

Roger Correia Pinheiro Silva

**Desenvolvimento e Análise de um Digitalizador Câmera-Projetor de Alta
Definição para Captura de Geometria e Fotometria**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional, da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre em Modelagem Computacional.

Orientador: Prof. D.Sc. Marcelo Bernardes Vieira

Coorientador: Prof. D.Sc. Sócrates de Oliveira Dantas

Juiz de Fora

2011

Silva, Roger Correia Pinheiro

Desenvolvimento e Análise de um Digitalizador Câmera-Projetor de Alta Definição para Captura de Geometria e Fotometria/Roger Correia Pinheiro Silva. – Juiz de Fora: UFJF/MMC, 2011.

XIV, 73 p.: il.; 29,7cm.

Orientador: Marcelo Bernardes Vieira

Coorientador: Sócrates de Oliveira Dantas

Dissertação (mestrado) – UFJF/MMC/Programa de Modelagem Computacional, 2011.

Referências Bibliográficas: p. 71 – 73.

1. Calibração Câmera Projetor. 2. Luz Estruturada.
3. Fotografia 3D. 4. Geometria Computacional. I. Vieira, Marcelo Bernardes *et al.* II. Universidade Federal de Juiz de Fora, MMC, Programa de Modelagem Computacional.

Roger Correia Pinheiro Silva

**Desenvolvimento e Análise de um Digitalizador Câmera-Projetor de Alta
Definição para Captura de Geometria e Fotometria**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional, da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre em Modelagem Computacional.

Aprovada em 26 de Agosto de 2011.

BANCA EXAMINADORA

Prof. D.Sc. Marcelo Bernardes Vieira - Orientador
Universidade Federal de Juiz de Fora

Prof. D.Sc. Sócrates de Oliveira Dantas
Universidade Federal de Juiz de Fora

Prof. D.Sc. Rodrigo Luis de Souza da Silva
Universidade Federal de Juiz de Fora

Prof. D.Sc. Anselmo Antunes Montenegro
Universidade Federal Fluminense

Prof. D.Sc. Saul de Castro Leite
Universidade Federal de Juiz de Fora

AGRADECIMENTOS

Aos meus pais, pelo apoio incondicional.

Aos amigos do GCG, que me ajudaram diretamente neste trabalho: Zé Luis, Rafael Alves, e Eduardo.

Aos amigos de sempre, Edão, Peçanha, McLeary, Cadim, Gustavo, Xandão, pela amizade.

Aos meu orientador, Marcelo Bernardes, pela paciência.

Ao Mestrado em Modelagem Computacional e ao Departamento de Ciência da Computação.

À CAPES pela bolsa de mestrado.

RESUMO

Um sistema câmera-projetor é capaz de capturar informação geométrica tridimensional de objetos e ambientes do mundo real. A captura de geometria em tal sistema baseia-se na projeção de luz estruturada sobre um objeto através do projetor, e na captura da cena modulada através da câmera. Com o sistema previamente calibrado, a deformação da luz projetada causada pelo objeto fornece a informação necessária para reconstruir a geometria do mesmo por meio de triangulação.

Este trabalho descreve o desenvolvimento de um digitalizador câmera-projetor de alta definição (com resoluções de até 1920x1080 e 1280x720); são detalhadas as etapas e processos que conduzem à reconstrução de geometria, como calibração câmera-projetor, calibração de cores, processamento da imagem capturada e triangulação. O digitalizador desenvolvido utiliza a codificação de luz estruturada (b, s) -BCSL, que emprega a projeção de uma sequência de faixas verticais coloridas sobre a cena. Este esquema de codificação flexível oferece um número variado de faixas para projeção: quanto maior o número de faixas, mais detalhada a geometria capturada.

Um dos objetivos deste trabalho é estimar o número limite de faixas (b,s) -BCSL possível dentro das resoluções atuais de vídeo de alta definição. Este número limite é aquele que provê reconstrução densa da geometria alvo, e ao mesmo tempo possui baixo nível de erro. Para avaliar a geometria reconstruída pelo digitalizador para os diversos números de faixas, é proposto um protocolo para avaliação de erro. O protocolo desenvolvido utiliza planos como objetos para mensurar a qualidade de reconstrução geométrica. A partir da nuvem de pontos gerada pelo digitalizador, a equação do plano para a mesma é estimada por meio de mínimos quadrados. Para um número fixo de faixas, são feitas cinco digitalizações independentes do plano: cada digitalização leva a uma equação; também é computado o plano médio, estimado a partir da união das cinco nuvens de pontos. Uma métrica de distância no espaço projetivo é usada para avaliar a precisão e a acurácia de cada número de faixas projetados.

Além da avaliação quantitativa, a geometria de vários objetos é apresentada para uma avaliação qualitativa. Os resultados demonstram que a quantidade de faixas limite para vídeos de alta resolução permite uma grande densidade de pontos mesmo em superfícies

com alta variação de cores.

Palavras-chave: Calibração Câmera Projetor. Luz Estruturada. Fotografia 3D.
Geometria Computacional.

ABSTRACT

A camera-projector system is capable of capturing three-dimensional geometric information of objects and real-world environments. The capture of geometry in such system is based on the projection of structured light over an object by the projector, and the capture of the modulated scene through the camera. With a calibrated system, the deformation of the projected light caused by the object provides the information needed to reconstruct its geometry through triangulation.

The present work describes the development of a high definition camera-projector system (with resolutions up to 1920x1080 and 1280x720). The steps and processes that lead to the reconstruction of geometry, such as camera-projector calibration, color calibration, image processing and triangulation, are detailed. The developed scanner uses the (b, s) -BCSL structured light coding, which employs the projection of a sequence of colored vertical stripes on the scene. This coding scheme offers a flexible number of stripes for projection: the higher the number of stripes, more detailed is the captured geometry.

One of the objectives of this work is to estimate the limit number of (b, s) -BCSL stripes possible within the current resolutions of high definition video. This limit number is the one that provides dense geometry reconstruction, and at the same has low error. To evaluate the geometry reconstructed by the scanner for a different number of stripes, we propose a protocol for error measurement. The developed protocol uses planes as objects to measure the quality of geometric reconstruction. From the point cloud generated by the scanner, the equation for the same plane is estimated by least squares. For a fixed number of stripes, five independent scans are made for the plane: each scan leads to one equation; the median plane, estimated from the union of the five clouds of points, is also computed. A distance metric in the projective space is used to evaluate the precision and the accuracy of each number of projected stripes.

In addition to the quantitative evaluation, the geometry of many objects are presented for qualitative evaluation. The results show that the limit number of stripes for high resolution video allows high density of points even on surfaces with high color variation.

Keywords: Camera Projector Calibration. Structured Light. 3D Photography.

Computational Geometry.

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Definição do Problema	15
1.2	Objetivos	16
1.3	Organização do Trabalho	16
2	TRABALHOS RELACIONADOS.....	18
2.1	Sistemas de Reconstrução por Luz Ativa	19
2.2	Codificação de Luz Estruturada	20
2.3	Digitalizadores de Luz Estruturada de Alta Resolução	22
2.4	Avaliação da Geometria Reconstruída	23
3	FUNDAMENTOS	25
3.1	Calibração Câmera-Projetor	25
3.1.1	<i>Método de Tsai</i>	29
3.2	Calibração de Cor	33
3.3	Codificação (b, s) -BCSL	34
4	DESENVOLVIMENTO DO DIGITALIZADOR TRIDIMENSIONAL.	38
4.1	Configuração do Sistema	38
4.2	Extração de Geometria e Fotometria	41
4.2.1	<i>Detecção de Fronteiras</i>	43
4.2.2	<i>Detecção de Fronteiras no Tempo</i>	46
4.3	Protocolo para Avaliação de Erro	47
5	RESULTADOS E ANÁLISE QUANTITATIVA.....	51
5.1	Resultados	51
5.2	Análise Quantitativa	55
5.2.1	<i>Experimento com Barreira de Vidro</i>	64
6	CONCLUSÃO	68
	REFERÊNCIAS	71

LISTA DE ILUSTRAÇÕES

3.1	Sistemas de coordenadas no processo de calibração [1].	26
3.2	Padrão de calibração plano xadrez com pontos característicos marcados.	27
3.3	Triangulação [2]. Projetor à esquerda e câmera à direita.	29
3.4	<i>Slides</i> projetados sobre o cachorro utilizando (6,2)-BCSL. Nas imagens, (a) e (c) são <i>slides</i> utilizados para codificação; (b) é o complementar de (a) e (d) é o complementar de (c).	35
3.5	Codificação (3,2)-BCSL [3].	36
4.1	Posicionamento de equipamentos e objetos.	39
4.2	Imagens capturadas para calibração de cores. Cada cor é projetada sobre um anteparo branco, como uma folha de papel.	40
4.3	Calibração câmera-projetor. (a) imagem capturada para calibração da câmera. (b) imagem capturada para calibração do projetor. (c) padrão projetado para calibração do projetor.	40
4.4	Sequência de passos na <i>pipeline</i> de reconstrução geométrica.	42
4.5	Localização de fronteiras. (a) imagem I_1 . (b) imagem \bar{I}_1 . (c) imagem de cores projetadas P_1 . (d) variação de intensidade dos canais R, G e B da imagem de diferença D_1 , para linha marcada em branco. As linhas verticais do gráfico indicam as fronteiras encontradas entre faixas.	44
4.6	Decodificação de fronteiras (6,2)-BCSL.	47
4.7	Ajuste de planos. (a) projeção de <i>slides</i> sobre o plano. (b) nuvem de pontos gerada pelo digitalizador. (c) plano estimado para a nuvem de pontos, minimizando erro quadrático.	48
4.8	Distância entre dois pontos do RP^2	49
5.1	Digitalização do boneco de cachorro. (a) textura. (b) nuvem de 33.298 pontos gerada pelo digitalizador. (c) composição da nuvem de pontos com a textura. (d) normais inferidas através de pontos vizinhos.	52

5.2	Caixa e garrafa. (a) cena capturada. (b) nuvem de 80.584 pontos gerada pelo digitalizador. (c) normais inferidas para caixa. (d) normais inferidas para garrafa.	52
5.3	Digitalização de livros coloridos. (a) textura. (b) nuvem de pontos gerada pelo digitalizador. (c) composição da nuvem de pontos com a textura da cena. .	53
5.4	Digitalização de um rosto. (a) cena capturada. (b) nuvem de 16.860 pontos gerada para projeção de 100 faixas. (c) nuvem de 29.128 pontos gerada para projeção de 200 faixas. (d) nuvem de 56.334 pontos gerada para projeção de 250 faixas.	54
5.5	Rosto reconstruído pelo digitalizador através da projeção de 250 faixas. . . .	54
5.6	Gráficos de erro para acurácia e precisão; configuração de câmera 1280x720x60. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.	56
5.7	Plano base e planos médios obtidos para configuração de câmera 1280x720x60. Plano base em cinza, planos médios em azul. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas.	56
5.8	Comparação entre planos médios e reconstruções independentes do plano, para configuração de câmera 1280x720x60. Planos médios em cinza; planos azuis, vermelhos e verdes representam reconstruções independentes do plano. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas.	57
5.9	Nuvens de pontos geradas para digitalização do plano, para configuração de captura 1280x720x60. (a) 50 faixas. Nuvem de 13.694 pontos. (b) 100 faixas. Nuvem de 27.162 pontos. (c) 150 faixas. Nuvem de 40.750 pontos. (d) 200 faixas. Nuvem de 54.319 pontos. (e) 250 faixas. Nuvem de 67.905 pontos. (f) 300 faixas. Nuvem de 78.668 pontos.	58
5.10	Gráficos de erro de acurácia e precisão para configuração de câmera 1920x1080x30. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.	59

5.11	Plano base e planos médios para configuração de câmera 1920x1080x30. Em cinza, plano base; em azul, planos médios. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas. (g) 350 faixas. (h) 400 faixas.	59
5.12	Comparação entre planos médios e reconstruções independentes do plano, para configuração de câmera 1920x1080x30. Em cinza, planos médios; em azul, verde e vermelho reconstruções independentes do plano. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas. (g) 350 faixas. (h) 400 faixas.	60
5.13	Nuvens de pontos geradas para digitalização do plano, para resolução de captura 1920x1080x30. (a) 50 faixas. Nuvem de 20.383 pontos. (b) 100 faixas. Nuvem de 40.635 pontos. (c) 150 faixas. Nuvem de 60.963 pontos. (d) 200 faixas. Nuvem de 81.275 pontos. (e) 250 faixas. Nuvem de 101.587 pontos. (f) 300 faixas. Nuvem de 121.871 pontos. (g) 350 faixas. Nuvem de 142.225 pontos. (h) 400 faixas. Nuvem de 156.776 pontos.	61
5.14	Posicionamento e orientação dos planos para digitalização. (a) sequência 1. (b) sequência 2.	62
5.15	Comparação do erro entre primeira e segunda sequência de medidas, para configuração de câmera 1280x720x60. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.	62
5.16	Comparação do erro entre primeira e segunda sequência de medidas, para configuração de câmera 1920x1080x30. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.	63
5.17	Passagens por zero marcadas, em branco, sobre a imagem I_1 do plano base. (a) barreira de vidro sem rotação. (b) barreira de vidro rotacionada 15 graus no sentido da câmera. (c) barreira de vidro rotacionada 30 graus no sentido do projetor.	66
5.18	Plano base e planos médios para medida de acurácia. Em cinza, plano base; em azul, planos médios. (a) rotação de 0 grau da barreira de vidro. (b) rotação de 15 graus da barreira de vidro. (c) rotação de 30 graus da barreira de vidro.	66

LISTA DE TABELAS

3.1	Tabela de decodificação para (3,2)-BCSL [3]	36
5.1	Acurácia e precisão para digitalização do plano através da barreira de vidro. .	65

1 INTRODUÇÃO

Um digitalizador 3D é um dispositivo capaz de capturar informação geométrica tridimensional de objetos e ambientes do mundo real. Ele tem aplicação em áreas diversas: na indústria, com medida de dimensões, controle de qualidade e engenharia reversa; na medicina, com *design* de próteses; no entretenimento, com a geração de modelos tridimensionais para jogos e filmes.

Um digitalizador câmera-projetor captura informação geométrica tridimensional de um objeto através da projeção de um padrão de luz conhecido (luz estruturada) sobre o mesmo. Com o sistema previamente calibrado, a deformação do padrão projetado causada pelo objeto fornece a informação necessária para reconstruir sua geometria.

Os padrões atuais de vídeo de alta definição tornaram sistemas de digitalização câmera-projetor mais atraentes: além do baixo custo, o aumento da resolução das imagens possibilita capturas e reconstruções geométricas mais detalhadas, com maior densidade de pontos.

Esta dissertação descreve o desenvolvimento de um digitalizador câmera-projetor de alta definição.

Neste trabalho também é feita uma análise de erro das geometrias reconstruídas pelo digitalizador. Para isto, um protocolo de erro é proposto.

1.1 Definição do Problema

O problema desta dissertação é estimar, para cada resolução, o número de faixas do código de luz estruturada (b, s) -BCSL que provê reconstrução geométrica de qualidade. Uma reconstrução de qualidade é aquela que consegue recuperar detalhes geométricos de uma cena ou objeto com precisão de subpixel, tem grande densidade de número de pontos e ao mesmo tempo, o nível de erro para a reconstrução deve ser baixo.

Para realizar esta estimativa, é proposto um protocolo de erro nesta dissertação. Trata-se do estabelecimento de uma métrica de distância entre planos no espaço projetivo, usada para avaliar a acurácia e a precisão das reconstruções.

O outro problema desta dissertação é examinar a possibilidade de uso de vídeo digital

de alta definição para reconstrução em tempo real.

1.2 Objetivos

O objetivo primário deste trabalho é obter métodos que permitam uma reconstrução precisa da geometria de um objeto ou cena, utilizando câmera-projetor de alta resolução.

Como objetivo secundário, deve-se avaliar o comportamento dos métodos de reconstrução obtidos para diversos números de faixas (b, s) -BCSL projetados, e também para diferentes resoluções de vídeo.

1.3 Organização do Trabalho

Este trabalho é dividido em 6 capítulos. O capítulo 2 apresenta trabalhos relacionados ao método de reconstrução por luz ativa. Este capítulo está dividido em três partes: sistemas de reconstrução por luz ativa, codificação de luz estruturada e avaliação da geometria reconstruída.

O capítulo 3 aborda os fundamentos que permitem a um sistema de estereoscopia ativa a extração de geometria de uma cena por triangulação. Os pontos-chave deste capítulo são os fundamentos matemáticos do processo de calibração câmera-projetor e o padrão de luz estruturada (b, s) -BCSL. Também é abordada a questão de calibração de cores, um passo necessário para uso do padrão (b, s) -BCSL.

O capítulo 4 está dividido em três seções. Na primeira são apresentadas as etapas necessárias para configuração do digitalizador. Trata-se de procedimentos que antecedem a captura de geometria da cena, como o posicionamento de equipamentos, controle da iluminação ambiente e captação de imagens para calibração do sistema de digitalização. A segunda seção detalha o processamento das imagens capturadas para reconstrução da geometria de uma cena; a fase final deste processamento é obtenção dos pontos tridimensionais da mesma. Na terceira e última seção é apresentado o protocolo de erro desenvolvido nesta dissertação para avaliação da geometria reconstruída pelo digitalizador. O protocolo é baseado no estabelecimento de uma métrica de distância entre planos no espaço projetivo, e avalia erros de acurácia e de precisão das reconstruções.

O capítulo 5 apresenta os resultados qualitativos e quantitativos deste trabalho. A geometria recuperada para vários objetos, de diferentes tipos de materiais, é apresentada

para uma avaliação qualitativa. A análise do erro de reconstrução para duas resoluções de imagem de captura, 1280x720 e 1920x1080, feita através do protocolo de erro proposto, também é apresentada.

Por fim, o capítulo 6 apresenta conclusões para os resultados obtidos e sugere como trabalhos futuros melhorias e extensões de funcionalidade para o digitalizador desenvolvido.

2 TRABALHOS RELACIONADOS

Sistemas de reconstrução geométrica utilizam métodos que podem ser categorizados em dois tipos básicos.

Métodos passivos não manipulam a luz que incide sobre uma cena. Entre estes, destaca-se a estereoscopia, que busca emular a forma como o sistema visual humano funciona. Em um sistema de visão estéreo, com duas ou mais câmeras em pontos de vista distintos, capturam-se imagens de uma cena; conhecida a transformação projetiva de cada câmera (isto é, aquela que mapeia pontos tridimensionais da cena para pontos bidimensionais do plano de projeção) é possível traçar uma reta no espaço que tem origem no centro ótico, passa por um pixel e atinge um ponto da cena. A interseção de pelo menos duas retas (uma para cada câmera distinta) é então utilizada para recuperar a profundidade de cada ponto. O problema maior em utilizar estereoscopia para reconstrução reside na identificação de correspondências entre os diversos pontos de vista; objetos com baixa variação de brilho ou com padrões periódicos causam grande dificuldade neste contexto. O desenvolvimento de algoritmos que reconheçam correspondências permanece como uma tarefa desafiadora [4].

Métodos ativos manipulam a luz incidente sobre a cena como forma de contornar o problema de correspondência. Sistemas que empregam estes métodos comumente substituem uma câmera por uma fonte de luz controlável (laser, projetor, etc.). O objetivo desta manobra é impor textura sobre a cena através da luz projetada para facilitar a localização de correspondências entre pontos de vista diferentes. A seção seguinte explora sistemas que usam este tipo de técnica.

É necessário esclarecer que em um sistema de reconstrução de geometria, seja ele passivo ou ativo, a etapa de calibração de dispositivos (câmeras, par câmera-laser, câmera-projetor e variações desses arranjos) é a primeira a ser realizada. Ela tradicionalmente envolve o uso de um objeto de dimensões e estrutura conhecida, que é capturado na imagem da câmera. O objetivo da calibração é estimar a transformação que relaciona pontos tridimensionais do objeto com pontos bidimensionais de sua imagem. Métodos de calibração comumente utilizados são constituídos por variações dos métodos de Tsai [5] e Zhang [6].

2.1 Sistemas de Reconstrução por Luz Ativa

Durante a década de 1970 surgiu o digitalizador baseado em laser. Neste esquema, uma série de espelhos fixos e rotativos é usada para direcionar um ponto único de laser que varre uma superfície. Uma câmera registra o movimento do ponto laser. A projeção 2D do ponto define uma reta entre o mesmo e o centro de projeção da câmera. O conhecimento prévio da angulação dos espelhos e da distância entre câmera e laser permite que a profundidade de cada ponto da superfície seja recuperada pela interseção desta reta com outra que tem origem no laser e termina no ponto da superfície atingido pelo mesmo. O problema com essa abordagem é a lentidão em deslocar um único ponto por toda a superfície para que a geometria possa ser obtida. Para lidar com essa dificuldade, digitalizadores que projetam uma faixa de laser foram desenvolvidos. Esta faixa, que tem origem no projetor de laser e atinge um objeto como uma linha, define um plano no espaço tridimensional. De maneira similar ao esquema anterior, a geometria de interesse é conseguida através das interseções entre o plano e o conjunto de retas que atingem o mesmo e se originam no centro de projeção da câmera. Este tipo de digitalização oferece medidas acuradas e uma geometria detalhada [7] [8]. É importante notar que a obtenção da geometria ainda implica na varredura completa da superfície pela faixa laser. Apesar disso, uma variedade de produtos comerciais utilizam essa técnica.

O trabalho de Bouguet e Perona [9] apresenta uma abordagem interessante para digitalização por luz ativa. Os autores descrevem o método de 'luz estruturada fraca', que usa a sombra projetada por um bastão para extração da geometria de um objeto. O interessante desta abordagem é que a fonte de luz que gera a sombra do bastão pode ser uma simples luminária, ou até mesmo o sol. O ponto chave do método é o uso de dois planos perpendiculares: o objeto de digitalização deve estar sobre um plano horizontal e à frente de um plano vertical. À medida que a sombra do bastão é deslocada para varrer toda a geometria do objeto, parte da mesma deve incidir sobre os planos e ser visível na imagem da câmera. Isto é essencial porque as parcelas da sombra sobre os planos horizontal e vertical permitem estimar o plano gerado pela sombra do bastão que intercepta o objeto. A geometria de interesse é reconstituída pela interseção do plano de sombra com os raios ópticos da câmera. É necessário explicitar que a posição relativa do plano horizontal em relação à câmera deve ser conhecida. Além disso, esta deve estar calibrada.

A ideia anteriormente apresentada sobre digitalização via laser pode ser ampliada. Ao invés de projetar uma faixa ou um ponto e deslocá-los sobre a geometria de um objeto, é possível projetar diversos pontos ou faixas de forma simultânea; isto é, pode-se realizar uma digitalização em área. Entretanto, para identificar de forma única cada ponto ou faixa, torna-se necessário codificá-los. Sensores que empregam essa tática geralmente substituem o laser por um projetor, e projetam um padrão de luz com estrutura conhecida sobre a cena que possibilita a triangulação, ou seja, o cálculo de profundidade. Tais sensores, como o desenvolvido no curso desta dissertação, são conhecidos como sistemas de luz estruturada. É fácil perceber que estes sistemas sofrem grande influência da luminosidade ambiente por utilizarem frequências de luz visível. Consequentemente, seu uso é recomendado somente em ambientes de luminosidade controlada, isto é, fechados e relativamente escuros. Uma alternativa para essa limitação é fazer uso de padrões que estão fora do espectro visível. Esse é o caso do Microsoft Kinect, um detector de movimentos para jogos que utiliza um padrão infravermelho. A luz infravermelha torna o sistema menos intrusivo e funcional em condições de luminosidade comuns [10].

2.2 Codificação de Luz Estruturada

A reconstrução da geometria em sistemas câmera-projetor é conseguida através da projeção de um padrão de luz estruturada conhecido sobre o objeto de interesse; o objetivo desse procedimento é estabelecer pontos de correspondência entre os sistemas de coordenadas do projetor e da câmera. A projeção impõe textura sobre o objeto e atribui códigos para pontos do mesmo na imagem capturada pela câmera. A posterior decodificação de cada ponto permite recuperar sua posição em coordenadas do projetor. Com o sistema previamente calibrado, a transformação que relaciona os sistemas de coordenadas da câmera e do projetor é conhecida; isto significa que a posição dos mesmos pontos, dessa vez no sistema da câmera, pode ser encontrada. A geometria do objeto é então reconstituída por meio de triangulação.

O trabalho de Salvi *et al.*[11] faz uma extensa revisão das diversas estratégias usadas na codificação de luz estruturada, e propõe uma classificação das técnicas existentes. A **codificação espacial** agrupa técnicas onde o código de um ponto específico é extraído dos pontos adjacentes. Além disso, as estratégias desse grupo tendem a concentrar todo

esquema de codificação (que geralmente é baseado em variações de intensidade ou cor) em um único padrão. Pode-se perceber então que o passo de decodificação torna-se difícil se a superfície do objeto que se deseja reconstruir não for suave. Um exemplo dessa abordagem é o trabalho desenvolvido por Kawasaki [12], que estabelece um padrão de linhas verticais e horizontais em formato de grade. A unicidade de uma região específica é codificada no espaço entre linhas horizontais azuis, ao passo que as linhas verticais vermelhas são igualmente espaçadas. A detecção dos pontos de cruzamento das linhas e posterior comparação de distâncias para pontos vizinhos determinam a posição dos primeiros no padrão projetado.

A **codificação temporal** é baseada na projeção sucessiva de padrões sobre a superfície do objeto. O código para cada pixel é formado pela sequência de valores de luminância ou crominância à ele associado em cada um dos padrões projetados. Usualmente, este padrão é uma série de faixas, e apenas valores de luminância (0 ou 1) são utilizados para cada faixa. A acurácia da codificação é diretamente relacionada ao número de padrões projetados, uma vez que cada padrão da sequência possui maior número de faixas (isto é, maior resolução). Já que o código de um pixel não está completo até que toda sequência de padrões seja projetada, pode-se notar que a limitação desse tipo de codificação reside em aplicações de tempo real. Apesar disso, Rusinkiewicz *et al.*[13] desenvolveram um sistema de captura de modelos tridimensionais a 60Hz. Isto foi possível porque uma codificação híbrida, temporal-espacial, foi utilizada. Uma sequência de projeção de quatro padrões foi empregada; contudo, a codificação nesse esquema não reside nas faixas, mas nas transições entre elas. Além disso, é necessário rastrear cada transição nos quadros do vídeo de captura. Seguindo a mesma linha de codificação híbrida, Ishii [14] desenvolveu um sistema de codificação flexível que se adapta às características da imagem capturada. Seu esquema de codificação se comporta como codificação temporal em regiões da imagem onde não há movimento, e como codificação espacial onde há grande variação de brilho (isto é, movimento).

A **codificação contínua baseada em deslocamento de fase** projeta um padrão cinza, cuja intensidade é modulada por uma função senoidal, sobre a superfície. Dessa maneira, cada ponto ao longo de uma linha paralela ao eixo de codificação pode ser caracterizado por um único valor de fase. Superfícies não planas causam deformação no padrão; essa deformação é registrada como desvio de fase, e fornece informação sobre

a geometria. O padrão deve ser deslocado e projetado diversas vezes (mínimo de três, de acordo com Salvi[11]) para que a criação de um mapa de fase relativo seja possível. Este tipo de codificação é capaz de gerar geometria densa, mas devido à necessidade da projeção de padrões deslocados no tempo não é adaptado para sistemas de tempo real. Trabalhos nesta área procuraram diminuir a quantidade de padrões projetados usando múltiplas frequências[15], ou utilizar o mínimo de padrões possível [16].

O presente trabalho utiliza a codificação (b, s) -BCSL (em inglês, Boundary Coded Structured Light) descrita por Sá *et al.* [3]. Trata-se de uma codificação temporal-espacial híbrida, baseada na projeção de padrões de faixas coloridas complementares. De forma semelhante a proposta por Rusinkiewicz [13], a codificação ocorre na transição entre faixas; cada transição é unicamente identificada pela cor à sua esquerda e à sua direita. O método oferece vantagens em relação às abordagens anteriores: grande resolução (pode-se, por exemplo, obter 900 transições distintas), consegue extrair a textura da cena durante a aquisição de geometria (isto ocorre pela projeção de padrões de cor complementares), e possibilidade de uso em aplicações de tempo real.

2.3 Digitalizadores de Luz Estruturada de Alta Resolução

Entre os digitalizadores de luz estruturada que procuram realizar uma captura geométrica de alta resolução, isto é, de alta densidade de pontos, o trabalho de Fechteler *et al.* [17] utiliza uma câmera fotográfica de 8.2 *megapixel* (resolução de aproximadamente 3520x2344). Neste sistema, um único padrão de faixas coloridas horizontais é projetado sobre a cena para extração de geometria; uma segunda imagem, iluminada apenas com luz branca, é utilizada para captura de textura. Neste trabalho, o resultado da digitalização de um rosto mostra uma reconstrução geométrica detalhada, com grande quantidade de pontos. Todavia, os autores não mencionam o número de pontos gerados, nem a possibilidade do uso do sistema para reconstrução em tempo real.

Guhring, em 2001, propôs um tipo de **codificação temporal** cujo foco é a obtenção de nuvens de pontos densas [18]. Tipicamente, a codificação temporal é baseada da projeção sucessiva de padrões de faixas verticais brancas e pretas sobre a cena, onde cada padrão tem um número crescente de faixas. A alteração proposta por Guhring é adicionar uma

segunda etapa de projeção de padrões; trata-se da projeção de um padrão periódico de faixas, sucessivas vezes, onde o mesmo é deslocado para direita a cada projeção. Esta segunda etapa possibilita uma codificação com acurácia de *subpixel*, além de atribuir um código individual para cada *pixel* da imagem da câmera.

A codificação proposta por Guhring é capaz de gerar nuvens de pontos densas (isto ocorre pela atribuição de um código para cada *pixel* da imagem capturada pela câmera), e possui acurácia de *subpixel*; por isto, ela é atraente para implementação de digitalizadores comerciais. Este é o caso do digitalizador Breuckmann triTOS-HE, que utiliza uma câmera de 1.4 *megapixel* (resolução aproximada 1384x1036), e é empregado para digitalização de objetos de arte [19] e documentação de escavações arqueológicas [20].

A limitação dos sistemas de digitalização que fazem uso da codificação de Guhring reside em aplicações de tempo real. Isto ocorre porque é necessária a projeção de um grande número de padrões de luz estruturada sobre a cena: em seu trabalho, Guhring relata a projeção de 10 padrões para extração de geometria.

Reconstrução tridimensional em alta resolução e tempo real é o objetivo do trabalho de Zhang [21]. Nele, um sistema de digitalização que utiliza **codificação contínua baseada em deslocamento de fase** foi desenvolvido. Neste sistema câmera e projetor estão sincronizados, e a taxa de captura é de 120 quadros por segundo. Como é necessária a projeção de três padrões em sequência para obtenção da geometria da cena, o sistema reconstrói formas tridimensionais a taxa de 40 quadros por segundo; a resolução de reconstrução é de 532x500 pontos por quadro de vídeo. Este sistema, entretanto, não captura textura colorida da cena, e para capturar vídeo a taxa de 120 quadros por segundo utiliza uma câmera de alta velocidade.

2.4 Avaliação da Geometria Reconstruída

Entre os trabalhos que se preocupam em avaliar a precisão e a acurácia da geometria reconstruída por sistemas câmera-projetor, Trobina [22] apresentou um modelo de erro para codificação temporal de luz estruturada. Foi demonstrado que o ponto principal para reconstrução com baixo erro está na localização acurada das transições entre faixas do padrão projetado na imagem da câmera. Duas formas para encontrar as transições com acurácia de subpixel foram propostas. A primeira é encontrar as passagens por

zero da segunda derivada da imagem capturada, ortogonalmente às faixas projetadas. A segunda é projetar os padrões normal e inverso (positivo e negativo); as transições são encontradas através dos pontos de interseção dos perfis de intensidade (cinza) dos dois padrões; interpolação linear é usada entre os pixels mais próximos de cada interseção para cálculo do ponto de passagem por zero. Resultados experimentais demonstraram que a segunda opção é a melhor. É importante notar que esse resultado está incorporado ao código (b, s) -BCSL. Neste caso, os padrões normal e inverso correspondem à projeção de cores complementares.

Em Feng *et al.*[23], o erro de reconstrução é analisado com foco nos métodos de calibração. Os autores projetaram uma faixa laser sobre um objeto de dimensões conhecidas, e calibraram a câmera com três métodos: Tsai [5], Zhang [6] e o método de transformação linear direta, que não considera distorção provocada pela lente da câmera [24]. A soma dos quadrados das diferenças entre pontos conhecidos e pontos medidos pelos três métodos mostrou que Tsai e Zhang retornaram valores bastante próximos, e foram mais acurados que o método de transformação linear direta.

O artigo de Salvi *et al.*[11] analisa a acurácia de reconstrução dos grupos de codificação de luz estruturada apresentados anteriormente nesta dissertação. Os autores implementaram um método representativo de cada grupo, e testaram a acurácia de reconstituição de um plano de cada um deles. A base para comparação de reconstrução é a equação do plano conseguida em cada teste, que foi obtida por análise de componentes principais (PCA), através das nuvens de pontos geradas. Foi concluído que os melhores resultados, isto é, aqueles de reconstrução densa e que apresentaram menor erro médio e desvio padrão, foram conseguidos por Guhring[18], representante da codificação temporal, e por Pribanic *et al.*[15], representante da codificação baseada em deslocamento de fase.

Nesta dissertação, a acurácia do sistema construído com (b, s) -BCSL é analisada. O objetivo é explorar a resolução flexível oferecida por esse código dentro das resoluções do padrão atual de vídeo digital. Pretende-se, dessa forma, estimar a resolução limite possível, aquela que provê reconstrução densa com baixo nível de erro. Para tanto, uma abordagem semelhante à de Salvi *et al.*[11] é utilizada. Um plano é capturado e reconstruído, e a qualidade da reconstrução é medida pela comparação com o plano ajustado por regressão linear a partir dos pontos 3D conseguidos pelo sistema.

3 FUNDAMENTOS

Este capítulo apresenta os fundamentos do sistema de reconstrução geométrica desenvolvido. São eles: calibração do sistema câmera-projetor, calibração de cor, e codificação de luz estruturada (b, s) -BCSL.

3.1 Calibração Câmera-Projetor

A etapa de calibração precede a captura de imagens que possibilita a aquisição da geometria de uma cena. É nesta etapa que deve-se estimar a posição e orientação dos dispositivos (câmera e projetor) em relação a um sistema de coordenadas global, e também obter os parâmetros da transformação projetiva perspectiva que mapeia pontos tridimensionais do espaço para pontos bidimensionais do plano de projeção.

O primeiro passo para obtenção da posição na imagem que corresponde a um ponto no espaço é a mudança de referencial tridimensional. Deve-se expressar um ponto $\vec{p}_w = (x_w, y_w, z_w)$, do sistema de coordenadas do mundo (SCM), no sistema de coordenadas da câmera (SCC). Este último tem origem no centro ótico da câmera; seu plano xy é paralelo ao plano de projeção e o eixo z corresponde ao eixo ótico. A mudança de referencial é descrita por uma matriz de rotação \mathbf{R} e um vetor de translação \vec{t} .

O passo seguinte é a aplicação de uma transformação projetiva que leva um ponto do SCC em sua projeção bidimensional no sistema de coordenadas da imagem (SCI), situado sobre o plano de projeção; a origem deste sistema 2D é a projeção ortogonal do eixo ótico. Este passo é dividido em duas partes. Inicialmente, a projeção perspectiva é realizada considerando a câmera segundo o modelo *pinhole*. Neste modelo, a projeção (x_u, y_u) é ideal, sem distorção causada pela lente. Em seguida é feita a correção através de um modelo de distorção radial. A projeção corrigida é denotada por (x_d, y_d) .

Por fim, é necessário realizar a transformação do SCI para o sistema de coordenadas em pixels (SCP). Este último sistema bidimensional trata das especificidades de cada equipamento, do registro de um raio luminoso em seus sensores, que geralmente estão dispostos como uma matriz retangular. As coordenadas do sistema são expressas em pixels, e a origem do mesmo é comumente o canto superior esquerdo da imagem formada. Um

ponto neste sistema tem coordenadas (x_f, y_f) .

A Equação 3.1 é a transformação final, em coordenadas homogêneas, que relaciona pontos tridimensionais do espaço e pontos bidimensionais da imagem da câmera. Ela é o resultado da composição de todas as transformações anteriormente descritas.

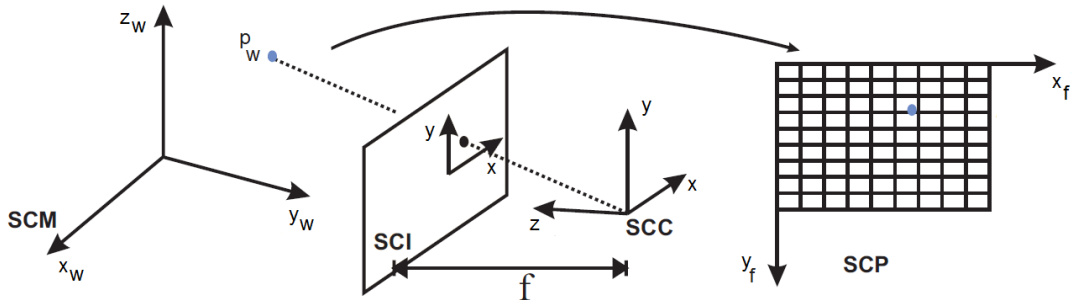


Figura 3.1: Sistemas de coordenadas no processo de calibração [1].

$$\begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} \cong \begin{bmatrix} f s_x & \tau & c_x \\ 0 & f s_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \vec{t} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3.1)$$

Na Equação 3.1, a matriz \mathbf{K} , composta pelos parâmetros f , s_x , s_y , τ , c_x e c_y , é chamada de matriz intrínseca. Ela retrata características inerentes da câmera:

- f é a distância focal.
- s_x e s_y são fatores de escala para a direção horizontal e vertical, respectivamente. Eles descrevem o número de pixels por unidade de comprimento.
- τ é a tangente do ângulo que as colunas da imagem formam com as linhas. Idealmente, $\tau = 0$.
- c_x e c_y representam a projeção do eixo ótico na imagem formada; usualmente, coordenadas do centro da imagem.

Já a matriz de rotação \mathbf{R} e o vetor de translação \vec{t} são chamados de parâmetros extrínsecos:

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (3.2)$$

$$\vec{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (3.3)$$

A estimativa dos parâmetros intrínsecos e extrínsecos é feita por meio de pontos característicos dos padrões de calibração. Estes são objetos de dimensões e estruturas conhecidas, que após serem capturados na imagem da câmera permitem a obtenção de pares correspondentes de pontos (x_w, y_w, z_w) e (x_f, y_f) para a Equação 3.1. Um padrão bastante comum é o plano xadrez da Figura 3.2: é fácil obter pontos do SCM se o tamanho do lado de cada quadrado é conhecido; pontos correspondentes do SCP são encontrados pelo processamento de imagem que busca os pixels das bordas dos quadrados.

É necessário explicitar a Equação 3.1 expressa uma visão geral do problema de calibração, sem detalhamento de etapas. Os métodos de calibração mais acurados, como Tsai [5] e Zhang [6], modelam a distorção causada pela lente da câmera. Isto transforma as relações lineares da Equação 3.1 em um problema não linear; o motivo é a adição de um termo de distorção radial ao sistema. Para resolver este novo sistema, um esquema de otimização, como o algoritmo de Levenberg-Marquadt [25] [26], é empregado.

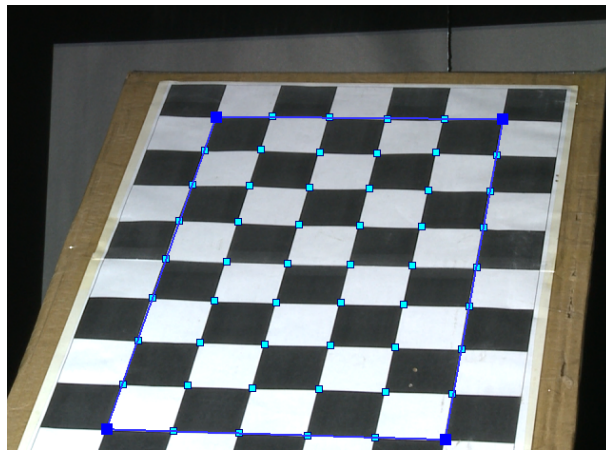


Figura 3.2: Padrão de calibração plano xadrez com pontos característicos marcados.

Em um sistema composto por uma câmera e um projetor, a calibração do último acontece de forma semelhante à da primeira. A câmera, já calibrada, captura a imagem de um padrão projetado. Este padrão deve ter estrutura e dimensões conhecidas; um exemplo é a imagem de um plano xadrez, cujo tamanho do lado de cada quadrado tem um número conhecido de pixels. O propósito é novamente estabelecer pares de pontos correspondentes

(x_w, y_w, z_w) e (x_f, y_f) para a Equação 3.1. Neste caso, (x_f, y_f) são coordenadas dos pontos de interesse da imagem usada para projeção. Já os pontos (x_w, y_w, z_w) do mundo são expressos no SCC. Isto significa que a rotação e a translação calculadas para o projetor são relativas à câmera. De maneira geral, se \mathbf{R}_c é a rotação da câmera em relação ao mundo e \mathbf{R}_p é a rotação do projetor em relação ao mundo, a rotação \mathbf{R}_{cp} , relativa entre câmera e projetor é

$$\mathbf{R}_{cp} = \mathbf{R}_p \mathbf{R}_c^{-1} \quad (3.4)$$

Se \vec{t}_c é a translação da câmera em relação ao mundo e \vec{t}_p é a translação do projetor em relação ao mundo, a translação relativa é

$$\vec{t}_{cp} = \vec{t}_p - (\mathbf{R}_{cp} \vec{t}_c) \quad (3.5)$$

O cálculo da profundidade de um ponto da cena ocorre através da interseção do plano de luz projetado com o raio que tem origem no centro ótico da câmera (Fig. 3.3). Conhecidas a rotação \mathbf{R}_{cp} e a translação \vec{t}_{cp} , a transformação que leva um ponto $\vec{p}_c = (x_c, y_c, z_c)$, do sistema de coordenadas da câmera, em um ponto $\vec{p}_p = (x_p, y_p, z_p)$ do sistema de coordenadas do projetor é

$$\vec{p}_p = \mathbf{R}_{cp} \vec{p}_c + \vec{t}_{cp} \quad (3.6)$$

A profundidade z_c pode ser calculada se a coordenada x_p , da projeção de \vec{p}_p , e a projeção $\vec{p}_f = (x_f, y_f, 1)$ de \vec{p}_c são conhecidas [2]:

$$z_c = \frac{t_x - x_p t_z}{\langle x_p \vec{r}_z - \vec{r}_x, \vec{p}_f \rangle} \quad (3.7)$$

Na Equação 3.7, t_x e t_z são elementos de \vec{t}_{cp} (Eq. 3.3), \vec{r}_x é um vetor formado pela primeira linha de \mathbf{R}_{cp} (Eq. 3.2), e \vec{r}_z é um vetor formado pela terceira linha da mesma matriz.

O sistema de reconstrução de geometria desenvolvido no presente trabalho permite que diversos métodos de calibração sejam empregados; a calibração está desacoplada do sistema, que necessita apenas das matrizes intrínsecas e extrínsecas da câmera e do projetor, independente do método utilizado, para reconstrução geométrica. Por motivos práticos, como a calibração rápida, o método de Tsai foi utilizado.

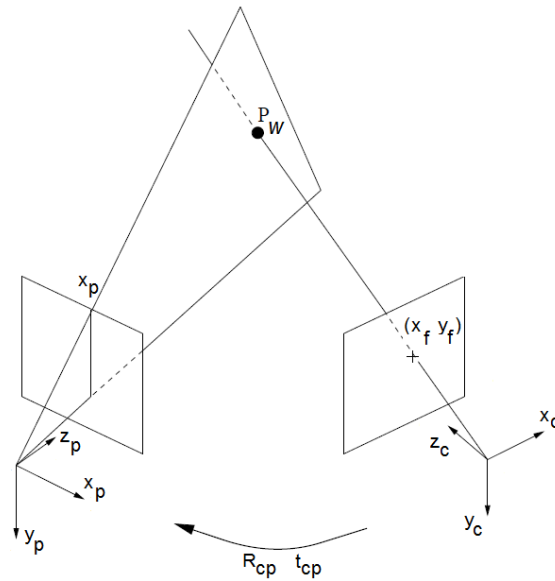


Figura 3.3: Triangulação [2]. Projetor à esquerda e câmera à direita.

3.1.1 Método de Tsai

O método de Tsai [5], proposto em 1987, tem como objetivo prover uma calibração de câmera versátil, que seja acurada (por isto inclui modelagem da distorção das lentes) e rápida (não deve incluir muitos parâmetros não-lineares). O método é dividido em duas etapas. Na primeira, os parâmetros extrínsecos que relacionam o SCM e o SCC são estimados: é calculada a matriz de rotação \mathbf{R} e as componentes t_x e t_y do vetor de translação \vec{t} . Na segunda etapa é realizado o cálculo da distância focal f e da componente t_z de \vec{t} .

O método assume como verdadeiras as seguintes premissas [27]:

- SCM e SCC possuem a mesma orientação. Então, \mathbf{R} é de fato uma matriz de rotação, ou seja, suas colunas e linhas formam vetores unitários e ortogonais.
- Os pontos característicos do padrão de calibração se encontram no plano $Z_w = 0$ do SCM.
- Os parâmetros intrínsecos da câmera, com exceção da distância focal f , são conhecidos. Contudo, é possível utilizar o método mesmo que estes parâmetros não estejam disponíveis. Neste caso, podemos considerar que as coordenadas do centro óptico, c_x e c_y , correspondem ao ponto do centro da imagem capturada pela câmera; que as colunas da imagem são perfeitamente perpendiculares às linhas ($\tau = 0$); e que o fatores de escala s_x e s_y são iguais à 1, ou seja, os *pixels* são quadrados.

O método de Tsai utiliza o plano xadrez (Figura 3.2) como padrão de calibração; sejam $(x_{wi}, y_{wi}, 0)$ as coordenadas conhecidas do i -ésimo ponto característico do padrão no SCM, e (x_{fi}, y_{fi}) as coordenadas, também conhecidas, deste mesmo ponto no SCP. As coordenadas (x_{di}, y_{di}) , correspondentes à projeção do i -ésimo ponto no SCI, considerando a distorção causada pela lente, são dadas por

$$(x_{di}, y_{di}) = (x_{fi} - c_x, y_{fi} - c_y). \quad (3.8)$$

A relação entre os pontos do SCI que consideram a distorção causada pela lente, (x_{di}, y_{di}) , e os pontos de projeção ideal (sem distorção), (x_{ui}, y_{ui}) , é

$$(x_{ui}, y_{ui}) = (1 + k_1 r_i^2)(x_{di}, y_{di}), \quad (3.9)$$

onde $r_i = \sqrt{x_{di}^2 + y_{di}^2}$, e k_1 é um parâmetro de valor desconhecido, chamado coeficiente de distorção radial, que depende da lente da câmera.

O modelo de distorção de lentes da Equação 3.9 é simplificado, pois considera apenas um tipo de distorção da lente, a radial; modelos mais completos consideram também a distorção tangencial. Contudo, segundo Tsai [5], este modelo simplificado é suficientemente acurado para a maioria das aplicações industriais.

Se a matriz de rotação \mathbf{R} (Equação 3.2) e o vetor de translação \vec{t} (Equação 3.3) são os parâmetros extrínsecos que relacionam o SCM e o SCC, e f é a distância focal da câmera, a relação entre os pontos (x_{wi}, y_{wi}, z_{wi}) do mundo e a projeção ideal (x_{ui}, y_{ui}) dos mesmos no SCI é

$$\begin{bmatrix} x_{ui} \\ y_{ui} \\ 1 \end{bmatrix} \cong \begin{bmatrix} fr_{11} & fr_{12} & fr_{13} & ft_x \\ fr_{21} & fr_{22} & fr_{23} & ft_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix}. \quad (3.10)$$

Com a informação de que $z_{wi} = 0$, a Equação 3.10 pode ser reescrita como

$$x_{ui} = f \frac{r_{11}x_{wi} + r_{12}y_{wi} + t_x}{r_{31}x_{wi} + r_{32}y_{wi} + t_z}, \quad (3.11)$$

$$y_{wi} = f \frac{r_{21}x_{wi} + r_{22}y_{wi} + t_y}{r_{31}x_{wi} + r_{32}y_{wi} + t_z}. \quad (3.12)$$

A partir da razão entre x_{wi} e y_{wi} (Equações 3.11 e 3.12), e também da relação entre (x_{di}, y_{di}) e (x_{wi}, y_{wi}) (Equação 3.9), é possível encontrar a seguinte relação:

$$\frac{x_{di}}{y_{di}} = \frac{r_{11}x_{wi} + r_{12}y_{wi} + t_x}{r_{21}x_{wi} + r_{22}y_{wi} + t_y}. \quad (3.13)$$

Rearranjando, obtemos

$$y_{di}r_{11}x_{wi} + y_{di}r_{12}y_{wi} + y_{di}t_x - x_{di}r_{21}x_{wi} - x_{di}r_{22}y_{wi} = x_{di}t_y. \quad (3.14)$$

Dividindo cada termo da Equação 3.14 por t_y :

$$y_{di}x_{wi} r_{11} t_y^{-1} + y_{di}y_{wi} r_{12} t_y^{-1} + y_{di} t_x t_y^{-1} - x_{di}x_{wi} r_{21} t_y^{-1} - x_{di}y_{wi} r_{22} t_y^{-1} = x_{di}. \quad (3.15)$$

Sejam $u_1 = r_{11} t_y^{-1}$, $u_2 = r_{12} t_y^{-1}$, $u_3 = t_x t_y^{-1}$, $u_4 = r_{21} t_y^{-1}$ e $u_5 = r_{22} t_y^{-1}$. Assim, a Equação 3.15 pode ser rearranjada:

$$\begin{bmatrix} y_{di}x_{wi} & y_{di}y_{wi} & y_{di} & -x_{di}x_{wi} & -x_{di}y_{wi} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = x_{di}. \quad (3.16)$$

Para resolver este sistema, são requeridos n pontos característicos do plano de calibração; n deve assumir 5 como valor mínimo, para que o sistema seja determinado. Contudo, geralmente são utilizados muitos pontos característicos, de forma que $n > 5$. Assim, o sistema linear possui um número maior de equações do que variáveis desconhecidas, e pode ser resolvido como um problema de mínimos quadrados: é necessário calcular u_1 , u_2 , u_3 , u_4 , e u_5 , a partir das coordenadas (x_{di}, y_{di}) e (x_{wi}, y_{wi}) conhecidas.

Encontrados os valores de u_1 , u_2 , u_3 , u_4 e u_5 , as componentes t_x e t_y de \vec{t} , e a matriz de rotação \mathbf{R} podem ser computadas através das condições de ortonormalidade desta matriz:

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1, \quad (3.17)$$

$$r_{21}^2 + r_{22}^2 + r_{23}^2 = 1, \text{ e} \quad (3.18)$$

$$r_{11} r_{21} + r_{12} r_{22} + r_{13} r_{23} = 0. \quad (3.19)$$

Como $u_1 = r_{11} t_y^{-1}$, $u_2 = r_{12} t_y^{-1}$, $u_3 = t_x t_y^{-1}$, $u_4 = r_{21} t_y^{-1}$ e $u_5 = r_{22} t_y^{-1}$, as Equações 3.17, 3.18, e 3.19 podem ser reescritas, respectivamente, como

$$(u_1^2 + u_2^2) t_y^2 + r_{13}^2 = 1, \quad (3.20)$$

$$(u_4^2 + u_5^2) t_y^2 + r_{23}^2 = 1, \text{ e} \quad (3.21)$$

$$(u_1 u_4 + u_2 u_5) t_y^2 + r_{13} r_{23} = 0. \quad (3.22)$$

Se r_{13} e r_{23} são eliminadas do sistema, a seguinte equação biquadrada é obtida:

$$A t_y^4 - B t_y^2 + 1 = 0. \quad (3.23)$$

Na Equação 3.23, $A = (u_1^2 + u_2^2 + u_4^2 + u_5^2)$ e $B = (u_1 u_5 - u_2 u_4)^2$; se $A = 0$, então $t_y^2 = \frac{1}{B}$. Caso contrário,

$$t_y^2 = \frac{B - \sqrt{B^2 - 4A}}{2A}. \quad (3.24)$$

Encontrado t_y , os valores de r_{11} , r_{12} , t_x , r_{21} , e r_{22} podem ser obtidos diretamente através das relações anteriormente definidas para u_1 , u_2 , u_3 , u_4 e u_5 . Os valores para r_{13} e r_{23} podem ser facilmente calculados explorando a restrição de que as linhas da matriz de rotação devem formar vetores unitários. Como existe a restrição de ortogonalidade, os valores da terceira linha, r_{31} , r_{32} e r_{33} , podem ser obtidos através o produto vetorial dos vetores formados pelas duas primeiras linhas de R .

No segundo estágio do método de Tsai é realizado o cálculo aproximado da distância focal f e da componente t_z de \vec{t} . Inicialmente, consideramos $k_1 = 0$: neste caso, pela Equação 3.9, $(x_{di}, y_{di}) = (x_{ui}, y_{ui})$; daí, as Equações 3.11 e 3.12 podem ser reescritas,

respectivamente, como

$$(r_{11}x_{wi} + r_{12}y_{wi} + t_x)f - x_{di}t_z = x_{di}(r_{31}x_{wi} + r_{32}y_{wi}), \text{ e} \quad (3.25)$$

$$(r_{21}x_{wi} + r_{22}y_{wi} + t_y)f - y_{di}t_z = y_{di}(r_{31}x_{wi} + r_{32}y_{wi}). \quad (3.26)$$

Nesta etapa do método são conhecidas a matriz de rotação \mathbf{R} , as componentes t_x e t_y do vetor de translação e as coordenadas (x_{di}, y_{di}) e (x_{wi}, y_{wi}) dos n pontos característicos do plano de calibração. Para determinar os valores das incógnitas t_z e f , as Equações 3.25 e 3.26 são resolvidas como um novo problema de mínimos quadrados.

O passo final do método é realizar o cálculo exato de f e t_z , e também considerar a distorção das lentes através da determinação do coeficiente k_1 . Com a adição da distorção radial, as Equações 3.25 e 3.26 se tornam, respectivamente

$$(r_{11}x_{wi} + r_{12}y_{wi} + t_x)f - x_{di}[1 + k_1(x_{di}^2 + y_{di}^2)](r_{31}x_{wi} + r_{32}y_{wi} + t_z) = 0, \text{ e} \quad (3.27)$$

$$(r_{21}x_{wi} + r_{22}y_{wi} + t_y)f - y_{di}[1 + k_1(x_{di}^2 + y_{di}^2)](r_{31}x_{wi} + r_{32}y_{wi} + t_z) = 0. \quad (3.28)$$

Estas últimas equações não lineares podem ser resolvidas com o método de Levenberg-Marquadt, adotando como solução inicial $k_1 = 0$ e os valores aproximados de f e t_z anteriormente calculados.

3.2 Calibração de Cor

Os sensores da câmera e do projetor normalmente respondem de modo diferente à variações de brilho. Isto significa que as cores possuem coordenadas distintas nos sistemas dos dois dispositivos. Considere ainda a luminosidade ambiente, e que cada objeto da cena iluminada é composto por um tipo de material que reflete a luz de acordo com propriedades inerentes. Como a codificação (b, s) -BCSL é baseada no uso de cores, fica evidente a necessidade de uma transformação que aproxime a cor capturada pela câmera, modulada pela cena, àquela projetada pelo projetor; isto é, uma calibração de cor é requerida.

Uma transformação linear simples, que oferece boa aproximação, é dada por

$$[r_p \ g_p \ b_p]^T = \mathbf{M} \cdot [r_c \ g_c \ b_c]^T, \quad (3.29)$$

onde $[r_p \ g_p \ b_p]^T$ é a cor projetada, $[r_c \ g_c \ b_c]^T$ é a cor capturada e \mathbf{M} é uma matriz 3x3 que leva o espaço de cor da imagem capturada pela câmera para o espaço de cor do projetor.

Se b cores são utilizadas na codificação, \vec{c}_i é a i -ésima cor projetada e \vec{r}_i é a respectiva i -ésima cor de resposta, a matriz de calibração \mathbf{M} pode ser computada por meio de mínimos quadrados, como mostra a Equação 3.30.

$$[\vec{c}_1^T \ \vec{c}_2^T \ \dots \ \vec{c}_{b+1}^T, \vec{c}_{b+2}^T] = \mathbf{M} \cdot [\vec{r}_1^T \ \vec{r}_2^T \ \dots \ \vec{r}_{b+1}^T, \vec{r}_{b+2}^T] \quad (3.30)$$

As cores preto e branco são representadas por \vec{c}_{b+1}^T e \vec{c}_{b+2}^T . Elas são responsáveis pelos limites inferior e superior, respectivamente, na minimização expressa pelo sistema da Equação 3.30.

O digitalizador desenvolvido aplica a matriz \mathbf{M} sobre toda imagem recebida da câmera. Este é um procedimento que antecede a localização de bordas entre as faixas verticais.

3.3 Codificação (b, s) -BCSL

O padrão de luz estruturada (b, s) -BCSL (em inglês, Boundary Coded Structured Light) tem como objetivo prover alta resolução sem impor fortes restrições sobre as características fotométricas das cenas capturadas.

Neste contexto, prover alta resolução significa projetar um grande número de faixas verticais coloridas sobre a cena. Cada transição entre duas faixas adjacentes tem um código único, dado pela cor da esquerda e da direita. Quanto maior o número de transições, necessariamente únicas, sobre um objeto, mais detalhada é a geometria obtida. Existem duas formas de controlar a quantidade de faixas (e consequentemente, de transições) projetadas com (b, s) -BCSL: o parâmetro b denota o número de cores utilizadas para as faixas; no total, seis cores são permitidas: as primárias vermelho (R), verde (G), azul (B) e as respectivas complementares, ciano (C), magenta (M) e amarelo (Y). Já o parâmetro s denota o número de *slides* S de faixas diretamente relacionados à codificação e decodificação das transições. Para $s = 1$ somente um *slide* é necessário, e cada transição é unicamente identificada por duas cores. Para $s > 1$ as restrições temporais tornam-se

maiores. Mais *slides* devem ser projetados para que a geometria de uma cena possa ser reconstituída; isto significa que objetos que se movimentam rapidamente podem sofrer erros de detecção. Se $s = 2$, por exemplo, dois slides, S_1 e S_2 , são requeridos. Cada transição entre faixas passa então a ser identificada por quatro cores, esquerda e direita, em S_1 e S_2 . O número de diferentes códigos possíveis para transições é dado por $[b(b-1)]^s$, assumindo que duas faixas consecutivas não podem ter a mesma cor.

A codificação (b, s) -BCSL emprega *slides* de cores complementares para lidar com restrições fotométricas de uma cena. A projeção de um *slide* S é sempre seguida pela projeção do seu complementar \bar{S} , de maneira que uma cor de S sempre tem a complementar na mesma posição em \bar{S} . Isto é equivalente ao uso de um código binário (0 ou 1) para cada canal de cor. Então, para o caso $s = 2$, são projetados S_1, \bar{S}_1, S_2 e \bar{S}_2 em sequência; cada transição é identificada pela cor da esquerda e direita em S_1 e S_2 , e os *slides* complementares eliminam restrições relacionadas à propriedades de reflexão da cena. Estes últimos são essenciais para localização acurada da borda entre cores durante o processamento das imagens que antecede a reconstrução.

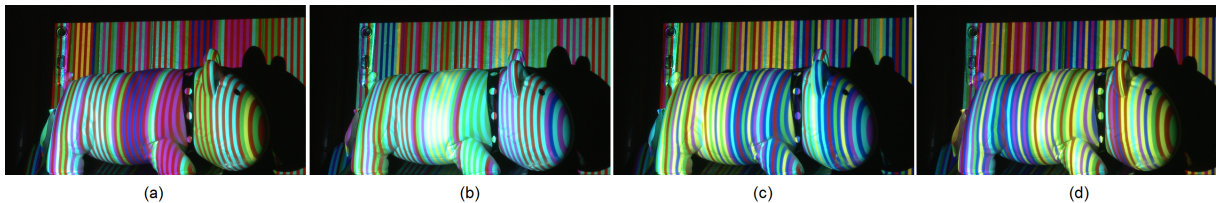


Figura 3.4: *Slides* projetados sobre o cachorro utilizando $(6,2)$ -BCSL. Nas imagens, (a) e (c) são *slides* utilizados para codificação; (b) é o complementar de (a) e (d) é o complementar de (c).

Nesta dissertação, a codificação $(6,2)$ -BCSL é utilizada. Ela permite 900 transições únicas, e impõe um mínimo de restrição na coerência temporal da cena.

O problema de geração de uma sequência particular de faixas de b cores e s slides é modelado através da descoberta de um caminho euleriano em um grafo G . Este grafo possui b^s vértices; cada um corresponde à uma escolha das b cores para uma posição específica nos s *slides*. Como exemplo, o $(3,2)$ -BCSL tem $b = 3$ e $s = 2$. Neste cenário G tem nove vértices. É o que acontece na Figura 3.5(a). Nela, o vértice 01 corresponde à projeção, na mesma posição, da cor 0 no primeiro *slide* e da cor 1 no segundo.

As arestas de G indicam as transições possíveis entre duas faixas consecutivas. Não são permitidas transições que repetem a mesma cor, consecutivamente, em um mesmo

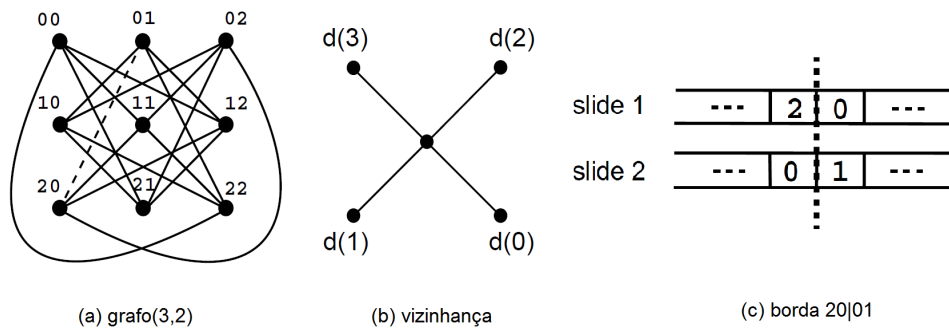


Figura 3.5: Codificação (3,2)-BCSL [3].

slide: no grafo da Figura 3.5 não existe uma conexão entre o vértice 01 e 02, porque isto significa a repetição da cor 0 para faixas adjacentes do primeiro *slide*.

Para o caso geral da codificação (b, s) -BCSL, cada vértice possui $(b - 1)^s$ vizinhos. A Figura 3.5(b) exemplifica a vizinhança de um vértice para o caso (3,2); apesar da Figura 3.5 não mostrar um grafo orientado, o sentido da passagem entre vértices indica transições distintas. Por isso, é adequado interpretar que cada vértice de G tem um conjunto de $(b - 1)^s$ arcos incidentes e outro de $(b - 1)^s$ arcos no sentido inverso.

Como já dito, o problema de geração de uma sequência de s *slides* para projeção é modelado através da descoberta de um caminho sem arcos repetidos no grafo orientado G . O número máximo de faixas é conseguido quando todo o caminho euleriano é percorrido. Nesta dissertação, um caminho completo de 900 transições para o (6,2)-BCSL é usado; sequências menores de faixas utilizam apenas uma seção do mesmo.

vértices	d(0)	d(1)	d(2)	d(3)
V(00)	0	3	6	9
V(01)	14	17	19	11
V(02)	28	34	22	24
V(10)	26	29	18	21
V(11)	1	31	33	35
V(12)	15	4	8	13
V(20)	16	23	32	12
V(21)	27	5	7	25
V(22)	2	10	20	30

Tabela 3.1: Tabela de decodificação para (3,2)-BCSL [3]

Decodificar o código de uma transição de cores através das imagens da câmera significa encontrar a posição do arco correspondente no caminho euleriano do grafo G . O algoritmo de decodificação faz uso de uma tabela que permite o cálculo dessa posição em $O(1)$.

Para o caso $b = 3$ e $s = 2$, considere a Tabela 3.1, cujo grafo correspondente é exibido na Figura 3.5. Cada linha da tabela corresponde a um vértice de G ; as colunas representam os vizinhos. Se uma transição do vértice v_i para v_j foi detectada, para encontrar a posição da mesma no caminho euleriano é necessário determinar a localização de v_j na vizinhança de v_i . Como exemplo, o arco que tem origem no vértice 11 e atinge o vértice 02, que é o vizinho $d(2)$ de 11 (Fig. 3.5(b)), é o trigésimo terceiro do caminho. Isto significa que esta é a transição de número 33 nos *slides* de cor.

Com a decodificação de cada transição detectada, é possível calcular a posição de cada uma em coordenadas do projetor. Como o sistema está calibrado, a posição das mesmas transições, desta vez no sistema de coordenadas da câmera, pode ser obtida. Isto permite o estabelecimento da correspondência que leva à reconstrução geométrica.

4 DESENVOLVIMENTO DO DIGITALIZADOR TRIDIMENSIONAL

Este capítulo está dividido em três seções. A primeira descreve detalhadamente os procedimentos necessários para configuração do sistema de digitalização. Trata-se de etapas preparatórias, anteriores à captura de imagens da cena. São discutidas as questões de preparação do ambiente, posicionamento de equipamentos, calibração de cor e calibração câmera-projetor.

A segunda seção, extração de geometria e fotometria, trata da sequência de passos da *pipeline* de reconstrução geométrica. São detalhados os algoritmos utilizados pelo digitalizador tridimensional desenvolvido.

A terceira seção descreve o protocolo de erro desenvolvido para avaliar a geometria reconstruída pelo digitalizador.

4.1 Configuração do Sistema

A Figura 4.1 é uma foto da configuração do digitalizador. Na montagem do sistema, deve-se garantir que os volumes de visualização da câmera e do projetor se interceptem. Caso contrário, não há como triangular coordenadas. Neste trabalho, a distância entre câmera e projetor é de cerca de 50cm. A distância em relação ao objeto de reconstrução é aproximadamente 150cm.

A câmera utilizada, Sony PMW-EX3, captura imagens na resolução máxima 1920x1080; ela possui saída *IEEE 1394-A* para transferência de vídeo em tempo real e entrada *genlock* para sincronização com um sinal externo. O projetor, Benq W1000, projeta em resolução 1920x1080, e tem baixa latência na troca de cores, por ser um modelo com processamento digital de luz (em inglês, DLP). Estas características tornam o sistema apto para geração de vídeo 3D em alta resolução através do (b, s) -BCSL. Para isto, um sincronizador é necessário. Ele garante que um *slide* de cores projetado corresponda à uma imagem



Figura 4.1: Posicionamento de equipamentos e objetos.

capturada, isto é, que captura e projeção se iniciem ao mesmo tempo; conseqüentemente, o sensor da câmera estará exposto às cores de um *slide* durante todo tempo de sua projeção. Devido ao alto custo, não foi possível utilizar um sincronizador neste trabalho. Dito isto, esclarece-se que as imagens utilizadas nesta dissertação são extraídas manualmente dos vídeos capturados e gravados no formato MPEG-4.

Para a calibração de cores, o balanço de branco da câmera deve ser ajustado às condições de luminosidade do ambiente. Este deve estar tão livre quanto possível de fontes externas de iluminação; é importante que esta condição se mantenha no momento da captura de imagens de cenas para reconstrução. Para ajustar o branco da câmera, projeta-se a cor branca sobre um anteparo também branco, como uma folha de papel. Somente este anteparo deve estar enquadrado no visor da câmera; pode-se ajustar o foco para isto. Se o ajuste é bom, cores da cena vistas diretamente pelo olho são próximas àquelas vistas através do visor da câmera.

A calibração de cores ocorre após o balanço de branco. O passo inicial neste processo é projetar cada uma das cores usadas na codificação (b, s) -BCSL sobre um anteparo branco não especular, e capturar as imagens resultantes; imagens para preto e branco também devem ser capturadas. É necessário que somente o anteparo, iluminado pelas cores usadas, seja visível nas imagens obtidas. Novamente, pode-se ajustar o foco da câmera para isto.

O algoritmo de calibração de cores percorre os pixels das imagens capturadas para cada cor. Para cada posição (i, j) nas imagens captadas, valores dos pixels de preto, branco, vermelho, verde, azul, ciano, magenta e amarelo são usados para formar uma matriz 3×8

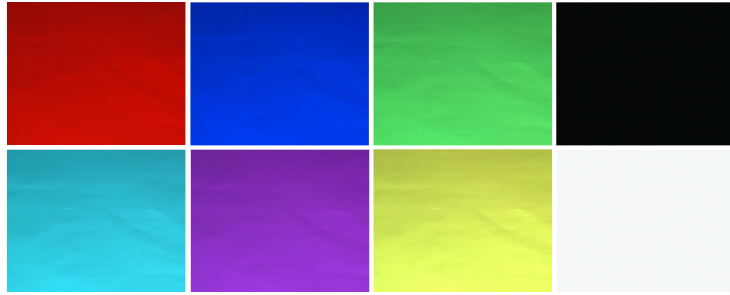


Figura 4.2: Imagens capturadas para calibração de cores. Cada cor é projetada sobre um anteparo branco, como uma folha de papel.

de cores de resposta. A matriz 3×8 de cores projetadas é formada por cores RGB ideais, isto é, preto $[0, 0, 0]^T$, branco $[255, 255, 255]^T$, vermelho $[255, 0, 0]^T$, verde $[0, 255, 0]^T$, azul $[0, 0, 255]^T$, ciano $[0, 255, 255]^T$, magenta $[255, 0, 255]^T$ e amarelo $[255, 255, 0]^T$. Uma matriz de calibração de cores é calculada, por mínimos quadrados, para cada matriz de cores de resposta, conforme descrito pela Equação 3.30. A matriz final de calibração de cores é dada pela média das matrizes obtidas.

A calibração câmera-projetor é a etapa seguinte na configuração do sistema. Primeiro captura-se a imagem de um padrão xadrez plano para calibração da câmera. Em seguida, para calibrar o projetor, captura-se a imagem de um padrão xadrez projetado sobre um anteparo branco. Este anteparo deve estar no mesmo lugar onde o plano de calibração para câmera foi posicionado.

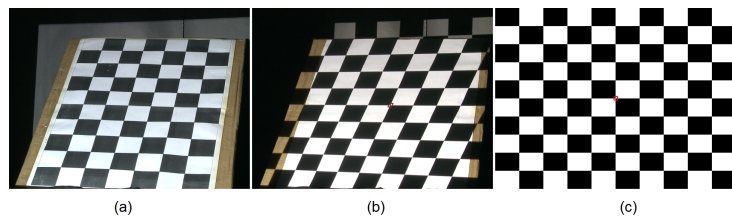


Figura 4.3: Calibração câmera-projetor. (a) imagem capturada para calibração da câmera. (b) imagem capturada para calibração do projetor. (c) padrão projetado para calibração do projetor.

O calibrador câmera-projetor utilizado neste trabalho necessita apenas das duas imagens capturadas para calcular todos os parâmetros intrínsecos e extrínsecos segundo o método de Tsai [5]. Para calibração da câmera é preciso conhecer o tamanho, em centímetros, dos lados dos retângulos que formam o padrão xadrez (Fig. 4.3(a)). Para calibração do projetor deve-se conhecer a mesma medida, porém em pixels; esta é relativa aos retângulos da imagem xadrez projetada (Fig. 4.3(c)). Contudo, esta etapa da confi-

guração do sistema está desacoplada do digitalizador desenvolvido. Outros calibradores, como o de Bouguet [28], que é uma implementação do método de Zhang [6], podem ser usados. O importante é que as matrizes intrínsecas e extrínsecas sejam obtidas.

É necessário esclarecer que para captura de cenas para reconstrução, o foco, a posição e a orientação da câmera e do projetor não devem sofrer reajustes após a calibração. Contudo, se câmera e projetor estão montados sobre uma superfície suficientemente firme, a mesma pode ser deslocada.

Com as condições de iluminação do ambiente controladas, e obtidas as matrizes de calibração de cor e calibração câmera-projetor, o sistema está configurado. Pode-se iniciar a captura de imagens iluminadas com o código (b, s) -BCSL para reconstrução de geometria.

4.2 Extração de Geometria e Fotometria

Entre a captura de uma cena iluminada com faixas geradas pelo código $(6, 2)$ -BCSL e a reconstrução tridimensional existe uma sequência de etapas a serem seguidas, como mostra a Figura 4.4.

Como entrada para a *pipeline* do digitalizador, são requeridas quatro imagens: I_1 , \bar{I}_1 , I_2 e \bar{I}_2 . As imagens complementares I_1 e \bar{I}_1 correspondem à capturas da cena iluminada pelos slides de cor S_1 e \bar{S}_1 , respectivamente. Do mesmo modo, I_2 e \bar{I}_2 são imagens complementares capturadas da cena iluminada por S_2 e \bar{S}_2 . Neste trabalho, uma quinta imagem, iluminada com branco, é usada para captura da textura da cena.

Inicialmente, as imagens I_1 , \bar{I}_1 , I_2 e \bar{I}_2 passam por uma filtragem passa-baixa (é utilizado um filtro gaussiano 5×5) para remoção de ruído. Em seguida, a matriz de calibração de cores é aplicada sobre cada uma delas. O resultado deste processamento são as imagens calibradas C_1 , \bar{C}_1 , C_2 e \bar{C}_2 .

O próximo passo é gerar, por meio de subtração, as imagens de diferença e imagens de cores projetadas. Estas imagens são empregadas para localização das fronteiras entre faixas e para a identificação correta das cores projetadas. Para o par imagens calibradas C_1 e \bar{C}_1 é gerada uma imagem de diferença, D_1 , e uma imagem de cores projetadas, P_1 . Do mesmo modo, para o par C_2 e \bar{C}_2 é gerada a imagem de diferença D_2 e a imagem de cores projetadas P_2 .

As imagens de diferença, D_1 e D_2 , e as imagens de cores projetadas, P_1 e P_2 , são

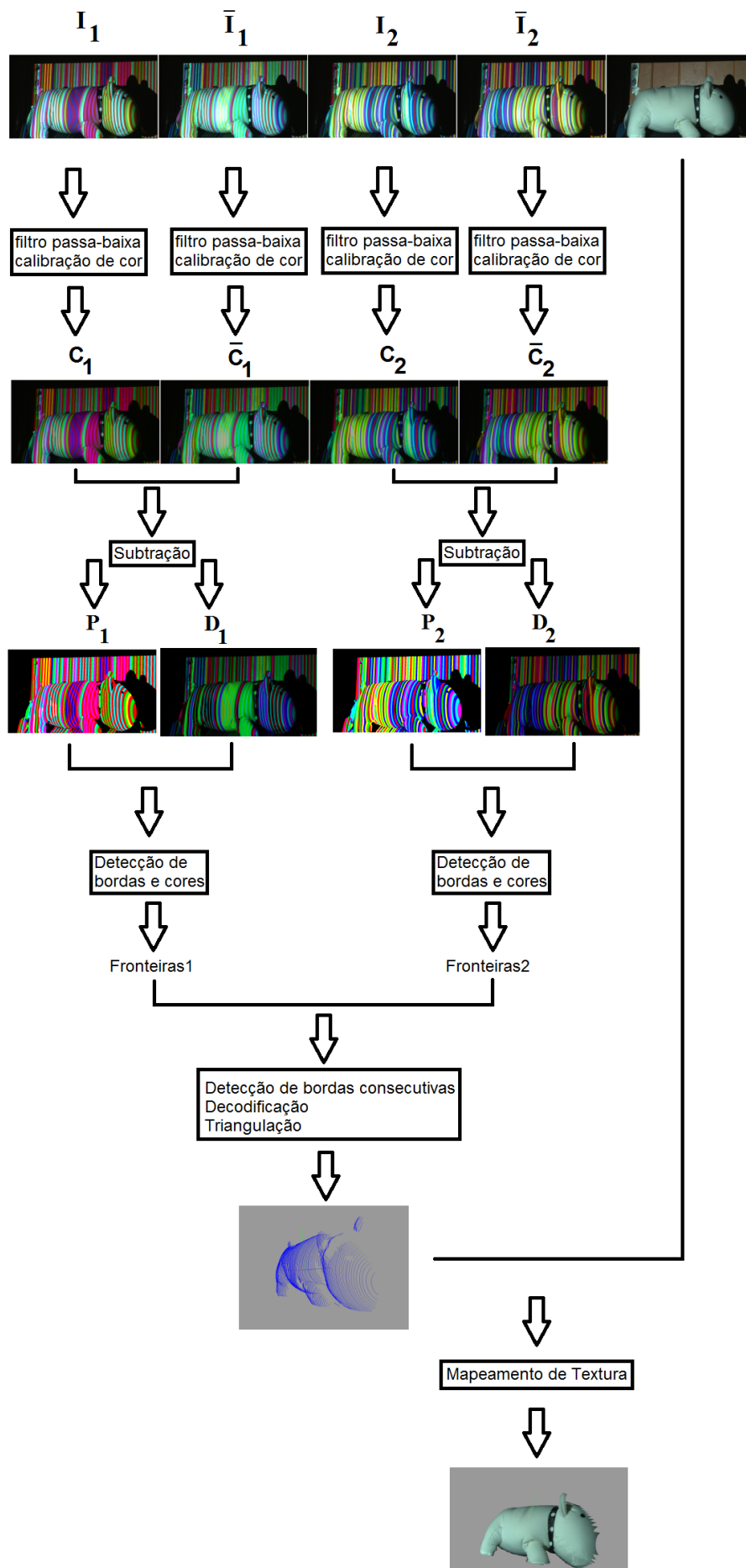


Figura 4.4: Sequência de passos na *pipeline* de reconstrução geométrica.

entradas para o algoritmo de detecção de fronteiras entre faixas verticais. As posições das fronteiras são identificadas através da busca de passagens por zero nos canais R, G e B das imagens de diferença. As cores à esquerda e à direita de cada fronteira são localizadas através de buscas nas imagens de cores projetadas. Ao final deste processamento, é formada uma lista de fronteiras para cada par de imagens complementares. Cada fronteira guarda a posição de uma transição entre as faixas verticais adjacentes, e também a cor à esquerda e à direita da transição.

O passo final para reconstrução geométrica é relacionar as fronteiras encontradas no primeiro par de imagens complementares com as fronteiras do segundo. Como as faixas são projetadas sempre sobre as mesmas posições, as fronteiras do primeiro e segundo par, em geral, estão em regiões próximas. Para cada fronteira do primeiro par, uma correspondente é buscada em uma vizinhança 7×7 na segunda lista de fronteiras. Encontrada a correspondência, quatro cores, duas de cada fronteira, são usadas para decodificação. Finalmente, com a posição da fronteira e o identificador da transição decodificada, é realizada a triangulação. Neste ponto da *pipeline* uma nuvem de pontos tridimensionais é obtida. A composição da nuvem com a textura da cena é o produto final do digitalizador.

4.2.1 Detecção de Fronteiras

Como dito, a localização das fronteiras entre faixas verticais é fundamentada na busca de passagens por zero nos canais de cor R, G e B das imagens de diferença (Figura 4.5d). As imagens de diferença, D_1 e D_2 , são geradas por meio da subtração das imagens complementares calibradas: $D_1 = \bar{C}_1 - C_1$, e $D_2 = \bar{C}_2 - C_2$.

Sejam D_n^R , D_n^G e D_n^B os canais de cor de uma imagem de diferença D_n , $n = 1, 2$. A detecção de fronteiras baseada somente na busca das posições (i, j) onde $D_n^R(i, j) = 0$ ou $D_n^G(i, j) = 0$ ou $D_n^B(i, j) = 0$ está sujeita a ruído. A primeira observação a ser feita na busca de passagens por zero é que as faixas projetadas pelo código (b, s) -BCSL são verticais. Isto significa que cada linha de D_n deve ser examinada de forma independente. Outro fato importante para detecção é que os pontos à esquerda e à direita de uma passagem por zero têm valores de sinais opostos, e entre eles existe alta variação de intensidade. Portanto, uma passagem por zero deve satisfazer, em pelo menos um canal

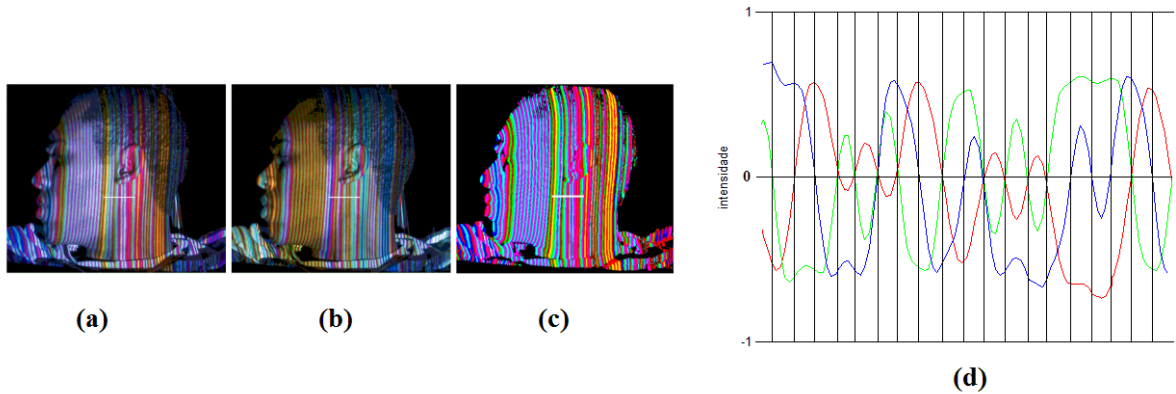


Figura 4.5: Localização de fronteiras. (a) imagem I_1 . (b) imagem \bar{I}_1 . (c) imagem de cores projetadas P_1 . (d) variação de intensidade dos canais R, G e B da imagem de diferença D_1 , para linha marcada em branco. As linhas verticais do gráfico indicam as fronteiras encontradas entre faixas.

de cor K , à seguinte restrição:

$$|D_n^K(i - \delta_1, j)| + |D_n^K(i + \delta_1, j)| \geq l \quad (4.1)$$

Na Equação 4.1, i denota as colunas da imagem de diferença D_n , e j as linhas. O parâmetro l , limiar de passagem, denota a variação de intensidade mínima que caracteriza uma passagem por zero. Nesta dissertação, l assume o valor 5. O parâmetro δ_1 é um pequeno deslocamento que descreve os pontos vizinhos à esquerda e à direita da passagem por zero. Não faz sentido adotar valores grandes para δ_1 . Tais valores podem causar a localização inacurada da passagem por zero, ou até mesmo uma localização errada, caso o valor de δ_1 ultrapasse a largura de uma faixa. Nesta dissertação δ_1 assume o valor 1, o menor possível.

Inicialmente, para cada linha j de D_n , as passagens por zero são buscadas e geradas para os canais de cor R, G e B de forma independente. Se a Equação 4.1 é atendida em uma posição (i, j) para $D_n^R(i, j)$ ou $D_n^G(i, j)$ ou $D_n^B(i, j)$, uma fronteira é gerada. Uma fronteira é definida como uma quintupla $f = (x, j, a, c_e, c_d)$:

- x é um número real que denota a coordenada horizontal da passagem por zero com precisão subpixel;
- j denota a linha da imagem D_n onde a passagem por zero foi encontrada;
- a é a amplitude da passagem por zero.

- c_e é a cor da faixa projetada à esquerda de x .
- c_d é a cor da faixa projetada à direita de x .

A amplitude a é dada por:

$$a = |D_n^K(i - \delta_1, j)| + |D_n^K(i + \delta_1, j)|. \quad (4.2)$$

A coordenada x é calculada por meio de interpolação:

$$x = \frac{D_n^K(i + \delta_1, j) \cdot (i - \delta_1) + D_n^K(i - \delta_1, j) \cdot (i + \delta_1)}{a}. \quad (4.3)$$

Mesmo buscando passagens por zero nos canais R, G e B separadamente, é possível que diversas fronteiras sejam detectadas para uma única transição entre faixas. Caso uma ou mais fronteiras f_2, f_3, \dots, f_n sejam identificadas, e a distância entre suas coordenadas de cruzamento x_2, x_3, \dots, x_n seja menor ou igual à uma distância δ_2 em relação à coordenada x_1 de uma fronteira f_1 previamente encontrada, uma única fronteira f_c , resultante da combinação linear de todas elas, é gerada. A coordenada combinada de cruzamento x_c é dada por

$$x_c = \frac{x_1 \cdot a_1 + x_2 \cdot a_2 + \dots + x_n \cdot a_n}{a_1 + a_2 + \dots + a_n}, \quad (4.4)$$

e a amplitude combinada a_c é calculada como

$$a_c = \frac{a_1^2 + a_2^2 + \dots + a_n^2}{a_1 + a_2 + \dots + a_n}. \quad (4.5)$$

Nesta dissertação o parâmetro δ_2 assume o valor 3; para este valor não há nenhuma transição redundante dentro de uma região de 3 pixels consecutivos. Valores grandes para δ_2 não devem ser usados; eles podem filtrar fronteiras referentes à transições entre faixas distintas.

O segundo passo da detecção de fronteiras é realizar uma filtragem conjunta das fronteiras detectadas em todos os canais de cor. A lista de fronteiras formada para uma linha da imagem é percorrida; para cada fronteira $f_i = (x_i, j_i, a_i, c_{ei}, c_{di})$ da lista, é buscada a fronteira $f_m = (x_m, j_m, a_m, c_{em}, c_{dm})$ de maior amplitude dentro da distância δ_2 . A posição

final de cruzamento por zero x_f computada para esta região da imagem D_n é

$$x_f = \frac{x_i \cdot a_i + x_m \cdot a_m}{a_i + a_m}. \quad (4.6)$$

Encontradas as fronteiras entre faixas, é preciso descobrir as cores à esquerda e à direita de cada uma delas. Isto é realizado através de uma busca na imagem de cores projetadas.

O objetivo das imagens de cores projetadas é capturar com exatidão (isto é, sem interferência de reflexões dos objetos da cena) as cores das faixas diretamente relacionadas à codificação (6,2)-BCSL; isto significa que P_1 deve capturar as cores das faixas projetadas em I_1 (Figura 4.5c), assim como P_2 deve capturar as cores das faixas de I_2 . A captura fiel destas cores é de vital importância para decodificação correta das fronteiras, e é possível por causa da projeção de faixas de cores complementares. A imagem de cores projetadas P_n , $n = 1,2$, é gerada através da imagem de diferença D_n :

- $P_n^R(i, j) = 0$, se $D_n^R(i, j) > 0$; 1 caso contrário.
- $P_n^G(i, j) = 0$, se $D_n^G(i, j) > 0$; 1 caso contrário.
- $P_n^B(i, j) = 0$, se $D_n^B(i, j) > 0$; 1 caso contrário.

Para uma fronteira $f = (x, j, a, c_e, c_d)$, a cor à esquerda e à direita são dadas por $c_e = P_n(x - \delta_2, j)$ e $c_d = P_n(x + \delta_2, j)$ respectivamente. Se $P_n(x - \delta_2, j)$ não é uma cor válida, como preto, δ_2 é decrementado, e uma nova verificação é feita; a busca segue até a posição $P_n(x - 1, j)$. Procedimento similar é realizado para encontrar a cor à direita: se a cor em $P_n(x + \delta_2, j)$ não é válida, δ_2 é decrementado, e a busca segue até $P_n(x + 1, j)$. Se cores válidas para esquerda e direita são encontradas, e elas são distintas, uma fronteira completa é formada; caso contrário, a mesma é descartada.

4.2.2 Detecção de Fronteiras no Tempo

O passo final para obtenção de pontos tridimensionais a partir das imagens capturadas é relacionar as fronteiras detectadas na imagem de diferença D_1 com as fronteiras de encontradas em D_2 ; para cada fronteira $f_p = (x_p, j_p, a_p, c_{ep}, c_{dp})$ da primeira lista de fronteiras, uma fronteira correspondente $f_q = (x_q, j_q, a_q, c_{eq}, c_{dq})$ é procurada em uma região 7x7 na segunda lista.

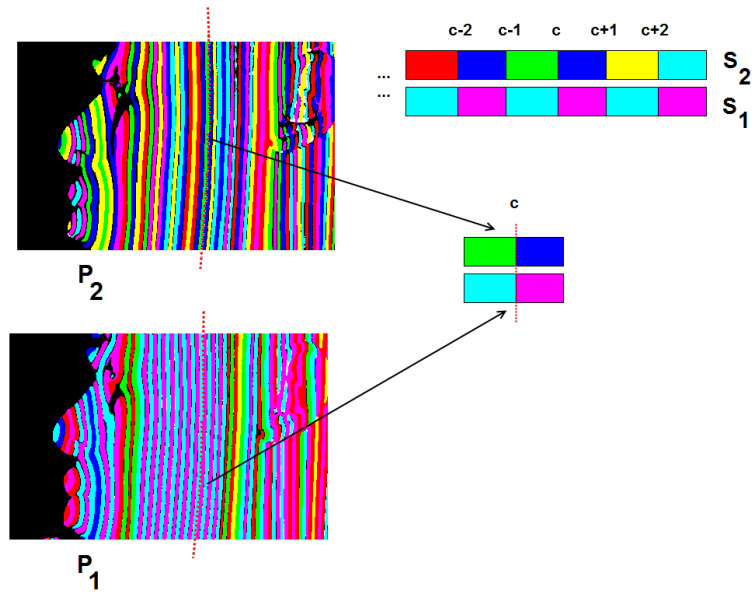


Figura 4.6: Decodificação de fronteiras (6, 2)-BCSL.

Se uma correspondência é encontrada, e as cores c_{ep} , c_{dp} , c_{eq} , c_{dq} formam um código (6, 2)-BCSL válido, a decodificação da fronteira é possível, e um ponto tridimensional pode ser gerado. Antes disto, uma última interpolação entre as posições x_p e x_q é realizada:

$$x_p = \frac{x_p \cdot a_p + x_q \cdot a_q}{a_p + a_q}. \quad (4.7)$$

Com o código c da fronteira, e as coordenadas x_p e j_p , o cálculo de profundidade para a mesma pode ser realizado de acordo com a Equação 3.7.

4.3 Protocolo para Avaliação de Erro

Para avaliar a geometria reconstruída pelo digitalizador tridimensional desenvolvido, um protocolo de erro é proposto no presente trabalho.

Optou-se por utilizar vários planos com diferentes orientações como objetos de digitalização para estimar a qualidade da reconstrução. A razão para isto é a praticidade pois a geometria simples do plano facilita a medição e comparação de pontos. Como estamos interessados em avaliar a digitalização de profundidades, o uso de objetos de superfícies curvas ou multifacetados complicaria a determinação do terreno verdade.

Apartir da nuvem de pontos gerada pelo digitalizador, a equação do plano para a mesma é estimada por meio de mínimos quadrados. Para comparar a semelhança entre

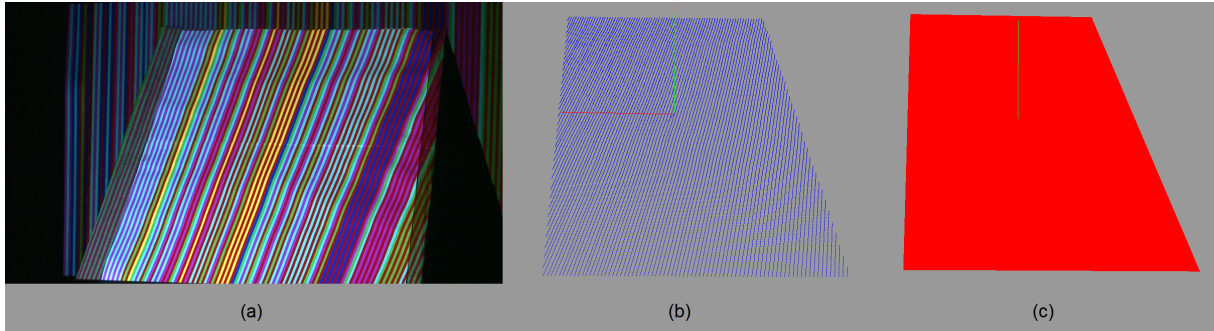


Figura 4.7: Ajuste de planos. (a) projeção de *slides* sobre o plano. (b) nuvem de pontos gerada pelo digitalizador. (c) plano estimado para a nuvem de pontos, minimizando erro quadrático.

dois planos $\pi_1 : a_1x + b_1y + c_1z = d_1$ e $\pi_2 : a_2x + b_2y + c_2z = d_2$ foi estabelecida uma métrica de distância no espaço projetivo:

$$d = \sqrt{\left(\frac{a_1}{d_1} - \frac{a_2}{d_2}\right)^2 + \left(\frac{b_1}{d_1} - \frac{b_2}{d_2}\right)^2 + \left(\frac{c_1}{d_1} - \frac{c_2}{d_2}\right)^2} \quad (4.8)$$

Na Equação 4.8, cada plano $\pi : ax + by + cz = d$ é considerado um ponto $P = (a, b, c, d) \in RP^3$. Além disso, $[a, b, c]$ é um vetor unitário em R^3 .

A abordagem de distância da Equação 4.8 é de visualização mais fácil quando uma analogia com RP^2 é feita: neste caso, se o plano $z = 1$ é considerado como plano afim euclidiano mergulhado em RP^2 , um ponto projetivo (x, y, z) pode ser identificado com $(x/z, y/z, 1)$, e a distância entre dois pontos, $(x_1/z_1, y_1/z_1, 1)$ e $(x_2/z_2, y_2/z_2, 1)$, é calculada como $\sqrt{\left(\frac{x_1}{z_1} - \frac{x_2}{z_2}\right)^2 + \left(\frac{y_1}{z_1} - \frac{y_2}{z_2}\right)^2}$ (Figura 4.8).

O objetivo do protocolo desenvolvido é medir a acurácia e a precisão do digitalizador construído com a codificação (b, s) -BCSL. A acurácia caracteriza a proximidade entre o valor real de determinada grandeza e o valor experimental obtido para a mesma; é, portanto, uma medida de exatidão. A precisão descreve o nível de variação entre as medidas experimentais; ou seja, precisão é uma medida de repetibilidade. Estabelecidos estes dois conceitos, é necessário discutir os possíveis resultados do protocolo:

- alta acurácia, alta precisão: o melhor resultado possível. Todas as reconstruções tem alto nível de exatidão, e a variabilidade entre elas é baixa. Isto indica alta confiabilidade.
- alta acurácia, baixa precisão: as reconstruções, em média, aproximam geometria alvo com exatidão. Contudo, as reconstruções individuais apresentam resultados

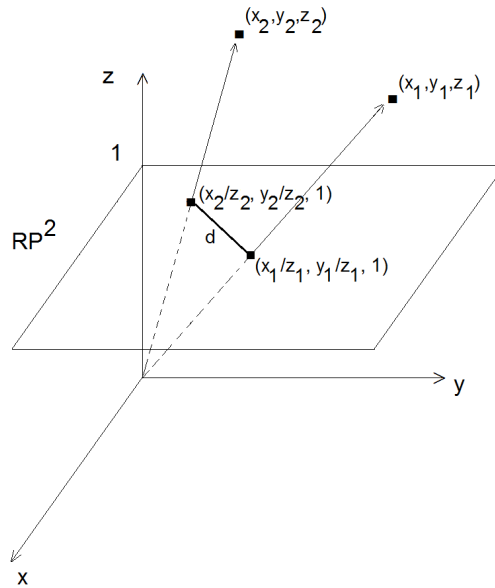


Figura 4.8: Distância entre dois pontos do RP^2 .

diversos. Isto significa que reconstruções com alta acurácia não são sempre repetidas; portanto, a confiabilidade está comprometida.

- baixa acurácia, alta precisão: este resultado significa que aproximações com baixa exatidão são constantemente repetidas. Configurações do digitalizador, como um número excessivo de faixas, que levem a este estado de erro sistemático não devem ser utilizadas.
- baixa acurácia, baixa precisão: as reconstruções, em média, tem baixa exatidão; além disso, a variabilidade entre elas é grande.

É necessário explicitar que neste contexto alta acurácia significa que a distância entre os planos comparados é pequena. Do mesmo modo, alta precisão significa baixa variabilidade entre as distâncias.

Existe uma outra observação a ser feita, em relação à unidade de distância do protocolo proposto. Como os pontos gerados pelo digitalizador tem centímetros como unidade de comprimento, e os coeficientes a, b, c, d das equações dos planos são gerados à partir destes pontos, a unidade dos coeficientes também é centímetros. A divisão de a, b e c por d na Equação 4.8 cancela as unidades de comprimento. Por isso, o cálculo de distância do protocolo de erro é adimensional.

Um dos objetivos desta dissertação é avaliar acurácia e precisão de variados números de faixas projetadas oferecidas pela codificação (6, 2)-BCSL. Para cada número de faixas

projetadas são capturados cinco conjuntos de imagens, contendo I_1, \bar{I}_1, I_2 e \bar{I}_2 , em tempos distintos e não consecutivos no vídeo. Cada conjunto leva à uma reconstrução individual e independente do plano.

A precisão da reconstrução (dado um número fixo de faixas projetadas) é calculada como o desvio padrão das distâncias dos cinco planos em relação ao plano médio. O **plano médio** é estimado a partir de todos os pontos obtidos pelo digitalizador, ou seja, pela união de todas as cinco nuvens de pontos. O plano médio minimiza o erro quadrático entre todos os pontos digitalizados usando um número fixo de faixas.

Para o cálculo da acurácia, utiliza-se um **plano base** de referência formado pelo plano médio da reconstrução com 50 faixas. Verificou-se manualmente a distância obtida pelo digitalizador para pontos marcados sobre padrão de calibração de dimensões $40cm \times 40cm$. Várias medidas foram realizadas com o auxílio de uma trena para validar distâncias e assim obter este plano base de referência. A acurácia da reconstrução é dada pela distância entre a equação deste plano base e o plano médio, dado um número fixo de faixas. Como procura-se explorar a projeção de diferentes números de faixas (6,2)-BCSL, é esperado que o erro eventualmente cresça a partir de um número de faixas limite; por isso, faz sentido usar como plano base aquele gerado por um número de faixas pequeno e seguro.

5 RESULTADOS E ANÁLISE QUANTITATIVA

Este capítulo apresenta o resultado da reconstrução de diversos objetos. Para mostrar a funcionalidade do digitalizador desenvolvido, objetos de diferentes tipos de materiais, e características geométricas distintas foram digitalizados; procurou-se utilizar um número diferente de faixas para cada objeto. Também é apresentado o resultado do estudo que relaciona número de faixas projetadas e resolução das imagens capturadas, realizado através do protocolo de erro desenvolvido.

5.1 Resultados

A Figura 5.1 demonstra o resultado da digitalização de um boneco de cachorro; sua superfície é branca, e portanto não impõe dificuldade para decodificação das faixas coloridas do código $(6, 2)$ -BCSL. Apesar disso, existem os detalhes pretos, como coleira e olhos. O tipo de material da coleira reflete as cores projetadas com intensidade suficiente para que a decodificação das bordas entre faixas seja possível. Contudo, o mesmo não ocorre com o olho; como consequência, não são gerados pontos tridimensionais para esta região.

Para este modelo, a resolução das imagens capturadas para reconstrução é de 1280×720 . Foram projetadas 150 faixas, e a nuvem resultante é composta por 33.298 pontos. A reconstituição geométrica captura as curvas suaves do corpo do cachorro, e também detalhes como as dobras na região da pata dianteira. As normais mostradas na Figura 5.1d são computadas através dos pontos 3D mais próximos.

A cena exibida pela Figura 5.2 foi capturada na resolução 1280×720 ; 170 faixas foram usadas para sua reconstrução. Há dois objetos com características distintas nesta cena.

A caixa é branca, com detalhes azuis e pretos. Novamente, os detalhes coloridos deste objeto não prejudicam a identificação de fronteiras e a extração de informação sobre sua geometria; isto pode ser notado pela preservação do ângulo reto entre as faces da caixa, como mostra a Figura 5.2c. A nuvem de pontos para caixa é formada por 8.586 pontos.

A garrafa é constituída de um material plástico, especular e preto. Apesar disso, uma

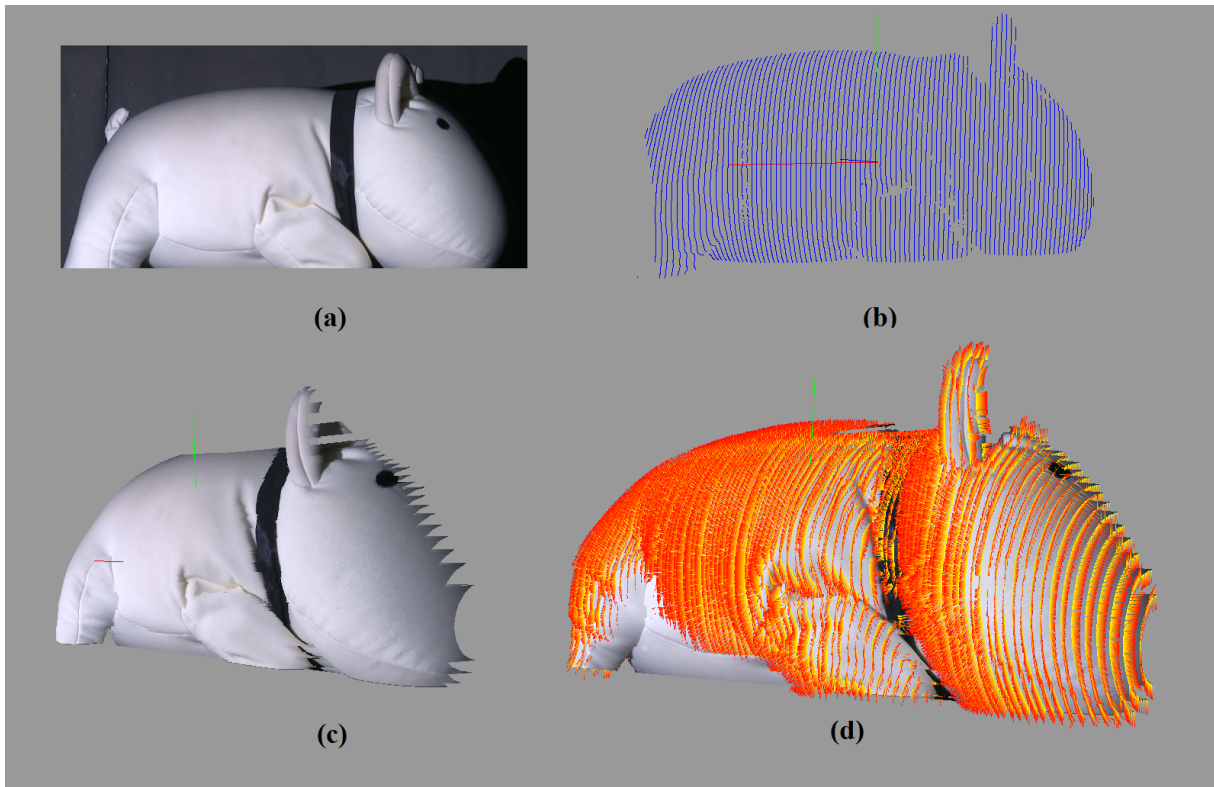


Figura 5.1: Digitalização do boneco de cachorro. (a) textura. (b) nuvem de 33.298 pontos gerada pelo digitalizador. (c) composição da nuvem de pontos com a textura. (d) normais inferidas através de pontos vizinhos.

nuvem de 4.614 pontos é reconstituída pelo digitalizador; isto foi possível pela proximidade entre garrafa e projetor, pela forte intensidade da luz projetada e pela detecção de fronteiras baseada em passagens por zero. A reconstrução da garrafa possibilita notar detalhes de seu relevo característico (Figura 5.2d).

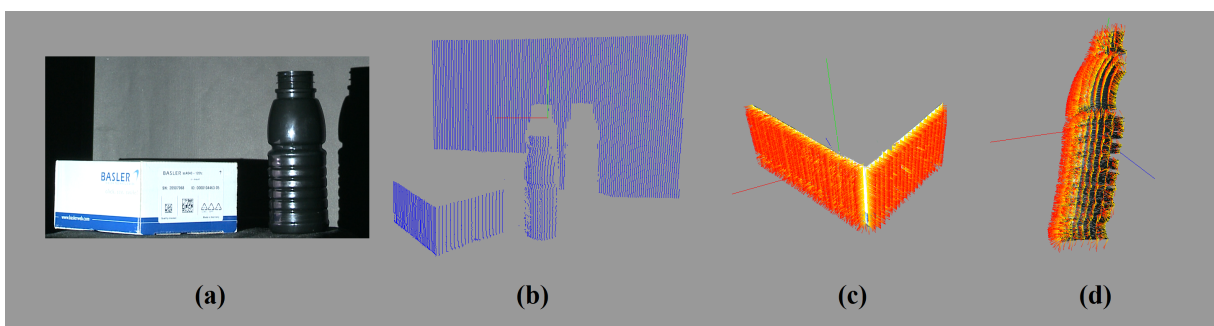


Figura 5.2: Caixa e garrafa. (a) cena capturada. (b) nuvem de 80.584 pontos gerada pelo digitalizador. (c) normais inferidas para caixa. (d) normais inferidas para garrafa.

A Figura 5.3 exemplifica a robustez do esquema de projeção de faixas complementares da codificação $(6,2)$ -BCSL. Os dois livros digitalizados apresentam geometria plana, simples, mas também superfícies com alta variação de cor; isto contudo não impede a

detecção de fronteiras entre as faixas projetadas e a geração de pontos tridimensionais à partir das mesmas (Figura 5.3b). As imagens capturadas para reconstrução do livro azul tem resolução de 1280x720; neste caso foram projetadas 200 faixas. A nuvem resultante é formada por 33.925 pontos. Para o livro laranja, a resolução de captura é 1920x1080; 220 faixas foram projetadas, e a nuvem resultante é composta por 66.456 pontos.

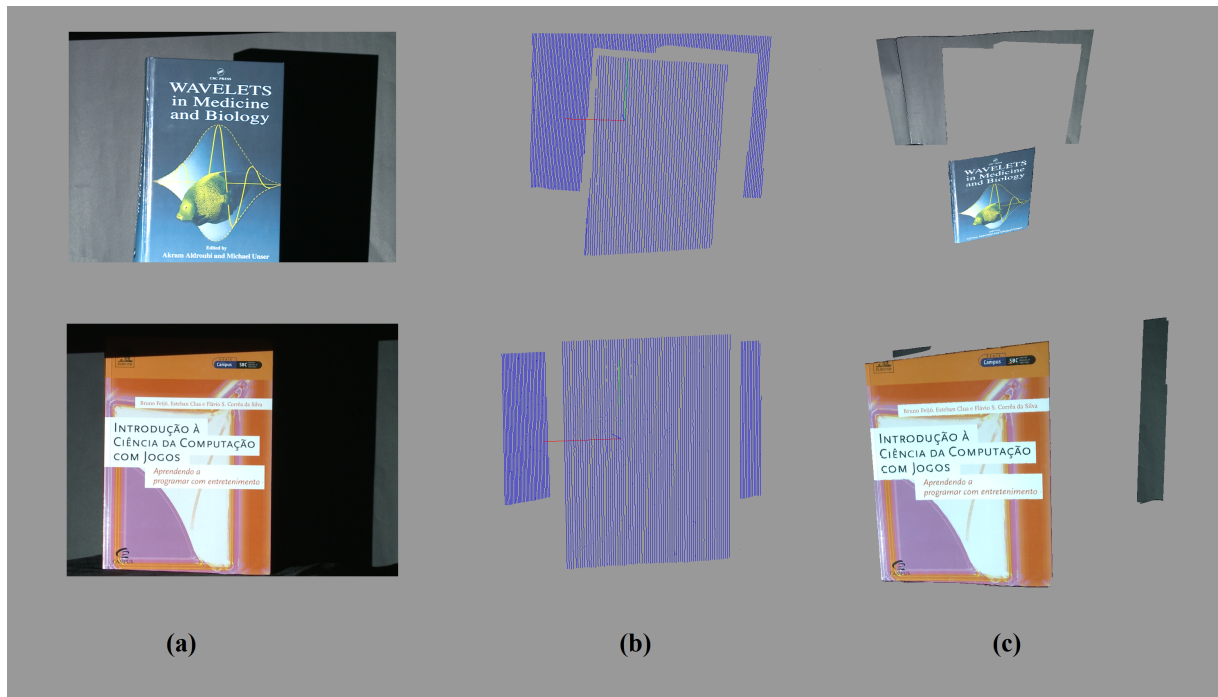


Figura 5.3: Digitalização de livros coloridos. (a) textura. (b) nuvem de pontos gerada pelo digitalizador. (c) composição da nuvem de pontos com a textura da cena.

A Figura 5.4 mostra três nuvens de pontos resultantes da digitalização de um rosto. Cada uma das nuvens foi gerada à partir da projeção de um número diferente de faixas: nas Figuras 5.4b e 5.4c foram utilizadas, respectivamente, 100 e 200 faixas; nestes casos a resolução das imagens capturadas para reconstrução é 1280x720. A nuvem do rosto gerada para projeção de 100 faixas tem 16.860 pontos, e a nuvem gerada para projeção de 200 faixas tem 29.128 pontos. Na Figura 5.4d, 250 faixas foram projetadas, a resolução de captura é 1920x1080, e a nuvem resultante tem 56.334 pontos. Através da comparação entre as nuvens de pontos é possível perceber que a projeção de um número maior de faixas resulta em uma captura geométrica mais detalhada; comparada às demais, a nuvem gerada para projeção de 250 faixas captura melhor o contorno do queixo, da boca, e detalhes da orelha.

Outra questão que pode ser notada na Figura 5.4 é que poucos pontos tridimensionais

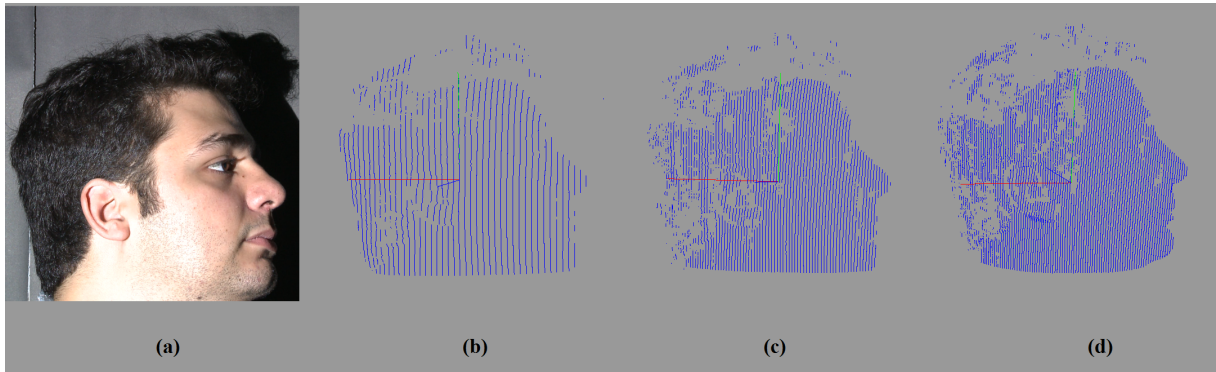


Figura 5.4: Digitalização de um rosto. (a) cena capturada. (b) nuvem de 16.860 pontos gerada para projeção de 100 faixas. (c) nuvem de 29.128 pontos gerada para projeção de 200 faixas. (d) nuvem de 56.334 pontos gerada para projeção de 250 faixas.

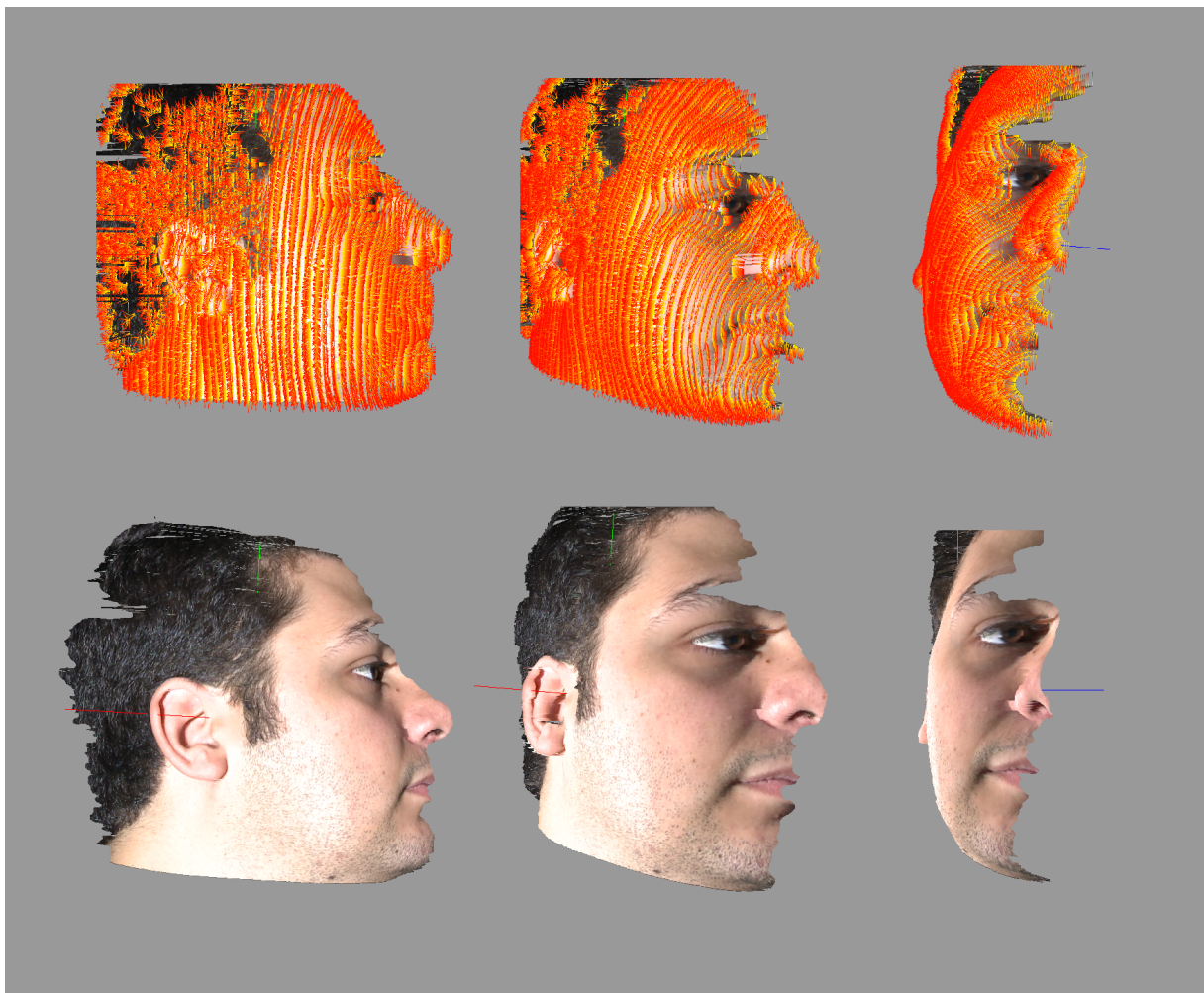


Figura 5.5: Rosto reconstruído pelo digitalizador através da projeção de 250 faixas.

foram gerados pelo digitalizador para região do cabelo; isto se deve à forma como os fios interagem com a luz projetada. Trata-se de uma limitação comum a todos os sistemas de reconstrução por luz ativa: a qualidade da geometria recuperada é sensível ao tipo de material que compõe a cena. Em especial, materiais translúcidos, especulares e que refletem a luz projetada com baixa intensidade são problemáticos.

A Figura 5.5 mostra a reconstrução do rosto conseguida com a projeção de 250 faixas. As normais inferidas permitem perceber a grande densidade de pontos da reconstrução.

5.2 Análise Quantitativa

Nesta seção é realizada a análise da relação entre número de faixas projetadas e resolução das imagens de captura para reconstrução geométrica; esta análise é feita através do protocolo de erro desenvolvido nesta dissertação. Como dito, o objetivo é estimar, para cada resolução, o número de faixas limite. Este número é aquele que provê reconstrução com grande densidade de pontos, e ao mesmo tempo possui baixo nível de erro.

A primeira sequência de medidas de acurácia e precisão foi tomada com o sistema câmera-projetor distanciado aproximadamente 1,50 metros do plano de digitalização. Neste caso, a posição e a orientação do plano são as mesmas empregadas pelo padrão plano xadrez na etapa de calibração do sistema. Variando o número de faixas projetadas sobre o plano, capturas de vídeos com duas configurações de câmera, 1280x720x60 (resolução de imagem 1280x720, taxa de captura de quadros de 60Hz) e 1920x1080x30 (resolução de imagem 1920x1080, taxa de captura de quadros de 30Hz), foram realizadas.

O resultado das medidas de erro de acurácia e precisão para a configuração 1280x720x60 pode ser observado na Figura 5.6.

As medidas de acurácia mostram erro baixo, com variações pequenas, para o intervalo entre 50 e 200 faixas. Para 250 faixas, entretanto, pode-se notar um salto significativo, em mais de uma ordem de grandeza, do erro de acurácia. Este aumento notável indica que o limite seguro de número de faixas para a configuração 1280x720x60 foi ultrapassado; para 300 faixas, o erro de acurácia sofre um novo salto qualitativo, o que significa que uma quantidade de faixas muito acima do limite está sendo empregada. Os resultados para acurácia também podem ser notados na comparação visual entre plano base e planos médios: enquanto os planos médios de 50, 100, 150 e 200 faixas (Figuras 5.7a, 5.7b, 5.7c e

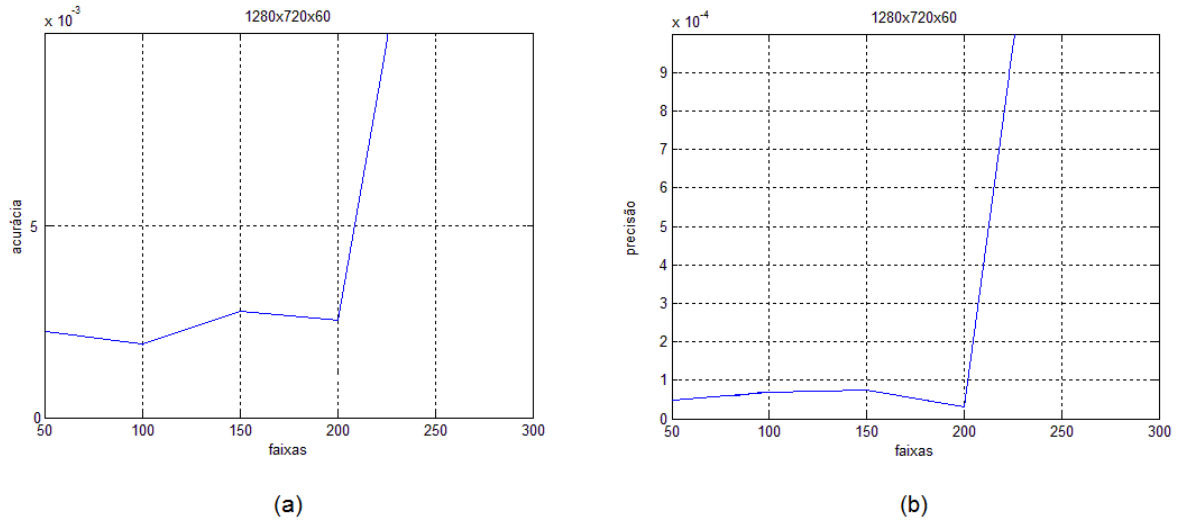


Figura 5.6: Gráficos de erro para acurácia e precisão; configuração de câmera 1280x720x60. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.

5.7d, respectivamente) são muito semelhantes ao plano base em inclinação e distância da origem, o plano médio de 250 faixas tem uma inclinação distinta (Figura 5.7e). O plano médio estimado para 300 faixas é completamente diferente do plano base (Figura 5.7f).

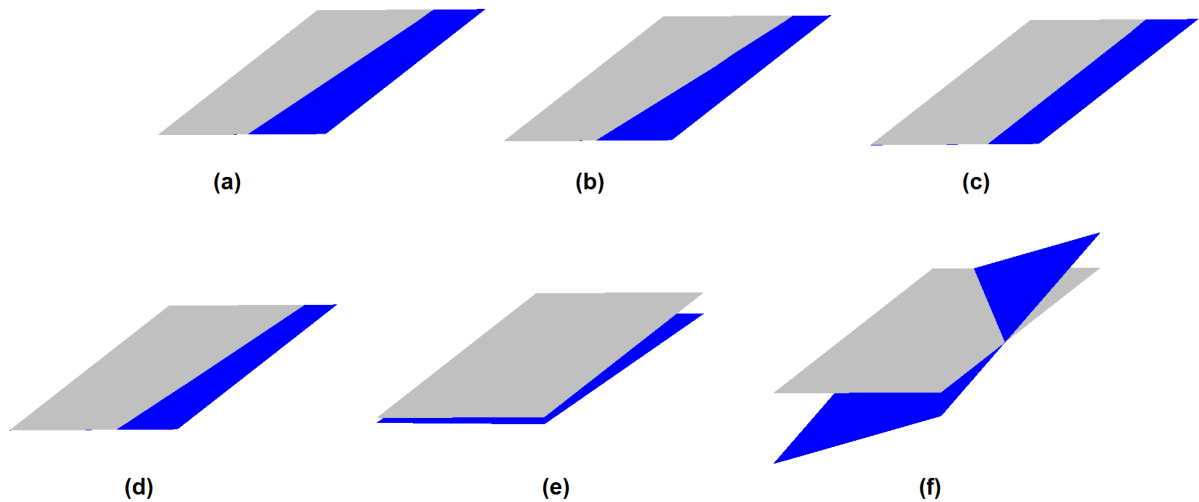


Figura 5.7: Plano base e planos médios obtidos para configuração de câmera 1280x720x60. Plano base em cinza, planos médios em azul. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas.

As medidas de precisão para configuração 1280x720x60 tem um comportamento semelhante às medidas de acurácia: oscilações pouco expressivas entre 50 e 200 faixas, e um salto significativo do erro para 250 faixas. Este resultado indica que reconstruções semelhantes do plano são repetidas para até 200 faixas.

A comparação visual entre planos médios e as reconstruções independentes do plano podem ser observadas na Figura 5.8: as reconstruções independentes do plano para projeção de até 200 faixas são similares ao plano médio em inclinação (nas Figuras 5.8a, 5.8b, 5.8c e 5.8d os planos estão praticamente sobrepostos); para 250 faixas é possível notar que os planos independentes possuem inclinações distintas à do plano médio (Figura 5.8e); esta situação se repete para projeção de 300 faixas (Figura 5.8f), onde a diferença entre inclinações é maior.

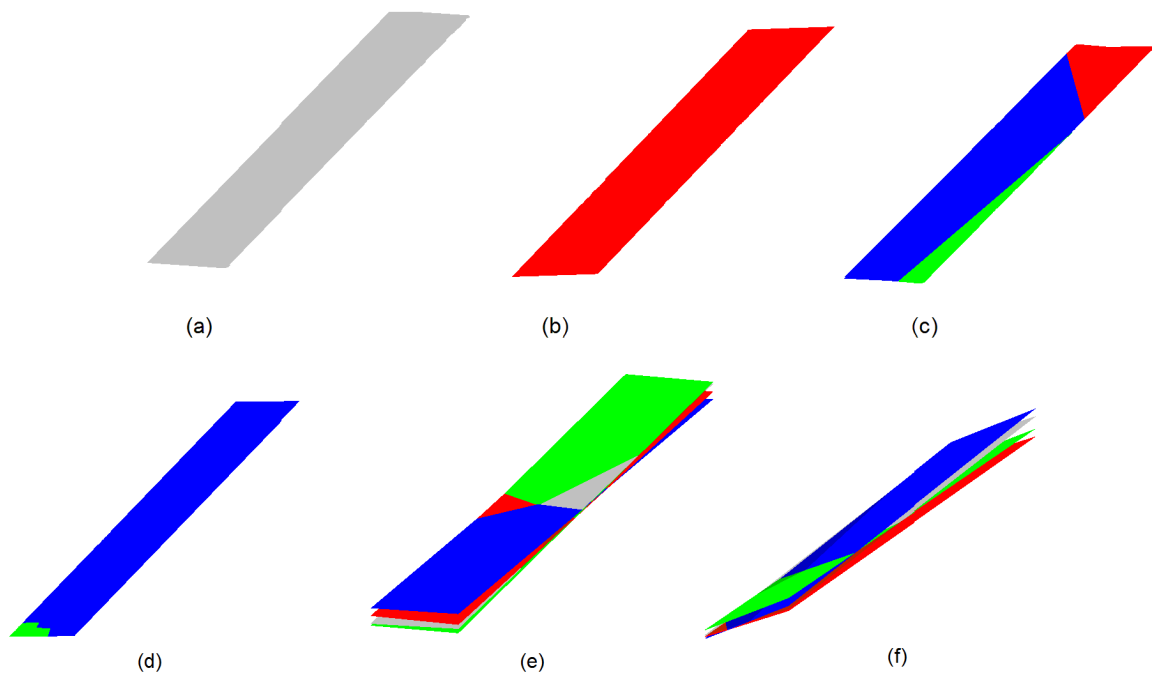


Figura 5.8: Comparação entre planos médios e reconstruções independentes do plano, para configuração de câmera 1280x720x60. Planos médios em cinza; planos azuis, vermelhos e verdes representam reconstruções independentes do plano. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas.

Os resultados das medidas de erro de acurácia e precisão para configuração 1280x720x60 mostram que reconstruções confiáveis, isto é, que possuem ao mesmo tempo bom nível de exatidão e também de repetibilidade, são obtidas para a projeção de até 200 faixas; isto também pode ser comprovado observando as nuvens de pontos reconstruídas (Figura 5.9). Para projeção de até 200 faixas (Figura 5.9d), os pontos tridimensionais gerados pelo digitalizador estão todos sobre o plano. Com 250 faixas projetadas (Figura 5.9e) ocorrem erros de decodificação; como consequência, alguns pontos são triangulados de maneira errônea, e recebem profundidades que estão à frente ou atrás do plano. Esta

situação também ocorre na projeção de 300 faixas, onde uma quantidade maior de pontos tem a coordenada z calculada de maneira errada.

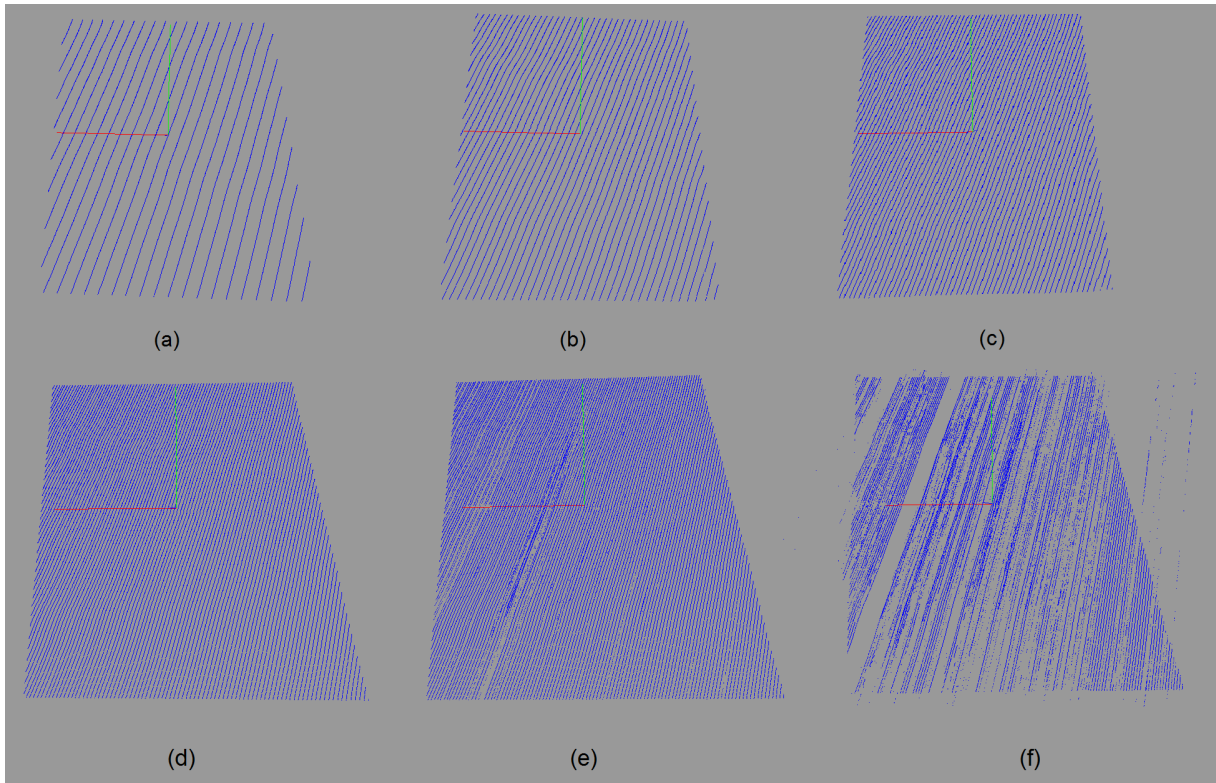


Figura 5.9: Nuvens de pontos geradas para digitalização do plano, para configuração de captura 1280x720x60. (a) 50 faixas. Nuvem de 13.694 pontos. (b) 100 faixas. Nuvem de 27.162 pontos. (c) 150 faixas. Nuvem de 40.750 pontos. (d) 200 faixas. Nuvem de 54.319 pontos. (e) 250 faixas. Nuvem de 67.905 pontos. (f) 300 faixas. Nuvem de 78.668 pontos.

As medidas de erro de acurácia e precisão para a configuração de câmera 1920x1080x30 podem ser observadas na Figura 5.10.

A curva que descreve o erro de acurácia para configuração 1920x1080x30 mostra comportamento estável, com leves variações, para até 300 faixas projetadas; para 350 faixas o erro sofre uma ascensão súbita em sua ordem de grandeza. A comparação entre plano base e planos médios (Figura 5.11) também descreve este comportamento: para até 300 faixas os planos médios estão sobrepostos ao plano base; os plano médios para projeção de 350 e 400 faixas apresentam inclinações e distâncias da origem distintas à do plano base.

A curva que caracteriza o erro de precisão para configuração 1920x1080x30 apresenta valores numéricos que estão na mesma ordem de grandeza para projeção de até 300 faixas; há oscilações, mas elas são pouco expressivas. Diferença significativa no erro de precisão surge na projeção de 350 faixas, onde o mesmo é amplificado em mais de uma ordem de

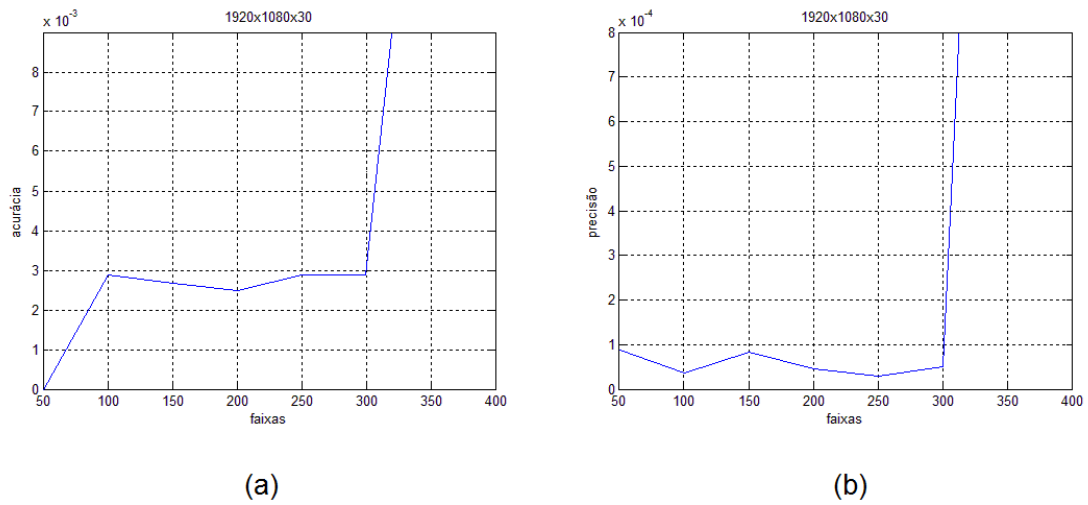


Figura 5.10: Gráficos de erro de acurácia e precisão para configuração de câmera 1920x1080x30. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.

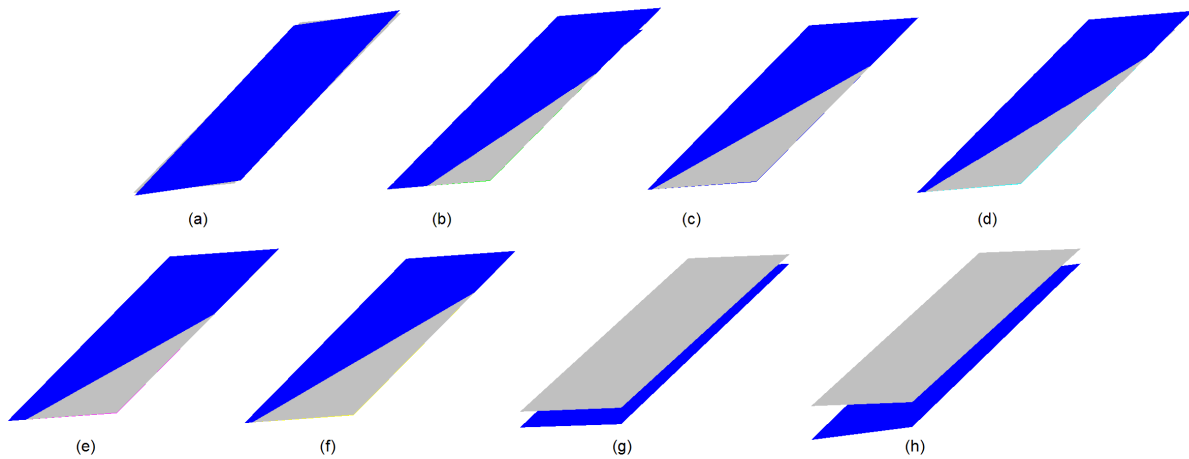


Figura 5.11: Plano base e planos médios para configuração de câmera 1920x1080x30. Em cinza, plano base; em azul, planos médios. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas. (g) 350 faixas. (h) 400 faixas.

grandeza. Este comportamento também pode ser observado através da diferença visual entre as reconstruções independentes do plano e os planos médios (Figura 5.12): para projeção de até 300 faixas os planos médios e os planos independentes possuem inclinações (isto é, vetor normal) muito similares; para 350 faixas é possível perceber inclinações diferentes.

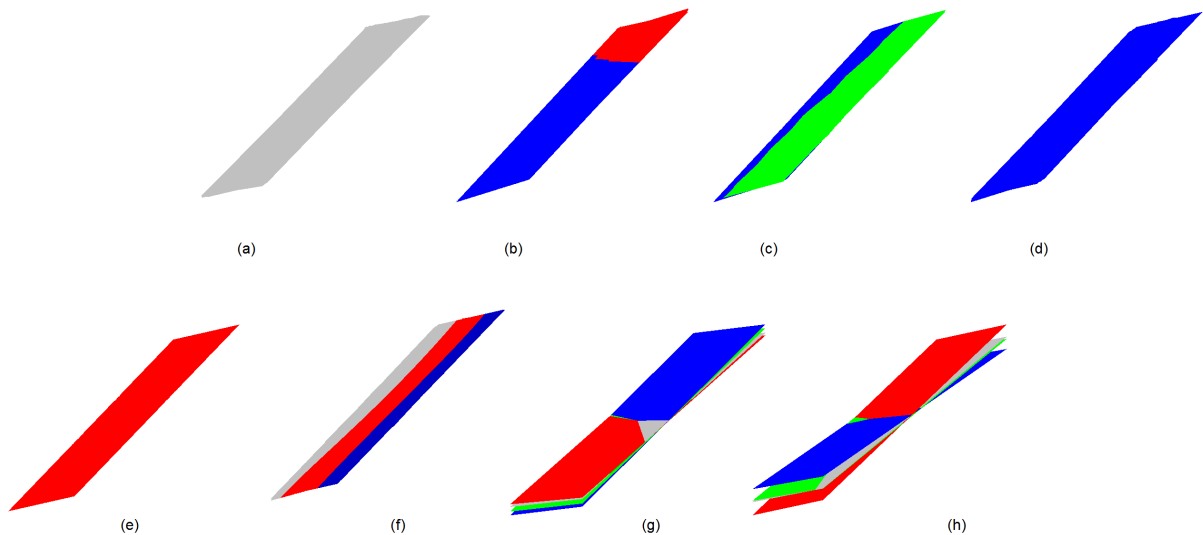


Figura 5.12: Comparação entre planos médios e reconstruções independentes do plano, para configuração de câmera 1920x1080x30. Em cinza, planos médios; em azul, verde e vermelho reconstruções independentes do plano. (a) 50 faixas. (b) 100 faixas. (c) 150 faixas. (d) 200 faixas. (e) 250 faixas. (f) 300 faixas. (g) 350 faixas. (h) 400 faixas.

As medidas de acurácia e precisão para configuração 1920x1080x30 indicam que reconstruções confiáveis, acuradas e precisas, para o plano são obtidas com a projeção de até 300 faixas. As nuvens de pontos geradas na digitalização do plano comprovam os resultados mostrados pelo protocolo; elas podem ser observadas na Figura 5.13. É possível notar que para a projeção de até 300 faixas todos os pontos tridimensionais gerados pelo digitalizador são triangulados corretamente, e estão sobre o plano. Erros de decodificação na resolução 1920x1080x30, que levam à cálculos incorretos de profundidade, começam a ocorrer com a projeção de 350 faixas. Isto confirma o limite estabelecido pelo protocolo, de 300 faixas.

Uma segunda sequência de medidas de erro para acurácia e precisão foi realizada com o plano de digitalização transladado cerca de 6cm para frente, e com uma diferença de inclinação de aproximadamente 15 graus em relação ao plano da primeira sequência de

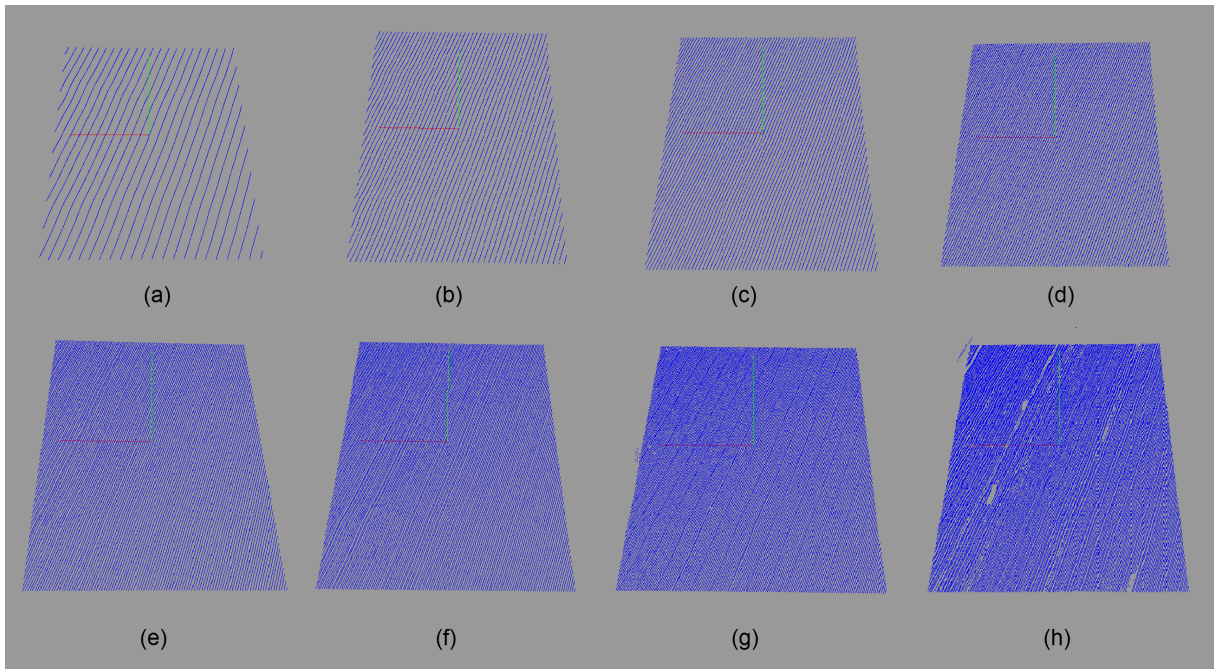


Figura 5.13: Nuvens de pontos geradas para digitalização do plano, para resolução de captura 1920x1080x30. (a) 50 faixas. Nuvem de 20.383 pontos. (b) 100 faixas. Nuvem de 40.635 pontos. (c) 150 faixas. Nuvem de 60.963 pontos. (d) 200 faixas. Nuvem de 81.275 pontos. (e) 250 faixas. Nuvem de 101.587 pontos. (f) 300 faixas. Nuvem de 121.871 pontos. (g) 350 faixas. Nuvem de 142.225 pontos. (h) 400 faixas. Nuvem de 156.776 pontos.

medidas (Figura 5.14). Isto foi feito para avaliar erros de acurácia e precisão fora do plano ideal de reconstrução; este plano é aquele utilizado na etapa de calibração do sistema, cuja posição e orientação foram reproduzidos na primeira sequência de medidas de erro. Para a segunda sequência de medidas, novamente, duas configurações de câmera foram testadas: 1280x720x60 e 1920x1080x30.

A comparação dos erros de acurácia e precisão entre a primeira e segunda sequência de medidas, para configuração de câmera 1280x720x60, pode ser observada na Figura 5.15.

De forma geral, para configuração de câmera 1280x720x60, a curva que descreve o erro de acurácia para a segunda sequência de medidas tem perfil semelhante à curva de acurácia da primeira sequência: baixo erro, com variações pequenas, para o intervalo entre 50 e 200 faixas, e um salto na ordem de grandeza para o erro de 250 faixas. A curva do erro de precisão para segunda sequência também é similar à curva de precisão da primeira sequência: baixo erro para até 200 faixas, e novamente um salto em ordem de grandeza para a projeção de 250 faixas.

A comparação das curvas de erro de acurácia entre sequência 1 e sequência 2, para

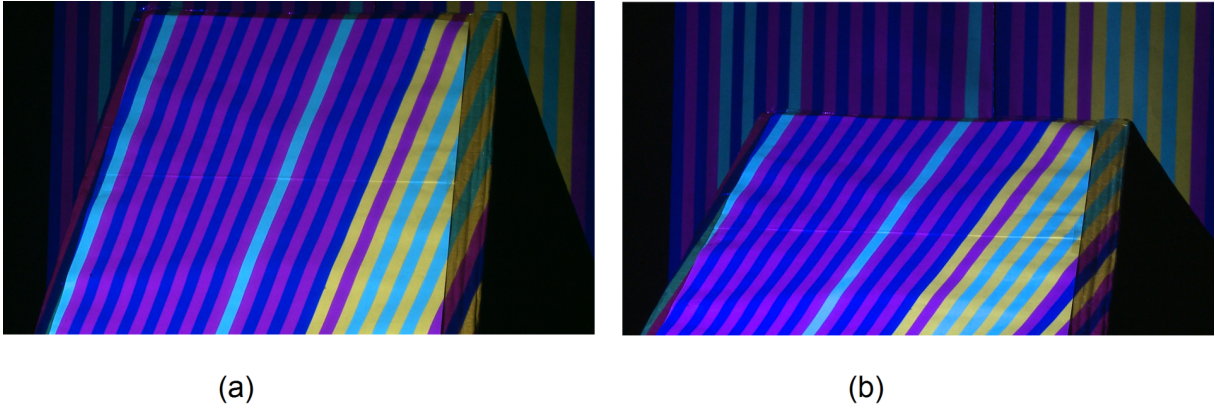


Figura 5.14: Posicionamento e orientação dos planos para digitalização. (a) sequência 1. (b) sequência 2.

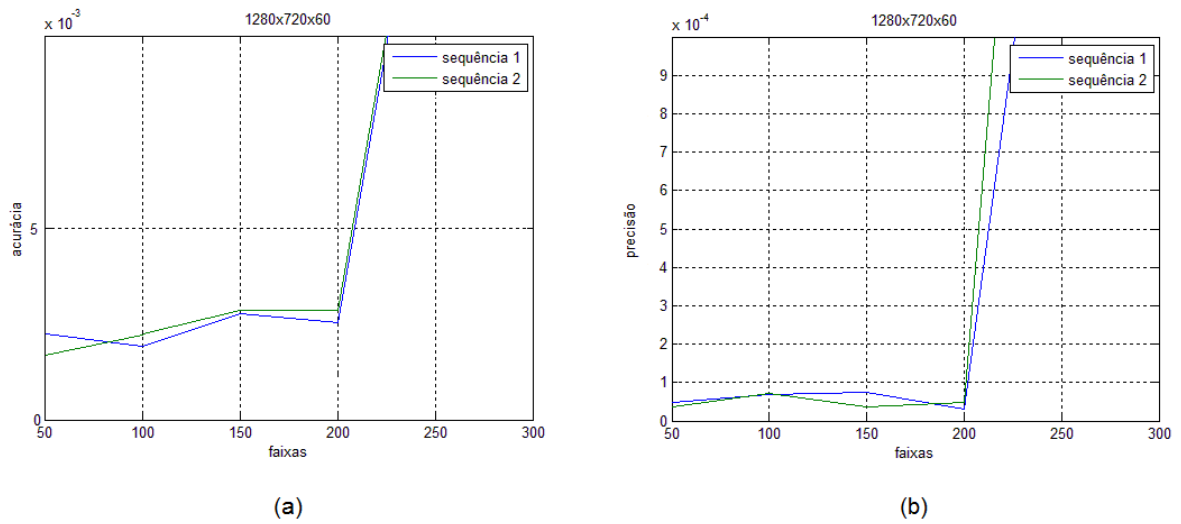


Figura 5.15: Comparação do erro entre primeira e segunda sequência de medidas, para configuração de câmera 1280x720x60. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.

configuração de câmera 1280x720x60, mostra que o erro da sequência 1, para maior parte do número de faixas projetadas, está abaixo do erro da sequência 2; uma possível explicação para este comportamento reside na etapa de calibração do sistema: é conhecido que o método de calibração de Tsai, utilizado neste trabalho, é acurado para pontos que estejam próximos ao plano de calibração. O plano digitalizado na segunda sequência de medidas, pelo deslocamento e inclinação, está fora do posicionamento do plano de calibração, ou seja, fora do plano ideal de reconstrução.

A comparação dos erros de precisão entre sequência 1 e sequência 2 mostra oscilações entre as duas curvas: os erros para sequência 1 não são sempre menores do que os erros da sequência 2. Isto sugere que a inclinação do plano na sequência 2 não afeta erros de precisão de maneira significativa.

De forma geral, dados os resultados de acurácia e precisão para sequência 1 e sequência 2, 200 faixas é o limite sugerido de confiabilidade que permite a repetição de reconstruções acuradas e precisas para configuração de câmera 1280x720x60.

A comparação dos erros de acurácia e precisão entre a primeira e segunda sequência de medidas, para configuração de câmera 1920x1080x30, pode ser observada na Figura 5.16.

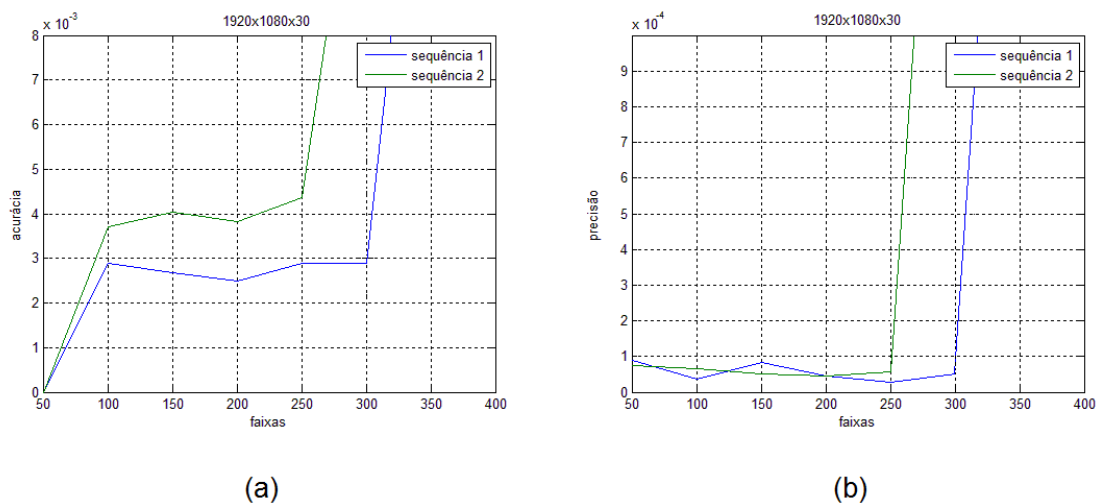


Figura 5.16: Comparação do erro entre primeira e segunda sequência de medidas, para configuração de câmera 1920x1080x30. (a) acurácia adimensional na detecção de planos. (b) precisão adimensional na detecção de planos.

Comparando as curvas de erro para acurácia de sequência 1 e sequência 2, para configuração de câmera 1920x1080x30, pode-se notar que a curva para sequência 1 está sempre abaixo da curva da sequência 2. Como discutido, uma possível razão para este compor-

tamento está no deslocamento e inclinação do plano na sequência 2, que o coloca fora do plano ideal de reconstrução. É possível notar também que o erro de acurácia da sequência 2, para projeção de 300 faixas, cresce subitamente em ordem de grandeza; na sequência 1 este crescimento é observado com a projeção de 350 faixas. Isto pode ser explicado pelo deslocamento aplicado ao plano da sequência 2: com a maior proximidade do projetor, as faixas que atingem o plano tem menor espessura, e por isso erros de decodificação aparecem antecipadamente.

A comparação das curvas de erro de precisão entre as duas sequências mostra oscilações para projeção de até 250 faixas: diferente do que ocorre com as medidas de erro para acurácia, os erros de precisão da sequência 1 não são sempre menores que os erros da sequência 2. É possível perceber, entretanto, que o erro de precisão da sequência 2 cresce em ordem de grandeza com a projeção de 300 faixas; isto não ocorre para sequência 1. Novamente, a explicação para este comportamento reside no deslocamento aplicado ao plano. Erros de decodificação acontecem mais cedo, por conta da maior proximidade do projetor, e conseqüentemente, menor largura das faixas que atingem o plano.

Então, para configuração de câmera 1920x1080x30, dados os resultados de acurácia e precisão da sequência 2, é mais seguro utilizar a projeção de 250 faixas para captura geométrica; a razão para isto é que raramente um objeto de digitalização estará posicionado exatamente sobre o plano ideal de reconstrução.

5.2.1 *Experimento com Barreira de Vidro*

Durante o desenvolvimento do digitalizador foi realizado um experimento de digitalização através de uma barreira de vidro. O objetivo é explorar a possibilidade de reconstituição de geometria em condições adversas para a estereoscopia ativa baseada na projeção de luz estruturada.

O experimento foi conduzido com uma placa de vidro de dimensões $1m \times 1m$ e $8mm$ de espessura, distanciada aproximadamente $50cm$ do sistema câmera-projetor e $80cm$ do plano de digitalização. Foram projetadas 150 faixas sobre o plano, e a configuração de captura da câmera foi de 1280x720x60. Em um segundo teste de reconstrução, utilizando as mesmas 150 faixas e configuração de captura 1280x720x60, foi aplicada uma rotação de 15 graus (sentido da câmera) sobre a placa de vidro; no terceiro e último teste a rotação aplicada sobre a placa foi de 30 graus, mas no sentido inverso (sentido do projetor).

A razão para aplicar diferentes rotações sobre a placa de vidro está relacionada ao fenômeno de refração. A direção dos raios de luz é alterada quando os mesmos entram em contato com a superfície de vidro; esta mudança de direção depende do ângulo de incidência formado entre os raios e a placa de vidro. Então, diferentes rotações da placa resultam em variações de direção distintas dos raios de luz; obviamente, mudanças de direção dos raios de luz podem gerar distorções na imagem formada nos sensores da câmera, e como consequência, afetar a reconstrução de geometria.

As medidas de acurácia e precisão para as reconstruções do plano através da barreira de vidro foram geradas com protocolo de erro desenvolvido nesta dissertação. Como o intuito do experimento é verificar a influência da placa de vidro sobre o erro de reconstrução geométrica, para medida de acurácia foi utilizado um plano base obtido sem barreira de vidro entre digitalizador e objeto digitalizado. Este plano base é o plano médio para projeção de 50 faixas.

As medida de erro para as três inclinações da placa de vidro podem ser observadas na Tabela 5.1:

rotação da placa de vidro	acurácia	precisão
0 grau	1.6×10^{-3}	1.3×10^{-4}
15 graus	4.2×10^{-1}	2.3×10^{-1}
30 graus	1.8×10^{-2}	1.1×10^{-4}

Tabela 5.1: Acurácia e precisão para digitalização do plano através da barreira de vidro.

Para rotação de 0 grau, a medida de acurácia mostra, surpreendentemente, baixo erro. De fato, a ordem de grandeza deste erro é a mesma da projeção de 150 faixas sem barreira de vidro (Figura 5.6a). A medida de precisão também indica baixo erro, que indica alta confiabilidade.

Para rotação de 15 graus é possível notar um salto de duas ordens de grandeza no erro de acurácia. Este comportamento pode ser explicado pela refração da luz que ocorre no vidro; na Figura 5.17 as passagens por zero estão marcadas, em branco, sobre uma imagem I_1 utilizada para obtenção do plano base; pode-se notar que para rotação de 15 graus (Figura 5.17b) as passagens estão deslocadas da posição ideal da transição entre faixas. Como evidenciado pela Equação 3.7, o deslocamento das coordenadas dos pontos de passagem por zero na imagem da câmera afeta o cálculo de profundidade. Além deste deslocamento, também foi observado que, em algumas reconstruções, alguns pontos foram

triangulados de forma incorreta, e receberam profundidades erradas, distantes do plano digitalizado. Este comportamento explica o grande salto no erro de precisão.

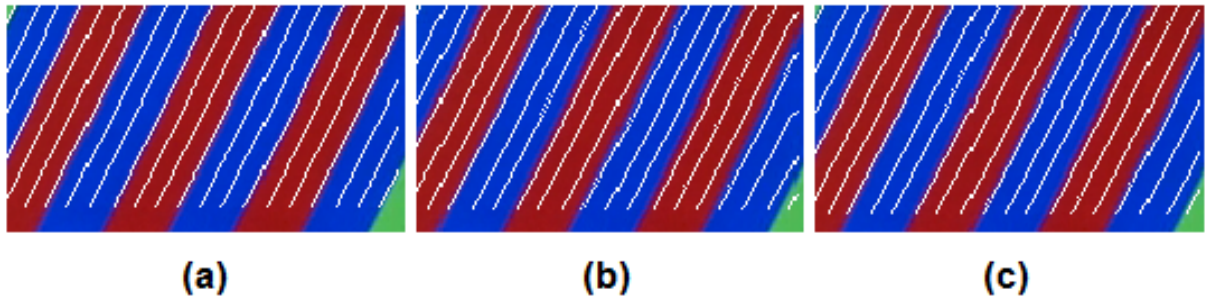


Figura 5.17: Passagens por zero marcadas, em branco, sobre a imagem I_1 do plano base. (a) barreira de vidro sem rotação. (b) barreira de vidro rotacionada 15 graus no sentido da câmera. (c) barreira de vidro rotacionada 30 graus no sentido do projetor.

Para rotação de 30 graus também foi verificado o deslocamento nas coordenadas de passagem por zero (Figura 5.17c). Como consequência, o erro de acurácia, quando comparado à rotação de 0 grau, é amplificado. O erro de precisão é baixo, está na mesma ordem de grandeza da rotação de 0 grau. Isto indica que reconstruções semelhantes do plano são repetidas. Contudo, todas as reconstruções são afetadas pelo deslocamento das passagens por zero.

A comparação visual entre plano base e planos médios confirma os resultados numéricos gerados pelo protocolo. Enquanto o plano médio para rotação de 0 grau está sobreposto ao plano base (Figura 5.18a), os planos médios para as rotações de 15 e 30 graus (Figuras 5.18b e 5.18c, respectivamente) são diferentes do plano base em distância da origem e inclinação.

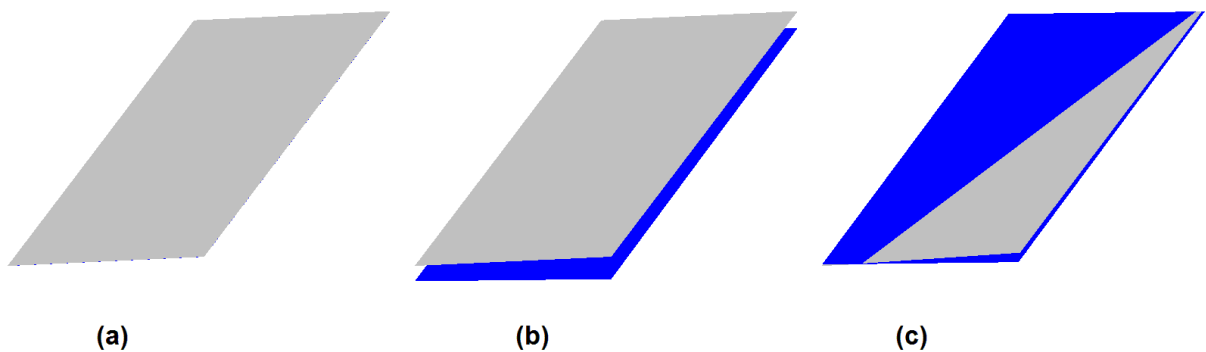


Figura 5.18: Plano base e planos médios para medida de acurácia. Em cinza, plano base; em azul, planos médios. (a) rotação de 0 grau da barreira de vidro. (b) rotação de 15 graus da barreira de vidro. (c) rotação de 30 graus da barreira de vidro.

Os erros de acurácia e precisão para as reconstruções do plano através da barreira de vidro mostram que, de forma geral, a reconstituição geométrica nestas condições é possível, mas a mesma é afetada pelo fenômeno de refração. Foi possível observar que a refração causa deslocamento nas coordenadas dos pontos de passagem por zero; o nível de deslocamento está relacionado a rotação aplicada sobre placa de vidro, posicionada entre digitalizador e objeto de digitalização. De maneira surpreendente, este deslocamento não foi observado para rotação de 0 grau, que obteve medidas de erro comparáveis à reconstrução sem barreira. As rotações de 15 e 30 graus, todavia, apresentam reconstruções com distorções; estas distorções não impossibilitaram a recuperação da forma do plano, mas obviamente afetaram erros de acurácia e precisão. Então, aplicações do digitalizador que exijam reconstruções acuradas não são possíveis com rotações de 15 e 30 graus.

6 CONCLUSÃO

Nesta dissertação foi desenvolvido um digitalizador câmera-projetor de alta definição (resoluções de 1920x1080 e 1280x720), baseado na projeção de luz estruturada, para a captura de geometria e fotometria.

Foram estudados e descritos os fundamentos matemáticos do processo de calibração câmera-projetor, responsável por estabelecer a transformação que relaciona os sistemas de coordenadas da câmera e do projetor. Foi apresentado o código de luz estruturada (b, s) -BCSL, baseado na projeção de faixas verticais coloridas, como forma de estabelecer pontos correspondentes para os sistemas de coordenadas dos dois dispositivos (câmera e projetor). Também foi apresentada uma solução para o problema de calibração de cores, necessária para aproximar uma cor projetada à uma cor capturada pela câmera, e assim assegurar a identificação correta das cores projetadas pelo código (b, s) -BCSL.

Foram discutidos e apresentados os procedimentos de configuração do sistema de digitalização: disposição e posicionamento de equipamentos, controle da iluminação ambiente e captura das imagens para calibração câmera-projetor e calibração de cores.

Também foi descrita a *pipeline* de processamento das imagens capturadas iluminadas pelas faixas do código (b, s) -BCSL. O ponto chave deste processamento é a localização das fronteiras entre faixas, com acurácia subpixel, que é baseada na descoberta dos pontos de passagens por zero nos canais de cor R, G e B das imagens capturadas e processadas. As fronteiras encontradas estabelecem pontos de correspondência entre câmera e projetor. Estes pontos passam pelo procedimento de triangulação, que calcula a coordenada z , ou seja, a profundidade, em relação à câmera. Os pontos triangulados formam uma nuvem que descreve a geometria da cena capturada e iluminada pelo padrão de cores (b, s) -BCSL.

Para demonstrar a funcionalidade do digitalizador desenvolvido, objetos de diferentes tipos de materiais foram reconstruídos. Os resultados destas digitalizações mostram que o digitalizador é capaz de reconstruir superfícies com alta variação de cores, e surpreendentemente até mesmo alguns objetos constituídos de materiais especulares e pretos, como ocorreu com a digitalização da garrafa (Figura 5.2). Contudo, obviamente, há limites para sistemas de reconstrução por luz ativa. Regiões de sombra, que não são atingidas pelo padrão de luz projetado, e certos tipos de materiais que refletem a luz projetada com

baixa intensidade causam imperfeições nas reconstruções. Neste trabalho, isto pode ser observado na digitalização do rosto: as regiões de sobancelha, cabelo e sombra causada pelo cílios apresentam falhas, isto é, são gerados poucos pontos tridimensionais para estas regiões.

A digitalização do rosto também serve como exemplo para demonstrar o efeito que o aumento de número de faixas e de resolução de imagem causa no nível de detalhe geométrico recuperado para um objeto digitalizado; claramente, tanto o aumento de resolução da imagem quanto do número de faixas projetadas promove a captura mais detalhada de geometria, e uma nuvem com maior quantidade de pontos. A nuvem do rosto gerada para projeção de 250 faixas, através de imagens de captura na resolução 1920x1080 tem 56.334 pontos, enquanto projeção de 100 faixas e captura na resolução 1280x720 gerou uma nuvem de 16.860 pontos. O número de pontos aproximadamente três vezes maior da primeira nuvem é explicado não só pelo aumento do número de faixas projetadas, mas também pelo maior número de linhas na imagem de 1920x1080.

O digitalizador também mostrou ser capaz de capturar características geométricas distintas, como as curvas suaves de um rosto, ou os ângulos retos de uma caixa.

Neste trabalho foi desenvolvido e proposto um protocolo de avaliação de erro. O mesmo é baseado em uma métrica de distância entre planos no espaço projetivo, e foi utilizado para avaliar a acurácia e a precisão de reconstrução de diversos números de faixas (b, s) -BCSL projetadas. Isto foi feito com o objetivo de estimar o número limite de faixas possível dentro das resoluções atuais de vídeo de alta definição. Este número é importante, pois identifica a quantidade de faixas para projeção que provê reconstrução de grande densidade de pontos da geometria alvo, e ao mesmo tempo possui baixo nível de erro.

Os resultados experimentais realizados com sistema câmera-projetor distanciado cerca de 1,50m do plano de digitalização mostram que para configuração de câmera 1280x720x60 (resolução de imagem 1280x720, taxa de captura de quadros de 60Hz), 200 faixas é o limite seguro sugerido que permite reconstruções acuradas e precisas. Para configuração de câmera 1920x1080x30 (resolução de imagem 1920x1080, taxa de captura de quadros de 30Hz) verificou-se que este limite é de 250 faixas. Estes resultados comprovam que imagens de alta resolução suportam um grande número de faixas projetadas. Como comparação, o sistema de geração de vídeo 3D desenvolvido por Vieira [29], que também emprega

codificação (6, 2)-BCSL, e captura de imagens na resolução 640x480, relata a projeção de 70 faixas para iluminação da cena.

Por fim, os resultados gerados pelo protocolo também mostram que, para maior parte do número de faixas projetadas, reconstruções do plano para sequência 2, em que o plano de digitalização está fora do plano ideal de reconstrução, tem erro de acurácia maior do que os planos da sequência 1, posicionados sobre o plano ideal. Apesar disso, a diferença de erro entre as duas sequências é pequena; a diferença notável observada neste sentido ocorreu na projeção de 300 faixas para resolução de imagem 1920x1080. Neste caso foi possível observar que os erros de precisão e acurácia da sequência 2 sofrem um salto qualitativo em ordem de grandeza, para pior, em comparação à sequência 1. Isto pode ser explicado pelo deslocamento aplicado ao plano na sequência 2. Com a maior proximidade do projetor, as faixas que atingem o plano tem menor espessura, e por isso erros de decodificação aparecem antecipadamente.

Existem pelo menos duas possibilidades para extensão do presente trabalho.

A primeira possibilidade é adaptar o digitalizador desenvolvido para geração de vídeo tridimensional em alta definição e tempo real. Como dito, câmera e projetor possuem as características necessárias para isto. A câmera, Sony PMW-EX3, possui saída IEEE 1394-A para transferência de vídeo em tempo real e entrada *genlock* para sincronização com um sinal externo. O projetor, Benq W1000, tem baixa latência na troca de cores, por ser um modelo com processamento digital de luz (em inglês, DLP). Contudo, como discutido, a geração vídeo em tempo real requer sincronia entre câmera e projetor. Para isto, é necessária a aquisição de um sincronizador, que garanta que captura e projeção se iniciem ao mesmo tempo, e que a projeção de um *slide* de cores corresponda à captura de uma imagem pela câmera.

Uma segunda possibilidade é estender a funcionalidade do digitalizador desenvolvido para captura de modelos tridimensionais completos de objetos ou cenas. Isto pode ser feito através da implementação do algoritmo *Iterative Closest Points* (ICP) [30]. Este é um algoritmo comumente empregado para integrar nuvens de pontos tridimensionais. Desta maneira, é possível gerar diversas nuvens de pontos para um mesmo objeto, cada uma de diferente perspectiva, de forma que todas as faces do objeto sejam capturadas. A posterior integração de todas as nuvens gera um modelo tridimensional completo do objeto. Uma abordagem similar a esta é usada no trabalho de Rusinkiewicz[13].

REFERÊNCIAS

- [1] CODA, C., *Um Sistema de Calibracao de Camera*, Master's Thesis, Universidade Federal de Alagoas, 2007.
- [2] BOUGUET, J. Y., "Stripe Extraction, Correspondence and Error Analysis", California Institute of Technology, 1998, class notes.
- [3] SA, A., CARVALHO, P. C., VELHO, L., "(b, s)-BCSL : Structured Light Color Boundary Coding for 3D Photography". In: *Proceedings of 7th International Fall Workshop on VISION, MODELING, AND VISUALIZATION*, 2002.
- [4] SEITZ, S. M., CURLESS, B., DIEBEL, J., SCHARSTEIN, D., SZELISKI, R., "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms". In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*, pp. 519–528, IEEE Computer Society: Washington, DC, USA, 2006.
- [5] TSAI, R., "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses", *Robotics and Automation, IEEE Journal of*, v. 3, n. 4, pp. 323–344, 1987.
- [6] ZHANG, Z., "A Flexible New Technique for Camera Calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, pp. 1330–1334, 2000.
- [7] JI, Z., LEU, M. C., "Design of optical triangulation devices", *Optics and Laser Technology*, v. 21, n. 5, pp. 339 – 341, 1989.
- [8] KEFERSTEIN, C. P., MARXER, M., "Testing bench for laser triangulation sensors", *Sensor Review*, v. 18, pp. 183–187, 1998.
- [9] BOUGUET, J. Y., PERONA, P., "3D Photography Using Shadows in Dual-Space Geometry", *Int. J. Comput. Vision*, v. 35, pp. 129–149, November 1999.
- [10] VARANASI, K., *Spatio-Temporal Modeling of Dynamic 3D Scenes from Visual Data*, Ph.D. Thesis, University of Grenoble, 2010.

- [11] SALVI, J., FERNANDEZ, S., PRIBANIC, T., LLADO, X., “A state of the art in structured light patterns for surface profilometry”, *Pattern Recogn.*, v. 43, pp. 2666–2680, August 2010.
- [12] HIROSHI KAWASAKI, H., FURUKAWA, R., SAGAWA, R., YAGI, Y., “Dynamic scene shape reconstruction using a single structured light pattern”. In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, june 2008.
- [13] RUSINKIEWICZ, S., HALL-HOLT, O., LEVOY, M., “Real-Time 3D Model Acquisition”, *ACM Transactions on Graphics (Proc. SIGGRAPH)*, v. 21, n. 3, pp. 438–446, July 2002.
- [14] ISHII, I., YAMAMOTO, K., DOI, K., TSUJI, T., “High-speed 3D image acquisition using coded structured light projection”. In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 925–930, 2007.
- [15] PRIBANIC, T., DZAPO, H., SALVI, J., “Efficient and Low-Cost 3D Structured Light System Based on a Modified Number-Theoretic Approach”, *EURASIP Journal on Advances in Signal Processing*, 2010.
- [16] ZHANG, S., ROYER, D., YAU, S.-T., “High-resolution, real-time-geometry video acquisition system”. In: *ACM SIGGRAPH 2006 Emerging technologies, SIGGRAPH '06*, ACM: New York, NY, USA, 2006.
- [17] FECHTELER, P., EISERT, P., RURAINSKY, J., “Fast and High Resolution 3D Face Scanning”. In: *Proceedings of the 14th International Conference on Image Processing (ICIP2007)*, San Antonio, Texas, USA, 16-19th September 2007.
- [18] GUHRING, J., “Dense 3-D surface acquisition by structured light using off-the-shelf components”. In: *Proc. Videometrics and Optical Methods for 3D Shape Measurement*, pp. 220–231, 2001.
- [19] AKCA, D., GRUEN, A., BREUCKMANN, B., LAHANIER, C., “High Definition 3-D Scanning of Arts Objects and Paintings”. In: *Proceedings of Optical 3-D Measurement Techniques VIII*, v. 2, pp. 50–58, Zurich, Switerland, July 2007.

- [20] MCPHERRON, S. P., GERNAT, T., HUBLIN, J.-J., “Structured light scanning for high-resolution documentation of in situ archaeological finds”, *Journal of Archaeological Science*, v. 36, n. 1, pp. 19 – 24, 2009.
- [21] ZHANG, S., *High-resolution, Real-time 3-D Shape Measurement*, Ph.D. Thesis, Stony Brook University, 2005.
- [22] TROBINA, M., *Error Model of a Coded-Light Range Sensor*, Tech. rep., Communication Technology Laboratory - Image Science Group, September 1995.
- [23] FENG, X., CAO, M., WANG, H., COLLIER, M., “The Comparison of Camera Calibration Methods Based on Structured-Light Measurement”. In: *Proceedings of the 2008 Congress on Image and Signal Processing, Vol. 2 - Volume 02*, pp. 155–160, IEEE Computer Society: Washington, DC, USA, 2008.
- [24] ABDEL-AZIZ, Y., KARARA, H. M., “Direct Linear Transformation into object space coordinates in close-range photogrammetry”. In: *Proceedings of Symp. Close-Range Photogrammetry*, pp. 1–18, 1971.
- [25] LEVENBERG, K., “A method for the solution of certain nonlinear problems in least squares”, *Quart. Appl. Math.*, v. 2, pp. 164–168, 1944.
- [26] MARQUADT, D., “An algorithm for the least-squares estimation of nonlinear parameters”, *SIAM J. Appl. Math.*, v. 11, pp. 431–441, 1963.
- [27] VELHO, L., CARVALHO, P., SOARES, E., SA, A., MONTENEGRO, A., PEIXOTO, A., RIVERA, L. A., *Fotografia 3D. 25 Coloquio Brasileiro de Matematica - IMPA: Rio de Janeiro*, 2005.
- [28] BOUGUET, J. Y., “Camera Calibration Toolbox for Matlab”, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html, accessed April 2, 2011.
- [29] VIEIRA, M. B., VELHO, L., SA, A., CARVALHO, P. C., “A Camera-Projector System for Real-Time 3D Video”. In: *Proceedings of PROCAMS - IEEE International Workshop on Projector-Camera Systems*, 2005.
- [30] BESL, P., MCKAY, N., “A Method for Registration of 3-D Shapes”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 14, pp. 239–256, 1992.