0.0029	18.5093	1.5240	1.4860	5.8695
11	0.0011	0.0002	18.3518	1.4913	1.4911	4.9170
12	0.0014	0.0008	17.7210	1.4787	1.4784	5.3691
13	0.0024	0.0017	17.1387	1.4685	1.4682	5.7746
14	0.0030	0.0022	17.0187	1.4674	1.4670	5.8799

Da mesma forma, foram realizados testes nas imagens de superfícies cilíndricas, correspondentes às imagens 1 a 7 e 8 a 14 das tabelas 4.11 e 4.12, respectivamente. Os resultados são mostrados na tabela 4.20.

Da análise dos resultados das tabelas 4.19 e 4.20, verifica-se que os ruídos introduzidos nas imagens afetam a precisão da reconstrução, cujos valores dos desvios-padrão na coordenada Z aumentam. O efeito dos ruídos nas imagens para reconstrução são percebidos, com maior intensidade, quando a distância entre a câmara e objeto são menores. Condição esta que é condicionada pelo efeito da ampliação (escala da imagem) do erro (ruído), de grandezas que estão no espaço imagem para o espaço objeto.

Tabela 4.20: Médias e Desvios-Padrão para as coordenadas nos testes (1 a 14 - superfície cilíndrica) realizadas em imagens com ruídos

Imagem	Média (mm)			Desvio Padrão (mm)		
	X	Y	Z	X	Y	Z
1	-0.0002	-0.0002	4.9798	0.7529	0.6924	1.8056
2	-0.0004	-0.0002	7.2993	0.7963	0.7872	2.3449
3	-0.0001	-0.0001	5.0340	0.7841	0.6948	2.1109
4	-0.0003	-0.0005	5.6289	0.7660	0.7224	1.5363
5	-0.0003	-0.0005	5.1986	0.7578	0.7025	1.7047
6	-0.0003	-0.0004	4.7897	0.7493	0.6841	1.9153
7	-0.0003	-0.0004	4.6960	0.7469	0.6798	1.9664
8	-0.0000	-0.0002	8.4979	0.7704	0.6640	3.2175
9	-0.0001	-0.0001	7.5818	0.7429	0.6471	3.6044
10	-0.0000	-0.0002	9.0865	0.7723	0.6924	2.6206
11	-0.0001	-0.0002	8.1024	0.7600	0.6603	3.3098
12	-0.0000	-0.0001	7.6623	0.7525	0.6487	3.7179
13	-0.0001	-0.0001	7.5169	0.7503	0.6450	3.8383
14	-0.0000	-0.0001	7.3807	0.7485	0.6419	3.9427
15	-0.0001	-0.0001	7.3524	0.7481	0.6413	3.9619

4.2 Análise dos resultados

A análise dos resultados obtidos nos experimentos realizados nas superfícies planas, esféricas e cilíndricas podem ser divididas em dois itens: o primeiro para a superfície plana, pois estes resultados já foram discutidos em Hasegawa e Tozzi [17], e o outro para as superfícies esféricas e cilíndricas.

4.2.1 Análise dos resultados da superfície plana

A análise dos resultados, obtidos nos experimentos realizados nas superfícies planas já foi discutido em Hasegawa e Tozzi [17]. Desta forma, reescrevendo e acrescentando algumas considerações, observando-se as tabelas 4.1 e 4.2, podem ser elaboradas as seguintes análises:

- há uma pequena melhora nos resultados obtidos no Experimento 1 em relação ao Experimento 2, resultado este que era esperado, pois o modelo proposto considera somente o componente difuso da luz;
- o processo simultâneo de calibração e reconstrução diminui o erro, como pode ser observado comparando-se os resultados dos Experimentos 1 e 3. Este fato pode ser verificado observando-se que, no experimento 3, nenhuma informação à respeito da superfície foi considerada enquanto que, no Experimento 1, com melhor precisão, as injunções foram estabelecidas em pontos da superfície;
- as experiências indicaram que a posição da câmara e da fonte de luz são parâmetros importantes. Isto pode ser verificado comparando-se o Experimento (1) ao (5), este último com melhores resultados. Estes experimentos indicam que a geometria (interseção dos raios), proporcionada pelos raios da luz e da câmara, determinam um melhor resultado para este tipo de superfície;
- o Experimento 4 indica que o modelo proposto apresenta melhores resultados do que aqueles que consideram a direção da fonte de luz constante. Isto pode ser visto pelo aumento das variâncias das coordenadas X e Y em relação aos obtidos no Experimento 1;
- quanto aos parâmetros de calibração da câmara, como pode ser visto nos experimentos 1 e 6 (Tabela 4.2), foi verificado que a quantidade de pontos, quando definidos com precisão (como no caso de imagens simuladas), não alteram os resultados.

4.2.2 Análise dos resultados para as superfícies esféricas e cilíndricas

Analisando-se as tabelas 4.7 a 4.10 e 4.15 a 4.18, pode-se tecer alguns comentários baseados nos resultados obtidos para as superfícies esféricas e cilíndricas:

- a posição da fonte de luz em relação ao objeto é um parâmetro importante, pois quanto mais próxima do objeto, melhor o resultado;
- a melhora proporcionada pela aproximação da fonte de luz é limitada até um certo nível. No caso dos testes realizados, a posição da fonte de luz para os melhores resultados variou entre 25% e 50% da distância de projeção, condição esta que pode ser verificada pelos gráficos mostrados nas figuras 4.6 e 4.7;
- notou-se que para posições da fonte de luz mais próximas do objeto do que os valores acima citados, os desvios-padrão dos erros tendem a aumentar. O aumento dos desvios-padrão dos erros podem ser provocados pelo fato de que a proximidade da fonte de luz na superfície a ser reconstruída, provoca uma maior variação dos valores de brilho uma vez que os ângulo entre a normal e a direção da fonte de luz apresenta uma variação maior, quando comparado com uma fonte de luz mais afastada.
- uma outra condição pode estar associada ao aumento do erro devido a proximidade da fonte de luz na superfície. A forma de amostragem utilizada, divisão uniforme dos *patches* na imagem, provoca *patches* irregulares e de diferentes tamanhos em superfícies inclinadas, condição esta que pode ser verificada nas figuras 4.4, 4.5, 4.9 e 4.10.
- inversamente, o resultado tende a piorar quando a fonte de luz se afasta do objeto, mantendo fixa a posição da câmara (observador), tendendo a estabilizar quando a fonte de luz está há duas vezes a distância da câmara, ou seja, quando a fonte de luz tende ao infinito. Esta condição pode significar que a posição da fonte de luz não influenciará nas determinações das intensidades, quando esta ultrapassar duas vezes a distância de projeção. Assim, neste caso, somente a direção da fonte de luz contribui para a determinação da intensidade;
- no caso das superfícies esféricas e cilíndricas, o resultado tende a piorar quando a fonte de luz (no mesmo plano do observador) se inclina em relação ao observador. O aumento dos erros, nestes casos, pode estar relacionado às posições da fonte de luz, observador e direção da normal, que pode proporcionar uma combinação geométrica predominante sobre as informações de *shading*;

- os teste indicam que o processamento considerando as arestas nas imagens, quando existentes, afetam positivamente os resultados. Isto pode ser verificado nas tabelas 4.9, 4.10, 4.17 e 4.18.
- os resultados mostram que, quanto mais próximo o observador do objeto, melhor a precisão do método, condição esta que era esperada, pois a proximidade do observador melhora a resolução e, conseqüentemente, a discretização do objeto.
- a tendência de erros sistemáticos verificados nos testes podem ter sido provocados pela perda de informações¹, inerente ao sistema de projeção perspectiva, devido a redução da dimensionalidade (3D para 2D).
- nas regiões onde a normal e a fonte de luz fazem um ângulo próximo de 90° (superfície mais inclinadas em relação a fonte de luz) os resultados tendem a piorar. Isto pode ser verificado, tomando a equação de *shading* 3.8, calculando-se a variação de $I_{(i,j)}$ em função do ângulo i

$$\left| \frac{\partial I_{(i,j)}}{\partial i} \right| = K_d \sin(i) \quad (4.12)$$

Verifica-se na equação 4.12, que a variação, $\frac{\partial I_{(i,j)}}{\partial i}$, aumenta com o aumento do ângulo i .

- os ruídos tendem a deteriorar o processo de reconstrução. O erro provocado pelo ruído, com desvio-padrão de 5 tons de cinza, não afeta tanto quanto um erro de posição do pixel na imagem, pois neste procedimento há uma mistura das informações de naturezas geométricas e radiométricas, sendo que as entidades geométricas são mais rígidas no processo de reconstrução.

Portanto, verifica-se que este procedimento, mesmo considerando somente um tipo de superfície (albedo constante), pode ser adaptado em reconstrução onde vários tipos de superfícies são imageadas, desde que uma segmentação e classificação na imagem tenham sido realizadas.

¹Quando duas imagens são tomadas, estéreo-visão, a informação perdida pode ser recuperada.

Capítulo 5

Conclusões

Neste capítulo, inicialmente, será realizada uma síntese do trabalho desenvolvido. Posteriormente, serão apresentadas as conclusões com base na análise dos resultados dos experimentos realizados. Para finalizar, serão discutidas algumas recomendações a partir das dificuldades encontradas no processamento do modelo proposto.

5.1 Síntese do Trabalho

Neste trabalho, foi proposto um algoritmo de reconstrução de superfícies, através da intensidade de cinza e das propriedades geométricas de uma imagem, sob a consideração da projeção perspectiva e com processamento iterativo. Foram utilizadas as equações de *shading* e de colinearidade. Ainda, com o intuito de eliminar a inconsistência do modelo matemático provocada pela falta de equações, foram, utilizadas informações auxiliares a respeito da superfície gerada, sem condicioná-la a uma forma geométrica. A superfície reconstruída foi considerada suave e formada por *patches* planares triangulares, descritos por pontos adjacentes, mantendo a correspondência com o pixel na imagem.

Para resolver o problema, foi desenvolvido um modelo matemático explícito para recuperação da informação 3D da superfície e dos parâmetros de calibração da câmara, realizando-se, assim, uma determinação simultânea dos parâmetros.

Devido ao uso de equações não lineares no modelo de reconstrução, o Método dos Mínimos Quadrados (MMQ) foi aplicado com a linearização do modelo por série de Taylor. Ainda, com o uso do MMQ, foi possível, sem aumentar muito o esforço computacional, inserir injunções (informações) de pontos de controle para definir um sistema de coordenadas (global)

e calibrar a câmara, permitindo ainda, a inserção de injunções de bordas.

Todas as etapas do trabalho foram programadas em linguagem C. As imagens foram geradas adotando-se o padrão ASCII/PBM do XV.

5.2 Conclusões

Os resultados dos experimentos realizados foram analisados, levando a uma avaliação do modelo proposto em relação ao nível de precisão atingido, às configurações problemáticas e às restrições do modelo.

O modelo é apropriado para calibração da câmara e reconstrução simultânea da superfície diferindo de outros processos encontrados na literatura, pela obtenção simultânea das coordenadas tridimensionais dos pontos e dos parâmetros de calibração da câmara, usando a projeção perspectiva. Ainda, para a solução dos sistemas de equações, não há qualquer restrição sobre a superfície, como observado em outros procedimentos.

Todas as equações do modelo, inclusive as relacionadas com as injunções das bordas, foram agrupadas de modo explícito, de forma a permitir a utilização do Método dos Mínimos Quadrados na solução do sistema.

O modelo proposto foi avaliado usando-se dados sintéticos, e os experimentos mais significativos foram apresentados.

As características principais do modelo proposto podem ser resumidas assim:

- o uso da projeção perspectiva, que torna possível a reconstrução de imagens reais;
- versatilidade em processar vários tipos de superfícies suaves, não sendo necessário realizar nenhuma condição sobre a forma (conhecimento prévio) da superfície;
- determinação simultânea dos parâmetros de calibração da câmara e das coordenadas dos pontos da superfície;
- versatilidade em realizar injunções nos parâmetros, tanto nos pontos como nos parâmetros de orientação;
- possibilidade de recuperação somente da geometria 3D da superfície, considerando-se conhecida a posição da câmara (injunção nos elementos de calibração da câmara); e
- reconstrução da superfície no sistema de coordenadas global (sistema do espaço objeto "real")

Verifica-se nos experimentos, que o erro cresce com o aumento da distância entre a câmara e/ou objeto, fator este que pode estar relacionado com a resolução (capacidade de distinguir a separação entre dois pontos) pois quanto mais próxima a câmara do objeto, maior será a resolução. Verifica-se, ainda, que a posição da fonte de luz é importante, sendo que quanto mais próxima do objeto, melhor a precisão obtida. Esta característica pode se deteriorar quando a fonte se aproxima muito da superfície.

Quando a fonte de luz se afasta do objeto enquanto a câmara permanece fixa, tende a piorar a precisão da reconstrução. Isto ocorre até um certo nível de afastamento, estabilizando-se em distâncias maiores. Este fato (estabilização do erro em distâncias maiores) pode estar relacionado à ortogonalização dos raios de luz, ou seja, os raios se tornam paralelos ou quase paralelos para distâncias grandes, condição que faz a posição da fonte de luz tornar-se indiferente no processamento (importando somente a direção da fonte).

A integração das informações de borda, na forma de injunções ao modelo proposto para solução do problema de SFS, reduz o erro de reconstrução. A inclusão das injunções de borda não acarreta aumento dos parâmetros incógnitos, aumentando somente o número de observações. Conseqüentemente, não aumenta a dimensão da matriz N (que precisa ter a inversa calculada neste procedimento de reconstrução). Portanto, a inclusão de informações de bordas no processamento não representa nenhum aumento significativo do esforço computacional.

O procedimento de reconstrução utilizado, processamento simultâneo de todas as coordenadas da imagem e dos parâmetros de calibração da câmara, tende a reduzir a influência de ruídos, produzindo resultados mais confiáveis e proporcionando melhores precisões. Entretanto, estes benefícios proporcionados pelo processamento simultâneo aumentam o esforço computacional e tornam a técnica dependente da memória RAM do computador. O esforço computacional está diretamente ligado ao processo de resolução do sistema equações pois, neste caso, se faz necessário inverter a matriz N (no processo sem otimização realiza-se n^3 operações) para se determinar os valores das incógnitas.

A partir dos resultados obtidos no processo de reconstrução, algumas expectativas de precisão podem ser estabelecidas:

- para objetos que não provocam silhuetas nas imagens, obtiveram erros entre 0,08% a 0,11% da distância média de projeção.
- para objetos que produzem silhuetas nas imagens, o erro variou de 0,11% a 0,16% da distância de projeção, sendo que, houve um caso (reconstrução do cilindro com

distância de projeção de 1800 mm) em que a discrepância entre os valores reais e medidos atingiu variações de 0,23% a 0,29% da distância de projeção.

Portanto, pode-se verificar que, para imagens com silhuetas produzidas por bordas, o efeito causado na reconstrução tende a piorar o resultado. Assim, a influência das bordas que provocam silhuetas nas imagens podem ser minimizadas, considerando-as no processamento.

5.3 Recomendações

Com base nas dificuldades encontradas no processamento de várias configurações, com o modelo de reconstrução proposto, pode-se comentar e sugerir algumas recomendações.

O processamento, por ser simultâneo, já comentado anteriormente, tem como limitação a capacidade de memória e o custo computacional, pois o processo é dependente da resolução de sistema de equações, o que requer muitas operações. Quanto à questão de capacidade de memória, pode-se tentar processar os dados recursivamente. Ainda, pode-se desenvolver um procedimento de diminuição do número de parâmetros, manipulando-se as equações de colinearidade e de *shading*.

O sistema formulado envolve observações de intensidade (*shading*) e da posição (geometria), requerendo, assim, uma padronização. Esta padronização foi realizada usando-se pesos diferentes nas observações, pesos estes que foram determinados empiricamente.

Desta forma, baseado nos experimentos realizados neste trabalho, algumas sugestões para trabalhos futuros podem ser dadas:

- desenvolvimento de uma fórmula empírica em função dos gráficos, para o posicionamento ótimo da fonte de luz e da câmara;
- desenvolvimento de um processamento que considere os efeitos da amostragem uniforme na imagem, que provoca *patches* irregulares quando há variação de inclinação na superfície;
- estudar a causa dos prováveis erros sistemáticos ocorridos na reconstrução da superfície e uma forma de eliminá-las, aumentando a precisão;
- solução com informações adicionais, por exemplo: imagens *range*, ultra-som, laser;
- combinação da solução com estéreo ou outros métodos de *shape from X* (Ex: *motion*, *optical flow*).

Apêndice A

Derivadas Parciais

A.1 Derivadas Parciais das Equações de Colineridade

Para cada uma das equações (3.6) e (3.7) existem três derivadas parciais, correspondentes às três incógnitas das coordenadas do ponto e mais seis correspondentes aos parâmetros de calibração da câmera:

$$\frac{\partial x_p}{\partial \omega} = \frac{x}{q}(-m_{33}DX + m_{32}DZ) + \frac{f}{q}(-m_{13}DY + m_{12}DZ)$$

$$\begin{aligned} \frac{\partial x_p}{\partial \phi} = & \frac{x}{q}[DX \cos(\phi) + DY(\sin(\omega) \sin(\phi) + DZ(-\sin(\phi) \cos(\omega))] + \\ & \frac{f}{q}[DX(-\sin(\phi) \cos(\kappa) + DY(\sin(\omega) \cos(\phi) \cos(\omega) + DZ(-\cos(\omega) \cos(\phi) \cos(\kappa))] \end{aligned}$$

$$\frac{\partial x_p}{\partial \kappa} = \frac{f}{q}(m_{21}DX + m_{22}DY + m_{23}DZ)$$

$$\frac{\partial x_p}{\partial X_c} = \frac{x}{q}m_{31} + \frac{f}{q}m_{11}$$

$$\frac{\partial x_p}{\partial Y_c} = \frac{x}{q}m_{32} + \frac{f}{q}m_{12}$$

$$\frac{\partial x_p}{\partial Z_c} = \frac{x}{q}m_{33} + \frac{f}{q}m_{13}$$

$$\frac{\partial x_p}{\partial X} = -\left(\frac{x}{q}m_{31} + \frac{f}{q}m_{11}\right)$$

$$\frac{\partial x_p}{\partial Y} = -\left(\frac{x}{q}m_{32} + \frac{f}{q}m_{12}\right)$$

$$\frac{\partial x_p}{\partial Z} = -\left(\frac{x}{q}m_{33} + \frac{f}{q}m_{13}\right)$$

$$\frac{\partial y_p}{\partial \omega} = \frac{y}{q}(-m_{33}DY + m_{32}DZ) + \frac{f}{q}(-m_{23}DY + m_{22}DZ)$$

$$\begin{aligned} \frac{\partial y_p}{\partial \phi} = & \frac{y}{q}[DX \cos(\phi) + DY(\sin(\omega) \sin(\phi) + DZ(-\sin(\phi) \cos(\omega))] + \\ & \frac{f}{q}[DX(\sin(\phi) \sin(\kappa) + DY(-\sin(\omega) \cos(\phi) \sin(\kappa) + DZ(\cos(\omega) \cos(\phi) \sin(\kappa))] \end{aligned}$$

$$\frac{\partial y_p}{\partial \kappa} = \frac{f}{q}(-m_{11}DX - m_{12}DY - m_{13}DZ)$$

$$\frac{\partial y_p}{\partial X_c} = \frac{y}{q}m_{31} + \frac{f}{q}m_{21}$$

$$\frac{\partial y_p}{\partial Y_c} = \frac{y}{q}m_{32} + \frac{f}{q}m_{22}$$

$$\frac{\partial y_p}{\partial Z_c} = \frac{y}{q}m_{33} + \frac{f}{q}m_{23}$$

$$\frac{\partial y_p}{\partial X} = -\left(\frac{y}{q}m_{31} + \frac{f}{q}m_{21}\right)$$

$$\frac{\partial y_p}{\partial Y} = -\left(\frac{y}{q}m_{32} + \frac{f}{q}m_{22}\right)$$

$$\frac{\partial y_p}{\partial Z} = -\left(\frac{y}{q}m_{33} + \frac{f}{q}m_{23}\right)$$

onde:

$$DX = X - X_c$$

$$DY = Y - Y_c$$

$$DZ = Z - Z_c$$

$$q = m_{31}DX + m_{32}DY + m_{33}DZ$$

A.2 Derivadas Parciais da Equação de *Shading*

A equação (3.8) proporciona nove derivadas parciais. Os nove parâmetros são referentes às três coordenadas dos três pontos envolvidos no cálculo do vetor normal sobre o plano definido por estes pontos.

$$\begin{aligned} & \frac{\partial I_{i,j}}{\partial X_{(i,j)}}; \frac{\partial I_{i,j}}{\partial Y_{(i,j)}}; \frac{\partial I_{i,j}}{\partial Z_{(i,j)}}; \\ & \frac{\partial I_{i,j}}{\partial X_{(i+1,j)}}; \frac{\partial I_{i,j}}{\partial Y_{(i+1,j)}}; \frac{\partial I_{i,j}}{\partial Z_{(i+1,j)}}; \\ & \frac{\partial I_{i,j}}{\partial X_{(i,j+1)}}; \frac{\partial I_{i,j}}{\partial Y_{(i,j+1)}} e \frac{\partial I_{i,j}}{\partial Z_{(i,j+1)}} \end{aligned}$$

com:

$$N_1 = (Y_{(i,j)} - Y_{(i+1,j)})(Z_{(i,j)} - Z_{(i,j+1)}) - (Y_{(i,j)} - Y_{(i,j+1)})(Z_{(i,j)} - Z_{(i+1,j)}),$$

$$N_2 = (X_{(i,j)} - X_{(i,j+1)})(Z_{(i,j)} - Z_{(i+1,j)}) - (X_{(i,j)} - X_{(i+1,j)})(Z_{(i,j)} - Z_{(i,j+1)}),$$

$$N_3 = (X_{(i,j)} - X_{(i+1,j)})(Y_{(i,j)} - Y_{(i,j+1)}) - (X_{(i,j)} - X_{(i,j+1)})(Y_{(i,j)} - Y_{(i+1,j)}),$$

$$D_n = \sqrt{N_1^2 + N_2^2 + N_3^2}$$

Para o ponto pivô da configuração tem-se:

$$L_1 = X_{(i,j)} - X_l, L_2 = Y_{(i,j)} - Y_l, L_3 = Z_{(i,j)} - Z_l,$$

$$D_l = \sqrt{(L_1)^2 + (L_2)^2 + (L_3)^2}$$

$$N_0 = N_1 L_1 + N_2 L_2 + N_3 L_3$$

- derivadas em relação a $X_{(i,j)}$:

$$\frac{\partial I_{i,j}}{\partial X_{(i,j)}} = \frac{[\frac{\partial N_0}{\partial X_{(i,j)}} D_n D_l] - [\frac{\partial D_n}{\partial X_{(i,j)}} D_l + \frac{\partial D_l}{\partial X_{(i,j)}} D_n] N_0}{D_n^2 D_l^2}, \quad (\text{A.1})$$

$$\begin{aligned} \frac{\partial N_0}{\partial X_{(i,j)}} &= \frac{\partial N_1}{\partial X_{(i,j)}} L_1 + \frac{\partial N_2}{\partial X_{(i,j)}} L_2 + \frac{\partial N_3}{\partial X_{(i,j)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial X_{(i,j)}} N_1 + \frac{\partial L_2}{\partial X_{(i,j)}} N_2 + \frac{\partial L_3}{\partial X_{(i,j)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial X_{(i,j)}} = 0,$$

$$\frac{\partial N_2}{\partial X_{(i,j)}} = Z_{(i,j+1)} - Z_{(i+1,j)},$$

$$\frac{\partial N_3}{\partial X_{(i,j)}} = Y_{(i+1,j)} - Y_{(i,j+1)},$$

$$\frac{\partial L_1}{\partial X_{(i,j)}} = 1,$$

$$\frac{\partial L_2}{\partial X_{(i,j)}} = \frac{\partial L_3}{\partial X_{(i,j)}} = 0,$$

$$\frac{\partial D_n}{\partial X_{(i,j)}} = [N_2(Z_{(i,j+1)} - Z_{(i+1,j)}) + N_3(Y_{(i+1,j)} - Y_{(i,j+1)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial X_{(i,j)}} = (X_{(i,j)} - X_L) D_l^{-1/2}$$

- derivadas em relação a $Y_{(i,j)}$:

$$\frac{\partial I_{i,j}}{\partial Y_{(i,j)}} = \frac{[\frac{\partial N_0}{\partial Y_{(i,j)}} D_n D_l] - [\frac{\partial D_n}{\partial Y_{(i,j)}} D_l + \frac{\partial D_l}{\partial Y_{(i,j)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.2})$$

$$\begin{aligned} \frac{\partial N_0}{\partial Y_{(i,j)}} &= \frac{\partial N_1}{\partial Y_{(i,j)}} L_1 + \frac{\partial N_2}{\partial Y_{(i,j)}} L_2 + \frac{\partial N_3}{\partial Y_{(i,j)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial Y_{(i,j)}} N_1 + \frac{\partial L_2}{\partial Y_{(i,j)}} N_2 + \frac{\partial L_3}{\partial Y_{(i,j)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial Y_{(i,j)}} = Z_{(i+1,j)} - Z_{(i,j+1)}$$

$$\frac{\partial N_2}{\partial Y_{(i,j)}} = 0$$

$$\frac{\partial N_3}{\partial Y_{(i,j)}} = X_{(i,j+1)} - X_{(i+1,j)}$$

$$\frac{\partial L_1}{\partial Y_{(i,j)}} = \frac{\partial L_3}{\partial Y_{(i,j)}} = 0$$

$$\frac{\partial L_2}{\partial Y_{(i,j)}} = 1$$

$$\frac{\partial D_n}{\partial Y_{(i,j)}} = [N_1(Z_{(i+1,j)} - Z_{(i,j+1)}) + N_3(X_{(i,j+1)} - X_{(i+1,j)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Y_{(i,j)}} = (Y_{(i,j)} - Y_L) D_l^{-1/2}$$

- derivadas em relação a $Z_{(i,j)}$:

$$\frac{\partial I_{i,j}}{\partial Z_{(i,j)}} = \frac{[\frac{\partial N_0}{\partial Z_{(i,j)}} D_n D_l] - [\frac{\partial D_n}{\partial Z_{(i,j)}} D_l + \frac{\partial D_l}{\partial Z_{(i,j)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.3})$$

$$\frac{\partial N_0}{\partial Z_{(i,j)}} = \frac{\partial N_1}{\partial Z_{(i,j)}} L_1 + \frac{\partial N_2}{\partial Z_{(i,j)}} L_2 + \frac{\partial N_3}{\partial Z_{(i,j)}} L_3 +$$

$$\frac{\partial L_1}{\partial Z_{(i,j)}} N_1 + \frac{\partial L_2}{\partial Z_{(i,j)}} N_2 + \frac{\partial L_3}{\partial Z_{(i,j)}} N_3$$

$$\frac{\partial N_1}{\partial Z_{(i,j)}} = Y_{(i,j+1)} - Y_{(i+1,j)}$$

$$\frac{\partial N_2}{\partial Z_{(i,j)}} = X_{(i+1,j)} - X_{(i,j+1)}$$

$$\frac{\partial N_3}{\partial Z_{(i,j)}} = 0$$

$$\frac{\partial L_1}{\partial Z_{(i,j)}} = \frac{\partial L_2}{\partial Z_{(i,j)}} = 0$$

$$\frac{\partial L_3}{\partial Z_{(i,j)}} = 1$$

$$\frac{\partial D_n}{\partial Z_{(i,j)}} = [N_1(Y_{(i,j+1)} - Y_{(i+1,j)}) + N_2(X_{(i+1,j)} - X_{(i,j+1)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Z_{(i,j)}} = (Z_{(i,j)} - Z_L) D_l^{-1/2}$$

- derivadas em relação a $X_{(i,j+1)}$:

$$L_1 = X_{(i,j+1)} - X_L, L_2 = Y_{(i,j+1)} - Y_L, L_3 = Z_{(i,j+1)} - Z_L$$

$$D_l = \sqrt{(L_1)^2 + (L_2)^2 + (L_3)^2}$$

$$N_0 = N_1 L_1 + N_2 L_2 + N_3 L_3$$

$$\frac{\partial I_{i,j}}{\partial X_{(i,j+1)}} = \frac{[\frac{\partial N_0}{\partial X_{(i,j+1)}} D_n D_l] - [\frac{\partial D_n}{\partial X_{(i,j+1)}} D_l + \frac{\partial D_l}{\partial X_{(i,j+1)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.4})$$

$$\frac{\partial N_0}{\partial X_{(i,j+1)}} = \frac{\partial N_1}{\partial X_{(i,j+1)}} L_1 + \frac{\partial N_2}{\partial X_{(i,j+1)}} L_2 + \frac{\partial N_3}{\partial X_{(i,j+1)}} L_3 +$$

$$\frac{\partial L_1}{\partial X_{(i,j+1)}} N_1 + \frac{\partial L_2}{\partial X_{(i,j+1)}} N_2 + \frac{\partial L_3}{\partial X_{(i,j+1)}} N_3$$

$$\frac{\partial N_1}{\partial X_{(i,j+1)}} = 0$$

$$\frac{\partial N_2}{\partial X_{(i,j+1)}} = Z_{(i+1,j)} - Z_{(i,j)}$$

$$\frac{\partial N_3}{\partial X_{(i,j+1)}} = Y_{(i,j)} - Y_{(i+1,j)}$$

$$\frac{\partial L_1}{\partial X_{(i,j+1)}} = \frac{\partial L_2}{\partial X_{(i,j+1)}} = \frac{\partial L_3}{\partial X_{(i,j+1)}} = 0$$

$$\frac{\partial D_n}{\partial X_{(i,j+1)}} = [N_2(Z_{(i+1,j)} - Z_{(i,j)}) + N_3(Y_{(i,j)} - Y_{(i+1,j)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial X_{(i,j+1)}} = 0$$

- derivadas em relação a $Y_{(i,j+1)}$:

$$\frac{\partial I_{i,j}}{\partial Y_{(i,j+1)}} = \frac{[\frac{\partial N_0}{\partial Y_{(i,j+1)}} D_n D_l] - [\frac{\partial D_n}{\partial Y_{(i,j+1)}} D_l + \frac{\partial D_l}{\partial Y_{(i,j+1)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.5})$$

$$\frac{\partial N_0}{\partial Y_{(i,j+1)}} = \frac{\partial N_1}{\partial Y_{(i,j+1)}} L_1 + \frac{\partial N_2}{\partial Y_{(i,j+1)}} L_2 + \frac{\partial N_3}{\partial Y_{(i,j+1)}} L_3 +$$

$$\frac{\partial L_1}{\partial Y_{(i,j+1)}} N_1 + \frac{\partial L_2}{\partial Y_{(i,j+1)}} N_2 + \frac{\partial L_3}{\partial Y_{(i,j+1)}} N_3$$

$$\frac{\partial N_1}{\partial Y_{(i,j+1)}} = Z_{(i,j)} - Z_{(i+1,j)}$$

$$\frac{\partial N_2}{\partial Y_{(i,j+1)}} = 0$$

$$\frac{\partial N_3}{\partial Y_{(i,j+1)}} = X_{(i+1,j)} - Y_{(i,j)}$$

$$\frac{\partial L_1}{\partial Y_{(i,j+1)}} = \frac{\partial L_3}{\partial Y_{(i,j+1)}} = \frac{\partial L_2}{\partial X_{(i,j+1)}} = 0$$

$$\frac{\partial D_n}{\partial Y_{(i,j+1)}} = [N_1(Z_{(i,j)} - Z_{(i+1,j)}) + N_3(X_{(i+1,j)} - X_{(i,j)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Y_{(i,j+1)}} = 0$$

- derivadas em relação a $Z_{(i,j+1)}$:

$$\frac{\partial I_{i,j}}{\partial Z_{(i,j+1)}} = \frac{[\frac{\partial N_0}{\partial Z_{(i,j+1)}} D_n D_l] - [\frac{\partial D_n}{\partial Z_{(i,j+1)}} D_l + \frac{\partial D_l}{\partial Z_{(i,j+1)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.6})$$

$$\begin{aligned} \frac{\partial N_0}{\partial Z_{(i,j+1)}} &= \frac{\partial N_1}{\partial Z_{(i,j+1)}} L_1 + \frac{\partial N_2}{\partial Z_{(i,j+1)}} L_2 + \frac{\partial N_3}{\partial Z_{(i,j+1)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial Z_{(i,j+1)}} N_1 + \frac{\partial L_2}{\partial Z_{(i,j+1)}} N_2 + \frac{\partial L_3}{\partial Z_{(i,j+1)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial Z_{(i,j+1)}} = Y_{(i+1,j)} - Y_{(i,j)}$$

$$\frac{\partial N_2}{\partial Z_{(i,j+1)}} = X_{(i,j)} - X_{(i+1,j)}$$

$$\frac{\partial N_3}{\partial Z_{(i,j+1)}} = 0$$

$$\frac{\partial L_1}{\partial Z_{(i,j+1)}} = \frac{\partial L_2}{\partial Z_{(i,j+1)}} = \frac{\partial L_3}{\partial Z_{(i,j+1)}} = 0$$

$$\frac{\partial D_n}{\partial Z_{(i,j+1)}} = [N_1(Y_{(i+1,j)} - Y_{(i,j)}) + N_2(X_{(i,j)} - X_{(i+1,j)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Z_{(i,j+1)}} = 0$$

- derivadas em relação a $X_{(i+1,j)}$:

$$L_1 = X_{(i+1,j)} - X_l L_2 = Y_{(i+1,j)} - X_l L_3 = Z_{(i+1,j)} - Z_l$$

$$D_l = \sqrt{(L_1)^2 + (L_2)^2 + (L_3)^2}$$

$$N_0 = N_1 L_1 + N_2 L_2 + N_3 L_3$$

$$\frac{\partial I_{i,j}}{\partial X_{(i+1,j)}} = \frac{[\frac{\partial N_0}{\partial X_{(i+1,j)}} D_n D_l] - [\frac{\partial D_n}{\partial X_{(i+1,j)}} D_l + \frac{\partial D_l}{\partial X_{(i+1,j)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.7})$$

$$\begin{aligned} \frac{\partial N_0}{\partial X_{(i+1,j)}} &= \frac{\partial N_1}{\partial X_{(i+1,j)}} L_1 + \frac{\partial N_2}{\partial X_{(i+1,j)}} L_2 + \frac{\partial N_3}{\partial X_{(i+1,j)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial X_{(i+1,j)}} N_1 + \frac{\partial L_2}{\partial X_{(i+1,j)}} N_2 + \frac{\partial L_3}{\partial X_{(i+1,j)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial X_{(i+1,j)}} = 0$$

$$\frac{\partial N_2}{\partial X_{(i+1,j)}} = Z_{(i,j)} - Z_{(i,j+1)}$$

$$\frac{\partial N_3}{\partial X_{(i+1,j)}} = Y_{(i,j+1)} - Y_{(i,j)}$$

$$\frac{\partial L_1}{\partial X_{(i+1,j)}} = \frac{\partial L_2}{\partial X_{(i+1,j)}} = \frac{\partial L_3}{\partial X_{(i+1,j)}} = 0$$

$$\frac{\partial D_n}{\partial X_{(i+1,j)}} = [N_2(Z_{(i,j)} - Z_{(i,j+1)}) + N_3(Y_{(i,j+1)} - Y_{(i,j)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial X_{(i+1,j)}} = 0$$

- derivadas em relação a $Y_{(i+1,j)}$:

$$\frac{\partial I_{i,j}}{\partial Y_{(i+1,j)}} = \frac{[\frac{\partial N_0}{\partial Y_{(i+1,j)}} D_n D_l] - [\frac{\partial D_n}{\partial Y_{(i+1,j)}} D_l + \frac{\partial D_l}{\partial Y_{(i+1,j)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.8})$$

$$\begin{aligned} \frac{\partial N_0}{\partial Y_{(i+1,j)}} &= \frac{\partial N_1}{\partial Y_{(i+1,j)}} L_1 + \frac{\partial N_2}{\partial Y_{(i+1,j)}} L_2 + \frac{\partial N_3}{\partial Y_{(i+1,j)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial Y_{(i+1,j)}} N_1 + \frac{\partial L_2}{\partial Y_{(i+1,j)}} N_2 + \frac{\partial L_3}{\partial Y_{(i+1,j)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial Y_{(i+1,j)}} = Z_{(i,j+1)} - Z_{(i,j)}$$

$$\frac{\partial N_2}{\partial Y_{(i+1,j)}} = 0$$

$$\frac{\partial N_3}{\partial Y_{(i+1,j)}} = X_{(i,j)} - X_{(i,j+1)}$$

$$\frac{\partial L_1}{\partial Y_{(i+1,j)}} = \frac{\partial L_3}{\partial Y_{(i+1,j)}} = \frac{\partial L_2}{\partial Y_{(i+1,j)}} = 0$$

$$\frac{\partial D_n}{\partial Y_{(i+1,j)}} = [N_1(Z_{(i,j+1)} - Z_{(i,j)}) + N_3(X_{(i,j)} - X_{(i,j+1)})] D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Y_{(i+1,j)}} = 0$$

- derivadas em relação a $Z_{(i+1,j)}$:

$$\frac{\partial I_{i,j}}{\partial Z_{(i+1,j)}} = \frac{[\frac{\partial N_0}{\partial Z_{(i+1,j)}} D_n D_l] - [\frac{\partial D_n}{\partial Z_{(i+1,j)}} D_l + \frac{\partial D_l}{\partial Z_{(i+1,j)}} D_n] N_0}{D_n^2 D_l^2} \quad (\text{A.9})$$

$$\begin{aligned} \frac{\partial N_0}{\partial Z_{(i+1,j)}} &= \frac{\partial N_1}{\partial Z_{(i+1,j)}} L_1 + \frac{\partial N_2}{\partial Z_{(i+1,j)}} L_2 + \frac{\partial N_3}{\partial Z_{(i+1,j)}} L_3 + \\ &\quad \frac{\partial L_1}{\partial Z_{(i+1,j)}} N_1 + \frac{\partial L_2}{\partial Z_{(i+1,j)}} N_2 + \frac{\partial L_3}{\partial Z_{(i+1,j)}} N_3 \end{aligned}$$

$$\frac{\partial N_1}{\partial Z_{(i+1,j)}} = Y_{(i,j)} - Y_{(i,j+1)}$$

$$\frac{\partial N_2}{\partial Z_{(i+1,j)}} = X_{(i,j+1)} - X_{(i,j)}$$

$$\frac{\partial N_3}{\partial Z_{(i+1,j)}} = 0$$

$$\frac{\partial L_1}{\partial Z_{(i+1,j)}} = \frac{\partial L_2}{\partial Z_{(i+1,j)}} = \frac{\partial L_3}{\partial Z_{(i+1,j)}} = 0$$

$$\frac{\partial D_n}{\partial Z_{(i+1,j)}} = [N_1(Y_{(i,j)} - Y_{(i,j+1)}) + N_2(X_{(i,j+1)} - X_{(i,j)})]D_n^{-1/2} e$$

$$\frac{\partial D_l}{\partial Z_{(i+1,j)}} = 0$$

Referências Bibliográficas

- [1] J. ALOIMONOS. Visual shape computation. *Proc. IEEE*, 76(8):899–915, 1988.
- [2] D. H. BALLARD and C. M. BROWN. *Computer Vision*. Prentice Hall, 1982.
- [3] S. T. BARNARD. Interpreting perspective images. *Artificial Intelligence*, 21:435–462, 1983.
- [4] M. BRADY. Computational approaches to image understanding. *Computing Surveys*, 14(1), March 1982.
- [5] M. J. BROOKS and B. K. P. HORN. Shape and source from shading. *In: Shape from Shading, BROOKS, M. J. and HORN, B. K. P.(Eds)*, pages 53–68, 1989.
- [6] A. M. BRUCKSTEIN. On shape from shading. *Computer Vision, Graphics, and Image Processing*, 44:139–154, 1988.
- [7] K. R. CASTLEMAN. *Digital Image Processing*. Prentice-Hall, Inc., Englewood Cliffs, 1979.
- [8] J. E. CRYER, P. S. TSAI, and M. SHAH. Integration of shape from shading and stereo. *Pattern Recognition*, 28(7):1033–1043, 1995.
- [9] E.R DAVIES. *Machine Vision: Theory, Algorithms, Practicalities*. Academics Press, 1990.
- [10] J.D. FOLEY, A. VAN DAM, S.K. FEINER, and S.K. HUGHES. *Computer Graphics, Principles and Practice*. Addison Wesley, 1990.
- [11] M. GALO and C.L. TOZZI. Surface reconstruction using multiple light sources and perspective projection. *Proc. IEEE - International Conference on Image Processing, Lausanne, Switzerland*, pages 309–312, september 1996.
- [12] CAMIL GEMAEL. *Introdução ao ajustamento de observações*. UFPR, 1994.
- [13] R.C. GONZALEZ. *Digital Image Processing*. Addison-Wesley Publishing Company, 1987.

- [14] R. HALL. *Illumination and Color in Computer Generated Imagery*. Springer-Verlag, New York, 1989.
- [15] R. M. HARALICK and L. G. SHAPIRO. *Computer and Robot Vision*. Addison Wesley, 1993.
- [16] WECHESLER HARRY. *Computational Vision*. Academic Press, Inc., 1990.
- [17] J. K. HASEGAWA and C.L. TOZZI. Shape from shading with perspective projection and camera calibration. *Computer and Graphics*, 20(3):351–364, june 1996.
- [18] B. K. P. HORN. Understanding image intensities. *Artificial Intelligence*, 8:201–231, 1977.
- [19] B. K. P. HORN. Height and gradient from shading. In: *Shape from Shading*, BROOKS, M. J. and HORN, B. K. P.(Eds), pages 173–214, 1989.
- [20] B. K. P. HORN and M. J. BROOKS. Shape from shading. In: *Shape from Shading*, BROOKS, M. J. and HORN, B. K. P.(Eds), pages 01–28, 1989.
- [21] B. K. P. HORN and M. J. BROOKS. The variational approach to shape from shading. *International Journal of Computer Vision*, 5:37–75, 1990.
- [22] B. K. P. HORN and R. W. SJOBERG. Calculating the reflectance map. In: *Shape from Shading*, BROOKS, M. J. and HORN, B. K. P.(Eds), pages 215–244, 1989.
- [23] K. IKEUCHI and B. K. P. HORN. Numerical shape from shading and occluding boundaries. In: *Shape from Shading*, BROOKS, M. J. and HORN, B. K. P.(Eds), pages 245–299, 1989.
- [24] BERND JAHNE. *Spatio-Temporal Image Processing*. Springer-Verlag. Berlin-Germany, 1993.
- [25] R. JAIN, R. KASTURI, and B. G. SCHUNCK. *Machine Vision*. McGraw-Hill Book Co - USA, 1985.
- [26] K.I. KANATANI and T.C. CHOU. Shape from texture: General principle. *Artificial Intelligence*, 38:1–48, 1989.
- [27] RON KIMMEL and ALFRED M. BRUCKSTEIN. Global shape from shading. *Computer Vision and Image Understanding*, 62(3):360–369, 1995.
- [28] RON KIMMEL and ALFRED M. BRUCKSTEIN. Tracking level sets by level sets: A method for solving the shape from shading problem. *Computer Vision and Image Understanding*, 62(2):47–58, 1995.

- [29] D. C. KNILL and D. KERSTEN. Learning a near-optimal estimator for surface shape from shading. *Computer Vision, Graphics, and Image Processing*, 50:75–100, 1990.
- [30] K. M. LEE and J. KUO. Shape from shading with perspective projection. *CVGIP:Image Understanding*, 59(2):202–212, march 1994.
- [31] E. M. MIKHAIL and F. ACKERMANN. *Observations and Least Squares*. IFP-A Run Donnelley Publisher, 1976.
- [32] J. OLIENSIS. Shape from shading as a partially well-constrained problem. *CVGIP:Image Understanding*, 54(2):163–183, september 1991.
- [33] M. A. PENNA. Local and semi-local shape from shading for a single perspective image of a smooth object. *CVGIP*, 46:346–366, 1989.
- [34] M. A. PENNA. A shape from shading analysis for a single perspective image of a polyhedron. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11(6), 1989.
- [35] C.J. POELMAN and T. KANADE. A paraperspective factorization method for shape and motion recovery. In: *Lecture Notes in Computer Science - Computer Vision - ECCV'94* Springer-Verlag, 2(801):97–108, 1994.
- [36] D.F. ROGERS. *Procedural elements for computer graphics*. McGraw Hill, Inc. - Singapore, 1985.
- [37] J. L. C. SANZ. *Advances in machine vision*. Springer-Verlag New York Inc, 1989.
- [38] K. SUGIHARA. An algebraic approach to shape-from-shading problems. *Artificial Intelligence*, 23:59–95, 1984.
- [39] R. SZELISKI. Fast shape from shading. *CVGIP - Image Understanding*, 53(2):129–153, march 1991.
- [40] F. ULUPINAR and R. NEVATIA. Shape from contour: Straight homogeneous generalized cylinders and constant cross section generalized cylinders. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 17(2):120–135, february 1995.
- [41] O. E. VEGA and Y. H. YANG. Shading logic: A heuristic approach to recover shape from shading. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15(6), 1993.
- [42] P. R. WOLF. *Elements of Photogrammetry*. McGraw-Hill Book Co - USA, 1983.
- [43] S.Y.K. YUEN. Shape from contour using symmetries. *Procedure: First European Conference on Computer Vision. Antibes*, 1990.