

UNIVERSIDADE FEDERAL FLUMINENSE
CENTRO TECNOLÓGICO
CURSO DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

JOÃO LUIZ FERNANDES

SHAPE FROM SHADING:
NOVAS ABORDAGENS A PARTIR DE MOVIMENTO E DE FOCO

NITERÓI
2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

JOÃO LUIZ FERNANDES

SHAPE FROM SHADING:
NOVAS ABORDAGENS A PARTIR DE MOVIMENTO E DE FOCO

Tese apresentada ao Curso de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Doutor. Área de Concentração: Computação Visual e Interfaces.

Orientador: Prof. Dr. JOSÉ RICARDO DE ALMEIDA TORREÃO

NITERÓI
2006

RESUMO

Shape from Shading é o processo de Visão Computacional que visa à estimação de forma a partir da gradação suave de intensidade luminosa (*shading*, ou sombreamento) registrada em uma única imagem de uma cena 3-D. Aqui, nós apresentamos duas novas abordagens para o *Shape from Shading*, com a característica comum de representarem extensões, para estimação de forma a partir de uma única entrada, de processos tradicionalmente baseados em múltiplas imagens. Em uma das abordagens, nós mapeamos no *Shape from Shading* o processo de Movimento Fotométrico, que normalmente emprega duas ou mais imagens de uma cena em rotação relativa à câmera. Isso é feito utilizando-se funções de Green de uma equação de casamento, para gerar, a partir da imagem de entrada, uma representação da cena rotacionada, desta forma simulando um par de movimento fotométrico para a imagem dada. Na outra abordagem, o *Shape from Shading* é mapeado a partir do processo de *Shape from Defocus*, que emprega duas ou mais imagens obtidas sob diferentes condições de foco. Neste caso, modelamos a imagem de *shading* como resultando de um processo de borramento, o que vai possibilitar a estimação de um parâmetro de desfocamento similar ao empregado no *Shape from Defocus*, permitindo a reconstrução da superfície representada. A validade destas novas abordagens para o *Shape from Shading* é comprovada pela análise tanto qualitativa quanto quantitativa dos resultados por elas proporcionados.

Palavras-chave: Visão Computacional. *Shape from Shading*. Função de Green. Movimento Fotométrico. *Shape from Defocus*.

ABSTRACT

Shape from Shading is a computer-vision shape estimation process based on the smooth intensity gradient ('shading') registered in a single image of a 3-D scene. Here we introduce two new approaches to Shape from Shading, with the common characteristic of extending, to single-input estimation, processes traditionally based on multiple images. In one of our approaches, Shape from Shading is mapped from the Photometric Motion process, which normally employs two or more images of a rotating scene. This is achieved through the use of the Green's function of a matching equation, in order to generate, from the single input, a representation of the rotated scene, thus simulating the photometric-motion pair to the given image. In the second approach, Shape from Shading is mapped from the Shape from Defocus process, which employs two or more images captured under different focus settings. In this case, we model the shading image as resulting from a blurring process, thus allowing the estimation of a defocus parameter similar to that used in Shape from Defocus, what leads to the reconstruction of the imaged scene. Both our new Shape from Shading approaches have been validated through a thorough qualitative and numerical analysis.

Keywords: Computer Vision. Shape from Shading. Green's Function. Photometric Motion. Shape from Defocus.

	Pág.
Capítulo I – Introdução	1
1.1 – Objetivos e contribuições.....	1
1.2 – Estrutura do trabalho.....	4
Capítulo II – Visão Computacional 3-D	6
2.1 – Introdução.....	6
2.2 – Visão Computacional 3-D.....	7
2.3 – Mapa de reflectância.....	7
2.3.1 – Mapa de reflectância lambertiano.....	10
2.3.2 – Mapa de reflectância especular.....	11
2.3.3 – Linearização do mapa de reflectância.....	13
2.4 – Técnicas para estimação de forma tridimensional.....	14
2.4.1 – Estéreo Fotométrico.....	14
2.4.2 – Estereoscopia.....	17
2.4.3 – <i>Shape from Shading</i> (SFS).....	19
2.4.4 – Estéreo Fotométrico Baseado em Disparidades (EFBD).....	21
2.4.5 – A abordagem da Função de <i>Green</i> para <i>Shape from Shading</i>	23
2.4.6 – Estrutura a partir do Movimento, Fluxo Óptico e Movimento Fotométrico.....	25
2.4.7 – <i>Shape from Defocus</i>	28
Capítulo III – Do Movimento Fotométrico ao <i>Shape from Shading</i>	30
3.1 – Introdução.....	30
3.2 – Nova abordagem para o <i>Movimento Fotométrico</i>	31
3.3 – <i>Green’s Function Photometric Motion</i> (GPM).....	47
3.3.1 – Experimentos.....	52

Capítulo IV – Do <i>Shape from Defocus</i> ao <i>Shape from Shading</i>.....	86
4.1 – Introdução.....	86
4.2 – <i>Shape from Defocus</i> : Abordagem de Pentland.....	88
4.3 – Nova abordagem para <i>Shape from Defocus</i> (SFD).....	90
4.3.1 – Modelagem do borramento por Função <i>Pillbox</i>	94
4.4 – A abordagem para uma única imagem.....	96
4.4.1 - Desenvolvimento do processo.....	96
4.4.2 - Abordagem Multifrequência.....	104
Capítulo V – Considerações Finais.....	138
Referências Bibliográficas.....	144
Publicações Relacionadas.....	152
Anexo I.....	153
Anexo II.....	155

Lista de Figuras

	Pág.
2.1 Esquema de formação e interpretação de imagens.....	6
2.2 Formação da imagem.....	8
2.3 Geometria da reflexão.....	10
2.4 Exemplo de mapa de reflectância.	12
2.5 Direção de iluminação.....	14
2.6 Estéreo Fotométrico – Exemplo.....	16
2.7 Intersecção das curvas de intensidade constante no espaço gradiente	16
2.8 Sistema simplificado de Visão Estéreo.....	18
2.9 Geometria epipolar horizontal para casamento entre pontos correspondentes.....	18
2.10 Reconstrução de Profundidades a partir do EFBD.....	22
2.11 Representação para o mapa de profundidades obtido através do EFBD.....	23
2.12 Reconstrução de Profundidades a partir do GSFS.....	25
2.13 O fluxo óptico não é sempre igual o campo de movimento.....	26
2.14 <i>Shape from Defocus</i> - Exemplo.....	29
3.1 Movimento Fotométrico.....	31
3.2 Movimento Fotométrico: Experimento (<i>Mozart</i>).....	37
3.3 Movimento Fotométrico: Experimento (<i>Mozart</i>).....	38
3.4 Movimento Fotométrico: Experimento (<i>Head1</i>).....	39
3.5 Movimento Fotométrico: Experimento (<i>Head1</i>).....	40
3.6 Movimento Fotométrico: Experimento (<i>Head2</i>).....	41
3.7 Movimento Fotométrico: Experimento (<i>Head2</i>).....	42
3.8 Movimento Fotométrico: Experimento (<i>Angel</i>).....	43
3.9 Movimento Fotométrico: Experimento (<i>Angel</i>).....	44

3.10	Movimento Fotométrico: Experimento (<i>Frog</i>).....	45
3.11	Movimento Fotométrico: Experimento (<i>Frog</i>).....	46
3.12	Resultado da filtragem da Imagem <i>Manequim</i> , de acordo com a equação (3.23).....	49
3.13	Mapa de profundidades estimado, através do GPM, a partir da imagem em 3.12 (a)..	50
3.14	Parâmetro k_0 em função do Erro Médio Gradiente e do Erro em Profundidades.....	57
3.15	Imagens utilizadas para o estudo comparativo experimental do GPM.	58
3.16	GPM: Análise quantitativa de erros.....	59
3.17	GPM. Experimento: Imagem <i>Mask2</i> - Mapas de profundidades.....	61
3.18	GPM. Experimento: Imagem <i>Mask2</i> - Mapas de orientações.....	62
3.19	GPM. Experimento: Imagem <i>Mask2</i> – Re-iluminação das superfícies estimadas.....	63
3.20	GPM. Experimento: Imagem <i>Mask1</i> - Mapas de profundidades.....	64
3.21	GPM. Experimento: Imagem <i>Mask1</i> - Mapas de orientações.....	65
3.22	GPM. Experimento: Imagem <i>Mask1</i> – Re-iluminação das superfícies estimadas.....	66
3.23	GPM. Experimento: Imagem <i>Mozart2</i> - Mapas de profundidades.....	67
3.24	GPM. Experimento: Imagem <i>Mozart2</i> - Mapas de orientações.....	68
3.25	GPM. Experimento: Imagem <i>Mozart2</i> – Re-iluminação das superfícies estimadas....	69
3.26	GPM. Experimento: Imagem <i>Vase</i> - Mapas de profundidades.....	70
3.27	GPM. Experimento: Imagem <i>Vase</i> - Mapas de orientações.....	71
3.28	GPM. Experimento: Imagem <i>Vase</i> – Re-iluminação das superfícies estimadas.....	72
3.29	GPM. Experimento: Imagem <i>Lena</i> - Mapas de profundidades.....	73
3.30	GPM. Experimento: Imagem <i>Lena</i> - Mapas de orientações.....	74
3.31	GPM. Experimento: Imagem <i>Lena</i> – Re-iluminação das superfícies estimadas.....	75
3.32	GPM. Experimento: Imagem <i>Head1</i> - Mapas de profundidades.....	76
3.33	GPM. Experimento: Imagem <i>Head1</i> - Mapas de orientações.....	77
3.34	GPM. Experimento: Imagem <i>Head1</i> – Re-iluminação das superfícies estimadas.....	78
3.35	GPM. Experimento: Imagem <i>Pepper</i> - Mapas de profundidades.....	79
3.36	GPM. Experimento: Imagem <i>Pepper</i> - Mapas de orientações.....	80
3.37	GPM. Experimento: Imagem <i>Pepper</i> – Re-iluminação das superfícies estimadas.....	81

3.38	GPM. Experimento: Imagem <i>Paolina</i>	82
3.39	GPM. Experimento: Imagem <i>Paolina</i>	83
3.40	GPM. Experimento: Imagem <i>Face2</i>	84
3.41	GPM. Experimento: Imagem <i>Face2</i>	85
4.1	Geometria da formação da imagem.....	87
4.2	O ângulo sólido formado pelos raios luminosos que chegam à lente, emanando de um mesmo ponto objeto, decresce com a distância.....	92
4.3	Nova abordagem para o <i>Shape from Defocus</i> : Experimento.....	94
4.4	<i>Shape from Shading</i> via <i>Defocus</i> (Componente DC). Experimento: <i>Lena</i>	101
4.5	<i>Shape from Shading</i> via <i>Defocus</i> (Componente DC). Experimento: <i>Paolina</i>	102
4.6	<i>Shape from Shading</i> via <i>Defocus</i> (Componente DC). Experimento: <i>Face</i>	103
4.7	Análise comparativa dos erros: GPM e SFS via <i>Defocus</i> (componente DC).....	104
4.8	Parte real da Função de Gabor em (4.40).....	105
4.9	Cortes 1-D da função de Gabor em (4.40).....	106
4.10	Cortes 1-D da Transformada da Função de Gabor em (4.41).....	107
4.11	Aplicação da abordagem multifrequência.....	108
4.12	Parâmetro ϕ em função do Erro Médio Gradiente e do Erro em Profundidades.....	110
4.13	Análise quantitativa dos erros: GPM e SFS via <i>Defocus</i>	113
4.14	SFS via <i>Defocus</i> . Experimento: <i>Mozart2</i>	114
4.15	SFS via <i>Defocus</i> . Experimento: <i>Mozart2</i>	115
4.16	SFS via <i>Defocus</i> . Experimento: <i>Lena</i>	116
4.17	SFS via <i>Defocus</i> . Experimento: <i>Lena</i>	117
4.18	SFS via <i>Defocus</i> . Experimento: <i>Paolina</i>	118
4.19	SFS via <i>Defocus</i> . Experimento: <i>Paolina</i>	119
4.20	SFS via <i>Defocus</i> . Experimento: <i>Pepper</i>	120
4.21	SFS via <i>Defocus</i> . Experimento: <i>Pepper</i>	121
4.22	SFS via <i>Defocus</i> . Experimento: <i>Face2</i>	122
4.23	SFS via <i>Defocus</i> . Experimento: <i>Face2</i>	123
4.24	SFS via <i>Defocus</i> . Experimento: <i>Mask1</i>	124

4.25	SFS via <i>Defocus</i> . Experimento: <i>Mask1</i>	125
4.26	SFS via <i>Defocus</i> . Experimento: <i>Mask2</i>	126
4.27	SFS via <i>Defocus</i> . Experimento: <i>Mask2</i>	127
4.28	SFS via <i>Defocus</i> . Experimento: <i>Head1</i>	128
4.29	SFS via <i>Defocus</i> . Experimento: <i>Head1</i>	129
4.30	SFS via <i>Defocus</i> . Experimento: <i>Head2</i>	130
4.31	SFS via <i>Defocus</i> . Experimento: <i>Head2</i>	131
4.32	SFS via <i>Defocus</i> . Experimento: <i>Vase2</i>	132
4.33	SFS via <i>Defocus</i> . Experimento: <i>Vase2</i>	133
4.34	SFS via <i>Defocus</i> . Experimento: <i>Manequim</i>	134
4.35	SFS via <i>Defocus</i> . Experimento: <i>Manequim</i>	135
4.36	SFS via <i>Defocus</i> . Experimento: <i>Penny</i>	136
4.37	SFS via <i>Defocus</i> . Experimento: <i>Penny</i>	137
5.1	Imagens de entrada utilizadas para avaliação quantitativa do erro.....	140
5.2	Análise quantitativa dos erros para as imagens da Figura 5.1.....	141
5.3	Paralelo entre nossa abordagem do SFS via <i>Defocus</i> e uma Rede Neural Artificial....	143

Lista de Siglas

Sigla	<i>Descrição</i>
SFS	Shape from Shading.
EFBD	Estéreo Fotométrico Baseado em Disparidades.
GSFS	<i>Green's Function Shape From Shading.</i>
GPM	<i>Green's Function Photometric Motion.</i>
SFD	<i>Shape from Defocus.</i>
SFS via <i>Defocus</i>	<i>Shape from Shading via Defocus</i>
PSF	Point Spread Function.

1.1 – Objetivos e contribuições

É inegável a importância da visão para o ser humano. Através dela o homem é capaz de reconhecer o seu ambiente e com ele interagir, podendo realizar uma série de tarefas inteligentes para as quais a informação visual é imprescindível. Assim, é fácil perceber o quanto é desejável prover também a máquina do sentido da visão, isto é, torná-la capaz de, a partir de imagens, analisar e interpretar as cenas representadas, permitindo-lhe lidar autonomamente com situações que hoje ainda requerem a intervenção humana. Com esse direcionamento caminham as pesquisas em Visão Computacional, cujo objetivo é o desenvolvimento de sistemas capazes de analisar imagens e produzir uma interpretação da situação representada, a partir das informações geométricas, topológicas ou físicas obtidas. Assim, se pudéssemos resumir o objetivo da Visão Computacional em uma única frase, diríamos ser este o de dotar a máquina da capacidade de *ver*, ou, melhor ainda, de emular a visão humana em toda a sua potencialidade.

A Visão Computacional, embora seja uma disciplina relativamente recente (seus primeiros desenvolvimentos datam da década de 60), representa um vasto campo de pesquisas. As aplicações nessa área são muito variadas, e vão desde a análise automática para o diagnóstico a partir de imagens médicas, até à navegação robótica em ambientes tridimensionais. No primeiro caso, dizemos tratar-se de Visão 2-D, onde a informação de profundidade das cenas não é relevante para a execução das tarefas propostas. Já no segundo caso, domínio da Visão 3-D, precisamos trabalhar em três dimensões, tendo em vista os objetivos impostos, que envolvem a recuperação da estrutura tridimensional da cena, intrinsecamente registrada em suas imagens. É no campo da Visão 3-D que concentramos o trabalho aqui reportado.

Muitos são os processos utilizados em Visão 3-D com o objetivo de extrair informações a partir de imagens. Esses processos visam, por exemplo, à obtenção de mapas de profundidades, orientação ou curvatura para as superfícies representadas, ou à estimação de velocidades em cenas dinâmicas, por exemplo. Alguns desses processos são *binoculares* –

isto é, requerem imagens capturadas de dois pontos de vista diferentes -, enquanto outros são *monoculares*, utilizando imagens obtidas a partir de uma única posição de observação.

A abordagem conhecida como *Shape from Shading* (SFS) [1] lida com uma única imagem de entrada, e objetiva estimar forma tridimensional a partir da gradação suave de intensidade luminosa registrada (o que se chama de *shading*, ou sombreamento). Uma extensão do SFS é o *Estéreo Fotométrico* [2], que utiliza duas ou mais imagens de *shading* monoculares, mas obtidas sob condições de iluminação diferentes. Já a *Estereoscopia* [3] baseia-se no emprego de duas ou mais imagens da cena obtidas através de câmeras localizadas em posições diferentes. Um mapa de deslocamentos relativos, ou disparidades, pode ser obtido quando se estabelece a correspondência entre pares de pontos, um em cada imagem, que representam o mesmo ponto na cena. Por um procedimento relativamente simples, pode-se recuperar informação de profundidade a partir dos mapas de disparidade. Ao contrário do SFS e do *Estéreo Fotométrico*, que se baseiam apenas nos aspectos *fotométricos* da formação das imagens, a *Estereoscopia* é um processo *geométrico*, pois utiliza informação de deslocamento dos pontos na cena.

Em nossa Dissertação de Mestrado [4], nós estudamos uma abordagem alternativa ao *Estéreo Fotométrico* tradicional, conhecida como *Estéreo Fotométrico Baseado em Disparidades* (EFBD), em que pares de imagens de estéreo fotométrico são casados como se fossem imagens estereoscópicas, para produzir um mapa de disparidades *fotométricas* (i.e., ligadas ao deslocamento de intensidades) que permite a reconstrução em profundidade a partir de imagens monoculares .

Este trabalho foi posteriormente estendido em [5], para o caso de uma única imagem de entrada, no processo denominado *Green's Function Shape From Shading* (GSFS). Nessa abordagem, utilizaram-se funções de Green de equações de casamento, de modo a gerar pares virtuais de estéreo fotométrico para a imagem dada, permitindo deste modo tratar um processo baseado em uma única imagem como se este contasse com múltiplas entradas.

Um outro processo envolvendo simultaneamente os aspectos geométricos e fotométricos da formação de imagens é o chamado *Movimento Fotométrico*, que foi desenvolvido originalmente por Pentland [6] e que difere fundamentalmente dos anteriores, pelo fato de empregar seqüências de imagens de cenas **dinâmicas**, para a estimação de forma.

A contribuição inicial do nosso trabalho de Tese, a ser reportada aqui, diz respeito à introdução de uma formulação alternativa para o *Movimento Fotométrico* [7], que, em muitos aspectos, segue um paralelo com a formulação alternativa para o *Estéreo Fotométrico* (o EFBD) introduzida na nossa Dissertação de Mestrado. A seguir, numa trajetória semelhante à que levou do *Estéreo Fotométrico Baseado em Disparidades* ao *Green's function shape from*

shading, nós demonstramos que a nossa formulação para o *Movimento Fotométrico* pode ser estendida para a estimação de forma a partir de **uma única** imagem de entrada, também com o uso de funções de Green de equações de casamento, através do processo chamado de *Green's Function Photometric Motion* (GPM) [8]. Desta forma, nós efetivamente chegamos a uma nova abordagem para o *Shape from Shading*, através do *Movimento Fotométrico*.

Por outro lado, a abordagem conhecida como *Shape from Defocus* [9] utiliza o desfocamento local das imagens para inferir forma 3-D, baseando-se no fato de que as condições inerentes aos dispositivos de captura fazem com que os pontos em uma cena, localizados a diferentes distâncias da câmera, sejam representados com diferentes graus de borramento. Quando imagens capturadas sob diferentes condições de abertura da lente estão disponíveis, é possível estimar uma medida de desfocamento e, a partir daí, obter uma representação tridimensional da cena.

Uma forma usual de modelar o processo de desfocamento é através da convolução da imagem, em janelas locais, com uma função Gaussiana, o que permite a estimação da medida de desfocamento (identificada com o desvio-padrão da Gaussiana) no domínio de Fourier. Neste contexto, uma outra contribuição do nosso trabalho de Tese é a demonstração de que o uso de Gaussianas *não normalizadas* - ao invés das normalizadas usualmente empregadas - pode ser vantajoso para o *Shape from Defocus*, já que permite a estimação de profundidades a partir da componente DC (frequência zero) das imagens, resultando em mapas de profundidade densos, e evitando também o efeito espúrio das altas frequências artificialmente introduzidas pelo uso de janelas. Mais importante ainda, empregando este novo modelo para o desfocamento, nós demonstramos que o processo de *Shape from Defocus* pode ser estendido para a estimação de forma a partir de **uma única** imagem de entrada, deste modo nos levando a uma segunda abordagem alternativa para o *Shape from Shading*, agora a partir de *defocus*.

As duas novas abordagens para o SFS, aqui apresentadas, provaram-se competitivas com propostas alternativas em uma série extensiva de experimentos, sendo que a abordagem baseada em *defocus* mostrou-se em geral superior ao GPM. Ambas as propostas são *unificadoras*, no sentido de que abordam o SFS a partir de técnicas e conceitos associados a outros processos de visão computacional. Do ponto de vista da *economia*, uma tal *unificação* dos processos constitui uma característica altamente desejável para os sistemas de visão. O SFS baseado em *defocus*, em particular, propõe uma inovação radical neste aspecto, ao modelar a formação de imagens de *shading* como um processo de desfocamento, desta forma prescindindo do conceito de *mapa de reflectância*, essencial para as abordagens tradicionais do *Shape from Shading*.

1.2. Estrutura do trabalho

O restante deste Trabalho apresenta-se organizado da seguinte forma:

- **Capítulo 2. Visão Computacional 3-D.**

Nesse capítulo, apresentamos conceitos básicos de Visão Computacional, e introduzimos tópicos em Visão 3-D indispensáveis para a compreensão do trabalho realizado. Em particular, nós descrevemos, sumariamente, certos processos para a estimação de forma 3-D que estão relacionados, direta ou indiretamente, com o nosso trabalho de Tese.

- **Capítulo 3. Do Movimento Fotométrico ao *Shape from Shading*.**

Aqui, introduzimos nossa nova formulação para o *Movimento Fotométrico*, cuja característica distintiva é a de ser baseada na variação de intensidades, devido ao movimento, em um ponto fixo no plano da imagem, e não, como na abordagem original do processo, em um dado ponto na superfície em movimento. Em seguida, nós mostramos como essa formulação pode ser estendida para o caso de reconstruções a partir de uma única imagem, através do uso das funções de Green de uma certa equação de casamento, numa nova abordagem denominada *Green's Function Photometric Motion*.

- **Capítulo 4. Do *Shape from Defocus* ao *Shape from Shading*.**

Nesse capítulo, apresentamos considerações gerais sobre o *Shape from Defocus*, e a nossa versão alternativa para o processo original introduzido por Pentland, desenvolvendo uma nova formulação que evita os problemas identificados neste último e que ainda permite a estimação de mapas densos de profundidade. Em seguida, nós estendemos essa nova abordagem para o caso da estimação de forma a partir de uma única imagem, introduzindo mais um algoritmo para o *Shape from Shading*, agora baseado em informações de desfocamento.

- **Capítulo 5. Considerações finais.**

Nesse capítulo, nós discutimos as contribuições do presente trabalho. Apresentamos, ainda, considerações finais sobre a pesquisa desenvolvida, sugerindo possibilidades de extensão em trabalhos futuros.

2.1. Visão Computacional

O alto custo, a limitada eficiência, o risco inerente às condições de operação e uma série de outras restrições associadas ao emprego de seres humanos em muitas atividades que requerem a análise, a interpretação e o tratamento de informações visuais são estímulos mais do que suficientes para o desenvolvimento de sistemas de visão autônomos, em que a precisão, a rapidez e a consistência de resposta da máquina funcionam como fatores decisivos no sentido de possibilitar ganhos substanciais de produtividade. Tais sistemas, que interpretam as informações obtidas a partir de imagens, são o objeto das pesquisas em Visão Computacional, e já vêm sendo utilizados em aplicações cada vez mais sofisticadas. (Figura 2.1) [10].

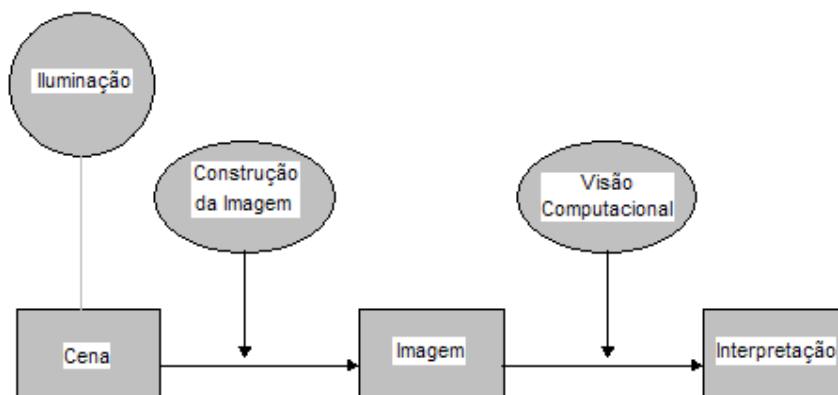


Figura 2.1. Esquema de formação e interpretação de imagens

A Visão Computacional, embora seja uma disciplina relativamente recente (seus primeiros desenvolvimentos datam da década de 60), representa um vasto campo de pesquisas. Neste trabalho, nós consideramos a subárea da Visão Computacional cujo objetivo é o de extrair, a partir das imagens, informações que viabilizem as atividades de reconhecimento, localização e navegação num ambiente tridimensional. Denotaremos tal subárea como Visão Computacional 3-D.

2.2. Visão Computacional 3-D

O termo Visão Computacional 3-D é normalmente empregado quando as informações registradas na imagem devem ser interpretadas como cenas tridimensionais. Assim, por exemplo, robôs móveis podem utilizar as informações de forma tridimensional, recuperadas a partir de imagens, com a finalidade de ultrapassar obstáculos ou ainda para manipular artefatos perigosos em uma linha de montagem.

Modernamente, podemos encontrar outras aplicações da Visão 3-D. A demanda por sistemas de Realidade Virtual com alta qualidade, e a inabilidade das técnicas padrões da Computação Gráfica em produzi-los a baixo custo, são razões suficientes para se aliar técnicas de Visão Computacional à síntese de imagens. Por exemplo, um problema substancial em Realidade Virtual é o de construir modelos de cenas complexas, como as existentes em cidades. Uma solução potencial é utilizar técnicas de Visão Computacional na recuperação de informações através de fotografias; assim, para se construir um modelo geométrico correto de um edifício com muitos detalhes, podemos utilizar várias fotografias do mesmo, e implementar algoritmos de extração de profundidades para a obtenção dos dados [11].

No presente trabalho, o nosso objetivo foi o de estudar a estimação de forma tridimensional a partir de imagens monoculares, ou seja, a recuperação da estrutura de uma cena a partir de imagens obtidas de um único ponto de vista. Com o intuito de proporcionar uma base sólida para a compreensão dos métodos que apresentaremos, começaremos por introduzir alguns tópicos importantes em visão 3-D.

2.3. Mapa de reflectância

Como desejamos extrair, a partir das intensidades de uma imagem, informações que nos permitam reconstruir a cena representada, devemos estabelecer o relacionamento existente entre estes dois domínios: imagem e cena. A imagem obtida de uma certa cena vai depender, evidentemente, da forma dos objetos ali presentes, de suas propriedades reflexivas, da posição de captura da imagem, e das fontes de iluminação utilizadas [10].

No processo de formação da imagem, a luz emitida pelas fontes de iluminação é refletida pelos objetos da cena e atinge os elementos sensores da câmera, gerando um registro que é, em cada ponto, proporcional à luz recebida do ponto correspondente na cena. Para sermos mais explícitos, vejamos alguns conceitos de radiometria relevantes para a nossa análise: chamamos de **Irradiância**, à quantidade de luz que chega a uma superfície, isto é, à potência luminosa por unidade de área (W/m^2 - Watts por metro quadrado) incidente na mesma. Já a potência luminosa emitida por uma superfície, por unidade de área e também por unidade de ângulo sólido ($\text{W}/\text{m}^2 \cdot \text{sr}$ - watts por metro quadrado por steradiano), é chamada de **Radiância**. O que acabamos de enunciar acerca do processo de formação da imagem pode ser resumido na afirmação de que *a irradiância da imagem é proporcional à radiância da cena* [10].

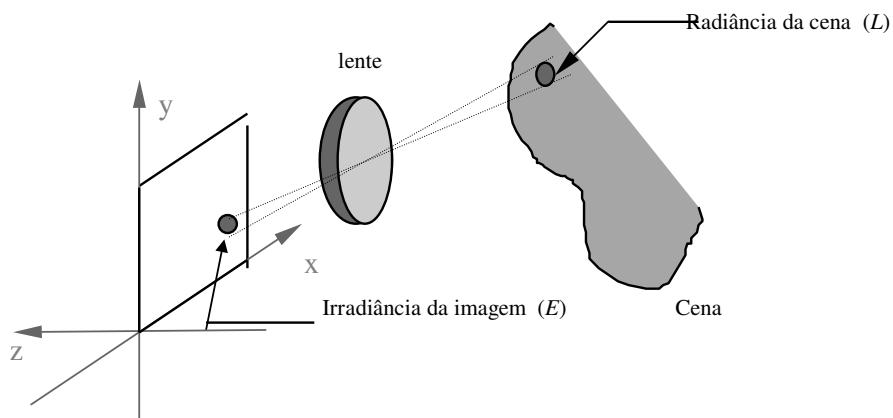


Figura 2.2. Formação da imagem

Se denotarmos por E a irradiância da imagem e por L a radiância da cena, temos então que, para regiões correspondentes (ver Figura 2.2):

$$E \propto L \quad (2.1)$$

Por outro lado, a radiância de cada ponto da cena depende da quantidade de luz incidente no mesmo e da fração desta luz que é refletida para a câmera. Essa dependência pode ser modelada através da **função de reflectância**, que relaciona a radiância de uma superfície à sua orientação local, às suas propriedades reflexivas e às condições de sua observação. Explicitamente, temos:

$$L = R(\hat{n}(x, y), \hat{s}, \hat{v}, \rho(x, y)) \quad (2.2)$$

onde R é o mapa de reflectância, \hat{n} é o vetor unitário normal à superfície, \hat{s} e \hat{v} são os vetores unitários que representam as direções de iluminação e observação e $\rho(x, y)$ é o albedo ou reflectividade do material da superfície.

Identificando a irradiância da imagem à intensidade registrada em cada ponto, obtemos então, a partir de (2.1) e (2.2), a relação:

$$I(x, y) = R(\hat{n}(x, y), \hat{s}, \hat{v}, \rho(x, y)) \quad (2.3)$$

Essa relação é conhecida como **Equação de Irradiância da Imagem**.

É importante observar que, em (2.3), consideramos que a fonte de luz, o objeto na cena e o observador estão distantes entre si. Assim, assumimos que a imagem representa uma projeção ortográfica da cena (eliminando o efeito de perspectiva), e também que \hat{s} e \hat{v} são vetores independentes da posição. Por outro lado, o albedo, $\rho(x, y)$ é um parâmetro dependente do material da superfície e que, em princípio, varia com a posição. Ele é tomado como um número entre zero e um, que indica a quantidade de luz que a superfície reflete em comparação com uma superfície ideal, com a mesma dependência geométrica [12]. Por simplicidade, passaremos a denotar o albedo simplesmente por ρ , assumindo-o uniforme ao longo da cena.

Uma representação alternativa para o mapa de reflectância pode ser dada em termos do gradiente da superfície. Se considerarmos que a superfície em questão pode ser representada como $Z(x, y)$, temos que $p = \partial Z / \partial x$ e $q = \partial Z / \partial y$ são as componentes do gradiente da superfície, significando as inclinações de $Z(x, y)$, nas direções x e y . Conforme demonstrado em [10], a direção perpendicular à superfície em um determinado ponto é dada pelo vetor $(-p, -q, 1)$. Dessa forma, podemos escrever o vetor \hat{n} (unitário e normal à superfície), em termos de p e q , como:

$$\hat{n} = (n_x, n_y, n_z) = \frac{(-p, -q, 1)}{\sqrt{1 + p^2 + q^2}} \quad (2.4)$$

sendo n_x , n_y e n_z as componentes da normal nas direções x , y e z , respectivamente.

Em termos de p e q , a equação (2.3) seria dada por:

$$I(x, y) = R(p(x, y), q(x, y), \hat{s}, \hat{v}, \rho(x, y))$$

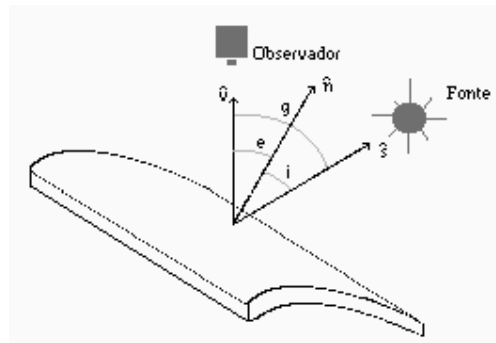


Figura 2.3. Geometria da reflexão. Ângulo de incidência “ i ” (entre as direções de iluminação e normal à superfície); ângulo de emissão “ e ” (entre as direções de observação e normal); ângulo fase “ g ” (entre as direções de observação e iluminação).

Vejamos agora as formas comumente usadas para o mapa de reflectância.

2.3.1. Mapa de reflectância lambertiano.

A superfície lambertiana ideal (ou difusora perfeita) é aquela em que a luz incidente é totalmente refletida, e de uma forma tal que a superfície aparece igualmente luminosa de qualquer ponto de vista. Por esta definição, podemos perceber que a reflexão lambertiana independe da direção de observação, e a intensidade da luz refletida depende apenas da direção da luz incidente. Para uma única fonte de iluminação, o mapa de reflectância lambertiano tem a forma:

$$R(\hat{n}, \hat{s}, \rho) = \begin{cases} \rho(\hat{n} \cdot \hat{s}), & \text{se } \hat{n} \cdot \hat{s} > 0 \\ 0, & \text{senão} \end{cases} \quad (2.5)$$

O produto $(\hat{n} \cdot \hat{s})$ é igual ao co-seno do ângulo de incidência (ver Figura 2.3). Desse modo, podemos escrever também

$$R = \rho \cos(i) \quad (2.6)$$

Como valores negativos para R não são definidos, a expressão acima vale para $0 \leq i \leq \pi/2$. Para ângulos de incidência que extrapolem $\pi/2$, R é igual a 0.

A equação (2.5) também pode ser escrita em termos do gradiente (p,q) , ou seja:

$$R(p,q) = \rho \left[\frac{1 + pp_s + qq_s}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}} \right] \quad (2.7)$$

onde expressamos \hat{n} de acordo com a equação (2.4), e denotamos a direção de iluminação pelo vetor unitário

$$\hat{s} = \frac{(-p_s, -q_s, 1)}{\sqrt{1 + p_s^2 + q_s^2}} \quad (2.8)$$

de forma que $(-p_s, -q_s)$ indica o gradiente de uma superfície perpendicular à direção de iluminação.

Ao espaço de todas as possíveis orientações (p,q) , denominamos espaço gradiente. É usual representarmos o mapa de reflectância no espaço gradiente como uma série de contornos para os quais $R(p,q) = \text{constante}$ (Figura 2.4).

2.3.2. Mapa de reflectância especular.

A - Superfície especular perfeita

Numa superfície perfeitamente especular, a luz incidente se reflete numa única direção, a direção especular perfeita, para a qual o ângulo de incidência é igual ao ângulo de emissão, e o vetor normal (\hat{n}) , o vetor de iluminação (\hat{s}) e o vetor de observação (\hat{v}) encontram-se todos no mesmo plano. Dessa forma, a reflectância é máxima para uma direção de observação igual à especular perfeita e nula nos outros casos, ou seja:

$$R(\hat{n}, \hat{s}, \hat{v}) = \begin{cases} 1, & \text{se } \hat{s} + \hat{v} = 2\hat{n} \\ 0, & \text{senão} \end{cases}$$

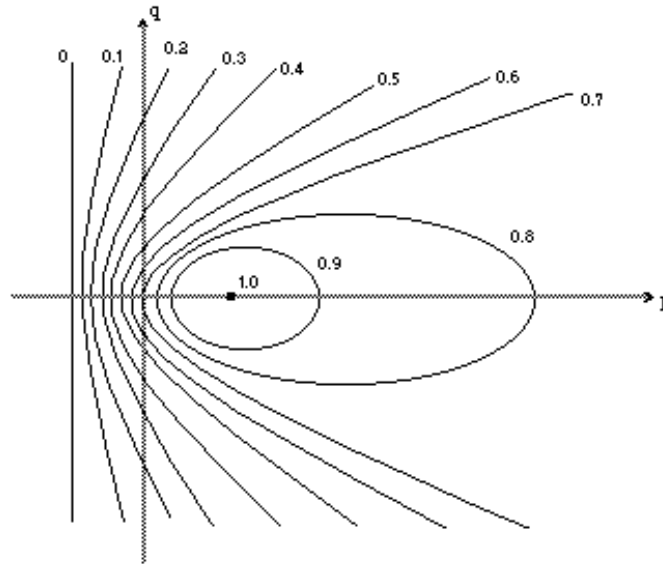


Figura 2.4. Exemplo de mapa de reflectância para uma superfície lambertiana, com direção de iluminação dada por $(p_s, q_s) = (1, 0)$. O número ao lado das curvas indica o valor de reflectância para as direções (p, q) correspondentes

B - Mapa de reflectância quase-especular.

Superfícies quase-especulares admitem reflexão num cone de direções em torno da direção especular perfeita. O seu mapa de reflectância pode ser modelado pela expressão:

$$R = \rho[\hat{v}_s \cdot \hat{v}]^m = \rho[2(\hat{n} \cdot \hat{s})(\hat{n} \cdot \hat{v}) - (\hat{v} \cdot \hat{s})]^m \quad (2.9)$$

onde \hat{v}_s é a direção de reflexão especular perfeita, $\hat{v}_s = 2\hat{n} - \hat{s}$. Notemos que $(\hat{v}_s \cdot \hat{v})$ representa o co-seno do ângulo entre a direção de observação \hat{v} e a direção \hat{v}_s ; assim, a reflectância máxima é obtida para $\hat{v} = \hat{v}_s$. A constante m modela a rapidez com que a reflexão quase-especular decai, à medida que \hat{v} se afasta da direção especular perfeita [3].

A relação (2.9) também pode ser escrita em termos dos ângulos de incidência, emissão e fase como:

$$R = \rho [2(\cos i) \cdot (\cos e) - (\cos g)]^m \quad (2.10)$$

C - Superfícies Lustrosas

A reflectância da maioria das superfícies pode ser caracterizada como uma composição de componentes especulares e difusas. Uma forma de mapa de reflectância geralmente utilizada para modelar esse tipo de comportamento é [12]:

$$R = \rho \left[(1-k)(\hat{n} \cdot \hat{s}) + k \frac{1}{2}(m+1)(2(\hat{n} \cdot \hat{s}) \cdot (\hat{n} \cdot \hat{v}) - (\hat{v} \cdot \hat{s}))^m \right] \quad (2.11)$$

onde a constante k , que pode assumir valores entre 0 e 1, vai determinar a influência relativa das componentes de reflexão difusa $(\hat{n} \cdot \hat{s})$ e quase-especular $(\frac{1}{2}(m+1)(2(\hat{n} \cdot \hat{s}) \cdot (\hat{n} \cdot \hat{v}) - (\hat{v} \cdot \hat{s}))^m)$.

2.3.3. Linearização do mapa de reflectância

Sob determinadas condições, podemos obter uma boa representação para o mapa de reflectância de uma superfície suave, através de uma aproximação linear. Tal aproximação vai permitir a definição de algoritmos simplificados para a estimação de forma tridimensional, conforme veremos adiante.

Para uma região da superfície com orientação média local (p_0, q_0) , uma aproximação linear para o mapa de reflectância pode ser obtida através da expansão em série de Taylor, em torno de (p_0, q_0) [13]. Dessa forma, a equação de irradiância da imagem pode ser escrita como

$$I(x, y) = R(p_0, q_0) + (p - p_0) \left. \frac{\partial R(p, q)}{\partial p} \right|_{p_0, q_0} + (q - q_0) \left. \frac{\partial R(p, q)}{\partial q} \right|_{p_0, q_0} \quad (2.12)$$

A equação (2.12) vai representar uma aproximação razoavelmente precisa do verdadeiro mapa de reflectância, para uma pequena faixa de valores de p e q em torno de (p_0, q_0) . Nota-se que, quanto mais oblíqua é a direção da iluminação sob a qual se observa a superfície, maiores se tornam os limites dessa faixa, e é possível afirmar que a qualidade da

aproximação linear obtida é diretamente proporcional à inclinação do vetor iluminante medida pelo ângulo σ na Figura 2.5 [13].

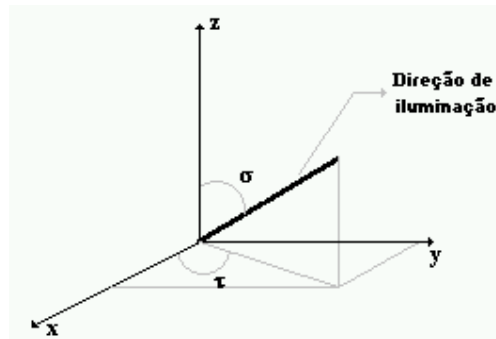


Figura 2.5. Direção de iluminação, representada pelos ângulos polares σ e τ .

2.4. Técnicas para estimação de forma tridimensional

Apresentaremos, agora, alguns processos comumente utilizados em Visão Computacional para a extração de informação de forma tridimensional.

2.4.1. Estéreo Fotométrico

O Estéreo Fotométrico é um processo de grande aplicabilidade prática em Visão Computacional, introduzido por Robert Woodham [14] em 1978. Ele consiste, basicamente, no emprego de duas ou mais imagens monoculares da cena, para, a partir da gradação de intensidade luminosa percebida (*shading*), estimar forma 3-D. As chamadas imagens de Estéreo Fotométrico são capturadas a partir de uma posição fixa, mas sob diferentes condições de iluminação da cena.

O objetivo do Estéreo Fotométrico é, a partir do conhecimento do mapa de reflectância das superfícies, extrair das suas imagens uma estimativa da orientação local em cada ponto observado. O mapa de orientações obtido pode ser representado graficamente através de um diagrama de agulhas, como o mostrado na Figura 2.6, que representa a superfície na cena por uma malha regular de pequenos vetores normais a cada um dos seus pontos. Cada agulha na Figura 2.6 é, portanto, a projeção, no plano da imagem, do vetor normal à superfície. O mapa de agulhas (como é conhecido esse diagrama) é uma representação da forma da superfície

observada, e pode ser usado na determinação de sua atitude espacial na cena [10]. No que se segue, vamos equacionar o problema do Estéreo Fotométrico.

Como procuramos a orientação em cada ponto, o problema resume-se a resolver as equações de irradiância da imagem

$$I_i(x, y) = R_i(p, q), \quad i = 1, 2, \dots \quad (2.13)$$

para as componentes p e q do gradiente da superfície em cada ponto.

A necessidade de se trabalhar com pelo menos duas imagens da cena, na ausência de informações adicionais, decorre do fato de que a equação de irradiância da imagem representa uma única restrição para os dois graus de liberdade, p e q , da orientação superficial. Analisada no espaço gradiente, por exemplo, a equação (2.13) descreve, para cada ponto (x, y) da imagem, uma linha como as representadas na Figura 2.7, já que, para uma dada intensidade luminosa, há uma infinidade de orientações (p, q) compatíveis, na superfície observada, definindo um contorno de intensidade constante. Apenas com a informação obtida a partir de uma segunda imagem é possível restringir a gama de possíveis orientações e, muitas vezes, dependendo da forma de $R(p, q)$, mesmo uma terceira imagem pode ser necessária, como ilustrado na Figura 2.7. Se denotarmos por $R1$ e $R2$ dois contornos no espaço gradiente, correspondentes às intensidades obtidas para as direções de iluminação (p_1, q_1) e (p_2, q_2) em um mesmo ponto (x, y) das imagens, percebemos que, pela geometria das curvas, há dois pares de orientações (p, q) possíveis para o ponto correspondente na cena. Com a inclusão de uma terceira imagem (direção de iluminação (p_3, q_3)), a intersecção da curva $R3$ com as outras duas vai caracterizar a orientação correta naquele ponto.

Há uma vantagem adicional na utilização de três imagens no Estéreo Fotométrico. Conforme vimos anteriormente, para uma superfície com propriedades de reflectância uniformes, a intensidade na imagem será o produto do albedo, ρ , com uma função da orientação local da superfície. A utilização da terceira imagem, pode permitir, portanto, não apenas a estimação do mapa de orientações da superfície, mas também do parâmetro ρ , conforme demonstrado em [10].

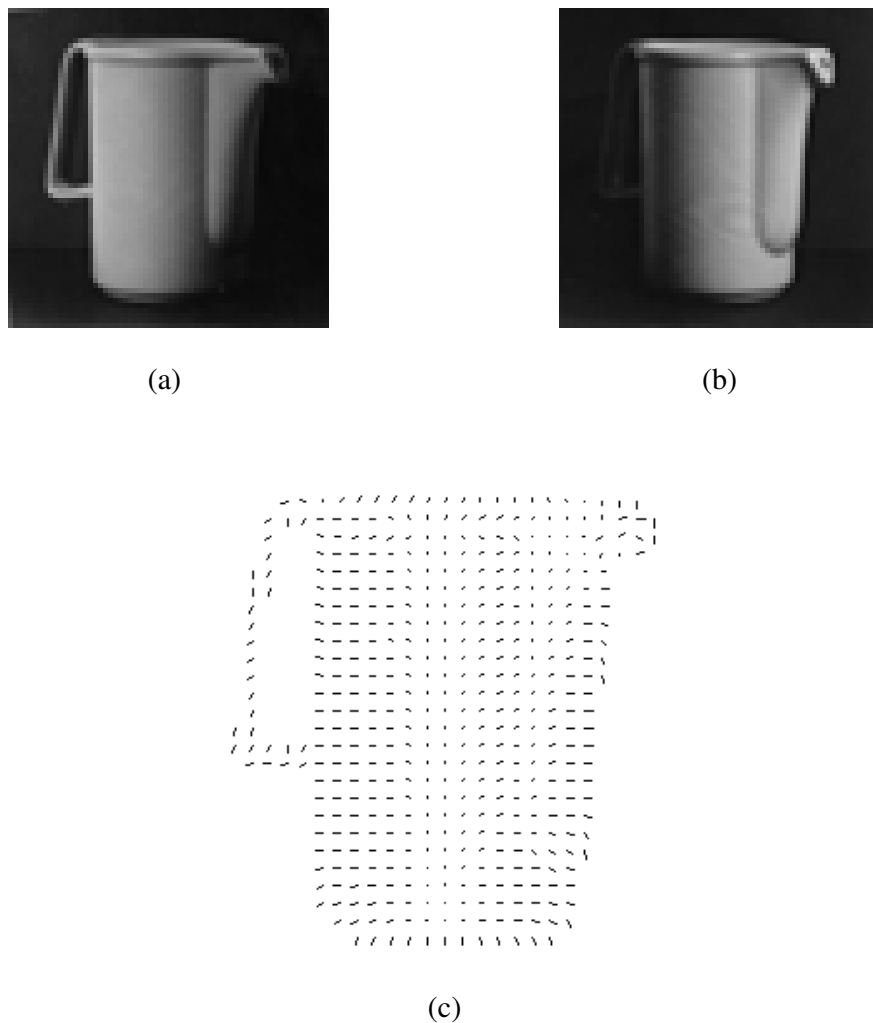


Figura 2.6. (a) e (b) Exemplo de imagens de Estéreo Fotométrico (c) Mapa de agulhas da superfície.

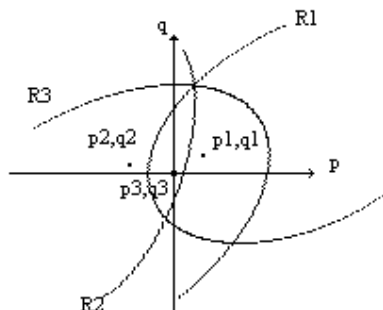


Figura 2.7. Intersecção das curvas de intensidade constante no espaço gradiente.

O Estéreo Fotométrico, como descrito, é um processo que requer condições extremamente controladas para sua aplicação. Além disso, o método pode não funcionar

bem para cenas que possuam objetos de reflectâncias diferentes (superfícies difusas e especulares, por exemplo).

2.4.2. Estereoscopia

Diferentemente do Estéreo Fotométrico, a Estereoscopia utiliza câmeras em posições diferentes, para a obtenção de suas imagens. Considerando um par de imagens da cena, podemos perceber, através da Figura (2.8), que um dado ponto da superfície observada vai se projetar em posições distintas nas duas imagens. Fazendo o “casamento” entre as imagens, é possível determinar, para cada ponto, o deslocamento acarretado pela mudança na posição de observação, ou seja, é possível obter o mapa de disparidades estereoscópicas. Esse mapa vai possibilitar o resgate das profundidades na cena, já que as disparidades estão relacionadas às posições tridimensionais dos pontos.

A Estereoscopia, também conhecida como Visão Estéreo ou Visão Binocular [3] é um processo muito importante, tanto em Visão Computacional como em visão biológica, já que pode ser empregado sob uma grande variedade de condições, além de não requerer equipamentos especiais. No entanto, o casamento estereoscópico não é simples de ser implementado e tende a ser custoso computacionalmente, o que tem até hoje impedido o surgimento de sistemas estereoscópicos eficientes.

Para tornar o processo de correspondência estereoscópica menos árduo, observa-se que há uma relação entre as projeções de um mesmo ponto da cena nos planos das duas imagens. Tal relação vai determinar que o casamento de um ponto (x_e, y_e) na imagem da esquerda somente se possa dar com pontos localizados ao longo de uma linha reta na imagem da direita. Essa reta, chamada de **linha epipolar**, é a projeção, no plano da segunda imagem, da linha que passa pelo ponto observado e pelo centro da lente da primeira câmera [12]. Se as câmeras forem posicionadas paralelamente e a uma mesma altura, a linha epipolar é uma reta horizontal, de mesma coordenada y_e que o ponto na imagem da esquerda (ver Figura 2.9).

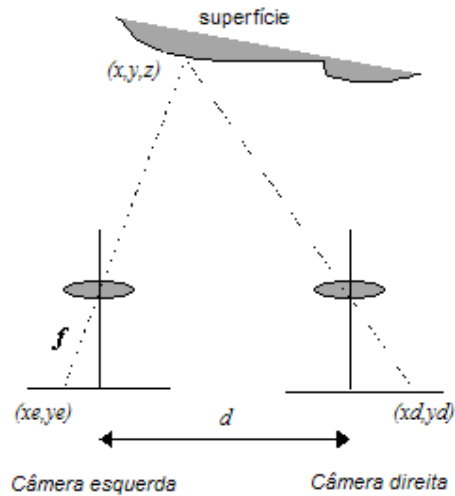


Figura 2.8. Sistema simplificado de Visão Estéreo. O ponto (x,y,z) da superfície é projetado em pontos de coordenadas diferentes nas duas imagens, (x_e,y_e) e (x_d,y_d) . “ f ” é a distância focal e “ d ” é a distância entre as câmeras.

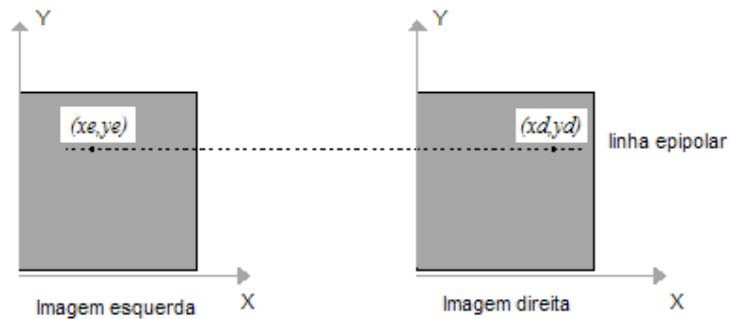


Figura 2.9. Geometria epipolar horizontal para casamento entre pontos correspondentes. Na Figura, a disparidade é calculada por $(x_d - x_e)$.

Para essa geometria de imageamento, é possível mostrar que a profundidade em cada ponto da cena pode ser relacionada ao mapa de disparidades estereoscópicas pela expressão:

$$Z(x,y) = \frac{f \cdot d}{D(x,y)} \quad (2.14)$$

onde f é a distância focal, d é a distância entre as câmeras e $D(x,y)$ é a disparidade correspondente ao ponto (x,y) na cena.

Correspondência Estereoscópica¹

Como podemos perceber, a implementação do casamento entre as duas imagens está no cerne do método de Visão Estéreo. Algumas técnicas utilizadas para esse fim são as seguintes:

A - Estéreo baseado em área.

Essa abordagem é baseada na correlação entre regiões nas duas imagens. A partir da escolha de uma região na primeira imagem, o procedimento é efetuado pela procura da região de melhor correspondência na outra, utilizando-se para isso alguma medida de similaridade, como por exemplo, a soma das diferenças absolutas entre as intensidades.

B - Estéreo baseado em características

Nesse processo, o casamento é realizado entre elementos das imagens que representam características extraídas da cena, como, por exemplo, as suas bordas. Como a correspondência é feita com um conjunto mais restrito de dados, esse método é mais rápido do que o baseado em áreas, tendendo a produzir, no entanto, mapas de disparidades esparsos, que obrigam à posterior utilização de alguma modalidade de interpolação [3], [15].

2.4.3 – *Shape from Shading* (SFS)

O processo de estimação de forma 3-D, a partir da gradação de intensidade luminosa percebida em uma única imagem sem textura, denomina-se *Shape from Shading* (SFS). Tradicionalmente, a estimação de forma através de *shading* tem sido baseada somente nos aspectos fotométricos do processo de formação das imagens, expressos através da função mapa de reflectância. Como vimos anteriormente, o mapa de reflectância incorpora as dependências fotométricas do SFS, ou seja, relaciona a irradiância de cada ponto da imagem ao gradiente no local correspondente na cena, através da equação de irradiância da imagem,

$$I(x, y) = R(p, q) \quad (2.15)$$

¹ As diversas técnicas para realização do casamento estereoscópico podem ser melhor apreciadas em [3], [15], [16], [17] e [18].

onde $I(x,y)$ denota a intensidade registrada no ponto (x,y) da única imagem de entrada e R é a função mapa de reflectância .

Pode-se, portanto, formular o problema do SFS como o de inverter a Equação (2.15) no sentido de obter-se o mapa de orientações da superfície observada. No entanto, é fácil notar que, diferentemente do Estéreo Fotométrico, nós temos agora, para cada ponto da imagem, duas incógnitas e apenas uma equação; dessa forma, torna-se imperativo inserir condições adicionais. Diversas abordagens para o SFS têm sido propostas, e o trabalho em [19] apresenta uma classificação bastante coerente para os métodos clássicos desenvolvidos. Naquele trabalho, sugere-se que as técnicas propostas podem ser divididas em quatro grupos: abordagens de minimização, de propagação, locais e lineares.

Nas abordagens de minimização, trata-se o SFS como um problema de otimização: Introduce-se uma função custo, cujos termos incorporam restrições ao problema, e a minimização deste funcional produz a solução, ou seja, as componentes do vetor gradiente. Normalmente, dois termos são utilizados no funcional: uma medida de o quanto uma dada solução deixa de satisfazer a equação de irradiância da imagem, e outra que avalia o quanto a solução se afasta de uma superfície suave. Outros termos também podem ser inseridos, como por exemplo, um termo de integrabilidade, que requer que uma dada solução se aproxime de uma superfície integrável (ou seja, $Z_{xy} = Z_{yx}$, onde os subscritos denotam diferenciação), ou ainda uma medida que garanta que o gradiente das intensidades da imagem reconstruída esteja próximo do gradiente das intensidades da imagem de entrada. Podem ser classificados dessa forma os trabalhos em [20-29]. O trabalho desenvolvido em [30] critica os métodos clássicos de minimização e utiliza termos restritivos mais complexos na função custo, possibilitando o resgate de maiores detalhes de forma.

As abordagens de propagação partem da informação sobre o gradiente da superfície em um conjunto de pontos e a propagam sobre toda a imagem. A abordagem de Horn [1], por exemplo, baseada em faixas características (descritas como linhas na imagem ao longo das quais a profundidade e a orientação na cena podem ser computadas a partir do conhecimento dessas informações em pontos dados) é essencialmente um método de propagação. De forma geral, as abordagens de propagação partem dos chamados pontos singulares, que são aqueles em que a irradiância da imagem é máxima. Os algoritmos descritos em [31-35] pertencem a essa classe.

Por outro lado, as abordagens classificadas como locais baseiam-se na suposição de que a superfície é localmente esférica. Enquadram-se nesse grupo o método de Pentland [36], que recupera informações de forma a partir da intensidade e de suas derivadas (primeira e segunda) e o de Lee e Rosenfeld [37], que utiliza o sistema de coordenadas da fonte de luz.

Por sua vez, as abordagens lineares utilizam aproximações lineares para o mapa de reflectância. Dentre essas, podemos citar o trabalho de Pentland [13], baseado na Transformada de Fourier do mapa de reflectância linearizado, e a de Tsai e Shah [38], que lineariza o mapa de reflectância em termos das profundidades.

2.4.4. Estéreo Fotométrico Baseado em Disparidades (EFBD)

Uma abordagem alternativa para o Estéreo Fotométrico padrão, proposta em [39] e posteriormente analisada em [12] e [4], sugeriu o emprego de técnicas de estereoscopia na estimação de forma a partir de imagens de Estéreo Fotométrico.

Esse método parte do princípio de que a mudança na iluminação da cena acarreta um deslocamento do padrão de irradiância sobre as superfícies observadas, que pode ser estimado a partir da correspondência (casamento) entre os pares de imagens de Estéreo Fotométrico. Os mapas de disparidades fotométricas assim obtidos incorporam informação sobre a forma das superfícies na cena, possibilitando a sua reconstrução. Assim, as imagens de Estéreo Fotométrico podem ser relacionadas de modo a produzir um mapa de disparidades semelhante ao obtido em Estereoscopia. Esse mapa de disparidades fotométricas, como é conhecido, pode ser utilizado para a reconstrução em profundidade da superfície observada (Figura 2.10), e também para a obtenção de informação sobre curvaturas.

O EFBD parte de duas relações, uma delas fotométrica e a outra geométrica. Considerando um par de imagens de Estéreo Fotométrico, I_1 e I_2 , a primeira relação é a equação de irradiância linearizada para a imagem diferença,

$$\Delta I \equiv I_1 - I_2 = k_0 + k_1 p + k_2 q \quad (2.16)$$

onde os coeficientes k_i representam os termos associados ao mapa de reflectância, de acordo com a equação (2.12). A segunda equação é uma expressão alternativa para ΔI , obtida a partir da relação de casamento das duas imagens de entrada, $I_2(x + u, y + v) = I_1(x, y)$, que, linearizada, torna-se

$$\Delta I = u \frac{\partial I_2}{\partial x} + v \frac{\partial I_2}{\partial y} \quad (2.17)$$

onde (u, v) denota o campo de disparidades, ou fluxo óptico.

Igualando (2.16) e (2.17), resulta uma equação diferencial em $Z(x,y)$, cuja solução pode ser obtida desde que o casamento de I_1 e I_2 seja efetuado ao longo da direção determinada pela razão $k_2/k_1 \equiv \gamma$, ou seja, para $v = \gamma u$. Nesse caso, temos

$$Z = \frac{uI_2}{k_1} - \frac{k_0(x + \gamma y)}{k_1(1 + \gamma^2)} \quad (2.18)$$

Um exemplo de estimação de forma utilizando esta abordagem aparece nas Figuras 2.10 e 2.11, abaixo.

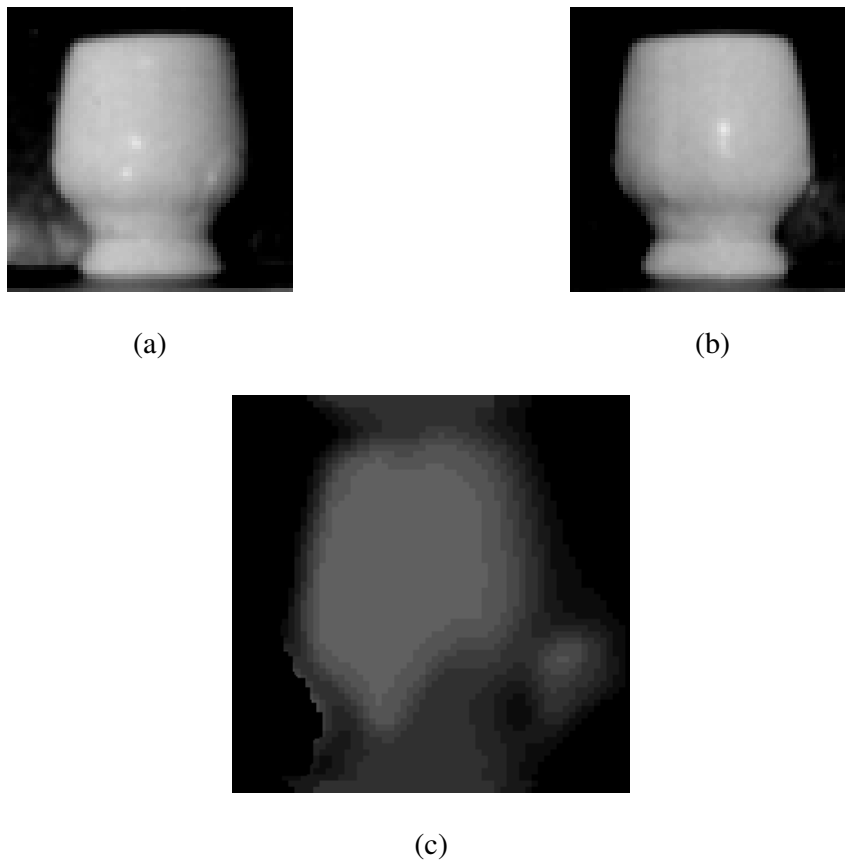


Figura 2.10. Reconstrução de Profundidades a partir do EFBD. (a) e (b) imagens de entrada, com vetores de iluminação $(-0.36,0,1)$ e $(0.36,0,1)$, respectivamente. (c) Mapa de disparidades fotométricas.

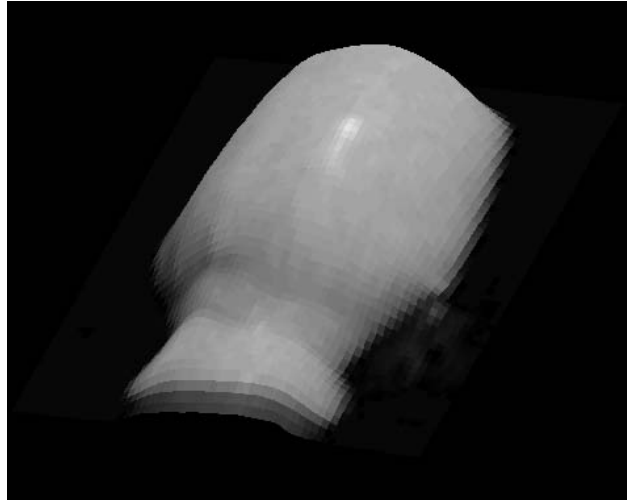


Figura 2.11. Representação para o mapa de profundidades, obtido através do EFBD, a partir das imagens ilustradas na Figura 2.10.

2.4.5 – A abordagem da Função de Green para *Shape from Shading*

No trabalho desenvolvido em [5], o EFBD foi estendido para o caso de uma única imagem de entrada (em um processo, portanto, de *Shape from Shading*), na abordagem denominada *Green's Function Shape From Shading* (GSFS). Nessa abordagem, utilizaram-se funções de Green de uma equação de casamento, de modo a gerar pares virtuais de estéreo fotométrico para a imagem dada, permitindo deste modo tratar um processo baseado em uma única imagem como se este contasse com múltiplas entradas.

Ao invés de se obter as disparidades a partir do casamento de duas imagens dadas, assumiu-se um mapa de disparidades conhecido e, a partir deste, estimou-se o correspondente à segunda imagem de entrada. Assim, expressando o casamento pela condição $I_2(x + u, y + v) = I_1(x, y)$ e tomando uma expansão em série de Taylor dessa equação até segunda ordem, vamos obter, quando $v = \gamma u$ como no EFBD ,

$$\frac{u^2}{2} \frac{\partial^2 I_2}{\partial x^2} + u \frac{\partial I_2}{\partial x} + I_2 = I_1 \quad (2.19)$$

onde $I_i = I_i(x, y + \gamma x)$, e com $\theta = \text{Arctan } \gamma = \text{Arctan } (v/u)$ indicando a direção de casamento. Considerando-se um mapa de disparidades uniforme, a imagem I_2 , que satisfaz (2.19), pode ser obtida, pelo método da função de *Green* [40], como

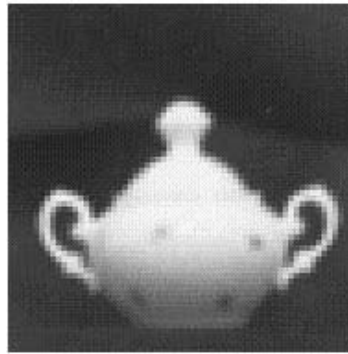
$$I_2(x, y + \gamma x) = \int_D G_u(x - x_0) I_1(x_0, y + \gamma x_0) dx_0 \quad (2.20)$$

onde D é o domínio da imagem. A função $G_u(x - x_0)$, chamada de Função de *Green*, é a solução para a equação (2.19) quando o termo do seu lado direito, I_1 , é substituído pela função Delta de Dirac, e pode ser dada por

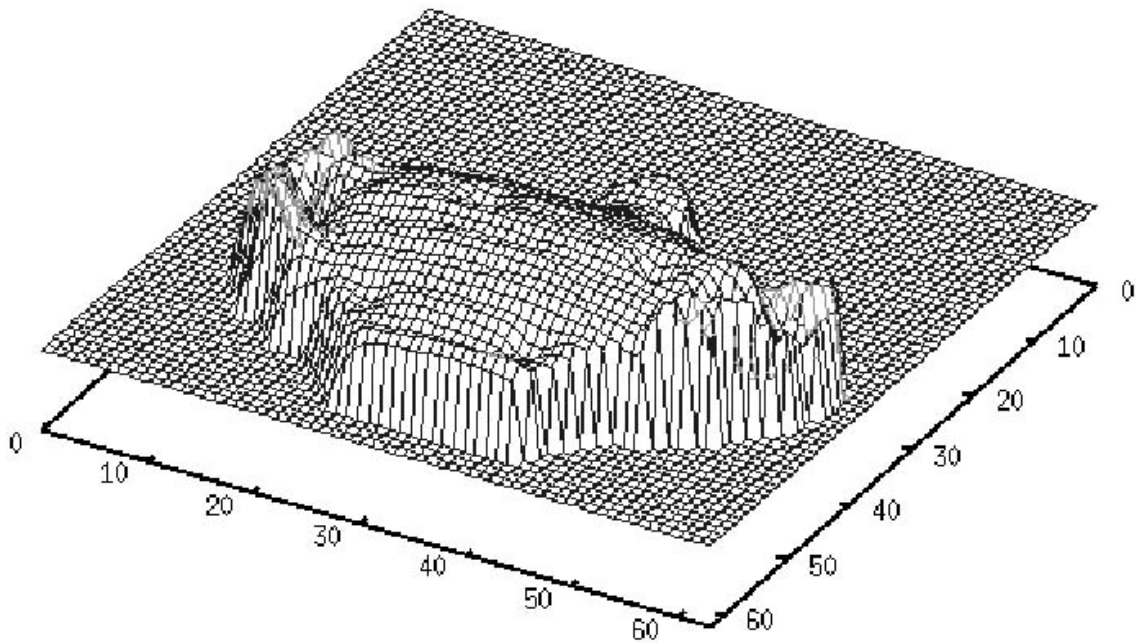
$$G_u(x - x_0) = \begin{cases} \frac{2}{u} \text{sen}\left(\frac{x - x_0}{u}\right) \exp\left[-\left(\frac{x - x_0}{u}\right)\right] & x > x_0 \\ 0 & x < x_0 \end{cases} \quad (2.21)$$

Assumindo-se irradiância linearizada para a imagem I_1 , e com I_2 obtida através de (2.20), é possível chegar, por um procedimento semelhante ao seguido no EFBD, a uma solução aproximada para a forma da superfície observada. É interessante observar que, na grande maioria dos casos, as reconstruções obtidas a partir do GSFS mostraram-se superiores às fornecidas pelo EFBD. Isto é um tanto surpreendente, uma vez que o segundo processo conta com duas imagens de entrada, e o primeiro apenas com uma. A explicação encontra-se no fato de que o EFBD requer o casamento das imagens, um processo sempre difícil e inexato. No GSFS, ao contrário, o par perfeito de casamento para a imagem dada (até segunda ordem em u e v) é obtido via função de *Green*, a partir da Equação (2.20).

Um exemplo de estimação de forma utilizando esta abordagem aparece na Figura 2.12, abaixo.



(a)



(b)

Figura 2.12. Reconstrução de Profundidades a partir do GSFS. (a) Imagem de entrada (b) Representação para o mapa de profundidades.

2.4.6 – Estrutura a partir do Movimento, Fluxo Óptico e Movimento Fotométrico

A recuperação da estrutura de uma cena a partir do movimento é um importante campo de pesquisas em Visão Computacional. O processo de estimação de forma, dada uma seqüência de imagens de uma cena em movimento relativo à câmera, é genericamente conhecido como “Estrutura a partir do Movimento” (*Structure from Motion*). O movimento relativo cena/câmera dá origem a um deslocamento das intensidades no plano da imagem, o

que pode ser usado para a estimação de velocidades, e, em certos casos, também para a recuperação das profundidades na cena.

Os algoritmos de estrutura a partir do movimento geralmente utilizam como dado de entrada o *fluxo óptico*, definido como o movimento do padrão de intensidades na imagem, em decorrência do movimento na cena. O fluxo óptico é um conceito relacionado ao de *campo de movimento*, mas distinto deste, já que o campo de movimento, definido como um campo de vetores que descreve a projeção, na imagem, de um movimento na cena, é um conceito puramente geométrico, enquanto o fluxo óptico envolve também aspectos fotométricos. Idealmente, podemos considerar que o fluxo óptico corresponde ao campo de movimento, mas nem sempre isso é verdadeiro, como podemos perceber pela Figura 2.13.

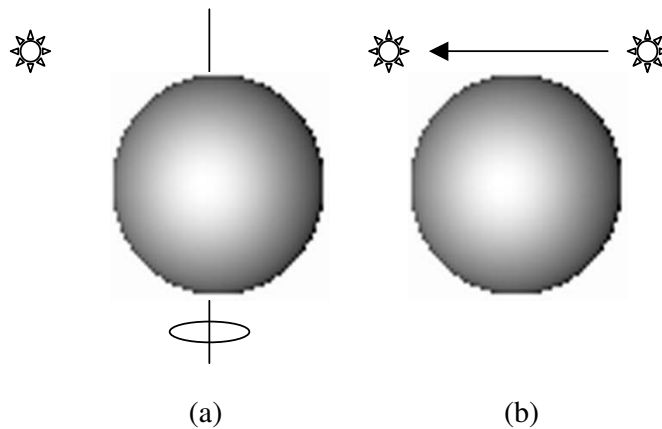


Figura 2.13. O fluxo óptico não é sempre igual ao campo de movimento. Em (a) uma esfera está sendo rotacionada sob iluminação constante (campo de movimento diferente de zero e fluxo óptico nulo). Em (b) uma esfera fixa é iluminada por uma fonte de luz que se movimenta (campo de movimento nulo e fluxo óptico diferente de zero) [10].

De qualquer modo, a hipótese da igualdade entre o fluxo óptico e o campo de movimento é uma suposição básica da maioria das abordagens de estrutura a partir do movimento (ou seja, assume-se que o deslocamento do mapa de irradiância sobre a imagem reflete exclusivamente a distorção projetiva devida ao movimento tridimensional, desprezando-se eventuais variações fotométricas). Uma outra suposição é a de que a irradiância se conserva, de modo que o movimento no plano da imagem possa ser estimado a partir de uma relação de casamento do tipo $I_1(x,y) = I_2(x+u,y+v)$, sendo I_1 e I_2 imagens sucessivas, com $[u(x,y),v(x,y)]$ agora representando o fluxo óptico.

Dentre as abordagens clássicas utilizadas na recuperação das componentes do fluxo óptico, podemos mencionar a descrita em [10], que considera a minimização de um funcional que incorpora a soma ponderada de dois termos: o primeiro derivado da suposição de conservação da irradiância, e o segundo de uma restrição de suavidade. Esse método recai em um processo iterativo que se refina em aproximações sucessivas para as componentes do fluxo óptico. Ainda em [10], podemos encontrar uma abordagem para recuperação do movimento relativo câmera/cena e da forma das superfícies observadas.

Em seu trabalho em [6], Alex Pentland demonstrou que o que ele definiu como *movimento fotométrico* (ou seja, a variação na intensidade em um determinado ponto da cena, como consequência do movimento) pode ser mais significativo do que os efeitos resultantes da distorção projetiva. Pentland, então, desenvolveu uma técnica baseada no movimento fotométrico, para a estimação de forma 3D. Em sua formulação, ele considerou projeção ortográfica e uma expansão em série de Taylor quadrática para o mapa de reflectância, suposto simétrico e separável (sem termos cruzados em p e q , e com mesmo coeficiente de expansão para os termos em p^2 e q^2). Sob tais condições, Pentland tentou determinar a variação de intensidade devida ao movimento.

Como a orientação de um ponto na cena varia com o seu movimento, não se pode empregar a equação de irradiância da imagem diretamente. Pentland então supôs que regiões de fluxo óptico aproximadamente linear podiam ser identificadas, permitindo efetuar-se a correspondência, em imagens sucessivas, dos pontos associados a uma mesma posição na superfície. Verificando que a diferença entre as intensidades destes pontos em registro era descrita por um mapa de reflectância linear, Pentland propôs a aplicação do seu *Linear Shape from Shading* [13], para a estimação de forma a partir das imagens diferença.

No próximo Capítulo, como uma primeira contribuição do nosso trabalho de Tese, nós desenvolveremos uma formulação alternativa àquela de Pentland [7] para o processo de Movimento Fotométrico, que, em muitos aspectos, segue um paralelo com a formulação para o *Estéreo Fotométrico Baseado em Disparidades*. Em seguida, nós demonstraremos como esta formulação pode ser estendida para a estimação de forma a partir de **uma única** imagem de entrada, também com o uso de funções de Green de equações de casamento, assim como no *Green's Function Shape from Shading*.

2.4.7 – Shape from Defocus

A impressão subjetiva que nós, seres humanos, temos é a de que vemos o mundo que nos rodeia em foco perfeito. Mas as imagens na nossa retina estão sempre “desfocadas”,

exceto na região conhecida como *fóvea* [41, 42]. O mesmo ocorre com a maioria dos sistemas ópticos artificiais. Quando consideramos um sistema óptico típico, temos que, teoricamente, uma lente somente proporciona foco preciso para objetos situados a uma certa distância da mesma. Na prática, observa-se que há uma região aquém e além dessa distância, na qual a nitidez da imagem é aceitável. Esse intervalo é denominado **profundidade de campo**. Considerando que essa medida é, normalmente, muito limitada, temos que pontos em uma cena localizados a diferentes distâncias com relação à lente são registrados, no plano da imagem, com diferentes desfocamentos. Assim, pode-se perceber a relação entre a geometria das superfícies na cena e o borramento característico, nas imagens resultantes, produzido pelo desfocamento.

O processo de *Shape from Defocus* aborda o problema de estimar a estrutura 3-D de uma cena a partir de um conjunto de imagens monoculares obtidas pela mudança de foco dos dispositivos de captura (variando-se portanto as regiões da cena que são capturadas em foco perfeito – ver Figura 2.13). O trabalho de Pentland [9] foi o primeiro em Visão Computacional a propor isto, demonstrando que duas imagens produzidas sob diferentes condições da abertura das câmeras poderiam proporcionar informações sobre as profundidades na cena. Em sua abordagem, Pentland modelou o processo de desfocamento através da convolução de uma imagem perfeitamente focada com uma função gaussiana cuja dimensão espacial é proporcional à medida de desfocamento. Considerando janelas correspondentes em um par de imagens capturadas sob diferentes condições abertura, Pentland obteve uma estimativa para a medida de desfocamento (identificada ao desvio-padrão da Gaussiana), a partir das Transformadas de Fourier das imagens de entrada. No entanto, utilizando este método, Pentland só se mostrou capaz de estimar descontinuidades em profundidades. Além disso, trabalhos posteriores apontaram outros problemas da sua abordagem, tais como a introdução de componentes espúrias de alta frequência (devido às descontinuidades artificiais introduzidas pelo uso de janelas) e o vazamento de irradiância (*irradiance leaking*) entre janelas vizinhas.

No Capítulo 4 do presente trabalho, nós retomaremos a abordagem de Pentland, tentando evitar os problemas identificados e desenvolvendo uma nova formulação para o *Shape from Defocus* que permite a estimação de mapas densos de profundidade. Seguindo a filosofia geral da nossa Tese, em seguida nós estenderemos essa formulação para o caso de **uma única** imagem de entrada, introduzindo, portanto, mais uma abordagem para o *Shape from Shading*, agora baseada em informações de desfocamento.

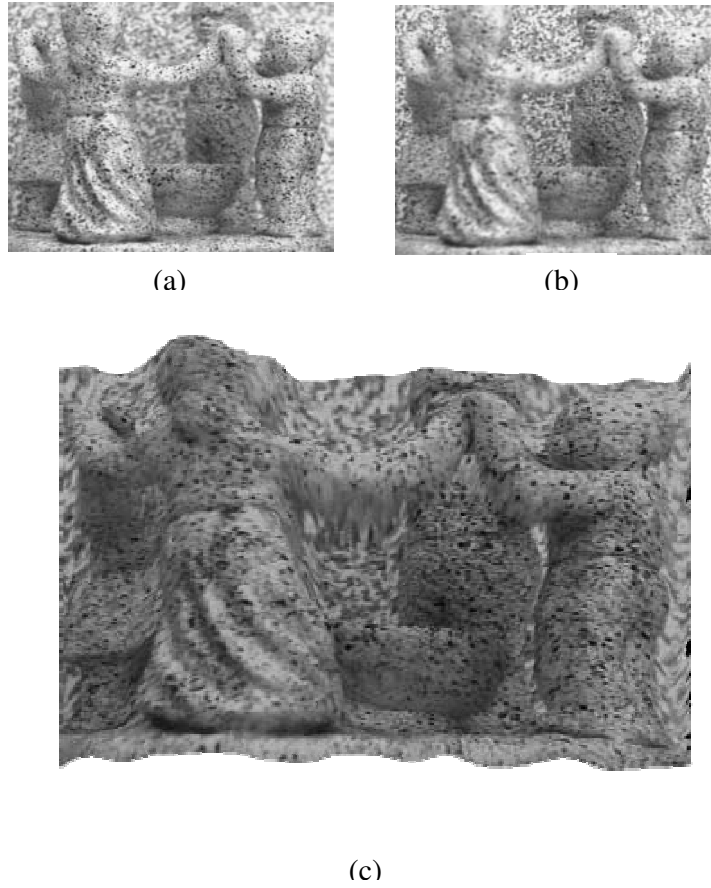


Figura 2.14. *Shape from Defocus.* (a) e (b) Imagens de entrada. Deve-se observar que os objetos na cena e o fundo encontram-se focados diferentemente nas duas imagens. (c) Representação 3-D para o mapa de profundidades estimado.

Do Movimento Fotométrico ao *Shape from Shading***3.1 – Introdução.**

O *Movimento Fotométrico* é um processo introduzido por Pentland [6], baseado em sua observação de que, para superfícies em rotação relativa à câmera, o efeito fotométrico do movimento (isto é, a variação de intensidade de cada ponto em deslocamento) podia ser mais significativo do que o efeito puramente geométrico, associado à distorção projetiva. Pentland assumiu que regiões de fluxo óptico aproximadamente linear podiam ser identificadas em uma seqüência de imagens da cena, permitindo que um *warping* prévio fosse executado, para pôr em registro as imagens sucessivas. Ele então propôs a aplicação do seu *Linear Shape from Shading*, para estimação de forma a partir da variação de intensidade dos pontos em movimento, obtida das imagens em registro.

Aqui, nós introduziremos uma formulação alternativa para o Movimento Fotométrico [7], cuja característica distintiva é a de ser baseada na variação de intensidades, devido ao movimento, em um ponto fixo no plano da imagem (vide Figura 3.1), e não, como na abordagem original, em um dado ponto na superfície em movimento. Assim como no EFBD, a nossa abordagem é desenvolvida a partir de duas relações para a imagem diferença em cada ponto, $\Delta I(x,y)$, que vão permitir a obtenção de uma expressão fechada para o mapa de profundidades da cena representada. A primeira delas resulta de considerações geométricas e fotométricas, e assumirá um papel semelhante ao da equação (2.16) do EFBD. A segunda, por sua vez, é uma relação de casamento de imagens, no espírito da equação (2.17), aplicável sob uma condição de não conservação de irradiância.

Embora nossa abordagem seja baseada em suposições um pouco mais restritivas do que as assumidas por Pentland (nós consideramos apenas rotações em torno dos eixos x e y), ela apresenta a vantagem de não requerer *warping* para o registro de pontos correspondentes na seqüência de imagens, o que conduz a uma implementação mais simples. Posteriormente,

nós mostraremos como essa formulação pode ser estendida para o caso de reconstruções a partir de uma única imagem, também através de Funções de Green. Essa nova aproximação, denominada *Green's Function Photometric Motion*, permite obter reconstruções tridimensionais de qualidade superior às fornecidas pelo GSFS.

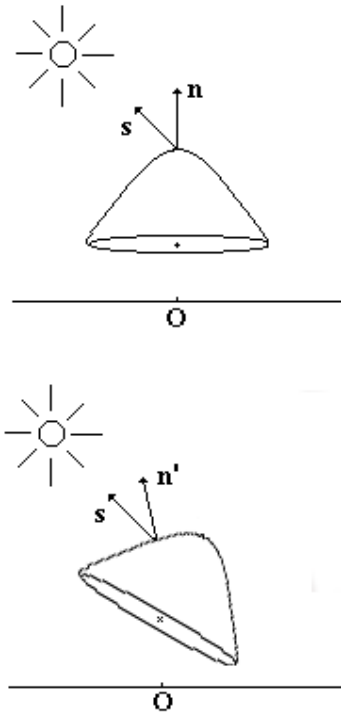


Figura 3.1. Em nossa abordagem para o *Movimento Fotométrico*, estamos interessados na variação de intensidade ΔI , num ponto qualquer no plano da imagem, devido ao movimento da superfície. Na Figura, os vetores \mathbf{n} e \mathbf{s} representam, respectivamente, a normal à superfície e a direção do iluminante.

3.2 – Nova abordagem para o *Movimento Fotométrico*

Vamos considerar uma superfície $Z(x,y)$, rotacionando em torno dos eixos x e y , cuja imagem é capturada sob projeção ortográfica. O movimento da superfície Z na cena dá origem a um fluxo óptico, que nós assumiremos idêntico ao campo de movimento $\mathbf{u} = (u,v)$, onde, pelas equações da rotação

$$u = \frac{dx}{dt} = BZ, \text{ e } v = \frac{dy}{dt} = -AZ \quad (3.1)$$

com A e B denotando as componentes da velocidade angular ao longo das direções x e y , respectivamente. A partir das equações do movimento, vamos obter ainda

$$\frac{dZ}{dt} \equiv \frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial x} \frac{dx}{dt} + \frac{\partial Z}{\partial y} \frac{dy}{dt} = Ay - Bx \quad (3.2)$$

e, a partir das equações anteriores,

$$\frac{\partial Z}{\partial t} = -B[Z(p + \gamma q) + (x + \gamma y)] \quad (3.3)$$

onde p e q são as componentes do gradiente da superfície, e onde definimos

$$\gamma = -\frac{A}{B} \equiv \frac{v}{u} \quad (3.4)$$

com a segunda identidade acima resultando de (3.1).

A equação (3.4) implica em que, para rotação uniforme, o campo de movimento será unidimensional, o fluxo óptico reduzindo-se a um mapa de disparidades.

Vamos assumir agora que a irradiância da imagem possa ser expressa, em termos do gradiente da superfície, através da forma linear

$$I = k_0 + k_1 p + k_2 q \equiv k_0 + k_1 (p + \gamma q) \quad (3.5)$$

Onde $k_2/k_1 = \gamma$ para o mesmo γ introduzido em (3.4). Isto significa que estamos assumindo que a superfície rotaciona em torno de um eixo perpendicular à projeção do “iluminante generalizado” – ou seja, o vetor (k_1, k_2, k_0) [13] – sobre o plano xy . Observe-se que a nossa introdução do parâmetro γ na análise do movimento fotométrico, guarda estreita consistência com o que foi feito no EFBD: lá, também esta constante foi definida como sendo a razão entre os coeficientes da expansão linear de um mapa de reflectância (no caso, associado à imagem diferença ΔI – ver Capítulo 2, seção 2.4.4), e escolheu-se a direção de casamento de acordo com ela (i.e., $v/u = k_2/k_1 = \gamma$).

A variação da irradiância devida ao movimento, em um ponto qualquer da imagem, pode ser expressa por

$$\Delta I \equiv I(t - \delta t) - I(t) = -\frac{\partial I}{\partial t} \delta t \quad (3.6)$$

onde δt é o fator de discretização temporal, que estaremos assumindo igual a 1. Usando (3.5) e (3.3), vamos obter

$$\frac{\Delta I}{k_1} = -\frac{\partial(p + \gamma q)}{\partial t} \equiv -\left(\frac{\partial}{\partial x} + \gamma \frac{\partial}{\partial y}\right) \frac{\partial Z}{\partial t} \quad (3.7)$$

e

$$\frac{\Delta I}{k_1} = \partial_\gamma \{B[Z(p + \gamma q) + (x + \gamma y)]\} \quad (3.8)$$

onde definimos

$$\partial_\gamma \equiv \frac{\partial}{\partial x} + \gamma \frac{\partial}{\partial y} \quad (3.9)$$

Novamente usando (3.5), obtemos

$$\Delta I = \partial_\gamma \{B[Z(I - k_0) + k_1(x + \gamma y)]\} \quad (3.10)$$

E, substituindo $B = u/Z$, da equação (3.1), finalmente chegamos a

$$\Delta I = \partial_\gamma \left[u(I - k_0) + \frac{k_1 u}{Z} (x + \gamma y) \right] \quad (3.11)$$

A equação (3.11) relaciona a forma da superfície na cena, $Z(x,y)$ aos seus observáveis: a seqüência de imagens e o fluxo óptico. Para obter, a partir dessa relação, uma estimativa da função Z , vamos introduzir agora, no espírito do EFBD (ver equação (2.17)), uma equação apropriada de casamento. Nós a propomos como

$$\Delta I = \left[u - (\partial_\gamma u) \left(\frac{x + \gamma y}{1 + \gamma^2} \right) \right] \partial_\gamma I \quad (3.12)$$

que tem a forma esperada para um casamento afim unidimensional, o que fica claro quando a reescrevemos como

$$\Delta I = [|u| - (x \cdot s) (\nabla |u| \cdot s)] \nabla I \cdot s \quad (3.13)$$

onde $x=(x,y)$, $|u|$ é a magnitude do vetor de disparidades, $u = u(I, \gamma)$, e $s = (\cos \theta, \sin \theta)$, é um versor que indica a sua direção, para $\tan \theta = \gamma$.

Igualando (3.11) e (3.12), nós obtemos uma equação diferencial em Z , cuja solução é dada por

$$Z(x, y) = - \frac{k_1(1 + \gamma^2)u}{(\partial_\gamma u)(I - k_0) + \frac{K}{(x + \gamma y)}} \quad (3.14)$$

como pode ser facilmente verificado por substituição direta, quando se despreza o termo em $\partial_\gamma^2 u$, sendo K uma constante arbitrária. Também é fácil perceber que o mesmo resultado seria obtido, caso os coeficientes do mapa de reflectância k_0 e k_1 , variassem em função de $(y - \gamma x)$.

Uma vez estimado o campo u , o mapa de profundidades da superfície observada pode ser reconstruído, utilizando-se (3.14), a partir de uma seqüência de imagens do seu movimento, conforme discutido em [7].

Algoritmo de casamento

De forma a utilizar a equação (3.14) para reconstrução de uma superfície, devemos, primeiro, obter um campo de fluxo óptico que satisfaça à condição de casamento afim dada pela equação (3.12). A estimação do fluxo óptico é um problema não-trivial, para o qual uma série de abordagens sofisticadas têm sido propostas [43,44,45]. Como nosso principal interesse reside na reconstrução de forma, e não no problema do fluxo óptico por si mesmo, utilizaremos um procedimento iterativo bem simples, que é uma extensão, para casamento afim, do conhecido algoritmo desenvolvido por Horn e Schunck [46]. Nesse processo, deve-se minimizar o funcional

$$\mathcal{E} = \mathcal{E}_s + K_c \mathcal{E}_c \quad (3.15)$$

onde \mathcal{E}_s é um termo que representa a condição de suavidade do fluxo óptico, dado por

$$\mathcal{E}_s = \int dx_\gamma \{ (u - \bar{u})^2 + (u_\gamma - \bar{u}_\gamma)^2 \} \quad (3.16)$$

e \mathcal{E}_c é um termo que representa a equação do fluxo óptico (3.12), dado por

$$\mathcal{E}_c = \int dx_\gamma \{ I_t - [u - u_\gamma x_\gamma] I_\gamma \}^2 \quad (3.17)$$

Nas equações acima, K_c é uma constante de ponderação, $I_t = -\Delta I$, $u_\gamma = \partial_\gamma u$, $I_\gamma = \partial_\gamma I$, $x_\gamma = (x + \gamma y)/(1 + \gamma^2)$, e \bar{u} e \bar{u}_γ representam médias espaciais (computadas sobre os primeiros vizinhos ao longo da direção de casamento) dos valores de u e u_γ .

Empregando cálculo variacional [40], nós minimizamos o funcional \mathcal{E} e encontramos o seguinte esquema iterativo para u (o lado direito é calculado na n -ésima iteração, para obter uma estimativa para u na iteração $n+1$):

$$\begin{aligned} u = \bar{u} + K_c \{ I_t - [u - u_\gamma x_\gamma] I_\gamma \} I_\gamma - \frac{K_c}{2\epsilon} \{ I_t - [u - u_\gamma x_\gamma] I_\gamma \}^{(-)} I_\gamma^{(-)} x_\gamma^{(-)} + \\ + \frac{K_c}{2\epsilon} \{ I_t - [u - u_\gamma x_\gamma] I_\gamma \}^{(+)} I_\gamma^{(+)} x_\gamma^{(+)} \end{aligned} \quad (3.18)$$

Aqui, ϵ é o fator de discretização espacial, e os sobrescritos $(-)$ e $(+)$ indicam que a expressão deve ser avaliada, respectivamente, à esquerda e à direita da posição corrente no domínio discretizado da imagem. Uma vez determinado u , a sua derivada u_γ pode ser estimada, em cada iteração, pela abordagem apresentada em [47], também baseada em funções de Green. Através da mesma aproximação, também podemos obter as derivadas I_γ presentes na equação (3.18).

Experimentos

Aqui apresentamos alguns experimentos com a nossa abordagem para o *Movimento Fotométrico* (Figuras 3.2 a 3.11). Cada par de imagens foi casado de acordo com (3.18), para obtenção do campo de fluxo óptico e de suas derivadas, como requerido para a estimação de profundidades através da equação (3.14), onde tomamos $K = 0$.

Deve ser ainda observado que, normalmente, os coeficientes k_0 e k_1 não são conhecidos, já que, no caso geral, não conhecemos uma expressão analítica para o mapa de reflectância da superfície. Assim, na nossa implementação do processo, as reconstruções foram obtidas a menos do fator multiplicativo k_1 e com k_0 escolhido empiricamente. Também deve ser observado que, se o iluminante generalizado fosse estimado, como descrito em [13], também poderíamos obter o parâmetro γ , da direção de casamento, já que $\gamma \equiv k_2/k_1$. Aqui, escolhemos $\gamma = 0$.

Como podemos perceber pelas figuras, em todos os casos, a forma estimada se aproxima da superfície real, mesmo considerando que há perda de informação ao longo das áreas de irradiância relativamente uniforme. A pequena gama de valores normalmente obtida para o mapa de profundidades é devida à aproximação linear utilizada para o mapa de reflectância, e também às pequenas rotações consideradas. Observamos que a abordagem de *Photometric Motion* de Pentland também apresenta limitações similares, como podemos inferir através dos experimentos reportados em [6].

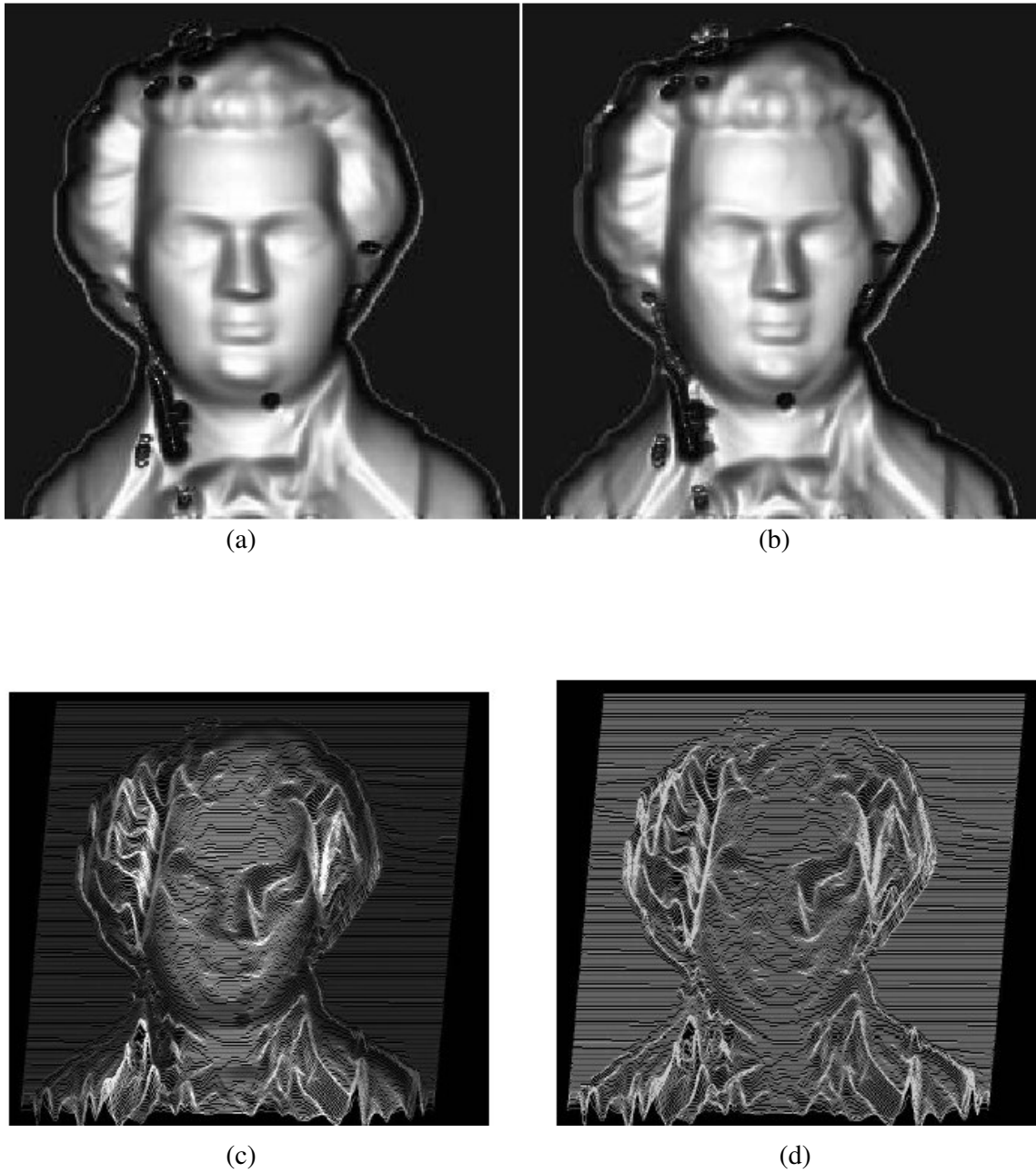
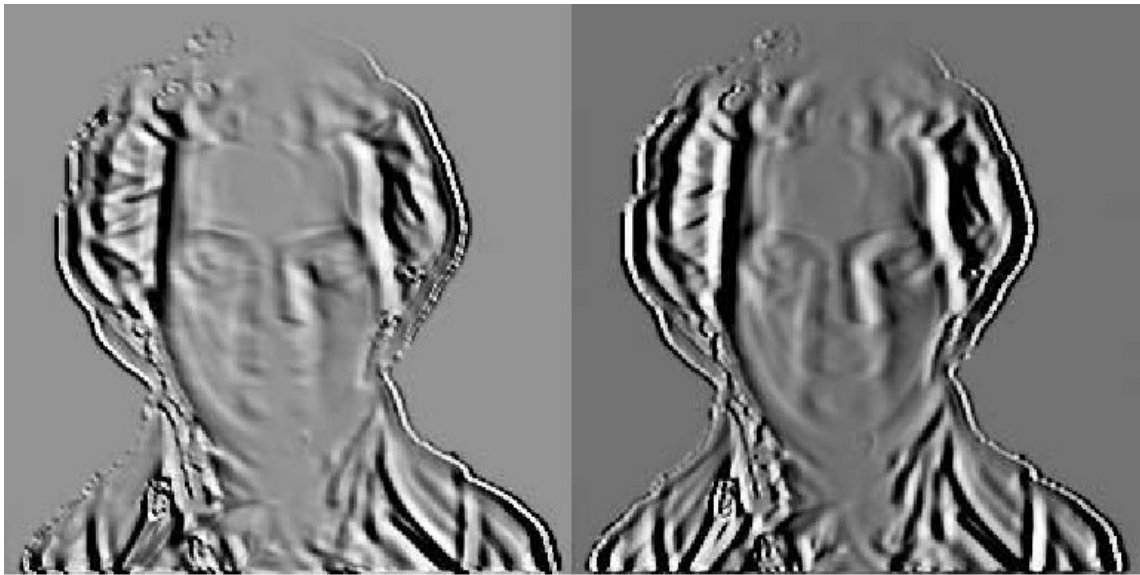
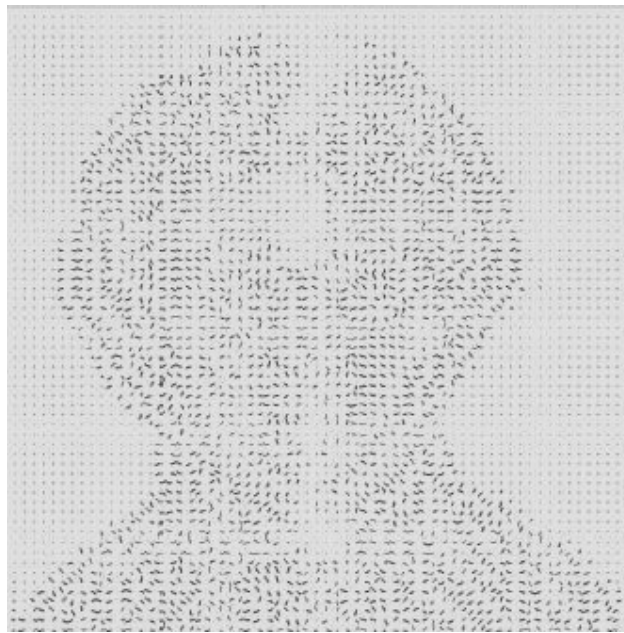


Figura 3.2. (a) e (b): Imagens de entrada obtidas a partir da re-iluminação do mapa real de profundidades (a) e de uma versão rotacionada deste mapa (b), supondo reflectância lambertiana e albedo uniforme. Direção de iluminação: $(0,0,1)$. (c): Profundidades estimadas, com o padrão de irradiância em (a) mapeado na representação. (d): O mesmo que em (c) sem o mapeamento de intensidades.



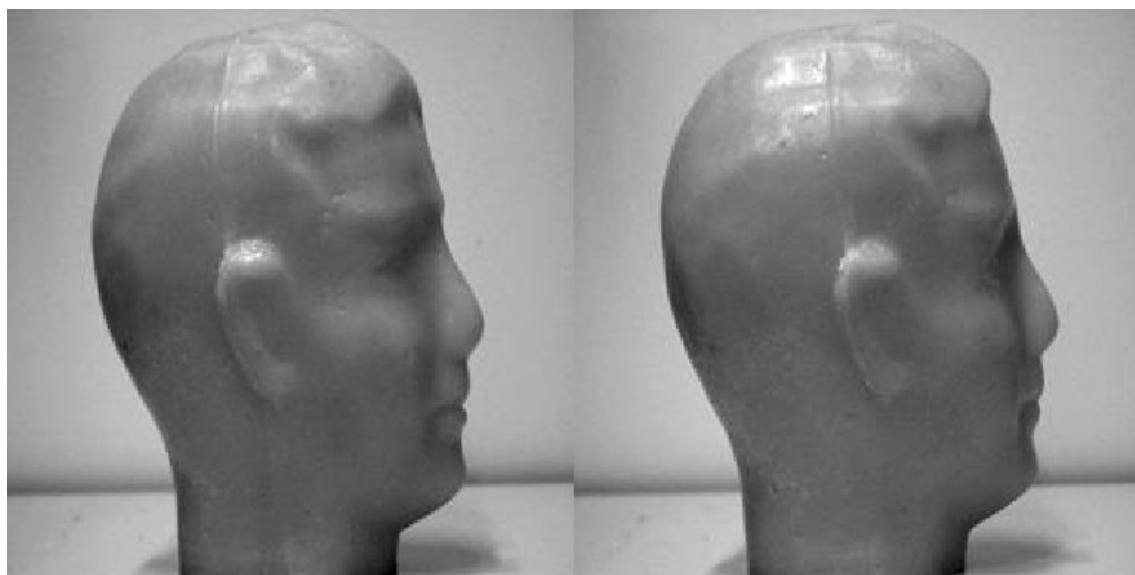
(a)

(b)



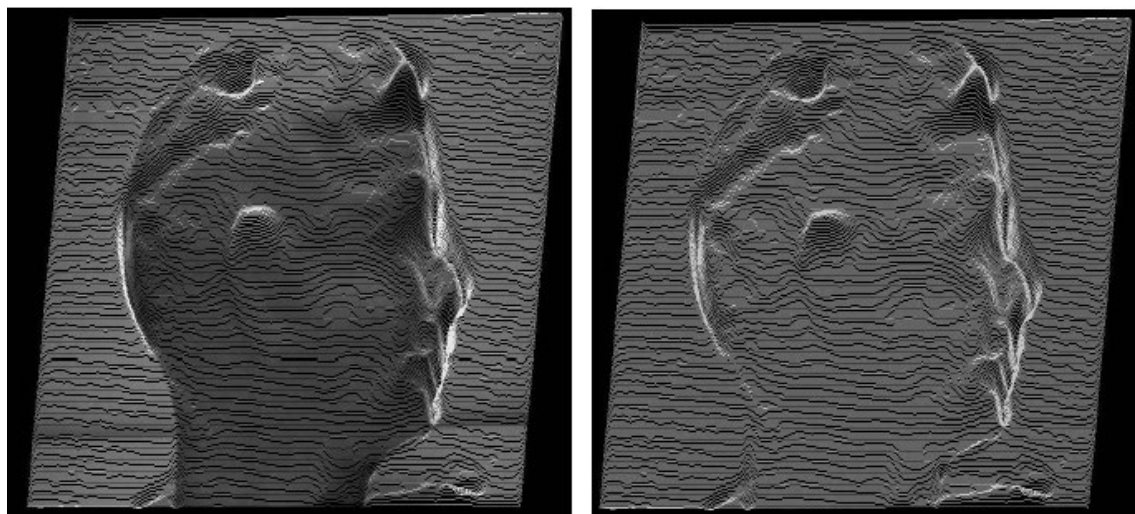
(c)

Figura 3.3. (a) e (b): Imagens obtidas a partir da re-iluminação da superfície estimada (Figura 3.2), supondo reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(1,2,1)$. (c): Mapa de orientações estimado.



(a)

(b)



(c)

(d)

Figura 3.4. (a) e (b): Imagens de entrada. (c): Profundidades estimadas, com o padrão de irradiância em (a) mapeado na representação. (d): O mesmo que em (c) sem o mapeamento de intensidades.



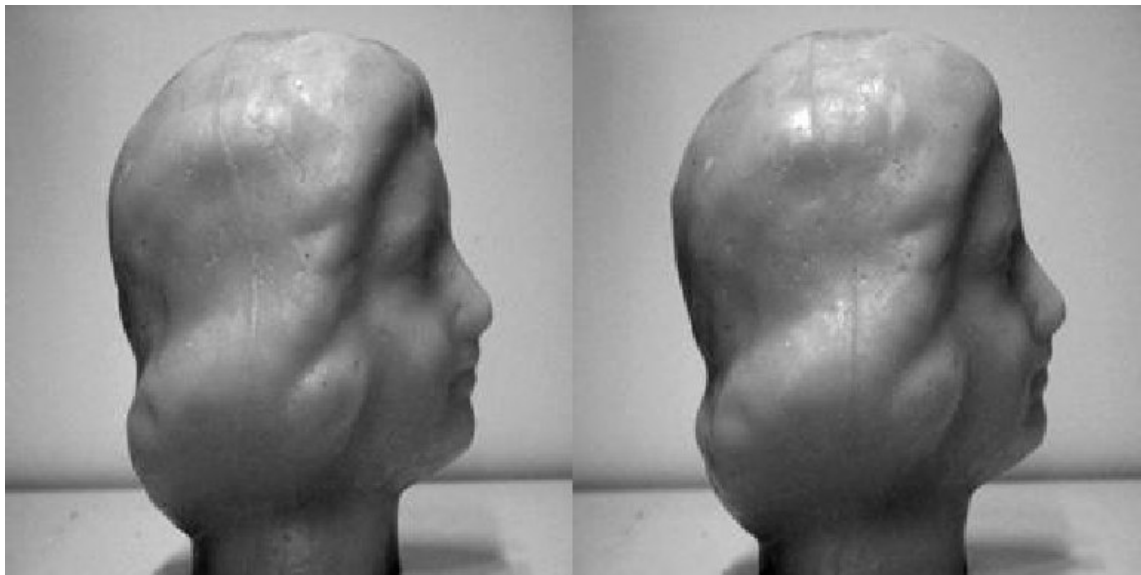
(a)

(b)



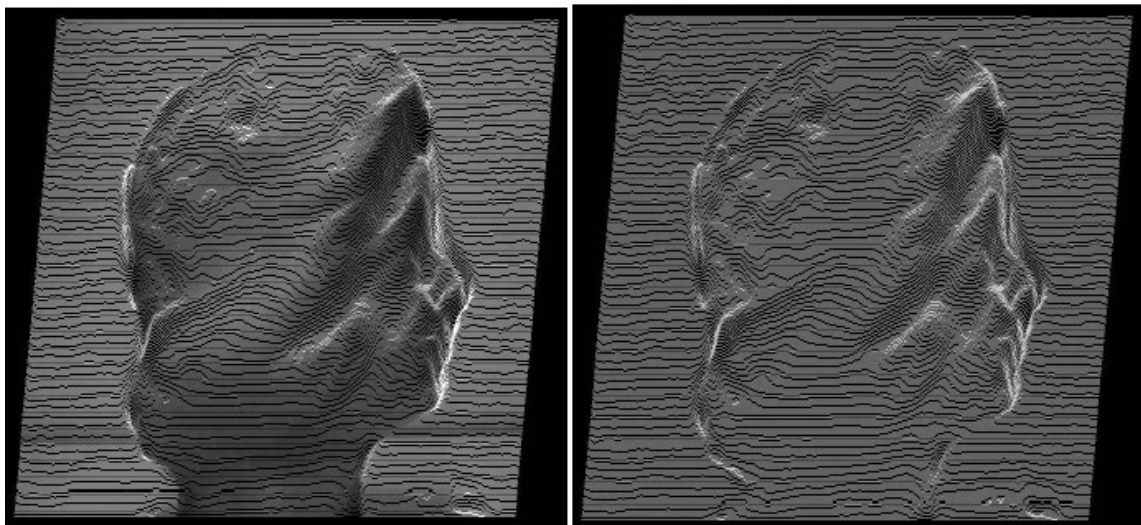
(c)

Figura 3.5. (a) e (b): Imagens obtidas a partir da re-iluminação da superfície estimada (Figura 3.4), supondo reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(0.5,2,1)$. (c): Mapa de orientações estimado.



(a)

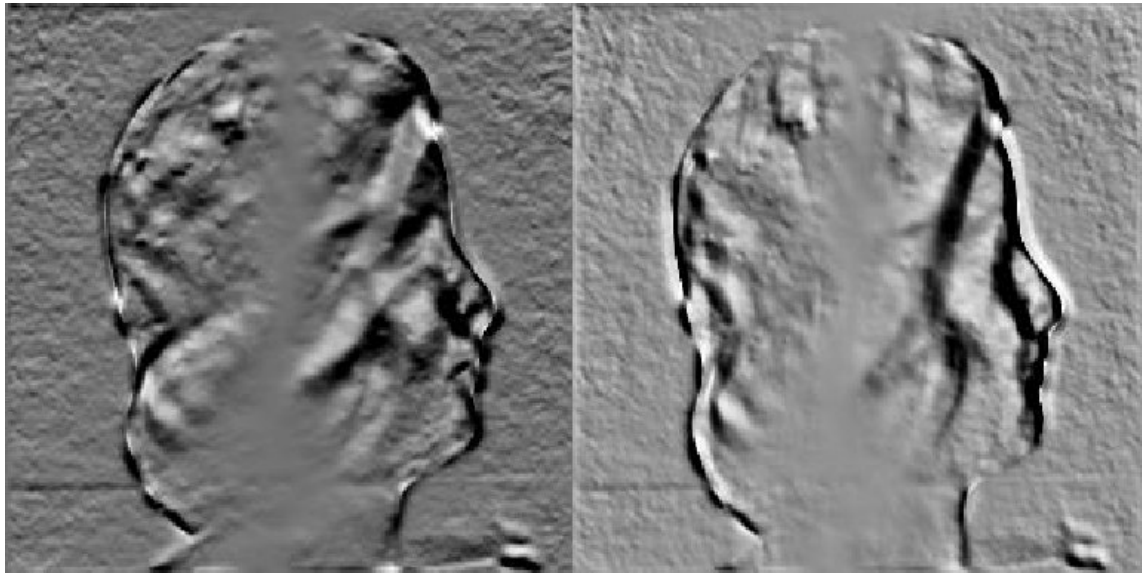
(b)



(c)

(d)

Figura 3.6. (a) e (b): Imagens de entrada. (c): Profundidades estimadas, com o padrão de irradiância em (a) mapeado na representação. (d): O mesmo que em (c) sem o mapeamento de intensidades.



(a)

(b)



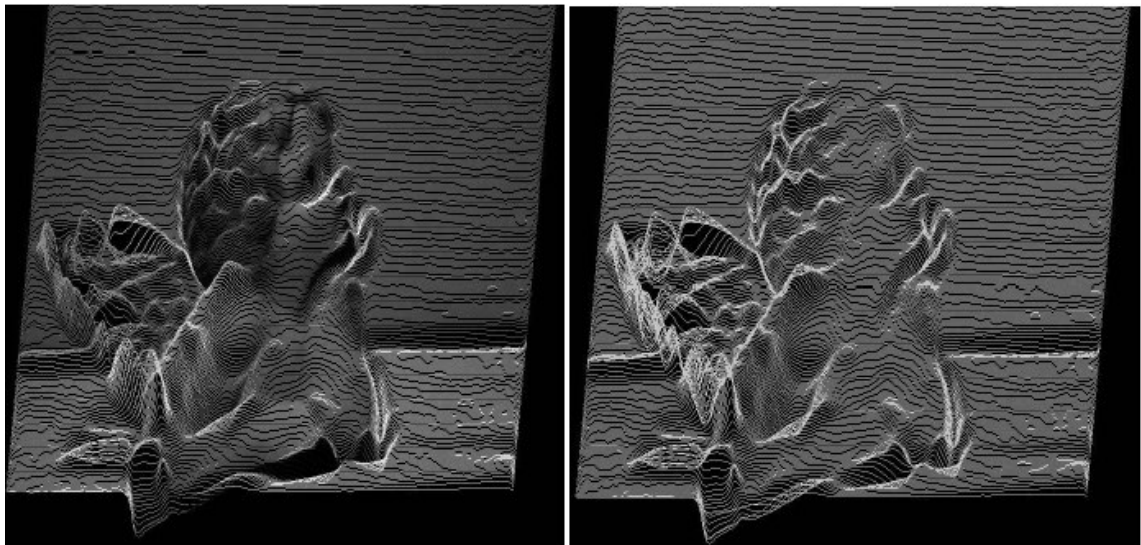
(c)

Figura 3.7. (a) e (b): Imagens obtidas a partir da re-iluminação da superfície estimada (Figura 3.6), supondo reflectância lambertiana e albedo uniforme. Direções de iluminação: $(0,2,1)$ e $(1,1,1)$. (c): Mapa de orientações estimado.



(a)

(b)



(c)

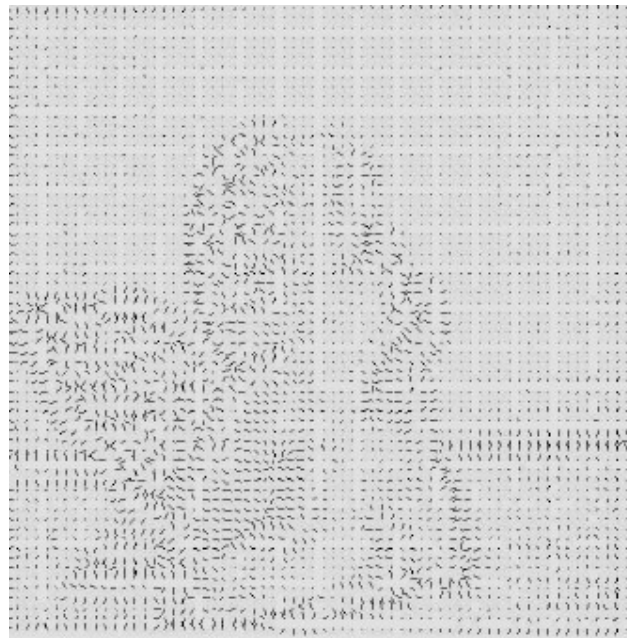
(d)

Figura 3.8. (a) e (b): Imagens de entrada. (c): Profundidades estimadas, com o padrão de irradiância em (a) mapeado na representação. (d): O mesmo que em (c) sem o mapeamento de intensidades.



(a)

(b)



(c)

Figura 3.9. (a) e (b): Imagens obtidas a partir da re-iluminação da superfície estimada (Figura 3.8), supondo reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(2,0,1)$. (c): Mapa de orientações estimado.

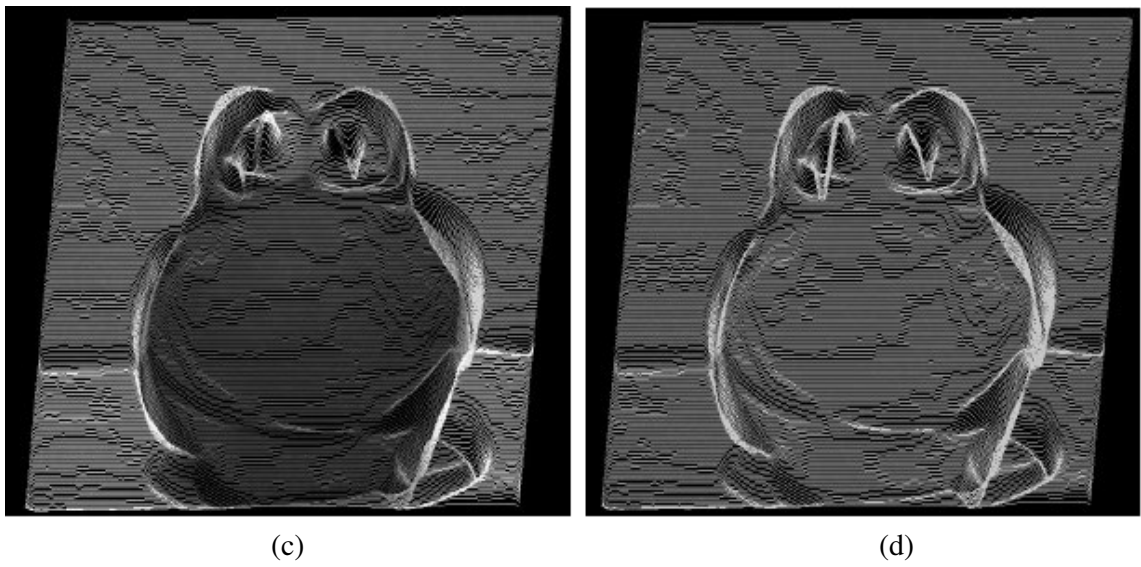
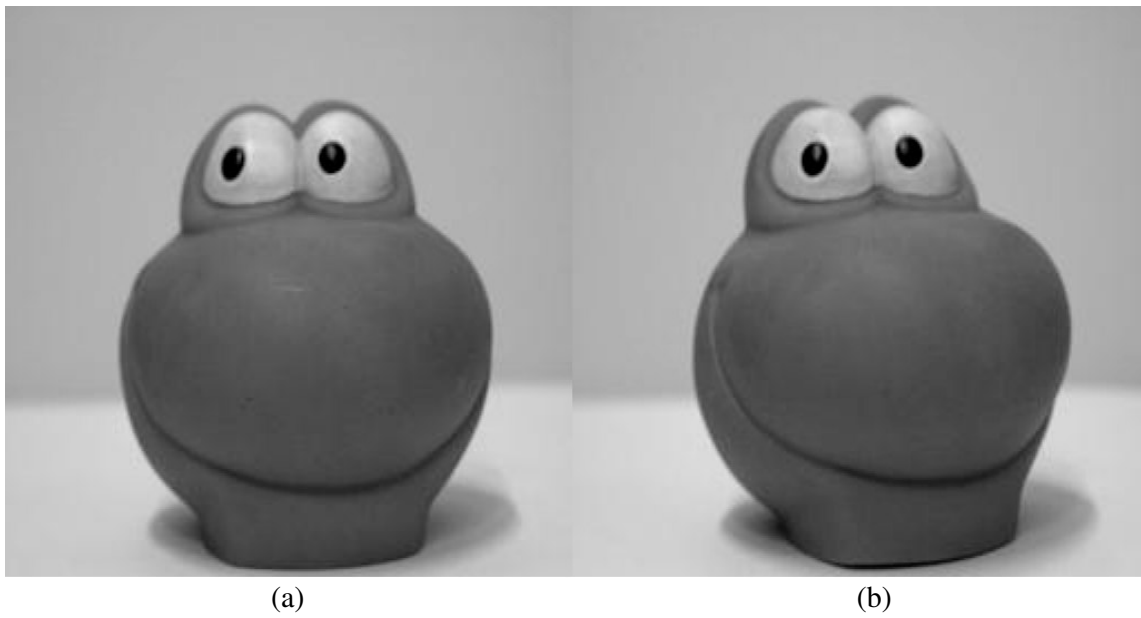
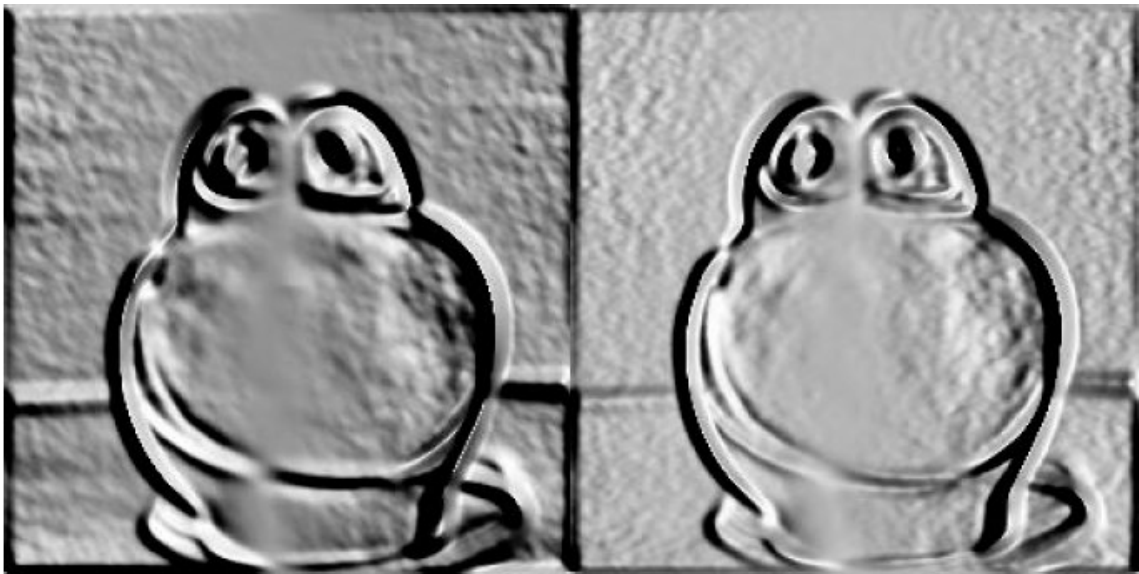
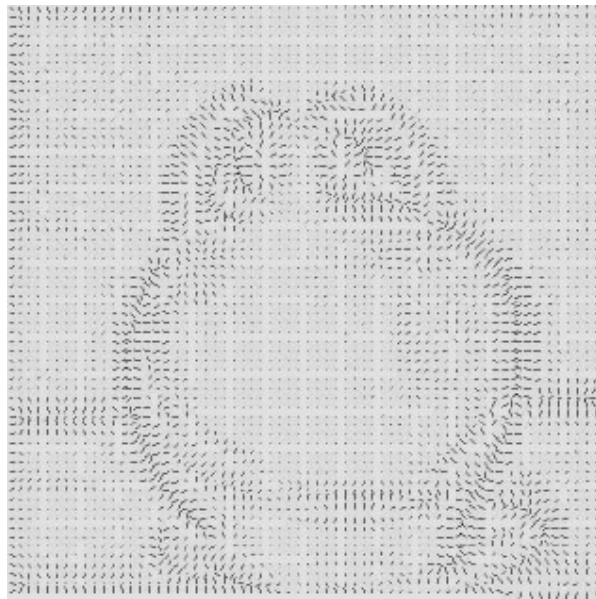


Figura 3.10. (a) e (b): Imagens de entrada. (c): Profundidades estimadas, com o padrão de irradiância em (a) mapeado na representação. (d): O mesmo que em (c) sem o mapeamento de intensidades.



(a)

(b)



(c)

Figura 3.11. (a) e (b): Imagens obtidas a partir da re-iluminação da superfície estimada (Figura 3.10), supondo reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(1,0,1)$. (c): Mapa de orientações estimado.

3.3 – Green’s Function Photometric Motion (GPM)

Nosso objetivo, agora, é estender a abordagem da seção anterior para o caso de uma única imagem de entrada. Esse processo é baseado na introdução de uma função de *Green* que, a partir da imagem de entrada I_1 , permita obter uma imagem transformada I_2 que se relacione à primeira de acordo com (3.12), de forma semelhante ao que foi feito na passagem do EFBD para o GSFS (vide Capítulo 2, Seção 2.4.5). Explicitamente, nós consideraremos um campo de disparidades com parâmetros u e $\partial_\gamma u$ uniformes, e geraremos a imagem que se casa perfeitamente à imagem dada, de acordo com (3.12). É importante observar que métodos exatos de *warping* poderiam ser utilizados na geração deste par de casamento, em lugar da função de Green. No entanto, as técnicas de *warping* são menos realísticas na simulação de movimento, por não introduzirem borramento (*motion blur*), e também não proporcionam uma relação mais estreita com processos da visão biológica, como a permitida pelas funções de Green, como veremos adiante.

Considerando que a expressão (3.12) é unidimensional, podemos, sem perda de generalidade, denotar a direção de casamento por x . Dessa forma, as imagens I_1 e I_2 serão relacionadas como

$$\Delta I \equiv I_1 - I_2 = (u_0 - u_1 x) \frac{\partial I_2}{\partial x} \quad (3.19)$$

onde u_0 e u_1 são constantes associadas, respectivamente, ao mapa de disparidades e à sua derivada em $x = 0$ (no caso geral, $u_1 = \partial_\gamma u$).

Vamos propor nossa função de *Green* como uma extensão da $G_u(x - x_0)$ em (2.21), numa forma remanescente a uma função de Gabor, isto é,

$$G(x, x_0) = \frac{2a}{\sigma^2} \text{sen}\left[\frac{a(x - x_0)}{\sigma^2}\right] \exp\left[-\frac{(x + a)^2 - (x_0 + a)^2}{2\sigma^2}\right] \quad (3.20)$$

para $x > x_0$, com $G = 0$ para $x < x_0$, onde os parâmetros σ e a se relacionam a u_0 e u_1 na forma $\sigma^2 = u_0^2 / u_1$ e $a = u_0 / u_1$, como mostraremos adiante.

No Anexo I deste trabalho, nós provamos que o filtro G é a parte imaginária do núcleo complexo

$$K(x, x_0) = N(x_0) W(x) W^*(x_0) \quad (3.21)$$

onde o asterisco denota conjugação complexa. O fator $N(x_0)$ é independente de x e $W(x)$ é a função de Gabor complexa

$$W(x) = \exp\left[i\frac{ax}{\sigma^2}\right] \exp\left[-\frac{(x+a)^2}{2\sigma^2}\right] \quad (3.22)$$

Quando $x, x_0 \ll a$, G tende à função G_u , para $u = \sigma^2/a$, mas, ao contrário de G_u , G não é invariante sob translação, nem normalizada (vide Seção 3.3.1).

Pelo método da função de Green, dada uma única imagem de entrada, $I_1(x)$, nós podemos gerar o seu par I_2 através de

$$I_2(x) = \int_D G(x, x_0) I_1(x_0) dx_0 \quad (3.23)$$

A relação entre I_1 e I_2 pode então ser facilmente obtida como

$$\frac{1}{2} \left(\frac{\sigma^2}{a}\right)^2 I_2'' + \left(\frac{\sigma^2}{a^2}\right) (x+a) I_2' + \left[1 + \frac{(x+a)^2 + (\sigma^2 - a^2)}{2a^2}\right] I_2 = I_1 \quad (3.24)$$

e, para valores grandes de a , de tal forma que $\sigma, |x| \ll a$, (3.24) pode ser aproximada por

$$\frac{1}{2} \left(\frac{\sigma^2}{a}\right)^2 I_2'' + \left(\frac{\sigma^2}{a} + \frac{\sigma^2}{a^2} x\right) I_2' + I_2 = I_1 \quad (3.25)$$

onde os apóstrofos indicam diferenciação com relação a x . Até primeira ordem em σ^2/a , (3.25) torna-se

$$\Delta I = \left(\frac{\sigma^2}{a} + x \frac{\sigma^2}{a^2}\right) \frac{\partial I_2}{\partial x} \quad (3.26)$$

Comparando as relações (3.26) e (3.19), vemos que a função de *Green* proposta leva a uma relação de casamento da forma desejada, com

$$u_0 = \frac{\sigma^2}{a} \text{ e } u_1 = -\frac{\sigma^2}{a^2} \quad (3.27)$$

Pode-se observar ainda, através da comparação com a equação (2.17), que essa nova abordagem da função de Green se baseia numa forma modificada da condição de casamento já considerada para o EFBD e o GSFS, estendendo-se essa condição para um campo de disparidades afim, ao invés de estritamente uniforme.

A Figura 3.12 apresenta uma imagem de entrada e a sua transformação de acordo com o Filtro de Green, para as direções de casamento $\gamma=0$ e $\gamma=0.175$. Ela mostra que esse filtro apresenta um efeito consistente com o nosso objetivo de emular rotações na cena: Não é difícil considerar as figuras 3.12 (b) e (c) como representando a mesma superfície em (a), mas rotacionada, respectivamente, em torno da vertical e em torno de um eixo passando pelo nariz do Manequim (para maiores detalhes sobre a utilização de funções de Green de equações de casamento afim para simulação de movimento, veja [48]). A Figura 3.13 apresenta a reconstrução tridimensional da superfície, obtida de acordo com a equação (3.14), para $K=0$, conforme discutido na Seção 3.3.1.

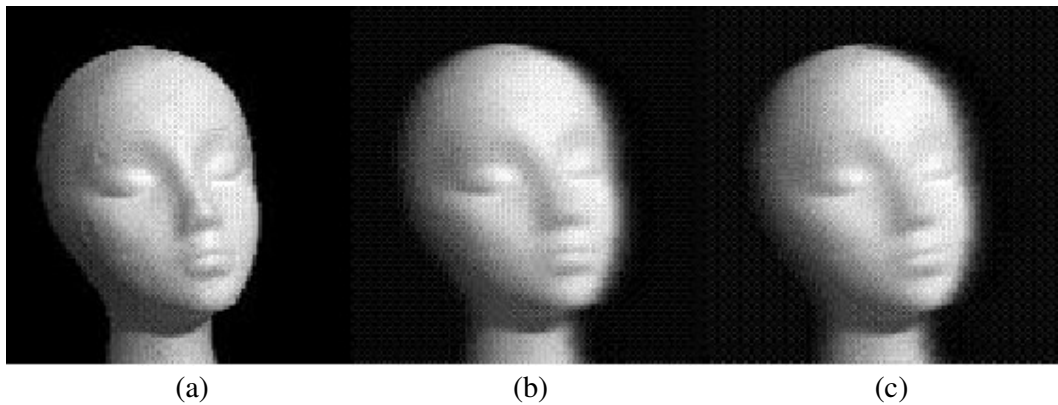


Figura 3.12. Imagem *Manequim* (a), e o resultado de sua filtragem, de acordo com (3.23), ao longo da direção horizontal (b) e ao longo da direção $+10^\circ$ (c). Não é difícil considerar (b) e (c) como representando a mesma superfície em (a), mas rotacionada em torno da vertical (b) e em torno de um eixo passando pelo nariz do Manequim (c). Parâmetros utilizados: $a=2$ e $\sigma=0.2$, nos dois casos.

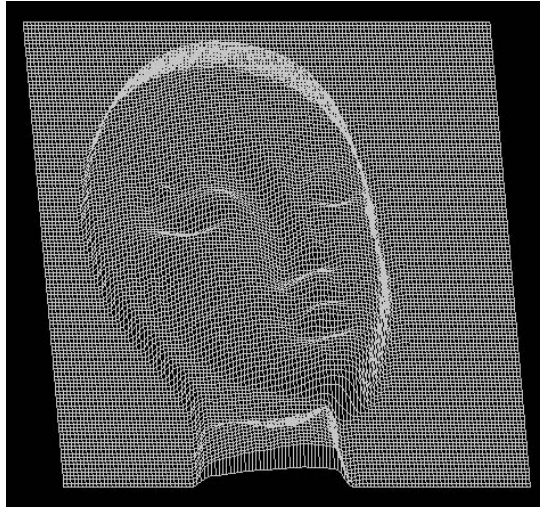


Figura 3.13. Representação para o mapa de profundidades estimado, através do GPM, a partir da imagem na Figura 3.12 (a) ($\gamma=0.175$, $a=2$ e $\sigma=0.1$).

Antes de apresentarmos nossos experimentos, é importante observar ainda que a forma I_2 na equação (3.23) não é única a satisfazer a condição de casamento desejada: outras possibilidades resultariam da combinação linear de I_2 com uma solução para a forma homogênea da equação (3.25). Tal solução homogênea pode ser obtida pela filtragem da imagem de entrada por uma função $H(x, x_0)$ similar a $G(x, x_0)$, mas com um co-seno substituindo a função seno na equação (3.20). As funções G e H apresentam uma propriedade relevante: quando filtram sinais modulados por funções de Gabor, tais como os que modelam o campo receptivo das chamadas células simples do córtex visual [49], elas produzem saídas similares, mas para versões espacialmente deslocadas do sinal de entrada. Isto pode ser provado da seguinte forma: usando a representação complexa

$$K(x, x_0) \equiv (H + iG)(x, x_0) = 2k \exp[ik(x - x_0)] \exp\left[-\frac{(x + a)^2 - (x_0 + a)^2}{2\sigma^2}\right]$$

onde $k \equiv a/\sigma^2$, teríamos, para o sinal filtrado

$$I_2(x) = \int_D K(x, x_0) I_1(x_0) dx_0 \quad (3.28)$$

onde aqui estamos assumindo

$$I_1(x) = \exp[i\omega x] \exp\left[-\frac{(x + \alpha)^2}{2\sigma^2}\right] I(x) \quad (3.29)$$

como o sinal de entrada modulado por uma função de Gabor.

Após manipulações triviais, a equação (3.28) se torna

$$I_2(x) = \exp[i(\omega x - \phi)] \exp\left[-\frac{(x + \alpha)^2}{2\sigma^2}\right] \bar{I}(x) \quad (3.30)$$

onde

$$\bar{I}(x) \equiv 2k \int_D \exp[i(a - \alpha)\xi / \sigma^2 - (a - \alpha)\xi / \sigma^2] I(x - \xi) d\xi \quad (3.31)$$

e $\phi = (a - \omega\sigma^2) / (\alpha - a)$.

Como pode ser facilmente verificado (compare (2.20) e (2.21), para $\gamma = 0$), temos que é possível identificar $\bar{I}(x) \equiv I(x - u)$, para $u = \sigma^2 / (a - \alpha)$. Pares de sinais como os das equações (3.29) e (3.30) são utilizados, em neurofisiologia, para modelar os campos receptivos das células simples do córtex visual associadas às entradas provenientes dos olhos direito e esquerdo, na visão estereoscópica. Nós temos, portanto, que o núcleo de Green complexo, $K = H + iG$, proporciona um *link* entre estes tipos de sinais que é exatamente aquele previsto pelo chamado *modelo da diferença de fase* para a codificação de disparidades [50], que assume que os campos receptivos das células dos olhos direito e esquerdo são descritos por um par de funções de Gabor complexas com uma envoltória comum, mas com diferentes fases. Acreditamos que essa característica do núcleo de Green complexo proporciona um certo grau de plausibilidade biológica para o nosso modelo de estimação de forma via GPM.

3.3.1 - Experimentos

Vamos apresentar aqui um estudo experimental do GPM. Dada um única imagem de entrada, I_1 , nós estimamos o mapa de profundidades das superfícies representadas, com base na fórmula

$$Z(x, y) = -\frac{k_1(1 + \gamma^2)u_0}{u_1(I_2 - k_0)} \quad (3.32)$$

que é a mesma da equação (3.14), para $K = 0$, com os parâmetros u_0 e u_1 substituindo, respectivamente, o campo de disparidades u e sua diferencial $\partial_\gamma u$, e com I_2 no lugar da imagem I . Deve-se observar que I_2 denota o par correspondente à imagem de entrada I_1 , obtida pela filtragem dessa última através da função de Green $G(x, x_0)$, de acordo com

$$I_2(x, y + \gamma x) = \int_D G(x, x_0) I_1(x_0, y + \gamma x_0) dx_0 \quad (3.33)$$

que é a generalização de (3.23) para um casamento na direção $\theta = \arctan \gamma$. Os parâmetros da função de Green, σ e a , estão relacionados a u_0 e u_1 de acordo com (3.27). Deve-se mencionar que, já que nós estamos simulando um casamento unidimensional, u_0 e u_1 – e também σ e a – podem, no caso geral, ser assumidos como variando perpendicularmente à direção de casamento, ou seja, como funções de $(y - \gamma x)$. Nos experimentos reportados aqui, esses parâmetros assumiram os valores $u_0 = 0.0018$ (que é da ordem de grandeza de um pixel, para o domínio da imagem considerado) e $u_1 = -0.0009$. Valores muito pequenos podem desestabilizar o filtro de Green, enquanto que valores muito grandes podem produzir um excesso de suavização.

Em relação à filtragem da imagem pela função de Green, é importante garantir um ganho unitário, já que a forma em (3.20) não é normalizada: Denotando por $Int(x)$ a sua integral sobre todo x_0 , para pequenos valores de σ^2/a nós obtemos

$$Int(x) \approx \frac{2a^2}{(x+a)^2 + a^2} \quad (3.34)$$

que tende a 1 para $|x| \ll a$. Assim, nós normalizamos a imagem I_2 (obtida através de (3.33)), por $\text{Int}(x)$. Como mostrado em [8], tal procedimento não altera a relação de casamento afim entre I_2 e a imagem de entrada, I_1 .

Já que estamos assumindo que o mapa de reflectância é desconhecido, também são desconhecidos os parâmetros k_0 , k_1 e também γ que é dado pela razão k_2/k_1 . Para a implementação do GPM, nós propusemos uma estratégia que torna o processo menos dependente destes parâmetros livres. Esse tratamento envolve a utilização da equação (3.32) em um processo iterativo, isto é, a partir da estimativa inicial para o mapa Z (obtida através de (3.32), para $\gamma = 0$, $k_1 = 1$ e um valor para k_0 empiricamente determinado, conforme descreveremos abaixo), nós empregamos uma regressão de mínimos quadrados, para obter os coeficientes k_0 , k_1 e k_2 do mapa de reflectância linear. Esses coeficientes são então utilizados para a obtenção de um novo k_0 e de uma nova direção de casamento e, portanto, de uma nova imagem I_2 , que, por sua vez, será utilizada na próxima estimação de $Z(x,y)$, e assim por diante. O final do processo é indicado quando não se verificam mais diferenças significativas entre as estimativas obtidas.

O procedimento iterativo pode ser resumido através da fórmula

$$Z^{n+1}(x, y) = -\frac{k_1^n [1 + (\gamma^n)^2] u_0}{u_1 (I_2^{n+1} - k_0^n)} \quad \text{para } n = 0, 1, \dots \quad (3.35)$$

onde I_2^{n+1} é obtida, a cada iteração, através da filtragem da imagem de entrada pela função de Green, ao longo da direção $\theta^n = \arctan \gamma^n$.

De forma a estimar os parâmetros do mapa de reflectância, vamos considerar que a imagem I_2 correspondente seja dividida em um certo número de regiões quadradas de m^2 pixels (com $m \geq 2$). Assumindo que, em cada uma destas regiões, aproximações lineares do tipo (3.5), em torno de uma orientação média, são válidas, nós teremos, para cada uma delas, m^2 equações em que as intensidades são conhecidas. A partir do mapa de profundidades corrente, Z , nós podemos obter, por simples derivação, uma estimativa das orientações (p e q) em cada ponto. Dessa forma, apenas os coeficientes lineares nas equações são incógnitas. Adotando uma numeração seqüencial para os pixels ($j = 1, 2, \dots, m^2$) em uma determinada região, nós obteremos

$$I_2(j) = k_0 + k_1 p(j) + k_2 q(j); \quad j = 1, 2, \dots, m^2 \quad (3.36)$$

As equações (3.36) nos proporcionam um modelo eficiente para a estimação dos coeficientes lineares k_0, k_1 e k_2 na região considerada, através do método de regressão por mínimos quadrados [51]. Assim, explicitando as equações (3.36), nós teremos

$$\begin{aligned}
 I_2(1) &= k_0 + k_1 p(1) + k_2 q(1) \\
 I_2(2) &= k_0 + k_1 p(2) + k_2 q(2) \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \\
 I_2(m^2) &= k_0 + k_1 p(m^2) + k_2 q(m^2)
 \end{aligned} \tag{3.37}$$

E reescrevendo o sistema (3.37) na forma matricial

$$\begin{bmatrix} I_2(1) \\ I_2(2) \\ \cdot \\ \cdot \\ \cdot \\ I_2(m^2) \end{bmatrix} = \begin{bmatrix} 1 & p(1) & q(1) \\ 1 & p(2) & q(2) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & p(m^2) & q(m^2) \end{bmatrix} \cdot \begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix} \tag{3.38}$$

Se denotarmos a matriz das componentes p e q por β , e denominarmos β^T a transposta de β , os valores de k_0, k_1 e k_2 podem ser encontrados a partir de

$$\begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix} = (\beta^T \cdot \beta)^{-1} \cdot \beta^T \cdot \begin{bmatrix} I_2(1) \\ I_2(2) \\ \cdot \\ \cdot \\ \cdot \\ I_2(m^2) \end{bmatrix} \tag{3.39}$$

Tal procedimento nos fornece estimativas para os coeficientes do mapa de reflectância requeridos a cada iteração, correspondendo a aproximações lineares em torno da orientação média da superfície, em cada uma das regiões consideradas.

Para o emprego na determinação de profundidades, através da equação (3.35), nós calculamos os k 's em todas as janelas ao longo da imagem e utilizamos a sua média aritmética como as nossas estimativas desses coeficientes. Tal procedimento é computacionalmente mais eficiente do que realizar o processo uma única vez, ao longo de toda a imagem.

De forma a escolher um valor apropriado para o k_0 a ser utilizado na estimativa inicial $Z^0(x,y)$, nós realizamos um estudo empírico sobre uma série de imagens com mapas de profundidades conhecidos, cujos resultados são apresentados na Figura 3.14. Nós consideramos duas medidas para a avaliação das superfícies estimadas: o erro médio do gradiente [38] e o erro médio em profundidades. O erro do gradiente foi tomado como a média, ao longo de toda a superfície, da medida

$$|p - p^r| + |q - q^r| \quad (3.40)$$

onde (p,q) e (p^r, q^r) representam, respectivamente, as orientações calculadas em cada ponto sobre a superfície estimada e a real. Os mapas de orientações foram obtidos utilizando-se diferenças à direita, para o cálculo das derivadas, sendo o mapa estimado de profundidades previamente normalizado pelo mapa real conhecido. Já o erro em profundidades [52] foi calculado como a média das diferenças absolutas entre o mapa real e o mapa estimado, ajustado para origem e escala. Também nesse caso, o mapa estimado foi normalizado pelo mapa real, antes do procedimento de ajuste.

As duas medidas de erro produziram resultados similares, como mostra a Figura 3.14: exceto por uma pequena faixa de valores, aproximadamente de $k_0 = -0.2$ a $k_0 = 1$, a qualidade das nossas reconstruções (obtidas para $\gamma = 0$ e $k_1 = 1$) não são muito sensíveis ao valor de k_0 utilizado, sendo que os melhores resultados, em média, foram obtidos para k_0 em torno de -0.5 . Este último valor foi, então, escolhido para a estimativa inicial do nosso processo iterativo. No entanto, nós observamos que os resultados finais eram também essencialmente independentes do k_0 inicial, desde que a faixa de valores problemática fosse evitada.

Implementação

Antes de apresentar nossos experimentos, ilustramos, a seguir, em pseudo-código, um esquema para implementação do GPM:

Início

- Entre com a imagem I_1 ;
- Defina os valores para os parâmetros σ e a da função de Green, de acordo com os valores desejados para u_0 e u_1 (equação (3.27));
- Defina os valores iniciais para os parâmetros k_0 , k_1 e k_2 (Em geral, tomamos $k_0=-0.5$, $k_1=1.0$, $k_2=0$ e, portanto, $\gamma=0$);
- Faça: $\gamma \leftarrow k_2 / k_1$;
- Defina o tamanho m das janelas, para estimação dos parâmetros k_i ($i = 0, 1$ e 2);
- Obtenha a imagem I_2 inicial, filtrando a imagem I_1 através da Função de Green, de acordo com as equações (3.20) e (3.33), ao longo da direção $\theta = \arctan \gamma$;
- Normalize a imagem I_2 pelo fator em (3.34);
- Obtenha o mapa de profundidades Z (inicial) através de (3.32);
- Faça: $N \leftarrow$ Número de janelas de tamanho $m \times m$ da imagem I_2 ;

Repita

- Faça: $Z_{ant} \leftarrow Z$;
- Obtenha os mapas para as componentes p e q do Gradiente da superfície estimada, diferenciando o mapa Z com relação a x e y ;
- Inicie variáveis de acumulação, S_0 , S_1 e S_2 , com zero;

Para cada janela j , de tamanho $m \times m$, da imagem I_2 faça

- Obtenha a matriz β (das componentes p e q) para a janela, de acordo com (3.38);
- Obtenha e matriz inversa e a matriz transposta de β ;
- Obtenha os coeficientes k_{0j} , k_{1j} e k_{2j} , de acordo com (3.39);

Variando i de 0 a 2 faça

- Faça: $S_i \leftarrow S_i + k_{ij}$;

Fim_Variando.**Fim_Para****Variando i de 0 a 2 faça**

- Faça: $k_i \leftarrow S_i / N$;

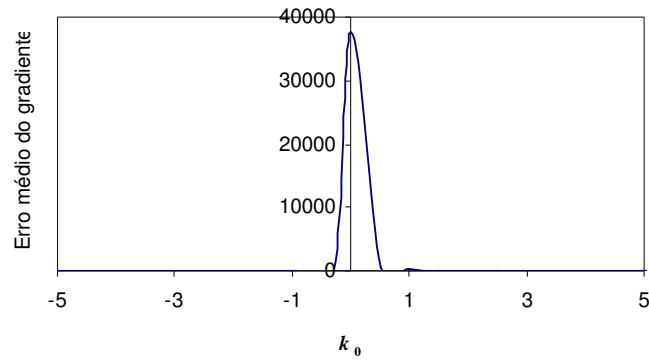
Fim_Variando.

- Faça: $\gamma \leftarrow k_2 / k_1$;
- Obtenha a nova imagem I_2 , filtrando a imagem I_1 através da Função de Green, de acordo com as equações (3.20) e (3.33), ao longo da direção $\theta = \arctan \gamma$;
- Normalize a imagem I_2 pelo fator em (3.34);
- Obtenha a nova estimativa para o mapa Z através de (3.32);
- Compare, através de diferenças absolutas, os mapas Z e Z_{ant} ;

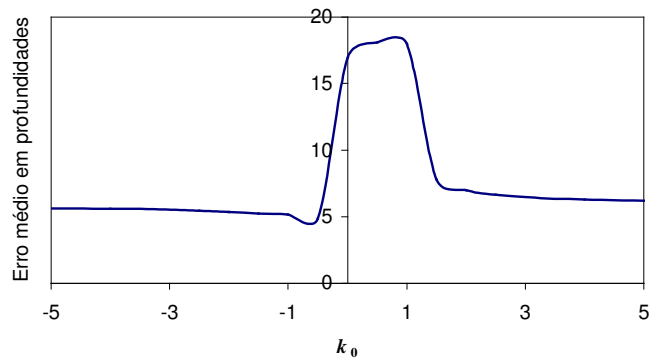
Até que não haja diferenças significativas entre os mapas Z e Z_{ant} .

- Retorne o mapa Z ;

Fim



(a)



(b)

Figura 3.14. Erro médio Gradiente (a) e Erro Médio em Profundidades (b), como funções do parâmetro k_0 , obtidos a partir de uma série de experimentos com um conjunto variado de imagens de superfícies lambertianas e quase-lambertianas. Utilizou-se $k_1 = 1$ e $\gamma = 0$.

A. Imagens Sintéticas.

Nós realizamos uma série de experimentos utilizando imagens sintéticas de superfícies conhecidas, com o objetivo de avaliar quantitativamente a qualidade das estimativas. No Capítulo 5 desse trabalho, nós mostraremos os resultados obtidos para todo o conjunto dessas imagens, em geral envolvendo três direções de iluminação distintas para cada superfície. Por ora, apresentaremos a análise de um subconjunto delas (Figura 3.15): a imagem do Mozart, com duas direções de iluminação distintas ($(0,0,1)$ – *Mozart1* e $(1,0,1)$ – *Mozart2*), e as imagens *Vase1*, *Penny* e *Mask1*, com direções de iluminação $(5,5,7)$, $(1,0,1)$ e $(0,0,1)$, respectivamente. Para uma avaliação comparativa dos nossos resultados, nós apresentamos, ainda, a análise quantitativa das reconstruções produzidas por dos algoritmos conhecidos de SFS: o de Tsai e Shah [38] e o de Lee e Rosenfeld [37]. Esses dois algoritmos, assim como o

GPM, são abordagens locais para SFS, sendo que o primeiro também assume uma função de reflectância linear. As duas abordagens são ainda tão facilmente implementáveis como o GPM. Por outro lado, ambas são explicitamente dependentes de informações sobre o mapa de reflectância, assumindo reflectância lambertiana e necessitando da direção de iluminação. O GPM, em contraste, não requer tais informações.

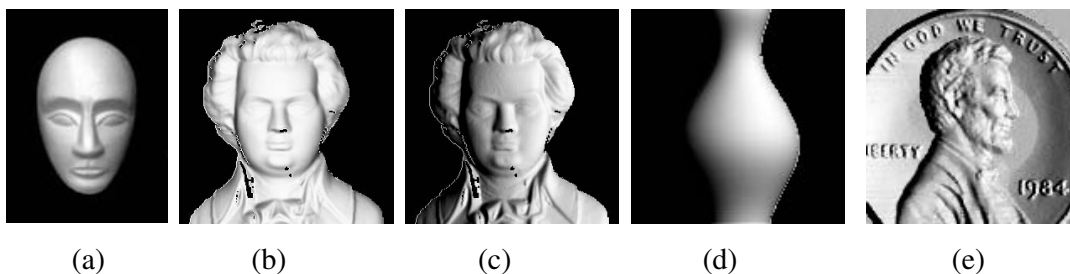


Figura 3.15. Imagens utilizadas para o estudo comparativo experimental do GPM: (a) *Mask1*, (b) *Mozart1*, (c) *Mozart2*, (d) *Vase1* e (e) *Penny*; para todas as imagens o mapa real de profundidades é conhecido.

Na Figura 3.16, nós mostramos os resultados quantitativos, obtidos com as três abordagens, em termos do erro médio gradiente e do erro médio em profundidades. Para o GPM, nós informamos as medidas de erro iniciais (identificadas como *Entrada*), que correspondem às estimativas de profundidade $Z^0(x,y)$, e as medidas de erro obtidas no final do processo iterativo (identificadas como *Saída*). Podemos observar que as estimativas iniciais já produzem resultados competitivos com aqueles de Tsai/Shah e Lee/Rosenfeld. Mesmo considerando que o principal ganho obtido, com a introdução de nosso processo iterativo, é a independência em relação ao conhecimento do verdadeiro mapa de reflectância das superfícies, a introdução dessa metodologia também possibilitou ganhos na precisão das reconstruções, como podemos avaliar através da Figura.

As Figuras 3.17 a 3.28 apresentam reconstruções produzidas pelo GPM a partir das imagens sintéticas. Nas mesmas figuras, mostramos também os resultados obtidos através dos algoritmos de Tsai/Shah e Lee/Rosenfeld. Como ilustrado nas Figuras 3.17 a 3.22, que se referem às imagens *Mask2* e *Mask1*, as reconstruções obtidas com o método de Lee/Rosenfeld mostraram-se visualmente competitivas com as nossas, enquanto que as estimativas produzidas por Tsai/Shah apresentaram-se muito ruidosas. A situação inversa ocorre para a imagem *Mozart2* (Figuras 3.23 a 3.25), onde podemos constatar que a estimativa produzida por Lee/Rosenfeld não se apresenta bem definida (embora o resultado numérico em 3.16 (b) não comprove exatamente essa situação, já que, na tabela, o erro médio em profundidades,

obtido por Lee/Rosenfeld para a imagem *Mozart2*, mostrou-se ligeiramente inferior ao do GPM). Como regra geral, as reconstruções obtidas através de Lee/Rosenfeld mostraram-se bastante sensíveis à direção de iluminação. Em nossos experimentos, essa abordagem somente proporcionou boas reconstruções para imagens com direção de iluminação central ou direções bem próximas a essa.

De forma geral, considerando a avaliação quantitativa mostrada na Figura 3.16 e os resultados ilustrados, podemos concluir que o GPM comportou-se melhor do que as duas abordagens aqui destacadas, com relação às imagens sintéticas.

<i>Imagens</i>	<i>Método</i>			
	<i>GPM</i>		<i>Lee/Rosenfeld</i>	<i>Tsai/Shah</i>
	<i>Entrada</i>	<i>Saída</i>		
<i>Mask1</i>	1,0	0,9	1,0	1,0
<i>Mozart1</i>	1,9	1,6	1,6	1,6
<i>Mozart2</i>	2,2	2,1	4,5	2,1
<i>Vase1</i>	0,9	0,6	0,7	0,6
<i>Penny</i>	1,3	1,1	2,2	1,6

(a)

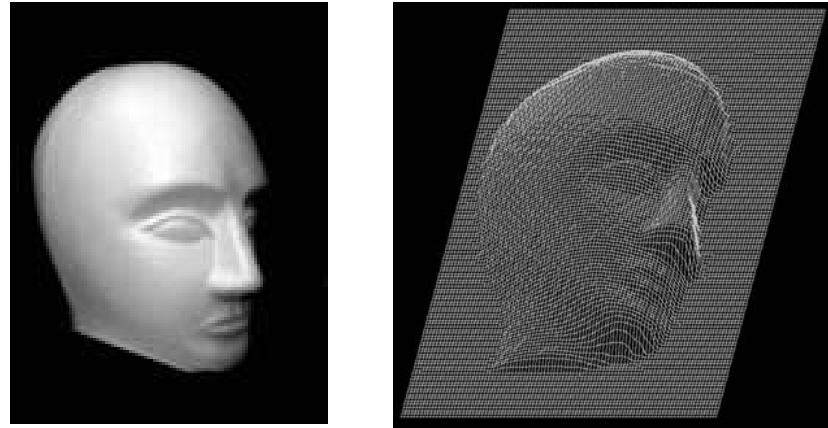
<i>Imagens</i>	<i>Método</i>			
	<i>GPM</i>		<i>Lee/Rosenfeld</i>	<i>Tsai/Shah</i>
	<i>Entrada</i>	<i>Saída</i>		
<i>Mask1</i>	7,7	5,4	8,0	7,9
<i>Mozart1</i>	7,6	7,1	7,2	7,6
<i>Mozart2</i>	16,6	12,7	12,3	15,9
<i>Vase1</i>	10,1	5,1	6,5	5,8
<i>Penny</i>	4,2	4,2	4,2	4,2

(b)

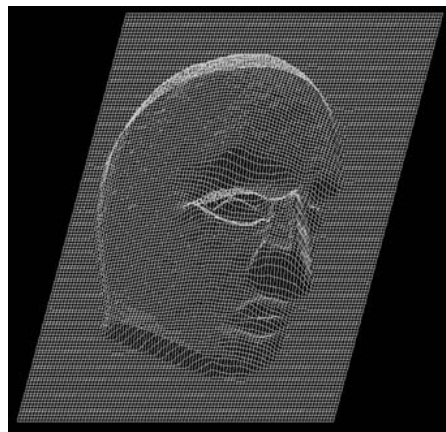
Figura 3.16. Análise quantitativa dos erros. (a): Erro médio gradiente das estimativas produzidas pelos três métodos. (b): O mesmo que em (a), para o erro médio em profundidades.

B. Imagens Reais.

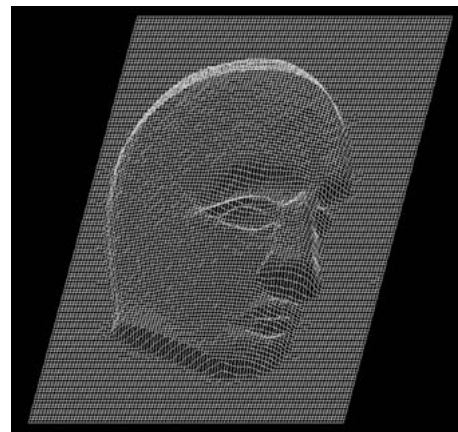
As Figuras 3.29 a 3.41 apresentam reconstruções produzidas pelo GPM a partir de imagens reais. Nós consideramos superfícies constituídas de diferentes materiais e imagens capturadas sob iluminação complexa. De forma a permitir a melhor avaliação das estimativas, nos casos das Figuras 3.29 a 3.37, nós mostramos também as reconstruções obtidas através dos algoritmos de Tsai/Shah e Lee/Rosenfeld. Deve-se observar que as imagens *Lena* e *Pepper* foram também consideradas no trabalho em [19]. Comparando nossas reconstruções com aquelas ilustradas no referido artigo, podemos constatar a superioridade das estimativas obtidas pelo GPM, principalmente no caso da *Lena*. A imagem *Manequim* também foi objeto de estudo em [38,53] e a imagem *Paolina* em [53]. Em ambos os casos as nossas reconstruções mostraram-se bem superiores às estimativas apresentadas nos citados artigos. Como regra geral, verificamos que o GPM produz estimativas superiores quando comparado aos seus precursores, o *Green's Function Shape from Shading* e nossa abordagem para o Movimento Fotométrico multi-imagem.



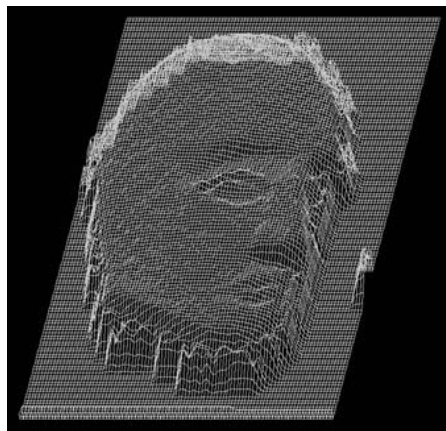
(a)



(b)



(c)



(d)

Figura 3.17. (a): Imagem de entrada (*Mask2*) e representação correspondente do mapa real de profundidades. (b),(c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Tsai/Shah e Lee/Rosenfeld, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.

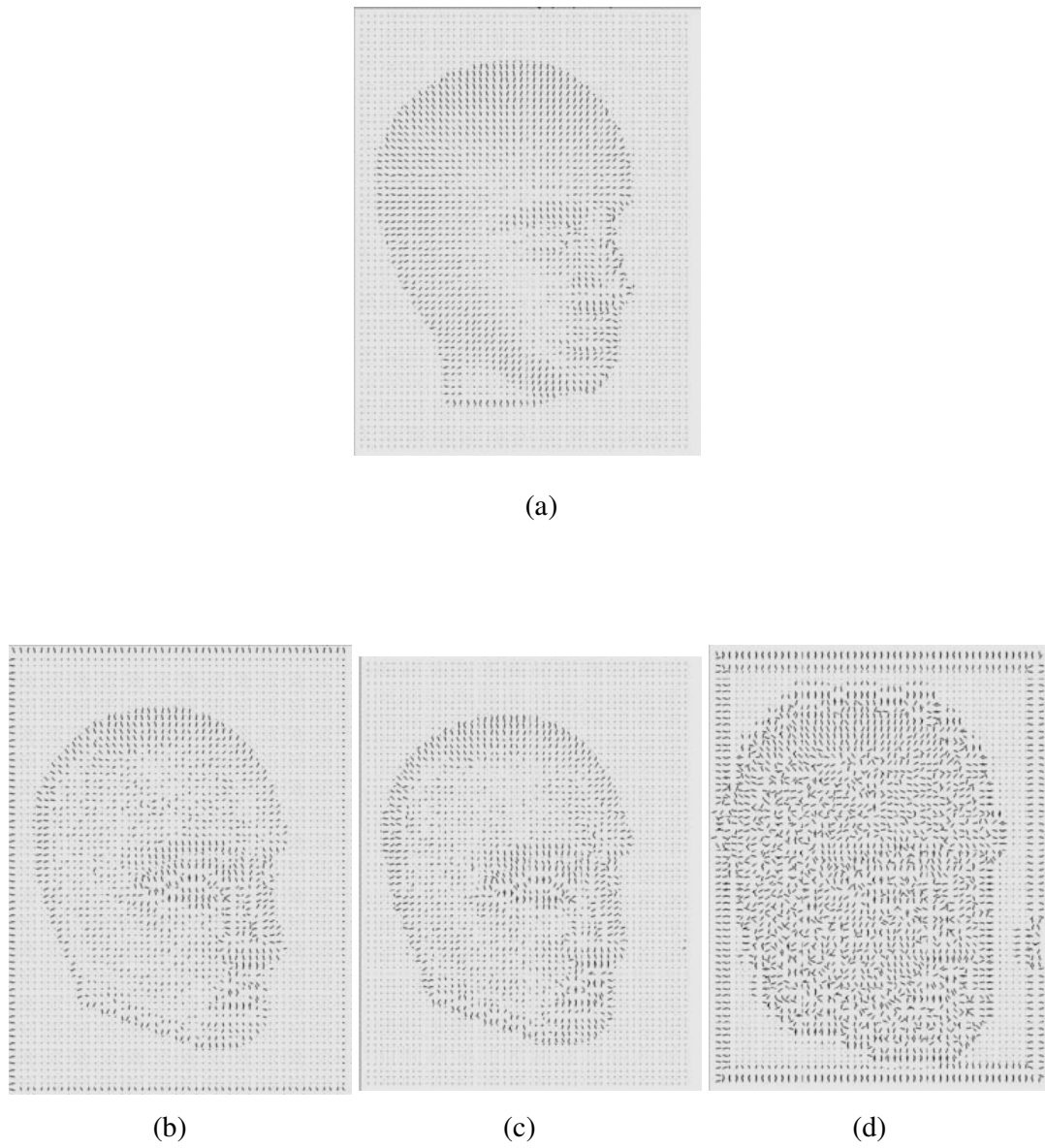


Figura 3.18. Mapas de orientações. (a): Obtido em relação às profundidades reais. (b), (c) e (d): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente .

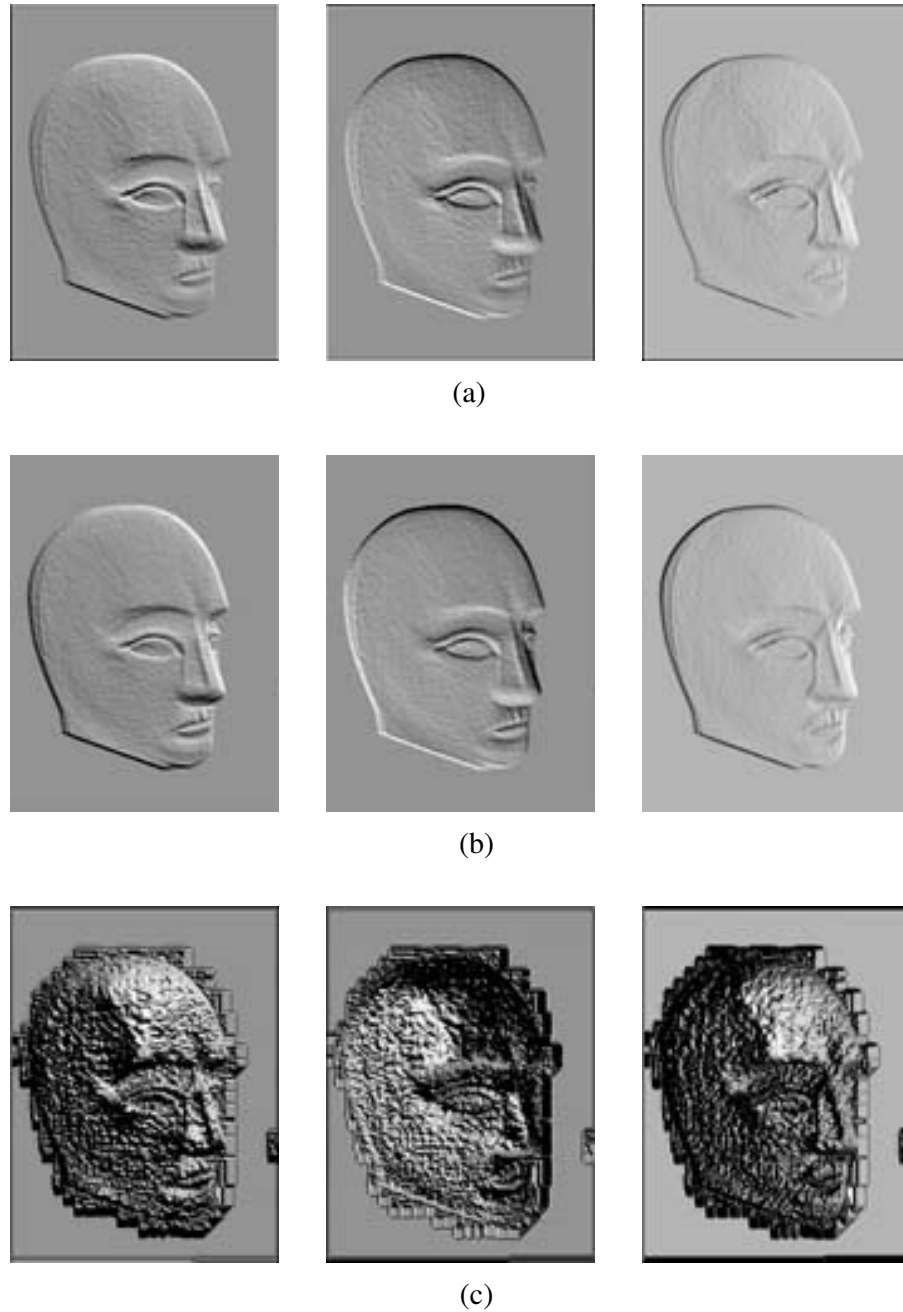
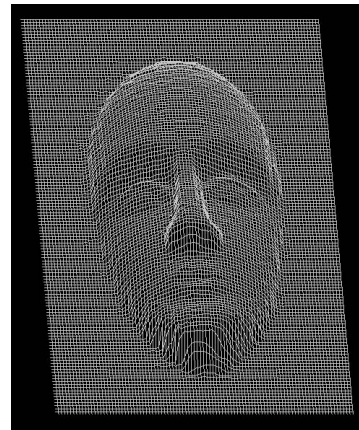
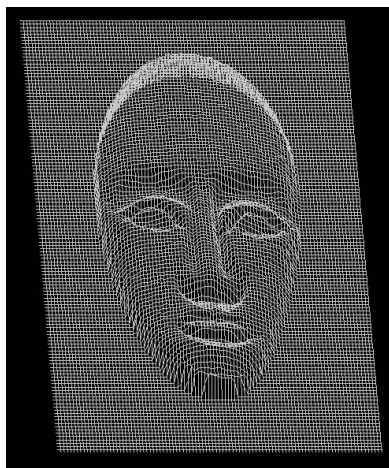


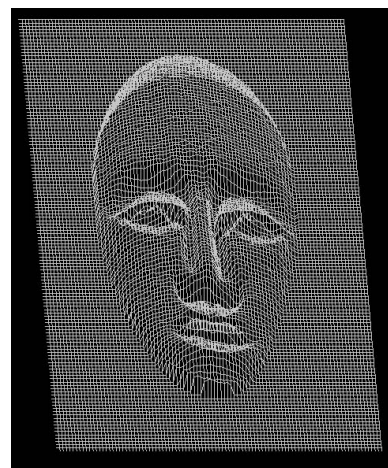
Figura 3.19. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.17), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Lee/Rosenfeld e (c): Tsai/Shah.



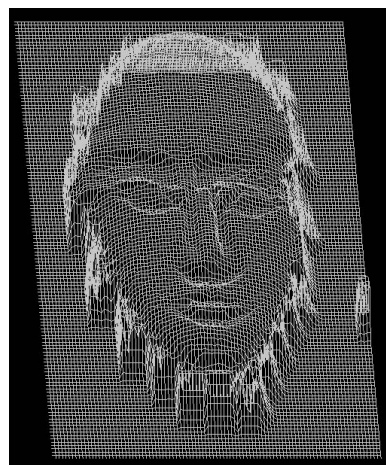
(a)



(b)

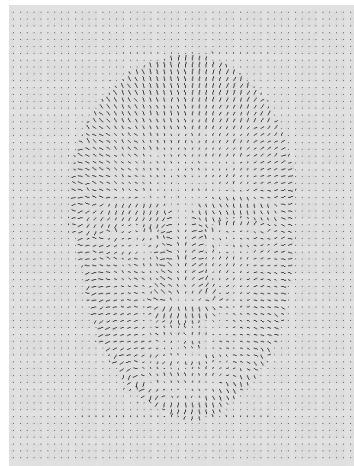


(c)

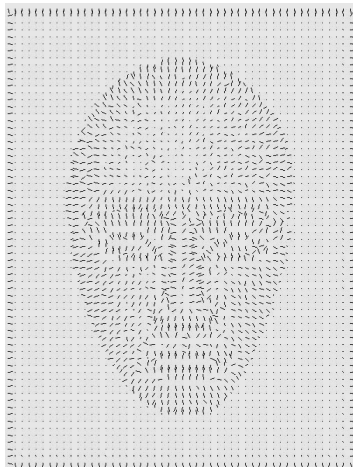


(d)

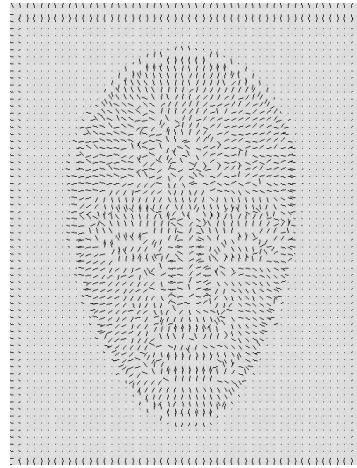
Figura 3.20. (a): Imagem de entrada (*Mask1*) e representação correspondente do mapa real de profundidades. (b),(c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Tsai/Shah e Lee/Rosenfeld, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.



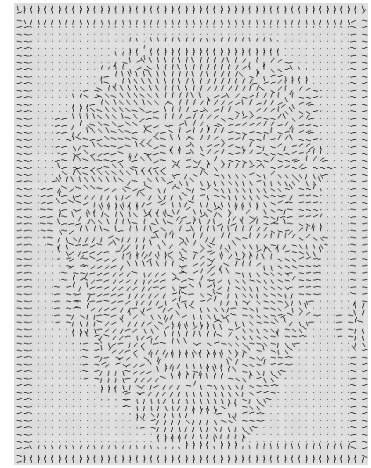
(a)



(b)



(c)



(d)

Figura 3.21. Mapas de orientações. (a): Obtido em relação às profundidades reais. (b), (c) e (d): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente .

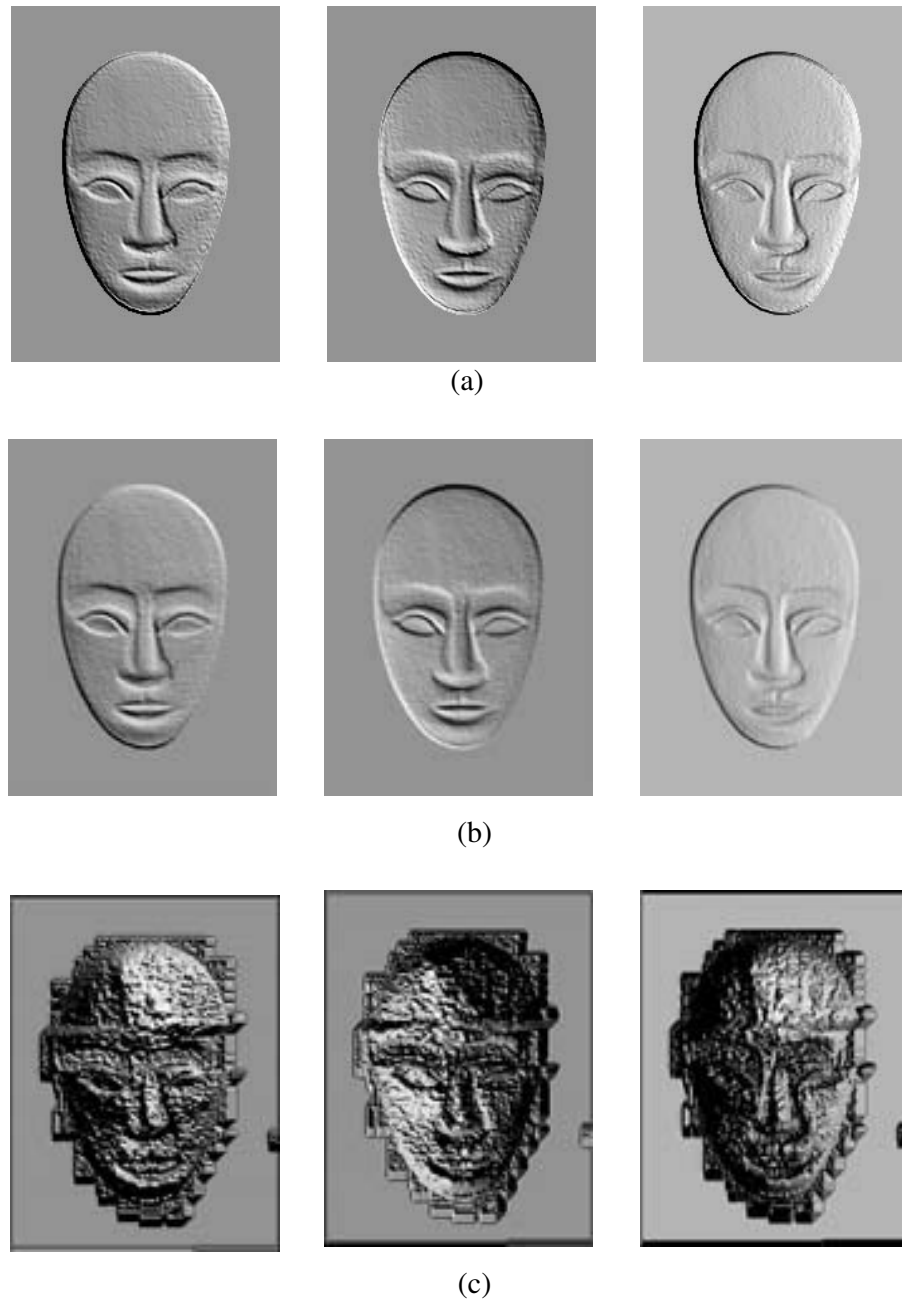
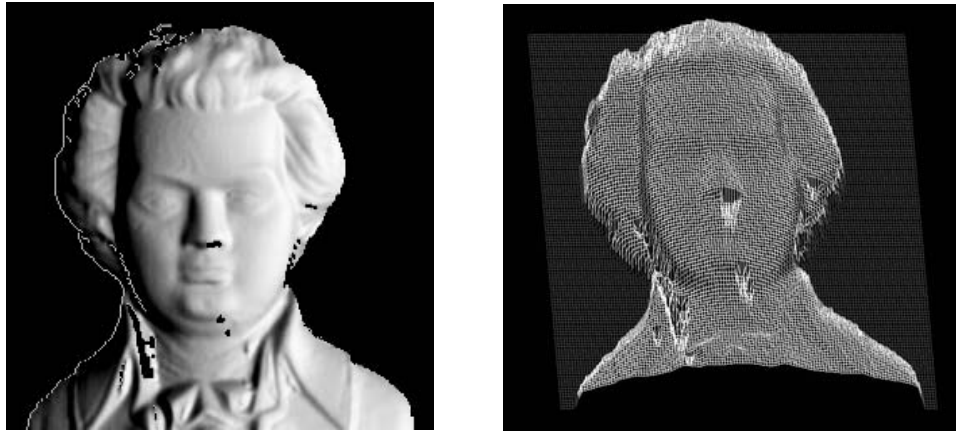
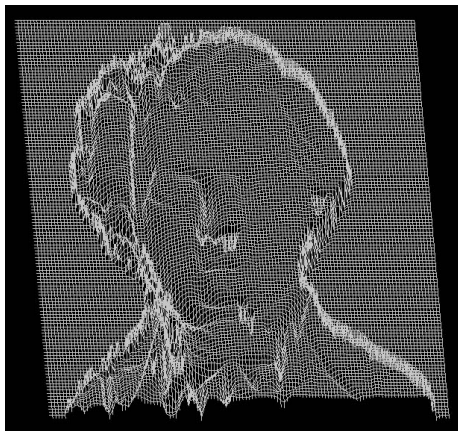


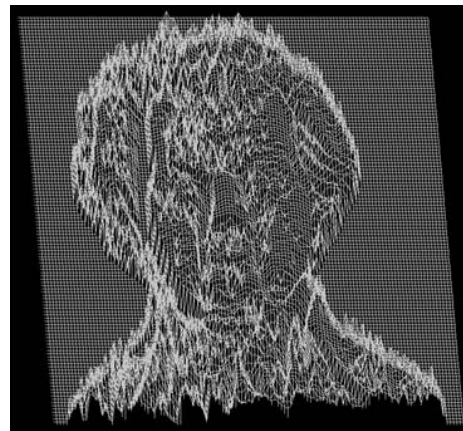
Figura 3.22. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.20), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Lee/Rosenfeld e (c): Tsai/Shah.



(a)



(b)

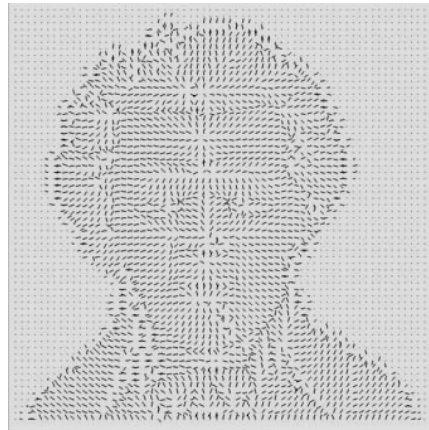


(c)

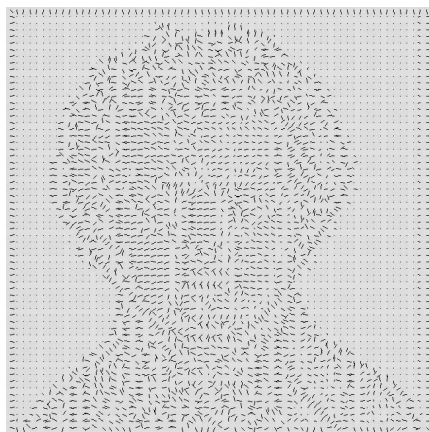


(d)

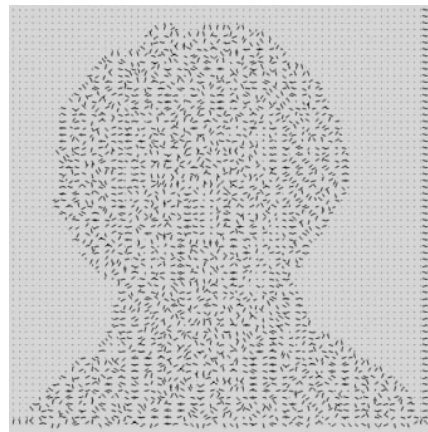
Figura 3.23. (a): Imagem de entrada (*Mozart2*) e representação correspondente do mapa real de profundidades. (b), (c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.



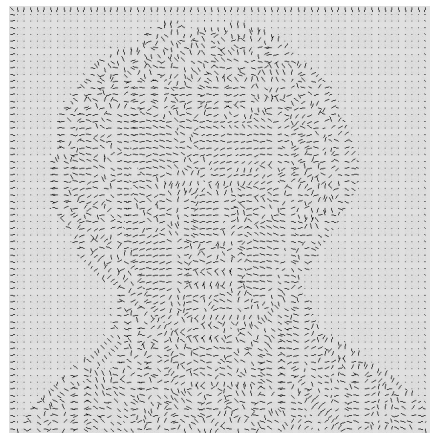
(a)



(b)

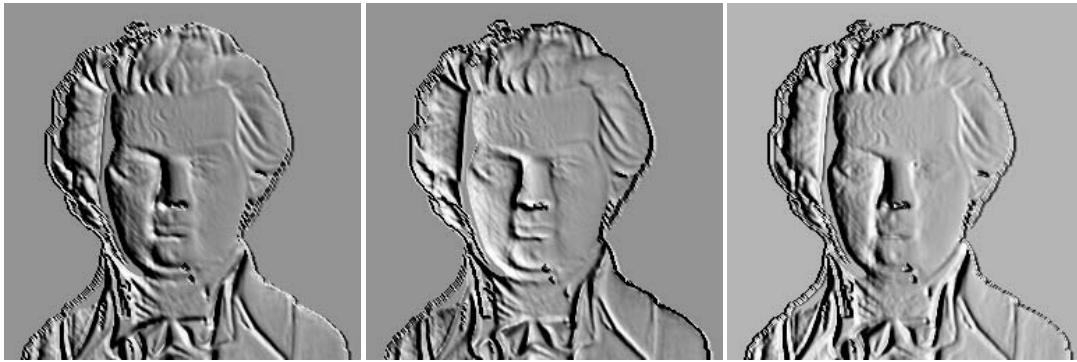


(c)

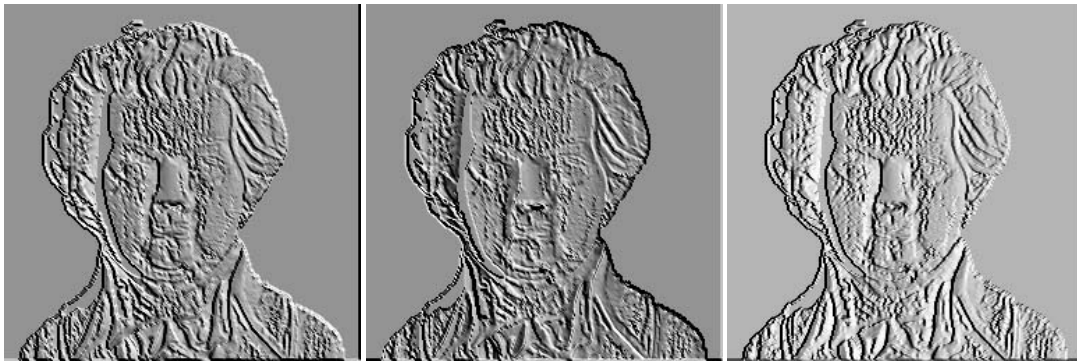


(d)

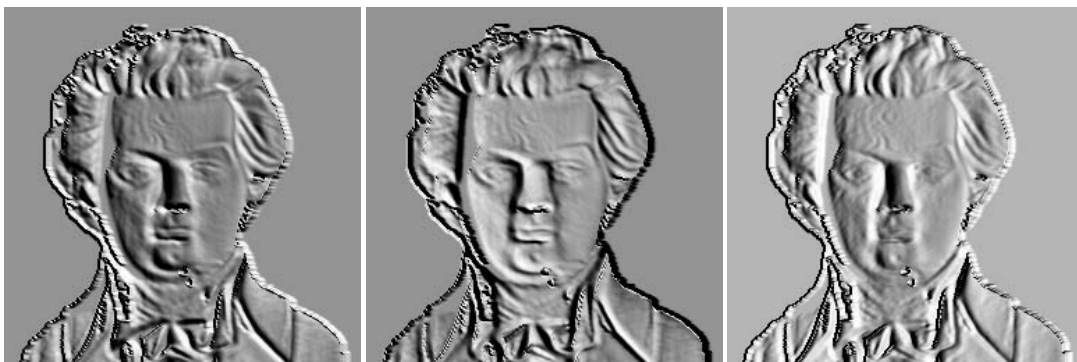
Figura 3.24. Mapas de orientações. (a): Obtido em relação às profundidades reais. (b), (c) e (d): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente .



(a)



(b)



(c)

Figura 3.25. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.23), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Método de Lee/Rosenfeld e (c): Método de Tai/Shah.

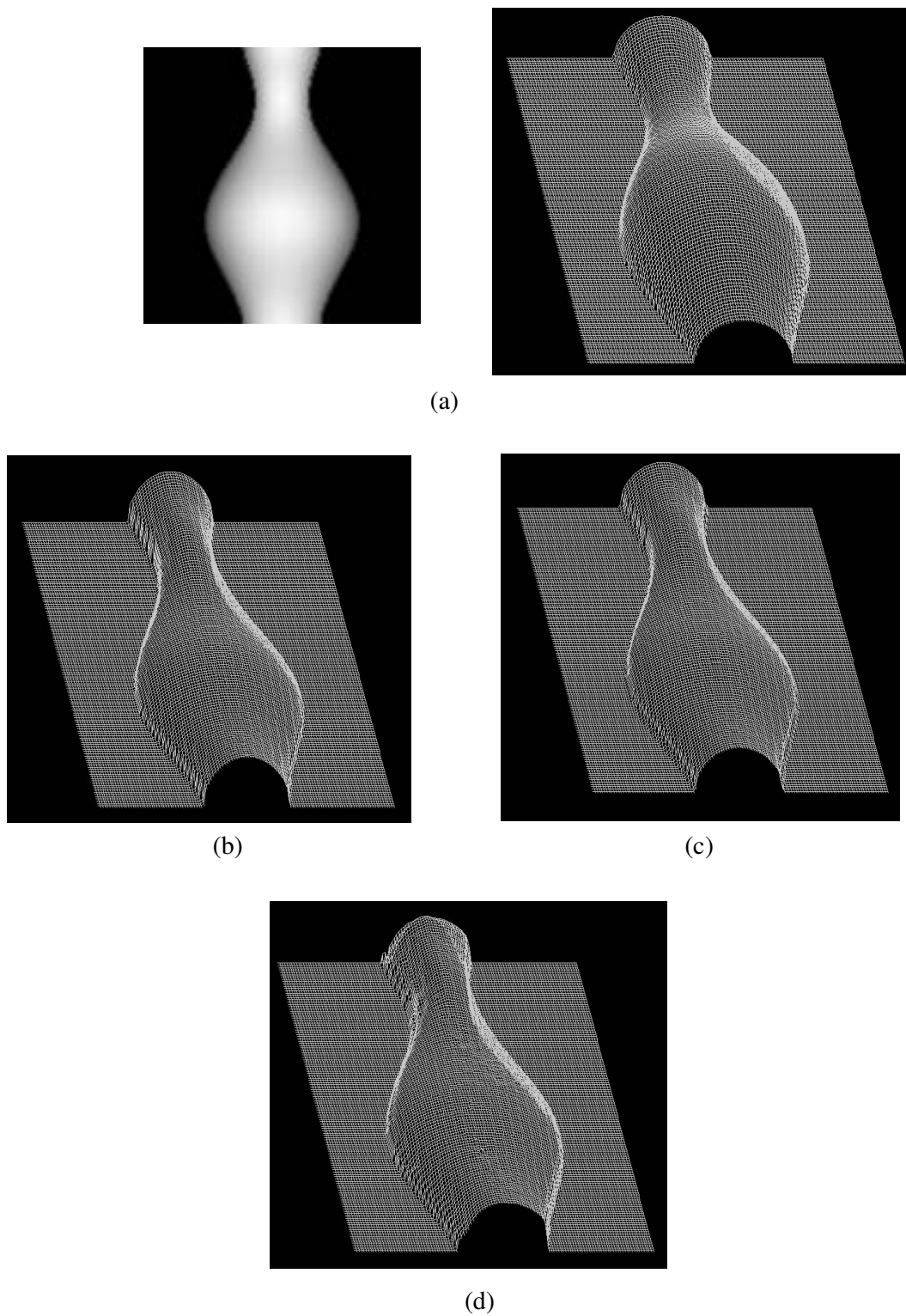


Figura 3.26. (a): Imagem de entrada (*Vase*) e representação correspondente do mapa real de profundidades. (b), (c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.

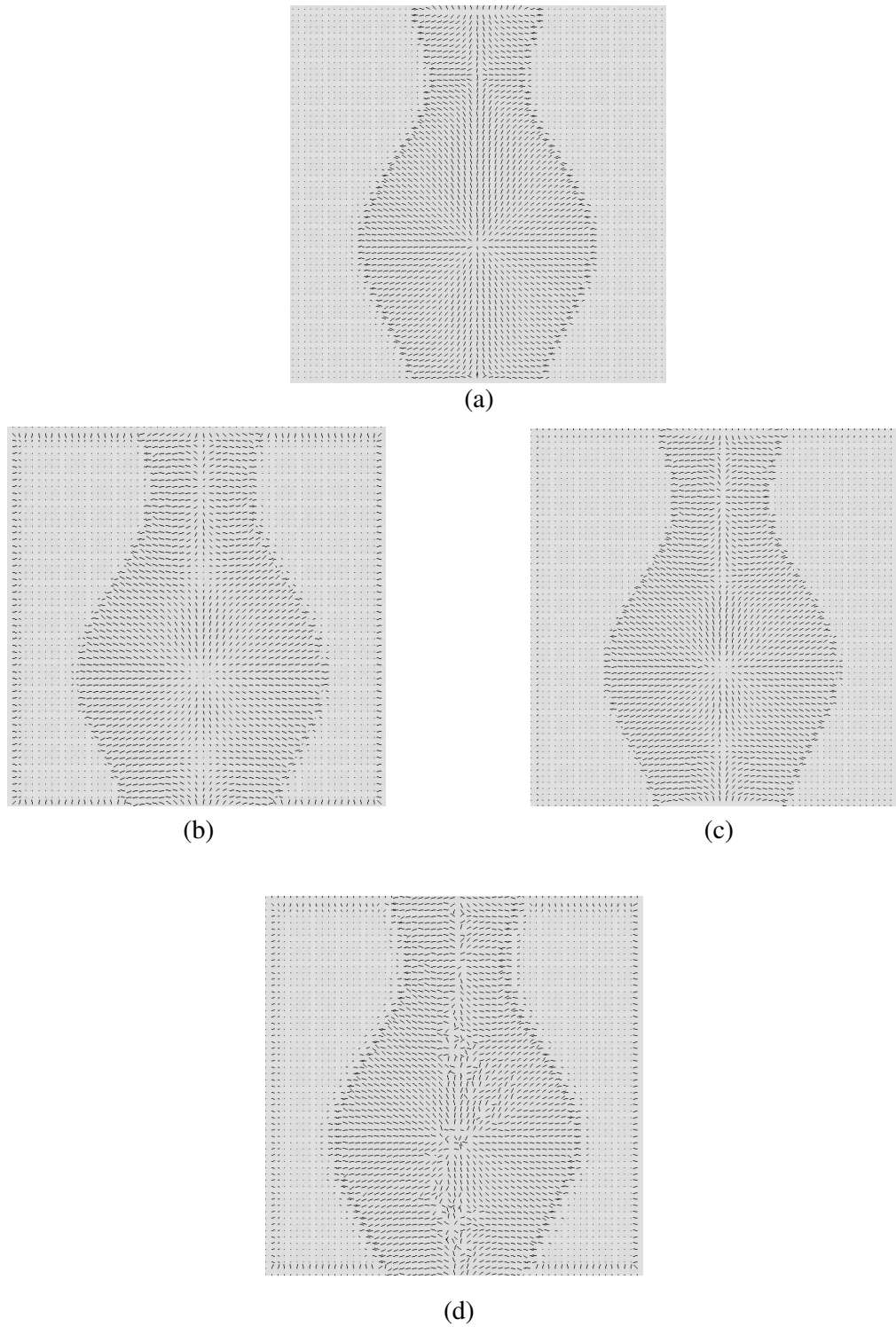
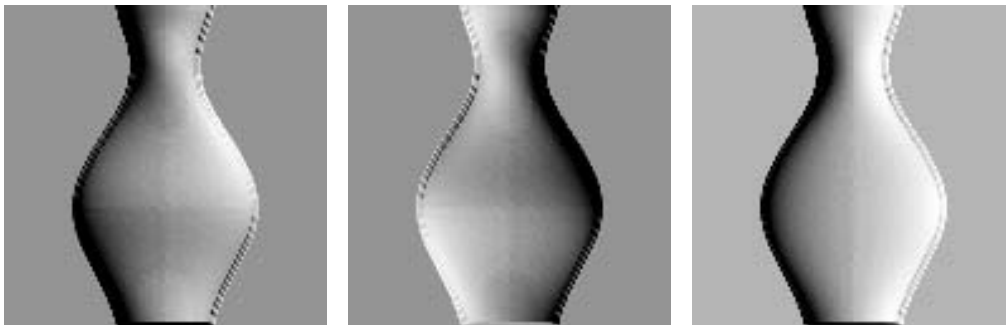


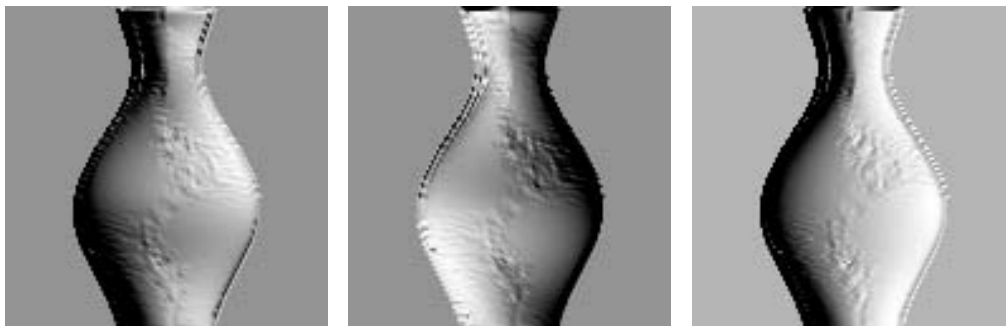
Figura 3.27. Mapas de orientações. (a): Obtido em relação às profundidades reais. (b), (c) e (d): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente .



(a)



(b)



(c)

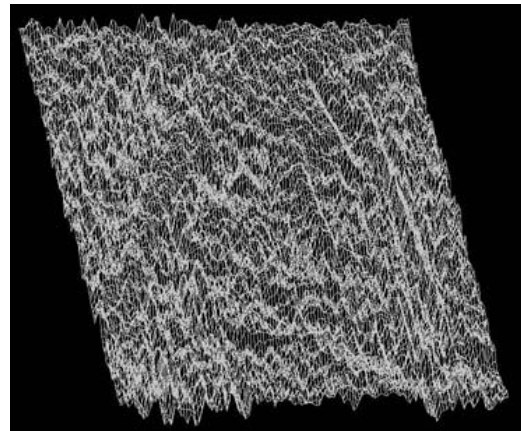
Figura 3.28. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.26), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Método de Lee/Rosenfeld e (c): Método de Tai/Shah.



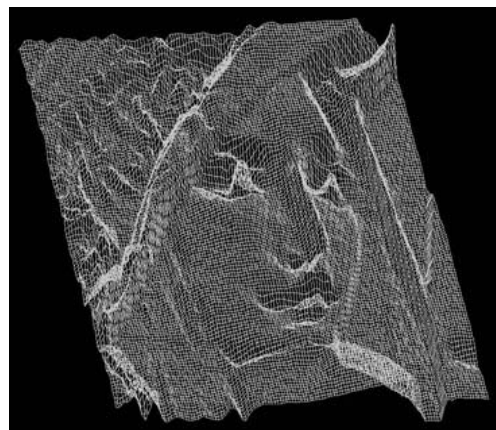
(a)



(b)

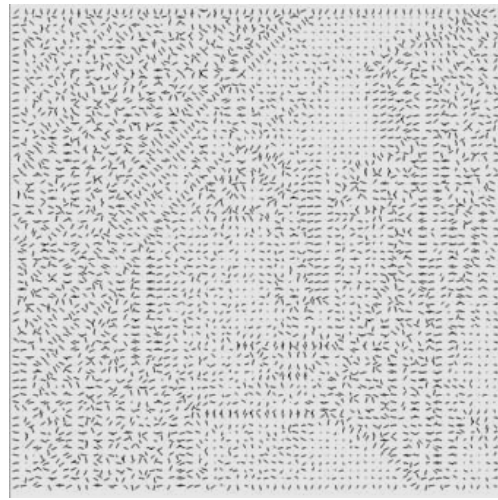


(c)

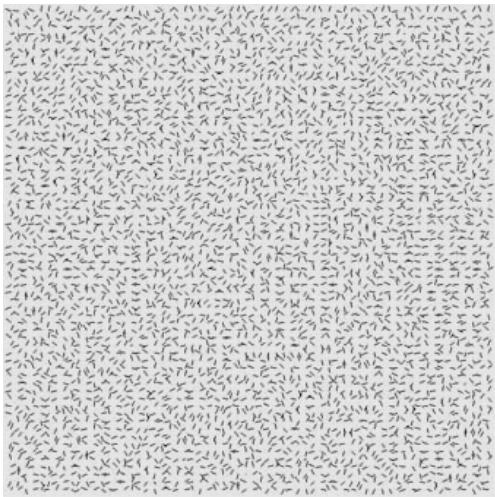


(d)

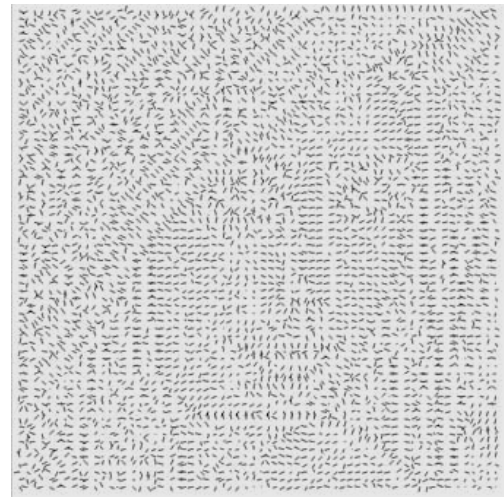
Figura 3.29. (a): Imagem de entrada (*Lena*). (b), (c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.



(a)



(b)



(c)

Figura 3.30. Mapas de orientações. (a), (b) e (c): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente.



(a)

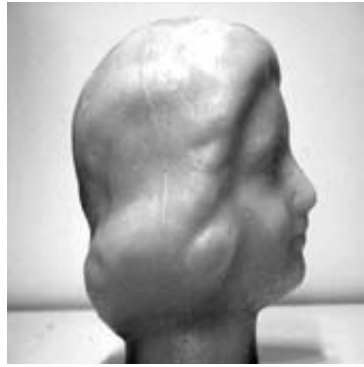


(b)

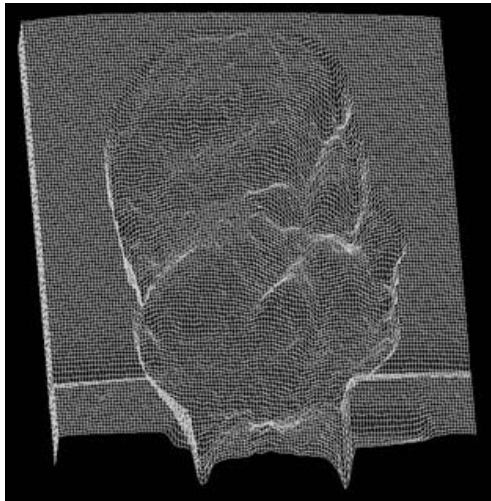


(c)

Figura 3.31. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.29), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Lee/Rosenfeld e (c): Tsai/Shah.



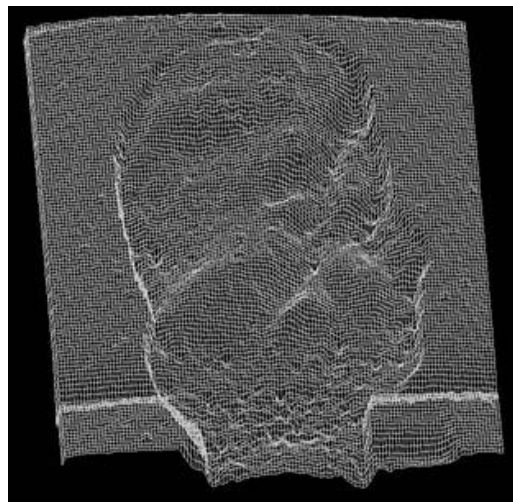
(a)



(b)

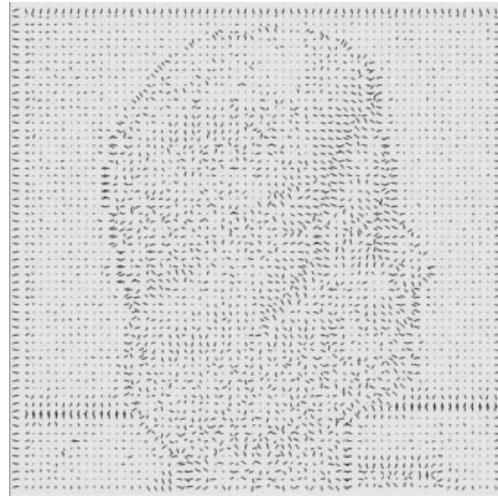


(c)

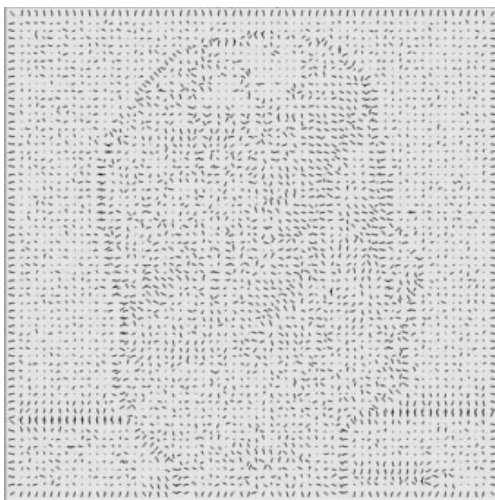


(d)

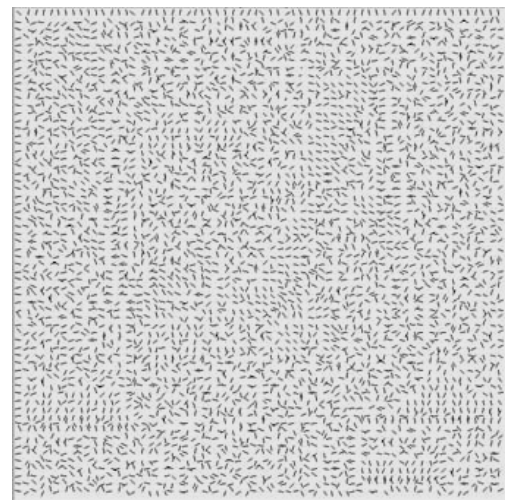
Figura 3.32. (a): Imagem de entrada. (*Head1*) (b),(c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.



(a)



(b)



(c)

Figura 3.33. Mapas de orientações. (a), (b) e (c): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente .

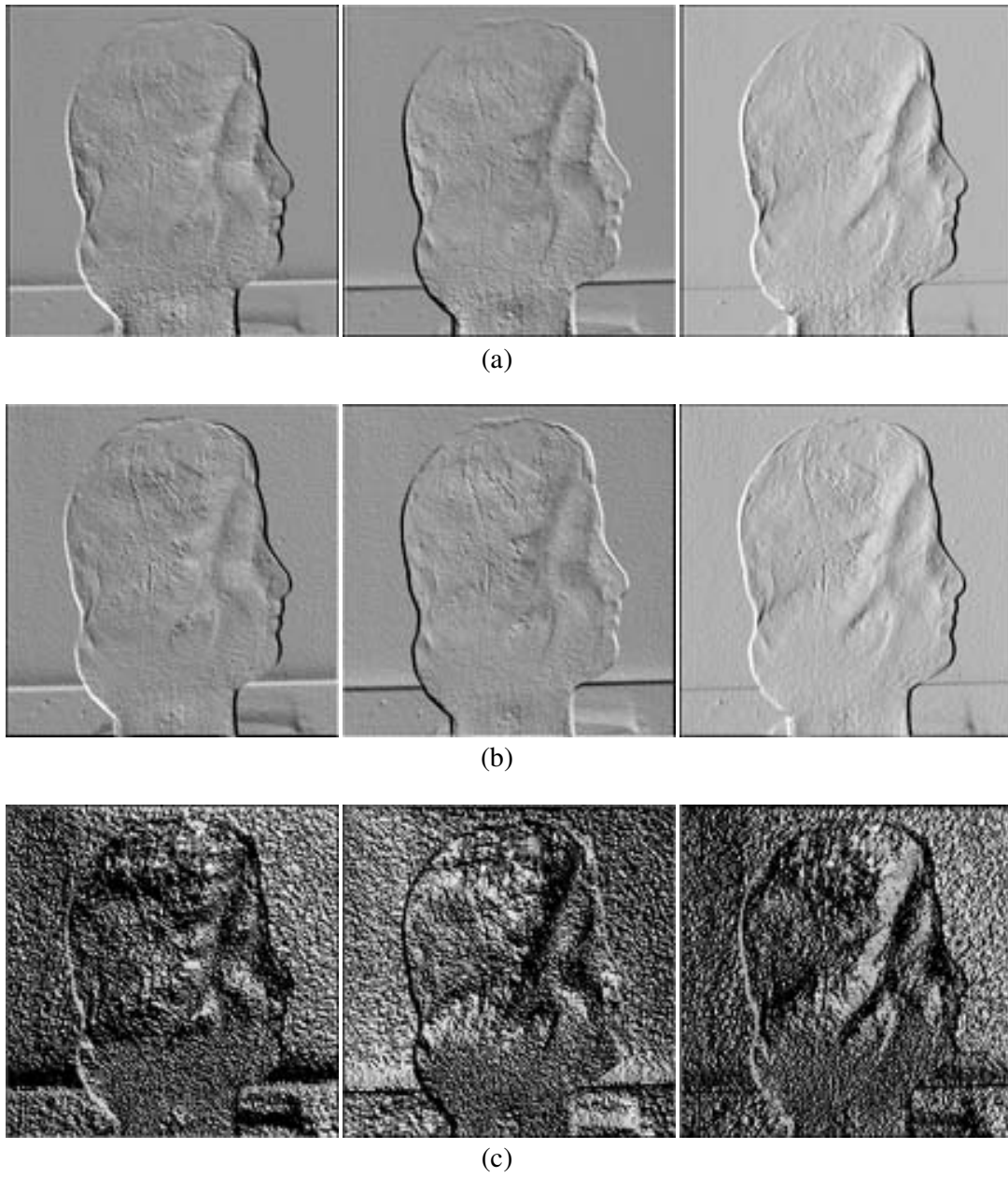
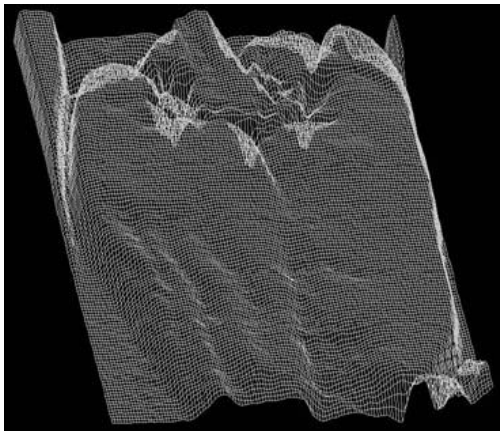


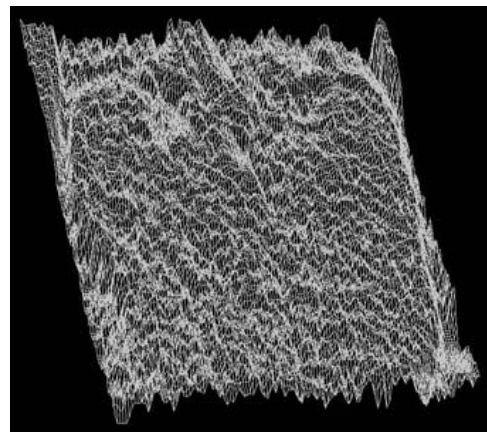
Figura 3.34. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.32), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Lee/Rosenfeld e (c) Tsai/Shah.



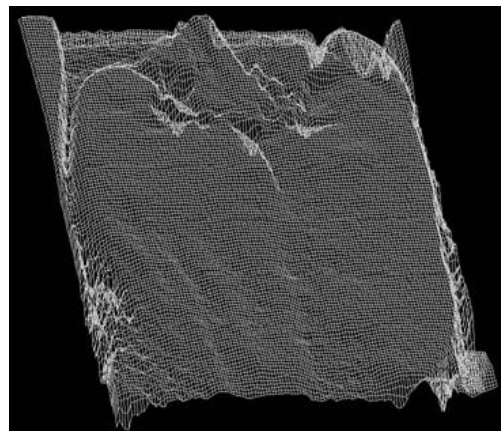
(a)



(b)

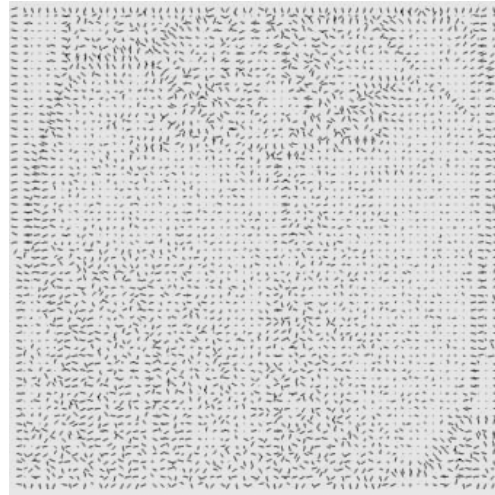


(c)

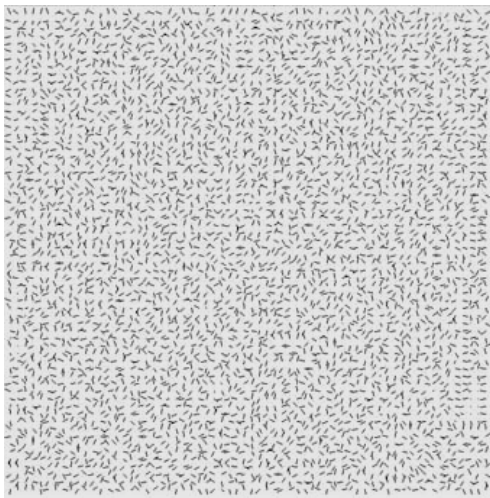


(d)

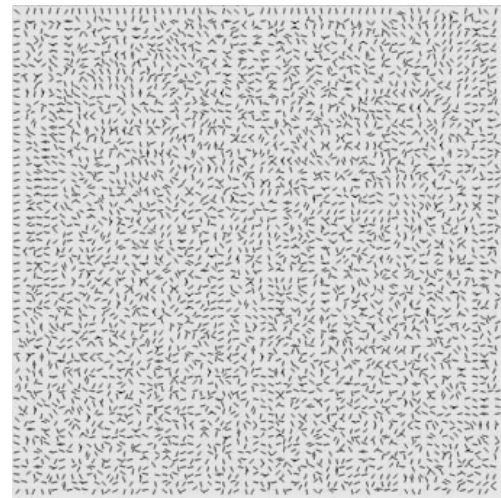
Figura 3.35. (a): Imagem de entrada (*Pepper*). (b), (c) e (d): Representações para os mapas estimados através do GPM e através dos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente. Parâmetros do GPM: $\sigma = 0.06$ e $a = 2$.



(a)



(b)



(c)

Figura 3.36. Mapas de orientações. (a), (b) e (c): Obtidos em relação às superfícies estimadas pelo GPM e pelos métodos de Lee/Rosenfeld e Tsai/Shah, respectivamente.

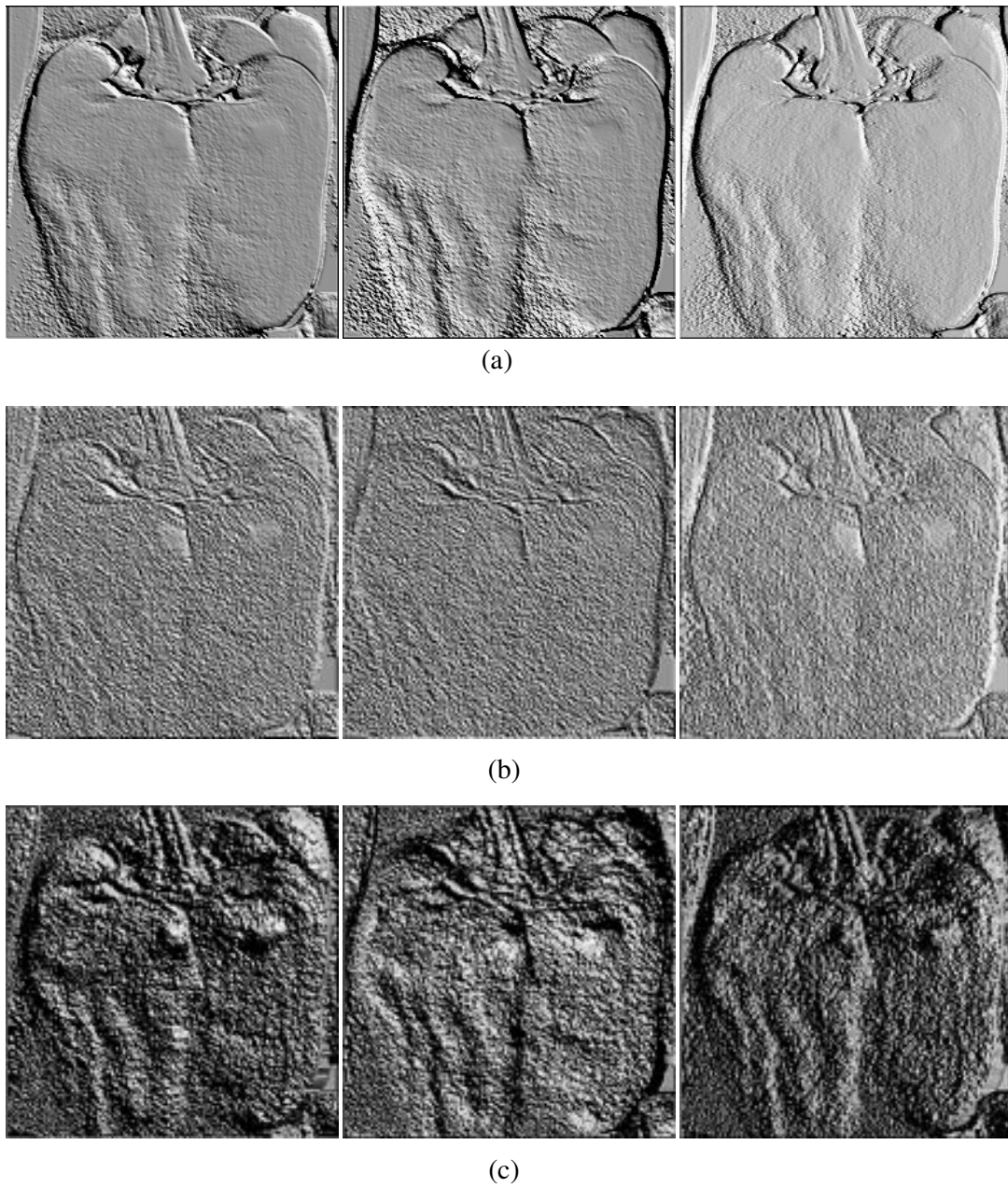
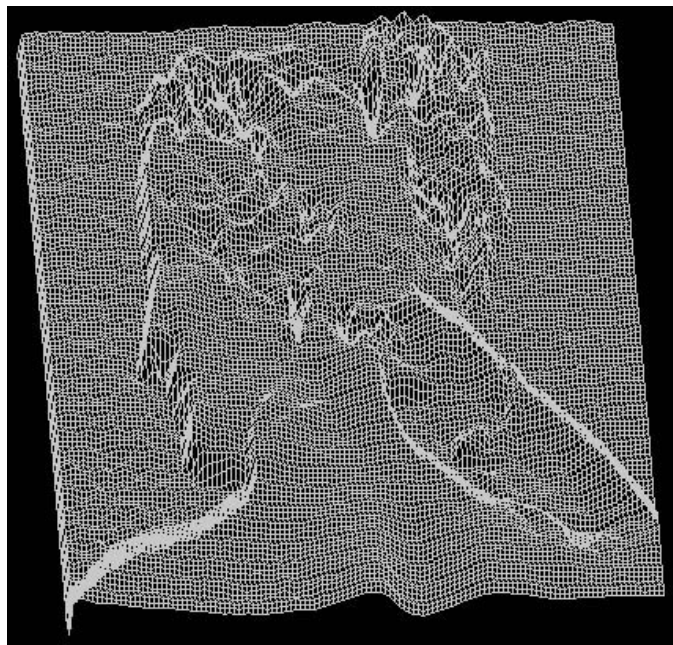
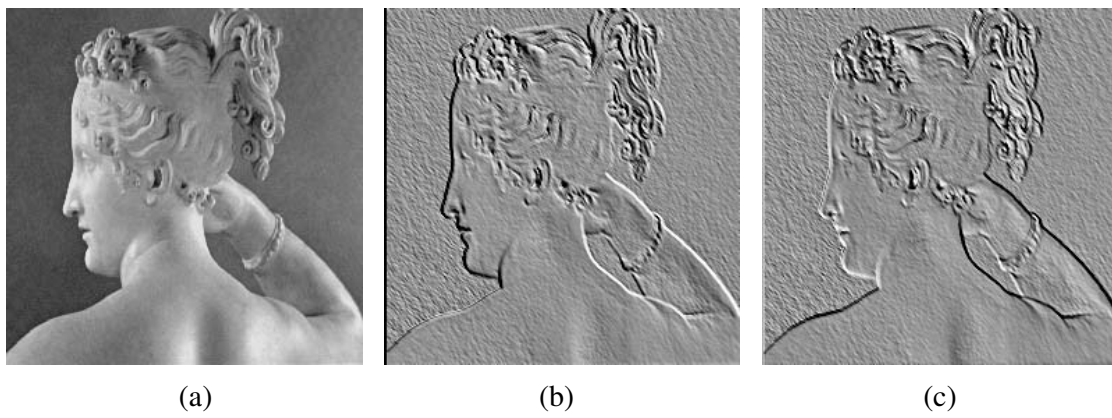
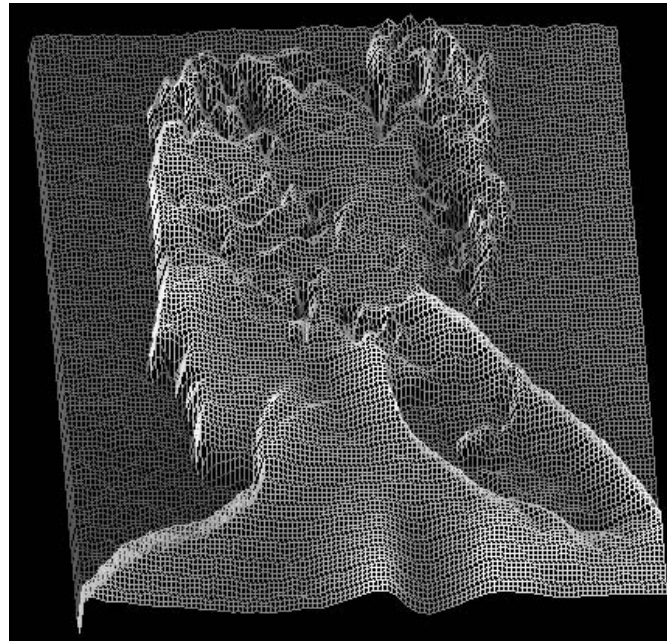


Figura 3.37. Imagens obtidas a partir da re-iluminação das superfícies estimadas (Figura 3.35), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$, $(-1,-1,1)$ e $(1,0,1)$. (a): GPM, (b): Lee/Rosenfeld e (c) Tsai/Shah.

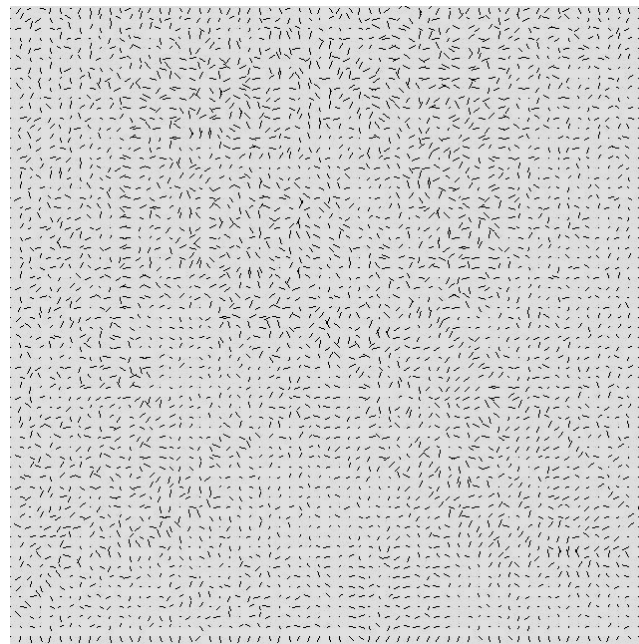


(d)

Figura 3.38. (a): Imagem de entrada (*Paolina*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 3.39. (a): Profundidades estimadas, referentes à imagem *Paolina*, com o padrão de irradiância em 3.38-(a) mapeado na representação. (b): Mapa de orientações estimado.

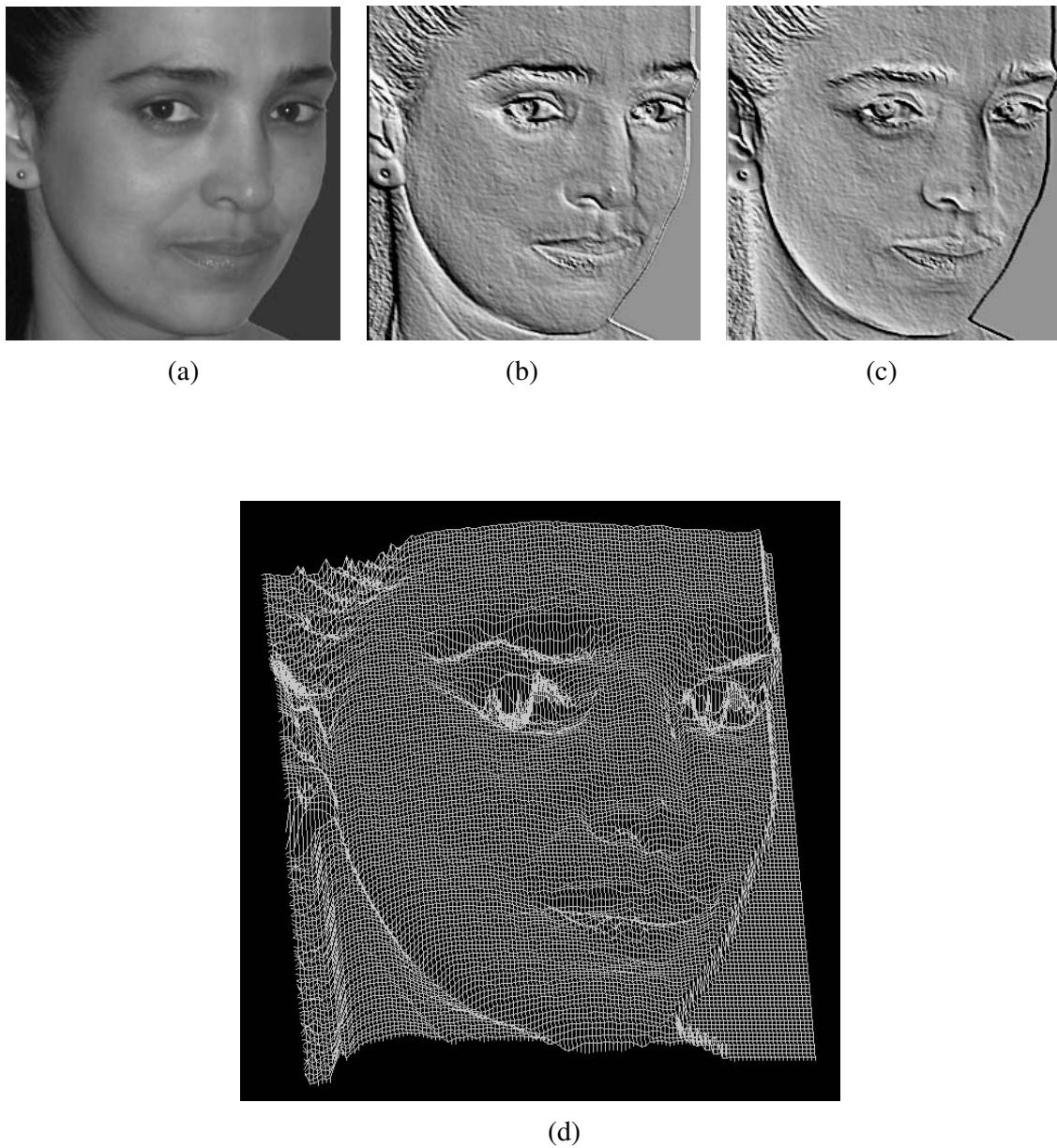
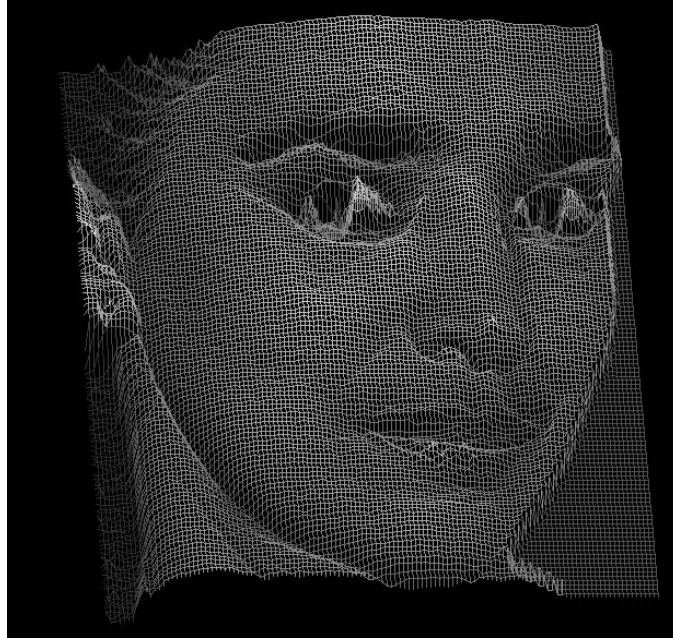
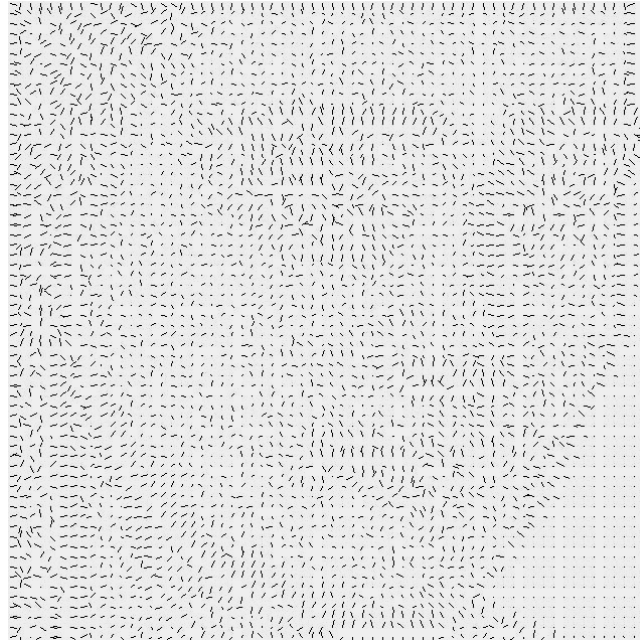


Figura 3.40. (a): Imagem de entrada (*Face2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 3.41. (a): Profundidades estimadas, referentes à imagem *Face2*, com o padrão de irradiância em 3.40-(a) mapeado na representação. (b): Mapa de orientações estimado.

Do *Shape from Defocus* ao *Shape from Shading***4.1 – Introdução**

Os processos de *Shape from Focus* e *Shape from Defocus* referem-se ao problema de estimar a estrutura 3-D de uma cena, a partir de um conjunto de imagens monoculares obtidas pela mudança de foco dos dispositivos de captura.

O conceito de foco está associado à nitidez com que uma imagem é representada, que depende não somente da geometria do sistema de captura, mas também da forma das superfícies na cena. Considerando-se a utilização de lentes ideais nos dispositivos de imageamento, temos que, conforme mostrado na Figura 4.1, somente os pontos na cena localizados a uma distância u_0 da lente serão representados em foco perfeito, de acordo com a conhecida equação das lentes:

$$\frac{1}{v_0} + \frac{1}{u_0} = \frac{1}{F} \quad (4.1)$$

onde v_0 é a distância da lente ao plano da imagem e F é a distância focal. Como podemos perceber na figura, um determinado ponto localizado a uma distância $u > u_0$, seria perfeitamente focado em $v < v_0$ e, portanto, sofreria um ‘espalhamento’ no plano da imagem em v_0 , sendo representado como um círculo. Se considerarmos que a lente possui raio de abertura r , o tamanho R do círculo de espalhamento (ou borramento), pode ser facilmente deduzido, por simples geometria, como

$$R = \frac{r}{v}(v_0 - v) \quad (4.2)$$

Pela equação (4.2), podemos perceber a relação direta entre o raio de borramento e o tamanho, ou abertura, da lente. Na prática, observamos ainda que há uma região aquém e

além da distância de foco perfeito, na qual a nitidez da imagem é aceitável, no sentido de que, para pontos nessa faixa, o diâmetro do círculo de borramento permanece menor do que a resolução do dispositivo de captura. Esse intervalo é denominado **profundidade de campo** e depende do tipo de sensor que é usado, mas, em qualquer caso, quanto maior a abertura da lente, menor será essa medida. Considerando que a profundidade de campo é, normalmente, muito limitada, pontos em uma cena localizados a diferentes distâncias em relação à câmera serão registrados, no plano da imagem, com diferentes desfocamentos. Dessa forma, pode-se perceber a relação entre a geometria da superfície e o borramento característico, nas imagens resultantes, produzido pelo desfocamento.

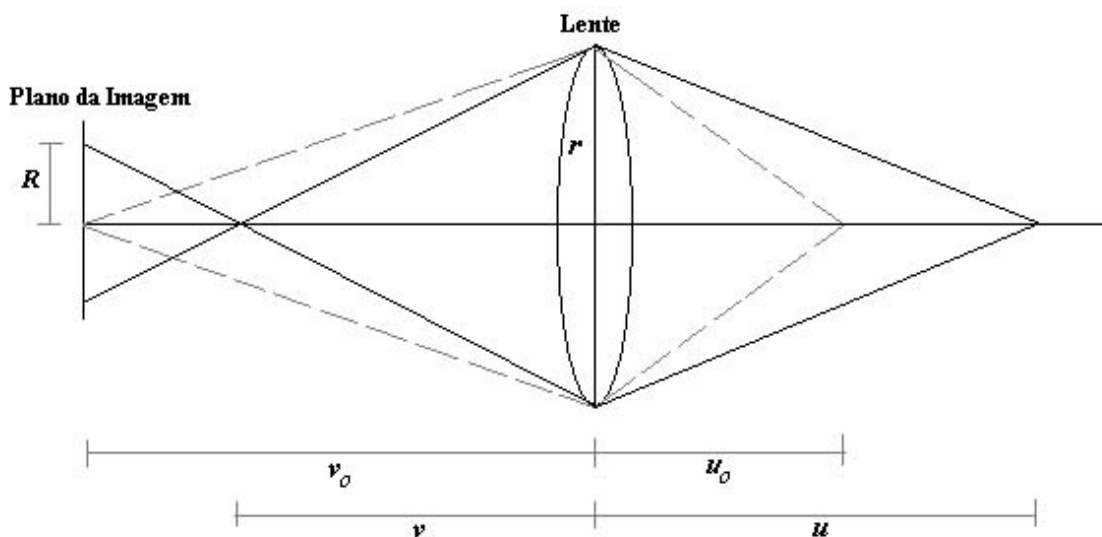


Figura 4.1. Geometria da formação da imagem. v_0 é a distância entre o plano da imagem e a lente, u_0 é a distância entre a lente e a posição do ponto-objeto perfeitamente focado e r é o raio da lente. Quando um ponto à distância $u > u_0$ é projetado através da lente, ele foca a uma distância $v < v_0$, assim, um círculo de borramento se forma no plano-imagem.

As técnicas que operam com *Shape from Focus* normalmente utilizam uma grande seqüência de imagens, tomadas sob diferentes condições de abertura da lente. Para se inferir a profundidade absoluta de um determinado ponto na cena, procede-se a uma busca, nas imagens disponíveis, da posição de melhor foco no ponto considerado [54]. Já as abordagens de *Shape from Defocus*, de forma geral, utilizam um pequeno número de imagens e um modelo para o processo de borramento como função da profundidade, buscando-se estimar a medida de desfocamento e, assim, utilizar essa informação para reconstrução tridimensional da cena. Nosso interesse aqui se restringe a esse último caso.

O trabalho de Pentland [9] foi o primeiro em Visão Computacional a propor a utilização de informação de desfocamento para se inferir forma 3D. Ele demonstrou que duas

imagens produzidas sob diferentes condições da abertura das câmeras poderiam proporcionar informações sobre as profundidades na cena. Desde então, muitos outros esquemas têm sido propostos. De forma geral, esses trabalhos podem requerer uma, duas ou mais imagens da cena, e podem operar no domínio espacial ou no domínio da frequência. Uma classificação interessante para os diferentes tipos de abordagem utilizados pode ser encontrada no artigo [55]. Segundo os seus autores, as abordagens mais comuns são:

- (i) aquelas que recuperam profundidades a partir de bordas [56,57,58],
- (ii) aquelas que utilizam bases polinomiais para modelar as imagens [55,59,60] e
- (iii) aquelas que utilizam filtragem inversa [9,56,61,62,63,64,65,66].

O método de Pentland [56], em que estamos interessados aqui, utiliza filtragem inversa e trabalha no domínio de Fourier, como veremos na seção a seguir.

4.2 – *Shape from Defocus: Abordagem de Pentland*

Como demonstrado por Pentland [56], a relação entre os mapas de profundidade e desfocamento de uma cena pode ser obtida, através de óptica geométrica, como

$$Z(x, y) = \frac{Fv_0}{v_0 - F - fR(x, y)} \quad (4.3)$$

onde, como já definimos, $R(x,y)$ denota o raio local de borramento (vide Figura 4.1), devido ao desfocamento, e o número- f é definido como a razão entre a distância focal e o raio de abertura da lente ($f = F/r$). De acordo com esta equação, profundidades podem ser estimadas a partir da medida de desfocamento R , uma vez conhecidos os parâmetros da câmera.

De forma usual, o processo de desfocamento é modelado através da convolução de uma imagem perfeitamente focada com uma função de espalhamento de ponto (PSF - *point spread function*), cuja dimensão espacial é proporcional ao parâmetro de desfocamento. Essa não é uma aproximação muito realística, já que o desfocamento é um processo espacialmente variante, mas, assumindo a função Gaussiana como PSF, Pentland considerou convoluções locais em janelas ao longo da imagem, de forma a quantificar o desfocamento dependente de profundidade.

Assim, Pentland assumiu que as intensidades, em uma janela local na imagem, podem ser expressas como a convolução

$$I(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right] * L(x, y) \quad (4.4)$$

onde $L(x, y)$ é a imagem idealmente focada, e σ denota a medida local de desfocamento, assumida uniforme sobre cada janela. No domínio da frequência, a equação (4.4) assume a forma

$$\tilde{I}(\omega) = \exp\left[-\frac{\sigma^2 \omega^2}{2}\right] \tilde{L}(\omega) \quad (4.5)$$

onde estamos usando as notações \tilde{I} e \tilde{L} para representar as transformadas de Fourier dos respectivos sinais espaciais, e onde $\omega \equiv \sqrt{\omega_x^2 + \omega_y^2}$ é a magnitude da frequência. Assim, considerando janelas correspondentes em um par de imagens, aqui denominadas I e I_0 , capturadas sob diferentes condições de foco, vamos obter

$$\frac{\tilde{I}(\omega)}{\tilde{I}_0(\omega)} = \exp\left[-\frac{\omega^2}{2}(\sigma^2 - \sigma_0^2)\right] \quad (4.6)$$

e então,

$$\sigma^2 = \sigma_0^2 - \frac{2}{\omega^2} \log \frac{\tilde{I}(\omega)}{\tilde{I}_0(\omega)} \quad (4.7)$$

Dados \tilde{I} e \tilde{I}_0 , Pentland sugeriu resolver a equação (4.7) através de uma regressão linear em ω^2 . No entanto, como é fácil verificar, não é possível utilizar a componente de Fourier de frequência zero (DC) neste processo, e mesmo componentes de baixa frequência podem inviabilizá-lo. Talvez por esta razão, Pentland, com base neste método, estimou apenas descontinuidades em profundidade, embora ele não tenha salientado tal limitação em seu artigo. Trabalhos posteriores apontaram ainda outros problemas com a sua abordagem, tais como a introdução de componentes espúrias de alta frequência (devido às descontinuidades artificiais introduzidas pelo uso de janelas) e o vazamento de irradiância (*irradiance leaking*) entre janelas vizinhas. A seguir, nós apresentaremos uma proposta de modificação do método

de Pentland que, além de procurar evitar estes problemas, permite a estimação de mapas densos de profundidade.

4.3 – Nova abordagem para *Shape from Defocus* (SFD)

A nossa proposta é a de introduzir uma modificação simples na formulação de Pentland, modelando o processo de desfocamento através da convolução com gaussianas **não normalizadas**. Além de parecer mais apropriado para levar em conta o vazamento de irradiância entre janelas vizinhas, este modelo nos permite estimar uma medida de desfocamento através da componente DC da Transformada de Fourier, obtida localmente em cada *patch* da imagem, evitando, portanto, as já mencionadas componentes espúrias de alta frequência, e permitindo a obtenção de um mapa denso de profundidades.

Assim, em lugar da equação (4.4), nós assumimos que as intensidades, em uma janela local na imagem, podem ser expressas como

$$I(x, y) = \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right] * L(x, y) \quad (4.8)$$

onde, novamente, $L(x,y)$ é a imagem idealmente focada, e σ denota a medida local de desfocamento. Considerando que sempre haverá um fluxo de irradiância (devido ao borramento) através das bordas de janelas vizinhas, é bem razoável não assumirmos normalização neste caso. Além disso, simples considerações de óptica geométrica nos levam a concluir que a intensidade luminosa total, capturada por uma lente, varia à medida que o ponto objeto se afasta da mesma. Para ilustrar essa afirmação, vide o esquema na Figura 4.2, onde mostramos o mesmo ponto objeto a diferentes distâncias com relação à lente. É fácil concluir que o ângulo sólido formado pelos raios luminosos emanados do ponto em questão, e que chegam à lente, decresce com a distância. Assim, a intensidade de um mesmo ponto objeto, capturado sob idênticas condições de abertura mas a diferentes distâncias da câmera, não irá se conservar, e é possível explicitar a dependência da irradiância da imagem com a posição, conforme abaixo:

Consideremos a situação representada na Figura 4.2, e vamos supor que o ponto P emita uma potência total unitária, por unidade de ângulo sólido, isotropicamente, no hemisfério de direções voltado para a lente (a sua radiância seria então dada por $\delta(x,y)$, onde δ é a função

delta de Dirac). A fração da potência emitida que atingirá a lente, quando o ponto P se encontrar a uma distância Z da mesma, será, portanto, numericamente igual ao valor do ângulo sólido compreendido pela lente quando vista daquela posição. Para a configuração da Figura 4.2, com o ponto P situado no eixo óptico, isto corresponderá a S/Z^2 , onde S é a área da lente. É, portanto, este valor total de irradiância que será redistribuído em um borramento, sempre que o ponto P não se encontrar na posição de foco perfeito, conforme a Figura 4.1 - ou seja, sempre que $Z \neq u_o$.

Vamos agora supor que, no borramento, a irradiância se distribua como uma função gaussiana. Esta será, por definição, a função de espalhamento de ponto associada ao desfocamento. É fácil concluir que tal gaussiana deve comportar um fator de normalização, N , da forma $N=S/(2\pi\sigma^2Z^2)$, de modo que a sua integral corresponda ao total da irradiância capturada pela lente. Neste caso, o processo de desfocamento será modelado pela equação

$$I(x, y) = \frac{S}{2\pi\sigma^2Z^2} \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right] * L(x, y) \quad (4.9)$$

Por outro lado, nós temos que a relação entre a medida de borramento e a posição do ponto objeto é dada pela expressão (4.3), com R identificado a σ . Resolvendo-a para este fator, nós vamos obter

$$\sigma = \frac{v_0 - F}{f} - \frac{Fv_0}{fZ} \quad (4.10)$$

e, portanto,

$$\sigma Z \approx -\frac{Fv_0}{f} = \text{constante} \quad (4.11)$$

sempre que $|Fv_0/Z| \gg |v_0 - F|$ - ou seja, assumindo $v_0 > F$, sempre que

$$Z \ll \frac{Fv_0}{v_0 - F} \quad (4.12)$$

Sob tal condição, que se torna menos restritiva à medida que v_0 se aproxima de F (ou seja, quando o sistema de captura está focado para longas distâncias), a equação (4.9) se torna equivalente à (4.8), já que o fator que multiplica a gaussiana na primeira é irrelevante.

Retornando, portanto, à equação (4.8), nós verificamos que, no domínio da frequência, esta assume a forma

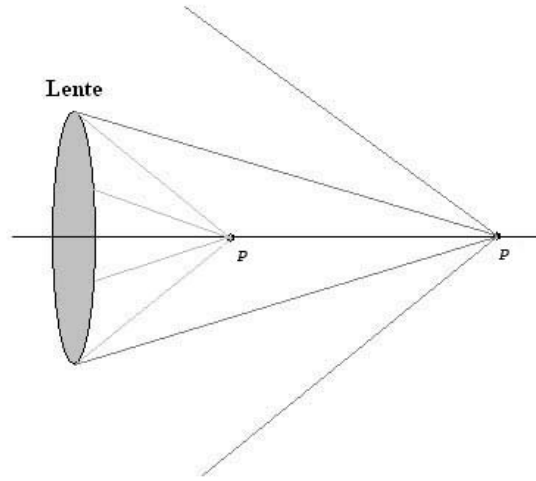


Figura 4.2. O ângulo sólido formado pelos raios luminosos que chegam à lente, emanando de um mesmo ponto objeto, decresce com a distância.

$$\tilde{I}(\omega) = 2\pi\sigma^2 \exp\left[-\frac{\sigma^2\omega^2}{2}\right] \tilde{L}(\omega) \quad (4.13)$$

onde, como antes, estamos usando as notações \tilde{I} e \tilde{L} para representar as transformadas de Fourier dos respectivos sinais espaciais, e onde ω é a magnitude da frequência. Novamente, se consideramos janelas correspondentes em um par de imagens capturadas sob diferentes condições de foco, vamos obter

$$\frac{\tilde{I}(\omega)}{\tilde{I}_0(\omega)} = \frac{\sigma^2}{\sigma_0^2} \exp\left[-\frac{\omega^2}{2}(\sigma^2 - \sigma_0^2)\right] \quad (4.14)$$

e então,

$$\log \frac{\tilde{I}(\omega)}{\tilde{I}_0(\omega)} = \log \frac{\sigma^2}{\sigma_0^2} - \frac{\omega^2}{2} (\sigma^2 - \sigma_0^2) \quad (4.15)$$

É fácil ver que, diferentemente da abordagem de Pentland, a relação (4.15) permite a estimação de σ , em função de σ_0 , mesmo quando $\omega = 0$. Neste caso, obtemos

$$\sigma = \sigma_0 \sqrt{\frac{\tilde{I}(0)}{\tilde{I}_0(0)}} \quad (4.16)$$

A equação (4.16), quando computada sobre uma série de janelas ao longo de toda a imagem, vai definir a nossa medida de desfocamento dependente da posição. Utilizando-a na equação (4.3) com a identificação $R = \sigma$, esta pode ser facilmente reescrita sob uma forma que expressa a função Z , a menos de um fator multiplicativo, em termos de um único parâmetro livre:

$$Z(x, y) = \alpha \frac{1}{\phi - \sqrt{\frac{\tilde{I}(0)}{\tilde{I}_0(0)}}} \quad (4.17)$$

onde $\alpha = Fv_0 / f\sigma_0$ e $\phi = (v_0 - F) / f\sigma_0$. A relação (4.17) foi por nós utilizada para a estimação de profundidades, a menos do fator de escala α , na ausência de informações sobre os parâmetros do sistema de captura. Um resultado do processo encontra-se ilustrado na Figura 4.3. Neste caso, foram utilizadas janelas de 2 x 2 pixels para a estimação do desfocamento, com a imagem inicial, I_0 , assumida obtida sob desfocamento uniforme (i.e., σ_0 constante). O parâmetro livre, ϕ , foi obtido empiricamente (através de um estudo similar ao desenvolvido para a estimativa inicial do parâmetro k_0 , na abordagem iterativa do GPM – ver Capítulo 3 Seção 3.3.1).

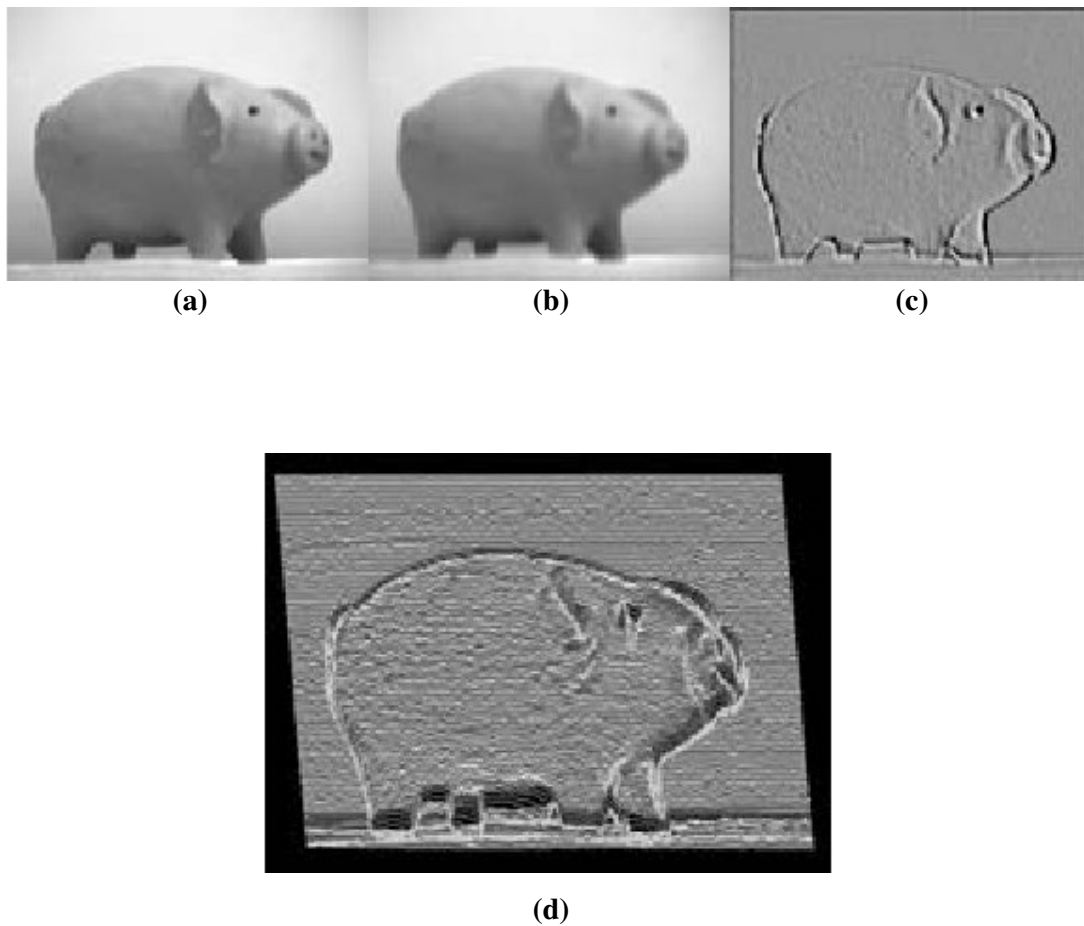


Figura 4.3. (a) e (b): Imagens de entrada, capturadas sob diferentes condições de abertura. (c): Imagem obtida a partir da re-iluminação da superfície estimada, em (d), assumindo reflectância lambertiana e albedo uniforme. Direção de iluminação: (1,1,1).

4.3.1 – Modelagem do borramento por Função *Pillbox*

Uma outra possibilidade que poderíamos ter considerado seria a utilização da chamada função *Pillbox* como PSF, o que, segundo certos autores, seria mais indicado para modelar o desfocamento. Nesse caso, em lugar da equação (4.8), assumiríamos que as intensidades, em uma janela local na imagem, podem ser expressas como

$$I(x, y) = h(x, y) * L(x, y) \quad (4.18)$$

onde $h(x,y)$ é a função *Pillbox* de espalhamento de ponto, não normalizada, dada por

$$h(x, y) = \begin{cases} 1, & \text{para } x^2 + y^2 \leq R^2 \\ 0, & \text{para } x^2 + y^2 > R^2 \end{cases} \quad (4.19)$$

sendo R o raio de borrimento (vide Figura 4.1), anteriormente identificado ao parâmetro σ da Gaussiana. Nosso objetivo, agora, seria o de estimar R .

Considerando a equação (4.18) no domínio da frequência, nós obtemos

$$\tilde{I}(\omega) = 2\pi R^2 \frac{J_1(R\omega)}{R\omega} \tilde{L}(\omega) \quad (4.20)$$

onde $J_1(R\omega)$ é a função de Bessel de primeira ordem [10]. Assumindo janelas correspondentes em um par de imagens capturadas sob diferentes condições de foco, temos, portanto, para $\omega = 0$,

$$\frac{\tilde{I}(0)}{\tilde{I}_0(0)} = \left(\frac{R}{R_0} \right)^2 \quad (4.21)$$

já que a função $J_1(z)/z$ tende a $1/2$, quando z tende a zero. Assim, explicitando o valor de R , obtemos

$$R = R_0 \sqrt{\frac{\tilde{I}(0)}{\tilde{I}_0(0)}} \quad (4.22)$$

que é um resultado equivalente ao expresso em (4.16). Dessa forma, verificamos que, mesmo sob um modelo alternativo para o borrimento, a equação (4.17) permanece válida para estimação das profundidades na cena.

4.4 – A abordagem para uma única imagem

Uma outra vantagem da nossa abordagem do *Shape from Defocus* é que ela admite a extensão para um processo de estimação de forma a partir de uma única imagem de *shading*, ou sombreamento.

As imagens ideais de sombreamento, tais como assumidas em processos de *Shape from Shading* (SFS), não apresentam um borrimento dependente das profundidades, já que se considera que elas foram capturadas sob projeção ortográfica. Conclui-se, portanto, que o processo de *Shape from Defocus* não poderia ser baseado em um tal tipo de imagens. Por outro lado, tendo em vista que o sombreamento (*shading*) é devido à variação de orientação das superfícies, e, portanto, às profundidades na cena, imagens *reais* de *shading*, capturadas sob abertura finita, devem, necessariamente, apresentar um certo desfocamento espacialmente variante, dependente da geometria da cena, sendo, dessa forma, adequadas para implementações do SFD. O nosso objetivo, aqui, será o de modelar a imagem de *shading* como resultando de um processo de borrimento, o que vai possibilitar a estimação de um parâmetro de desfocamento e, portanto, a aplicação do método acima descrito para a reconstrução da superfície representada.

4.4.1 - Desenvolvimento do processo

Vamos começar o desenvolvimento de nossa formulação apresentando uma identidade matemática. Para facilitar a exposição, consideraremos, inicialmente, sinais unidimensionais.

Temos que um sinal unidimensional qualquer, $I(x)$, pode ser representado como

$$I(x) = \int_{-\infty}^{\infty} d\omega_x \left\{ \exp[i\omega_x(x-a)] \exp\left[-\frac{(x-a)^2}{2\sigma^2}\right] * \exp[i\omega_x x] \right\} \quad (4.23)$$

onde σ e a são funções da frequência ω_x , dadas, respectivamente, por

$$\sigma(\omega_x) = \frac{1}{(2\pi)^{3/2}} |\tilde{I}(\omega_x)| \quad (4.24)$$

e

$$a(\omega_x) = -\frac{\varphi_I(\omega_x)}{\omega_x}, \quad \omega_x \neq 0 \text{ com } a(0) = 0 \quad (4.25)$$

Nas expressões acima, $|\tilde{I}(\omega_x)|$ representa o módulo da Transformada de Fourier do sinal $I(x)$, e $\varphi_I(\omega_x)$, o ângulo de fase correspondente.

Para justificar as relações em (4.24) e (4.25), vamos considerar a equação (4.23) no domínio da frequência: mudando a variável de integração para ω_x^0 e tomando a Transformada de Fourier dos dois lados da equação, obtemos

$$\begin{aligned} \tilde{I}(\omega_x) = \int_{-\infty}^{\infty} d\omega_x^0 \{ & \exp[-i\omega_x^0 a] \sqrt{2\pi} \sigma \exp[-\frac{\sigma^2}{2}(\omega_x - \omega_x^0)^2] \times \\ & \times \exp[-i(\omega_x - \omega_x^0)a] 2\pi \delta(\omega_x - \omega_x^0) \} \end{aligned} \quad (4.26)$$

ou seja,

$$\tilde{I}(\omega_x) = (2\pi)^{3/2} \int_{-\infty}^{\infty} d\omega_x^0 \{ \sigma \exp[-\frac{\sigma^2}{2}(\omega_x - \omega_x^0)^2] \exp[-i\omega_x^0 a] \delta(\omega_x - \omega_x^0) \}$$

onde δ representa a função delta de Dirac. Efetuando a integração sobre ω_x^0 , obtemos então, devido à propriedade de filtragem da δ ,

$$\tilde{I}(\omega_x) = (2\pi)^{3/2} \sigma \exp[-i\omega_x a] \quad (4.27)$$

e, explicitando σ ,

$$\sigma(\omega_x) = \frac{1}{(2\pi)^{3/2}} \tilde{I}(\omega_x) \exp[i\omega_x a] \quad (4.28)$$

Escolhendo adequadamente o parâmetro a , de forma que $a(\omega_x) = -\frac{\varphi_I(\omega_x)}{\omega_x}$, para todo

$\omega_x \neq 0$, com $a(0) = 0$, vamos obter

$$\sigma(\omega_x) = \frac{1}{(2\pi)^{3/2}} |\tilde{I}(\omega_x)| \quad (4.29)$$

o que demonstra as relações em (4.24) e (4.25).

Podemos também verificar a consistência matemática da equação (4.23), substituindo aí as expressões para σ e a . Explicitando a convolução em (4.23), vamos começar por reescrever esta equação como

$$I(x) = \int_{-\infty}^{\infty} d\omega_x \int_{-\infty}^{\infty} d\xi \left\{ \exp[i\omega_x(\xi - a)] \exp\left[-\frac{(\xi - a)^2}{2\sigma^2}\right] \exp[i\omega_x(x - \xi)] \right\} \quad (4.30)$$

ou ainda,

$$I(x) = \int_{-\infty}^{\infty} d\omega_x \left\{ \exp[i\omega_x x] \exp[-i\omega_x a] \int_{-\infty}^{\infty} d\xi \exp\left[-\frac{(\xi - a)^2}{2\sigma^2}\right] \right\} \quad (4.31)$$

Agora, resolvendo a integral interna em (4.31), obtemos

$$I(x) = \int_{-\infty}^{\infty} d\omega_x \left\{ \exp[i\omega_x x] \exp[-i\omega_x a] \sqrt{2\pi}\sigma \right\} \quad (4.32)$$

E substituindo o valor de σ deduzido em (4.29),

$$I(x) = \int_{-\infty}^{\infty} d\omega_x \left\{ \exp[i\omega_x x] \frac{1}{2\pi} |\tilde{I}(\omega_x)| \exp[-i\omega_x a] \right\} \quad (4.33)$$

Agora, como tomamos $\omega_x a = -\varphi_I(\omega_x)$, o lado direito de (4.33) torna-se

$$\frac{1}{(2\pi)} \int_{-\infty}^{\infty} d\omega_x \exp[i\omega_x x] \tilde{I}(\omega_x) \quad (4.34)$$

que é exatamente a definição de $I(x)$ por Transformada Inversa de Fourier, comprovando a consistência matemática da representação proposta.

Para um sinal bidimensional, a equação (4.23) assumiria a forma:

$$I(x, y) = \iint_{-\infty}^{\infty} d\omega_x d\omega_y \left\{ \exp[i\omega_x(x-a) + i\omega_y(y-b)] \exp\left[-\frac{(x-a)^2 + (y-b)^2}{2\sigma^2}\right] * \right. \\ \left. * \exp[i(\omega_x x + \omega_y y)] \right\} \quad (4.35)$$

onde, agora,

$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|} \quad (4.36)$$

e

$$\omega_x a(\omega_x, \omega_y) + \omega_y b(\omega_x, \omega_y) = -\varphi_I(\omega_x, \omega_y) \quad (4.37)$$

conforme demonstramos no Anexo II deste trabalho, por um procedimento similar ao desenvolvido acima.

Um caso particular interessante da equação (4.35) é aquele em que aproximamos $I(x,y)$ apenas pela componente de frequência zero da integral à direita. Dessa forma, $I(x,y)$ se reduz à convolução de uma Gaussiana não-normalizada por uma função uniforme: $f(x, y) = 1$. Nesta aproximação, a relação (4.35) expressa um sinal qualquer como o ‘borramento’ (no sentido introduzido na Seção 4.3) de uma imagem totalmente ‘branca’. Isto pode ser utilizado como um interessante modelo para a representação de imagens de *shading* com albedo unitário, em que elas são vistas como resultando do desfocamento de uma superfície uniforme. Por outro lado, abre ainda a possibilidade do emprego da nossa abordagem de SFD para a extração de *shape from shading*. Nesse caso, considerando a relação (4.35) sobre janelas na imagem, e utilizando apenas a componente de frequência zero, obteremos, a partir de (4.36),

$$\sigma = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(0,0)|} \quad (4.38)$$

como a estimativa de desfocamento em cada janela.

De forma semelhante à da Seção 4.3, utilizando (4.38) na equação (4.3), esta pode ser então reescrita em uma forma que expressa o mapa de profundidades, a menos de um fator multiplicativo, em termos de um único parâmetro livre:

$$Z(x, y) = \alpha \frac{1}{\phi - \sqrt{|\tilde{I}(0,0)|}} \quad (4.39)$$

onde, agora, $\alpha = (2\pi)^{3/2} Fv_0 / f$ e $\phi = (2\pi)^{3/2} (v_0 - F) / f$ representariam parâmetros de um sistema virtual de captura de imagens. A relação (4.39) pode ser utilizada para a estimação da forma das superfícies representadas.

Exemplos de aplicação dessa nova abordagem para o SFS podem ser vistos nas Figuras 4.4 a 4.6. Para estimação de forma, utilizamos a equação (4.39), implementada em janelas 2 x 2 ao longo das imagens. Nas Figuras, com o intuito de validar o nosso modelo, apresentamos também as intensidades das imagens de entrada, recuperadas a partir da equação (4.35) e dos parâmetros σ obtidos de (4.38). Os mapas de profundidade foram gerados a menos de um fator multiplicativo e com o único parâmetro livre, ϕ , empiricamente escolhido, como no exemplo da Figura 4.3.

Embora os nossos experimentos tenham demonstrado a aplicabilidade da abordagem de reconstrução baseada apenas na componente DC, nós verificamos, através da análise dos erros incorridos, que, na maioria dos casos, os resultados não estão ainda competitivos com aqueles gerados pelo GPM. Isto pode se comprovado através dos quadros na Figura 4.7, onde procuramos avaliar numericamente a qualidade das reconstruções obtidas a partir das imagens apresentadas na Figura 3.15 (a) a (e), baseadas em cenas cujos mapas de profundidade são conhecidos. Tais reconstruções são comparadas com as anteriormente obtidas a partir do GPM, utilizando os mesmos índices de erro definidos no Capítulo 3.



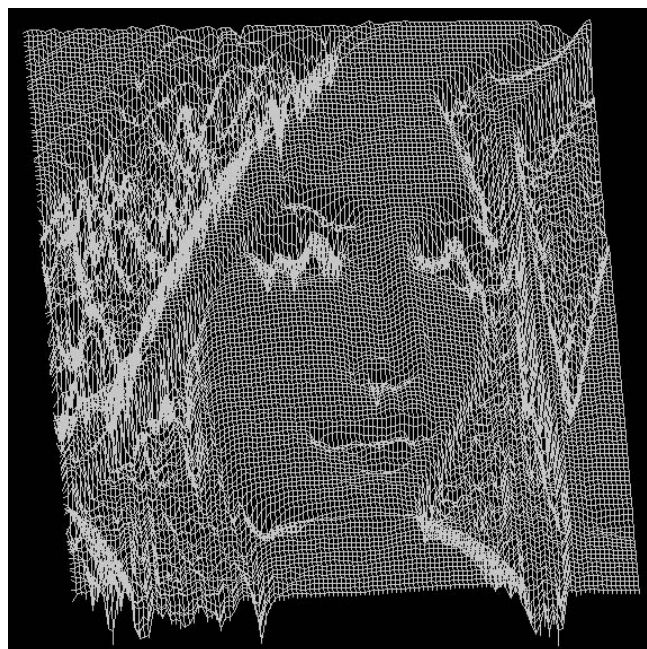
(a)

(b)

(c)



(d)



(e)

Figura 4.4. (a): Imagem de entrada (*Lena*). (b) e (c): Imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Intensidades da imagem de entrada, obtidas a partir das equações (4.35) e (4.38).



(a)

(b)

(c)



(d)



(e)

Figura 4.5. (a): Imagem de entrada (*Paolina*). (b) e (c): Imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Intensidades da imagem de entrada, obtidas a partir das equações (4.35) e (4.38).

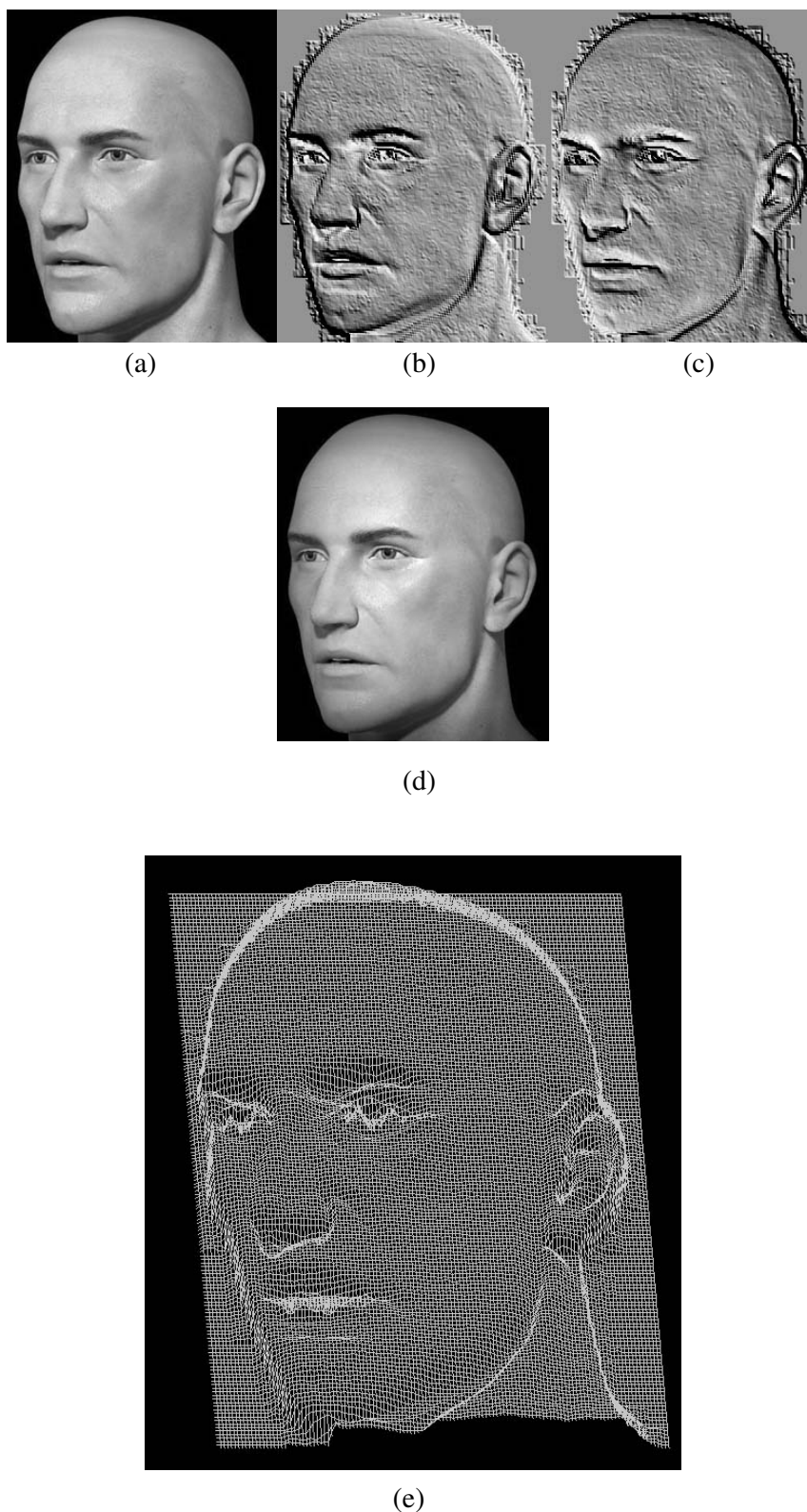


Figura 4.6. (a): Imagem de entrada (*Face*). (b) e (c): Imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Intensidades da imagem de entrada, obtidas a partir das equações (4.35) e (4.38).

<i>Imagens</i>	<i>Método</i>	
	<i>GPM</i>	<i>SFS</i> <i>Via Defocus (DC)</i>
<i>Mask1</i>	0,9	2,8
<i>Mozart1</i>	1,6	3,9
<i>Mozart2</i>	2,1	4,7
<i>Vase1</i>	0,6	1,6
<i>Penny</i>	1,1	1,8

(a)

<i>Imagens</i>	<i>Método</i>	
	<i>GPM</i>	<i>SFS</i> <i>Via Defocus (DC)</i>
<i>Mask1</i>	5,4	7,3
<i>Mozart1</i>	7,1	7,6
<i>Mozart2</i>	12,7	14,3
<i>Vase1</i>	5,1	7,9
<i>Penny</i>	4,2	4,4

(b)

Figura 4.7. Análise comparativa dos erros produzidos pelas abordagens do GPM e do SFS via *Defocus* (componente DC, apenas). As imagens utilizadas para estimação dos mapas de profundidades são aquelas da Figura 3.15 (a) a (e). (a): Erro médio gradiente calculado a partir das estimativas produzidas pelos dois métodos. (b): Erro médio em profundidades.

4.4.2 - Abordagem Multifrequência

Aqui, nós retomamos a equação (4.35), no intuito de tentar refinar a nossa abordagem. Assim, interpretando aquela equação no caso geral, nós vamos assumir que a imagem de *shading* possa ser modelada como resultando da composição de uma série de ‘borramentos generalizados’ - num sentido a ser esclarecido abaixo - de ondas planas por filtros de Gabor, e que tais ‘borramentos’ possam ser interpretados como desfocamentos de uma superfície subjacente, gerados de forma semelhante aos desfocamentos físicos produzidos por um sistema óptico.

Os filtros de Gabor (senóides moduladas por gaussianas) vêm há um longo tempo sendo utilizados em análise de imagens, devido à sua propriedade de localização ótima no espaço e na frequência [67], e têm sido empregadas também como modelo para os campos receptivos de células do córtex visual [49,68].

Na equação (4.35), esses filtros aparecem sob a forma complexa

$$g_a(x, y) = \exp[i\omega_x^0(x-a) + i\omega_y^0(y-b)] \exp\left[-\frac{(x-a)^2 + (y-b)^2}{2\sigma^2}\right] \quad (4.40)$$

A Figura 4.8 mostra representações para a parte real de (4.40), e, na Figura 4.9, podemos ver exemplos de cortes transversais das suas partes real e imaginária.

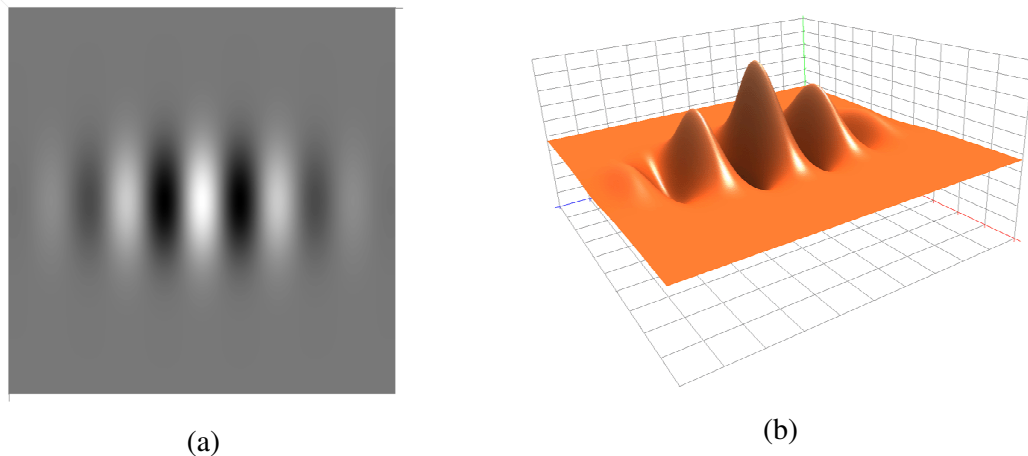


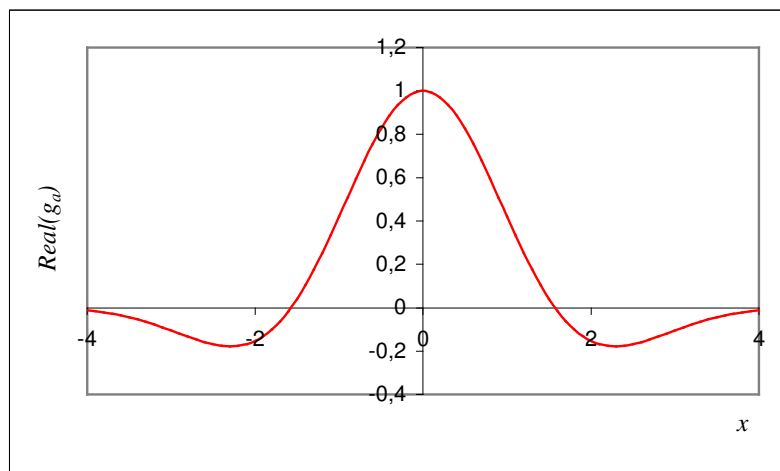
Figura 4.8. Parte real da Função de Gabor em (4.40). (a) Vista superior com representação em intensidades. (b) Representação 3-D.

A Transformada de Fourier, \tilde{g}_a , da função em (4.40) é também uma função de Gabor no domínio da frequência, cujo módulo é uma gaussiana centrada em (ω_x^0, ω_y^0) :

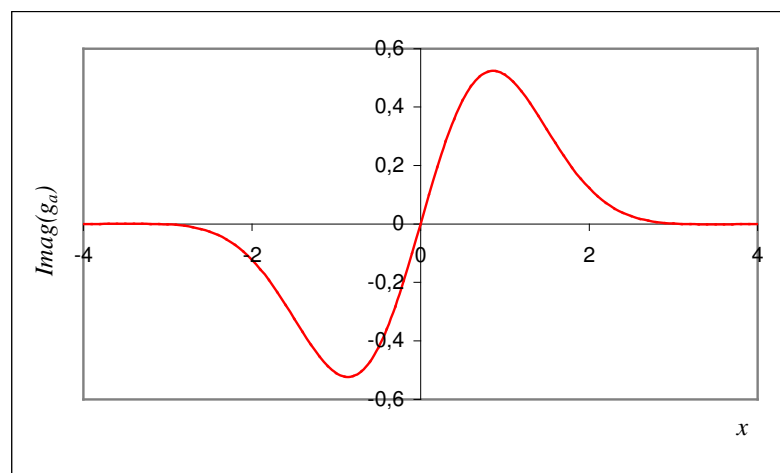
$$\tilde{g}_a(\omega_x, \omega_y) = 2\pi\sigma^2 \exp\left[-\sigma^2 \frac{(\omega_x - \omega_x^0)^2 + (\omega_y - \omega_y^0)^2}{2}\right] \exp[-i(\omega_x a + \omega_y b)] \quad (4.41)$$

Na Figura 4.10 podemos ver exemplos de cortes transversais do módulo e da fase de \tilde{g}_a . É fácil concluir que o efeito da filtragem de um sinal pela função de Gabor em (4.40) é o de esmaecer progressivamente as suas componentes de frequência, à medida em que estas se afastam de (ω_x^0, ω_y^0) . No caso da filtragem de um sinal geral, multi-frequência, a função de

Gabor se comporta, portanto, como um filtro passa-faixa. Como aparece na equação (4.35), porém, que é o caso que nos interessa, o filtro de Gabor de frequência (ω_x^0, ω_y^0) atua exatamente sobre um sinal com a frequência exclusiva (ω_x^0, ω_y^0) . O seu efeito é o de esmaecer ou de realçar, dependendo do valor de σ , aquela única componente de frequência. Assim, a mesma interpretação dada à componente DC da equação (4.35) pode ser estendida às suas componentes de uma frequência qualquer, e estamos justificados em falar, também aqui, do ‘borramento’ (no sentido da Seção 4.3) de uma imagem de uma única frequência não-nula. Da mesma forma, permanece válida a interpretação de tal borramento como resultante de um desfocamento, devido à variação de profundidade, semelhante ao que seria produzido por um sistema físico de captura de imagens.

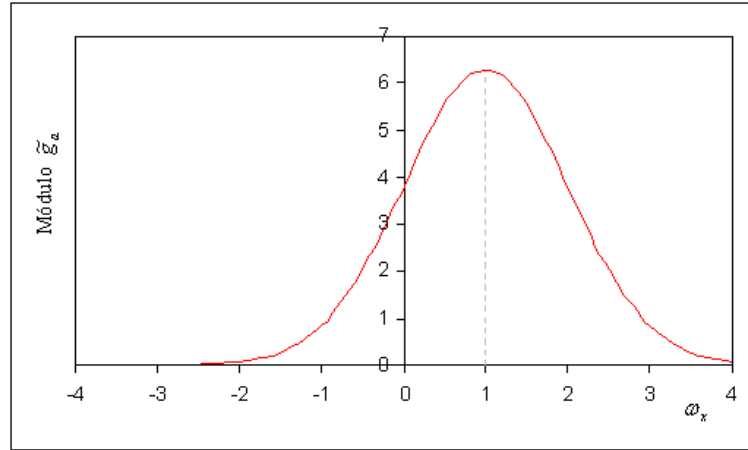


(a)

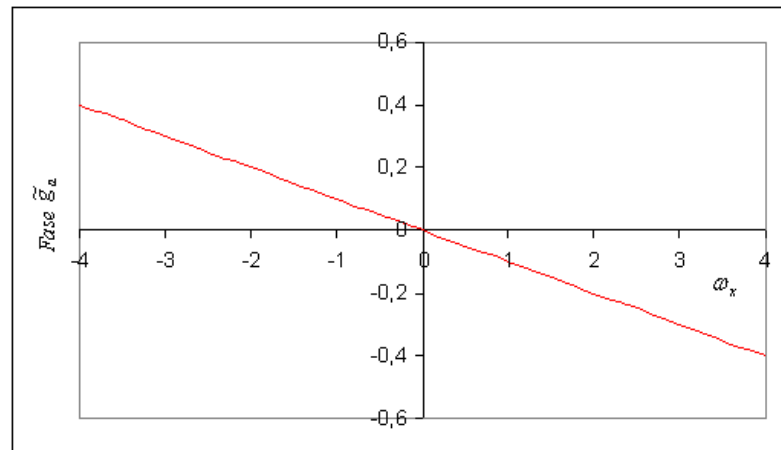


(b)

Figura 4.9. Cortes 1-D da função de Gabor em (4.40), ao longo da direção x , para $\omega_x^0 = 1$, $a = 0$ e $\sigma = 1$. (a) Parte real. (b) Parte Imaginária.



(a)



(b)

Figura 4.10. Cortes 1-D da Transformada da Função de Gabor em (4.41), ao longo da direção ω_x , para $\omega_x^0 = 1.0$, $a = 0.1$ e $\sigma = 1.0$. (a) Módulo (b) Ângulo Fase (em radianos).

Em vista disso, podemos admitir que a equação (4.3) permanece válida para a estimação de profundidades, e, recordando que o parâmetro σ é função da frequência (ver (4.36)), podemos obter um mapa $Z(x,y)$ para cada componente (ω_x, ω_y) , ou seja,

$$Z_{(\omega_x, \omega_y)}(x, y) = \alpha \frac{1}{\phi - \sqrt{|\tilde{I}(\omega_x, \omega_y)|}} \quad (4.42)$$

onde α e ϕ são assumidos independentes da frequência.

A nossa proposta é, então, a de obter uma estimativa global do mapa de profundidades da cena, utilizando as estimativas parciais para uma gama de frequências. Conforme verificamos em nossos experimentos, uma forma bastante robusta para o mapa final de profundidades é aquela obtida consolidando-se todas as estimativas parciais numa média ponderada, onde cada reconstrução recebe o peso que a frequência correspondente tem na imagem, ou seja,

$$Z(x, y) = \sum_{\omega_x=0}^{\omega_x, \text{máx}} \sum_{\omega_y=0}^{\omega_y, \text{máx}} Z_{(\omega_x, \omega_y)}(x, y) \frac{|\tilde{I}(\omega_x, \omega_y)|}{SI} \quad (4.43)$$

onde $SI = \sum_{\omega_x=0}^{\omega_x, \text{máx}} \sum_{\omega_y=0}^{\omega_y, \text{máx}} |\tilde{I}(\omega_x, \omega_y)|$, sendo $\omega_x, \text{máx}$ e $\omega_y, \text{máx}$ as máximas componentes x e y das frequências consideradas no somatório. Com o objetivo de minimizar o ruído decorrente das altas frequências, normalmente nós utilizamos até as 5 componentes de frequência mais baixa.

Implementação do processo

Para implementação do processo acima sumariado, nós avaliamos a profundidade em cada pixel da imagem, aplicando a abordagem multifrequência a uma janela englobando o ponto considerado. Assim, o nosso modelo trabalha com uma janela quadrada deslizando ao longo da imagem, sempre centrada no pixel de interesse (a Figura 4.11 ilustra o processo). A estimativa global para o mapa de profundidades é então obtida, em cada ponto, consolidando-se as estimativas de todas as frequências em uma média, como em (4.43).

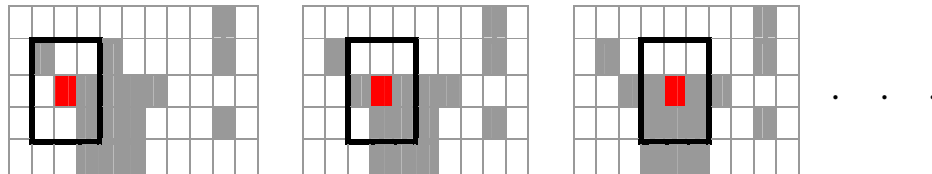


Figura 4.11. Para aplicação da abordagem multifrequência, consideramos janelas em translação ao longo da imagem, centradas no pixel de interesse.

Inicialmente, no cômputo da Transformada de Fourier, nós assumimos o mesmo peso para todos os pixels em uma dada janela, mas uma outra possibilidade que consideramos foi a ponderação gaussiana das intensidades de entrada. Ou seja, nós calculamos as transformadas em cada janela como

$$\tilde{I}(\omega_x, \omega_y) = \iint_{\text{janela}} \exp[-i(\omega_x x + \omega_y y)] [I(x, y) \times W(x, y)] dx dy \quad (4.44)$$

onde $W(x, y)$ é o fator de ponderação, escolhido como uma Gaussiana centrada no pixel de interesse. Nesse caso,

$$W(x, y) = \frac{1}{2\pi\sigma'_x\sigma'_y} \exp[-(\frac{x'^2}{2\sigma_x'^2} + \frac{y'^2}{2\sigma_y'^2})] \quad (4.45)$$

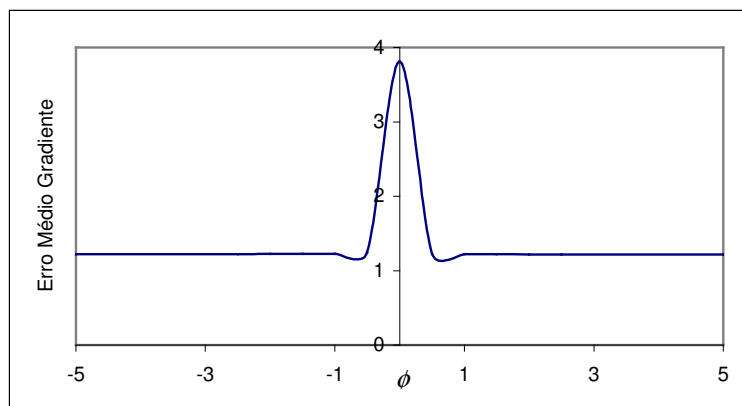
onde $x' = x \cos \theta + y \sin \theta$ e $y' = -x \sin \theta + y \cos \theta$. O ângulo θ dá a orientação do eixo maior da Gaussiana, e $\sigma_x'^2$ e $\sigma_y'^2$ são as respectivas variâncias em x' e y' . Dessa forma, a equação (4.44) se torna

$$\tilde{I}(\omega_x, \omega_y) = \iint_{\text{janela}} \exp[-i(\omega_x x + \omega_y y)] \frac{1}{2\pi\sigma'_x\sigma'_y} \exp[-(\frac{x'^2}{2\sigma_x'^2} + \frac{y'^2}{2\sigma_y'^2})] I(x, y) dx dy \quad (4.46)$$

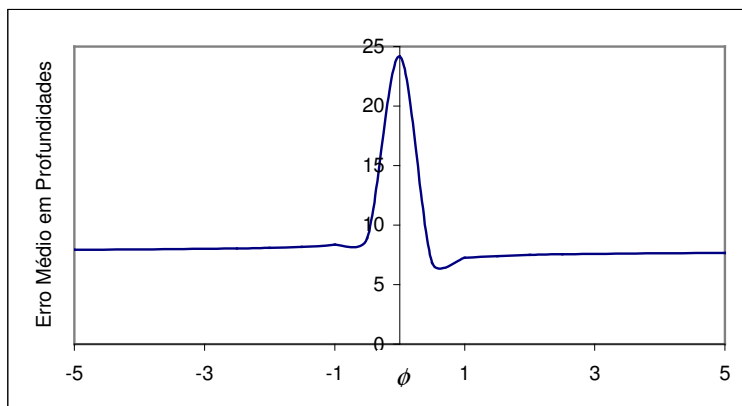
Os valores para σ'_x e σ'_y foram escolhidos de acordo com o tamanho da vizinhança considerada em torno de cada pixel da imagem. Supondo uma superfície Gaussiana centrada no pixel de interesse, ou seja, no centro da janela correspondente, esses valores foram selecionados de modo que a área da janela englobasse, pelo menos, 95% do volume total sob a Gaussiana, considerando-se sempre $\sigma'_x \neq \sigma'_y$. Para a obtenção da nossa estimativa global de profundidade, empregamos diferentes ângulos θ de orientação, consolidando as várias estimativas em uma média aritmética global.

Para o emprego da equação (4.42) na estimação do mapa $Z(x, y)$, a menos do fator de escala α , nós precisamos escolher empiricamente o parâmetro ϕ . Com este objetivo, nós efetuamos uma análise do erro incorrido nas reconstruções para diferentes valores de ϕ , utilizando o mesmo conjunto de imagens empregado na avaliação do parâmetro k_0 do GPM (ver Capítulo 3, Seção 3.3.1). As mesmas medidas de erro foram também utilizadas: erro

médio gradiente e erro médio em profundidades. De forma muito similar ao observado com o GPM, o processo não se mostrou muito sensível ao valor do parâmetro livre, exceto por uma pequena faixa de valores (aproximadamente entre -0.5 e 0.5), como mostram os gráficos na Figura 4.12.



(a)



(b)

Figura 4.12. Erro médio Gradiente (a) e Erro Médio em Profundidades (b), como funções do parâmetro ϕ , obtidos a partir de uma série de experimentos com um conjunto variado de imagens de superfícies lambertianas e quase-lambertianas. Utilizou-se α igual a 1.

Ilustramos, a seguir, em pseudo-código, um esquema para implementação da nossa abordagem de *Shape from Shading* via *Defocus*:

Início

- Entre com a imagem I_1 ;
- Defina o tamanho m das janelas deslizantes;
- Defina o número máximo de frequências a serem utilizadas nas direções x e y , ou seja, $\omega_x máx$ e $\omega_y máx$.
- Defina os valores para os parâmetros α e ϕ (Em geral, utilizamos $\alpha = 1$ e $\phi = 2$);
- Defina o número de orientações a serem utilizadas (N) e os valores θ_i correspondentes, com referência à função gaussiana usada na ponderação das intensidades de entrada;
- Defina também os valores para os parâmetros σ'_x e σ'_y da função gaussiana;
- Crie a matriz Z_{teta} , para acumular as estimativas de profundidades e inicie Z_{teta} com zero;
- Crie a matriz SI de acumulação dos módulos das transformadas;
- Crie a matriz Z' , para representar a estimativa de profundidades em cada orientação;

Para cada orientação θ Faça

- Inicie a matriz SI com zero;
- Inicie a matriz Z' com zero;

Para cada pixel (x,y) da imagem de entrada Faça

- Crie uma subimagem I formada pelos pixels de I_1 correspondentes a uma janela de tamanho $m \times m$, centrada em (x,y) ;
- Obtenha a Transformada de Fourier da subimagem I , observando que as intensidades desta entram no processo com ponderação gaussiana, de acordo com as equações (4.44) e (4.45);

Variando a frequência (ω_x, ω_y) de $(0,0)$ a $(\omega_x máx, \omega_y máx)$

- Obtenha o módulo da Transformada de I em (ω_x, ω_y) - ou seja, $|\tilde{I}(\omega_x, \omega_y)|$;
- Faça: $SI(x,y) \leftarrow SI(x,y) + |\tilde{I}(\omega_x, \omega_y)|$;
- Obtenha $Z_{(\omega_x, \omega_y)}(x, y)$, que representa a estimativa de profundidades em cada frequência, de acordo com a equação (4.42);

Fim_Variando**Variando a frequência (ω_x, ω_y) de $(0,0)$ a $(\omega_x máx, \omega_y máx)$**

- Faça: $Z'(x, y) \leftarrow Z'(x, y) + Z_{(\omega_x, \omega_y)}(x, y) \frac{|\tilde{I}(\omega_x, \omega_y)|}{SI(x, y)}$;

Fim_Variando

- Faça: $Z_{teta}(x,y) \leftarrow Z_{teta}(x,y) + Z'(x,y)$;

Fim_Para

Fim_Para

Para cada pixel (x,y) da imagem de entrada

- Faça: $Z(x,y) \leftarrow Z_{teta}(x,y) / N$;

Fim_Para

- Retorne o mapa de profundidades Z ;

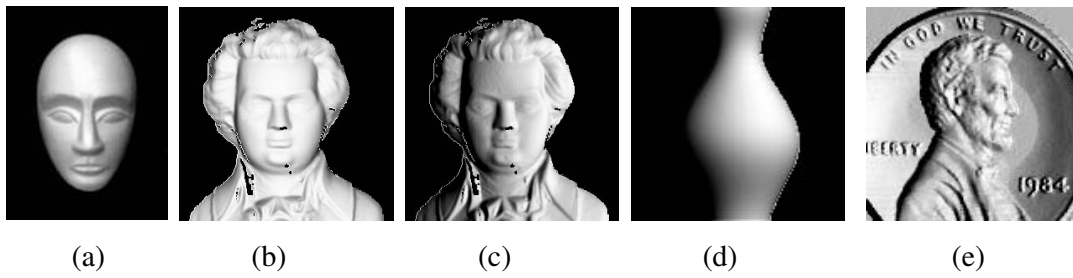
Fim

Experimentos

Os experimentos das Figuras 4.14 a 4.37 ilustram a aplicação do processo. Para cada imagem de entrada, nós consideramos janelas quadradas de 5 x 5 pixels, perfazendo um total de 25 entradas para cada pixel da imagem, utilizadas para o cômputo das Transformadas de Fourier com pesos gaussianos como em (4.46). Os parâmetros σ'_y e σ'_x assumiram valores em torno de 1 e 2 pixels, respectivamente. Para cada orientação do fator de ponderação gaussiano (θ igual a 0° e θ igual a $\pm 45^\circ$) nós utilizamos, no máximo, as cinco componentes de mais baixa frequência da Transformada. Estas foram empregadas na obtenção das estimativas dos mapas de profundidade da cena, de acordo com as equações (4.42) e (4.43), utilizadas com $\alpha = 1$ e $\phi = 2$. O mapa Z final foi obtido pela média aritmética dos resultados para cada orientação. Logicamente, um número maior de orientações poderia ter sido considerado, mas nós verificamos, de forma empírica, que o acréscimo destas ao processo apenas aumentava o seu custo computacional, sem trazer melhoras significativas às reconstruções.

Todos os mapas estimados mostraram-se competitivos com aqueles produzidos pelo GPM, como podemos perceber através dos índices na Figura 4.13 (f) e (g), pelos quais procuramos avaliar numericamente a qualidade das reconstruções produzidas, utilizando imagens de cenas cujos mapas de profundidades são conhecidos. Na tabela, aparecem os índices obtidos tanto com o uso da ponderação gaussiana (SFS via *Defocus* – Pesos Gaussianos) como sem ponderação (SFS via *Defocus* – Pesos Uniformes), verificando-se que, em quase todos os experimentos, a incorporação dos pesos gaussianos mostrou-se vantajosa. Também aqui, como podemos observar nas figuras 4.14 (d), 4.16 (d) e 4.18 (d), com o intuito de validar nosso modelo, recuperamos as intensidades das imagens de entrada, utilizando os parâmetros $\sigma(\omega_x, \omega_y)$, $a(\omega_x, \omega_y)$ e $b(\omega_x, \omega_y)$, obtidos através de (4.36) e (4.37), na equação (4.35).

Com relação aos experimentos com imagens sintéticas, vale destacar a reconstrução do Mozart para iluminação (1,0,1) – Figuras 4.14 e 4.15. Com a nova abordagem, produziu-se um mapa mais uniforme, em comparação com aqueles tipicamente obtidos a partir dessa imagem, que em geral apresentam a superfície com um gradiente irreal de profundidades ao longo da horizontal, devido ao elevado gradiente da irradiância da imagem naquela direção.



<i>Imagens</i>	<i>Método</i>		
	<i>GPM</i>	<i>SFS via Defocus</i> <i>Pesos Gaussianos</i>	<i>SFS via Defocus</i> <i>Pesos Uniformes</i>
<i>Mask1</i>	0,9	0,9	0,9
<i>Mozart1</i>	1,6	1,5	1,7
<i>Mozart2</i>	2,1	2,2	2,6
<i>Vase1</i>	0,6	0,5	0,9
<i>Penny</i>	1,1	0,6	0,7

(f)

<i>Imagens</i>	<i>Método</i>		
	<i>GPM</i>	<i>SFS via Defocus</i> <i>Pesos Gaussianos</i>	<i>SFS via Defocus</i> <i>Pesos Uniformes</i>
<i>Mask1</i>	5,4	7,2	8,1
<i>Mozart1</i>	7,1	6,7	7,6
<i>Mozart2</i>	13,3	9,4	9,5
<i>Vase1</i>	5,1	4,2	4,2
<i>Penny</i>	4,2	4,2	4,3

(g)

Figura 4.13. Análise quantitativa dos erros. (a) a (d): Imagens utilizadas para estimação dos mapas de profundidades: (a) *Mask1*, (b) *Mozart1*, (c) *Mozart2*, (d) *Vase1* e (e) *Penny*; para todas as imagens o mapa real de profundidades é conhecido. (e): Erro médio gradiente das estimativas produzidas pelos três métodos. (f): O mesmo que em (e), para o erro médio em profundidades.

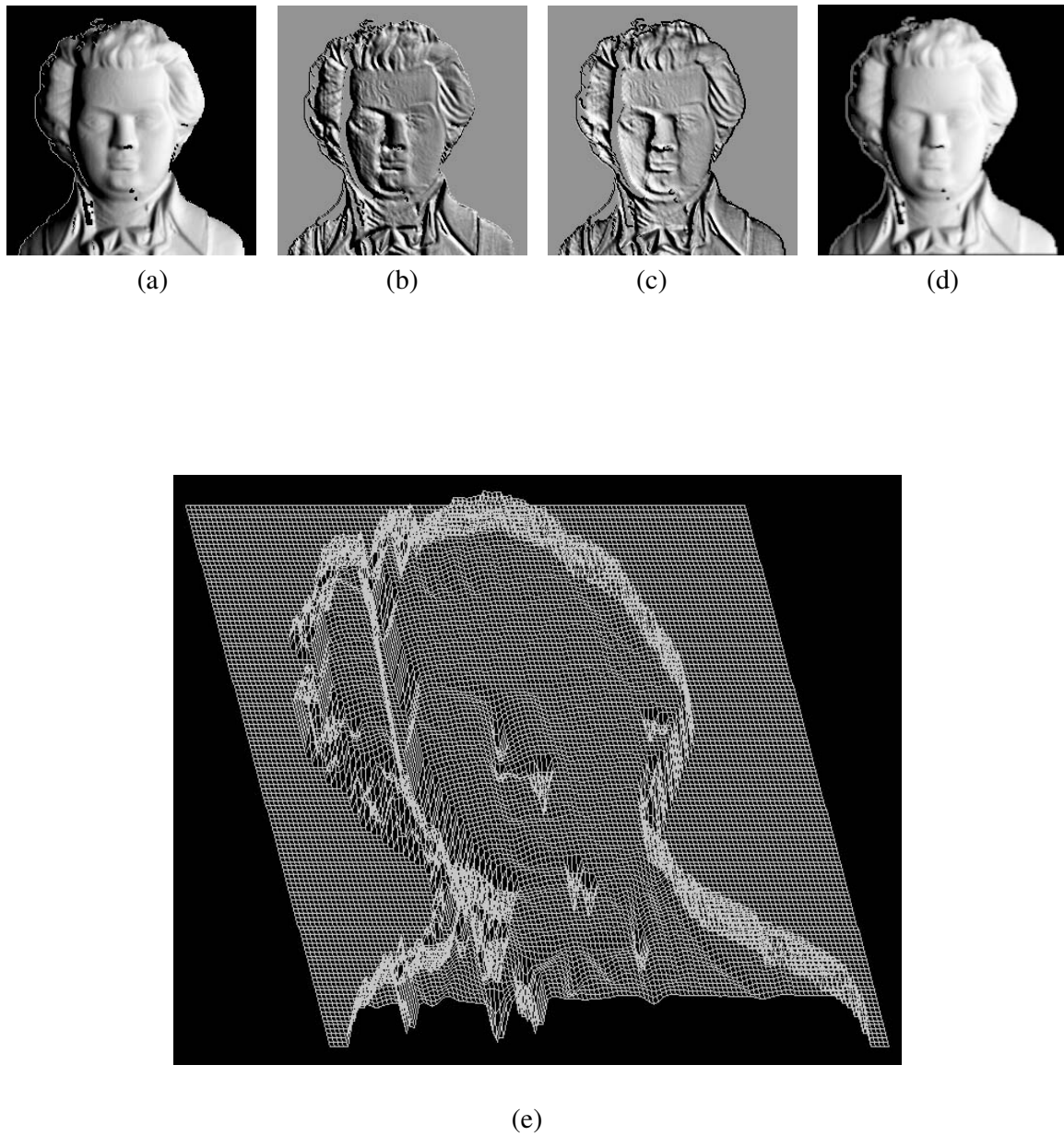
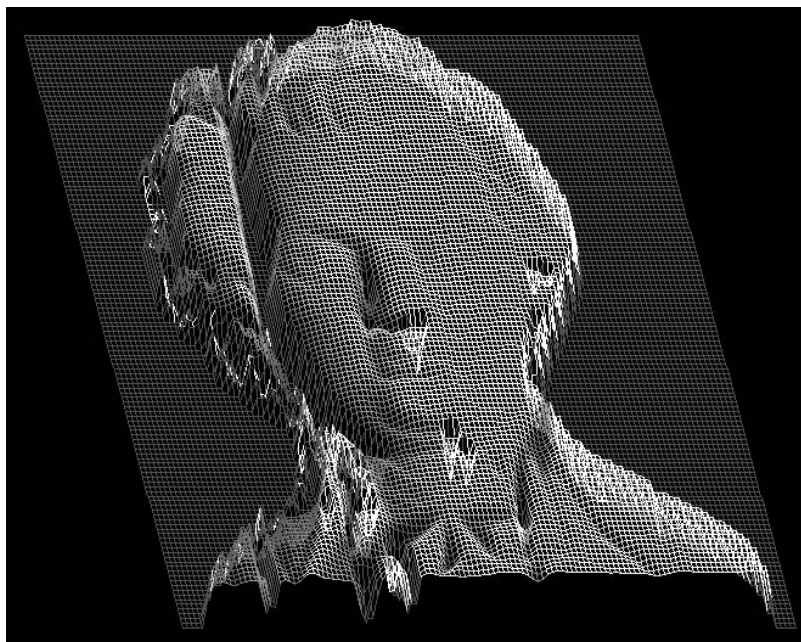
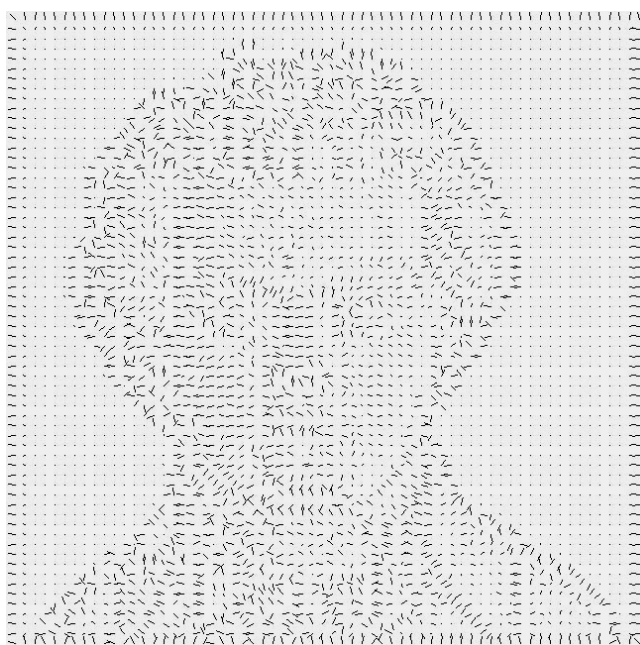


Figura 4.14. (a): Imagem de entrada (*Mozart2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Recuperação das intensidades originais da imagem de entrada, utilizando as equações (4.35), (4.36) e (4.37).



(a)



(b)

Figura 4.15. (a): Profundidades estimadas, referentes à imagem *Mozart2*, com o padrão de irradiância em 4.14-(a) mapeado na representação. (b): Mapa de orientações estimado.

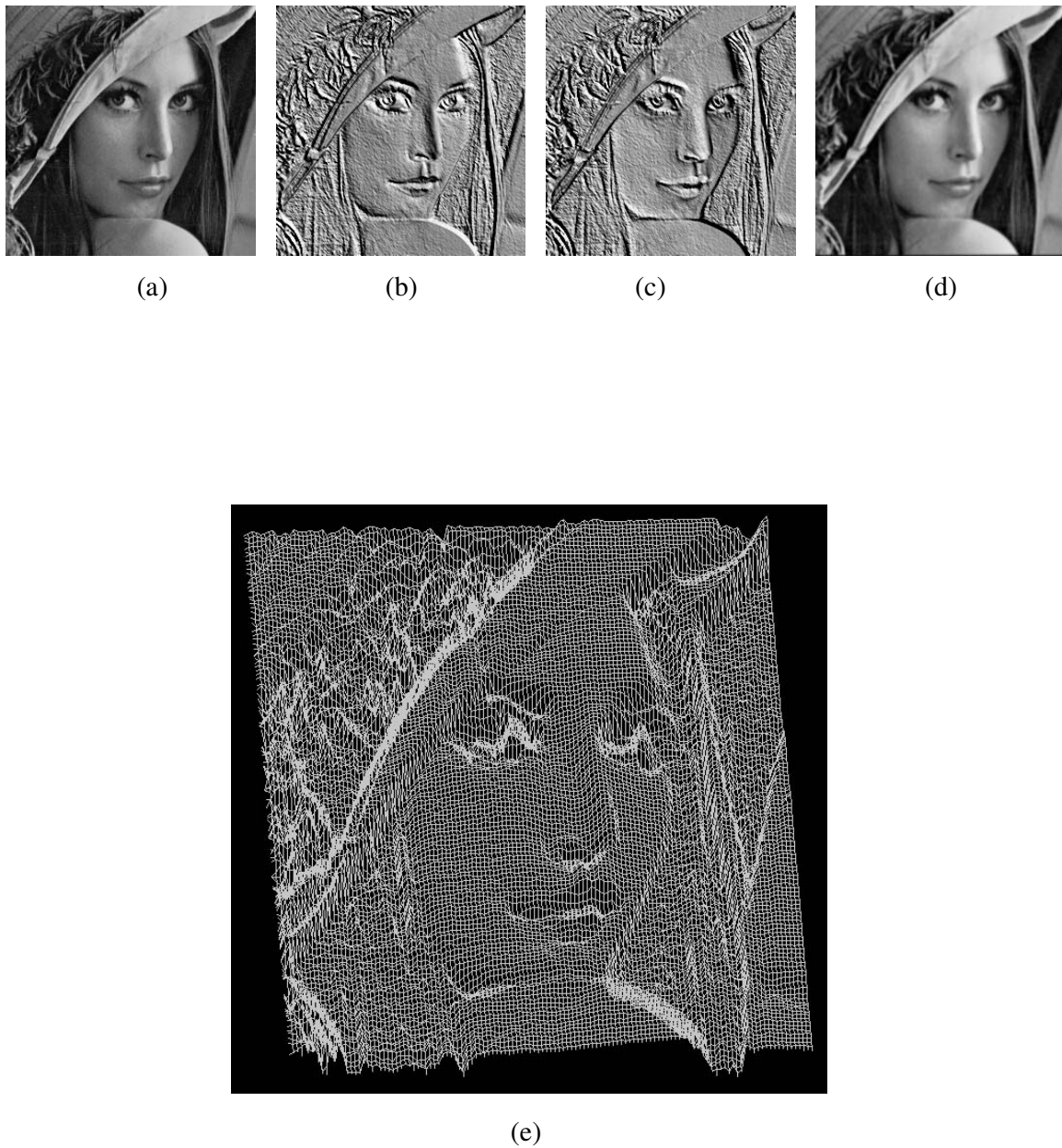
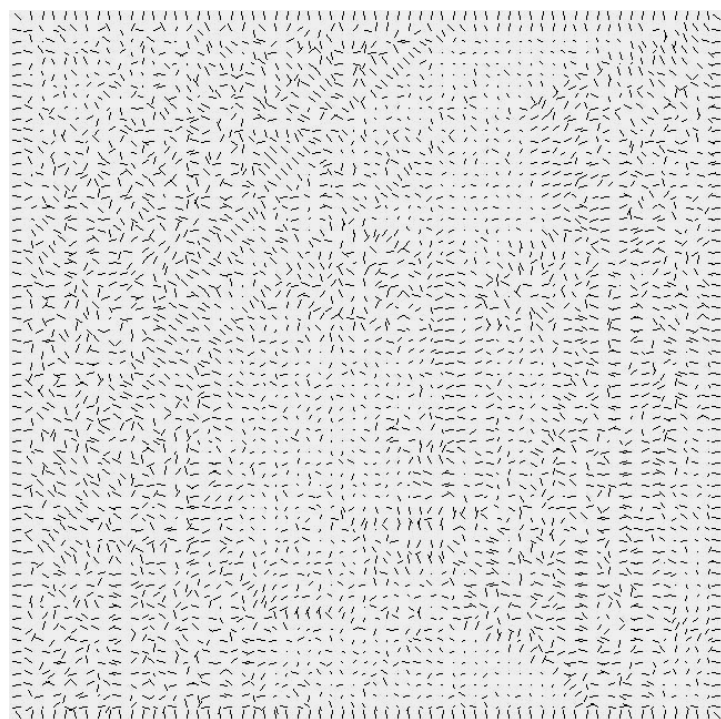


Figura 4.16. (a): Imagem de entrada (*Lena*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Recuperação das intensidades originais da imagem de entrada, utilizando as equações (4.35), (4.36) e (4.37).



(a)



(b)

Figura 4.17. (a): Profundidades estimadas, referentes à imagem *Lena*, com o padrão de irradiância em 4.16-(a) mapeado na representação. (b): Mapa de orientações estimado.

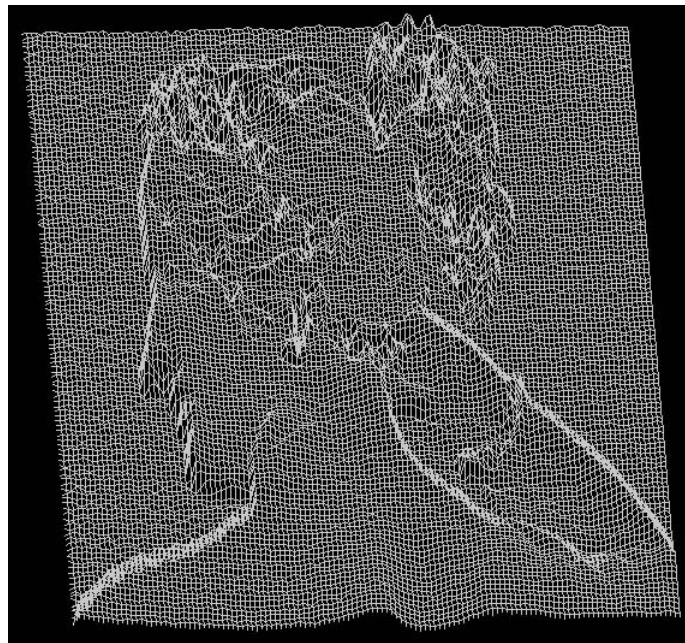


(a)

(b)

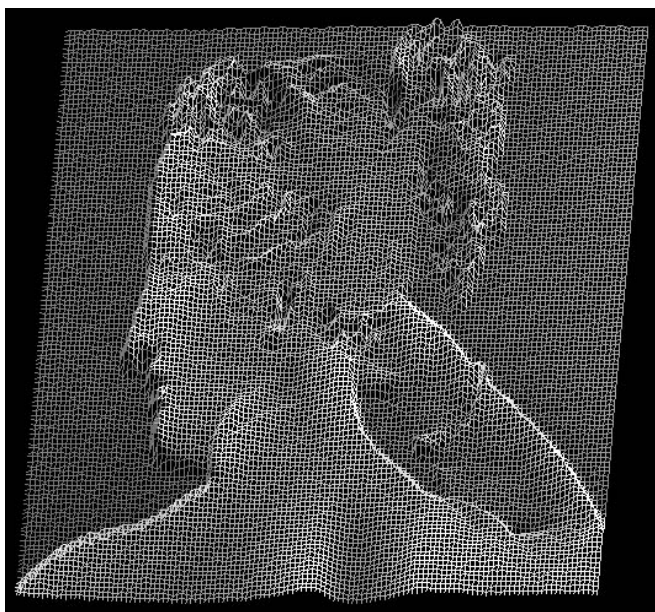
(c)

(d)

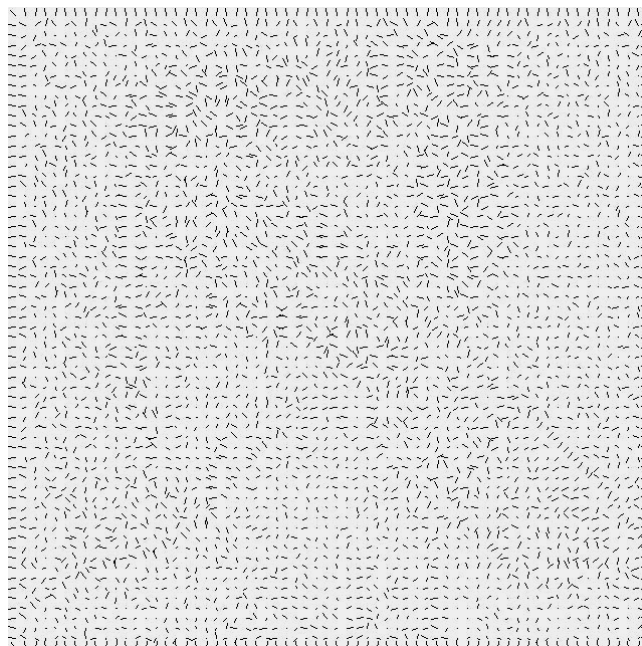


(e)

Figura 4.18. (a): Imagem de entrada (*Paolina*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (e), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$. (d): Recuperação das intensidades originais da imagem de entrada, utilizando as equações (4.35), (4.36) e (4.37).



(a)



(b)

Figura 4.19. (a): Profundidades estimadas, referentes à imagem *Paolina*, com o padrão de irradiância em 4.18-(a) mapeado na representação. (b): Mapa de orientações estimado.

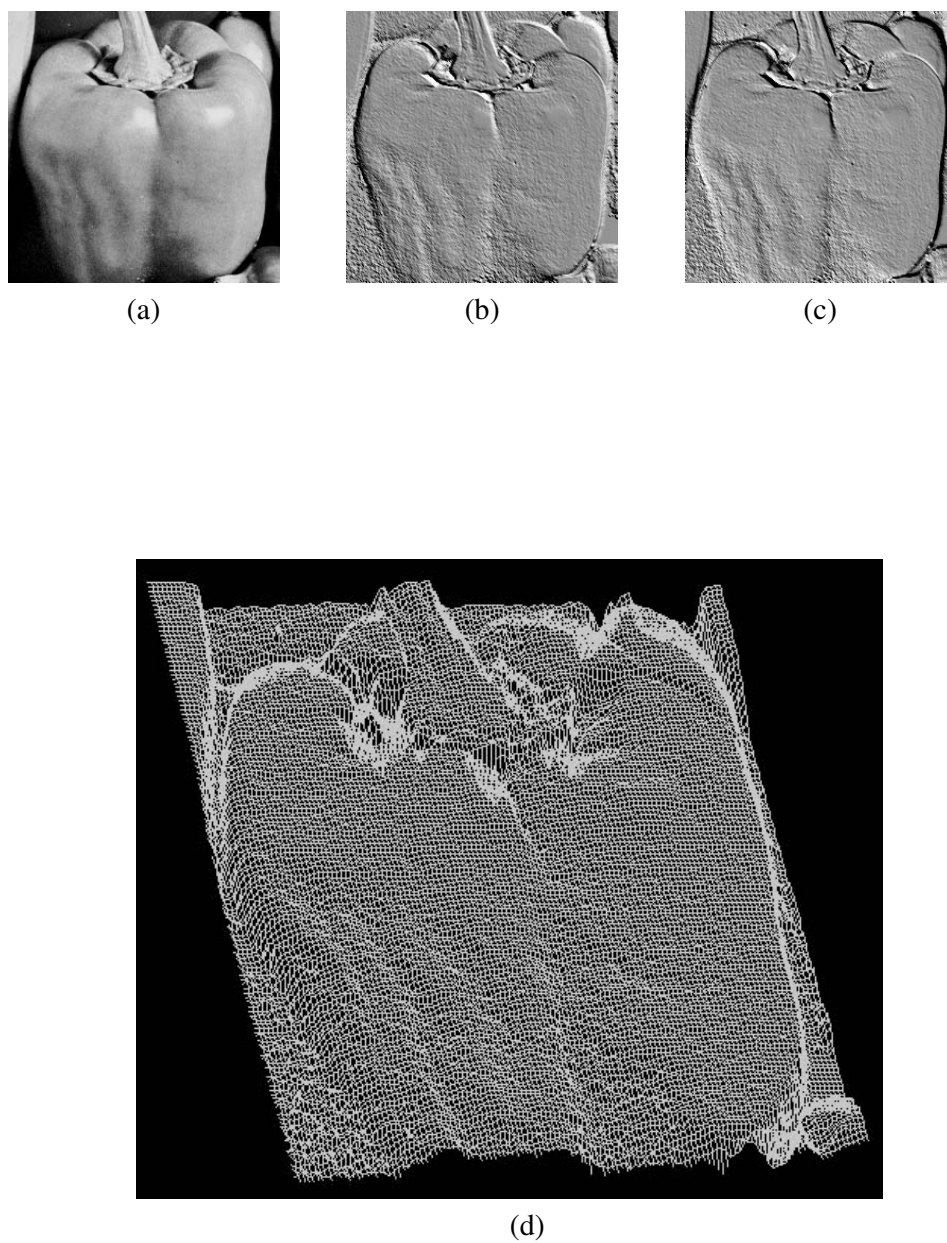
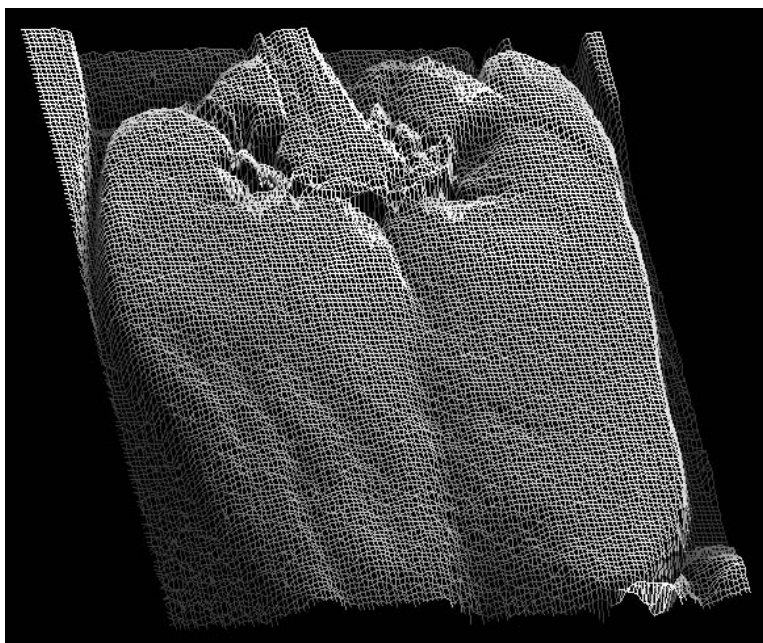
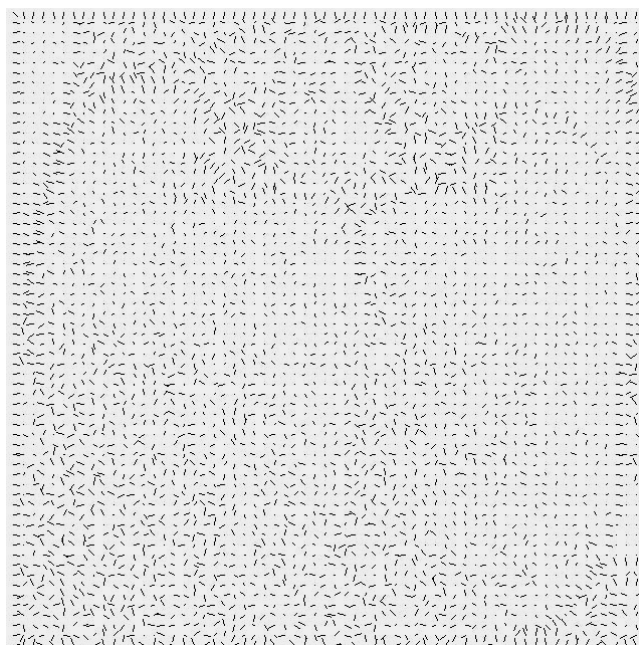


Figura 4.20. (a): Imagem de entrada (*Pepper*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.21. (a): Profundidades estimadas, referentes à imagem *Pepper*, com o padrão de irradiância em 4.20-(a) mapeado na representação. (b): Mapa de orientações estimado.

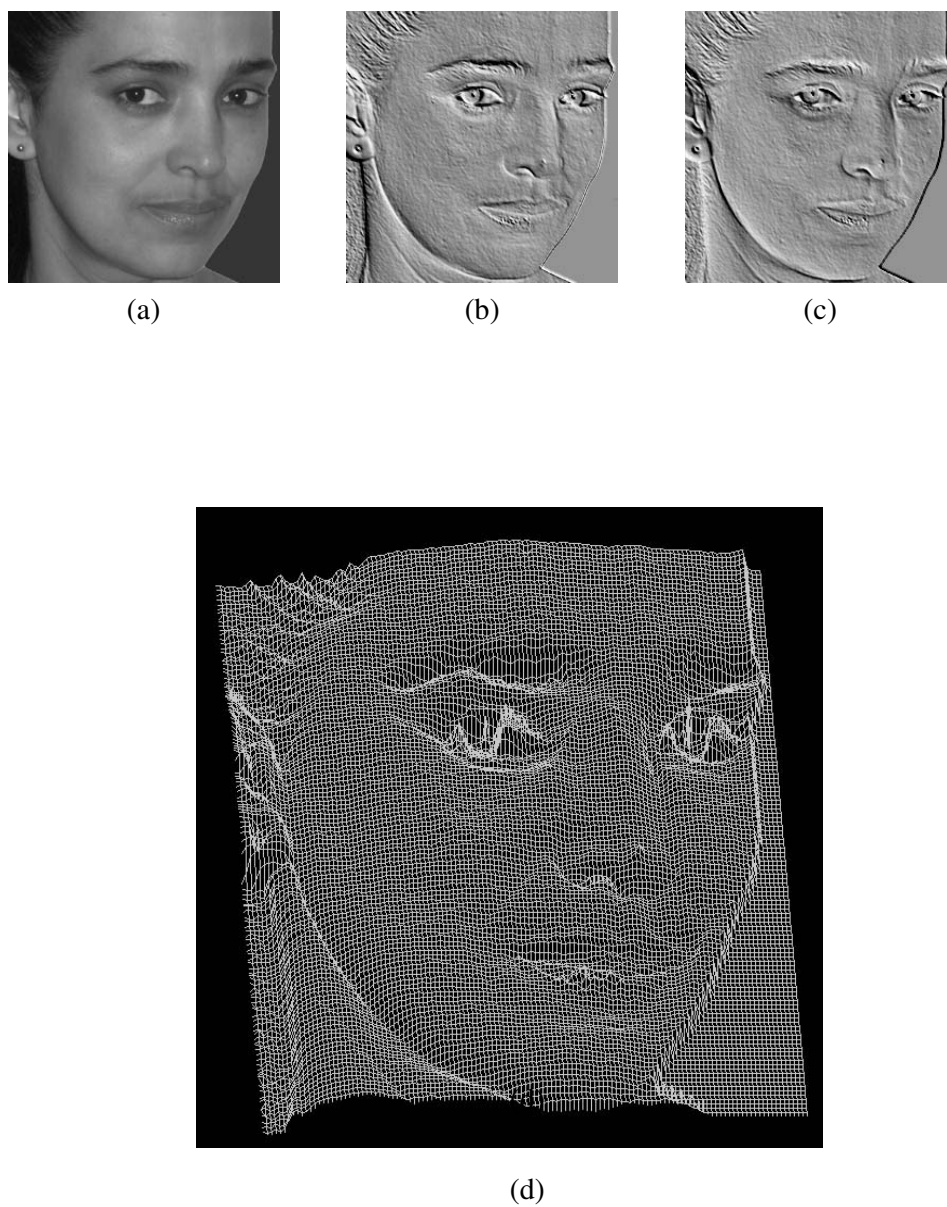
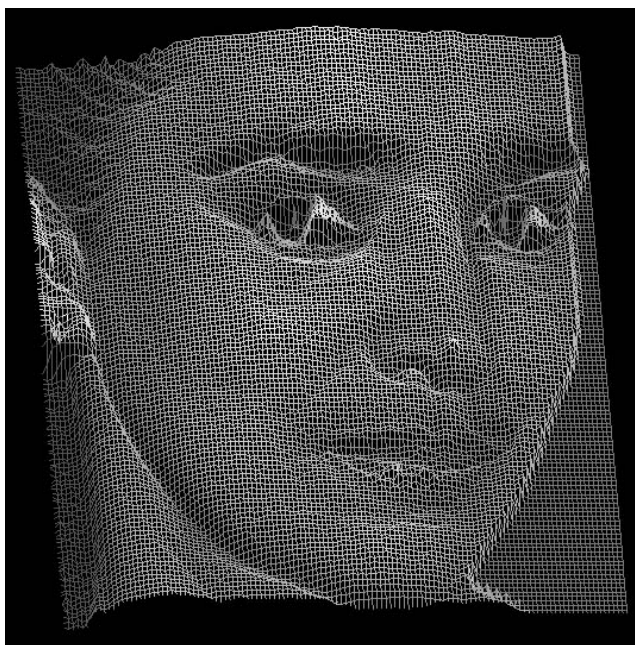


Figura 4.22. (a): Imagem de entrada (*Face2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.23. (a): Profundidades estimadas, referentes à imagem *Face2*, com o padrão de irradiância em 4.22-(a) mapeado na representação. (b): Mapa de orientações estimado.

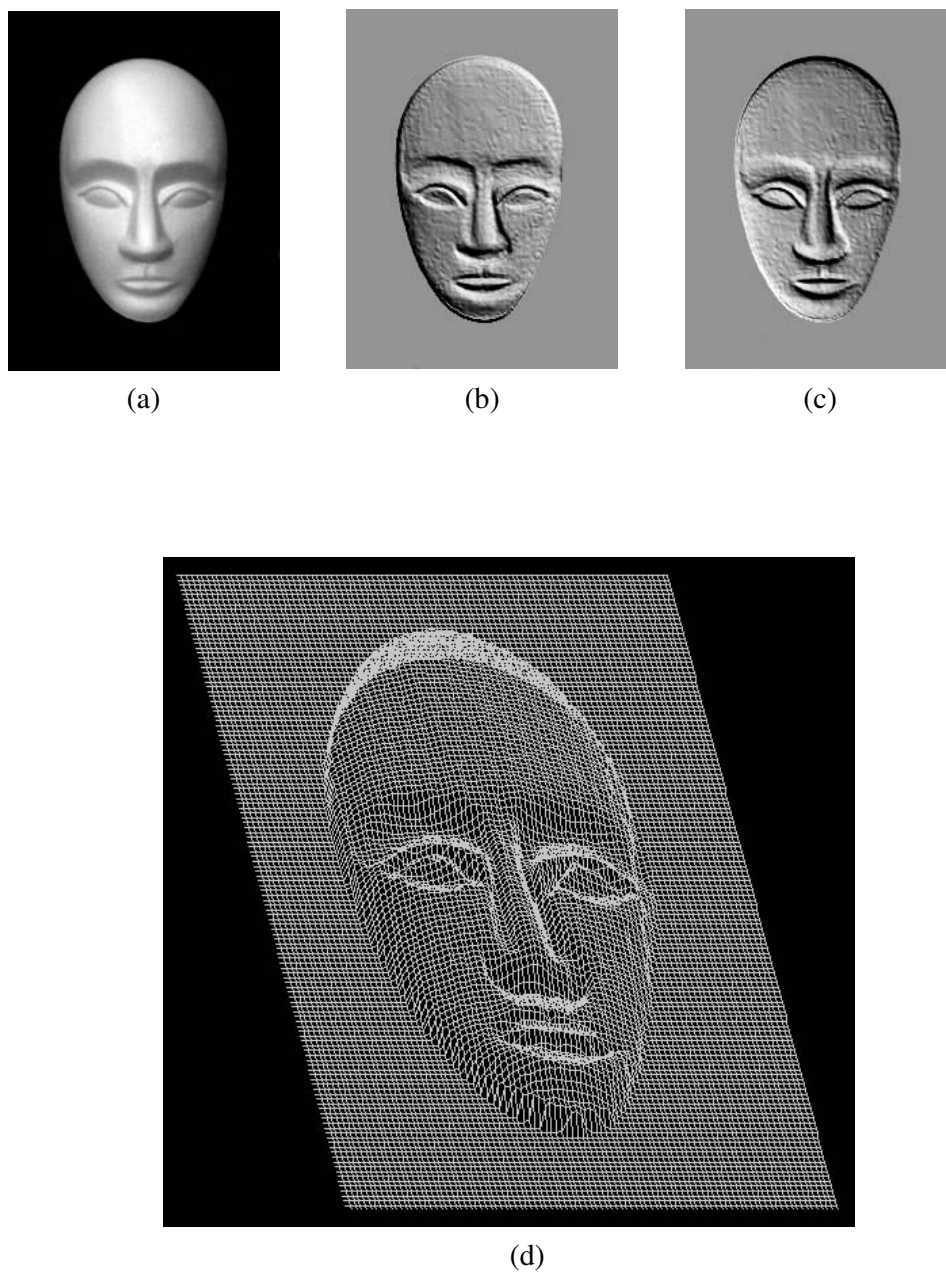
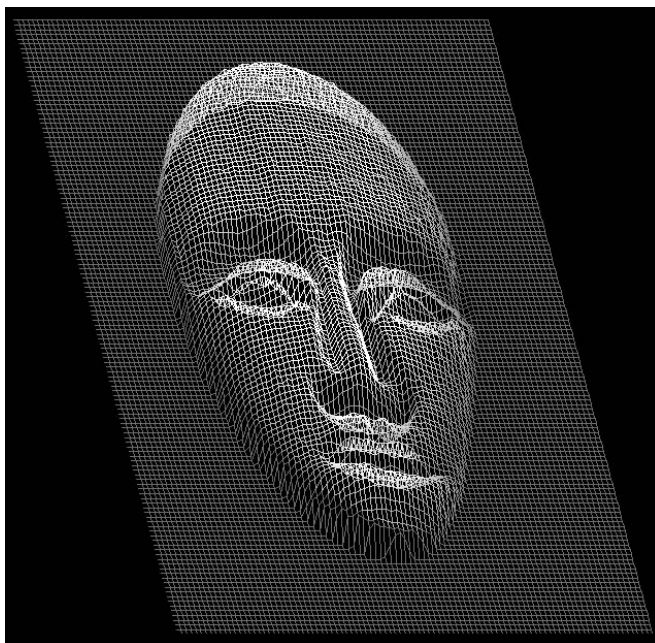
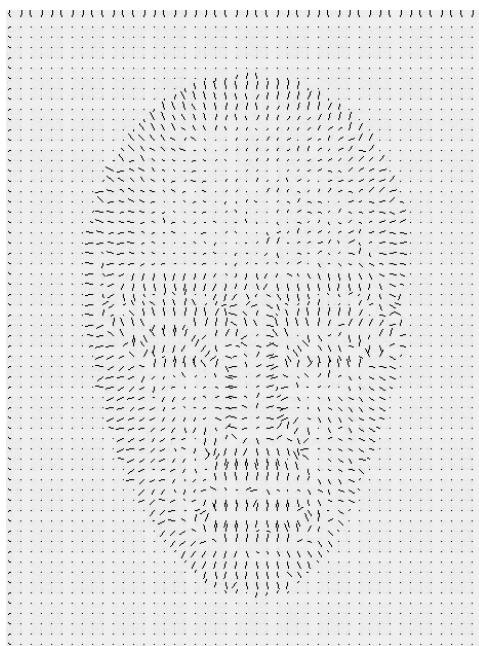


Figura 4.24. (a): Imagem de entrada (*Mask1*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.25. (a): Profundidades estimadas, referentes à imagem *Mask1*, com o padrão de irradiância em 4.24-(a) mapeado na representação. (b): Mapa de orientações estimado.

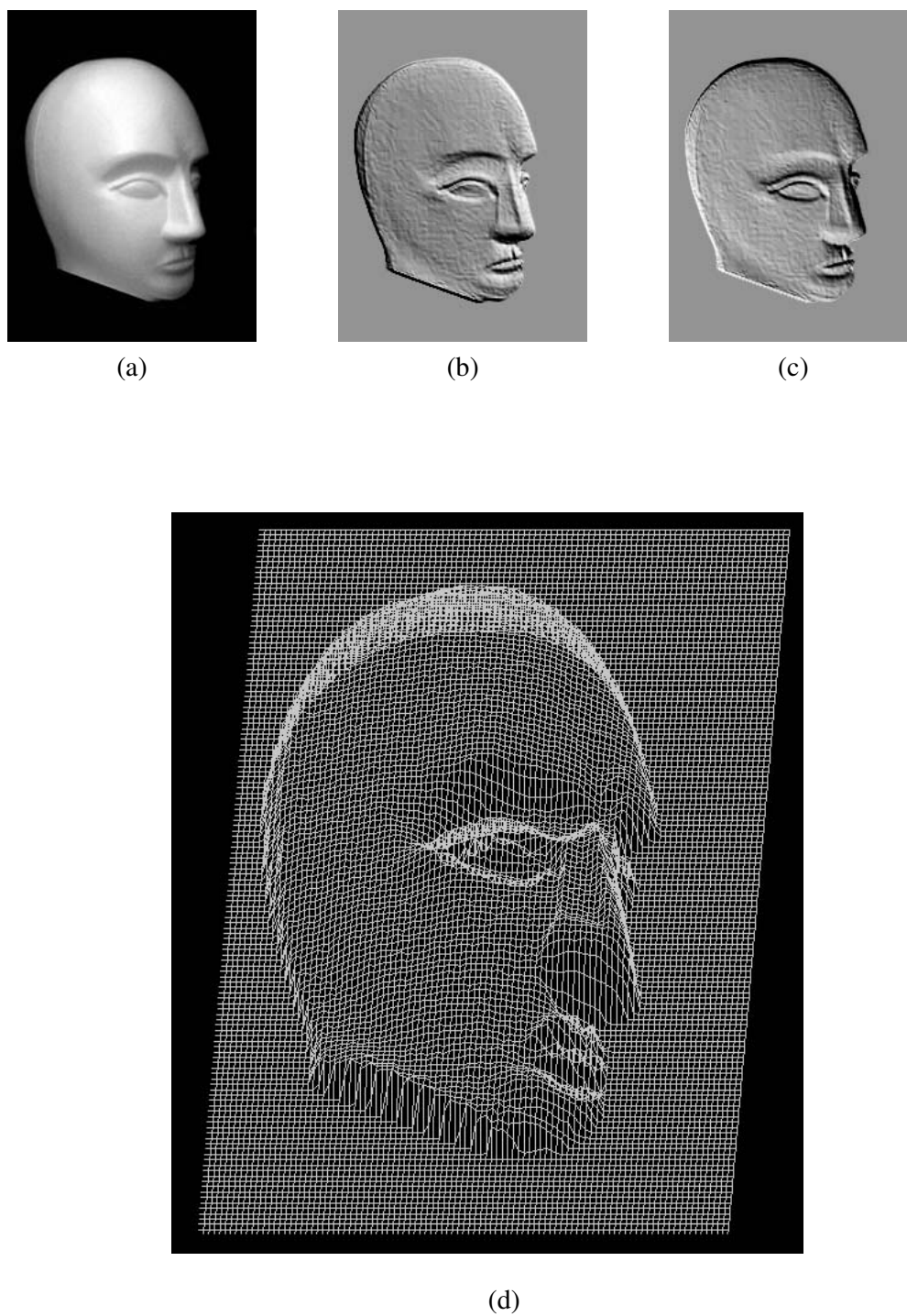
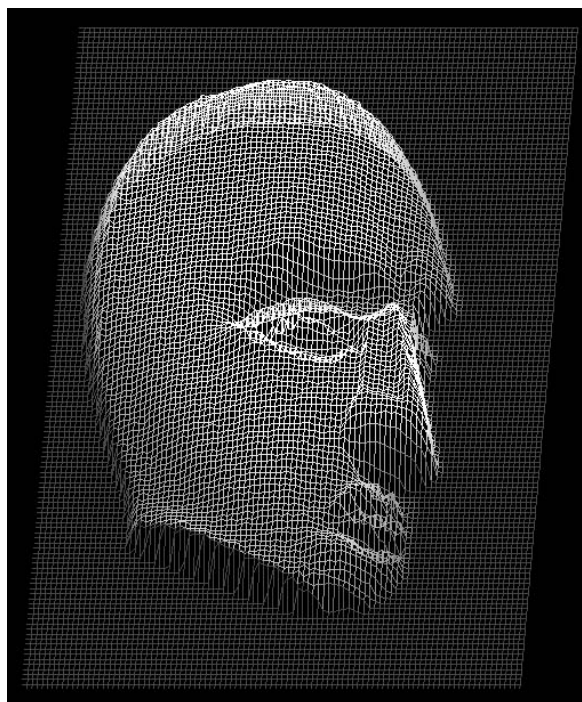
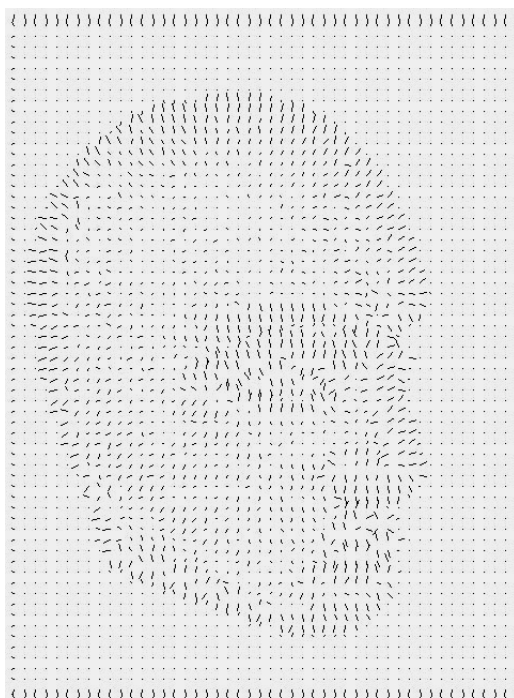


Figura 4.26. (a): Imagem de entrada (*Mask2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.27. (a): Profundidades estimadas, referentes à imagem *Mask2*, com o padrão de irradiância em 4.26-(a) mapeado na representação. (b): Mapa de orientações estimado.

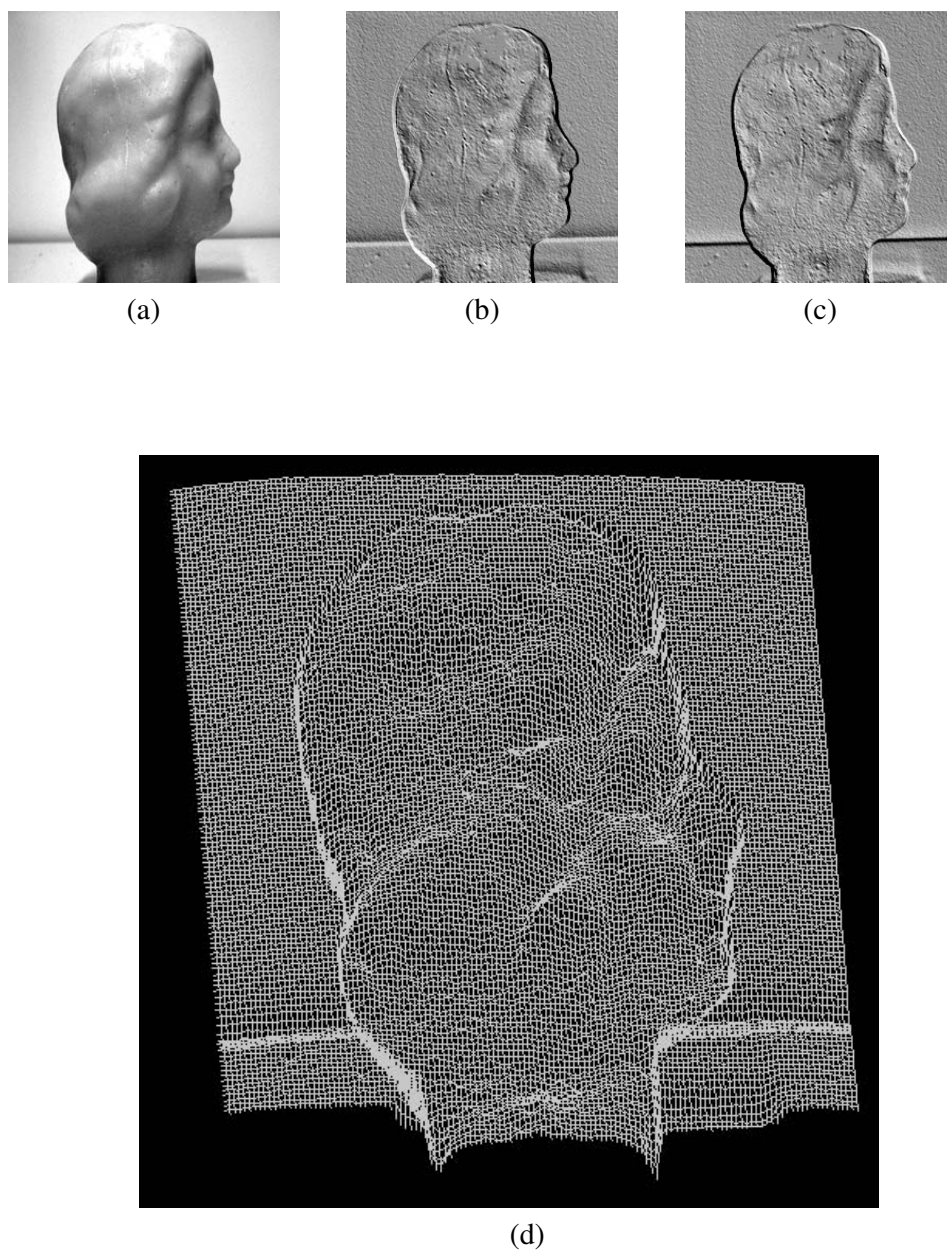
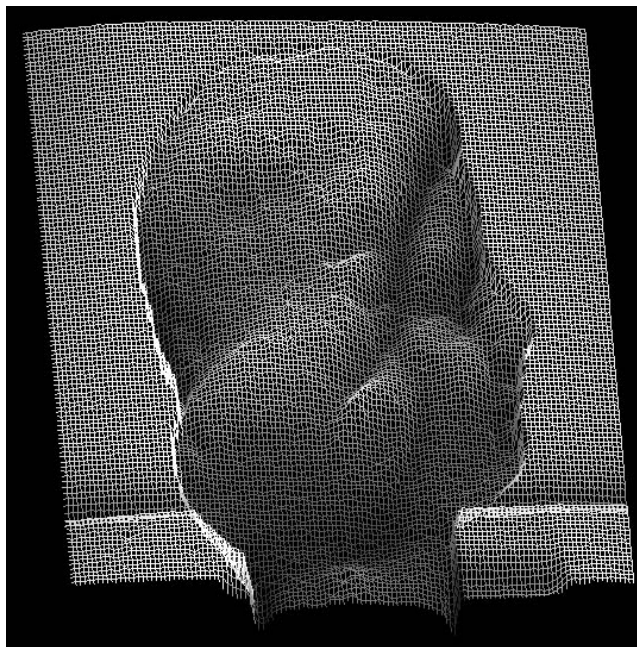
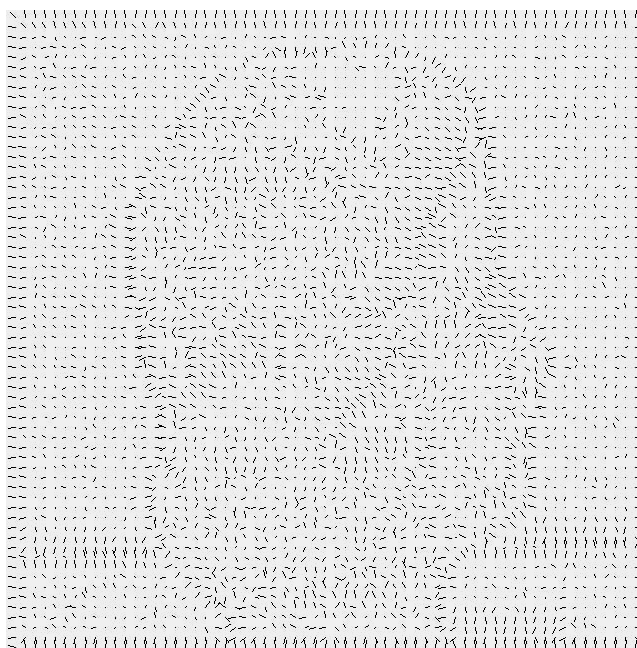


Figura 4.28. (a): Imagem de entrada (*Head1*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.29. (a): Profundidades estimadas, referentes à imagem *Head1*, com o padrão de irradiância em 4.28-(a) mapeado na representação. (b): Mapa de orientações estimado.

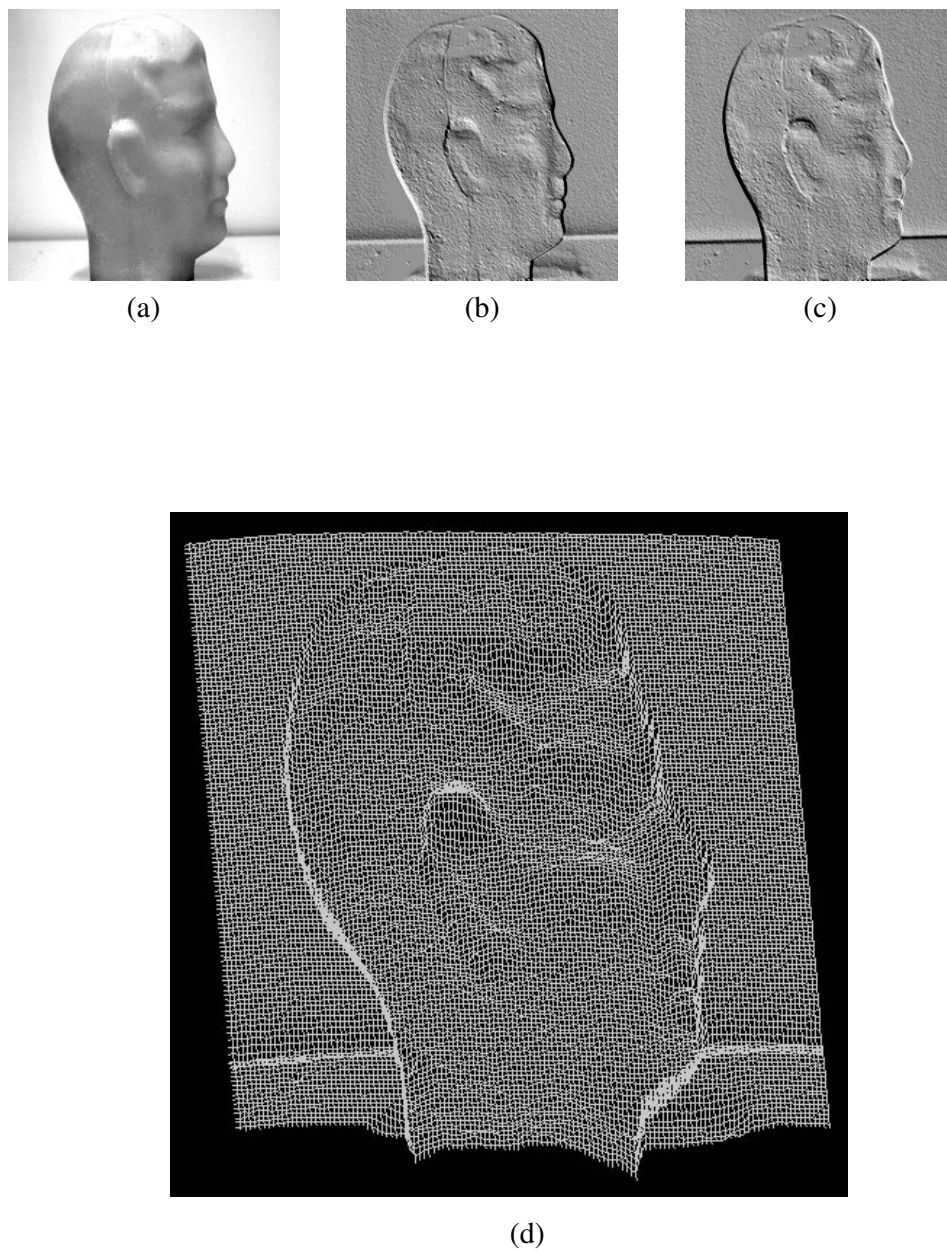
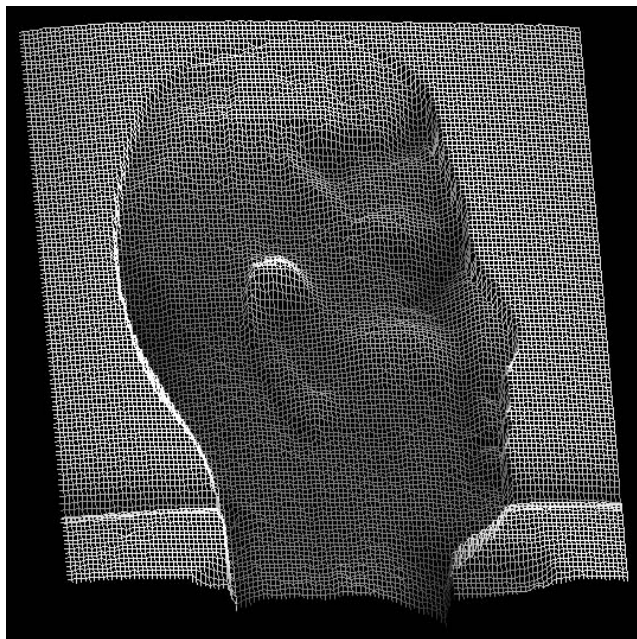


Figura 4.30. (a): Imagem de entrada (*Head2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.31. (a): Profundidades estimadas, referentes à imagem *Head2*, com o padrão de irradiância em 4.30-(a) mapeado na representação. (b): Mapa de orientações estimado.

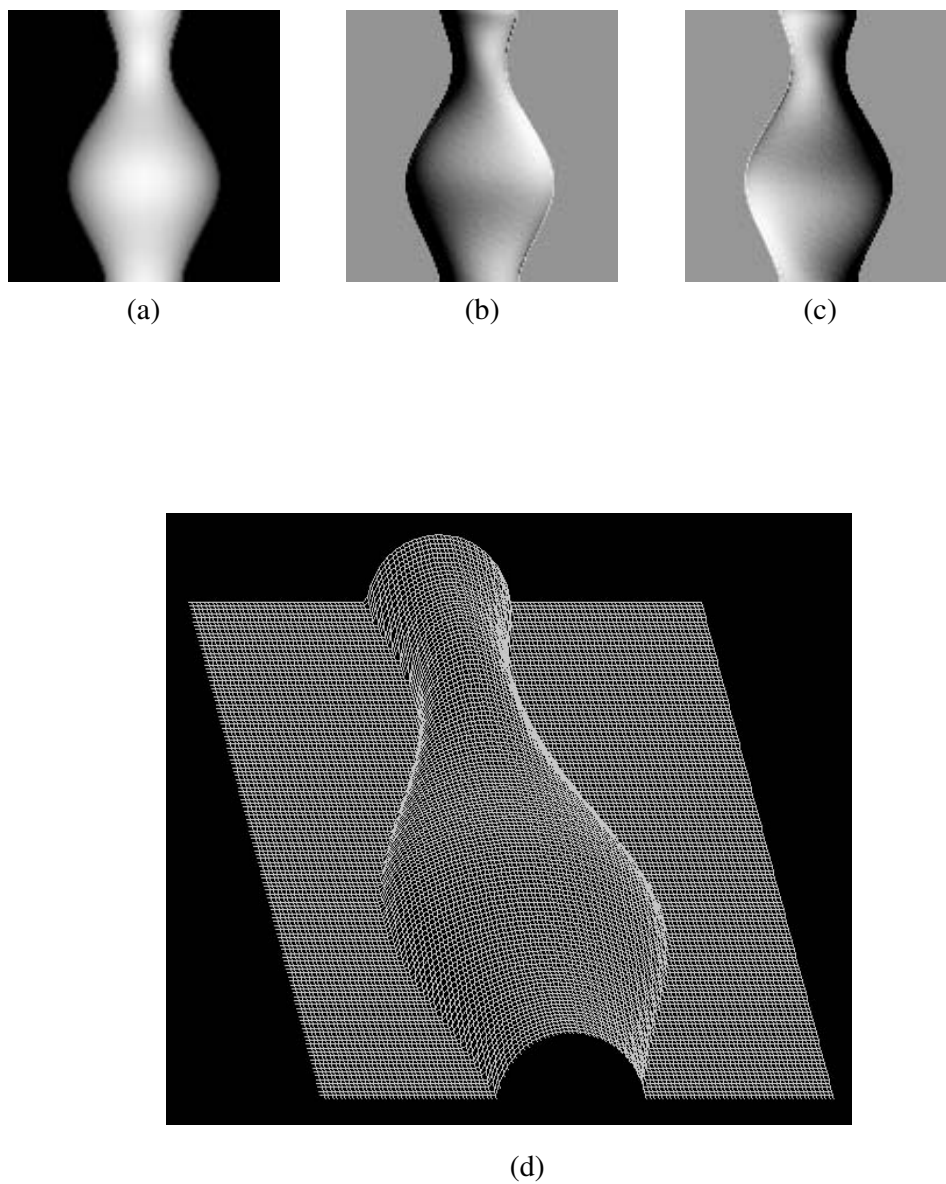
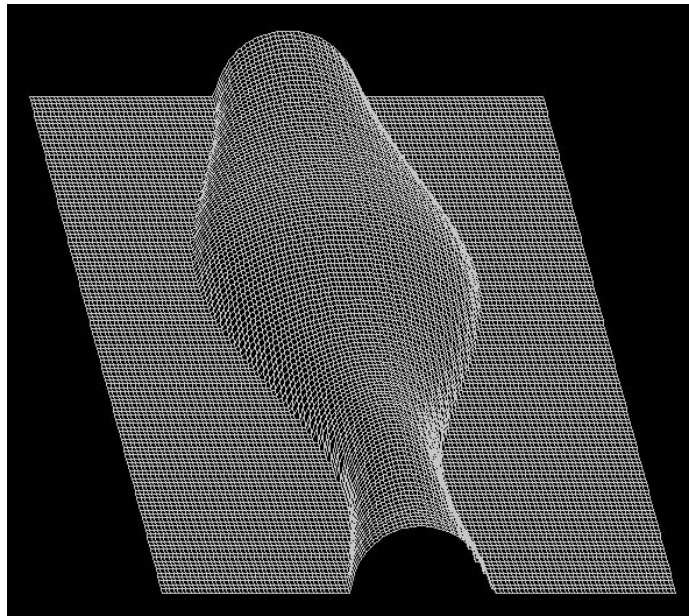
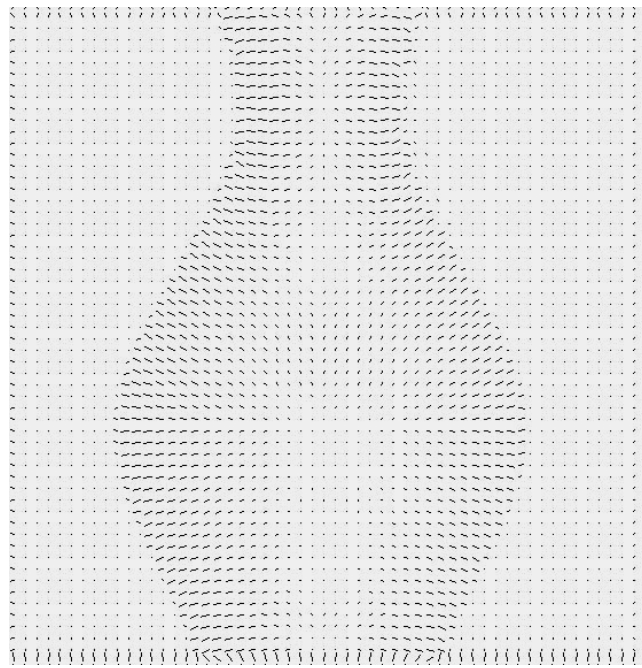


Figura 4.32. (a): Imagem de entrada (*Vase2*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.33. (a): Nova vista para o mapa de Profundidades estimadas, referentes à imagem *Vase2*.
(b): Mapa de orientações estimado.

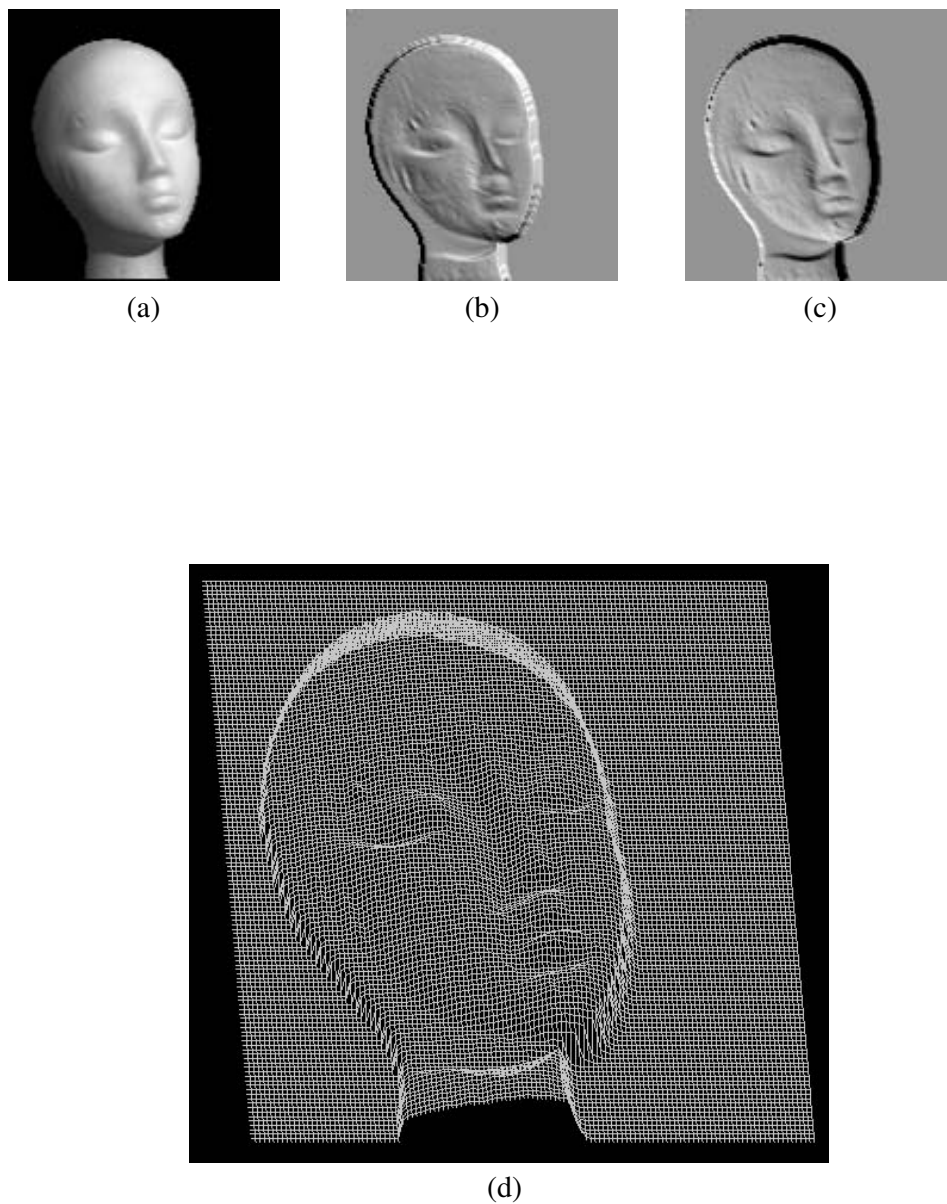
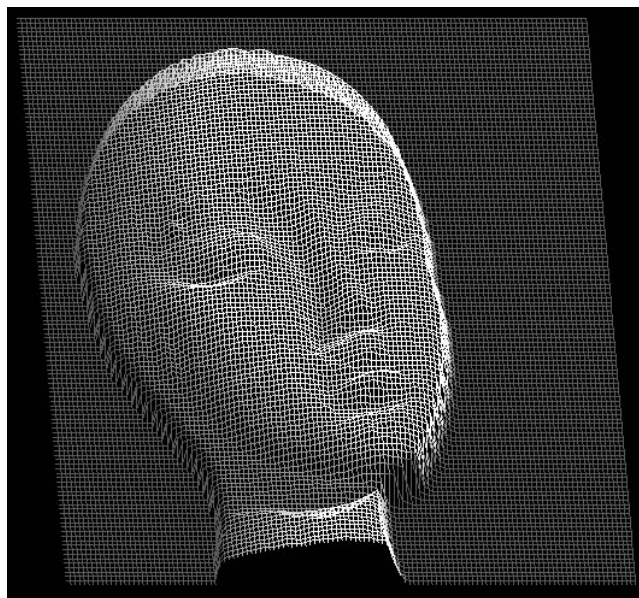
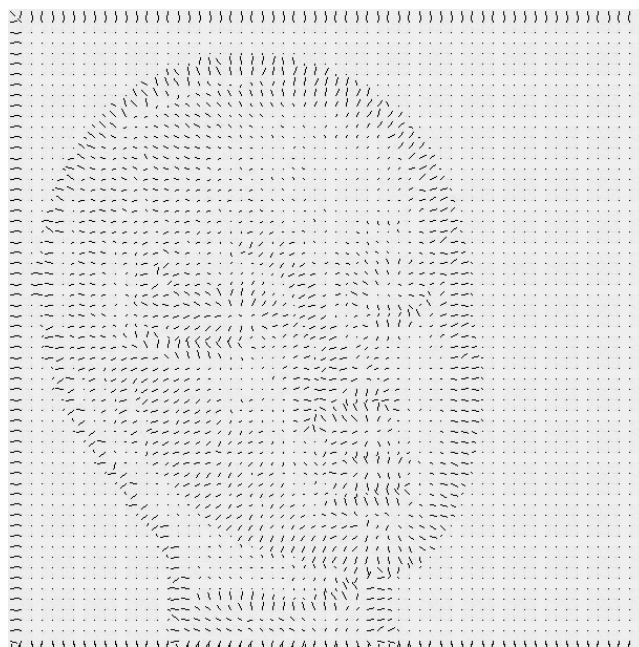


Figura 4.34. (a): Imagem de entrada (*Manequim*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.35. (a): Profundidades estimadas, referentes à imagem *Manequim*, com o padrão de irradiância em 4.34-(a) mapeado na representação. (b): Mapa de orientações estimado.

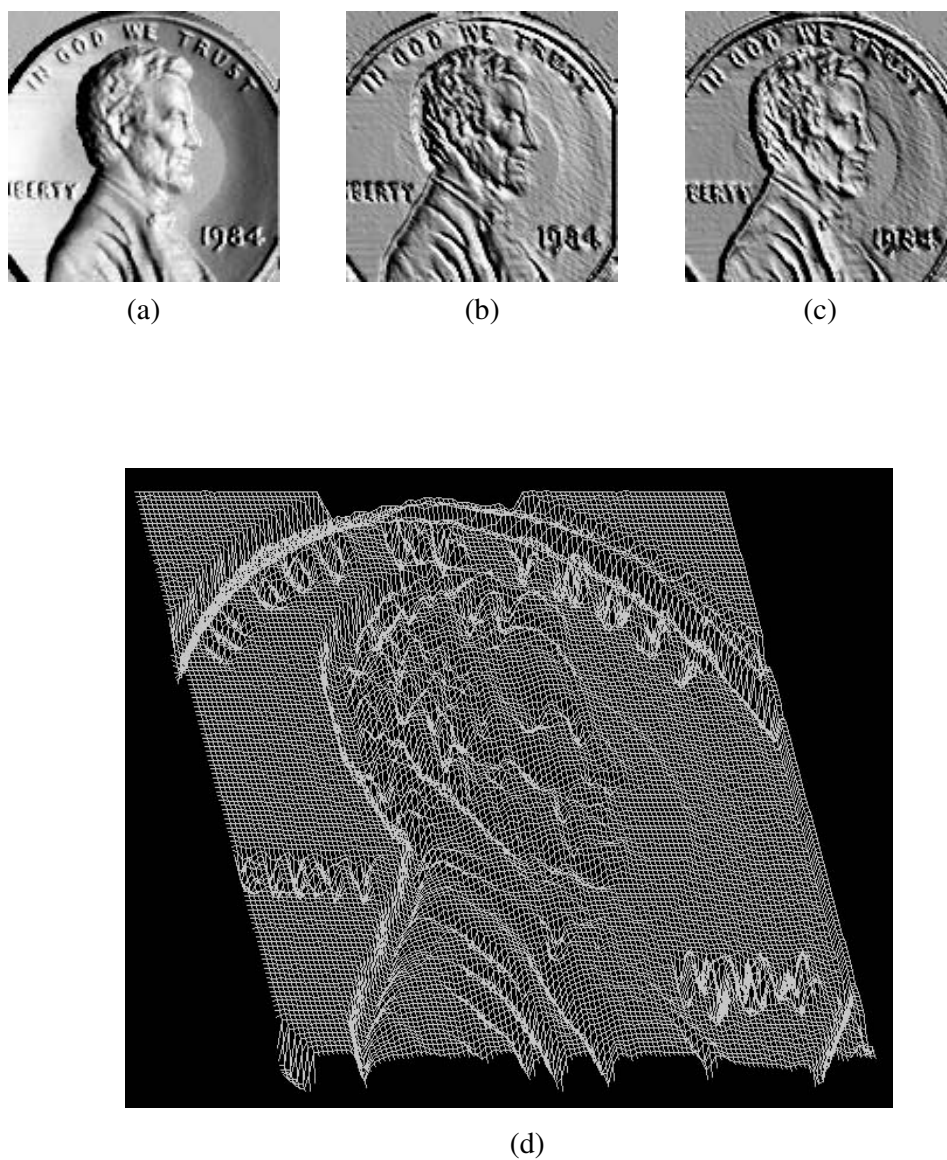
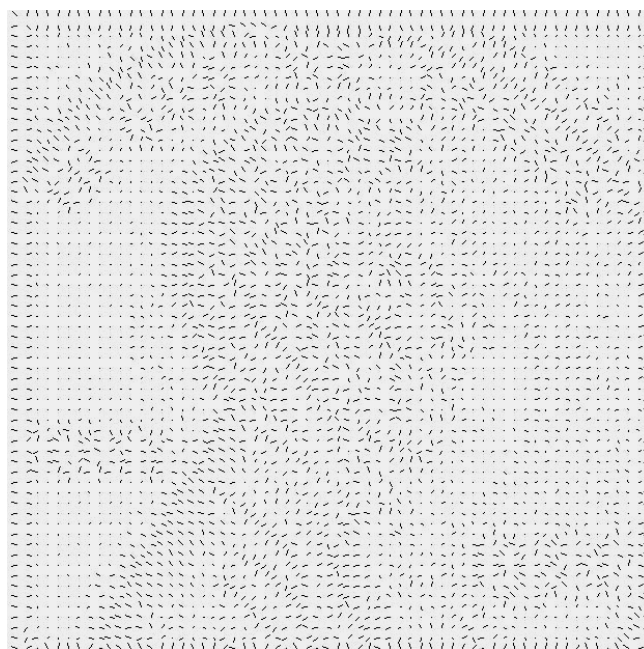


Figura 4.36. (a): Imagem de entrada (*Penny*). (b) e (c): imagens obtidas a partir da re-iluminação da superfície estimada, em (d), reconstruídas sob reflectância lambertiana e albedo uniforme. Direções de iluminação: $(1,1,1)$ e $(-1,-1,1)$.



(a)



(b)

Figura 4.37. (a): Profundidades estimadas, referentes à imagem *Penny*, com o padrão de irradiância em 4.36-(a) mapeado na representação. (b): Mapa de orientações estimado.

CAPÍTULO V

Considerações Finais

As principais contribuições do nosso trabalho de Tese, aqui apresentadas, podem ser sumariadas da seguinte forma:







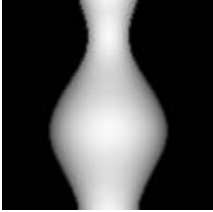
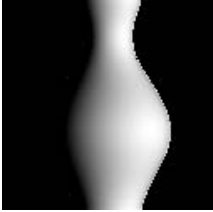

i) Apresentamos uma nova formulação para o *Movimento Fotométrico*, estendendo, para cenas dinâmicas, a mesma abordagem geométrico/fotométrica introduzida no EFBD. A partir de uma seqüência de imagens de uma superfície em rotação, consideramos a variação de intensidade em um ponto fixo no plano da imagem, e não, como na formulação original desenvolvida por Pentland [6], em uma posição fixa na superfície em movimento. Para descrever tal variação de intensidades, introduzimos, de forma similar ao EFBD, uma relação de casamento e uma equação baseada no mapa de reflectância, possibilitando, assim, uma solução aproximada para o mapa de profundidades da superfície. Comparada com a abordagem original do *Movimento Fotométrico*, a nossa tem a vantagem de não requerer *warping* prévio das imagens, sendo, portanto, bem mais simples de implementar.

ii) Partindo da nova formulação para o *Movimento Fotométrico*, estendemos esse tratamento, via Função de *Green* de uma equação de casamento afim, para um processo baseado em uma única imagem de entrada (GPM). Sob condições similares, nossa abordagem para o *Movimento Fotométrico* (com base em múltiplas entradas) produz reconstruções inferiores às obtidas através do GPM. Esse fato pode ser facilmente entendido, já que, no GPM, geramos um par perfeito para a imagem de entrada (através da função de *Green*), sob a condição de casamento afim dada pela equação (3.12). Como vimos na Seção 3.2, a abordagem baseada em imagens reais de uma cena em movimento exige a resolução do problema, nada trivial, de se obter uma estimativa para o fluxo óptico afim. Independentemente do tipo de algoritmo de casamento utilizado, não é difícil aceitar que o par artificialmente gerado, no GPM, produza melhores resultados. Os experimentos relacionados neste trabalho confirmam tal expectativa.

iii) Modificamos a abordagem de Pentland [56] para o *Shape from Defocus* (baseado em múltiplas entradas), considerando funções gaussianas não-normalizadas, ao invés de normalizadas, para a modelagem do processo de desfocamento. O emprego de gaussianas não normalizadas é justificado pela análise do processo de geração das imagens, como aqui demonstrado, e possibilitou a estimação da medida de desfocamento a partir da componente DC da Transformada de Fourier da imagem (obtida localmente em cada *patch*). Além de evitar as componentes indesejáveis de alta frequência, introduzidas pela divisão da imagem em janelas, isto possibilitou a estimação de um mapa denso de profundidades, o que a abordagem de Pentland não permitia.

iv) Propusemos a extensão da nossa abordagem do *Shape from Defocus* para o caso da estimação de forma a partir de uma única imagem de *shading*, numa abordagem, portanto, de *Shape from Shading via Defocus*. Tal extensão foi inspirada pela constatação de que a geração de uma imagem de *shading* pode ser formalmente modelada como um processo de desfocamento semelhante ao produzido por um sistema óptico. A medida de desfocamento associada a um tal processo (modelado através de funções de Gabor não normalizadas) pode ser estimada localmente assim como no *Shape from Defocus* multi-imagem, desta forma possibilitando a estimação das profundidades da cena representada.

Através da experimentação exaustiva realizada ao longo deste trabalho, podemos constatar que as duas novas abordagens aqui introduzidas produzem resultados bastante competitivos com os de abordagens tradicionais do SFS. Alguns desses resultados já foram discutidos neste documento, em seus capítulos 3 e 4. Para uma avaliação mais completa, apresentamos aqui todo o conjunto dos experimentos realizados. Na Figura 5.1, abaixo, mostramos a série de imagens sintéticas consideradas. Para os experimentos com as figuras *Mozart*, *Penny* e *Vase*, empregamos três direções de iluminação: *S1* (0,0,1), *S2* (1,0,1) e *S3* (5,5,7). Para as figuras *Mask1* e *Mask2*, consideramos apenas a iluminação *S1*. Na Figura 5.2, apresentamos os nossos resultados comparativos, relacionando os índices de erro obtidos para estimativas referentes às nossas abordagens para o SFS e às de Lee/Rosenfeld e Tsai/Shah. Deve-se observar que utilizamos as mesmas medidas de erro descritas no Capítulo 3 – Seção 3.3.1: erro médio do gradiente e erro médio em profundidades. Pela análise numérica dos resultados mostrados na Figura 5.2, podemos perceber que, numa média geral considerando todos os experimentos, nossas abordagens apresentaram melhores estimativas.

<i>Imagens</i>	<i>Iluminações</i>		
	<i>S1</i>	<i>S2</i>	<i>S3</i>
<i>Mozart</i>			
<i>Penny</i>			
<i>Vase</i>			



<i>Iluminação</i>	<i>Imagens</i>	
	<i>Mask1</i>	<i>Mask2</i>
<i>S1</i>		

Figura 5.1. Imagens de entrada, utilizadas para avaliação quantitativa do erro nas estimativas de profundidade. *S1*, *S2* e *S3* referem-se, respectivamente, às direções de iluminação (0,0,1), (1,0,1) e (5,5,7).

Imagens		Métodos			
Nome	Ilumin.	GPM	SFS via Defocus	Lee/Rosenfeld	Tsai/Shah
Mozart	S1	1,6	1,5	1,6	1,6
	S2	2,1	2,2	4,5	2,1
	S3	2,2	2,1	4,3	2,6
Penny	S1	0,9	1,1	1,2	1,4
	S2	1,1	0,6	2,2	1,6
	S3	1,1	0,6	2,2	1,6
Vase	S1	0,3	0,3	0,2	0,5
	S2	0,6	0,5	0,8	0,6
	S3	0,6	0,5	0,7	0,6
Mask1	S1	0,9	0,9	1,0	1,0
Mask2	S1	1,0	1,0	1,0	1,0

(a)

Imagens		Métodos			
Nome	Ilumin.	GPM	SFS via Defocus	Lee/Rosenfeld	Tsai/Shah
Mozart	S1	7,1	6,7	7,2	7,6
	S2	12,7	9,4	12,3	15,9
	S3	12,7	9,1	12,7	15,4
Penny	S1	4,1	4,1	4,1	4,1
	S2	4,2	4,2	4,2	4,2
	S3	4,2	4,2	4,2	4,2
Vase	S1	2,3	2,3	2,3	3,3
	S2	5,7	4,8	3,4	6,7
	S3	5,1	4,2	6,5	5,8
Mask1	S1	5,4	7,2	8,0	7,9
Mask2	S1	7,7	8,5	8,9	8,7

(b)

Figura 5.2. Análise quantitativa dos erros, para as imagens da Figura 5.1. (a): Erro médio do gradiente das estimativas produzidas pelos quatro métodos. (b): O mesmo que em (a), para o erro médio em profundidades. Para nosso método de SFS via *Defocus*, o processo foi implementado utilizando pesos gaussianos, conforme descrito no Capítulo 4 – Seção 4.4.2.

Com relação à implementação dos processos, destacamos, no GPM, o modelo iterativo para estimação dos coeficientes lineares do mapa de reflectância. Esse procedimento não só

permitiu uma independência em relação ao conhecimento do verdadeiro mapa de reflectância da superfície, mas também possibilitou ganhos com relação à qualidade das reconstruções. Já em nosso método para SFS via *Defocus*, obtivemos ganhos expressivos através da introdução da abordagem multifrequência. A utilização de outras componentes de frequência, além da componente DC, na composição da estimativa final para o mapa de profundidade, melhorou consideravelmente as nossas estimativas – levando a que, na maioria dos experimentos, essa abordagem produzisse melhores resultados do que o GPM. Ainda com relação à implementação dos processos, acrescentamos que, mesmo sendo um processo iterativo, o GPM produz suas estimativas com um tempo de processamento bem menor do que aquele necessário ao SFS via *Defocus*. Isto pode ser justificado considerando-se que o último processo utiliza Transformada de Fourier da imagem de entrada, o que provoca um acréscimo no tempo de processamento.

É interessante observar ainda que a estimação das profundidades, na nossa abordagem do SFS via *Defocus*, se dá de acordo com o esquema da Figura 5.3 abaixo, que é formalmente o mesmo utilizado para a obtenção da saída de uma rede neural artificial [69]. Explicitamente, a resposta Z , para uma certa frequência e em um dado ponto da imagem, é obtida a partir da soma de contribuições das entradas (intensidades) na vizinhança deste ponto, cada uma multiplicada pelo fator de ponderação P . Esta soma ponderada, que define o equivalente ao nível de ativação (S) do neurônio, é então submetida a uma transformação não-linear, representada pela função de ativação f , gerando assim a resposta $Z=f(S)$. No caso do SFS via *Defocus*, o nível de ativação corresponde à transformada de Fourier $\tilde{I}(\omega_x, \omega_y)$, calculada de acordo com a equação (4.44), identificando-se portanto os pesos P como sendo $\exp[-i(\omega_x x + \omega_y y)] W(x,y)$ dentro da janela, e nulos fora dela. A função de ativação, por sua vez, corresponde ao lado direito da equação (4.42), ou seja,

$$f(S) = \alpha \frac{1}{\phi - \sqrt{|S|}} \quad (5.1)$$

Note-se que, em nossos experimentos, nós utilizamos $\phi=2$, e $\alpha=1$. Para tal escolha de parâmetros, nós podemos identificar $F(S)$ como sendo uma aproximação, quando $|S|$ assume valores pequenos, da comumente empregada função logística,

$$y(S) = \frac{1}{1 + e^{-|S|}} \quad (5.2)$$

O estimador de profundidades correspondente ao SFS via *Defocus* difere, portanto, do modelo tradicional das redes neurais artificiais, essencialmente pela utilização de pesos complexos e pela presença do módulo do nível de ativação na função f . Explorar este paralelo com redes neurais é uma das extensões do presente trabalho que podemos considerar futuramente, por exemplo, investigando a possibilidade de aprendizado no nosso esquema. Isto poderia ser útil para a obtenção de configurações mais adequadas para os pesos $W(x,y)$.

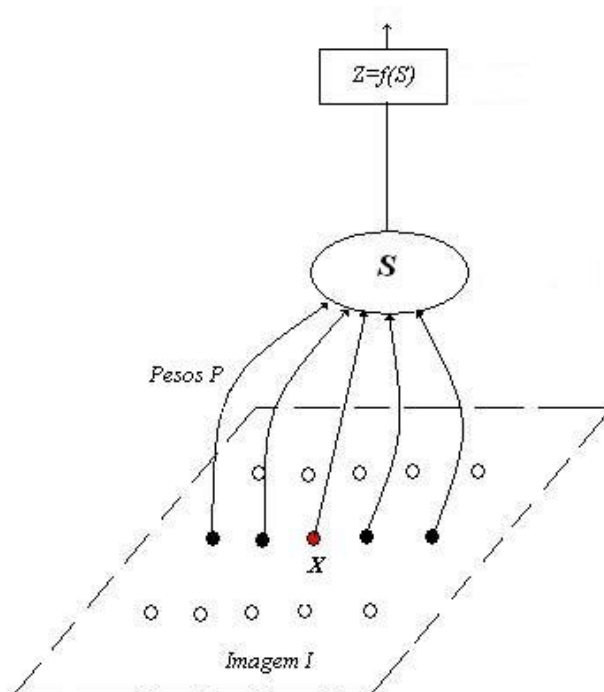


Figura 5.3. Para estimação da Profundidade Z em um determinado ponto X da imagem de entrada, estamos associando a esse ponto um elemento da Rede, que calcula o nível de ativação S , com base na Transformada de Fourier das intensidades (com ponderação P) em uma vizinhança de X .

Uma outra possibilidade de extensão imediata do nosso trabalho é o refinamento da estimação de forma através do SFS baseado em *Defocus*. Como a condição $\sigma Z = \text{constante}$, utilizada na equação 4.9, é apenas aproximada, a normalização da função Gaussiana que modela o processo de desfocamento depende, na verdade, de Z . Uma proposta para o refinamento dos nossos mapas de profundidade seria portanto a de empregar as estimativas preliminares para Z , obtidas com a suposição de $\sigma Z = \text{constante}$, para alterar as estimativas subsequentes da medida de desfocamento σ , levando a um processo iterativo de estimação das profundidades.

Referências Bibliográficas

- [1] B.K.P. Horn, "Shape from shading: a method for obtaining the shape of a smooth opaque object from one view", PhD thesis, MIT, 1970.
- [2] R.J. Woodham, "Photometric method for determining surface orientation from multiple images", *Optical Engineering* 19 (1) (1980) 139-144.
- [3] Y. Shirai, "Three-dimensional computer vision", Springer-Verlag, 1987.
- [4] J.L.Fernandes, "Contribuições ao estudo do estéreo fotométrico baseado em disparidades", Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal Fluminense, Niterói, RJ, 1997.
- [5] J.R.A. Torreão, "A Green's function approach to shape from shading", *Pattern Recognition* 34 (2001) 2367-2382.
- [6] A. Pentland, "Photometric motion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9) (1991) 879-890.
- [7] J.R.A. Torreão, J.L.Fernandes and H.C.G. Leitão, "A novel photometric motion approach", *Proceedings of Sibgrapi 2003*, pp. 294-298, IEEE Computer Society, October 2003.
- [8] J.R.A. Torreão and J.L. Fernandes, "From photometric motion to shape from shading", *Proceedings of SIBGRAPI 2004*, pp. 186-191, IEEE Computer Society, October 2004.

- [9] A. P. Pentland, "Depth of scene from depth of field", In Proc. Image Understanding Workshop, pp. 253-259, 1982.
- [10] B.K.P. Horn, "Robot Vision", Cambridge, MA: MIT Press, 1986.
- [11] A. H. Watt and F. Policarpo, "The computer image", Addison-Wesley Pub Co (Net), 1998.
- [12] B. M. de Carvalho, "Estimação de forma tridimensional a partir do estéreo fotométrico baseado em disparidades", Dissertação de Mestrado apresentada à Pós-Graduação em Ciências da Computação do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, 1995.
- [13] A.P. Pentland, "Linear shape from shading", International Journal of Computer Vision 4 (1990) 153-162.
- [14] R. J. Woodham, "Photometric stereo: a reflectance map technique for determining surface orientation from a single view", Image Understanding Systems and Industrial Applications, 155: 136-143, August 1978.
- [15] S. T. Barnard, "Stochastic stereo matching over scale", International Journal of Computer Vision, 3: 17-32, 1989.
- [16] P. Fua, "A parallel stereo algorithm that produces dense depth maps and preserves image features", Rapports de Recherche 1369, INRIA, january 1991.
- [17] W. Eric and L. Grimson, "Computing stereopsis using feature point contour matching", In Azriel Rosenfeld, editor, Techniques for 3-D Machine Perception, pages 75-111, 1986.
- [18] S. T. Barnard, "Stereo matching by hierarchical, microcanonical annealing". In 10th International Joint Conference in Artificial Intelligence, pages 832-835, 1987.
- [19] R. Zhang, P.-S. Tsai, J.E. Cryer, M. Shah, "Shape from shading: A survey", IEEE Trans. PAMI {21(8)} (1999) 690-706.

- [20] K. Ikeuchi and B.K.P. Horn, "Numerical shape from shading and occluding boundaries", *Artificial Intelligence*, 17(1-3):141-184,1981.
- [21] M.J. Brooks, B.K.P. Horn, "Shape and source from shading", *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 932-936, 1985.
- [22] R.T. Frankot, R. Chellappa, "A method for enforcing integrability in shape from shading algorithms", *IEEE Trans. PAMI* 10 (1988) 439-451.
- [23] B.K.P. Horn, "Shape from shading", *International Journal of Computer Vision*, pages 37-75, 1989
- [24] R. Szeliski, "Fast shape from shading", *Computer Vision, Graphics, Image Processing: Image Understanding*, 53:129-153, 1991.
- [25] O.E.Vega and Y.H. Yang. "Shading logic: a heuristic approach to recover shape from shading". *IEEE Transactions on Pattern Abalysis and Machine Intelligence*, 15(6):592-597, 1993.
- [26] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):680-702, 1991.
- [27] Y. G. Leclerc and A.F. Bobick, "The direct computation of height from shading", *In IEEE Proceedings of Computer Vision and Pattern Recognition*, pages 552-558, 1991.
- [28] J. Malik and D. Maydan, "Recovering three dimensional shape from a single image of curved objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 1989.
- [29] K. M. Lee and C. C. J. Kuo, "Shape from shading with a linear triangular element surface model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):815-822, 1993.

- [30] P.L. Worthington, E.R. Hancock, "New constraints on data-closeness and needle map consistency for shape-from-shading", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(12) (1999) 1250-1267.
- [31] E. Rouy and A. Tourin, "A viscosity solutions approach to shape from shading", *SIAM Journal on Numerical Analysis*, 29(3):867-884, 1992.
- [32] P. Dupuis and J. Oliensis, "Direct method for reconstructing shape from shading", In *IEEE Proceedings of Computer Vision and Pattern Recognition*, pages 720-721, June 1993.
- [33] J. Oliensis and P. Dupuis, "A global algorithm for shape from shading", In *Proceedings International Conference on Computer Vision*, pp. 692-701, 1993.
- [34] M. Bichsel and A.P. Pentland, "A simple algorithm for shape from shading", In *IEEE Proceedings of Computer Vision and Pattern Recognition*, pages 459-465, 1992.
- [35] R. Kimmel and A. M. Bruckstein, "Shape from shading via level sets", *Israel Institute of Technology, CIS Report 9209*, 1992
- [36] A.P. Pentland, "Local shading analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (1984) 170-187.
- [37] C.H. Lee, A. Rosenfeld, "Improved methods of estimating shape from shading using the light source coordinate system", *Artificial Intelligence* 26 (1985) 125-143.
- [38] P.S. Tsai, M. Shah, "Shape from shading using linear approximation", *Image and Vision Computing* 12(8) (1994) 487-498.
- [39] J. R. A. Torreão and C. J. L. Pimentel, "Disparity-based photometric stereo", In *PANEL 92*, pages 1166-1181, September, 1992.

- [40] A.L. Fetter and J.D. Walecka, "Theoretical Mechanics of Particles and Continua". McGraw-Hill, New York, 1980.
- [41] H. Crane, "A theoretical analysis of the visual accommodation system in humans", NASA Ames Research Center, Final Rep. NAS 2-2760, 1966.
- [42] M. Born and E. Wolf, "Principles of optics", London: Pergamon, 1965.
- [43] J. Barron, D. Fleet and S. Beauchemin, "Performance of optical flow techniques", International Journal of Computer Vision 12(1) (1994) 43-79.
- [44] N. Gupta and L. Kanal, "Gradient based image motion estimation without computing gradients", International Journal of Computer Vision 22(1) (1997) 81-101.
- [45] G.P. Stein and A. Sashua, "Model-based brightness constraints: on direct estimation of structure and motion", IEEE Transactions on Pattern Analysis and Machine Intelligence 22(9) (2000) 992-1015.
- [46] B.K.P. Horn and B.G. Schunck, "Determining Optical Flow", Artif. Intell. 17 (1981) 185-203.
- [47] J.R.A. Torreão, M.S. Amaral, "Signal differentiation through a Green's function approach", Pattern Recognition Letters 23 1755-1759 (2002).
- [48] P.E. Ferreira Jr., J.R.A. Torreão and P.C.P. Carvalho, "Data-based motion simulation through a Green's function approach," Proc. XVII Brazilian Symposium on Computer Graphics and Image Processing, pp. 193-199, IEEE Computer Society, October 2004.
- [49] S. Marcelja, "Mathematical description of the responses of simple cortical cells," Journal of the Optical Society of America, vol. 70, pp. 1297-1300, 1980.
- [50] I. Ohzawa, G.C. DeAngelis, and R.D. Freeman, "Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors," Science, vol.

249, pp. 1037-1041, 1990.

- [51] A.P. Pentland, "Finding illuminant direction", *Optical Society of America* 4(1982) 448-455.
- [52] D. Samaras and D. Metaxas, "Incorporating illumination constraints in deformable models for shape from shading and light direction estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(2) (2003) 247-264.
- [53] J.R.A.Torreão, "Geometric-photometric approach to monocular shape estimation", *Image and Vision Computing* 21(2003) 1045-1061.
- [54] H.N. Nair, C.V.Stewart, "Robust focus ranging", *IEEE Computer Vision and Pattern Recognition*, pp 309-314, 1992.
- [55] D. Ziou, F. Deschenes, "Depth from defocus in spatial domain", *Comp. Vision and Image Understanding* 81 (2001) 143-165.
- [56] A.P. Pentland, "A new sense for depth of field", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(4) (1987) 523-531.
- [57] S.H.Lai, C.W. Fu and S. Chang, "A generalized depth estimation algorithm with a single image", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14 (1992) 405-411.
- [58] M. Subbarao and N. Gurumoorthy, "Depth recovery from blurred edges", In *Proceedings of IEEE, International Conference on Computer Vision and Pattern Recognition*, pp. 498-503, 1988.
- [59] M. Subbarao and G. Surya, "Depth from Defocus: a spatial domain approach", *Int. J. Comput. Vision* 13 (1994) 271-294.
- [60] G. Surya, "Three-dimensional scene recovery from image defocus". Ph. D. Thesis, Department of Electrical Engineering, SUNY at Stony Brook, 1994.

- [61] A.P. Pentland, T. Darrell, M. Turk and W. Huang, "A simple, real-time range camera", In Proceedings of IEEE, International Conference on Computer Vision and Pattern Recognition, pp. 256-261, 1989.
- [62] Y. Xiong and S.A. Shafer, "Depth from focusing and defocusing", In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 68-73, 1993.
- [63] J. Ens and P. Lawrence, "An investigation of methods for determining depth from focus", IEEE Trans. PAMI 15(2) (1993) 97-108.
- [64] M. Gokstrop, "Computing depth from out-of-focus blur using a local frequency representation", In Proceedings of IEEE, International Conference on Computer Vision and Pattern Recognition, pp. 153-158, 1994.
- [65] A. Horii, "Depth from defocusing", Tech. Report ISRN KTH/NA/P-92/16-SE, Royal Institute of Technology, Sweden, 1992.
- [66] A.N. Rajagopalan and S. Chaudhuri, "Space-variant approaches to recovery of depth from defocused images", Comput. Vision and Image Understanding, 68(1997), 309-329
- [67] P. Moreno, A. Bernardino and J.S. Victor, "Gabor Parameter Selection for Local Feature Detection", 2nd Iberian Conference on Pattern Recognition and Image Analysis, Estoril, Portugal, June 7-9, 2005.
- [68] J. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles", Vision Research 20 847-856 (1980).
- [69] R. Schalkoff, "Pattern recognition: statistical, structural and neural approaches", John Wiley & Sons, Inc., Singapore, 1992.
- [70] J.R.A. Torreão, J.L. Fernandes, "Single-image shape from defocus", Proceedings of Sibgrapi 2005, pp. 241-246, IEEE Computer Society, October 2005.

- [71] J.R.A. Torreão, J.L. Fernandes, "Matching photometric stereo images", *Journal of the Optical Society of America A* 15(12)(1998) 2966-2975.

Publicações Relacionadas

- J.R.A. Torreão, J.L.Fernandes and H.C.G. Leitão, “A novel photometric motion approach”, Proceedings of Sibgrapi 2003, pp. 294-298, IEEE Computer Society, October 2003.
- J.R.A. Torreão and J.L. Fernandes, “From photometric motion to shape from shading”, Proceedings of SIBGRAPI 2004, pp. 186-191, IEEE Computer Society, October 2004.
- J.R.A. Torreão and J.L. Fernandes, “Single-image shape from defocus”, Proceedings of Sibgrapi 2005, pp. 241-246, IEEE Computer Society, October 2005.
- J. R.A. Torreão, J. L. Fernandes and H.C.G. Leitão, “A novel approach to photometric motion”, Image and Vision Computing 25 (2007), pp. 126-135.

Aqui, vamos deduzir a função de Green da equação diferencial (3.24). Identificando $k \equiv a/\sigma^2$, vamos escrever essa equação como

$$\frac{1}{2}I_2'' + \left(\frac{x+a}{\sigma^2}\right)I_2' + \frac{1}{2}\left[k^2 + \frac{(x+a)^2 + \sigma^2}{\sigma^4}\right]I_2 = k^2 I_1 \quad (\text{I.1})$$

A função de Green, $G(x, x_0)$, deverá satisfazer

$$\frac{1}{2}G''(x, x_0) + \left(\frac{x+a}{\sigma^2}\right)G'(x, x_0) + \frac{1}{2}\left[k^2 + \frac{(x+a)^2 + \sigma^2}{\sigma^4}\right]G(x, x_0) = k^2 \delta(x - x_0) \quad (\text{I.2})$$

Considerando que a função impulso $\delta(x - x_0)$ se anula identicamente para $x > x_0$ ou $x < x_0$, $G(x, x_0)$ será a solução homogênea para (I.2) desde que $x \neq x_0$. Por outro lado, em $x = x_0$, ela deverá ser de tal forma que o lado esquerdo de (I.2) produza a descontinuidade associada à função δ . Integrando os dois lados da equação com relação a x , em torno de $x = x_0$, e assumindo que $G(x, x_0)$ é contínua, concluímos que deve existir uma descontinuidade na derivada da função em x_0 , ou seja,

$$G'(x_{0+}, x_0) - G'(x_{0-}, x_0) = 2k^2 \quad (\text{I.3})$$

onde $x_{0+} = x_0 + \varepsilon$, e $x_{0-} = x_0 - \varepsilon$, para um ε infinitesimal.

Verificamos então que uma função $G(x, x_0)$, satisfazendo tais condições e tendendo a zero no infinito, pode ser obtida como a parte imaginária do núcleo complexo:

$$K(x, x_0) = \begin{cases} N(x_0)W(x)W^*(x_0) & x > x_0 \\ 0 & x < x_0 \end{cases} \quad (\text{I.4})$$

onde $W(x)$ é a solução homogênea complexa para (I.2), tendendo a zero quando $x \rightarrow \infty$. $K(x, x_0)$ vai satisfazer, portanto, a equação (I.2) para $x > x_0$ ou $x < x_0$, faltando apenas determinar o fator $N(x_0)$, a partir da condição de descontinuidade de (I.3). É fácil verificar que este será dado por

$$N(x_0) = \frac{2k^2}{\text{Im}[W'(x_0)W^*(x_0)]} \quad (\text{I.5})$$

e vamos finalmente obter

$$G(x, x_0) = 2k^2 \frac{\text{Im}[W(x)W^*(x_0)]}{\text{Im}[W'(x_0)W^*(x_0)]} \quad (\text{I.6})$$

onde $W(x)$, a solução homogênea para (I.2), é a função de Gabor

$$W(x) = \exp[ikx] \exp\left[-\frac{(x+a)^2}{2\sigma^2}\right] \quad (\text{I.7})$$

Substituindo (I.7) em (I.6), vamos então obter

$$G(x, x_0) = 2k \sin[k(x - x_0)] \exp\left[-\frac{(x+a)^2 - (x_0+a)^2}{2\sigma^2}\right] \quad (\text{I.8})$$

que é parte imaginária do núcleo complexo em (3.21).

Aqui, vamos demonstrar que um sinal bidimensional qualquer, $I(x,y)$, pode ser representado como

$$I(x, y) = \iint_{-\infty}^{\infty} d\omega_x d\omega_y \left\{ \exp[i\omega_x(x-a) + i\omega_y(y-b)] \exp\left[-\frac{(x-a)^2 + (y-b)^2}{2\sigma^2}\right] * \right. \\ \left. * \exp[i(\omega_x x + \omega_y y)] \right\} \quad (\text{II.1})$$

onde,

$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|} \quad (\text{II.2})$$

e

$$\omega_x a(\omega_x, \omega_y) + \omega_y b(\omega_x, \omega_y) = -\varphi_{\tilde{I}}(\omega_x, \omega_y) \quad (\text{II.3})$$

Nas expressões acima, $|\tilde{I}(\omega_x, \omega_y)|$ representa o módulo da Transformada de Fourier do sinal $I(x,y)$, e $\varphi_{\tilde{I}}(\omega_x, \omega_y)$, o ângulo de fase correspondente.

Para justificar as relações em (II.2) e (II.3), vamos considerar a equação (II.1) no domínio da frequência: mudando as variáveis de integração para ω_x^0 e ω_y^0 , e tomando a Transformada de Fourier dos dois lados da equação, obtemos

$$\tilde{I}(\omega_x, \omega_y) = \iint_{-\infty}^{\infty} d\omega_x^0 d\omega_y^0 \left\{ \exp[-i(\omega_x^0 a + \omega_y^0 b)] (2\pi)\sigma^2 \times \right. \\ \times \exp\left[-\sigma^2 \frac{(\omega_x - \omega_x^0)^2 + (\omega_y - \omega_y^0)^2}{2}\right] \exp[-i[(\omega_x - \omega_x^0)a + (\omega_y - \omega_y^0)b]] \times \\ \left. \times (2\pi)^2 \delta(\omega_x - \omega_x^0, \omega_y - \omega_y^0) \right\} \quad (\text{II.4})$$

ou,

$$\begin{aligned} \tilde{I}(\omega_x, \omega_y) = (2\pi)^3 \iint_{-\infty}^{\infty} d\omega_x^0 d\omega_y^0 \left\{ \sigma^2 \exp\left[-\sigma^2 \frac{(\omega_x - \omega_x^0)^2 + (\omega_y - \omega_y^0)^2}{2}\right] \times \right. \\ \left. \times \delta(\omega_x - \omega_x^0, \omega_y - \omega_y^0) \exp[-i(\omega_x a + \omega_y b)] \right\} \end{aligned} \quad (\text{II.5})$$

onde δ representa a função delta de Dirac. Efetuando a integração sobre ω_x^0 e ω_y^0 , obtemos então, devido à propriedade de filtragem da δ ,

$$\tilde{I}(\omega_x, \omega_y) = (2\pi)^3 \sigma^2 \exp[-i(\omega_x a + \omega_y b)] \quad (\text{II.6})$$

e, explicitando σ ,

$$\sigma^2(\omega_x, \omega_y) = \frac{1}{(2\pi)^3} \tilde{I}(\omega_x, \omega_y) \exp[-i(\omega_x a + \omega_y b)] \quad (\text{II.7})$$

Escolhendo adequadamente o parâmetros a e b de forma que $\omega_x a + \omega_y b = -\varphi_I(\omega_x, \omega_y)$, vamos obter

$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|} \quad (\text{II.8})$$

o que demonstra as relações em (II.2) e (II.3).

Podemos também verificar a consistência matemática da equação (II.1), substituindo aí as expressões para σ , a e b . Explicitando a convolução em (II.1), vamos começar por reescrever esta equação como

$$\begin{aligned} I(x, y) = \iint_{-\infty}^{\infty} d\omega_x d\omega_y \iint_{-\infty}^{\infty} d\xi d\eta \left\{ \exp[i(\omega_x x + \omega_y y)] \exp[-i(\omega_x a + \omega_y b)] \times \right. \\ \left. \times \exp\left[-\frac{(\xi - a)^2 + (\eta - b)^2}{2\sigma^2}\right] \right\} \end{aligned} \quad (\text{II.9})$$

ou ainda,

$$\begin{aligned}
I(x, y) = & \iint_{-\infty}^{\infty} d\omega_x d\omega_y \{ \exp[i(\omega_x x + \omega_y y)] \exp[-i(\omega_x a + \omega_y b)] \times \\
& \times \iint_{-\infty}^{\infty} d\xi d\eta \exp[-\frac{(\xi - a)^2 + (\eta - b)^2}{2\sigma^2}] \} \quad (\text{II.10})
\end{aligned}$$

e, resolvendo a integral interna em (II.10),

$$I(x, y) = \iint_{-\infty}^{\infty} d\omega_x d\omega_y \{ \exp[i(\omega_x x + \omega_y y)] \exp[-i(\omega_x a + \omega_y b)] 2\pi\sigma^2 \} \quad (\text{II.11})$$

Substituindo o valor de σ deduzido em (II.8), teremos

$$I(x, y) = \iint_{-\infty}^{\infty} d\omega_x d\omega_y \{ \exp[i(\omega_x x + \omega_y y)] \frac{1}{(2\pi)^2} \exp[-i(\omega_x a + \omega_y b)] |\tilde{I}(\omega_x, \omega_y)| \} \quad (\text{II.12})$$

como tomamos $\omega_x^0 a + \omega_y^0 b = -\varphi_T(\omega_x^0, \omega_y^0)$, o lado direito de (II.12) torna-se

$$\frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} d\omega_x d\omega_y \exp[i(\omega_x x + \omega_y y)] \tilde{I}(\omega_x, \omega_y) \quad (\text{II.13})$$

que é exatamente a definição de $I(x, y)$ por Transformada Inversa de Fourier, comprovando a consistência matemática da representação proposta.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)