

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA MECÂNICA**

**ALINHAMENTO DE IMAGENS DE PROFUNDIDADE NA  
RECONSTRUÇÃO 3D DE OBJETOS DE FORMA LIVRE**

**LANDECIR ALVES DE ALBUQUERQUE**

**ORIENTADOR: JOSÉ MAURÍCIO SANTOS TORRES DA MOTTA**

**DISSERTAÇÃO DE MESTRADO  
EM SISTEMAS MECATRÔNICOS**

**PUBLICAÇÃO: ENM.DM – 09 A/06  
BRASÍLIA/DF: AGOSTO - 2006**

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA MECÂNICA**

**ALINHAMENTO DE IMAGENS DE PROFUNDIDADE NA  
RECONSTRUÇÃO 3D DE OBJETOS DE FORMA LIVRE**

**LANDECIR ALVES DE ALBUQUERQUE**

**DISSERTAÇÃO DE MESTRADO SUBMETIDA AO  
DEPARTAMENTO DE ENGENHARIA MECÂNICA DA  
FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE  
BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA  
A OBTENÇÃO DO GRAU DE MESTRE EM SISTEMAS  
MECATRÔNICOS.**

**APROVADA POR:**

---

**JOSÉ MAURÍCIO SANTOS TORRES DA MOTTA, PhD, (ENM/UnB)  
(Orientador)**

---

**CARLOS HUMBERTO LLANOS QUINTERO, Doutor, (ENM/UnB)  
(Examinador Interno)**

---

**DÍBIO LEANDRO BORGES, PhD, (CIC/UnB)  
(Examinador Externo)**

**BRASÍLIA/DF, 9 de AGOSTO de 2006.**

## FICHA CATALOGRÁFICA

ALBUQUERQUE, LANDECIR ALVES DE	
Alinhamento de Imagens de Profundidade na Reconstrução 3D de Objetos de Forma Livre [Distrito Federal] 2006.	
xviii, 125p., 297 mm (ENM/FT/UnB, Mestre, Sistemas Mecatrônicos, 2006). Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.	
Departamento de Engenharia Mecânica.	
1. Reconstrução 3D	2. Visão Computacional
3. Imagens de Profundidade	4. Alinhamento de Dados 3D
I. ENM/FT/UnB	II. Título (série)

## REFERÊNCIA BIBLIOGRÁFICA

ALBUQUERQUE, L. A. (2006). Alinhamento de Imagens de Profundidade na Reconstrução 3D De Objetos de Forma Livre. Dissertação de Mestrado em Sistemas Mecatrônicos, Publicação ENM.DM-09A/06, Departamento de Engenharia Mecânica, Universidade de Brasília, Brasília, DF, 125p.

## CESSÃO DE DIREITOS

AUTOR: Landecir Alves de Albuquerque.

TÍTULO: Alinhamento de Imagens de Profundidade na Reconstrução 3D de Objetos de Forma Livre.

GRAU: Mestre ANO: 2006

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

---

Landecir Alves de Albuquerque  
Landmailbox@gmail.com  
Bairro Generoso, Avenida Sabino José da Costa, nº 09.  
79.300-000, Corumbá /MS - Brasil.

## **AGRADECIMENTOS**

Agradeço primeiramente ao meu orientador, o professor José Maurício S. T. Motta, por toda ajuda, todo horizonte de conhecimentos que me proporcionou, toda a paciência, todos os conselhos, incentivos e atenção que dedicou à minha orientação.

Ao professor Sadek Crisóstomo Absi Alfaro, uma pessoa de valor inestimável e que recebe a todos de braços abertos.

À Universidade de Brasília, uma instituição que respeito e admiro.

E a todo o departamento de Engenharia Mecânica da UnB, e em especial ao GRACO (Grupo de Automação e Controle), porque uma pesquisa é resultado do trabalho direto e indireto de muitas pessoas.

## **DEDICATÓRIA**

Este trabalho é dedicado a todas as pessoas que amam a pesquisa. Um pesquisador é antes de tudo um trabalhador incansável e perspicaz, que conhece a fundo o alto preço do domínio do conhecimento.

## **RESUMO**

### **ALINHAMENTO DE IMAGENS DE PROFUNDIDADE NA RECONSTRUÇÃO 3D DE OBJETOS DE FORMA LIVRE**

**Autor: Landecir Alves de Albuquerque**

**Orientador: José Maurício Santos Torres da Motta**

**Programa de Pós-Graduação em Sistemas Mecatrônicos**

**Brasília, agosto de 2006**

A reconstrução 3D de modelos digitais a partir de imagens de profundidade obtidas com digitalizadores que não efetuam contato constitui-se atualmente o estado-da-arte para a construção de modelos. A abordagem consiste de três etapas. Na primeira o mundo real é amostrado através de imagens de profundidade, que nada mais são do que mapas com informações de distâncias medidas a partir de um visualizador a pontos na cena. A grande vantagem dessas imagens é que a informação 3D está explícita, poupando-se o trabalho de recuperá-la, o que representa um sério entrave para muitas aplicações e é inevitável quando se utiliza imagens de intensidade. A etapa subsequente é o alinhamento ou registro de imagens, que surge devido ao fato do digitalizador estar limitado a adquirir apenas porções de superfície por vez, em virtude do caráter unidirecional das varreduras. Isso leva a necessidade de se adquirir múltiplas imagens de profundidade variando-se as poses entre o digitalizador e o objeto. As imagens são adquiridas centradas no sistema de coordenadas do sensor e o processo de alinhamento consiste em trazê-las juntas dentro de um sistema de coordenadas comum ou global. A última etapa é a integração de vistas e a reconstrução da superfície do objeto virtual. Nesta dissertação é apresentado o estado tecnológico atual de todas as etapas da reconstrução 3D. A pesquisa tem por foco a segunda etapa da reconstrução, visto que esse é um problema de extrema importância para a qualidade do modelo final, pois um alinhamento incorreto gera distorções no modelo que o desaprovam para muitas aplicações que demandam precisão, como as presentes na engenharia, onde é comum a necessidade de se construir modelos tridimensionais com precisão suficiente para serem usados em sistemas de manufatura ou simulação numérica de performance de máquinas e componentes em operação, como turbinas e escoamentos em dutos não circulares. Nesse sentido também é apresentada uma implementação do algoritmo ICP para o alinhamento de múltiplas imagens de profundidade obtidas através de varredura laser. A metodologia utilizada foi submeter o algoritmo a testes com imagens de profundidades reais e dados sintéticos contaminados com variadas porcentagens de ruído.

## **ABSTRACT**

### **ALIGNMENT OF RANGE IMAGES IN THE 3D RECONSTRUCTION OF FREE-FORM OBJECTS**

**Author: Landecir Alves de Albuquerque**

**Supervisor: José Maurício Santos Torres da Motta**

**Programa de Pós-Graduação em Sistemas Mecatrônicos**

**Brasília, August of 2006**

The 3D reconstruction of digital models from range images acquired by non-contact digitizers is currently the state of the art for the construction of models. The approach consists of three stages. In the first one the real world is sampled in the range images, which are nothing more than maps with information on measured distances from a viewer to points in the scene. The great advantage of these images is that the 3D information is explicit, saving the work of retrieving it, which represents a serious obstacle for many applications and is inevitable when using intensity images. The subsequent step is the alignment or register of images, which is due to the fact that the scanner is limited only to acquire portions of surface at a time, as a result of the unidirectional nature of the sweeps. It leads to need to acquire multiple range images varying poses between the scanner and the object. The images are acquired on the sensor system of coordinates and the process of alignment is to bring them together within a global or common system of coordinates. The last step is the integration of views and reconstruction of the surface of the virtual object. In this dissertation is presented the current state of technology all stages of 3D reconstruction. The research has focused on the second stage of reconstruction, since this is an issue of extreme importance to the quality of the final model, because a misaligned creates distortions in the model that disapprove it for many applications that require precision, such as those present in engineering, where is the common need to build three-dimensional models with sufficient accuracy to be used in manufacturing systems or numerical simulation performance of machines in operation and components such as turbines and flows in non-circular ducts. In this sense is also presented an implementation of ICP algorithm for alignment of multiple range images acquired with laser scanning. The methodology used was subject the algorithm to tests with real range images and synthetic data contaminated with varying percentages of noise.

## SUMÁRIO

<b>1 – INTRODUÇÃO</b> .....	01
1.1 - CONSIDERAÇÕES INICIAIS .....	01
1.2 - MOTIVAÇÕES E JUSTIFICATIVAS PARA ESTE TRABALHO.....	03
1.3 - DELIMITAÇÃO DA PESQUISA.....	05
1.3.1 - Objetivo geral .....	05
1.3.2 - Objetivos específicos.....	05
1.4 - METODOLOGIA .....	06
1.5 - CONTRIBUIÇÕES DESTE TRABALHO.....	07
1.6 - ORGANIZAÇÃO DA DISSERTAÇÃO E APRESENTAÇÃO DOS CAPÍTULOS SUBSEQÜENTES .....	08
<b>CAPÍTULO 2 – MODELAGEM GEOMÉTRICA E RECONSTRUÇÃO 3D</b> .....	09
2.1 - ABORDAGENS PARA A CONSTRUÇÃO DE MODELOS .....	09
2.2 - ETAPAS DA RECONSTRUÇÃO 3D A PARTIR DE IMAGENS DE PROFUNDIDADE.....	12
2.2.1 - Aquisição de dados .....	12
2.2.2 - Registro de imagens de profundidade .....	13
2.2.3 - Integração de vistas, reconstrução de superfície e visualização.....	13
2.3 - A RECONSTRUÇÃO 3D E ÁREAS RELACIONADAS .....	14
2.4 - APLICAÇÕES DE MODELOS DIGITAIS 3D .....	17
2.4.1 - Aplicações na engenharia reversa, prototipagem rápida e manufatura industrial.....	17
2.4.2 - Aplicação na navegação 3D .....	20
2.4.3 - Aplicações na área médica e afins .....	21
2.4.4 - Aplicações em sítios de internet, comércio eletrônico e mundos virtuais.....	22
2.4.5 - Aplicação na indústria de entretenimento .....	23
2.4.6 - Aplicações nas artes em geral e arqueologia digital.....	24
2.4.7 - Aplicações na arquitetura e construção civil.....	24
2.4.8 - Aplicações nas áreas militares, de segurança e defesa.....	25
2.4.9 - Aplicações geomáticas .....	25

<b>CAPÍTULO 3 – AQUISIÇÃO DE DADOS</b> .....	26
<b>3.1 - O PROCESSO DE FORMAÇÃO DE IMAGENS</b> .....	26
3.1.1 - O modelo de câmera pinhole .....	26
<b>3.2 - GEOMETRIA DA FORMAÇÃO DE IMAGEM</b> .....	27
3.2.1 - Projeção de perspectiva .....	27
3.2.2 - Equação de perspectiva fraca.....	28
3.2.3 - Projeção ortogonal .....	29
<b>3.3 - IDENTIFICAÇÃO DE SISTEMAS DE COORDENADAS</b> .....	29
<b>3.4 - PARÂMETROS GEOMÉTRICOS DE CÂMERA</b> .....	31
3.4.1 - Parâmetros intrínsecos.....	31
3.4.2 - Parâmetros extrínsecos .....	32
<b>3.5 - MUDANÇA SISTEMA DE COORDENADAS</b> .....	32
<b>3.6 - REPRESENTAÇÕES PARA A ROTAÇÃO</b> .....	33
3.6.1 - Representação de rotações com ângulos de Euler.....	34
3.6.2 - Representação de rotações com um ângulo e um eixo arbitrário .....	35
3.6.3 - Representação de rotações utilizando quaternions .....	35
<b>3.7 - REPRESENTAÇÃO DE TRANSLAÇÕES</b> .....	36
<b>3.8 - COORDENADAS HOMOGÊNEAS</b> .....	36
<b>3.9 - ASPECTOS DE CALIBRAÇÃO DE CÂMERAS</b> .....	37
<b>3.10 - A PERCEPÇÃO DE PROFUNDIDADE E A INFORMAÇÃO 3D ABSOLUTA</b> .....	38
<b>3.11 - TIPOS DE IMAGENS DIGITAIS</b> .....	39
3.11.1 - Imagens de intensidade .....	39
3.11.2 - Imagens de profundidade .....	40
<b>3.12 - TAXONOMIA DOS MÉTODOS DE AQUISIÇÃO DE FORMA 3D E PROFUNDIDADE ABSOLUTA</b> .....	41
<b>3.13 - PRINCÍPIOS DE IMAGEAMENTO DE PROFUNDIDADE</b> .....	42
3.13.1 - Medição de profundidade pelo princípio TOF .....	42
3.13.2 - Medição profundidade pelo princípio da triangulação.....	43
3.13.3 - Medição de profundidade explorando mudança de fase ou frequência .....	43
<b>3.14 - MÉTODOS DE AQUISIÇÃO QUE EFETUAM CONTATO</b> .....	44
<b>3.15 - MÉTODOS DE AQUISIÇÃO QUE NÃO EFETUAM CONTATO</b> .....	45
3.15.1 - Tomografia Computadorizada (CT) .....	45

3.15.2 - Radares de microondas e sonares.....	45
3.16 - MÉTODOS ÓPTICOS PASSIVOS .....	46
3.16.1 - Estereoscopia .....	46
3.16.2 - Focagem e desfocagem ativa.....	47
3.16.3 - Forma a partir de sombreamento .....	48
3.16.4 - Forma a partir de movimento .....	48
3.17 - MÉTODOS ÓPTICOS ATIVOS .....	50
3.17.1 - Iluminação e propriedades do laser.....	50
3.17.2 - Iluminação estruturada .....	50
3.17.3 - Radar baseado em laser (Ladar).....	52
3.17.4 - Interferometria de Moiré .....	52
3.18 - PLANEJAMENTO DO DESLOCAMENTO DO DIGITALIZADOR .....	53
3.19 - INCERTEZAS EM IMAGENS DE PROFUNDIDADE .....	53
<b>CAPÍTULO 4 – REGISTRO DE IMAGENS DE PROFUNDIDADE .....</b>	<b>55</b>
4.1 - DEFINIÇÃO DO PROBLEMA DO REGISTRO DE IMAGENS DE PROFUNDIDADE.....	55
4.2 - ALGUNS CONCEITOS IMPORTANTES .....	58
4.2.1 - Estimação de pose, combinação e correspondência .....	58
4.2.5 - Complexidade computacional .....	58
4.2.5 - Heurísticas e meta-heurísticas.....	62
4.2.6 - O conceito de máxima verossimilhança, o método dos mínimos quadrados e estimação robusta .....	62
4.2.6 - Otimização, convexidade de uma função e matriz de covariância .....	63
4.2.6 - Decomposição em valores singulares.....	64
4.3 - REVISÃO DOS MÉTODOS DE REGISTRO NA LITERATURA.....	64
4.3.1 - Classificações para os métodos de registro .....	65
4.4 - TRABALHOS PRÉVIOS ENVOLVENDO A CONSTRUÇÃO DE MODELOS E O PROBLEMA DO REGISTRO .....	68
4.5 - O ALGORITMO ICP .....	71
4.6 - ETAPAS ESSENCIAIS DO REGISTRO UTILIZANDO O ALGORITMO ICP	
4.6.1 - Pré-alinhamento .....	72
4.6.2 - Seleção de pontos para formar pares correspondentes .....	72

4.6.3 - Soluções para o problema da correspondência .....	73
4.6.4 - Estimação de parâmetros de movimento rígido .....	76
4.6.5 - Métrica de erro e minimização .....	79
4.6.6 - Critérios de parada .....	79
4.6.7 - Pré-processamento e rejeição de dados discrepantes.....	80
4.7 - FORMULAÇÃO MATEMÁTICA DO REGISTRO ENTRE DUAS IMAGENS UTILIZANDO ICP.....	81
4.8 - PSEUDOCÓDIGO DO ALGORITMO ICP .....	83
4.9 - ESTRATÉGIAS DE ALINHAMENTO DE MÚLTIPLAS VISTAS PARA FORMAR UM MODELO COMPLETO .....	83
4.10 - VALIDAÇÃO DE REGISTRO.....	85

## **CAPÍTULO 5 – INTEGRAÇÃO DE IMAGENS, RECONSTRUÇÃO DE SUPERFÍCIE E VISUALIZAÇÃO .....**

5.1 - FORMAS DE REPRESENTAÇÃO DE SUPERFÍCIE.....	86
5.2 - ABORDAGENS PARA A RECONSTRUÇÃO DE SUPERFÍCIE.....	88
5.3 - VISUALIZAÇÃO DO MODELO.....	90

## **CAPÍTULO 6 – PROCEDIMENTOS E RESULTADOS**

<b>EXPERIMENTAIS.....</b>	<b>92</b>
6.1 - PROCEDIMENTOS EXPERIMENTAIS .....	93
6.1.1 - Especificações do algoritmo ICP implementado .....	93
6.1.2 - Considerações sobre o elemento ruído .....	94
6.2 - RESULTADOS EXPERIMENTAIS COM DADOS 3D SINTÉTICOS DOS VÉRTICES DE UM CUBO .....	95
6.2.1 - Primeiro teste: dados sem ruído .....	95
6.2.2 - Segundo teste: 12,5% de ruído .....	97
6.2.3 - Terceiro teste: 25% de ruído .....	98
6.2.4 - Quarto teste: 50% de ruído .....	100
6.3 - RESULTADOS EXPERIMENTAIS COM IMAGENS DE PROFUNDIDADE REAIS .....	103
6.3.1 - Primeiro teste.....	103
6.3.2 - Segundo teste.....	106

<b>CAPÍTULO 7 – ANÁLISE DOS RESULTADOS.....</b>	<b>110</b>
<b>7.1 - CONSIDERAÇÕES COM RELAÇÃO AOS TESTES COM O CUBO.....</b>	<b>110</b>
<b>7.2 - CONSIDERAÇÕES COM RELAÇÃO AOS TESTES COM IMAGENS DE PROFUNDIDADE REAIS.....</b>	<b>110</b>
<b>7.3 - ASPECTOS QUE INFLUENCIAM A CONVERGÊNCIA DO ALGORITMO.....</b>	<b>111</b>
<b>CAPÍTULO 8 – CONCLUSÕES E SUGESTÕES PARA PESQUISAS FUTURAS .....</b>	<b>113</b>
<b>8.1 - CONCLUSÕES .....</b>	<b>113</b>
<b>8.2 - SUGESTÕES PARA PESQUISAS FUTURAS .....</b>	<b>114</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>116</b>
<b>APÊNDICES .....</b>	<b>124</b>
<b>APÊNDICE A - DIAGRAMA DA IMPLEMENTAÇÃO .....</b>	<b>125</b>

## LISTA DE TABELAS

Tabela 3.1 - Sistemas de coordenadas identificados na Figura 3.4 .....	31
Tabela 3.2 - Problemas de orientação em calibração de câmeras.....	37
Tabela 4.1 - Funções limitantes superiores freqüentemente utilizadas .....	61
Tabela 4.2 - Vantagens e desvantagens de representações de rotação .....	78
Tabela 4.3 - Pseudocódigo do algoritmo ICP .....	83
Tabela 6.1 - Dados obtidos nos testes de execução do algoritmo do ICP com dados de dois cubos rotacionados 45° em relação ao eixo Z.....	101
Tabela 6.2 - Ângulos de rotação estimados e o erro RMS correspondente a cada Iteração .....	102

## LISTA DE FIGURAS

Figura 2.1 - Reconstrução a partir de imagens de intensidade. ....	10
Figura 2.2 - Robô de soldagem auxiliado por visão .....	15
Figura 2.3 - Relacionamento entre as áreas de processamento de imagens, computação gráfica e visão computacional .....	16
Figura 2.4 - Regra da engenharia reversa .....	18
Figura 2.5 - Peça produzida por prototipagem rápida .....	19
Figura 2.6 - Engenharia reversa do modelo CAD de um rotor de turbina .....	20
Figura 2.7 - Robô móvel inteligente Athene II .....	21
Figura 2.8 - Esquema que evidencia a abrangência dos processos envolvidos em um sistema composto por reconstrução 3D e prototipagem rápida em medicina	22
Figura 2.9 - Modelagem 3D na indústria de entretenimento .....	23
Figura 2.10 - Projeto de reconstrução 3D da estátua David de Michelangelo .....	24
Figura 3.1 - O modelo de câmera pinhole .....	26
Figura 3.2 - Projeção de perspectiva.....	28
Figura 3.3 - Projeção ortogonal. ....	29
Figura 3.4 - Sistemas de coordenadas de um sistema de aquisição. ....	30
Figura 3.5 - Relação entre o sistema da câmera e do mundo. ....	33
Figura 3.6 - Rotação utilizando ângulos de Euler.....	34
Figura 3.7 - Rotação utilizando um ângulo e um eixo arbitrário. ....	35
Figura 3.8 - Estrutura de dados .....	40
Figura 3.9 - Taxonomia de Curless para os métodos de aquisição de forma 3D .....	42
Figura 3.10 - Modulação em frequência .....	43
Figura 3.11 - Uma típica máquina de medição de coordenadas (CMM) .....	44
Figura 3.12 - Braço Articulado. ....	45
Figura 3.13 - Geometria da estereoscopia .....	47
Figura 3.14 - Recuperação de geometria de superfície a partir de dicas em imagens .....	49
Figura 3.15 - Configuração de um digitalizador baseado em iluminação estruturada .....	51
Figura 3.16 - Configuração básica do padrão geométrico e da interferometria de Moiré .....	53
Figura 4.1 - Geometria do processo de aquisição por múltiplas vistas .....	56
Figura 4.2 - Geometria da aquisição de duas vistas.....	57
Figura 4.3 - $f(n)$ é $O(g(n))$ .....	60

Figura 4.4 - $f(n)$ é $\Theta(g(n))$ .....	60
Figura 4.5 - $f(n)$ é $\Omega(g(n))$ .....	60
Figura 4.6 - Exemplo de uma função não convexa .....	63
Figura 4.7 - Detecção de contorno em uma peça automotiva .....	69
Figura 4.8 - Correspondência ponto a ponto em 2D.....	74
Figura 4.9 - Correspondência ponto a plano em 2D.....	76
Figura 4.10 - Acúmulo de erros de registro em pares representado em 2D .....	84
Figura 5.1 - De imagens de profundidade a superfície de profundidade .....	88
Figura 5.2 - Renderização em malha de facetas triangulares de uma imagem de profundidade utilizada nos procedimentos experimentais deste trabalho .....	91
Figura 6.1 - Ângulos rotação estimados em função da iteração no experimento dos cubos sem ruído .....	95
Figura 6.2 - Erro de alinhamento RMS em função da iteração no experimento dos cubos sem ruído .....	95
Figura 6.3 - Renderização do alinhamento dos cubos sem ruído .....	96
Figura 6.4 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 12,5% de ruído .....	97
Figura 6.5 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 12,5% de ruído .....	97
Figura 6.6 - Renderização do alinhamento dos cubos com 12,5% de ruído.....	98
Figura 6.7 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 25% .....	98
Figura 6.8 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 25% de ruído .....	99
Figura 6.9 - Renderização do alinhamento dos cubos com 25% de ruído.....	99
Figura 6.10 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 50% de ruído .....	100
Figura 6.11 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 50% de ruído .....	100
Figura 6.12 - Renderização do alinhamento dos cubos com 50% de ruído.....	101
Figura 6.13 - Ângulos rotação estimados em função da iteração no 1º teste de alinhamento do par de imagens de profundidade .....	103

Figura 6.14 - Erro de alinhamento RMS em função da iteração no 1º teste de alinhamento do par de imagens de profundidade .....	103
Figura 6.15 - Renderização como pontos do alinhamento do par de imagens de profundidade no 1º teste .....	106
Figura 6.16 - Ângulos rotação estimados em função da iteração no 2º teste de alinhamento do par de imagens de profundidade .....	106
Figura 6.17 - Erro de alinhamento RMS em função da iteração no 2º teste de alinhamento do par de imagens de profundidade .....	107
Figura 6.18 - Renderização como pontos do alinhamento do par de imagens de profundidade no 2º teste .....	109

## LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES

2D	- Bidimensional
3D	- Tridimensional
CAD	- Computer Aided Design (Projeto Assistido por Computador)
CAE	- Computer-aided engineering (Engenharia Assistida por Computador)
CAM	- Computer Aided Manufacturing (Manufatura Assistida por Computador)
CCD	- Charge Coupled Device (Dispositivo de Carga Acoplada)
CFD	- Computational Fluid Dynamics (Dinâmica dos Fluidos Computacional)
CMM	- Coordinate Measuring Machine (Máquina de Medição de Coordenadas)
CT	- Computer Tomography (Tomografia Computacional)
DARCES	- Data Aligned Rigidity Constrained Exhaustive Search (Alinhamento de Dados por Busca Exaustiva Confinada por Rigidez)
GRACO	- Automation and Control Group (Grupo de Automação e Controle)
IBM	- International Business Machines
ICP	- Iterative Closest Point (Ponto Mais Próximo Iterativo)
Internet	- Rede Mundial de Computadores
LADAR	- Laser detection and ranging (Detecção e Determinação de Distância por Laser)
LASER	- Light Amplification by Stimulated Emission of Radiation (Amplificação de Luz por Emissão Estimulada de Radiação)
MATLAB	- Matrix Laboratory (Laboratório Matricial)
MRI	- Magnetic resonance imaging (Imageamento por Ressonância Magnética)
OpenGL	- Open Graphics Library (Biblioteca Gráfica Aberta)
Pixel	- Picture Element (Elemento de Imagem)
RADAR	- Radio Detecting and Ranging (Detecção e Determinação de Distância por Radiofrequência)
RANSAC	- Random Sample Consensus (Consenso de Amostra Aleatória)
RMS	- Root mean square (Raiz da Média dos Quadrados)
SEM	- Scanning Electron Microscope (Microscópio de Varredura Eletrônica)
SIM	- Surface Interpenetration Measure (Medida de Interpenetração de Superfície)

SONAR	- Sound Navigation and Ranging (Navegação e determinação de Distância pelo Som)
STL	- Formato de Arquivo para Esteriolitografia
STM	- Scanning Tunneling Microscope (Microscópio de Varredura por Tunelamento)
SVD	- Singular Value Decomposition (Decomposição em Valores Singulares)
TOF	- Time of Flight (Tempo de Percurso)
UnB	- University of Brasília (Universidade de Brasília)
VFSR	- Very Fast Simulated Reannealing (Têmpera Muito Rápida Simulada)
VRML	- Virtual Reality Modeling Language (Linguagem de Modelagem para Realidade Virtual)

# 1 – INTRODUÇÃO

Neste capítulo são apresentados os esclarecimentos sobre o tema, a relevância da proposta, a motivação, os objetivos gerais e específicos, a metodologia, e as contribuições deste trabalho.

## 1.1 - CONSIDERAÇÕES INICIAIS

Na atualidade há uma demanda crescente por modelos geométricos digitais tridimensionais de objetos ou mesmo ambientes do mundo real. Constata-se que essa procura tem sido impulsionada pelo fato de tais modelos estarem se tornando recursos indispensáveis para a execução de inúmeras tarefas de caráter geométrico, gráfico e visual em computador. Uma vasta gama de áreas e aplicações podem ser beneficiadas com a disponibilidade desses modelos. Alguns exemplos são: aplicações de reconhecimento, inspeção, representação, planejamento, visualização, navegação, simulação e análise; presentes nas mais variadas áreas do conhecimento humano, como *manufatura industrial, medicina, robótica, artes, arqueologia digital, indústria de entretenimento, segurança, defesa*, e inúmeras outras.

Para se obter um modelo digital é necessário recorrer-se a técnicas de modelagem 3D. As técnicas empregadas historicamente na engenharia tem sido baseadas na síntese de modelos utilizando-se programas modeladores conhecidos como *ferramentas CAD (Computer Aided Design)*, ou na construção de modelos a partir de dados obtidos com um digitalizador de contato. Entretanto, a mais sofisticada abordagem para a construção de modelos digitais na atualidade consiste da utilização de imagens que codificam a geometria tridimensional do objeto e que são obtidas através de equipamentos que efetuam digitalização sem contato. Digitalizadores baseados em laser são especialmente atrativos porque são compactos, seguros e geram uma grande e densa quantidade de dados. Ou seja, são ideais para se digitalizar com riqueza de detalhes objetos que apresentam estrutura geométrica complexa.

Essa mudança de paradigma é explicada por uma combinação de fatores. No passado, os equipamentos que podiam medir diretamente coordenadas 3D de superfícies por um modo sem contato eram tidos como caros, imprecisos e restritos a digitalização de objetos de

tamanho bastante reduzido. Abordagens tradicionais de medição por contato eram mais precisas e acessíveis economicamente. Entretanto, a evolução tecnológica de componentes eletrônicos, o barateamento dos equipamentos de digitalização, e o aumento do poder de processamento de computadores são fatores que estão tornando cada vez mais popular a obtenção de imagens de profundidade por meio de digitalizadores 3D ópticos ativos, e o processamento e visualização de quantidades massivas de dados. Imagens de profundidade, por sua vez, nada mais são do que mapas com informações de distâncias medidas a partir de um observador (sensor ou câmera) a pontos na cena. Conseqüentemente, a informação 3D é explícita, evitando-se cálculos necessários para sua recuperação; necessidade esta que é um sério entrave para muitas aplicações e inevitável quando se faz uso de imagens de intensidade.

A abordagem da construção de modelos conhecida por *reconstrução 3D* é uma verdadeira tecnologia de modelagem que compreende uma variedade de técnicas, equipamentos e conceitos, geralmente compreendidas em três etapas. A primeira etapa é a de *aquisição de dados*, onde detalhes da geometria ou estrutura 3D do objeto físico real são amostrados através de imagens adquiridas com o emprego de um digitalizador adequado para aplicação. Como os dados de interesse são medições de profundidades, sensores que fazem a leitura direta dessas informações, como os digitalizadores ópticos ativos, levam vantagem sobre outros digitalizadores. Outras virtudes desses instrumentos incluem a rapidez, flexibilidade e a produção de mapas de dados com alta resolução e alta taxa de amostragem. Mas, em contrapartida, são instrumentos de leitura unidirecional. Por esse motivo, diversas imagens precisam ser tomadas a partir de diferentes pontos de vista para que a superfície completa do objeto seja capturada. Esse empecilho cria a necessidade de uma segunda etapa na reconstrução: o *alinhamento* ou *registro de múltiplas imagens de profundidade* adquiridas do objeto. Tecnicamente, as imagens encontram-se desalinhadas no sentido de que cada uma é adquirida centrada em um sistema de coordenadas diferente, problema esse que é ocasionado pelas mudanças de pose entre o sensor e o objeto no processo de aquisição. Essa etapa é de extrema importância para a qualidade do modelo final. O alinhamento incorreto gera distorções no modelo que o desaprovam para diversas aplicações que exigem exatidão ou fidelidade com o objeto sensoriado. Esta fase também é o foco desta pesquisa. A última etapa da reconstrução é a *integração das vistas e a reconstrução da superfície do modelo*. Esta última etapa tem o propósito de finalizar integrar as imagens em uma estrutura única e global e reconstruir a superfície do modelo,

de modo que se tenha uma representação computacional completa e em alto nível do objeto real.

## 1.2 - MOTIVAÇÕES E JUSTIFICATIVAS PARA ESTE TRABALHO

- *A reconstrução 3D a partir de imagens de profundidade é atualmente o estado-da-arte para aquisição de modelos digitais.* Mas intrinsecamente também é uma tecnologia complexa, onde o sucesso no objetivo de construir um modelo está associado a diversos fatores. Dentre esses fatores citam-se: os requisitos da aplicação alvo, como exatidão geométrica, realismo visual, ou rapidez na aquisição do modelo, a escolha de uma entre muitas opções de digitalização, adversidades como a complexidade topológica inerente a superfícies de forma livre, a degradação da qualidade dos dados sensorizados por medições errôneas, a solução do difícil problema da correspondência e a indisponibilidade de informações importantes para a reconstrução. Portanto, é um grande desafio, principalmente para a engenharia, a busca do domínio, aprimoramento e o avanço tecnológico das diversas etapas e técnicas dessa metodologia para a construção de modelos.

Uma vasta gama de aplicações pode fazer uso de modelos computacionais. Essa é uma das principais provas da relevância dessa proposta na atualidade. Como exemplo, citam-se casos na *engenharia* e na *indústria*. A reconstrução é especialmente importante para a engenharia para se obter modelos para três grandes aplicações: *simulação numérica CFD (Computational Fluid Dynamics)*, *prototipagem rápida* e *navegação robótica*.

Pode ser fator decisivo para muitos projetos de engenharia a disponibilidade de um modelo computacional 3D preciso de um objeto físico sólido de forma livre para tarefas como, testes, análises, aprimoramentos ou simulações na fase de desenvolvimento ou reengenharia de um protótipo. Um exemplo recorrente acontece com máquinas construídas em décadas passadas que estão passando a não mais atender a metas de produtividade estabelecidas no projeto. Essas máquinas podem estar deterioradas pelo desgaste natural imposto pelo longo tempo de uso, ou precisam ser submetidas a processos de modernização para compensar o obsolescência em relação a tecnologias mais atuais. Nesses casos, é comum não se dispor das especificações de projeto para a construção de um novo protótipo, e a abordagem mais sofisticada é fazer a engenharia reversa das peças ou

componentes para se obter o modelo 3D do objeto físico. Já no cenário tecnológico industrial, modelos digitais estão no centro das soluções de problemas gerados por um estado de inovações constantes do mercado. Os meios de fabricação e desenvolvimento são forçados a responder adequadamente a essas mudanças, pois a competitividade ou mesmo a sobrevivência de empresas são fatores que dependem de metas como a alta qualidade e o desenvolvimento rápido e a baixo de custo de novos produtos. É necessário, por exemplo, que fatores importantes, como a quantidade dos dados utilizados para a reconstrução e o tempo gasto na sua aquisição, sejam otimizados na fase de prototipagem, visto que ambos representam custo variável para o fluxo de caixa de uma empresa. Nessa empreitada as indústrias estão empregando práticas com alta capacidade de adaptação, rapidez, e baseadas em modelos, como a *manufatura flexível* e a *prototipagem rápida*. Além do fator “mercado”, o crescente nível de automação e informatização que levou a consolidação de paradigmas *CAE (Computer-aided engineering)*, *CAD (Computer Aided Design)* e *CAM (Computer Aided Manufacturing)* no passado recente, também está levando a uma clara convergência para uso de tecnologias baseadas em modelos 3D por reconstrução. Indústrias como a automotiva e a aeroespacial, que já utilizam modelos *CAD* de longa data, reforçam essa perspectiva.

Portanto, qualquer avanço nos processos de desenvolvimento e fabricação, tanto nas práticas da engenharia como no chão-de-fábrica das indústrias, tendem a causar um grande impacto em toda a cadeia produtiva.

- *Alinhamento é uma etapa crucial da reconstrução 3D, principalmente para aquisição de modelos de alto nível.* Inserido em todo o contexto da construção de modelos, está o problema chave do registro ou alinhamento de dados 3D. Muitos pesquisadores, principalmente da área da visão computacional, têm empreendido esforços em busca da solução desse complexo e fundamental problema da reconstrução. Avanços especificamente nesse tema podem beneficiar novas e promissoras aplicações. Dois exemplos são: aplicações de tempo real, como a navegação 3D de robôs móveis; ou de alta precisão, como a construção de modelos de alto nível para tarefas como a simulação numérica, que é uma prática comum na engenharia.

## **1.3 - DELIMITAÇÃO DA PESQUISA**

### **1.3.1 - Objetivo geral**

O objetivo geral deste trabalho é fazer uma revisão de todos os aspectos relacionados com a tecnologia da construção de modelos digitais a partir de imagens de profundidade. A meta é proporcionar uma visão panorâmica do tema através de um levantamento comparativo de elementos como a importância, o estado tecnológico, as aplicações e as diversas técnicas e equipamentos empregados, com respeito às três etapas reconstrução 3D. A intenção não é esgotar o assunto, visto que é bastante vasto e em pleno estado de desenvolvimento.

### **1.3.2 - Objetivos específicos**

- Adquirir conhecimento para a construção de modelos computacionais completos, para serem usados em aplicações de simulação numérica na engenharia.
- Avaliar a viabilidade do emprego de imagens de profundidade obtidas por digitalizadores ópticos em substituição a abordagem tradicional de se usar digitalizadores de contato.
- Implementar o algoritmo *ICP (Iterative Closest Point)* com o propósito de alinhar múltiplas imagens de profundidades obtidas por varredura laser. Esse objetivo também se constitui o conteúdo prático desta dissertação. O motivo para escolha desse algoritmo está no fato de ser um método vastamente empregado para a proposta de alinhamento, tendo em vista a sua facilidade de adaptação e resultados já apresentados. Já a justificativa para o foco na segunda etapa está no planejamento do projeto de reconstrução. Considerando as 3 etapas do processo, a solução da primeira pode ser contornada utilizando-se imagens de profundidade disponibilizadas gratuitamente em bancos de imagens na *Internet*. Essa medida precisa ser tomada para permitir a execução de testes de algoritmos até que um digitalizador sob medida seja construído. A terceira etapa pode ser resolvida exportando-se a nuvem de pontos alinhados para que um software modelador reconstrua a superfície com uma representação adequada para aplicação. A segunda etapa, no entanto, é uma das mais críticas para projetos que tem a exatidão como foco, e não dispõe de soluções gerais.

Muitas propostas promissoras e até de grande originalidade intelectual tem aparecido para solucionar o problema do alinhamento na reconstrução 3D. Entretanto, como esse problema é altamente dependente da aplicação, essas abordagens precisam de análise teórica e muitos testes para se comprovar a eficiência empiricamente. O *ICP*, por outro lado, já vem sendo submetido a melhoramentos constantes por muitas pesquisas e está se estabelecendo como um método clássico para a proposta de alinhamento de dados 3D.

- Avaliar o comportamento do algoritmo *ICP* quanto à convergência para os valores corretos de alinhamento quando na presença de dados discrepantes. A estratégia é submeter o algoritmo, na configuração proposta originalmente, a testes com dados contaminados com variadas porcentagens de ruído introduzido artificialmente.
- Estabelecer conclusões com base na análise dos resultados apresentados e sugerir aprimoramentos futuros para o algoritmo.

#### **1.4 - METODOLOGIA**

Os meios para ser atingir os objetivos desta pesquisa baseiam-se no estudo de métodos teóricos consagrados pela literatura especializada, e procedimentos puramente experimentais.

Um dos objetivos principais deste trabalho consiste da implementação do algoritmo *ICP* para alinhamento de dados, e nesse sentido, a metodologia é, antes de tudo, um estudo de casos, caracterizando uma metodologia indutiva. Essa abordagem baseia-se no fato de que o conceito de “bom alinhamento” é relativo e multidependente, o que faz com que o resultado final seja condicionado a diversas variáveis, como: requisitos da aplicação e características dos dados disponíveis, dentre muitas outras. Diante disso, uma fórmula única de emprego geral é pouco provável e talvez impossível de ser concebida, restando a alternativa de avaliar situações de interesse.

Tecnicamente os procedimentos consistem da implementação e realização de sessões de testes experimentais para avaliar a convergência do algoritmo em questão. Esse algoritmo apresenta poucos passos, é de fácil entendimento e implementação. Mas como qualquer trabalho na área visual, e principalmente no espaço tridimensional, existe por trás uma

grande quantidade de algoritmos e recursos complexos que precisam ser empregados para se atingir os resultados desejados. Algumas das implementações necessárias são: os mais diversos tipos de implementações para efetuar operações de computação matricial, como armazenamento e manipulação de dados 3D através de matrizes multidimensionais; operações básicas com matrizes como adição, multiplicação, cálculo do posto e traço; algoritmos de busca, ordenação e geração de números aleatórios; técnicas de visualização de modelos 3D através de imagens de intensidade para possibilitar a avaliação visual dos resultados e a comparação com os parâmetros encontrados; bibliotecas para converter pontos em malhas de polígonos; métodos numéricos como decomposição matricial por *SVD*; e inúmeras outras implementações da álgebra linear numérica, estatística, geometria computacional e métodos heurísticos. Ressalta-se também que dificuldades a mais são impostas pela dimensão tridimensional do espaço dos dados que são trabalhados, e a complexidade inerente às formas das superfícies de objetos de forma livre.

## **1.5 - CONTRIBUIÇÕES DESTE TRABALHO**

- A primeira e mais geral contribuição deste trabalho é uma revisão sobre o estado-da-arte da reconstrução 3D.
- Propor soluções para uma aplicação que apresenta grande carência de modelos e que tem por prioridade a exatidão. A aplicação em questão é a da simulação numérica CFD na engenharia.
- Implementar o algoritmo para a construção de modelos através do alinhamento de múltiplas imagens de profundidade. O código implementado pode ser adaptado para incorporar novas heurísticas a suas etapas.
- Explorar uma brecha encontrada nas pesquisas acerca do tema. A reconstrução 3D é uma abordagem altamente sofisticada, mas devido a incertezas nas medições, há ainda uma grande preocupação quanto à qualidade de imagens de profundidade para certas propostas, como a construção de modelos exatos. Essas incertezas constituem-se um grande obstáculo e acrescenta importância ao fator “convergência” do algoritmo de alinhamento.

- Um parecer sobre a viabilidade da utilização de digitalizadores laser para a engenharia reversa de modelos *CAD*.

- Por fim, uma contribuição é também apresentar um parecer sobre a aplicabilidade prática dessa implementação.

## **1.6 - ORGANIZAÇÃO DA DISSERTAÇÃO E APRESENTAÇÃO DOS CAPÍTULOS SUBSEQÜENTES**

Esta dissertação esta organizada de acordo com a seqüência das etapas da reconstrução 3D.

O capítulo 2 discorre sobre o campo da modelagem 3D de um modo geral, suas abordagens e áreas correlatas.

O capítulo 3 visa introduzir os conceitos e técnicas relacionadas com a aquisição de dados 3D.

O capítulo 4 é o foco desta pesquisa. Neste capítulo os métodos para registro de imagens 3D são analisados. É apresentado o estado-da-arte, são discutidos os avanços e as referências na literatura.

O capítulo 5 introduz noções sobre a fusão de imagens e a reconstrução da superfície de um modelo digital e a necessidade de visualização do modelo construído.

O capítulo 6 explica a metodologia empregada e os resultados obtidos nos experimentos com a implementação.

O capítulo 7 faz uma análise dos resultados obtidos.

O capítulo 8 trata das conclusões e propõe futuras direções de pesquisa acerca do tema.

## CAPÍTULO 2 – MODELAGEM GEOMÉTRICA E RECONSTRUÇÃO 3D

Este capítulo tem o propósito de apresentar uma visão geral do tema da reconstrução 3D através da discussão de vários aspectos teóricos e práticos. O capítulo se inicia com os tipos de abordagens para a construção de modelos. Depois são descritas as etapas da reconstrução 3D e discutidas as ramificações nas áreas congêneres. O capítulo termina com a apresentação de uma pequena amostra da extensa gama de áreas e aplicações que podem ser beneficiadas com a tecnologia em estudo.

### 2.1 - ABORDAGENS PARA A CONSTRUÇÃO DE MODELOS

Um *modelo geométrico* computacional é uma representação digital da estrutura geométrica de um objeto existente fisicamente ou idealizado para uma dada aplicação (Requicha, 1999; Shum *et al.*, 1997). Sua disponibilidade permite a realização de inúmeras tarefas gráficas e visuais em computador, como simulação, análise, aperfeiçoamento, reconhecimento, inspeção, planejamento, dentre outras. *Modelagem geométrica* por sua vez é o processo de obtenção desses modelos. Neste trabalho, o termo *objeto* refere-se ao assunto ou tema específico a ser modelado; *forma-livre* refere-se ao fato de não se impor restrições a forma desse objeto, e termos recorrentes como *vista* e *varredura* ganham muitas vezes generalidade de imagem de profundidade.

Historicamente o problema da modelagem 3D tem sido abordado de duas maneiras. A primeira consiste em se construir, sintetizar ou literalmente desenhar o assunto (objeto ou cena) tridimensionalmente em computador. Programas de computador para modelagem geométrica conhecidos como ferramentas CAD tem sido usados para esse propósito. Essa abordagem popularizou-se como uma prática da engenharia, mas as técnicas utilizadas por esses softwares são tradicionalmente alvo de pesquisa da computação gráfica. A inadequação da síntese de modelos começa a ficar evidente nas situações em que o objeto a ser modelado apresenta o que pode ser classificado como um grau de dificuldade geométrica elevado. A complexidade geométrica é inerente aos objetos existentes no mundo real. Entende-se por isso que a topologia de suas superfícies não está restrita a um conjunto de classes de formas como planos, poliedros, ou esferas, por exemplo. Além de impor restrições geométricas ao objeto a ser modelado, a modelagem *CAD* também é um

processo totalmente dependente de intervenção humana, demorado, trabalhoso e caro. Proporcionalmente a complexidade do modelo, essa abordagem tende a utilizar muitas horas de mão-de-obra especializada, influenciando variáveis que determinam a competitividade de empresas, como tempo e custo. No entanto, essa abordagem sempre foi e provavelmente continuará a ser dominante, principalmente para se fazer o modelo de entrada de um protótipo e para projetos que se enquadram em áreas mais específicas, como a modelagem de sólidos, por exemplo.

A segunda abordagem é a reconstrução a partir de imagens. A reconstrução utilizando medições diretas de informações de profundidade é a solução para muitas das limitações da síntese de formas. O fundamento dessa abordagem é amostrar o mundo real através de imagens tomadas por sensores e combiná-las de modo a reconstruir o modelo computacionalmente. Essa abordagem é análoga à forma como a visão humana trabalha, pegando as imagens da retina e criando um modelo 3D do mundo no cérebro. As imagens tomadas podem ser do tipo que codifica informações de intensidade luminosa, chamadas de *imagens de intensidade* (2D), ou do tipo que codifica medições de distância de um visualizador a pontos na cena, chamadas de *imagens de profundidade* (2,5D). Esse tipo de abordagem é alvo de muitas pesquisas da comunidade da visão computacional, e em outras palavras, é uma virtualização do mundo real.

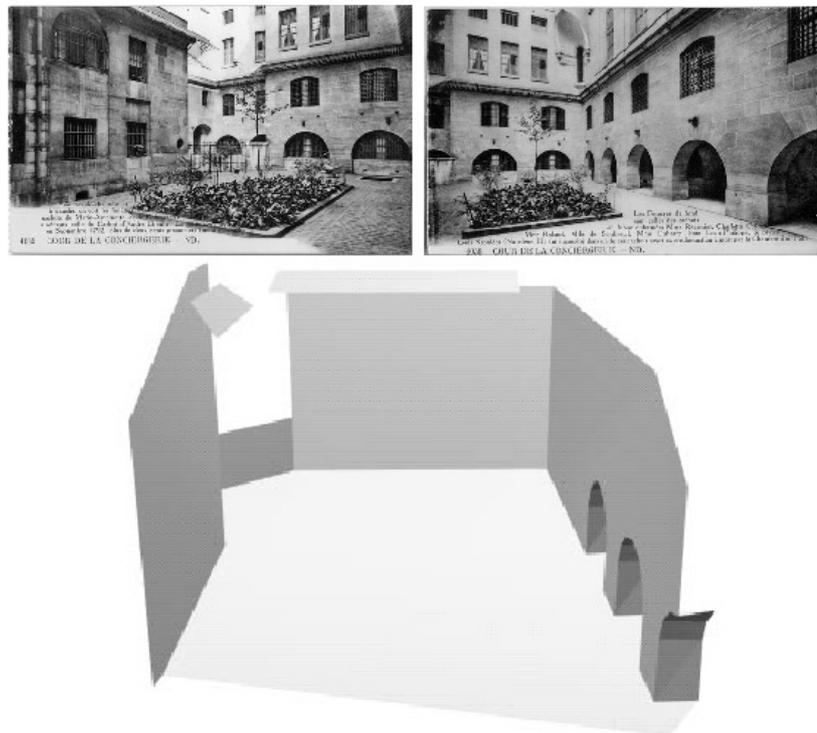


Figura 2.1 - Reconstrução a partir de imagens de intensidade.

Muitas razões qualificam a reconstrução 3D utilizando dados de profundidade como a abordagem mais sofisticada para a modelagem baseada em imagens. A primeira é que imagens de intensidade são projeções do mundo real com a perda da informação de profundidade, portanto, usar medições 3D diretas poupa-se o trabalho de ter de se recuperar a informação 3D a partir de imagens 2D. Por outro lado, obter dados para reconstrução com métodos que efetuam contato apresenta algumas desvantagens, como o fato de geralmente produzirem mapas com baixa resolução de amostragem, ou esparsos. Outro complicador é o já citado problema de que objetos existentes no mundo real e que precisam ser modelados geralmente possuem geometria complexa. Para se chegar a modelos de alto nível o digitalizador precisa retornar muitos pontos e com alta resolução para que sejam captados muitos detalhes de superfície, como curvaturas, concavidades com dimensões muito pequenas, arestas agudas, e muitas outras. Além disso, o estado tecnológico atual associado aos custos mais acessíveis dos equipamentos para aquisição e ao poder de processamento de quantidades massivas de dados 3D estão agindo favoravelmente para tornar mais popular e acessível a abordagem da reconstrução com base em imagens de profundidade obtidas com digitalizadores baseados em laser. Ressalta-se também que, na empreitada da construção de um modelo, quanto mais informações estiverem disponíveis, mais provável é de se ter bons resultados. O trabalho em laboratório com iluminação controlada, sistema de aquisição calibrado, com planejamento de vistas, e com o controle de todos os fatores em conta no processo, podem propiciar maior possibilidade de sucesso na construção do modelo. Outra exigência da tecnologia da reconstrução é um domínio e conhecimento técnico considerável de todas as suas etapas, pois são necessárias habilidades e conhecimentos no tocante a modelos matemáticos, conhecimento sobre propriedades físicas que interferem no ambiente, exercício de programação exaustiva para a implementação de estruturas de dados complexas para armazenamento de grande quantidade de dados, técnicas de visualização de modelos 3D através de imagens de intensidade, algoritmos de computação geométrica, construção e calibração de equipamentos especializados como sistemas robóticos para posicionamento de sensores e digitalizadores, e exaustivos testes em bancada.

## **2.2 - ETAPAS DA RECONSTRUÇÃO 3D A PARTIR DE IMAGENS DE PROFUNDIDADE**

Três etapas podem ser identificadas na Reconstrução: a *aquisição de dados de profundidade*, o *alinhamento ou registro de imagens*, e a *integração de imagens e reconstrução de superfície*. A qualidade do modelo final está diretamente relacionada com o sucesso de cada uma dessas etapas.

### **2.2.1 - Aquisição de dados**

Uma grande variedade de dispositivos podem ser empregados para se digitalizar objetos ou cenas. Dependendo da aplicação, as profundidades envolvidas na digitalização podem variar do nível microscópico a extensões de ordem quilométrica. Cenas muito grandes podem abranger os domínios de áreas como a fotogrametria e sensoriamento remoto, alguns exemplos nesse sentido ocorrem na utilização de imagens adquiridas através de imagens SAR (*Synthetic Aperture Radar*) de satélites artificiais e a construção de modelos 3D de cidades. Esses dispositivos também podem fornecer dados em vários formatos como imagens de intensidade, medições de profundidade isoladas e até os atraentes mapas de profundidade completos, conhecidos como imagens de profundidade. Mapas de profundidade podem ser obtidos também através de técnicas baseadas em imagens de intensidade, como a estereoscopia, entretanto, este método é pouco recomendável para muitas aplicações, como as que têm por prioridade a precisão do modelo.

Digitalizadores ópticos 3D baseados em laser particularmente já atingiram um estado de desenvolvimento tecnológico que os qualificam para uso em aplicações que exigem modelos de alto nível de qualidade. Já existem também digitalizadores comerciais que retornam um modelo completo, entretanto, são muito limitados e termos de volume de trabalho e exatidão. Para projetos de engenharia é comum a necessidade de se construir um digitalizador sob medida, empregando-se para esse propósito recursos como diodos lasers e algum dispositivo para movimentar o objeto ou a câmera, como um manipulador robótico, por exemplo.

### 2.2.2 - Registro de imagens de profundidade

Por mais que os digitalizadores tenham evoluído, uma limitação que ainda não teve uma solução definitiva é o fato de, em geral, esses dispositivos só poderem varrer o objeto a partir de única direção. Por isso, múltiplas imagens de profundidades precisam ser adquiridas variando-se pontos de vista, de modo que seja possível adquirir todas as informações de superfície necessárias para reconstruir o modelo. A varredura completa do objeto resulta em uma seqüência de imagens de profundidade, sendo que cada uma foi adquirida centrada no sistema de coordenadas do sensor. Tecnicamente, alinhar as imagens é o mesmo que trazê-las juntas dentro de um sistema de coordenadas comum ou global, que pode ser tanto um sistema de coordenadas centrado no objeto quanto o sistema de coordenadas de qualquer uma das vistas. O alinhamento correto é necessário para evitar distorções no modelo final. O que torna possível essa operação é a recuperação das transformações de movimento rígido, ou seja, as rotações e/ou translações sofridas pelo sensor ou pelo objeto, e a reaplicação desses movimentos às imagens de modo a alinhá-las.

### 2.2.3 - Integração de vistas, reconstrução de superfície e visualização

Preferivelmente que uma nuvem de pontos, o produto final desejado da reconstrução para muitas aplicações é um modelo monolítico com uma representação de superfície de alto nível. O desejável é que a superfície do objeto seja uma réplica da superfície do objeto real, sem brechas ou distorções, produzindo o que costuma ser chamado de *modelo à prova d'água*. Essas preocupações fazem parte da última etapa da reconstrução: a *integração de vistas e reconstrução de superfície*. Essa etapa enfrenta vários obstáculos associados principalmente com a qualidade dos dados de profundidade digitalizados. Problemas como ruído, regiões não digitalizadas pelo sensor (*occlusões*), e baixa resolução de amostragem (mapas esparsos), podem afetar tanto a etapa de alinhamento como a etapa de reconstrução de superfície, gerando imperfeições no modelo. A primeira forma, e muitas vezes a única, de se avaliar o modelo final é evidentemente a crítica visual. Mas não é possível enxergar o modelo estando na forma de uma estrutura de dados 3D na memória do computador, é necessário visualizá-lo através de fotografias. A visualização do objeto consiste da *renderização* de pontos 3D para imagens de intensidade, e isso nada mais é do que a projeção dos pontos de uma dimensão 3D para uma dimensão 2D, exatamente como

acontece na projeção de perspectiva de um ponto do mundo real a um ponto na retina ou num plano de imagem. As técnicas para essa tarefa são oriundas principalmente da computação gráfica e é necessário determinar várias especificações como cor, ponto de vista de visualização, iluminação e refletância. Bibliotecas gráficas como o *OpenGL* fornecem rotinas para esse tratamento. Ressalta-se também que digitalizadores de profundidade não captam atributos como cor e refletância. Para incluir esses atributos a um modelo é necessária solução de um problema conhecido como *mapeamento de textura*.

### **2.3 - A RECONSTRUÇÃO 3D E ÁREAS RELACIONADAS**

A reconstrução 3D é um campo tecnológico emergente em franca expansão e desenvolvimento. É justamente por ser uma área muito jovem e em pleno estado de formação, não há muito consenso quanto às suas terminologias e domínios, suscitando muitas debates e controvérsias quando se procura definições. Os únicos consensos estão no fato de ser uma abordagem de modelagem geométrica e estar intimamente relacionada com a visão computacional. A multidisciplinaridade também é uma característica dessa tecnologia, pois ela agrega uma diversidade recursos encontrados nos mais variados campos de conhecimento, variando de instrumentos de digitalização, a práticas da engenharia e métodos da ciência da computação; juntas essas técnicas constituem o que muitos pesquisadores chamam de *técnicas de fotografia 3D*. Devido em muito a essa diversidade, é difícil situar o tema em uma área específica. Grande parte das técnicas empregadas nessa tecnologia vem da visão computacional, mas não se pode dizer que ela pertence exclusivamente aos domínios dessa área. As fases de registro e reconstrução de superfície, por exemplo, são permeados por problemas de modelagem de dados e estimação, que utilizam ferramentas computacionais oriundas em grande parte da computação gráfica a estatística.

A *visão computacional* é uma área da ciência da computação que procura recuperar e interpretar informações de cenas do mundo real a partir de imagens. Sua meta é criar um modelo do mundo real a partir de projeções contidas em imagens (Jain *et al.*, 1995). É nesse contexto que a construção de modelos digitais tridimensionais se insere. É também uma matéria multidisciplinar e em pleno desenvolvimento, que no seu exercício de recuperação de cena agrega técnicas de diversas outras áreas como: matemática, estatística, processamento de imagens, computação gráfica, reconhecimento de padrões, inteligência

artificial, psicofísica, dentre outras. Por ser um ramo científico relativamente novo, muitos consensos ainda estão sendo estabelecidos com respeito a diversos aspectos. Muitos pesquisadores, por exemplo, consideram certas terminologias, como *análise de imagens*, *entendimento de cena*, *análise de cena* e *visão de máquinas* como sinônimos de visão computacional (Jain *et al.*, 1995; Trucco e Verri, 1998). Existe, por exemplo, diferenças de ponto de vista em relação a área denominada *visão de máquinas*. Batchelor e Whelan defendem que a visão computacional é um ramo da ciência da computação, enquanto a visão de máquinas é uma área de especialização dentro da *engenharia de sistemas* e que não implica na implementação de suas técnicas exclusivamente em computador. Uma prova disso é que técnicas da visão computacional podem ser implementados fisicamente com equipamentos eletrônicos, geralmente com o objetivo de se obter maior velocidade de processamento em relação às mesmas técnicas implementadas computacionalmente (Batchelor e Whelan, 1997). Já Guda e co-autores definem o campo da visão de máquinas como a aplicação das técnicas da visão computacional a automação industrial (Guda *et al.*, 2000). Algumas das tarefas visuais freqüentemente automatizadas na indústria são: problemas de inspeção, reconhecimento, montagem e guiagem de equipamentos auxiliada por imagens.

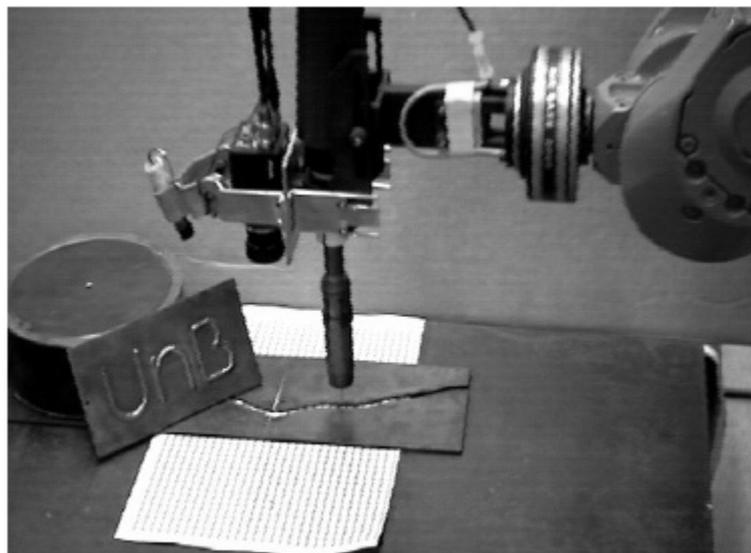


Figura 2.2 - Robô de soldagem auxiliado por visão, GRACO/UnB.  
No detalhe, câmera CCD acoplada à ferramenta.

Não é incomum também a confusão entre os objetivos da visão computacional com outras duas disciplinas intimamente relacionadas: a computação gráfica e o processamento de

imagens. Qualquer dúvida pode ser esclarecida analisando-se as informações que correspondem a “*entrada*” e “*saída*” dessas áreas.

Enquanto o objetivo da *computação gráfica* é construir imagens a partir de primitivas ou descrições, como pontos, retas, curvas, superfícies, etc, a visão computacional obedece a uma estratégia inversa, extraindo primitivas e descrições de uma imagem.

Já a área do *processamento de imagens* trata da transformação de uma imagem em outra imagem. Algumas tarefas típicas dessa área são a filtragem, restauração, realce, compressão e segmentação de imagens.

Em suma, a visão computacional trata da análise de imagens, a computação gráfica da síntese de imagens, e o processamento de imagens da transformação de imagem em imagem.

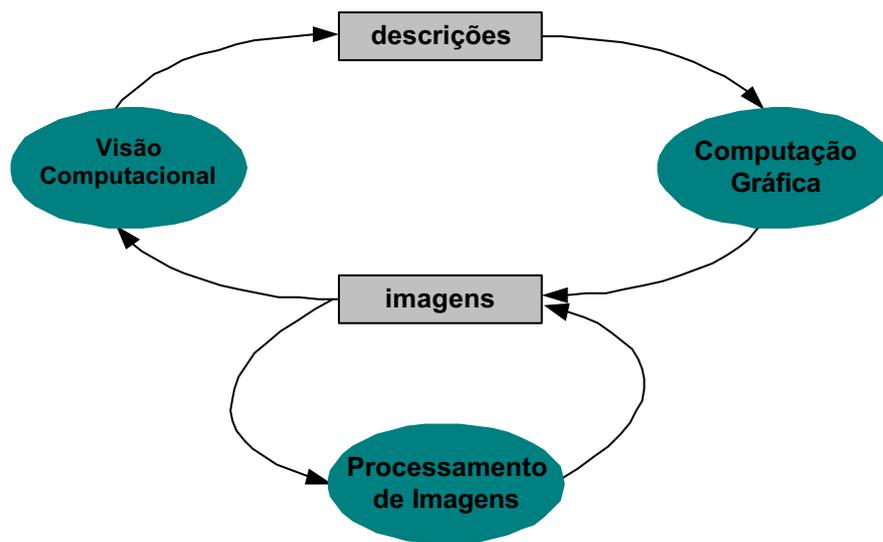


Figura 2.3 - Relacionamento entre as áreas de processamento de imagens, computação gráfica e visão computacional.

No sentido de que sistemas de visão computacional geralmente são compostos de componentes de percepção (o que implica em sensoriamento) e entendimento (que implica em análise de imagens), a visão computacional é freqüentemente considerada parte integrante de uma área maior, que é a *inteligência artificial* (Jain *et al.*, 1995). A inteligência artificial, por sua vez, compreende três etapas: *percepção, cognição e ação*.

Tarefas como extração e a análise de símbolos são permeadas pelo raciocínio e inteligência. É também opinião generalizada de pesquisadores, que o fato de o conhecimento sobre a inteligência humana ser incipiente constitui um dos principais motivos de estar ainda muito distante o dia da visão computacional emular com fidelidade o funcionamento do sistema de visão humano.

## **2.4 - APLICAÇÕES DE MODELOS DIGITAIS 3D**

Grande parte da literatura relacionada à reconstrução 3D se inicia com uma listagem das aplicações dessa tecnologia (Pulli, 1997; Curless, 1997; Reed, 1998; Silva e Bellon, 1995). São inúmeras as áreas que podem fazer uso de modelos digitais. Alguns exemplos são: a navegação 3D, engenharia reversa, inspeção industrial, realidade virtual, indústria de entretenimento, arte, arqueologia digital e medicina. Prevê-se que muitas aplicações ainda devam surgir, em virtude dessa área ser relativamente jovem e estar em pleno e rápido desenvolvendo. Nesta seção será demonstrada apenas uma amostra da variedade de aplicações.

### **2.4.1 - Aplicações na engenharia reversa, prototipagem rápida e manufatura industrial**

O termo *engenharia reversa* remete muitas vezes à prática de violações de direitos de propriedade sobre programas de computador. No entanto, é antes de tudo um conceito e um recurso lícito, quando feita dentro da legalidade, e muito recorrente na engenharia. Num processo direto parte-se do projeto na forma de um modelo computacional para se chegar ao produto ou protótipo físico. A engenharia reversa é o processo inverso: dispõe-se de um produto e deseja-se chegar ao seu projeto. No contexto da reconstrução, dispõe-se de um objeto físico real, e deseja-se chegar ao seu modelo computacional. Esse modelo guarda intrinsecamente na sua geometria as especificações do projeto (Eggert *et al.*, 1998; Cerrada *et al.*, 1990; Page *et al.*, 2003). A reconstrução 3D concretiza a engenharia reversa no sentido virtualizar o modelo físico.

A abordagem tradicional para o desenvolvimento de protótipos na engenharia é construir primeiro um modelo 3D em software *CAD* e depois exportá-lo em um formato de arquivo

especifico, como *STL*, para uma máquina *CAM*, para que produto seja manufaturado empregando algum processo de fabricação. Durante a fase de desenvolvimento de um protótipo um objeto pode viajar várias vezes entre o modelo físico no mundo real e o modelo virtual computacional, e vice-versa, num ciclo de aprimoramento composto por duas etapas claras de engenharia reversa e impressão 3D, e que se repete até que se chegue ao resultado desejado.

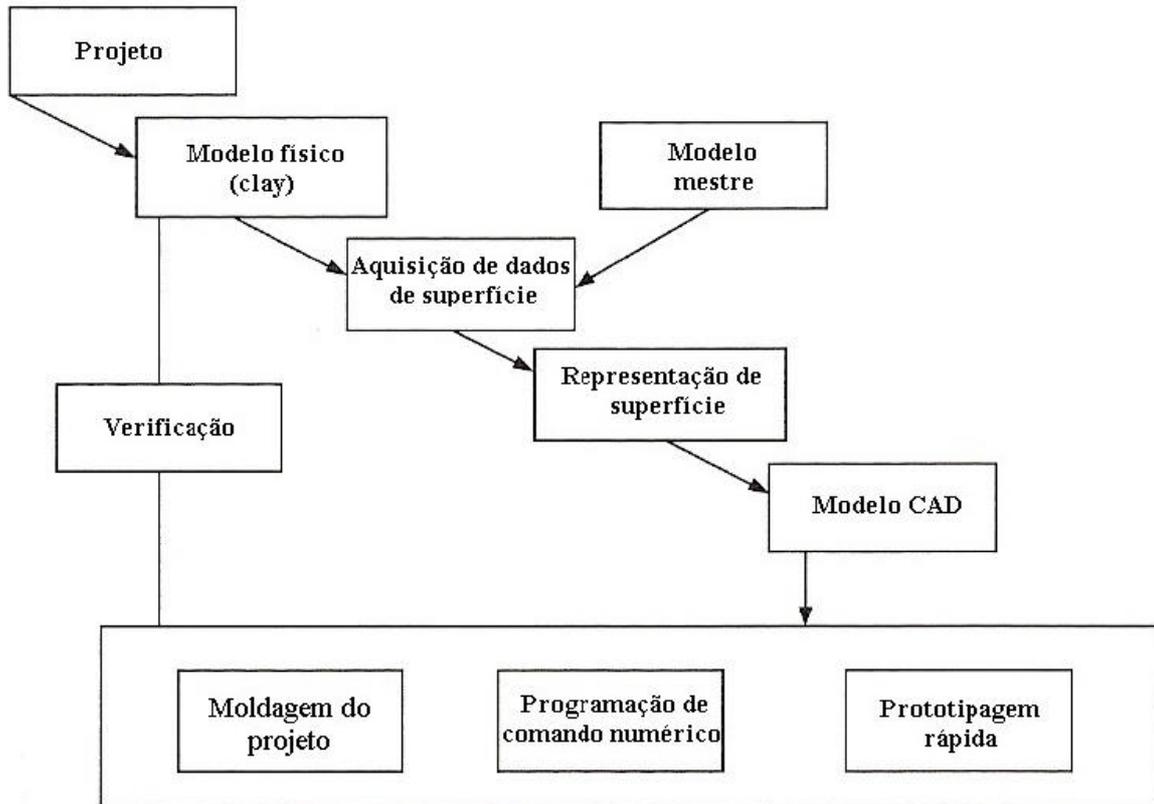


Figura 2.4 - Regra da engenharia reversa (modificado - Sinha e Jain, 1994).

A *prototipagem rápida* é o processo de fabricação mais sofisticado que pode ser usado para fabricação, pois é composta de tecnologias de impressão 3D que empregam modelos computacionais tridimensionais.

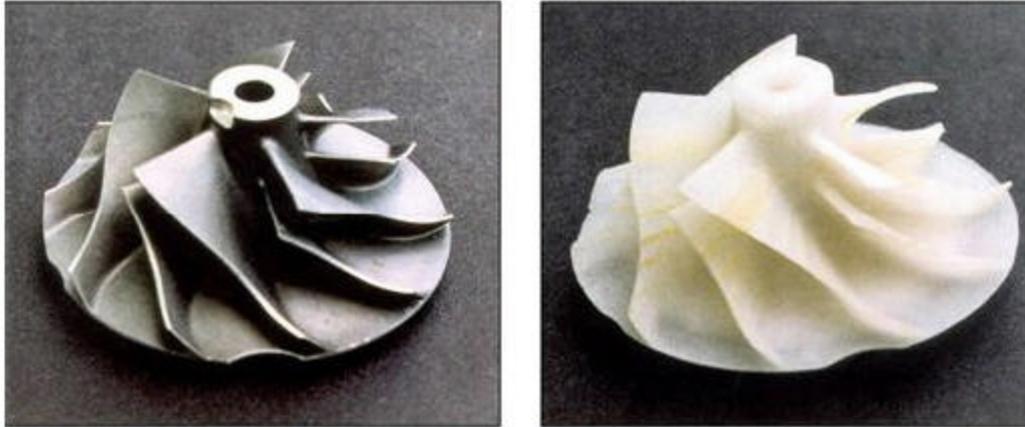


Figura 2.5 - Peça produzida por prototipagem rápida. À esquerda: objeto físico. À direita, protótipo construído a partir do arquivo *STL*.

É praticamente uma regra não existir o modelo computacional de objetos construídos em décadas passadas, tais como manufaturados em geral, produtos eletrônicos ou peças mecânicas, que foram concebidos no início ou mesmo antes do advento do computador. Nos casos mais graves nem mesmo existem as especificações do projeto em forma escrita. Por isso, uma importante aplicação da reconstrução na atualidade está na construção de modelos voltados para propósitos como a refabricação, a reengenharia, ou simplesmente a montagem um inventário virtual. Um exemplo concreto disso ocorre com componentes de turbinas de hidrelétricas, principalmente as de grande porte. Além da própria deterioração ocasionada pelo longo tempo em operação, as pás dos rotores dessas máquinas funcionam em regime de alta pressão, e por isso estão sujeitas a erosão por cavitação. Turbinas mais antigas precisam da manutenção ou mesmo da substituição dessas peças. A refabricação desses componentes passa por diversos procedimentos e estudos do comportamento dinâmico através análise, visualizações, e muitas horas de simulações computacionais. No nível mais técnico, abstrações como a geração de malhas de elementos finitos a partir de formas 3D são necessárias para a simulação dinâmica do escoamento de fluidos no interior da turbina em funcionamento. Aplicações dessa magnitude exigem modelos fiéis ao objeto físico, o que implica em exatidão e alta resolução de pontos. Modelos com essas características são considerados de alto nível, e só podem ser obtidos por reconstrução 3D.



Figura 2.6 - Engenharia reversa do modelo *CAD* de um rotor de turbina, *GRACO/UnB*. À esquerda, o corte vertical de uma turbina Kaplan. Ao centro, uma maquete de uma pá de turbina. À direita, o modelo computacional do rotor da turbina obtido através de um digitalizador de contato.

#### 2.4.2 - Aplicação na navegação 3D

Um dos propósitos para os quais robôs são concebidos é substituir a mão-de-obra humana em diversas tarefas, principalmente as de caráter monótonas, insalubres ou que envolvem algum tipo de risco a integridade física de uma pessoa. Da mesma forma que os seres humanos utilizam massivamente o sentido da visão no dia-a-dia, é de grande importância dotar robôs, ou qualquer máquina com capacidade de mobilidade, com sistemas sensoriamto baseados em imagens para que possam perceber o ambiente em que estão inseridos e assim tornar possível o desvio de obstáculos durante a locomoção. Esse é um problema típico de navegação, onde o mapeamento tridimensional de um ambiente pode proporcionar uma melhor movimentação. Já não é incomum robôs móveis estarem sendo dotados de sensores de longa profundidade, como radares baseados em laser (*ladares*), para a tarefa específica de mapear ambientes em 3D. Nesse tipo de aplicação a reconstrução deve ser feita geralmente em tempo real, e a meta principal, portanto, é a rapidez na obtenção dos modelos. Ressalta-se também que a percepção do ambiente é um tema altamente complexo por si só, mas é apenas um primeiro passo no problema da navegação 3D. A forma de utilização dos mapas, o planejamento de trajetória, a tomada de decisões e o controle de movimentos também são grandes desafios ainda pouco pesquisados e que precisam ser tratados e nessa aplicação (Florczyk, 2005).



Figura 2.7 - Robô móvel inteligente *Athene II*, usado principalmente para estudos de navegação em ambiente interno e aprendizagem de máquina (Bischoff e Graefe, 1998).

#### 2.4.3 - Aplicações na área médica e afins

A reconstrução 3D é uma ferramenta valiosa para a medicina e outras áreas relacionadas. Modelos digitais podem auxiliar no planejamento médico, diagnóstico e interpretação por imagem, decisão terapêutica e em procedimentos cirúrgicos (Paiva *et al.*, 1999). Modelos digitais para medicina geralmente têm a característica de serem reconstruídos a partir de dados sensoriados na forma de fatias bidimensionais, no caso de imagens de tomografias computadorizadas (*CT*), de ressonância magnética (*MRI*), ultra-som ou medicina nuclear (Souza *et al.*, 2001; Paiva *et al.*, 1999). Uma nova geração de digitalizadores tomográficos está evoluindo no sentido de se reconstruir modelos tridimensionais texturizados de pacientes em tempo real. Outras aplicações incluem o uso de modelos 3D de animais para estudo em laboratório, projetos conhecidos como *cobaias virtuais*, e modelos virtuais da anatomia humana, que são projetos conhecidos como *homem virtual*. Na odontologia uma aplicação é a construção de modelos para a fabricação de próteses ou moldes de arcadas dentárias.

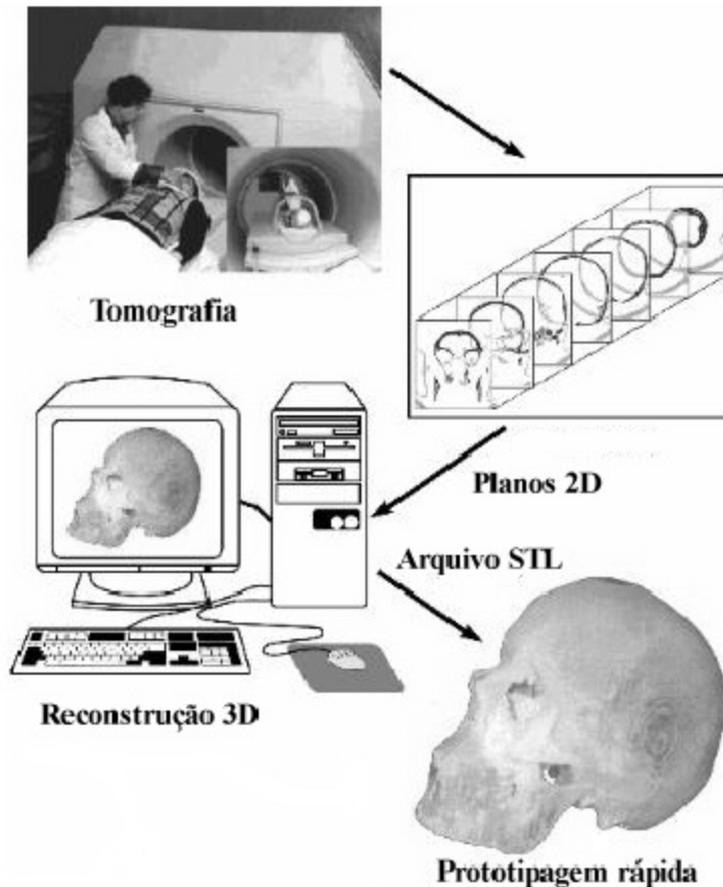


Figura 2.8 - Esquema que evidencia a abrangência dos processos envolvidos em um sistema composto por reconstrução 3D e prototipagem rápida em medicina (Souza *et al.*, 2001).

#### 2.4.4 - Aplicações em sítios de internet, comércio eletrônico e mundos virtuais

Com o surgimento da linguagem *VRML* (*Virtual Reality Modeling Language*), programas para navegação na Internet ganharam a capacidade de interpretar e visualizar modelos 3D. Aparentemente essa aplicação se dissemina com lentidão, em parte devido ao seu desconhecimento do grande público. A proposta dos modelos digitais nesse contexto é agregar realismo e interatividade a elementos que já são puramente visuais, como as páginas de *Internet*. Um dos segmentos que mais pode ser beneficiado é o do comércio eletrônico, que é um mercado em grande ascensão, devido ao fato de um estabelecimento comercial virtual atingir um público muito maior que o limitado à região geográfica de estabelecimento físico, e por proporcionar custos de manutenção bem menores comparados aos custos de manutenção de um comércio convencional. Outro propósito dos modelos na

*Internet* é o povoamento de mundos virtuais, cita-se a construção de *avatares* para personificação em ambientes de relacionamento virtuais.

#### **2.4.5 - Aplicação na indústria de entretenimento**

A indústria de entretenimento é uma das indústrias que mais se beneficiaram e ajudaram a impulsionar o desenvolvimento das técnicas de modelagem 3D. Os modelos para esse campo têm a característica de serem desenvolvidos para serem apreciados por olhos humanos, caracterizando uma proposta de visualização contemplativa, diferentemente da meta alçada para os modelos em projetos de engenharia. Nesse sentido, o motivo mais forte para se utilizar modelos virtuais no cinema e em jogos tridimensionais é a necessidade de se criar um elemento visual que não existe no mundo real, e que seja principalmente *realístico* o suficiente para dar a impressão de ser real. O realismo nesse caso tem o significado de alta fidelidade com o mundo real. Hoje em dia, o nível de realismo já se chegou a um ponto que muitas vezes é difícil identificar em um filme quais objetos, personagens ou cenários são reais e quais são virtuais. O campo específico da animação 3D apresenta um grande consumo de modelos digitais.



Figura 2.9 - Modelagem 3D na indústria de entretenimento.

#### **2.4.6 - Aplicações nas artes em geral e arqueologia digital**

Algumas áreas como a arqueologia digital e as artes em geral precisam virtualizar objetos físicos com propósitos como preservação e restauração de patrimônio histórico e obras de

arte, ou construção de museus ou acervos virtuais. Alguns exemplos de trabalhos nessa área são: o projeto *Michelangelo Digital*, desenvolvido pela Stanford University (Levoy *et al.*, 2000); o projeto *Great Buddha*; e os projetos *Pietà* e *Eternal Egypt*, conduzidos pelo grupo *IBM T. J. Watson researcher center*. Um ponto chave nesse tipo de reconstrução está nas exigências da digitalização, que precisa ser feita preferivelmente com o emprego de um digitalizador que não efetue contato, visto que objetos de arte ou arqueológicos são geralmente frágeis, fixos, muito grandes, ou não permitem mudança de localização. A superfície desses objetos também é muito rica em cores, texturas e detalhes, como marcas de instrumentos de corte, que só digitalizadores com uma alta resolução podem captar. Além de permitir a reprodução de cópias precisas como as obras de museus de cera, acervos virtuais 3D também servem de contingência para prevenir a perda total e a descaracterização de obras raras, que sabidamente, são passíveis de sinistros como incêndios, furtos e degradação por intempéries da natureza.

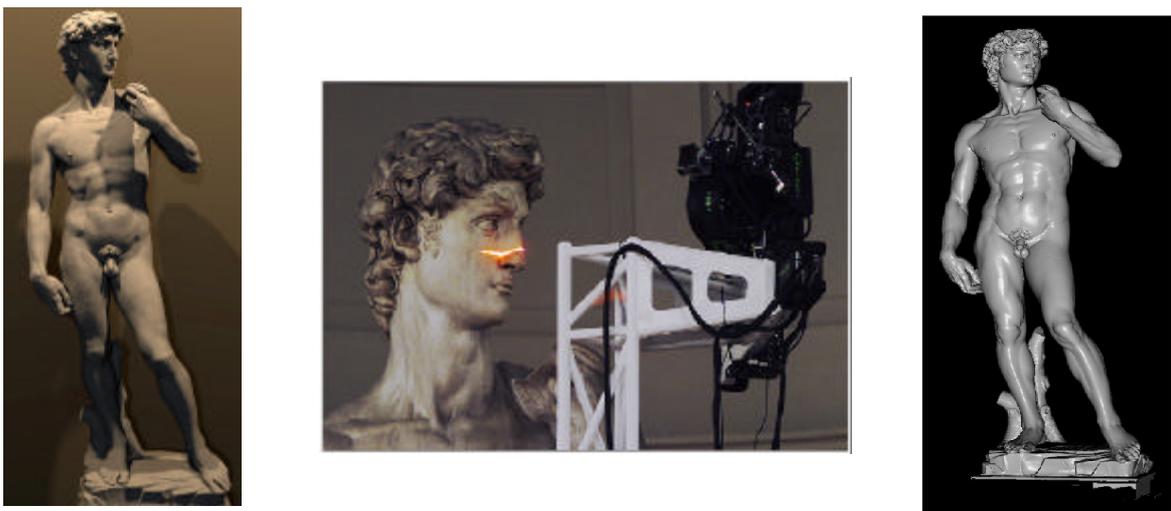


Figura 2.10 - Projeto de reconstrução 3D da estátua David, de Michelangelo (modificado - Levoy, 2003).

#### 2.4.7 - Aplicações na arquitetura e construção civil

A reconstrução de modelos pode facilitar a elaboração de plantas tridimensionais de edificações, bem como visualização de projetos arquitetônicos armazenados em 3D, e até mesmo auxiliar na concepção artística e planejamento gráfico de construções.

#### **2.4.8 - Aplicações nas áreas militares, de segurança e defesa**

Modelos computacionais têm potencial para se tornarem ferramentas auxiliares sofisticadas para áreas como a criminalística. O emprego da digitalização de dados e a reconstrução de modelos permitem abordagens mais realísticas e precisas para a reconstituição do cenário de um crime ou para testes de balística. Futuramente também poderão ser montadas bases de dados de modelos para os mais variados propósitos, como reconhecimento biométrico e criptografia tridimensional baseada em modelos. Outra aplicação é na formação de banco de modelos para a localização, detecção e reconhecimento de alvo, ou para a montagem de ambientes de simuladores para treinamento.

#### **2.4.9 - Aplicações geomáticas**

A reconstrução 3D de modelos se insere também no contexto de aplicações geomáticas como geoprocessamento, geologia ou na indústria petrolífera. Com base em fotos aéreas ou de sensoriamento remoto é possível fazer a reconstrução terrenos de grandes extensões, como cidades, estradas de rodagem e leitos de oceanos.

## **CAPÍTULO 3 – AQUISIÇÃO DE DADOS**

A proposta deste capítulo é fazer um levantamento de uma variedade de técnicas e conceitos relacionados com a aquisição de dados para a construção de modelos digitais tridimensionais. São apresentados princípios básicos sobre tipos de imagens e a geometria de formação. São discutidas as vantagens e desvantagens na utilização de dispositivos imageadores para certas aplicações e apresentadas diversas técnicas de aquisição de forma 3D e medição de profundidade absoluta, com foco na digitalização 3D baseada em laser. Ressalva-se que o conteúdo deste capítulo representa apenas uma pequena amostra de um tema extremamente vasto e em vigoroso estado de desenvolvimento tecnológico.

### **3.1 - O PROCESSO DE FORMAÇÃO DE IMAGENS**

Os aspectos fundamentais envolvidos no processo de formação de uma imagem são: a *radiometria*, que compreende a relação entre o brilho de um ponto no plano de imagem como função da iluminação e de propriedades de superfícies da cena; a *geometria* de formação de imagem, que especifica sistemas de coordenadas e as transformações de pontos entre referenciais, e a *digitalização e quantização* da imagem. Este capítulo se concentrará no aspecto da geometria, pois é de extrema importância para o problema da reconstrução.

#### **3.1.1 - O modelo de câmera pinhole**

Para estudar a geometria de aquisição será utilizado neste trabalho um modelo simples e clássico de câmera de intensidade, conhecido desde épocas remotas da fotografia. Com os avanços tecnológicos esse modelo tem se limitado ao uso como recurso artístico em técnicas de pintura. O modelo em questão é conhecido como *câmera pinhole*, e consiste de uma câmera escura com um pequeno orifício, por onde são captadas amostras de raios luminosos refletidos pelas superfícies da cena, e que atingem o plano oposto ao plano que contém o orifício, formando uma imagem invertida (Figura 3.1). Se o plano de formação da imagem for revestido de uma superfície fotossensível como um filme fotográfico, ou por uma matriz de fotossensores eletrônicos, como *CCD* ou *CMOS*, a imagem é registrada. Esse modelo não contém lentes e, portanto, não serão modeladas as distorções causadas por esses elementos.

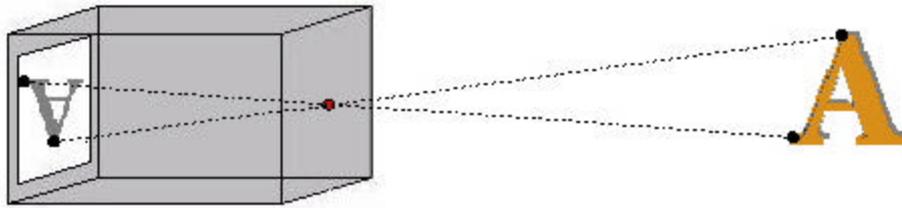


Figura 3.1 - O modelo de câmera pinhole.

## 3.2 - GEOMETRIA DA FORMAÇÃO DE IMAGEM

Transformações geométricas exercem um papel importante na geometria de sistemas de aquisição. Um tipo de transformação fundamental para a formação de imagens de intensidade são as projeções, que basicamente são transformações que mapeiam pontos entre sistemas de coordenadas com dimensões diferentes. Há diversos tipos de projeções, a depender, por exemplo, do tipo de superfície de formação de imagem, que pode ser plana, cilíndrica ou esférica (Persiano e Oliveira, 1989). Dois tipos de transformações sobre superfícies planas são de especial interesse neste estudo: a *projeção de perspectiva* e a *projeção ortogonal*.

### 3.2.1 - Projeção de perspectiva

A projeção de perspectiva é o modelo que governa a transformação do mundo real em imagem de intensidade na câmera pinhole (Figura 3.2). Pontos tridimensionais da cena são representados, ou melhor, projetados em um meio bidimensional, que é chamado de plano de formação de imagem, ou simplesmente *plano de imagem*. Determinar essa relação é um problema de calibração de câmeras. Nesse tipo de projeção os raios de luz formam linhas concorrentes convergindo para o ponto  $O$ , que é chamado de *centro óptico* ou *centro de projeção*. A projeção de perspectiva altera medidas e formas reais de um objeto no mundo real; regiões mais próximas parecem maiores, e regiões mais longínquas parecem menores. O modelo que relaciona as coordenadas  $X$ ,  $Y$ , e  $Z$  de um ponto  $P$  observado na cena aos pontos  $x'$  e  $y'$  do plano de imagem é o da equação de perspectiva:

$$(x', y') = \left( f \frac{X}{Z}, f \frac{Y}{Z} \right) \quad (3.1)$$

onde:

$Z$  é a profundidade do objeto;

$f$  é o comprimento focal, ou a distância do plano de imagem ao centro óptico.

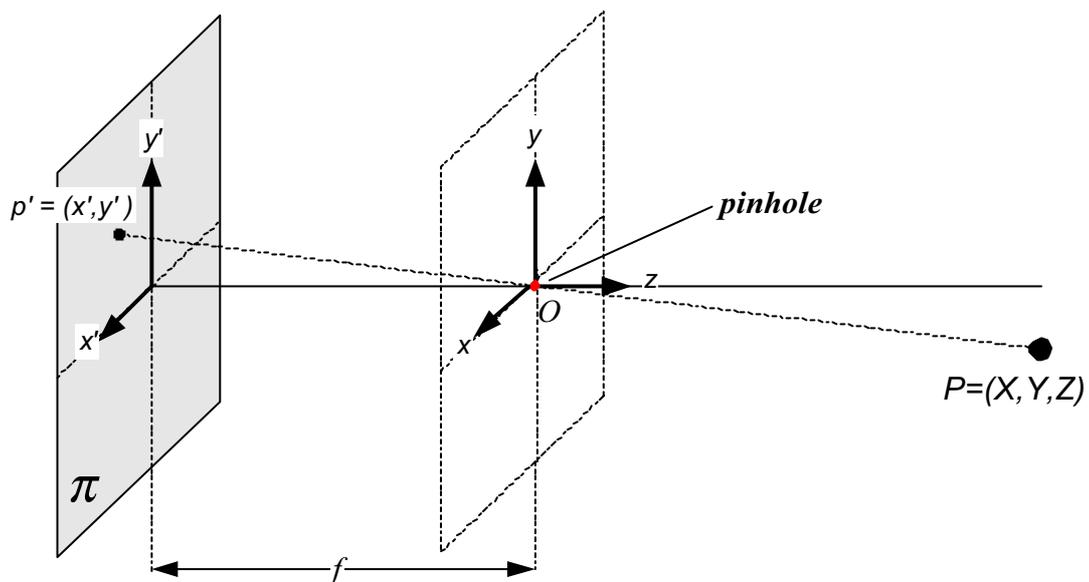


Figura 3.2 - Projeção de perspectiva.

### 3.2.2 - Equação de perspectiva fraca:

Pelo fato de  $\frac{1}{Z}$  ser não linear, é conveniente linearizar a projeção reescrevendo a Equação

3.1 na forma de perspectiva fraca:

$$(x', y') \approx \left( \frac{f}{\bar{Z}} X, \frac{f}{\bar{Z}} Y \right) \quad (3.2)$$

Onde:

$\bar{Z}$  é a profundidade média;

$f$  é o comprimento focal.

### 3.2.3 - Projeção ortogonal

A projeção ortogonal é um tipo de projeção muito utilizado em projetos gráficos em computador, para visualizar objetos preservando escala e forma, sem a deformação causada pela perspectiva. Esse tipo de projeção é um caso especial da projeção de perspectiva quando  $f \rightarrow \infty$ . Na projeção ortogonal os raios luminosos que partem da cena são paralelos ao eixo  $Z$  e atingem sempre perpendicularmente o plano de imagem. Matematicamente, pontos  $P = (X, Y, Z)$  nas coordenadas do mundo são mapeados para pontos  $(x', y')$  no plano de imagem, na forma de:

$$(x', y') = (X, Y) \quad (3.3)$$

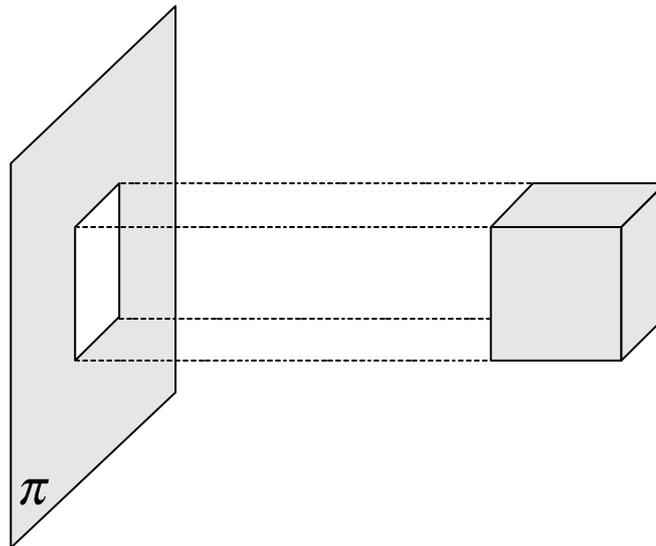


Figura 3.3 - Projeção ortogonal.

### 3.3 - IDENTIFICAÇÃO DE SISTEMAS DE COORDENADAS

Para entender a geometria envolvida em um sistema de aquisição é necessário identificar ou anexar os sistemas de coordenadas envolvidos no processo (Figura 3.4). O sistema deve ser visualizado pelo ponto de vista gráfico e modelado algebricamente, caracterizando os pontos das imagens como vetores e descrevendo cada sistema de coordenadas por uma base ou gerador formado por um conjunto de vetores linearmente independentes. O número de vetores da base informam a dimensão do espaço.

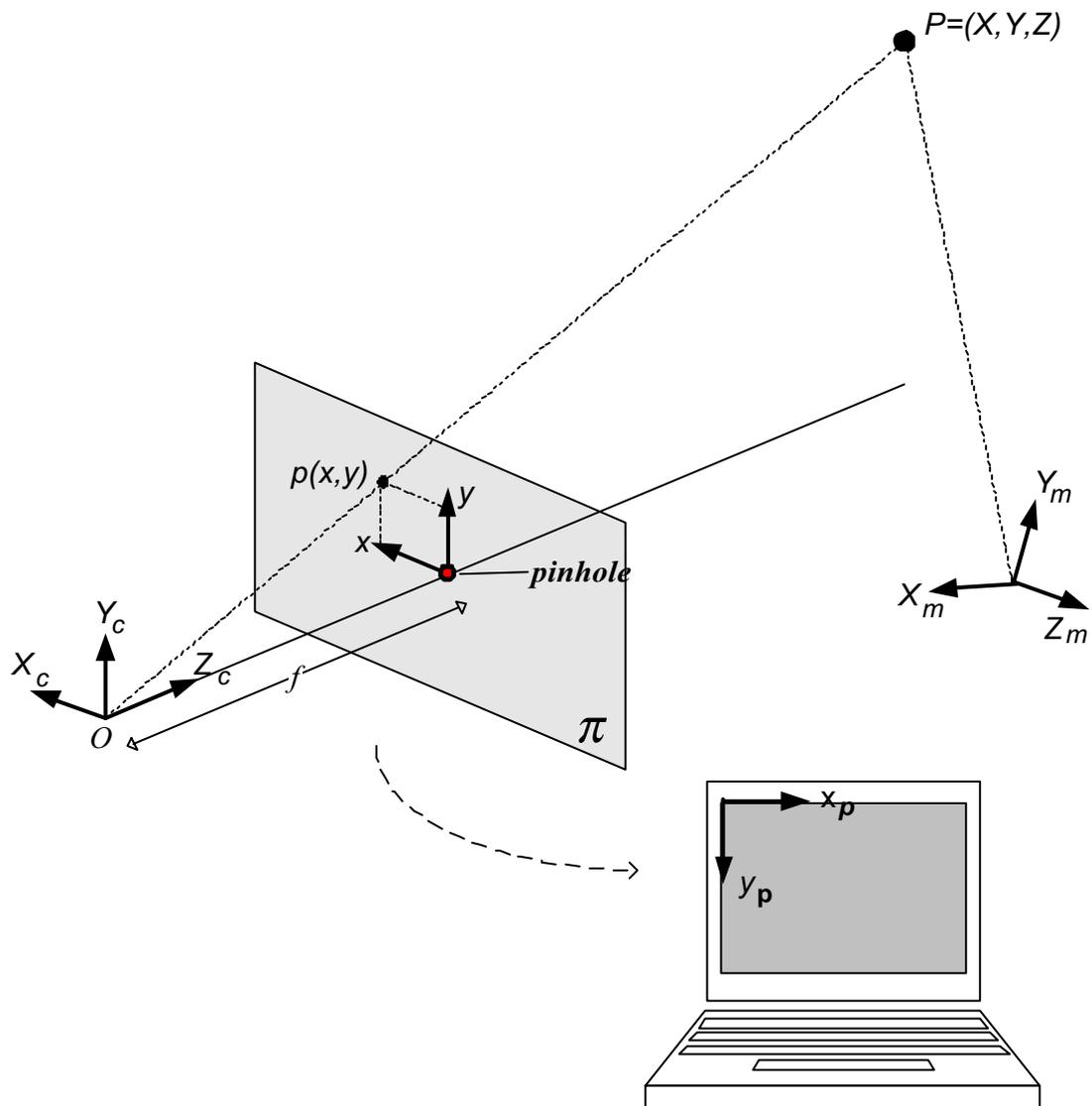


Figura 3.4 - Sistemas de coordenadas de um sistema de aquisição.

Os referenciais identificados na Figura 3.4, bem como a dimensão de cada espaço, a base correspondente e a descrição são sumarizados na Tabela 3.1.

Tabela 3.1 - Sistemas de coordenadas identificados na Figura 3.4.

Sistema de coordenadas	Dimensão	Base	Descrição
da cena, ou absoluto	tridimensional	$(X_m, Y_m, Z_m)$	É o sistema de referência utilizado para descrever pontos na cena ou mundo real.
da câmera	tridimensional	$(X_c, Y_c, Z_c)$	Sistema de referência com origem no centro óptico da câmera.
de imagem	bidimensional	$(x', y')$	Descreve pontos projetados no plano de formação da imagem
em pixel	bidimensional	$(x', y')$	Sistema de coordenadas utilizado para localizar pixels de imagens. As coordenadas são inteiras. Em geral, programas de processamento gráfico consideram a origem desse sistema no canto superior esquerdo. O processamento de imagens geralmente é feito nesse sistema de coordenadas.

### 3.4 - PARÂMETROS GEOMÉTRICOS DE CÂMERA

Parâmetros geométricos são especificações que determinam a geometria da câmera.

#### 3.4.1 - Parâmetros intrínsecos

Determinam a geometria interna da câmera e consistem do comprimento focal, tamanho dos pixels e posição do ponto principal.

### 3.4.2 - Parâmetros extrínsecos

Parâmetros extrínsecos determinam a posição e orientação do sistema de referência da câmera com relação ao sistema de coordenadas do mundo. Um vetor de translação fornece a posição espacial, e uma matriz de rotação fornece a orientação angular espacial da câmera.

### 3.5 - MUDANÇA SISTEMA DE COORDENADAS

Sempre que mais de um referencial está em jogo, torna-se evidente o problema de se efetuar a transição de coordenadas entre referenciais. Essa é a essência de problemas de alinhamento e calibração. Para mudar de sistema de coordenadas é necessário que uma transformação geométrica seja aplicada a um ponto ou vetor, como ficou evidenciado com as projeções. Entretanto, ao contrário das projeções, existem transformações que efetuam mapeamentos de pontos entre espaços de mesma dimensão e não alteram a forma de um objeto, preservando algebricamente a norma e o tamanho de um vetor, e por isso são conhecidas como *transformações de corpos rígidos*. Essas transformações são: a *translação*, que efetua mudança de posição e é representada por um vetor de translação.

$$T = [T_X, T_Y, T_Z]^T \quad (3.4)$$

e a *rotação*, que efetua mudança de orientação e é representada por uma matriz de rotação de ordem 3.

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (3.5)$$

Considerando o P do espaço na Figura 3.5, ele pode ser descrito em dois sistemas de coordenadas ortogonais: no sistema de coordenadas da cena ou do mundo com coordenadas  $(X_m, Y_m, Z_m)$ , ou no sistema de coordenadas da câmera, com coordenadas  $(X_c, Y_c, Z_c)$ . A transformação **T** do ponto *P* entre esses dois sistemas de coordenadas é determinada pela equação do movimento de corpos rígidos:

$$P_c = RP_m + T \quad (3.6)$$

onde:

$R$  é uma matriz de rotação;

$T$  é um vetor de translação.

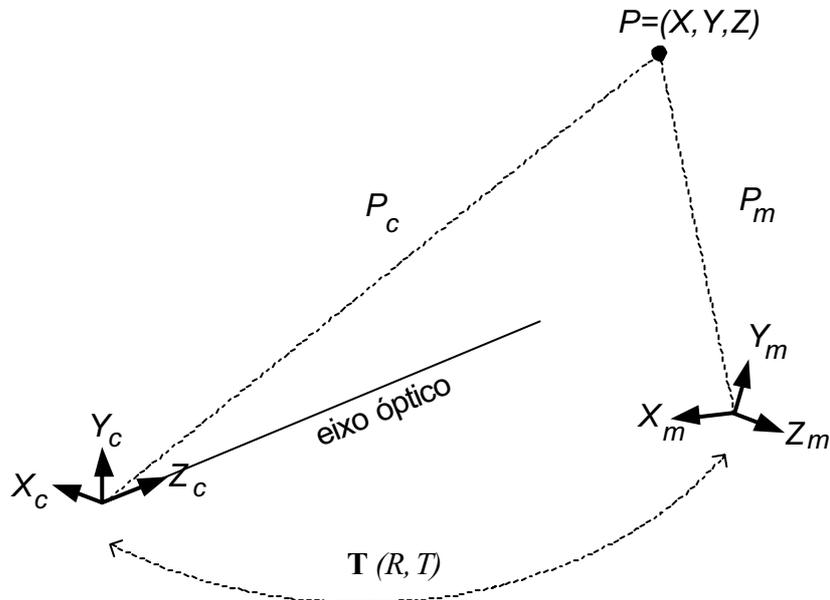


Figura 3.5 - Relação entre o sistema da câmera e do mundo.

### 3.6 - REPRESENTAÇÕES PARA A ROTAÇÃO

Já foram propostas variadas formas para representar rotações, mas 3 formas são especialmente populares em áreas como visão computacional, robótica e computação gráfica: representação com *ângulos de Euler*, utilizando um ângulo e um eixo arbitrário, e utilizando. A escolha da melhor forma de representação para uma dada aplicação depende de muitos aspectos particulares, sendo que a melhor opção para uma implementação em computação gráfica pode não ser a mais adequada para uso em um sistema físico em robótica, por exemplo. Essa transformação é representada por uma matriz com 3 graus de liberdade, que quando aplicada a um ponto transforma as coordenadas desse ponto em outro ponto descrito em outro sistema de coordenadas.

### 3.6.1 - Representação de rotações com ângulos de Euler

A representação com ângulos de Euler utiliza ângulos e eixos fixos. A transformação é composta pelas rotações  $\alpha$  em relação ao eixo  $X$ ,  $\beta$  em relação ao eixo  $Y$ , e  $\gamma$  em relação ao eixo  $Z$ , de um sistema de coordenadas. Existem muitas convenções para a utilização de ângulos de Euler. Para ordenar algebricamente a seqüência de ângulos e eixos dessa representação costuma-se empregar a *regra da mão direita*. Por essa regra, imagine-se que o eixo sendo segurado pela mão direita, e considera-se o sentido da rotação determinado pelo encolhimento dos dedos em volta do eixo, e o sentido anti-horário das rotações como positivo.

$$R(\alpha, \beta, \gamma) = R_X(\alpha)R_Y(\beta)R_Z(\gamma) \quad (3.7)$$

A matriz de rotação para ângulos de Euler em 3.7 apresenta a seguinte forma:

$$\begin{bmatrix} \cos(\beta)\cos(\gamma) & \sin(\alpha)\sin(\beta)\cos(\alpha)+\cos(\alpha)\sin(\gamma) & -\cos(\alpha)\sin(\beta)\cos(\gamma)+\sin(\alpha)\sin(\gamma) \\ -\cos(\beta)\sin(\gamma) & -\sin(\alpha)\sin(\beta)\sin(\gamma)+\cos(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\beta)\sin(\gamma)+\sin(\alpha)\cos(\gamma) \\ \sin(\beta) & -\sin(\alpha)\cos(\beta) & \cos(\alpha)\cos(\beta) \end{bmatrix} \quad (3.8)$$

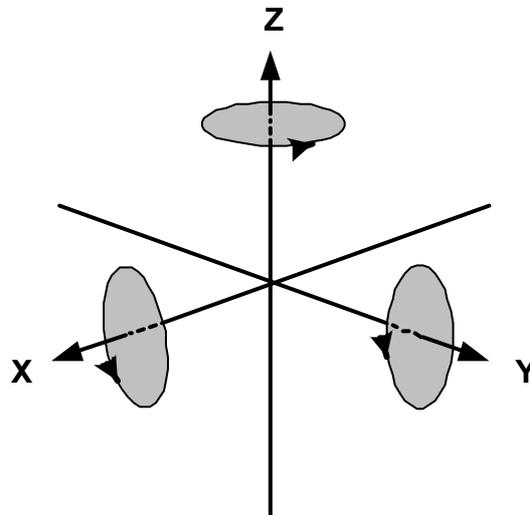


Figura 3.6 - Rotação utilizando ângulos de Euler.

### 3.6.2 - Representação de rotações com um ângulo e um eixo arbitrário

Uma série de rotações também pode ser representada por uma única rotação em relação a um eixo arbitrário. Qualquer rotação pode ser especificada por um ângulo  $\theta$  e um eixo de rotação arbitrário  $n = [n_1, n_2, n_3]^T$ . O primeiro especifica o sentido e a quantidade de rotação, e o segundo especifica a direção de rotação.

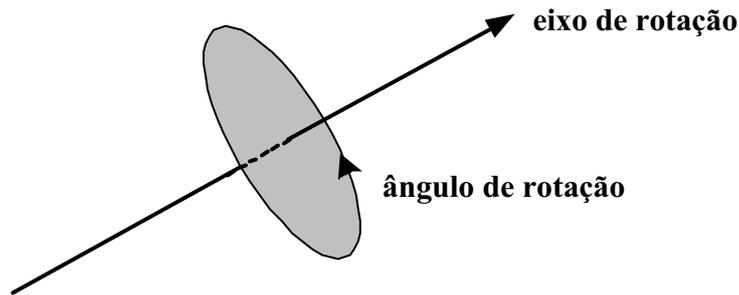


Figura 3.7 - Rotação utilizando um ângulo e um eixo arbitrário.

### 3.6.3 - Representação de rotações utilizando quaternions

Quaternions estão para a rotação no espaço tridimensional como os números complexos estão para a rotação no plano. Um quaternion  $\mathbf{q}$  apresenta 4 elementos e 3 graus de liberdade, na forma:

$$\mathbf{q} = [q_0, q_1, q_2, q_3]^T \quad (3.9)$$

$$\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \quad (3.10)$$

Onde:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1 \quad (3.11)$$

Seja o eixo  $n$  da representação eixo-ângulo e um ponto  $P = [X, Y, Z]^T$  qualquer no espaço.

A representação na forma de quaternions num referencial  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  fica:

$$\mathbf{q} = \cos(\theta/2) + (n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k})\text{sen}(\theta/2) \quad (3.12)$$

$$P = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k} \quad (3.13)$$

Uma rotação do ponto  $P$  para uma nova orientação  $P_1$  é dada por:

$$P_1 = \mathbf{q} P \mathbf{q}^* \quad (3.14)$$

onde:

$\mathbf{q}^*$  é o conjugado de um quaternion

### 3.7 - REPRESENTAÇÃO DE TRANSLAÇÕES

Como já foi mencionado, a representação de translações é bastante simples, bastando para tal apenas um vetor. Define-se a translação no espaço cartesiano tridimensional de um ponto  $P_1 = [X_1, Y_1, Z_1]^T$  para uma nova posição  $P_2 = [X_2, Y_2, Z_2]^T$ , como o deslocamento do ponto  $P_1$  quantificado e direcionado por um vetor de translação  $T = [T_X, T_Y, T_Z]^T$ . Isto é:

$$P_2 = P_1 + T \quad (3.15)$$

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \quad (3.16)$$

### 3.8 - COORDENADAS HOMOGÊNEAS

Em situações práticas, principalmente em implementações computacionais, é comum a necessidade de se manipular mais de uma transformação, por isso torna-se mais conveniente uma estrutura compacta e unificada para representá-las. A representação conhecida como *coordenadas homogêneas* viabiliza esse tipo de tratamento. Por esse meio é possível concatenar-se várias matrizes de transformação tais como rotação, translação, escala e projeções em apenas uma matriz de transformação, bastando para isso multiplicar as respectivas matrizes de modo a obter uma única matriz de transformação. Ao invés de se utilizar 3 matrizes de rotação, 3 de translação, por exemplo, pode-se obter uma única matriz de transformação que combina e efetua o trabalho das 6. Para isso é necessário elevar a ordem das matrizes de 1 unidade para que o vetor de translação se torne uma matriz quadrada. Salienta-se que matrizes, em geral, não comutam e, portanto, a ordem do

produto pode alterar a transformação resultante. As transformações de rotação e translação já vistas combinadas em uma única matriz A em coordenadas homogêneas, fica:

$$A = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_X \\ r_{21} & r_{22} & r_{23} & T_Y \\ r_{31} & r_{32} & r_{33} & T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.17)$$

### 3.9 - ASPECTOS DE CALIBRAÇÃO DE CÂMERAS

O problema da calibração de câmeras consiste em se estimar os valores dos parâmetros intrínsecos e extrínsecos de câmera. Um sistema de bem calibrado é tão importante para se extrair medidas em fotogrametria quanto para se construir um sistema de visão para reconstrução. A base desse conceito é a mesma por trás do problema do registro de imagens de profundidade. A caracterização dos problemas típicos encontrados na calibração de câmeras são listados na tabela abaixo.

Tabela 3.2 - Problemas de orientação em calibração de câmeras.

Natureza do problema	Caracterização
orientação exterior	Consiste em se determinar a posição e orientação de uma câmera em um sistema de coordenadas absoluto a partir de projeções de pontos de calibração na cena.
orientação interior	Consiste em se determinar a geometria interna de uma câmera, incluindo a constante de câmera, a localização do ponto principal e correções para distorção de lentes.
orientação absoluta	Consiste em se determinar a transformação entre dois sistemas de coordenadas ou a posição e orientação de uma câmera em um sistema de coordenadas absoluto a partir de coordenadas de calibração na cena.
orientação relativa	Consiste em se determinar a posição e orientação entre duas câmeras a partir de projeções de pontos de calibração da cena.

### **3.10 - A PERCEPÇÃO DE PROFUNDIDADE E A INFORMAÇÃO 3D ABSOLUTA**

Toda pessoa tem uma noção intuitiva do que significa dizer que algo ou alguma forma gráfica é tridimensional, embora muitas vezes não saiba explicar precisamente essa percepção em palavras. Tecnicamente, o mundo real é tridimensional porque para descrever qualquer ponto são necessárias três coordenadas em relação a um referencial.

Considerando que, tanto os olhos humanos quanto uma câmera convencional de intensidade são sensores 2D, e que o mundo real se projeta com a perda de uma dimensão tanto sobre a retina humana quanto sobre o plano de imagem de uma câmera, é inevitável o questionamento a respeito de como a visão humana consegue inferir a tridimensionalidade, dispondo apenas de imagens bidimensionais. Tanto para a visão humana quanto para a visão computacional isso é um complexo problema de percepção 3D, o qual define-se como o problema de recuperar a dimensão de profundidade perdida, a partir de uma ou múltiplas imagens (Shah, 1997). É tão espantosa a perfeição com a qual a visão humana consegue resolver esse problema que pode levar ao equívoco de se achar que é um problema de fácil conceituação matemática e a implementação física e computacional. Entretanto, resolver o problema da visão teoricamente e implementá-la com eficiência similar ao da visão humana é uma meta ainda muito distante de ser alcançada, devido em grande parte ao fato de ser ainda bastante desconhecida a forma como o cérebro trabalha o problema da visão. Existem muitas dúvidas e questionamentos em aberto, que somente muitas pesquisas envolvendo diferentes áreas podem levar a uma melhor compreensão do problema.

O fato é que a perda da informação 3D é uma perda severa para qualquer sistema de visão, tanto biológico quanto artificial. É uma informação tão valiosa que precisa ser recuperada ou estimada. Em essência, para recuperar a estrutura 3D completa de um objeto é, necessita-se no nível mais elementar, da recuperação da informação que falta em cada ponto: a informação 3D, que nada mais é que uma informação de distância entre um ponto na cena e um visualizador. Essa coordenada é abolida devido ao tipo de transformação: a projeção de perspectiva. No caso da visão humana, devido ao fato dos olhos serem sensores passivos 2D, não é possível medir diretamente a informação de profundidade; e o cérebro precisa usar alternativamente uma série de artifícios, emulados pela visão computacional, para recuperar a estrutura tridimensional de uma cena. Esses artifícios

exploram dicas ou pistas abundantes em imagens de intensidade para se obter a forma 3D. Algumas dessas dicas são: textura, movimento, contorno e sombreado, dentre varias outras. Essa abordagem exige cálculos para se obter a informação de profundidade ou a forma 3D que se encontra implícita. Exemplos de técnicas nesse sentido são: forma a partir de sombreado, para uma única imagem; triangulação estereoscópica, para pares de imagens; e fluxo óptico e métodos de fatorização para fluxos de vídeo (Curless, 1997). Essas técnicas são tratadas na seqüência deste capítulo.

### 3.11 - TIPOS DE IMAGENS DIGITAIS

Uma imagem digital é uma matriz bidimensional de números (Trucco e Verri, 1998). Cada célula dessa matriz, ou elemento atômico de uma imagem, é chamado de pixel; palavra esta que é um acrônimo de *picture element* (elemento de figura). Imagens podem ser adquiridas da cena e armazenadas em uma superfície fotossensível como um *filme fotográfico*, ou numa matriz de fotossensores eletrônicos tipo *CDD* ou *CMOS*, além de também poderem ser sintetizadas em computador. Quanto à natureza da informação que o pixel armazena, as imagens digitais podem ser de dois tipos: *imagens de intensidade*, que codificam intensidade luminosa; e *imagens de profundidade*, que codificam medições de distâncias.

#### 3.11.1 - Imagens de intensidade

Imagens digitais de intensidade são matrizes bidimensionais que codificam em cada pixel um valor de uma função inteira de intensidade luminosa que depende da iluminação da cena e das propriedades de refletância das superfícies de objetos (Gonzalez e Woods, 2000):

$$F(x,y) = I(x,y) * R(x,y) \quad (3.18)$$

Onde  $1 \leq x \leq m$  e  $1 \leq y \leq n$ .

Na forma matricial tem-se:

$$F = \begin{bmatrix} f(1,1) & \dots & f(1,n) \\ \vdots & \vdots & \vdots \\ f(m,1) & \dots & f(m,n) \end{bmatrix} \quad (3.19)$$

### 3.11.2 - Imagens de profundidade

Uma *imagem de profundidade* é uma imagem digital onde cada pixel ao invés de armazenar um valor de intensidade luminosa guarda uma informação de profundidade (Figura 3.8). Essa informação é, mais especificamente, um valor de distância de um ponto visível na cena a um sistema de referência conhecido do sensor. Outras denominações também encontradas para essas imagens são: *perfis de superfície*, *mapas de profundidade*, *mapas xyz* (nuvem de pontos), *imagens  $r_{ij}$*  (onde os valores de profundidade são regularmente espaçados nas direções dos eixos da imagem), e *imagens 2.5D*, (Jain *et al.*, 1995; Trucco e Verri, 1998). Em virtude dessas imagens guardarem a estrutura geométrica 3D de uma cena na forma de informação 3D explícita, pura, a recuperação desse dado não é necessária, o que é uma característica muito desejável para muitas aplicações de visão 3D.

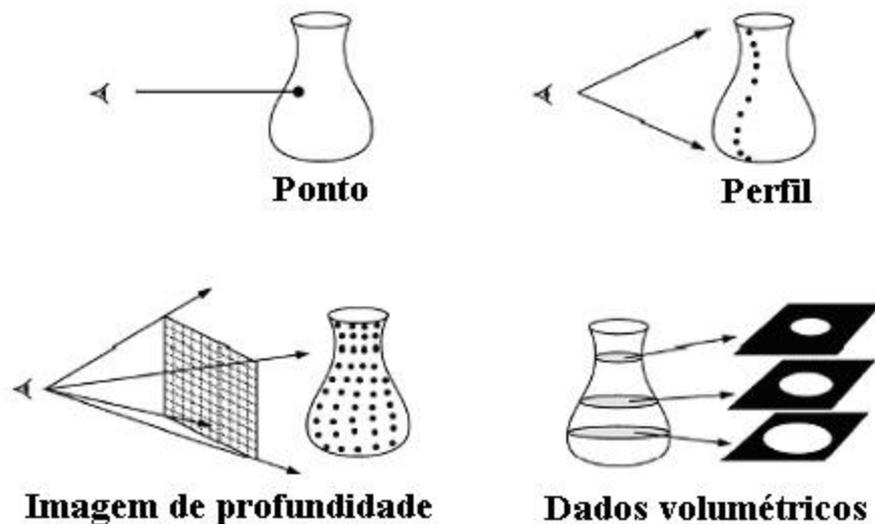


Figura 3.8 - Estrutura de dados (modificado - Curless, 2000).

Existem ainda imagens que possuem a capacidade de codificar estruturas volumetricamente, na forma de seções bidimensionais (Figura 3.8); um exemplo são as imagens tomográficas, a quais não são tridimensionais inerentemente, mas permitem a reconstrução de um modelo utilizando métodos adequados (Jähne, 2002).

### 3.12 TAXONOMIA DOS MÉTODOS DE AQUISIÇÃO DE FORMA 3D E PROFUNDIDADE ABSOLUTA

Um dos atrativos que está popularizando a modelagem por reconstrução na atualidade é a quantidade de digitalizadores 3D disponíveis no mercado, em comparação com alguns anos atrás. A variedade de métodos de aquisição de imagens existente, seja de intensidade ou de profundidade, é tão vasta e complexa que compõe um campo de pesquisa a parte. A capacidade de digitalização inclui dimensões díspares como as presentes no nível microscópico e no âmbito do sensoriamento remoto. Os digitalizadores utilizam diferentes métodos de sensoriamento para captar a distância do sensor a pontos na cena, como toque, envio de sondas ópticas, sondas acústicas, dentre outras. A escolha do digitalizador mais adequado para a aplicação depende de vários fatores como: o tamanho do objeto, características da superfície do objeto e custos associados ao projeto. Digitalizadores que utilizam iluminação coerente, tal como laser de baixa intensidade, são bastante atrativos, versáteis e produzem mapas de profundidade com rapidez e alta taxa de amostragem. Em aplicações como robótica móvel, radares baseados em laser são a solução mais sofisticada para modelagem 3D de ambientes. Na indústria em geral os digitalizadores mais populares são os baseados em triangulação óptica com iluminação ativa.

Como mencionado anteriormente, não é sempre que a informação 3D é recuperada na forma explícita. Digitalizadores 3D tem a capacidade de retornar medições diretas, mas técnicas que utilizam imagens de intensidade retornam muitas vezes apenas informações sobre a estrutura 3D na forma de orientações de superfície.

Já foram propostas algumas classificações que englobam o conjunto de técnicas para se obter a informação de profundidade direta ou indiretamente (Varady *et al.*, 1997; Curless, 1997). Mas em geral os pesquisadores da área classificam as técnicas em *passivas* e *ativas*. Abordagens passivas não interagem com o objeto, enquanto métodos ativos fazem contato com o objeto ou projetam alguma forma de energia sobre ele. Outras duas categorias bem gerais também podem ser identificadas: métodos de aquisição de forma *com contato*, e *sem contato* (Curless, 1997). Na prática, o custo dos equipamentos dos métodos ativos é superior ao dos métodos passivos; entretanto, os métodos passivos exigem softwares mais complexos para reconstrução, o que acaba encarecendo também os sistemas que usam esses métodos.

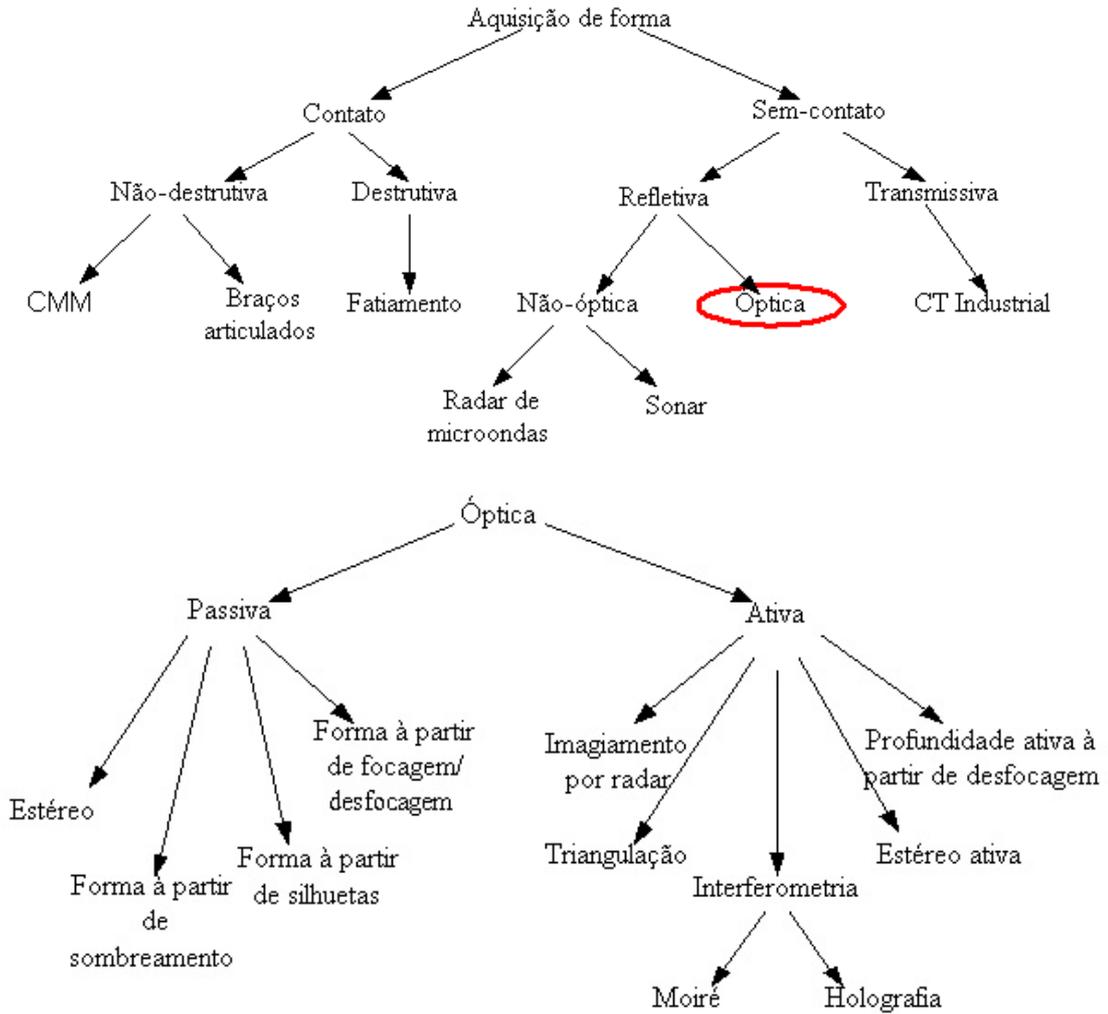


Figura 3.9 - Taxonomia de Curless para os métodos de aquisição de forma 3D (modificado - Curless, 2000).

### 3.13 - PRINCÍPIOS DE IMAGEAMENTO DE PROFUNDIDADE

#### 3.13.1 - Medição de profundidade pelo princípio TOF

Digitalizadores que utilizam o princípio TOF (*time-of-flight*) medem profundidade avaliando as variáveis envolvidas na fórmula da distância percorrida.

$$d = v.t \tag{3.20}$$

onde:

$v$  é a velocidade de propagação do pulso;

$t$  é o tempo de percurso do pulso.

A variável *velocidade* pode ser especificada como a velocidade do som no ar ou como a velocidade da luz. A variável *tempo* é determinada fazendo-se a leitura do tempo gasto pelo pulso no percurso entre a sua emissão, a sua reflexão pela superfície do objeto, e o seu retorno ao receptor do dispositivo (Borenstein *et al.*, 1996). O pulso emitido pode ser ultrassônico, de radiofrequência (RF), ou óptico. A profundidade absoluta é obtida dividindo-se a distância percorrida por dois.

### 3.13.2 - Medição profundidade pelo princípio da triangulação

Métodos de triangulação, tanto *passivos* quanto *ativos*, exploram relações de proporção trigonométricas básicas existentes em triângulos. Pela *lei dos senos*, se o comprimento de um lado e dois ângulos interiores de um triângulo são conhecidos, essas informações são suficientes para se determinar os comprimentos dos outros dois lados e o ângulo restante (Conrad e Sampson, 1990).

### 3.13.3 - Medição de profundidade explorando mudança de fase ou frequência

Digitalizadores que utilizam técnicas de mudança de fase ou frequência emitem uma onda contínua de amplitude ou frequência modulada ao invés de pulsos, e detectam a mudança de fase ou frequência do sinal refletido, a qual é proporcional à distância ao objeto.

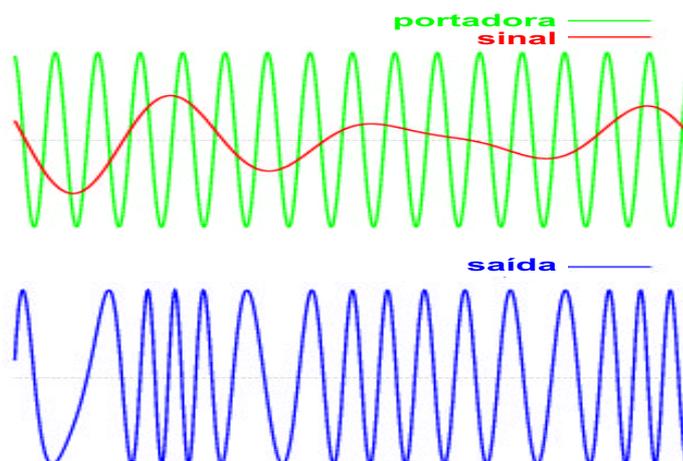


Figura 3.10 - Modulação em frequência (modificado - Wikipédia, 2006).

### 3.14 - MÉTODOS DE AQUISIÇÃO QUE EFETUAM CONTATO

Métodos de contato capturam parâmetros de forma efetuando contato físico direto com a superfície do objeto, ou se aproximando do objeto e fazendo a leitura com um dispositivo óptico de leitura acoplado. Ou seja, nem sempre precisam necessariamente tocar fisicamente o objeto (Conrad e Sampson, 1990). Sensores tácteis e de toque são os dispositivos usados para medir os parâmetros do contato entre o digitalizador e a superfície do objeto. Uma matriz de sensoriamento táctil pode ser considerado um grupo de sensores de toque.

Digitalizadores de contato têm emprego amplo e de longa data na engenharia, e estão longe de caírem em desuso. Esses instrumentos têm a virtude de serem altamente precisos, precisão submilimétrica muitas vezes, e até no nível atômico (abaixo de 0,1 nm), no caso de técnicas como *SEM* e *STM* (Sinha e Jain, 1994). São tradicionalmente empregados na engenharia, em tarefas como *engenharia reversa* de modelos *CAD*, *metrologia* ou *inspeção dimensional*. A meta deste último é determinar se um modelo sólido de uma peça manufaturada reúne suas especificações de projeto (Spitz, 1999).

As desvantagens dos digitalizadores de contato são a lentidão, o fato de precisarem tocar a superfície de objetos, o que não é recomendável em muitas aplicações, e principalmente, o fato de gerarem dados com baixa resolução de amostragem (esparsos) e a uma baixa taxa de amostragem. Exemplos desses digitalizadores são os *CMM* (*Coordinate Measuring Machines*) (Figura 3.11), e os braços articulados (Figura 3.12).

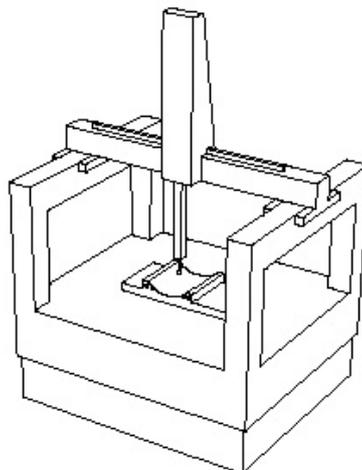


Figura 3.11 - Uma típica máquina de medição de coordenadas (*CMM*) (Spitz e Requicha, 2000).

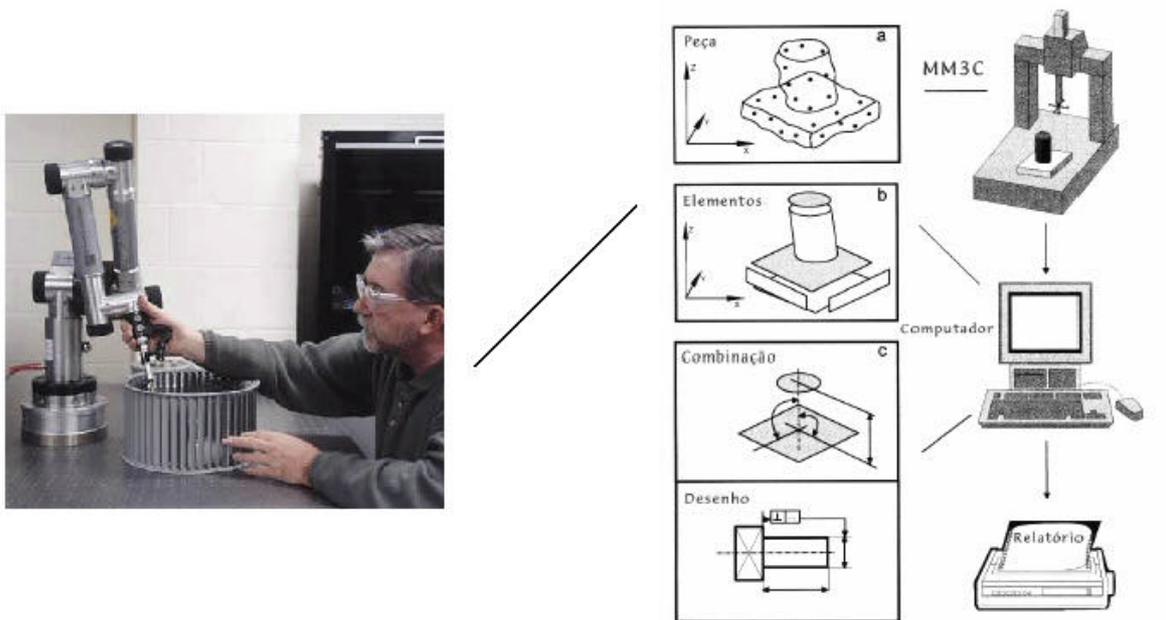


Figura 3.12 - Braço Articulado.

### 3.15 - MÉTODOS DE AQUISIÇÃO QUE NÃO EFETUAM CONTATO

Métodos de aquisição *sem contato* são aqueles que exploram uma energia existente no ambiente ou projetam alguma forma de energia seguida pela gravação da energia *refletida* ou *transmitida*. Esses métodos podem ser tanto *ópticos* como *não-ópticos*.

#### 3.15.1 - Tomografia Computadorizada (CT)

A tomografia computadorizada é um exemplo de método de aquisição de *dados volumétricos* por abordagem transmissiva usada principalmente em aplicações médicas.

#### 3.15.2 - Radares de microondas e sonares

Radares e sonares são dois instrumentos de sensoriamento que fazem uso dos princípios de mudança de fase ou frequência e *TOF*, respectivamente.

Radares de microondas são instrumentos muito empregados em aplicações que trabalham com imageamento de longa distância, em áreas como sensoriamento remoto, embora o uso de radar óptico de curta profundidade também seja factível.

Sonares são vastamente empregados como sensores embarcados em robótica móvel, devido principalmente ao seu menor custo em relação a outros sensores para essa proposta. Entretanto, eles são pouco recomendados para a captura de geometria 3D de objetos ou ambientes por retornarem dados pouco precisos e muito ruidosos. Uma etapa de filtragem dos dados retornados por esses sensores é altamente recomendável.

### 3.16 - MÉTODOS ÓPTICOS PASSIVOS

As técnicas ópticas passivas obtêm a informação 3D indiretamente, a partir da análise de imagens de intensidade. Isso é possível porque essas imagens são muito ricas em informações ou dicas visuais que levam a profundidade absoluta ou orientações de superfície. Um conjunto de técnicas conhecidas como *forma a partir de X*, onde o *X* representa um conjunto de técnicas como sombreamento, movimento, textura, e outras; estimam orientação de superfície local preferivelmente que profundidade absoluta (Jain *et al.*, 1995). Comparativamente a métodos ativos, essas técnicas exigem algoritmos complexos ao invés de equipamentos de imageamento especializado, um raro exemplo são câmeras autofocus. A seguir são apresentadas algumas das diversas técnicas ópticas passivas.

#### 3.16.1 - Estereoscopia

O imageamento estéreo ou por *triangulação passiva* é uma técnica muito usada em visão computacional e que emprega uma técnica crucial para a visão humana recuperar profundidade. Nessa técnica são utilizadas tipicamente duas câmeras de intensidade *calibradas*. A informação de profundidade de um ponto é extraída por meio de triangulação. A estimativa de profundidade de um ponto na cena é resolvido quando são conhecidos a *linha de base* (distância entre os centros das câmeras) e a *disparidade* (distância entre os pontos correspondentes quando as imagens são sobrepostas) (Figura 3.13). A linha de base está disponível se o sistema está calibrado. Entretanto, encontrar a disparidade exige o tratamento do difícil problema da correspondência, que pode ser entendido como: dado um pixel em uma imagem, encontrar o seu par correspondente na outra imagem, sendo que os dois correspondem ao mesmo ponto na cena. A escolha da linha de base deve levar em conta que, quanto mais longa, mais exata a estimativa, e quanto mais curta, mais confiavelmente os pixels podem ser combinados (Pulli, 1997).

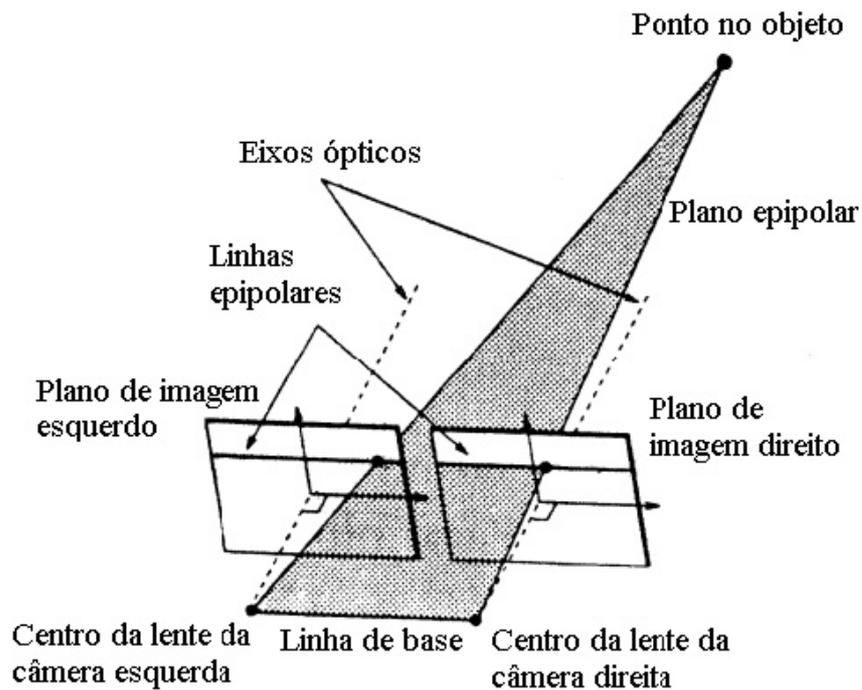


Figura 3.13 - Geometria da estereoscopia (modificado - Jain *et al.*, 1995).

Os problemas com mapas de profundidade obtidos por estereoscopia estão associados com a baixa taxa de amostragem, a necessidade de cálculos para recuperar a informação 3D e a imprecisão na estimação de profundidades absolutas. Essas características podem representar um grande empecilho para diversas aplicações, como navegação robótica e engenharia reversa. Por outro lado, é um assunto bastante estudado, é uma técnica de baixo custo, e a informação de cor já está disponível. A confiabilidade vem sendo aumentada através da utilização de mais de duas câmeras.

### 3.16.2 - Focagem e desfocagem ativa

Profundidade ou forma a partir de *desfocagem* é o problema de reconstruir o mapa de profundidade de uma cena, dado um conjunto de imagens coletadas a partir de uma mesma câmera, mas com diferentes configurações ópticas (comprimento focal, por exemplo).

Profundidade a partir de *focagem* por sua vez é o problema de reconstruir a profundidade da cena mudando ativamente o conjunto óptico da câmera até que o ponto de interesse esteja em foco. Sabendo que o ponto está em foco, a informação sobre a sua profundidade

é obtida através da lei de lentes finas. Profundidade a partir de desfocagem é dividida em passiva e ativa. No primeiro caso não há intervenção na cena, enquanto no segundo a idéia é artificialmente criar textura sobre a superfície por meio de iluminação estruturada (Favaro, 2000).

### **3.16.3 - Forma a partir de sombreamento**

No imageamento projetivo a profundidade absoluta é perdida, mas a orientação das superfícies da cena pode ser recuperada através das mudanças das intensidades dos pixels na imagem.

Forma a partir de sombreamento trata com a recuperação de forma a partir de uma variação gradual do sombreamento na imagem (Zhang *et al.*, 1999). A orientação em questão refere-se a aspectos como normais de superfície, gradiente e inclinação de superfície. A normal de superfície, em particular, pode ser recuperada a partir do sombreamento porque a radiância de superfícies depende do ângulo de incidência da fonte de iluminação (Jähne, 2002).

### **3.16.4 - Forma a partir de movimento**

Movimento é uma importante dica visual para se inferir a estrutura 3D de objetos. A obtenção de forma 3D por essa abordagem é considerada uma extensão da estereoscopia para uma seqüência de imagens (Jain *et al.*, 1995).

As similaridades entre esses dois métodos estão no fato de utilizarem parâmetros extrínsecos de câmera e na exploração de correspondências entre imagens. As diferenças estão nas disparidades, que são bem menores no movimento de câmera devido a altas taxas de aquisição de quadros, e no deslocamento relativo 3D que não é necessariamente causada por uma única transformação rígida 3D (Trucco e Verri, 1998).

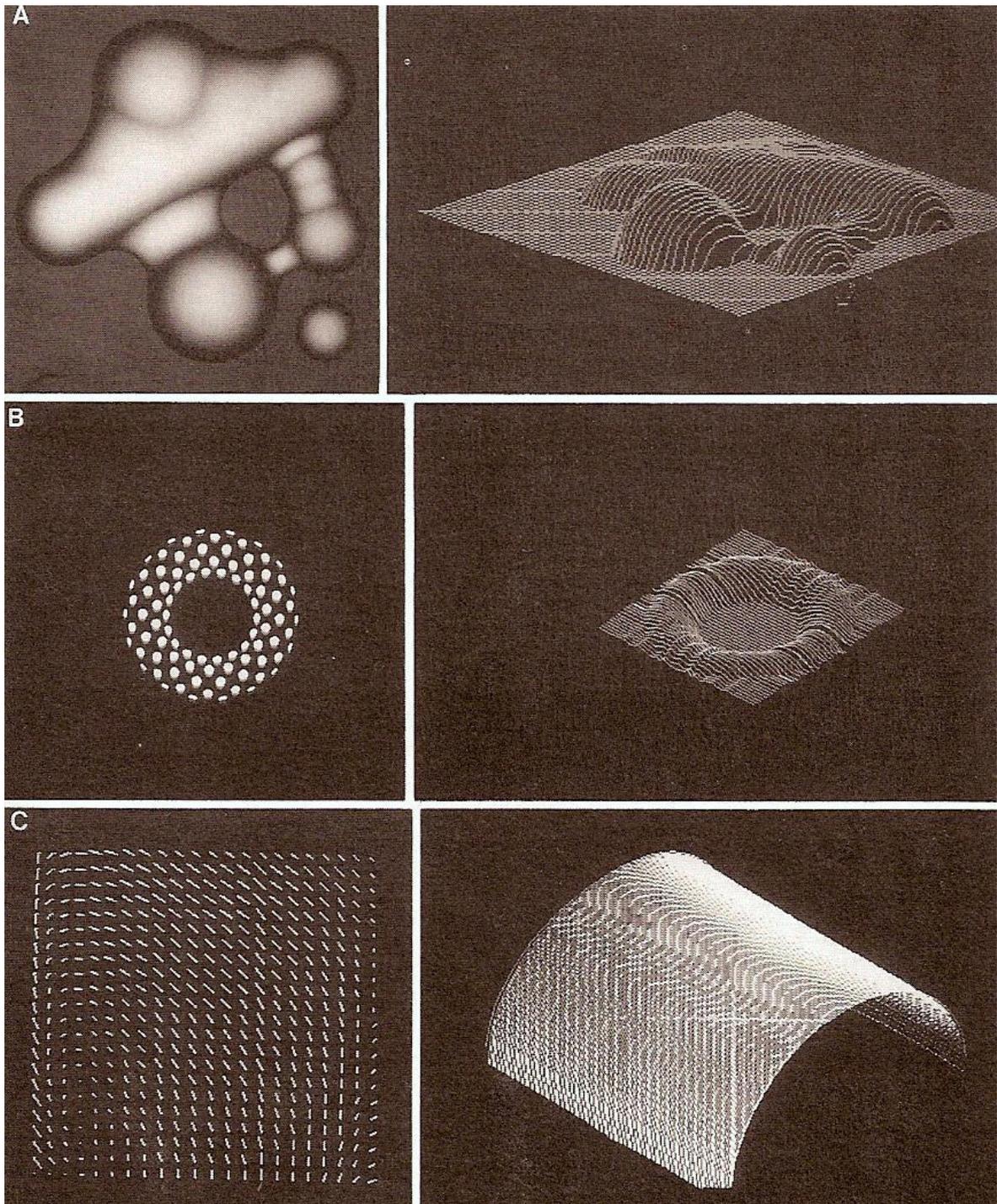


Figura 3.14 - Recuperação de geometria de superfície a partir de dicas em imagens: (a) a partir de sombreamento; (b) a partir de padrões; (c) a partir de movimento. A coluna a esquerda mostra a imagem (ou campo de movimento da imagem); a coluna a direita mostra a reconstrução de superfície (modificado - Aloimonos e Rosenfeld, 1993).

### 3.17 - MÉTODOS ÓPTICOS ATIVOS

Técnicas ópticas ativas obtém geometria de superfícies controlando a iluminação incidente sobre o ambiente ou objeto. Para isso são utilizadas sondas de energia que atingem a superfície do objeto. Em virtude dessa interação, as técnicas ativas são altamente afetadas por propriedades de superfície, como refletância, transparência, brilho, textura, cores não refletidas. Dentre as técnicas ópticas ativas, as mais populares são as baseadas em *iluminação estruturada*, também conhecida como triangulação ativa; e imageamento por radar laser (*LADAR*).

#### 3.17.1 - Iluminação e propriedades do laser

Um elemento importante nas técnicas ópticas ativas é a iluminação. A iluminação pode ser feita tanto com luz monocromática, constituída por apenas uma cor, quanto por luz policromática, constituída por uma mistura de cores.

A utilização do laser para iluminação é preferível por causa das seguintes propriedades:

- a luz é *monocromática*, ou seja, ela contém um comprimento de onda específico de luz (uma cor específica);
- a luz é *coerente*, ou seja, ela é “organizada” , cada fóton se move juntamente com os outros. Isso significa que todos os fótons têm frentes de onda que são iniciadas em uníssono;
- a luz é bem *direcionada*, ou seja, tem um feixe muito estreito e é muito forte e concentrada (Weschler, 2006).

#### 3.17.2 - Iluminação estruturada

O imageamento por iluminação estruturada ou triangulação ativa funciona de modo semelhante a estereoscopia, entretanto, uma das câmeras é substituída por uma fonte de iluminação que emite um padrão de luz. Diversas técnicas podem ser empregadas na iluminação: pontos, padrão de pontos, folha de luz, etc. O tipo de luz mais recomendado,

no entanto, é o laser, em virtude da precisão. Os pixels da imagem que representam o perfil do objeto iluminado podem ser facilmente identificados, pois são mais brilhantes que os outros na imagem. Também podem ser utilizados filtros na câmera para esse propósito. Na configuração física, a câmera e o projetor de luz precisam estar calibrados. Precisam ser conhecidos o deslocamento entre a projetor de luz e o eixo focal da câmera, e o ângulo de incidência da luz (Figura 3.15). Dessa forma é possível obter a informação de profundidade por triangulação.

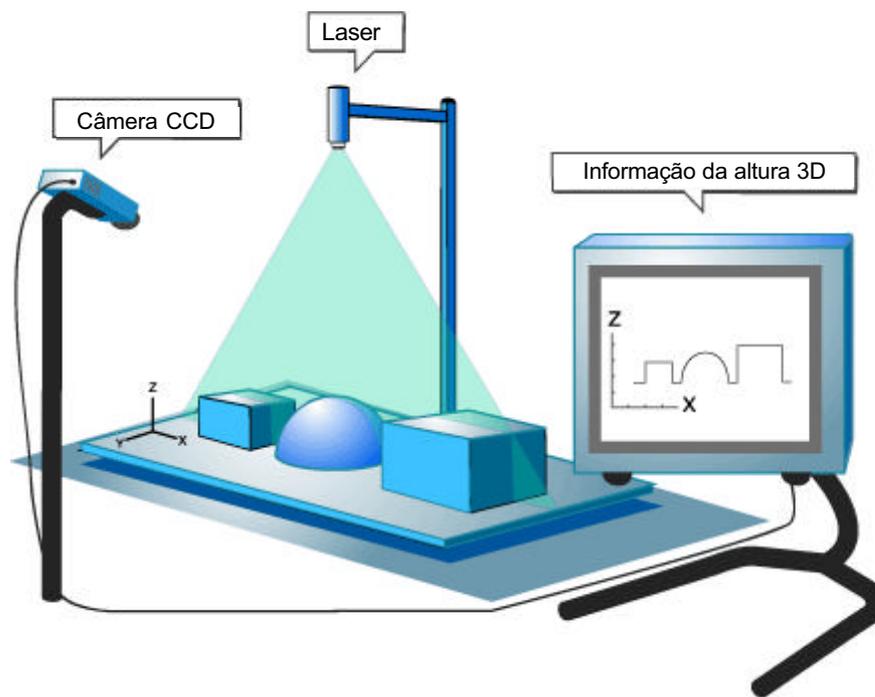


Figura 3.15 - Configuração de um digitalizador baseado em iluminação estruturada (modificado - Stockeryale, 2006).

A grande idéia por trás dos digitalizadores baseados em iluminação estruturada é a de substituir o difícil problema da correspondência encontrado na triangulação passiva por um simples problema de detecção. As desvantagens são a baixa velocidade de aquisição, devido à natureza seqüencial das aquisições, a oclusão de dados onde o ponto laser está oculto da câmera pelo próprio objeto, e dados perdidos ou errôneos devido a especularidade da superfície do objeto (Forsyth, e Ponce, 2003). Sistemas de iluminação estruturada são recomendados para aplicações de curta profundidade e são os mais utilizados em aplicações de visão nas indústrias.

### 3.17.3 - Radar baseado em laser (*Ladar*)

Para longa profundidade, que é uma característica freqüente em aplicações como navegação robótica ou fotogrametria, é bastante sugestiva a utilização de *ladares*. Um *LADAR* é um instrumento cujo sensoriamento utiliza o princípio *TOF*, de *tempo de percurso*, para recuperar profundidade (Schwarte, 1990). Uma grande vantagem desse tipo de sensor é a versatilidade de poder funcionar embarcado em um veículo móvel. No entanto, por utilizarem circuitos mais sofisticados, apresentam geralmente preço mais elevado comparado a outros sensores para a mesma proposta.

### 3.17.4 - Interferometria de Moiré

O padrão interferométrico de Moiré é obtido pela sobreposição de dois outros padrões. O primeiro consiste de uma franja de luz regularmente espaçada, produzida pela colocação de uma grelha na frente de um projetor de luz que ilumina o objeto. Não há restrição quanto ao uso de iluminação coerente ou não. A superfície do objeto modula a fase do padrão de luz, e a informação de profundidade pode ser extraída demodulando o padrão refletido. Isso é feito utilizando-se na câmera de aquisição um segundo padrão semelhante ao que foi projetado. A configuração consiste de um projetor, grelha e câmera calibrados. Variações da técnica dependem da forma de obtenção da franja, as quais podem ser produzidas colocando-se a grelha o mais próximo possível do objeto e cobrindo toda a extensão em frente ao projetor e a câmera, método que é chamado de *Moiré de sombra* (Coelho e Tavares, 2003; Vuylsteke *et al.*, 1990); alternativamente podem ser colocadas máscaras na câmera e no projetor: *Moiré de projeção* (Vuylsteke *et al.*, 1990).

Essa técnica apresenta boa resolução, entretanto, sofre do mesmo problema de extração de forma a partir de imagens de intensidade, é necessário processamento computacional para recuperar a informação 3D implícita. Para se obter distâncias absolutas entre as franjas é necessário ter uma distância de referência, caso contrário só se obtém distâncias relativas.

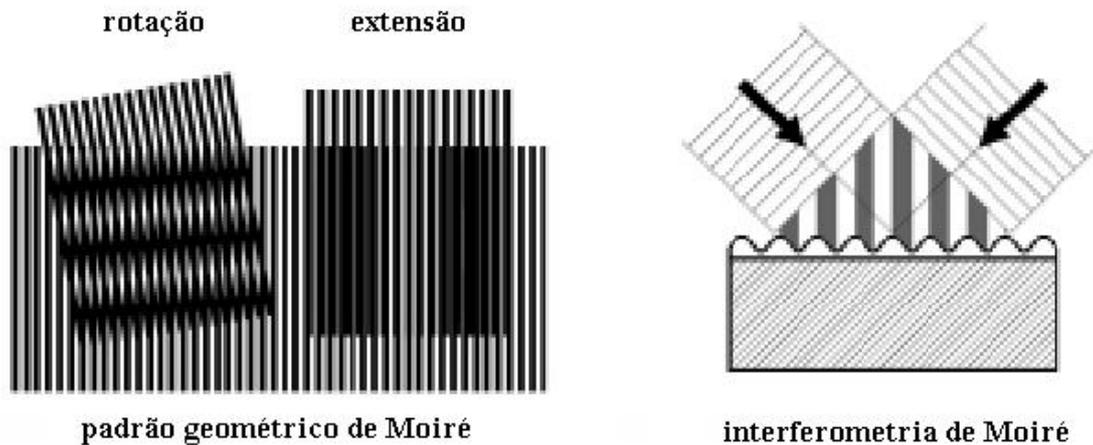


Figura 3.16 - Configuração básica do padrão geométrico e da interferometria de Moiré (modificado - Stockmann e Naumann, 2005).

### 3.18 - PLANEJAMENTO DOS PONTOS DE VISTA DO DIGITALIZADOR

O problema de se determinar as poses do digitalizador de modo a otimizar o processo de aquisição é popularmente conhecido como problema da *próxima melhor vista* (*Next Best View - NBV*) (Connolly, 1985; Pito e Bajcsy, 1995; Reed, 1998). Naturalmente, esse problema deve ser solucionado de modo automático, com uma abordagem algorítmica.

Planejar os pontos de vista do digitalizador é uma tarefa essencial para se agregar novas e importantes informações a cada nova varredura, prevenindo assim as indesejáveis oclusões e, conseqüentemente, problemas nas fases posteriores da reconstrução. Outra meta é minimizar o número de imagens de profundidade a serem adquiridas, tanto porque isso diminui o tempo de modelagem e poupa memória (Wong *et al.*, 1999; Banta *et al.*, 2000), quanto porque existe um custo financeiro associado à aquisição de cada imagem.

### 3.19 - INCERTEZAS EM IMAGENS DE PROFUNDIDADE

O ponto fraco das imagens de profundidade tem sido as incertezas nas medições. As quais podem estar presentes nos três eixos:  $X$ ,  $Y$  e  $Z$ . Outras desvantagens são:

- o modelo de projeção do sensor pode não ser totalmente de perspectiva, próximo das bordas costuma ser ortogonal, ocasionando erros nos eixos  $X$ ,  $Y$ . Esses erros geram

problemas de amostragem. O espaçamento desordenado entre os dados pode criar problemas para algoritmos alinhamento e geração de malha;

- erros no eixo Z podem ser aditivos ou compostos. Iluminação coerente tende a propiciar o temido ruído especular;
- a inexatidão de medições cresce proporcionalmente à distância dos pontos medidos. Quanto maior a profundidade, menos exata;
- a distribuição de erros em sensores 3D é heterocedástico e anisotrópico, caracterizando dados multiestruturados (Williams e Bennamoun, 2001).

Uma etapa de pré-processamento é importante para a atenuação de ruído, mas com o comprometimento de se evitar a perda de informações geométricas. O conhecimento sobre o modelo de ruído do sensor e da relação sinal/ruído podem ajudar na filtragem e na escolha do estimador de regressão na fase de alinhamento.

## CAPÍTULO 4 – REGISTRO DE IMAGENS DE PROFUNDIDADE

A construção de modelos digitais utilizando-se múltiplas imagens de profundidade da superfície de um objeto exige que uma inevitável etapa de alinhamento de dados 3D seja resolvida. Na comunidade da *visão computacional* esse problema de alinhamento é popularmente conhecido como *registro de imagens*. Este capítulo trata do detalhamento, e das abordagens para resolução desse problema.

### 4.1 - DEFINIÇÃO DO PROBLEMA DO REGISTRO DE IMAGENS DE PROFUNDIDADE

A necessidade do registro de imagens surge na tarefa da construção de modelos digitais quando são empregados na fase de aquisição de dados uma certa classe de digitalizadores 3D que só podem fazer a leitura do perfil do objeto a partir de uma única direção. Esse é o caso dos digitalizadores ópticos, e a direção é a que está visível para o sensor. A consequência dessa limitação é que somente uma porção de superfície pode ser adquirida a cada varredura, o que representa um sério entrave para a construção de modelos digitais completos, onde é necessária a captura de toda a superfície do objeto. Para contornar esse obstáculo a solução é efetuar diversas varreduras variando-se as poses entre o sensor e o objeto até que toda a superfície seja digitalizada (Faugeras, 1993; Hartley e Zisserman, 2000; Faugeras e Luong, 2004). Para alterar as poses, mudanças de orientação são feitas aplicando-se movimentos de *rotação*, e mudanças de posição são efetuados aplicando-se movimentos de *translação*. Como tratado no capítulo anterior, esses movimentos são conhecidos como *transformações de corpos rígidos*, e não alteram a forma do objeto na imagem. É indiferente o objeto permanecer parado e o sensor movimentar-se, ou vice-versa, ambos os procedimentos terão o mesmo efeito nas imagens tomadas. O movimento pode ser aplicado tanto ao objeto, rotacionando-o com uma plataforma giratória, quanto ao sensor, controlado-o com um manipulador robótico, por exemplo.

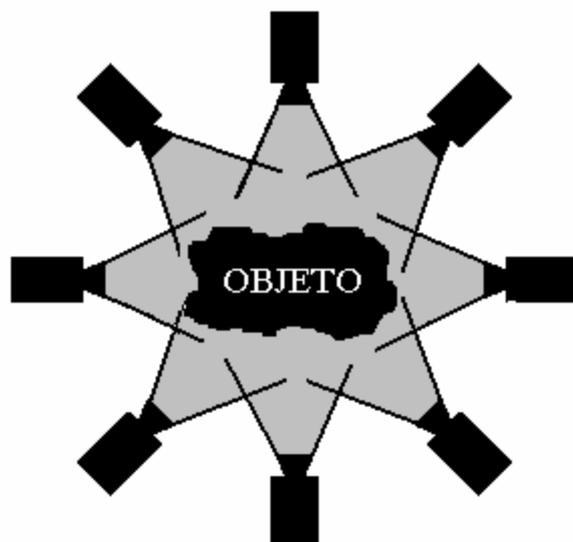


Figura 4.1 - Geometria do processo de aquisição por múltiplas vistas.

Pela Figura 4.1 percebe-se intuitivamente que a idéia de se formar um modelo completo adquirindo-se múltiplas imagens a partir de diferentes pontos de vista é análoga à montagem de um quebra-cabeça geométrico, onde informações contidas nas próprias peças precisam estar disponíveis para possibilitar o encaixe. Na reconstrução, a disponibilidade da informação sobre a correspondência entre imagens vizinhas permite uma melhor combinação das peças. Para atingir essa meta, em um sistema de aquisição calibrado, basta se tomar cuidado para que sejam criadas regiões de sobreposição entre imagens adjacentes. Com isso é possível “produzir” pontos correspondentes que servirão de guia para “arrastar” uma imagem para cima da outra. Esse procedimento é necessário porque não é possível a fusão imediata das imagens logo após o final do processo de aquisição. O resultado não seria um modelo, mas uma representação distorcida e inexata da realidade. Uma prova prática desse problema pode ser constatada pegando-se dois pontos correspondentes, um em cada imagem, esses pontos representam o mesmo ponto na superfície do objeto, mas não apresentam as três coordenadas exatamente iguais. Portanto, uma etapa de alinhamento ou registro precisa ser resolvida para que todas as imagens sejam casadas. A essência desse problema está na compreensão dos sistemas de coordenadas envolvidos na geometria do processo de aquisição de dados.

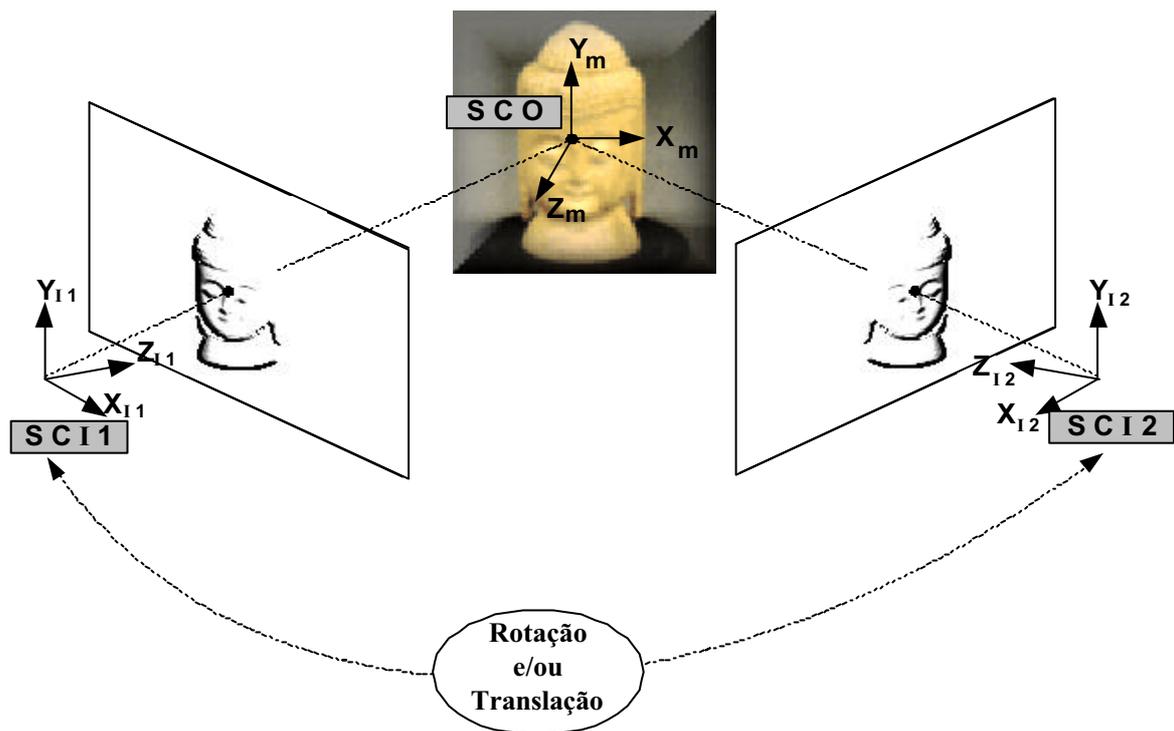


Figura 4.2 - Geometria da aquisição de duas vistas. O objeto encontra-se centrado no sistema de coordenadas SCO (sistema de coordenadas do objeto). Já as imagens, são tomadas centradas no sistema de coordenadas do sensor, SCI1 (sistema de coordenadas da imagem 1), e SCI2 (sistema de coordenadas da imagem 2).

Pode ser verificado na Figura 4.2 que o sensor tem seu próprio sistema de coordenadas, e que cada imagem é adquirida centrada nesse referencial. O objeto também tem o seu próprio referencial, no qual ele está centrado. Sabe-se que a relação entre dois sistemas referenciais fica determinada recuperando-se as transformações espaciais que efetuam o mapeamento dos pontos entre eles. Entretanto, entre duas imagens existe uma classe de transformações possíveis, e somente um par de parâmetros é o ótimo. Implementações computacionais podem enfrentar grandes desafios para recuperar esses parâmetros a partir dos dados das imagens. Por isso faz sentido definir o objetivo geral do problema em questão como o de trazer dois conjuntos de dados dentro do *melhor alinhamento possível*, e o objetivo específico do registro como o de *estimar os parâmetros das transformações espaciais* dentro desse processo de otimização.

## 4.2 - ALGUNS CONCEITOS IMPORTANTES

### 4.2.1 - Estimação de pose, combinação e correspondência

*Estimação de pose* é o processo de determinar a posição e orientação de um objeto descrito em um sistema de coordenadas em relação a outro sistema de coordenadas (Rosenhahn, 2003; Zhang, 1995).

O registro se preocupa em calcular, ou determinar explicitamente, os valores das transformações rígidas: rotação e translação, da estimação de pose.

O *problema da combinação de superfícies* consiste em se encontrar o registro espacial entre duas superfícies que maximiza suas similaridades de formas (Potmesil, 1983). O problema da combinação implica o método para se estabelecer correspondência.

O *problema da correspondência* consiste em se fazer associação ou determinar o mapeamento entre padrões presentes em imagens (Rosenhahn, 2003).

### 4.2.5 - Complexidade computacional

Um algoritmo é uma seqüência finita de instruções estritamente claras, que quando obedecidas, fornecem ao final um resultado para uma dada entrada. Um programa é a codificação de um algoritmo em uma linguagem de programação.

Considerando que um algoritmo é implementado com o propósito de executar uma tarefa ou resolver um problema, um aspecto importante é a sua efetividade na busca da solução. Um algoritmo pode ser desvantajoso em relação a outros, inadequado, ou mesmo inviável de ser utilizado em situações práticas de certas aplicações. Felizmente, é comum existir mais de um algoritmo para solucionar um determinado problema. A escolha do melhor algoritmo vai depender de um fator determinante: o *desempenho*. É nesse contexto que surge o termo *complexidade de algoritmos*, que é definido como a quantidade de trabalho ou esforço requerido por um algoritmo (Toscani e Veloso, 2001). Diversos critérios podem ser empregados para se avaliar o desempenho de um algoritmo, alguns exemplos são: a exatidão, a portabilidade e a clareza do código. Mas dois critérios em especial são ótimos

indicadores de desempenho: a *complexidade de tempo*, que se refere ao tempo de execução de um algoritmo; e a *complexidade de espaço*, que se refere ao espaço requerido em memória por um algoritmo.

Uma preocupação é encontrar uma forma de medir a complexidade, assunto este que é tratado pelo campo conhecido como *análise de algoritmos* ou *análise de desempenho* (Horowitz *et al.*, 1998). Salienta-se, no entanto, que se obtém em geral somente uma aproximação da complexidade, visto que esse é um aspecto que depende de uma série de fatores, e por esse motivo uma predição exata é quase impossível de ser obtida (D'harbe, 2005).

Uma forma de medir a complexidade é através da experimentação, cronometrando o tempo de execução e medindo a quantidade de memória utilizada. Entretanto, isso proporciona uma noção muito relativa de desempenho. Os resultados fornecidos por essa abordagem são muitos pontuais, devido ao tamanho do domínio das entradas, que pode ser muito vasto para ser avaliado. As medições variam também de máquina para máquina devido a diferenças de arquitetura de componentes eletrônicos de dos programas de computadores.

A segunda forma é fazer uma medição teórica através de uma análise matemática. Essa medição consiste em se obter uma função que quantifique as operações fundamentais efetuadas pelo algoritmo, antes mesmo dele ser implementado.

Não é difícil perceber que a complexidade está fortemente associada com o tamanho das entradas fornecidas para o algoritmo. Diante disso, as entradas de maior magnitude são também as de maior interesse, porque são nessas situações que a complexidade torna-se um fator sensível e o seu conhecimento ganha ainda mais importância. Convencionou-se descrever a complexidade para esses casos, chamados de *comportamento assintótico*, através de funções limitantes, caracterizadas por *ordens assintóticas*. São vários os tipos de limites assintóticos, mas os mais importantes são: o superior, o inferior e o exato, que são representados pelas notações  $O$ ,  $\Theta$  e  $\Omega$ , respectivamente.

- *Limite assintótico superior*: define-se  $f(n) = O(g(n))$  (lê-se big Oh), se, e somente se, existem constantes  $c$  e  $n_0$  tal que  $f(n) \leq c \cdot g(n)$  para todo  $n$ , com  $n \geq n_0$ ;

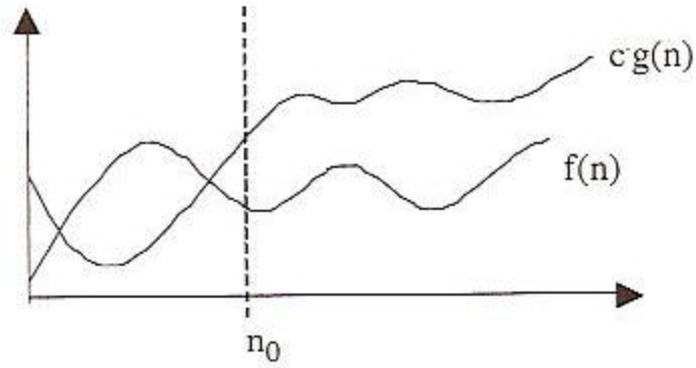


Figura 4.3 -  $f(n)$  é  $O(g(n))$  (Toscani e Veloso, 2001).

- *Complexidade exata*: define-se  $f(n) = \Theta(g(n))$ , se, e somente se, existem constantes  $c$ ,  $c'$  e  $n_0$  tal que  $c.g(n) \leq f(n) \leq c'.g(n)$  para todo  $n$ , com  $n \geq n_0$ .

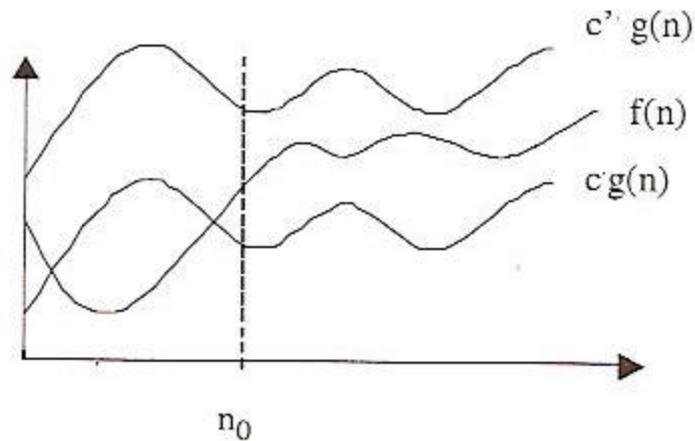


Figura 4.4 -  $f(n)$  é  $\Theta(g(n))$  (Toscani e Veloso, 2001).

- *Limite assintótico inferior*: define-se  $f(n) = \Omega(g(n))$ , se, e somente se, existem constantes  $d$  e  $n_0$  tal que  $g(n) \leq d.f(n)$  para todo  $n$ , com  $n \geq n_0$ ;

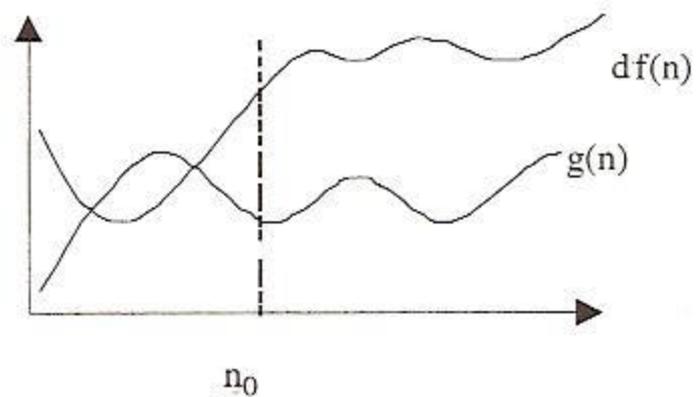


Figura 4.5 -  $f(n)$  é  $\Omega(g(n))$  (Toscani e Veloso, 2001).

A notação  $O$  (big Oh) é a mais comumente referenciada como indicador de complexidade. Por essa notação, dizer que a um algoritmo tem complexidade de espaço  $O(n^2)$ , é o mesmo que dizer que a quantidade de memória utilizada é no máximo uma função quadrática da entrada (D'eharbe, 2005).

Tabela 4.1 - Funções limitantes superiores freqüentemente utilizadas.

$O(g(x))$	<b>Descrição</b>
$O(1)$	Constante
$O(\log n)$	Logarítmica
$O(n)$	Linear
$O(n^2)$	Quadrática
$O(n^3)$	Cúbica
$O(n^k)$ , onde k é uma constante	Polinomial
$O(e^n)$	Exponencial

Outro fato que não pode ser negligenciado quanto ao aspecto desempenho computacional é a complexidade inerente ao próprio problema. No tocante a visão computacional, é da natureza das imagens, principalmente as imagens de profundidade, serem estruturas densamente povoadas. Por isso muitos problemas de natureza global em visão computacional tem como ponto sensível a complexidade computacional. Para se ter uma idéia do problema, basta verificar o custo do tratamento computacional de tarefas corriqueiras dentro da visão computacional. Por exemplo, uma operação de limiarização em uma imagem de intensidade pode ser aplicada em um único pixel sem a necessidade de manipulação de outros pixels. Uma operação de detecção de contorno precisa manipular uma janela de pixels para limiarizar um único pixel. Já uma operação de estabelecimento de correspondências pode impor complexidades preocupantes, pois, no pior caso, para cada pixel de uma imagem é necessário que se opere sobre todos os pixels de outra imagem.

#### 4.2.5 - Heurísticas e meta-heurísticas

Algoritmos com complexidade polinomial são muito desejáveis para muitas aplicações. Essa preocupação também é uma constante dentro da visão computacional. As *heurísticas* procuram satisfazer essa necessidade. Essencialmente, heurísticas são abordagens que tem por meta resolver um problema tendo por diretriz o melhor resultado com o menor esforço. Do ponto de vista computacional, são métodos que possibilitam resolver problemas dentro de uma complexidade de tempo polinomial. As *meta-heurísticas*, por sua vez, são heurísticas de proposta geral. Alguns exemplos de meta-heurísticas são o método da *têmpera simulada*, que simula a técnica de recozimento de metais empregado na metalurgia, e *algoritmos genéticos*, que simulam os aspectos do comportamento evolucionário de sistemas biológicos. Esses métodos fornecem boas soluções, mas paga-se o preço da não garantia da solução ótima (Viana, 1998). Métodos heurísticos são ferramentas importantes para a solução de diversos problemas dentro da visão computacional, principalmente em situações adversas onde a complexidade é muita alta.

#### 4.2.6 - O conceito de máxima verossimilhança, o método dos mínimos quadrados e estimação robusta

Basta olhar para uma população de dados que geralmente tem-se uma noção sobre a similaridade entre eles. Essa similaridade é caracterizada pelo princípio da *máxima verossimilhança*, que é um método que visa estimar valores que maximizam uma função de densidade de probabilidade (Press *et al.*, 1992).

O método dos *mínimos quadrados* é um método de otimização vastamente utilizado para estimar parâmetros de modelos de funções a partir de dados discretos. No ajuste de curvas, o procedimento de estimação começa por escolher-se um modelo matemático que se assemelha aos dados. Isso nos remete ao conceito de máxima verossimilhança. A segunda etapa é montar uma função mérito, que é também chamada de *função objetivo ou função de custo*, que mede a similaridade entre dois conjuntos de dados, de acordo com um estimador. A última etapa é encontrar, a partir da função de custo, os parâmetros que especificam esse modelo.

Estimadores são considerados robustos quando é garantido que o erro de estimação é menor que o tolerado para uma aplicação (Medioni e Kang, 2004). Nesse sentido, robustez é caracterizada por uma insensibilidade a dados discrepantes, a critério de um *ponto de colapso (breack down point)*, ou uma *função de influência*. Métodos robustos começaram a ser investigados nas décadas recentes, nos domínios das mais variadas disciplinas, como estatística, visão computacional, engenharia de controle, e outros.

#### 4.2.6 - Otimização, convexidade de uma função e matriz de covariância

Otimização é um campo da matemática que trata da maximização e minimização de funções. A busca de extremos de funções como os máximos e mínimos globais e locais são o alvo dessa disciplina.

É de grande interesse para problemas de otimização o conhecimento sobre a convexidade de uma função, pois funções não convexas apresentam múltiplos pontos de máximo e mínimo, que podem representar grandes obstáculos para a convergência de métodos numéricos de otimização que procuram por um máximo ou mínimo global. Um algoritmo de busca pode travar em um extremo local e enganar o experimentador.

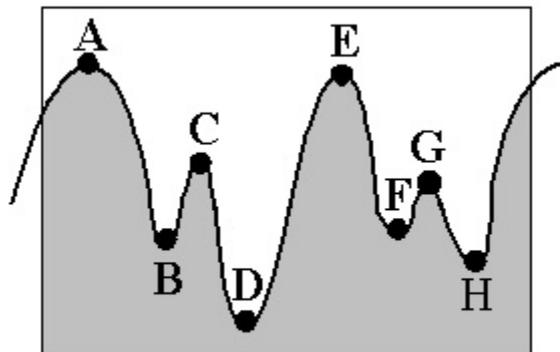


Figura 4.6 - Exemplo de uma função não convexa. Os pontos A, B, C, D, E, F e G são extremos da função. O ponto D é um mínimo global enquanto os pontos B, F e H são mínimos locais. O ponto A é um máximo global, enquanto os pontos C, E e G são máximos locais.

Uma *matriz de covariância* é uma matriz resultante do produto entre duas matrizes de resíduos em relação a uma média ou *centróide*. Dados  $n$  conjuntos de variáveis  $\{X_1\}, \dots, \{X_n\}$ , a matriz de covariância de primeira ordem é definida por:

$$V_{ij} = cov(x_i, x_j) \equiv \langle (x_i - \mu_i)(x_j - \mu_j) \rangle \quad (4.1)$$

onde:

$\mu_i$  e  $\mu_j$  são médias (Weisstein, 2006).

#### 4.2.6 - Decomposição em valores singulares

A *decomposição em valores singulares (SVD)* é um dos métodos mais atraentes para a decompor matrizes e se obter informações úteis sobre espaços vetoriais e sistemas lineares, em virtude da sua estabilidade numérica. A decomposição de uma matriz  $A$  de ordem  $m \times n$ , onde  $m > n$ , em valores singulares, fica:

$$A = UDV^T \quad (4.2)$$

onde:

$U$  e  $V$  são matrizes ortogonais;

$D$  é uma matriz diagonal com valores singulares da matriz  $A$  (Weisstein, 2006).

### 4.3 - REVISÃO DOS MÉTODOS DE REGISTRO NA LITERATURA

Problemas de alinhamento de dados são recorrentes em diversas áreas de pesquisa, e graças a esse fato, muitas soluções já foram propostas e hoje são empregadas na reconstrução 3D. Algumas dessas áreas e a situação em que o problema do registro aparece são: a integração de informações tomadas de diferentes sensores, que ocorre com frequência em robótica; a identificação de mudanças em imagens tomadas em diferentes momentos ou sob diferentes condições, que é comum na medicina; o problema de efetuar a combinação entre uma varredura e o modelo de um objeto 3D previamente armazenado, que acontece com frequência em problemas de reconhecimento, inspeção, rastreamento de objetos em tempo

real; e o problema de calibração de câmeras em visão computacional (Brown, 1992; Simon, 1996; Simon, 1997).

Embora a visão computacional concentre grande parte dos estudos sobre registro na atualidade, as primeiras abordagens do problema surgiram no âmbito da fotogrametria, para resolver o problema de *orientação absoluta* em aquisição estéreo (Thompson, 1958; Schut, 1960). Desde então uma quantidade razoável de métodos já foi proposta, e constantemente estão sendo pesquisados novos métodos, justamente porque o problema do registro é muito particular e o seu sucesso está associado a muitos fatores, como:

- *Requisitos da aplicação*: metas como exatidão do modelo, em aplicações que exigem modelos de alto nível como a simulação numérica na engenharia, e a medicina; rapidez na aquisição do modelo, como é a meta de aplicações de tempo real como a navegação robótica, ou realismo na visualização, como é a meta de aplicações voltados para o entretenimento, como filmes e jogos 3D.
- *Topologia dos dados*: métodos e abordagens de registro dependem de aspectos relacionados às propriedades das superfícies digitalizadas. Esse fato refere-se mais precisamente à complexidade topológica ou geométrica que as superfícies podem impor aos dados. Por exemplo, efetuar o registro entre dois planos é bem diferente de se efetuar o registro entre duas superfícies fractais (Rusinkiewicz e Levoy, 2001).
- *Informações disponíveis*: quando o processo é controlado, desde a etapa de aquisição de dados, muitas informações importantes estão disponíveis e elas podem auxiliar na solução de problemas ou prevenir adversidades nas diversas fases da reconstrução. Exemplos dessas informações são: os parâmetros extrínsecos de calibração e o modelo do ruído introduzido pelo sensor.

#### **4.3.1 - Classificações para os métodos de registro**

Vários critérios podem ser utilizados para se classificar métodos de registro. Eles podem variar quanto à forma que estabelecem correspondência, quanto à exigência ou não de uma estimativa de movimento inicial, quanto à abordagem de otimização ou a qualidade do alinhamento final. Mas na prática, percebe-se que o que difere os métodos de registro de

um modo geral é a forma como efetuam a busca pela transformação final de alinhamento. Nesse sentido, podem ser verificados dois tipos de abordagens: a primeira baseia-se na busca direta pelos parâmetros de movimento em todo o espaço paramétrico. Nesse caso não há a exigência de uma solução inicial, que pode não estar disponível para muitas aplicações, mas o esforço computacional tende a ser bastante grande, visto que o espaço paramétrico geralmente é muito vasto também para ser esquadrinhado; o resultado também depende fortemente da heurística de busca utilizada, e mesmo boas heurísticas não são garantia de sucesso. O outro tipo de abordagem consiste em se buscar a transformação de alinhamento refinando-se uma solução inicial, como faz o algoritmo ICP. A introdução dessa restrição reduz o espaço de busca, mas outras restrições geralmente são necessárias para se ter uma certa garantia de convergência.

Já foram apresentadas algumas classificações para a grande quantidade de métodos de registro já propostos, não havendo, entretanto, muito consenso a esse respeito.

Blais e Levine classificaram os métodos dentro de dois grupos: métodos de *laço aberto* e de *laço fechado* (Blais e Levine, 1995). Métodos de laço aberto utilizam exclusivamente os parâmetros de movimento fornecidos por sistemas de aquisição calibrados para alinhar as vistas. De fato, nas primeiras tentativas de construção de modelo a solução mais evidente era controlar o processo de aquisição de modo que os movimentos estivessem disponíveis posteriormente para o alinhamento das vistas (Bhanu, 1984). No entanto, resultados práticos provam que mesmo sistemas bem calibrados fornecem apenas uma aproximação para os parâmetros ótimos (Pulli, 1997). Os métodos de *laço fechado* por sua vez empregam métodos computacionais que em algum momento utilizam os dados das próprias imagens para encontrar os parâmetros ótimos de movimento.

Chen e co-autores identificaram 4 tipos de registro (Chen *et al.*, 1999). O primeiro tipo foi chamado de *métodos baseados em características*. O termo “característica” apresenta um caráter muito generalista em visão computacional, no caso específico do problema do registro, características são preferencialmente primitivas ou propriedades geométricas presentes nas áreas de sobreposição entre as imagens e invariantes a movimento rígido de rotação e translação, tais como: pontos, retas, planos, cilindros, cones, triângulos, normais de superfície, e outros; também podem ser incorporadas, como marcadores fiduciais, no caso de imagens de intensidade. Explorando-se características correspondentes é possível

obter uma estimativa do movimento espacial que alinha as imagens. Esse tipo de abordagem é bastante popular em problemas de reconhecimento de objetos, mas não são exatas o suficiente para certas aplicações; problemas como inexatidão na comparação ou mesmo a inexistência de características proeminentes comprometem a precisão do registro ou mesmo impedem a sua realização por essa abordagem (Goshtasby, 1998). O segundo tipo de método identificado pelos autores foi chamado de *abordagens de pré-modelagem*. Essas abordagens convertem uma nuvem de pontos em uma representação de mais alto nível, como uma malha de polígonos, por exemplo, e efetuam o registro utilizando essa representação. O terceiro grupo de métodos foi chamado de *abordagens iterativas*. Nessas abordagens empregam-se métodos numéricos para a busca da transformação ótima através do refino iterativo de uma estimativa de movimento inicial. Abordagens iterativas são rápidas, fáceis de implementar e baseiam-se na minimização de uma função de custo que mede a qualidade do alinhamento. Entretanto, algumas hipóteses precisam ser satisfeitas para se ter garantia mínima de convergência, a principal é ter-se uma boa estimativa do movimento inicial. O ICP se insere nesse tipo de abordagem. É mais comum abordagens iterativas trabalharem sobre pontos, mas nada impede que outras características de mais alto nível sejam exploradas por esses métodos. O último tipo identificado pelos autores são as *abordagens de otimização*, as quais se fazem a busca pela transformação ótima analisando o espaço de transformações. Para tal são empregados métodos de otimização de proposta geral, provenientes de áreas como a computação evolucionária, otimização estocástica, visão computacional robusta e controle ótimo.

Borges e co-autores dividiram os métodos de registro 3D dentro duas categorias: *técnicas baseadas em verificação de correspondência*, que numa primeira etapa resolvem o problema da correspondência, e numa segunda etapa calculam a transformação rígida de alinhamento; e *técnicas de minimização*, que procuram as transformações de alinhamento minimizando uma função de erro que mede a similaridade entre vistas (Borges *et al.*, 1994).

Silva e Bellon propuseram dois grupos para os métodos de registro: *técnicas de registro aproximado*, que utilizam um registro aproximado como restrição para um registro ótimo; e *técnicas de registro ótimo*, que utilizam heurísticas para efetuar buscas no espaço paramétrico e não precisam de estimativa de movimento inicial (Silva e Bellon, 1995).

#### 4.4 - TRABALHOS PRÉVIOS ENVOLVENDO A CONSTRUÇÃO DE MODELOS E O PROBLEMA DO REGISTRO

Potmesil desenvolveu um trabalho para a construção de modelos utilizando um sistema de aquisição calibrado e com iluminação controlada para adquirir múltiplas imagens 2D, além de também empregar técnicas de processamento de imagens. O processo de reconstrução consistiu da parametrização da superfície do objeto projetando-se uma grade 2D de linhas isoparamétricas ortogonais, seguida da gravação da superfície parametrizada em pelo menos duas imagens 2D. Segmentos de superfície foram reconstruídos utilizando-se linhas extraídas das imagens e o conhecimento da geometria do sistema de aquisição. Os mesmos segmentos foram organizados hierarquicamente dentro de uma estrutura de dados *quadtree* para posterior combinação e fusão dentro de um modelo completo do objeto (Potmesil, 1983).

Bhanu construiu um sistema para aquisição de dados e análise de cena combinando dados de profundidade contidos em uma seqüência de imagens adquiridas reorientado-se um objeto por ângulos conhecidos em cima de uma plataforma giratória, a qual era um dispositivo integrante de um digitalizador de triangulação ativa baseado em laser. A combinação de formas para reconstruir o modelo empregou uma técnica de rotulação estocástica hierárquica para maximizar um critério de medição da ambigüidade e inconsistência da classificação. Os resultados da combinação foram utilizados para determinar a orientação do objeto no espaço (Bhanu, 1984).

Faugeras e Hebert desenvolveram um sistema de visão considerando três aspectos principais: a representação, o reconhecimento, e a localização de objetos 3D. E onde eles recomendam efetuar a tarefa de reconhecimento simultaneamente com a da localização. A pesquisa começou com a amostragem da superfície de uma peça automotiva utilizando-se um digitalizador a laser. Empregando *transformada de Hough* e *crescimento de regiões*, os dados foram segmentados em um conjunto de regiões  $[P(U_i)]_{i=1,\dots,n}$ , onde  $U_i$  são vetores com os parâmetros de primitivas tais como planos ou quádricas, utilizadas para representar superfícies. As descrições foram hierarquizadas em um grafo para ser empregado no reconhecimento do objeto. Baseado em uma lista  $((M_1S_1), \dots, (M_P S_P))$  de primitivas do modelo e da cena foi criado o critério de consistência  $g(M)$ :

$$g(M) = \underset{T}{\text{Min}} \sum \left\| U_i - T(U'_j) \right\|^2 \quad (4.3)$$

para avaliar a qualidade da combinação. Uma vez que os parâmetros das primitivas da cena estão centrados no sistema de coordenadas do objeto e os parâmetros das primitivas do modelo estão centrados no sistema de coordenadas do visualizador, o problema da localização da peça no espaço consistiu do cálculo do deslocamento rígido que sobrepõe a descrição do modelo à descrição da cena. Para tal foi foram propostos algoritmos não iterativos baseados em pontos e planos, e um algoritmo iterativo baseado em linhas, esses algoritmos representavam rotações entre dois referenciais através de quaternions (Faugeras e Herbert, 1983; Faugeras e Hebert, 1986).

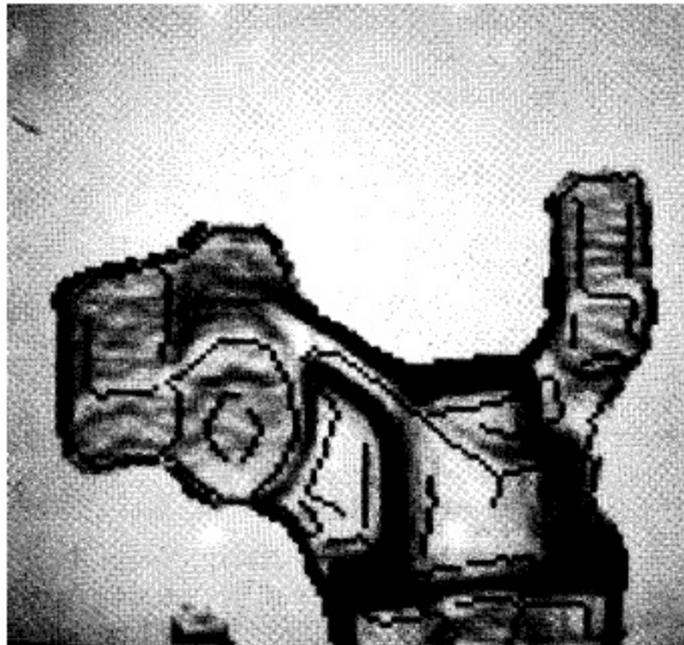


Figura 4.7 - Detecção de contorno em uma peça automotiva (modificado - Faugeras e Hebert, 1986).

Arun e co-autores propuseram uma solução não iterativa para o problema da estimação de movimento no registro entre dois conjuntos de dados. Diferentemente de Faugeras e Hebert, que utilizaram quaternions, os autores utilizaram *SVD* de uma matriz de covariância para obter matrizes ortonormais, cujo produto resulta em uma matriz de rotação entre os sistemas de coordenadas dos conjuntos de dados das imagens (Arun *et al.*, 1987).

Besl e McKay estabeleceram a base e cunharam o termo ICP para o algoritmo que vem se estabelecendo como o método mais popular para alinhamento de dados 3D. O algoritmo por eles proposto é um método numérico iterativo baseado em mínimos quadrados, que alinha dois conjuntos de dados através do refinamento de uma estimativa de movimento inicial (Besl e McKay, 1992). Esse algoritmo é utilizado neste trabalho para o alinhamento de imagens para a construção de modelos e é descrito em detalhes na seqüência deste capítulo.

Chen e Medioni apresentaram uma implementação com o mesmo conceito de minimização iterativa do *ICP* de Besl e Mckay. A diferença está na utilização da heurística de combinação conhecida por *ponto a plano* para estabelecer correspondências ao invés da heurística *ponto a ponto* utilizada no *ICP* original (Chen e Medioni, 1992). A idéia dessa heurística não era totalmente nova, visto que Potmesil já havia utilizado esse expediente no seu trabalho.

Zhang apresentou, aproximadamente na mesma época que Besl e Mckay e Chen e Medioni, um algoritmo iterativo com características semelhantes a do *ICP*. A proposta era efetuar a combinação entre curvas de forma livre representadas por um conjunto de pontos encadeados reduzindo gradativamente a distancia media entre os pontos através de uma estimativa de mínimos quadrados (Zhang, 1992).

Blais e Levine utilizaram uma aparelhagem de aquisição de imagens calibrada e aproveitaram os parâmetros de calibração extrínsecos para registrar as imagens. Os autores utilizaram um método denominado *calibração reversa* do sensor para resolver o problema da correspondência. O método de otimização estocástica *VFSR (Very Fast Simulated Reannealing)* foi usado para minimizar a função de custo (Blais e Levine, 1995).

Silva e Bellon efetuaram registro simultâneo de múltiplas vistas sem pré-alinhamento e com pouca sobreposição utilizando algoritmos genéticos. O mesmo trabalho também propôs a métrica *SIM (Surface Interpenetration Measure)* para avaliação da precisão do registro (Silva e Bellon, 1995).

Chen e co-autores propuseram o método para registro robusto denominado *DARCES* (*Data Aligned Rigidity Constrained Exhaustive Search*), inspirado na técnica *RANSAC* (*Random Sample Consensus*), oriunda da visão computacional robusta. O processo consiste na busca exaustiva do alinhamento ótimo, considerando que três pontos correspondentes é a quantidade mínima de dados para se efetuar registro, e explorando a restrição da invariância de forma determinada pelo movimento rígido (Chen *et al.*, 1999).

Dorai construiu modelos completos a partir de dados corrompido com ruído. A pesquisa voltou-se para o problema do registro ótimo entre vistas e para tal derivou um *estimador de variância mínima* para medir a qualidade do registro entre duas vistas (Dorai *et al.*, 1994; Dorai *et al.*, 1997).

Pulli construiu modelos digitalizando a cor e a geometria de um objeto com um sistema estereoscópico a partir de múltiplas vistas arbitrárias. O registro em pares foi utilizado como restrição para um registro simultâneo de modo a distribuir igualmente os erros de registros entre as vistas (Pulli, 1997).

Rusinkiewicz e Levoy fizeram um estudo focado na classificação e comparação de variantes do algoritmo *ICP*, além de apresentar uma variante com configuração voltada para o registro em alta velocidade. A explosão combinatorial de possibilidades de variantes foi resolvida estipulando-se uma configuração de referencia como metodologia de comparação. Seis estágios foram identificados no algoritmo: a *seleção* de pontos, a *combinação* de pontos, *ponderação* de pares correspondentes, a *rejeição* de dados discrepantes, a *métrica de erro* e a *minimização* da função de custo (Rusinkiewicz e Levoy, 2001).

#### **4.5 - O ALGORITMO ICP**

O algoritmo *ICP* se tornou popular por ser uma solução simples para um problema complexo como o do registro 3D. O método é composto de etapas para as quais podem ser ajustadas heurísticas para torná-lo mais rápido ou mais exato. Essa flexibilidade o torna bastante atrativo quanto ao aspecto da generalidade e capacidade de adaptação a requerimentos de diferentes aplicações, permitindo que muitos melhoramentos já tenham

sido propostos a ele ao longo dos anos, dando origem assim a uma verdadeira família de algoritmos variantes do *ICP* original.

O *ICP* é essencialmente um método numérico que refina iterativamente uma transformação inicial. A iteração é um processo de otimização que visa minimizar uma função objetivo de mínimos quadrados entre pares de pontos correspondentes que mede a similaridade entre duas imagens. A partir dessa função, parâmetros de movimento rígido são estimados e reaplicados a uma das imagens para aproximá-las. A iteração é repetida até que um critério de parada seja satisfeito.

A primeira hipótese para se aplicar o *ICP* é que as imagens adjacentes tenham regiões de sobreposição de dados. Uma restrição que foi imposta aos dados no *ICP* original, para prevenir a geração de falsos pares, e que é pouco realista com relação a situações práticas, é a de que uma imagem deveria ser subconjunto da outra (Besl e Mckay, 1992). E uma exigência para a convergência é a de que uma “*boa estimativa*” dos parâmetros de registro inicial seja fornecida como entrada para que o algoritmo o refine (Besl e Mckay, 1992; Rusinkiewicz e Levoy, 2001). Por fim, uma peculiaridade é a de que o algoritmo converge monotonicamente para um mínimo local (Besl e Mckay, 1992).

## **4.6 - ETAPAS ESSENCIAIS DO REGISTRO UTILIZANDO O ALGORITMO ICP**

### **4.6.1 - Pré-alinhamento**

O pré-alinhamento é o problema da determinação de um registro inicial, o qual pode ser obtido de variadas formas, inclusive utilizando-se outros métodos de registro considerados não-ótimos. Como já mencionado, se o sistema de aquisição está calibrado, os parâmetros extrínsecos de calibração podem ser uma alternativa para o alinhamento inicial do *ICP* (Pulli, 1997); ou pode-se usar, por exemplo, um método de combinação de características para esse fim.

### **4.6.2 - Seleção de pontos para formar pares correspondentes**

A seleção de pontos é a etapa de escolha de pontos de interesse para efetuar combinação. Esses pontos são chamados de *pontos de controle* e devem estar na região de sobreposição das imagens. Eles podem ser escolhidos aleatoriamente, ou com base em algum critério de

subamostragem; em apenas uma, ou nas duas imagens envolvidas. Os pares correspondentes na reconstrução com imagens de profundidade são análogos aos pares conjugados da estereoscopia, a diferença é que nesta última, a correspondência tem por meta determinar a matriz fundamental que relaciona duas imagens de intensidade calibradas.

#### **4.6.3 - Soluções para o problema da correspondência**

O problema da correspondência é reconhecidamente um dos maiores desafios presentes no campo da visão computacional, e ganha ainda mais dificuldade quando a dimensão do espaço do conjunto de dados é tridimensional. Na reconstrução a combinação é um elemento-chave porque afeta decisivamente duas metas que se tornam conflitantes devido à complexidade computacional do problema: a qualidade e a velocidade do registro.

- **Complexidade do problema da correspondência**

De acordo com Besl e McKay, a busca por correspondência consome aproximadamente 95% do tempo de execução (Besl e McKay, 1992). A complexidade computacional de tempo é da ordem de  $O(n^2)$  comparações para se encontrar  $N$  pareamentos efetuando-se busca exaustiva pelo ponto mais próximo (Eggert *et al.*, 1998). Muitas alternativas podem ser tomadas para se acelerar essa busca, uma das mais eficientes é reduzir a complexidade utilizando uma estrutura de dados hierárquica, como uma árvore binária, tipo octree ou kd-tree, por exemplo; esta última apresenta complexidade de tempo da ordem  $O(N \log N)$ .

- **Tipos de abordagens**

A escolha do melhor elemento, ou elementos, para se efetuar correspondência depende de aspectos como a topologia dos pontos, a disponibilidade das características escolhidas e a dificuldade para sua extração (Li, 1999). Pode se escolher uma combinação baseada em pontos ou combinação baseada em outras características de similaridade entre imagens.

A nuvem de pontos de uma imagem de profundidade é na verdade a discretização de superfícies. Portanto, outras primitivas geométricas ou representações de superfície

diferente de pontos podem ser exploradas, citam-se: conjuntos de segmentos de linhas, curvas implícitas, curvas paramétricas, conjuntos de triângulos, superfícies implícitas e superfícies paramétricas (Besl e McKay, 1992). Características não geométricas como intensidade, cor ou textura podem auxiliar no estabelecimento da correspondência se disponíveis (Pulli, 1997; Rusinkiewicz e Levoy, 2001).

Explorar características diferentes de pontos costumam aumentar a exatidão em certos casos, mas também tendem tornar mais complexo o tratamento matemático do problema. As dificuldades com as abordagens baseadas em características estão relacionados com a disponibilidade, exatidão, representação, extração, classificação e catalogação dessas características (Bhanu, 1984; Faugeras e Herbert, 1983; Chen e Medioni, 1992).

- **Correspondência ponto a ponto**

No *ICP* original foi utilizada uma combinação de superfícies baseada em pontos. Essa abordagem é especialmente atrativa devido à facilidade de implementação e por proporcionar atraentes soluções de forma fechada. No entanto, efetuar a busca pelo ponto mais próximo em outra imagem é uma heurística, e como tal, não há garantias de que o ponto mais próximo seja realmente o ponto correspondente. Quando a busca retorna um falso positivo, um pareamento errado é formado e precisa ser descartado, para não comprometer a função de custo.

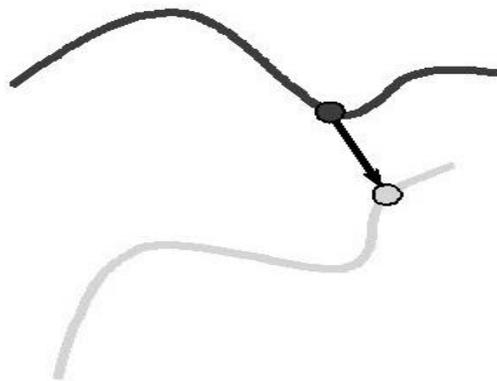


Figura 4.8 - Correspondência ponto a ponto em 2D (modificado – ISR/Coimbra, 2001).

- **Correspondência ponto a plano**

Uma alternativa ainda pouco explorada é no sentido de se utilizar propriedades de superfície de ordem superior, como gradiente ou normais de superfície, para orientar a busca pelo ponto mais próximo na correspondência. Chen e Medioni utilizaram a abordagem conhecida por correspondência *ponto a plano* ou *lançamento de normal* (*normal shooting*), que consiste em se montar uma função de custo medindo-se a distância entre superfícies na direção da normal ao ponto em uma imagem até ao plano em outra imagem (Chen e Medioni, 1992; Rusinkiewicz e Levoy, 2001).

Considerando dois conjuntos de pontos  $p_i$  e  $q_j$  pertencentes a duas superfícies  $P$  e  $Q$  (Figura 4.6), respectivamente, o problema da correspondência entre eles é resolvido utilizando-se uma transformação inicial  $T^0$ , e em cada em cada iteração  $k$ , uma transformação prévia  $T^{k-1}$  disponível. A transformação de alinhamento  $T$  é a que minimiza a função de custo  $e$  formada pelas distâncias quadráticas  $d_s$  entre um ponto na superfície  $P$  a um plano tangente à superfície  $Q$ .

$$e^k = \sum_{i=1}^N d_s^2(T^k p_i, S_j^k) \quad (4.4)$$

Onde:

$S_j^k$  é o plano tangente a  $Q$  em  $q_j^k$ :

$$S_j^k = \{s \mid n_{qj}^k \cdot (q_j^k - s) = 0\}$$

$n_{qj}^k$  é o vetor normal a  $Q$  a partir de  $q_j^k$ :

$$q_j^k = (T^{k-1} \ell_i) \cap Q$$

e  $\ell_i$  representa uma linha normal a  $P$ , a partir de  $p_i$ :

$$\ell_i = \{a \mid (p_i - a) \times n_{p_i} = 0\}$$

Sendo ainda,  $\cdot$  um produto escalar e  $\times$  um produto vetorial.

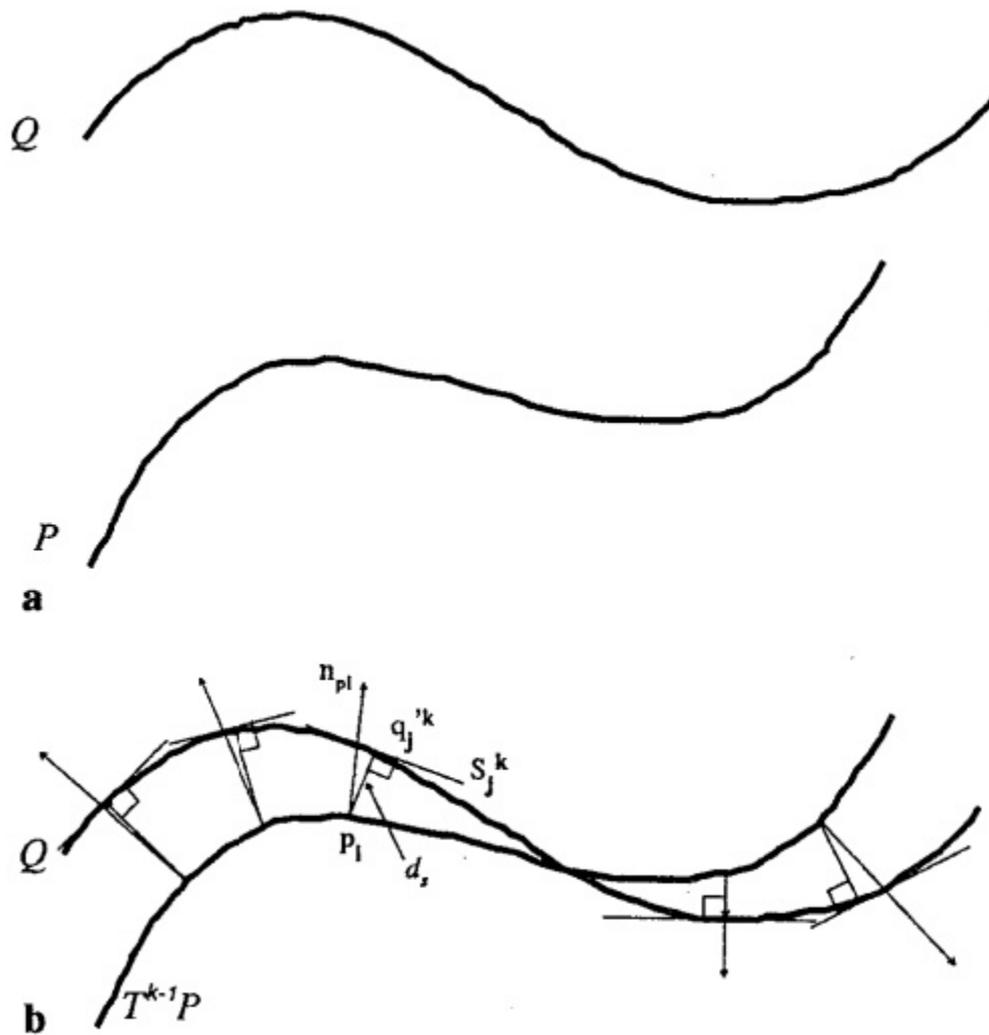


Figura 4.9 - Correspondência ponto a plano em 2D (Chen e Medioni, 1992). (a)  $Q$  e  $P$  antes de  $T^{k-1}$  ser aplicada. (b) distância ao plano tangente de  $Q$  (Chen e Medioni, 1992).

#### 4.6.4 - Estimação de parâmetros de movimento rígido

Em cada iteração o *ICP* precisa recuperar os parâmetros do movimento rígido 3D entre as imagens para poder reapplicá-los a uma delas, de modo a aproximá-las. Esse é um problema de estimação de pose, reconhecidamente complexo. A maior dificuldade encontra-se na estimação do parâmetro de rotação.

Um procedimento comum para a estimação é desacoplar os dois componentes, rotação e translação, para facilitar a estimação (Arun *et al.*, 1987). Esse artifício será especificado na seção que trata da formulação matemática do problema de alinhamento.

- **Estimação da rotação**

A estimação da rotação no registro é importante para alinhar os eixos dos sistemas de coordenadas envolvidos, de modo que eles fiquem com a mesma orientação.

Matrizes de rotação fazem parte de um grupo especial de matrizes não lineares chamado de *grupo ortogonal especial*  $SO(3)$ , com propriedades específicas que caracterizam a sua ortogonalidade. Seja  $R$  é uma matriz ortogonal de ordem  $N$ :

- o produto pela transposta fornece uma matriz identidade:

$$RR^T = I \quad (4.5)$$

- É ortogonal, ou seja, a transposta igual à inversa:

$$R^T = R^{-1} \quad (4.6)$$

- As linhas e também as colunas de  $R$ , devido a Equação (4.6), formam bases ortonormais.

- O determinante de  $R$  é igual a 1.

Existem variadas formas de se estimar rotação. Kanatani apresentou uma análise detalhada de variadas técnicas para ajuste de rotação 3D (Kanatani, 1994). Um trabalho similar foi feito por Eggert e co-autores (1997), onde eles fazem uma comparação entre quatro algoritmos de estimação de transformação rígida 3D: solução envolvendo a *SVD de uma matriz*, envolvendo *matrizes ortonormais* (Horn *et al.*, 1988), envolvendo *quaternions unitários* (Horn, 1987), e solução envolvendo *dual quaternions*; sem encontrar diferenças significativas entre essas abordagens (Eggert *et al.*, 1997).

Muitos problemas enfrentados na estimação de pose ocorrem devido à existência de singularidades na representação. Algumas representações conduzem a algoritmos instáveis numericamente, ditos mal condicionados, onde uma pequena mudança no valor das variáveis causam grandes valores nos resultados (Kreyszig, 1998). Ângulos de Euler, por

exemplo, apresenta o problema da perda de um grau de liberdade devido ao alinhamento de eixos (*gimbal lock*).

Uma matriz de rotação pode ser obtida pelo produto de bases ortonormais. Para se chegar a essas bases emprega-se algum método de decomposição matricial, como *SVD* de uma matriz de covariância, por exemplo. O problema principal na estimação da rotação por esse caminho está em se determinar se a matriz estimada realmente representa uma matriz rotação, pois pode-se obter rotações impróprias (rotações com determinante diferente de 1) nos casos em que os dados contém ruído. Há a necessidade de se reforçar ortonormalidade nessa situação. Umeyama propôs a utilização de multiplicadores de Lagrange para o problema das rotações impróprias (Umeyama, 1991). Entretanto, a solução mais elegante para esse problema é empregar quaternions na representação de rotações, contudo, quaternions precisam ser normalizados para que representem rotações puras, o que pode ser inadequado em algumas aplicações.

Tabela 4.2 - Vantagens e desvantagens de representações de rotação.

<b>Representação de rotação</b>	<b>Vantagens</b>	<b>Desvantagens</b>
Ângulos de Euler	Fácil interpretação e utilização	Ângulos não são únicos para uma dada orientação
		Instável para ângulos pequenos devido a expressões trigonométricas na fórmula
		Conduz a algoritmos que não são bem condicionados, pois ângulos pequenos podem conduzir a grandes rotações
		Problema da perda de um grau de liberdade do sistema devido ao alinhamento de eixos ( <i>gimbal lock</i> )
Representação com ângulo e eixo arbitrário	Fácil interpretação e utilização	Conduz a algoritmos numéricos instáveis, da mesma forma que ângulos de Euler
		Expressões trigonométricas na fórmula
Quaternions	Conduz a algoritmos bem condicionados	O quaternion tem que ser normalizado para representar uma rotação pura
		Não é uma representação muito intuitiva à primeira vista

- **Estimação da translação**

A estimação da translação é necessária para alinhar as origens dos sistemas de coordenadas das imagens. A recuperação desse componente do movimento rígido é direta, desde que a matriz de rotação já tenha sido recuperada. O procedimento para isso é simples e é freqüentemente utilizado em calibração de câmeras, para tal basta efetuar uma subtração vetorial: a diferença entre os centróides (médias) dos dois conjuntos de dados.

#### **4.6.5 - Métrica de erro e a minimização**

A métrica de erro está intimamente relacionada com a abordagem escolhida para efetuar combinação (Brown, 1992). Os pares escolhidos em cada iteração determinam resíduos que compõe função de custo de mínimos quadrados a ser minimizada. A forma de medir esse resíduo influencia o tratamento da função de custo. A escolha da abordagem ponto a ponto proporciona atrativas soluções de forma fechada para a estimação da transformação de corpos rígidos que minimiza o erro. Para a abordagem ponto a plano, no entanto, não existem soluções de forma fechada e o tratamento matemático deve ser feito com métodos não lineares genéricos, como Levenberg-Marquardt, ou recorrendo-se a linearização (Rusinkiewicz. e Levoy, 2001).

A minimização da função de custo do *ICP* é feita iterativamente, reaplicando-se aos dados a pose estimada em cada iteração. Nesse processo o algoritmo pode enterrar em um mínimo local da função de custo, tendo como consequência um alinhamento impreciso, em consequência de aspectos como a qualidade do registro inicial e a presença de ruído.

#### **4.6.6 - Critérios de parada**

O *ICP* é um método iterativo, e como tal, variados critérios podem ser utilizados para parar o ciclo do algoritmo. A princípio, a parada deveria ocorrer quando um certo nível de acerto na sobreposição entre as imagens fosse atingido. Algum critério que indique esse estado podem ser arbitrariamente escolhido pelo experimentador, pode se utilizar um limiar, ou valor de tolerância, para os valores das rotações ou erros *RMS* calculados. O algoritmo fica iterando até que o limiar especificado seja atingido.

#### 4.6.7 - Pré-processamento e rejeição de dados discrepantes

Imagens de profundidade são estruturas que tendem a ser complexas topologicamente. Além desse fato, estão sujeitas a diversos problemas como partes ocultas por inacessibilidade da digitalização, ou medições errôneas introduzidas tanto pelo próprio sensor como pelas propriedades das superfícies digitalizadas. Esses percalços podem ser amenizados com uma etapa, não obrigatória, mas recomendável, de pré-processamento dos dados. Em suma, o que se pretende é prevenir dados discrepantes e seus efeitos quase sempre danosos para algoritmos de estimação. O desenvolvimento de filtros específicos para imagens de profundidade é um campo ainda muito pouco pesquisado, por isso, muitas técnicas utilizadas em imagens de intensidade e na estatística são adaptados para dados 3D. Dorai e co-autores (1998) utilizaram em imagens de profundidade uma abordagem de remoção de dados discrepantes baseada em duas etapas. Uma para a remoção de pixels isolados, e outra para remoção de ruído através de filtragem mediana. Segundo os autores, essa etapa de pré-processando melhorou notavelmente a exatidão dos resultados na fase de integração de vista (Dorai *et al.*, 1998).

O primeiro efeito dos dados discrepantes no *ICP* não sentidos claramente no pareamento de correspondências. Por esse motivo muitos melhoramentos têm sido propostos na fase de formação de correspondências no sentido de rejeitar falsos pares. Pode-se introduzir dispositivos para rejeitar dados discrepantes ou para amenizar os seus efeitos já “dentro” do laço do *ICP*, nas etapas de escolha de pontos de controle, combinação de pontos, e na função de custo. Os falsos pares são extremamente danosos para estimadores de mínimos quadrados. Apenas um dado discrepante é suficiente para comprometer funções de custo não robustas. Sem essa rejeição a função de custo pode ficar completamente comprometida, devido ao fato do *ICP* clássico ser baseado em uma função de custo de mínimos quadrados não-robusta. Uma alternativa matemática para amenizar o efeito desses dados é escolher uma função de custo robusta com um *breakdown* adequado para o nível de degradação dos dados. Mínimos quadrados ponderados e Estimadores M são algumas das soluções (Meer, 2004). Uma série de métodos para impedir o pareamento errôneo ou descartar pares falsos podem ser utilizados. Entretanto, tratando-se de imagens deve-se levar em conta o componente ruído, o qual pode é geralmente modelado como aditivo, mas outros modelos também podem ser encontrados.

Em resumo, o problema da presença de ruído ou dados discrepantes em imagens de profundidade são atacados de duas formas: trabalhando-se diretamente nos dados, efetuando a “limpeza” nos pontos de controle e pares correspondentes, ou manipulando a função de minimização, tornando-a mais robusta.

#### 4.7 - FORMULAÇÃO MATEMÁTICA DO REGISTRO ENTRE DUAS IMAGENS UTILIZANDO ICP

Considere dois conjuntos de pontos  $x_i$  e  $y_i$ , onde  $i = 1, \dots, N$ . A Equação (3.6) que modela o movimento rígido entre os dois conjuntos de pontos, reescrito em um sentido de mínimos quadrados, se transforma em uma somatória de resíduos na forma da Equação:

$$\sum_{i=1}^N \|x_i - (Ry_i + T)\|^2 \quad (4.7)$$

onde:

$R$  é uma matriz de rotação;

$T$  é um vetor de translação.

A Equação (4.7) é uma função de custo baseada em mínimos quadrados. Os parâmetros de rotação e translação ótimos são aqueles que transformam essa função de custo em um mínimo.

- **Desacoplando rotação e translação**

Se os dados estiverem com as posições alinhadas, então  $x_i$  e  $y_i$  tem o mesmo centróide, ou seja,  $\bar{x} = \bar{y}$  (Arun *et al.*, 1987), onde:

$$\begin{aligned} \bar{x} &= \sum_{i=1}^N \|x_i\| \\ \bar{y} &= \sum_{i=1}^N \|y_i\| \end{aligned} \quad (4.8)$$

Portanto, em termos de desvios  $p_i$  e  $q_i$ :

$$\begin{aligned} p_i &= x_i - \bar{x} \\ q_i &= y_i - \bar{y} \end{aligned} \quad (4.9)$$

Temos a função de custo da Equação (4.10):

$$\sum_{i=1}^N \|p_i - Rq_i\|^2 \quad (4.10)$$

A recuperação da rotação é reduzida ao problema de se determinar  $R^*$  que minimiza a Equação (4.10), e o problema da recuperação da translação consiste em se determinar  $T^*$  na Equação (4.11).

$$T^* = \bar{x} - R^* \bar{y} \quad (4.11)$$

Efetuando as devidas operações algébricas na Equação (4.10), têm-se a Equação (4.12), a partir da qual é possível se recuperar a matriz de rotação decompondo-a em duas bases ortonormais  $V$  e  $U$ , obtidas através da decomposição por *SVD* da matriz de covariância  $H$ :

$$H = \sum_{i=1}^N q_i p_i^T \quad (4.12)$$

$$H = UDV^T \quad (4.13)$$

E finalmente, chega-se a uma matriz de estimação  $K$  para a matriz de rotação:

$$K = VU^T \quad (4.14)$$

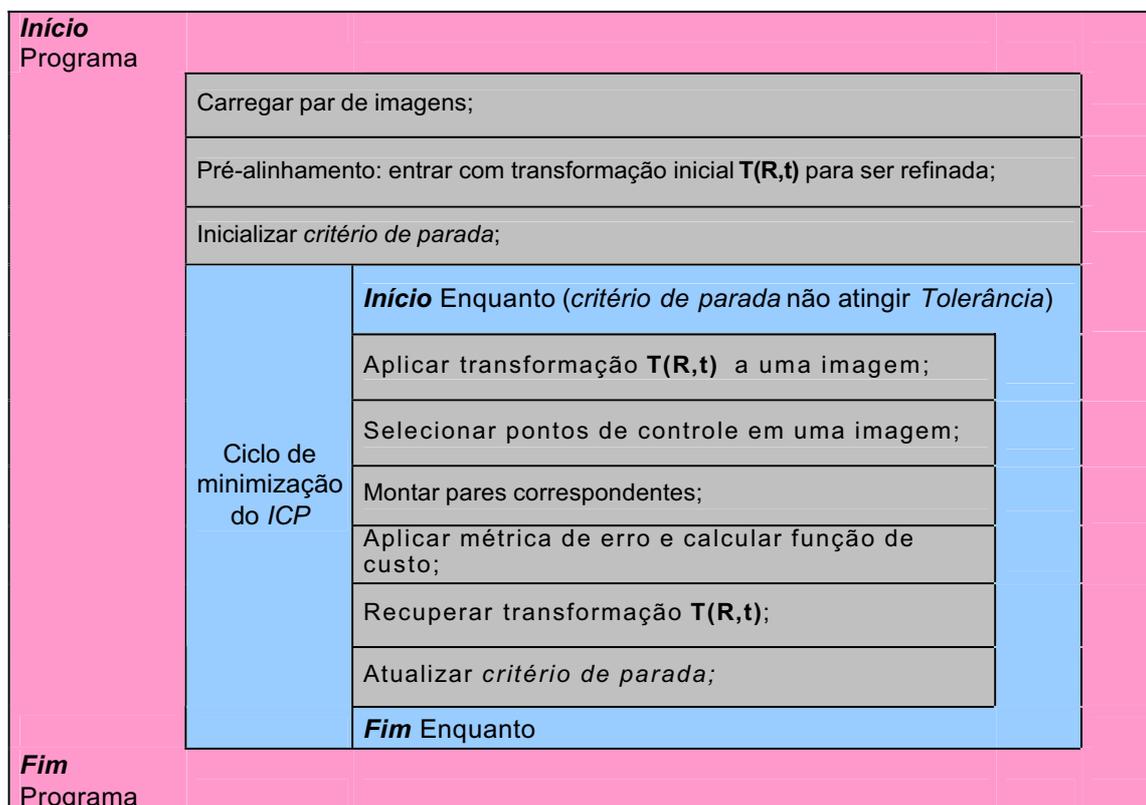
Para retornar uma rotação,  $K$  precisa ser uma matriz ortogonal, ou seja,  $R^T = R^{-1}$  e o seu determinante precisa ser igual a 1. Nesse caso faz-se  $R^*$  igual a  $K$ .

Se o determinante for igual a  $-1$ , a matriz retorna uma rotação imprópria, e os vetores coluna da matriz de rotação são postos para  $(v_1, v_2, -v_3)$ .

Esse processo todo é iterado até que seja satisfeito um critério de parada.

## 4.8 - PSEUDOCÓDIGO DO ALGORITMO ICP

Tabela 4.3 - Pseudocódigo do algoritmo *ICP*.



## 4.9 - ESTRATÉGIAS DE ALINHAMENTO DE MÚLTIPLAS VISTAS PARA FORMAR UM MODELO COMPLETO

Para formar um modelo completo é provável que muito mais que duas imagens precisem ser alinhadas, formando uma espécie de cadeia de imagens. Quanto a esse aspecto duas estratégias são utilizadas: *registro em pares ou local*, e o *registro global, simultâneo ou paralelo*.

A idéia do registro aos pares é alinhar um par de imagens, gerando uma nova imagem, chamada de *metavista*, para ser alinhada com a próxima da seqüência. Esse procedimento é repetido até que todas as vistas sejam registradas. O registro em pares é o mais praticado devido à facilidade de implementação e ao baixo custo computacional. Trabalha-se com

apenas duas vistas por vez. A desvantagem é uma maior inexatidão no resultado final. Os erros de registro de cada etapa são somados aos erros da etapa anterior, de forma que nas últimas etapas haverá uma concentração considerável de erros de registro. O efeito disso é um desalinhamento mais acentuado entre a primeira e a última vistas alinhadas, Figura 4.10 (Eggert *et al.*, 1997; Pulli, 1997; Pulli 1999).

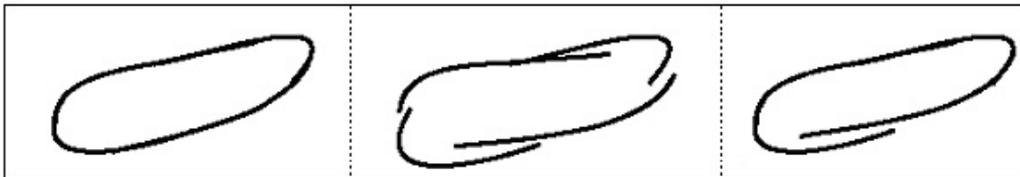


Figura 4.10 - Acúmulo de erros de registro em pares representado em 2D. À esquerda: objeto digitalizado. Ao centro: registro inicial das vistas. À direita: erros de registros em pares acumulados (modificado - Pulli, 1997).

O registro global, simultâneo ou paralelo de múltiplas imagens de profundidade resolve o problema de acúmulo de erros através do registro de todas as vistas ao mesmo tempo (Nishino e Ikeuchi, 2002; Stoddart e Hilton, 1996; Dorai *et al.*, 1998; Eggert *et al.* 1998). Nesse tipo de registro os erros são espalhados igualmente entre as vistas, resultando assim em um alinhamento mais homogêneo. Comparando as duas estratégias, o número de conjuntos de dados sobre os quais operam e como é tratado o problema da correspondência são as diferenças principais. Abordagens em pares operam sobre dois conjuntos de pontos, com correspondências *um para um* definida entre esses conjuntos. O registro global, por outro lado, envolve múltiplos conjuntos de pontos com múltiplos conjuntos de correspondência entre eles (Williams e Bennamoun, 2001). A desvantagem do registro global é o seu custo computacional muito alto, principalmente no tocante a complexidade de espaço; e naturalmente, o requerimento de grande quantidade de memória.

Há pesquisas também no sentido de um registro que pode ser considerado híbrido. Utilizando no registro global com uma etapa de registro em pares, como auxiliador no alinhamento inicial (Williams e Bennamoun, 2001; Pulli, 1997; Huber, 2002).

A quantidade relativamente pequena de estudos sobre essas estratégias, principalmente com respeito ao registro global sugere que outras soluções possam ser propostas.

#### 4.10 - VALIDAÇÃO DE REGISTRO

Além da simples avaliação visual, há claramente a necessidade de uma prova teórica para medir a qualidade do registro efetuado. Um ferramental para isso é de extrema importância, porque não existe nenhuma garantia teórica da convergência do algoritmo *ICP* para os valores corretos, mesmo para os casos sem ruído (Chen *et al.*, 1999). A validação de registro trata dessa necessidade e refere-se às formas de se avaliar se os parâmetros de movimento retornados convergiram para os valores corretos de alinhamento ao final da execução. Sem um valor comparativo, que existe quando se utiliza imagens de profundidade sintéticas, é difícil fazer essa avaliação e muitas restrições precisam ser impostas para se ter uma certa garantia de convergência. Ao passo que o registro é um problema de otimização, o erro *RMS* da função de custo é frequentemente utilizado como critério de parada e medidor da convergência de registro. No entanto, o decremento do valor *RMS* não é condição suficiente para garantir que o algoritmo convergiu para os valores corretos do movimento rígido. A função de custo do *ICP* é do tipo não-convexa, ao final o algoritmo pode ter atingido um mínimo local ao invés de um mínimo global. Há relativamente poucos estudos no sentido de se encontrar um meio adequado para se validar o registro. Uma análise matemática e estatística de modo a avaliar a convergência do algoritmo bem como o seu comportamento durante toda a execução seria salutar.

## **CAPÍTULO 5 – INTEGRAÇÃO DE IMAGENS, RECONSTRUÇÃO DE SUPERFÍCIE E VISUALIZAÇÃO**

Alguns dos propósitos finais do uso de um modelo são a sua simples visualização, a exportação para uma máquina de prototipagem rápida, ou como entrada para um software de simulação, otimização, inspeção ou reconhecimento. Muitas aplicações exigem uma representação de superfície mais adequada do que pontos, o que remete ao tema da representação de superfície.

Para se chegar ao modelo 3D final, pronto para uso em uma aplicação específica, dois problemas ainda precisam ser resolvidos, depois de as imagens de profundidade estarem devidamente alinhadas:

- As várias imagens de profundidade precisam ser fundidas para formar uma geometria única, monolítica;
- A superfície do objeto virtual precisa ser estimada com fidelidade a superfície do objeto real.

A ordem de realização dessas duas etapas pode variar (Pulli, 1997). E longe de ser um simples acabamento, esse passo final da reconstrução é tão importante e complexo quanto os anteriores. Muitas das técnicas empregadas nesta etapa são oriundas da computação gráfica.

Outro aspecto importante, não só durante esta última etapa, mas durante todas as fases de construção de um modelo, é a necessidade de visualizar dados 3D, comumente chamado de *renderização*.

### **5.1 - FORMAS DE REPRESENTAÇÃO DE SUPERFÍCIE**

Existem inúmeras formas de representar superfície. Áreas como a *modelagem de sólidos e computação gráfica* representam superfícies ajustando linhas e curvas a modelos CAD, para constituir estruturas em forma de arame (*wireframe*) ou representações de contorno *B-rep* (*boundary representation*), dentre outras formas de representação.

Com imagens de profundidade a estratégia é revestir a superfície do modelo com retalhos, que podem ser funções ajustadas ou malhas poligonais. A função de ajuste pode ter a forma *implícita*, *explícita* ou *paramétrica*; e a malha poligonal é geralmente formada por uma cadeia de facetas triangulares.

- *Funções explícitas*: os pontos que pertencem a uma superfície podem ser obtidos diretamente por uma função explícita, ou seja, através de um simples mapeamento. A função para uma imagem de profundidade tem a forma explícita:

$$z = f(x, y) \quad (5.1)$$

Essa representação plota “nuvens de pontos” e tem vantagem de ser de fácil implementação.

- *Funções implícitas*: as coordenadas  $z$  que especificam uma superfície podem não estar em evidência. Muitas vezes as coordenadas  $x$ ,  $y$  e  $z$  precisam satisfazer uma equação da forma:

$$f(x, y, z) = 0 \quad (5.2)$$

- *Forma paramétrica*: a superfície é especificada por três funções com domínio nos parâmetros  $(u, v)$ , ou seja:

$$z = f(x(u, v), y(u, v), z(u, v)) \quad (5.3)$$

- *Malha de polígonos*: desde que três pontos definem um plano, facetas triangulares podem ser ajustadas à superfície do objeto triangulando o conjunto de pontos.

Superfícies explícitas são boas para a plotagem gráfica, mas não são tão convenientes em tarefas de ajuste e modelagem. Já as curvas paramétricas e a representação poligonal são as técnicas que dominam a modelagem tridimensional na computação gráfica. Curvas paramétricas facilitam a descrição de curvas e superfícies em retalhos. Um tipo popular de representação de superfície paramétrica são as curvas *B-Splines* e *NURBS* (Piegl e Tiller,

1997). Esse tipo de representação, no entanto, só é recomendável se os pontos estão regularmente espaçados (Goshtasby, 1998).

Uma boa representação para objetos de forma livre deve acomodar todas as formas geométricas presentes em imagens de profundidade. A *geometria computacional* é o campo estudo que trata mais especificamente desses problemas geométricos, a partir de uma abordagem computacional.

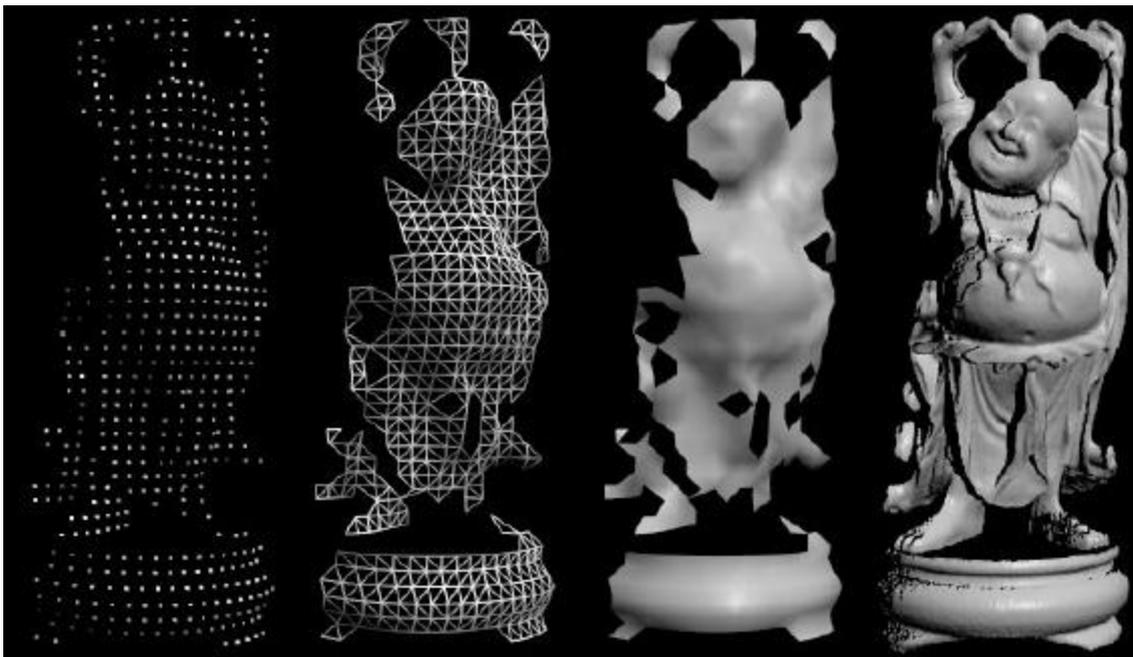


Figura 5.1 - De imagens de profundidade a superfície de profundidade. (a) Imagem de profundidade (subamostrada). (b) Depois de criar triângulos. (c) Renderização sombreada. (d) Superfície de profundidade de alta resolução (Curless, 1997).

## 5.2 - ABORDAGENS PARA A RECONSTRUÇÃO DE SUPERFÍCIE

Define-se a fase de integração ou fusão como o processo de ajustar uma representação de superfície *única* a partir de dados 3D de duas ou mais imagens (Chen e Medioni, 1992; Turk e Levoy, 1994; Chen e Medioni, 1995).

As imagens de profundidade guardam implicitamente a estrutura ou forma 3D de um objeto. Portanto, o objetivo da reconstrução de superfície é estimar uma réplica da topologia da superfície do objeto real a partir dos dados. Na solução desse problema estimação utiliza-se naturalmente métodos de interpolação e regressão em 3D.

Algoritmos de reconstrução precisam se comportar de forma robusta diante de inúmeros problemas associados com imagens profundidade. Citam-se:

- amostragem esparsa e irregular;
- oclusões ou brechas provenientes da geometria difícil de objetos de forma livre, como regiões encobertas por saliências da própria superfície do objeto ou regiões pontiagudas paralelas aos raios da iluminação;
- regiões digitalizadas fora do perímetro da superfície do objeto
- contaminação por ruído.

Existem propriedades que são muito desejáveis para um modelo e que influenciam o algoritmo de reconstrução de superfície a ser utilizado. Por exemplo, o modelo a ser obtido não pode ser ambíguo, ou seja, não podem representar dois objetos reais diferentes ao mesmo tempo. Isso é imprescindível em aplicações que fazem reconhecimento ou inspeção baseada em imagens. Curless e Reed listaram um conjunto de propriedades desejáveis para algoritmos de reconstrução de superfície (Reed, 1998; Curless, 1997).

- Representação de incertezas: algoritmo deveria considerar o modelo de erro do sensor.
- Eficiência de tempo e espaço: grande quantidade de imagens de profundidade é necessária para construir um modelo detalhado. Deve-se ter a preocupação de otimizar os parâmetros de tempo e espaço de máquina para que o algoritmo seja prático.
- Robustez: o algoritmo precisa ser robusto a dados discrepantes.
- Não impor restrições no tipo topológico: não pode haver restrições à forma do objeto. Por exemplo, não se pode assumir que ele é homeomórfico a um gênero geométrico, como uma esfera ou um poliedro.

- Habilidade para preencher buracos na reconstrução: o algoritmo deveria completar de forma coerente e inteligente os buracos causados por regiões inacessíveis pelo sensor, gerando um modelo a “prova d’água” e prazeroso esteticamente;
- Utilização de todos os dados: redundância pode ajudar na redução de ruído do sensor;
- Automação do planejamento de ponto de vista.

Duas abordagens principais tem sido adotadas para a reconstrução de superfície:  *fusão de malhas e reconstrução volumétrica*.

A reconstrução pode ser feita localmente, em cada imagem. Para isso utiliza-se os parâmetros de movimento obtidas na fase de registro para casar as malhas. Em seguida as regiões de sobreposição são costuradas utilizando um algoritmo de *zippering* (Turk e Levoy, 1994). Nas abordagens volumétricas o algoritmo percorre o conjunto de dados verificando quais são os *voxels* que pertencem à fronteira entre um espaço interior e exterior.

### 5.3 - VISUALIZAÇÃO DO MODELO

A primeira forma de avaliação de um modelo é naturalmente por meio visual. Uma vez que a descrição do modelo ainda está armazenada no computador na forma de um arquivo ou estrutura de dados contendo coordenadas tridimensionais, imagens de intensidade do modelo precisam ser geradas a partir desses dados para sua apresentação. O processo de apresentação ou exibição do modelo 3D através de imagens de intensidade é chamado de *renderização*. Matematicamente a renderização é uma projeção dos dados na forma 3D para pixels no plano de imagem 2D. Aplicações para efetuar essa projeção podem ser escritas usando funções de bibliotecas gráficas como *OpenGL* ou *Direct3D*. Parâmetros como o modelo de projeção (perspectiva ou ortogonal), definição de ponto de vista, tipo de iluminação, cor e refletância precisam ser especificados.

Existem vários métodos visualização, desde métodos com alto nível de realismo como o *traçado de raios (ray tracing)*, até forma mais simples, como renderizar um modelo 3D

através de imagens de intensidade transcrevendo-se propriedades de superfície para valores de intensidade.

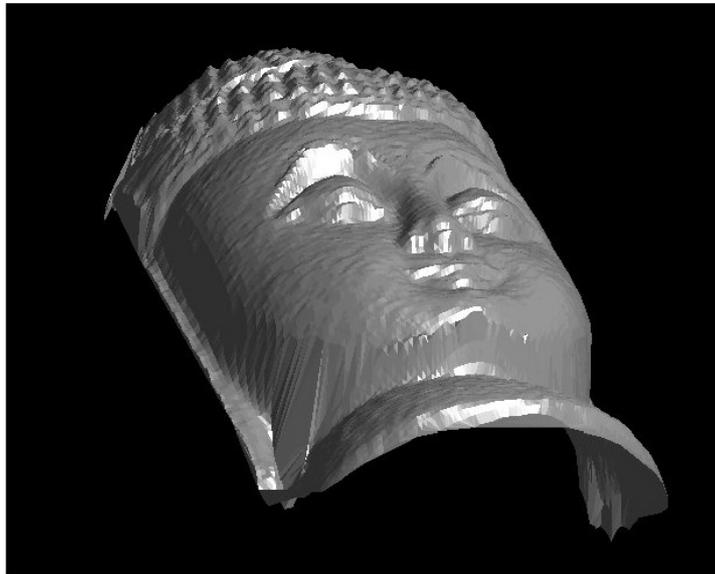


Figura 5.2 - Renderização em malha de facetas triangulares de uma imagem de profundidade utilizada nos procedimentos experimentais deste trabalho.

## CAPÍTULO 6 – PROCEDIMENTOS E RESULTADOS EXPERIMENTAIS

### 6.1 - PROCEDIMENTOS EXPERIMENTAIS

Como já mencionado neste trabalho, a metodologia básica para se atestar a convergência do modelo na reconstrução 3D consiste primariamente da simples crítica visual. À parte essa abordagem, pode-se verificar o comportamento dos dados ao longo da execução do algoritmo. O algoritmo tem sucesso se os valores forem monótonos decrescentes; no entanto, duas execuções podem não ter os mesmos resultados e o mesmo comportamento, devido ao caráter aleatório da amostragem de pontos.

Os procedimentos experimentais nesta pesquisa obedeceram a um roteiro de implementação do algoritmo *ICP* e posteriores testes com variações dos parâmetros de estimativa inicial, número de pontos de controle e níveis de ruído. As implementações totalizaram mais de 50 algoritmos divididos em módulos, que na verdade constituem-se bibliotecas de algoritmos voltados para uma finalidade comum, como registro ou renderização. Os testes foram realizados em dois cenários:

- Primeiro utilizando-se dados 3D sintéticos. Tais dados consistiram das coordenadas dos vértices de dois cubos. Os testes com esses dados tridimensionais simples e puros, ou seja, sem incertezas, serviram para possibilitar a introdução controlada de ruído, e assim se avaliar a qualidade dos pares obtidos e o comportamento da convergência do algoritmo com diferentes estimativas de movimento inicial. Nesse caso tem-se conhecimento sobre os parâmetros de movimento corretos e a relação sinal/ruído, e, portanto, é possível tirar conclusões sobre a convergência do *ICP* através de comparações entre os valores corretos e os resultados estimados. Nestes testes o algoritmo foi implementado em *MATLAB* e executado na presença de diferentes e variadas porcentagens de ruído, adicionados aos dados dos cubos na forma de pontos aleatórios nas proporções de: 0%, 12,5%, 25% e 50%.
- No segundo cenário foram efetuados testes com imagens de profundidade reais obtidas por meio de um digitalizador 3D baseado em laser. Essa é uma situação onde a quantidade de dados é bastante grande e há a presença de ruído. Para este caso o algoritmo *ICP* foi

implementado em linguagem C, pois a complexidade computacional de tempo e espaço, principalmente no aspecto da utilização de memória, desfavorece a implementação em *MATLAB*. As imagens utilizadas estão disponibilizadas publicamente na base de dados de imagens de profundidade *OSU 3D Database*, da *Universidade de Ohio* (Campbell e Flynn, 1998). O formato dessas imagens consiste de três matrizes de tamanho 200x200, ou seja, uma matriz para cada uma das coordenadas X, Y e Z. A fonte disponibilizou também mais uma matriz que serve como máscara de filtragem e que identifica os valores válidos como 1, e os valores que deveriam ser descartados como 0. De um total de 40.000 pontos por imagem, a primeira imagem ficou com 14.866 pontos e a segunda imagem com 14.523 pontos, após o descarte de pontos não confiáveis. Ressalta-se que, mesmo com esse descarte de pontos não confiáveis, podem ser identificados ainda uma grande quantidade de dados discrepantes nas imagens. As imagens foram obtidas com movimento de rotação de 20 graus entre varreduras sucessivas, e sem movimento de translação. Essa mesma orientação foi introduzida como estimativa do movimento inicial. Um total de 100 pontos de controle para o primeiro teste e 500 pontos de controle para o segundo teste foram escolhidos aleatoriamente em uma imagem. A iteração do algoritmo pára quando o valor do erro RMS calculado em cada iteração atinge um valor menor que um limiar de 30% do erro RMS inicial. Para efetuar o registro e viabilizar uma avaliação visual do modelo sendo reconstruído, dois importantes módulos precisaram ser implementados: um para efetuar o registro de fato, e outro para efetuar a renderização dos pontos das imagens em cada iteração.

### 6.1.1 - Especificações do algoritmo ICP implementado

- **Objetivo:** Alinhar dados 3D nos cenários de dados de um cubo e de imagens de profundidade reais obtidas por varredura laser.
- **Abordagem de otimização:** minimização iterativa de uma função objetivo através do método de mínimos quadrados.
- **Estratégia de alinhamento:** alinhamento em pares.

- **Amostragem de pontos:** escolha aleatória de N pontos de controle apenas em uma imagem. Esses pontos escolhidos são a entrada para o algoritmo de combinação que monta os pares de pontos correspondentes.
- **Método de combinação:** busca exaustiva pelo ponto mais próximo (ponto a ponto).
- **Método de estimação de pose:** recuperação da matriz de rotação através do produto de duas bases ortonormais obtidas por decomposição da matriz de covariância utilizando SVD; e vetor de translação recuperado através da diferença dos centróides das imagens.
- **Critério de parada:** limiar para rotação, no caso do cubo; e tolerância de 30% do erro RMS inicial, no caso das imagens de profundidade.

### 6.1.2 - Considerações sobre o elemento ruído

No primeiro cenário de testes o ruído foi introduzido nos dois cubos na mesma proporção, e consistiu basicamente de pontos aleatórios gerados com a distribuição normal:

$$N(\mu_c, \sigma_c)$$

Onde:

$\mu_c$  é a média dos dados do cubo;

$\sigma_c$  é o desvio padrão populacional dos dados do cubo.

No caso das imagens de profundidade reais nenhum conhecimento a priori sobre o modelo do ruído está disponível.

## 6.2 - RESULTADOS EXPERIMENTAIS COM DADOS 3D SINTÉTICOS DOS VÉRTICES DE UM CUBO

### 6.2.1 - Primeiro teste: dados sem ruído

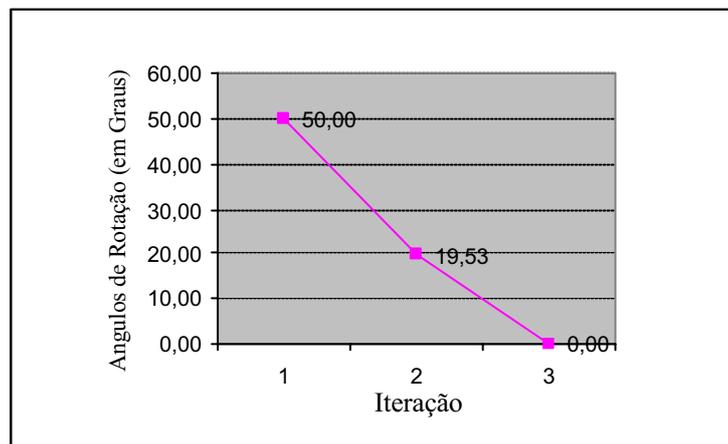


Figura 6.1 - Ângulos rotação estimados em função da iteração no experimento dos cubos sem ruído.

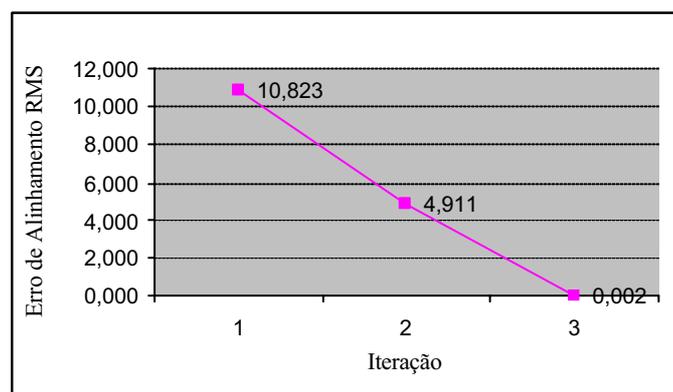
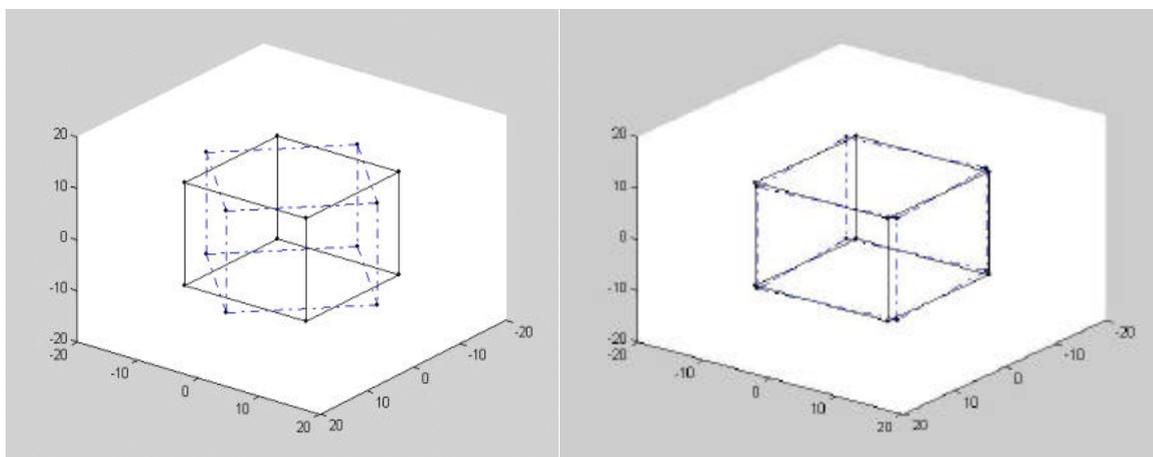
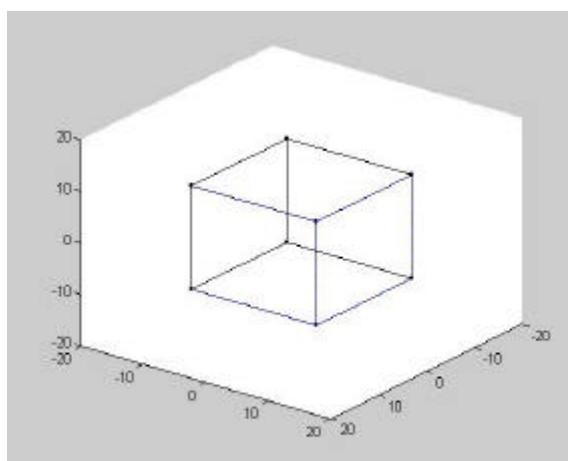


Figura 6.2 - Erro de alinhamento RMS em função da iteração no experimento dos cubos sem ruído.



(a)

(b)



(c)

Figura 6.3 - Renderização do alinhamento dos cubos sem ruído: (a) Cubos na primeira iteração; (b) cubos na segunda iteração; (c) cubos na última iteração.

### 6.2.2 - Segundo teste: 12,5% de ruído

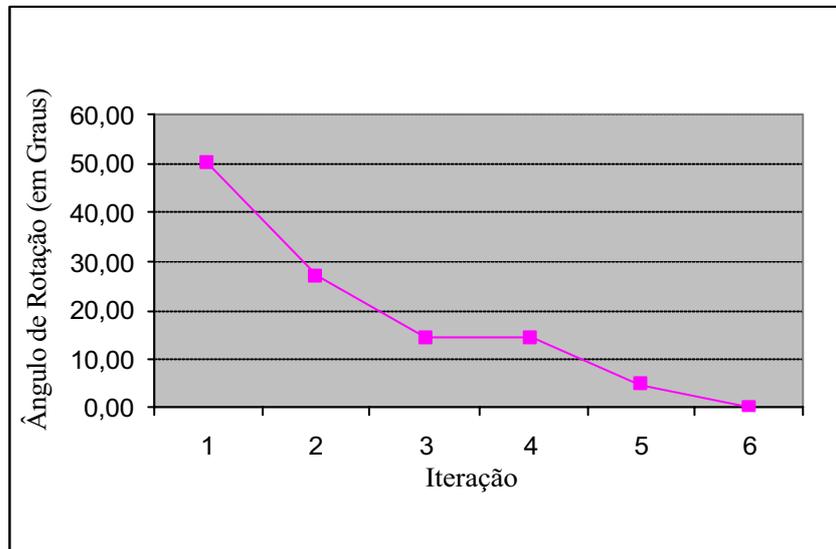


Figura 6.4 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 12,5% de ruído.

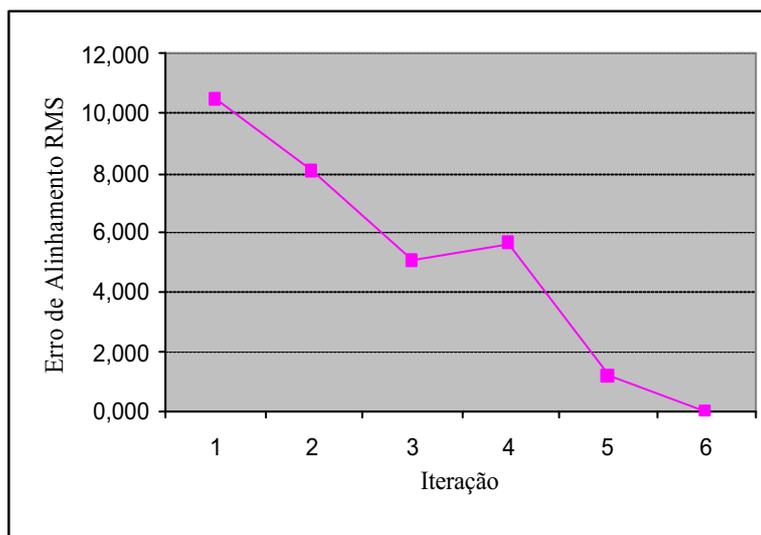
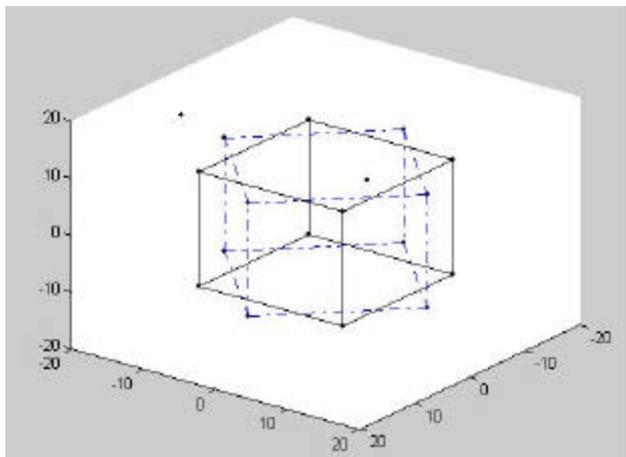
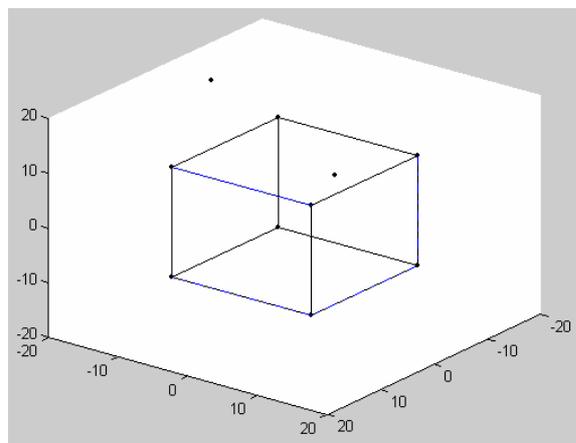


Figura 6.5 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 12,5% de ruído.



(a)



(b)

Figura 6.6 - Renderização do alinhamento dos cubos com 12,5% de ruído: (a) Cubos na primeira iteração, (b) cubos na última iteração.

### 6.2.3 - Terceiro teste: 25% de ruído

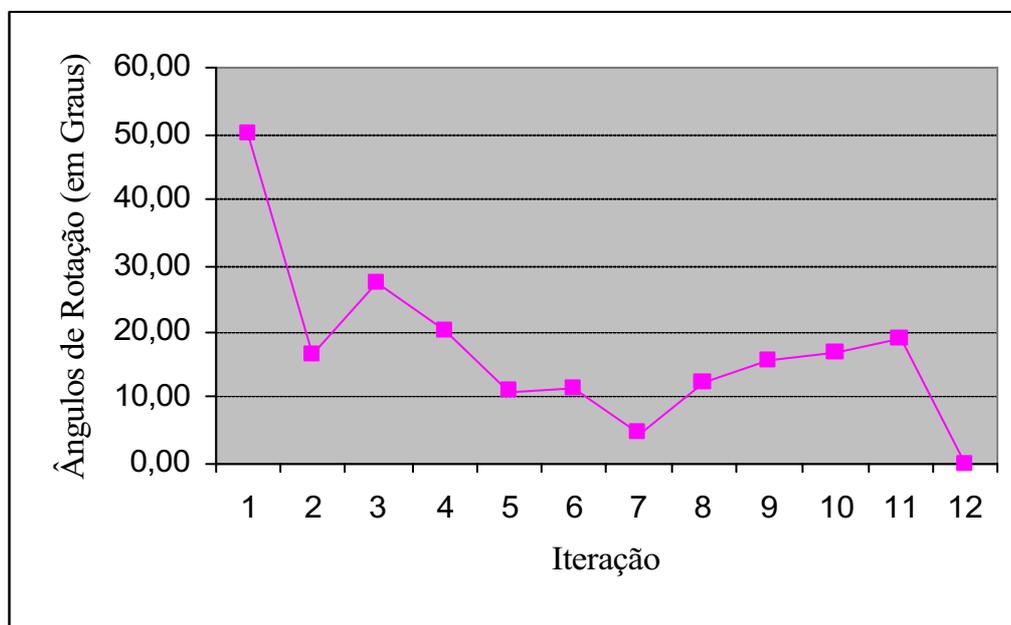


Figura 6.7 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 25%.

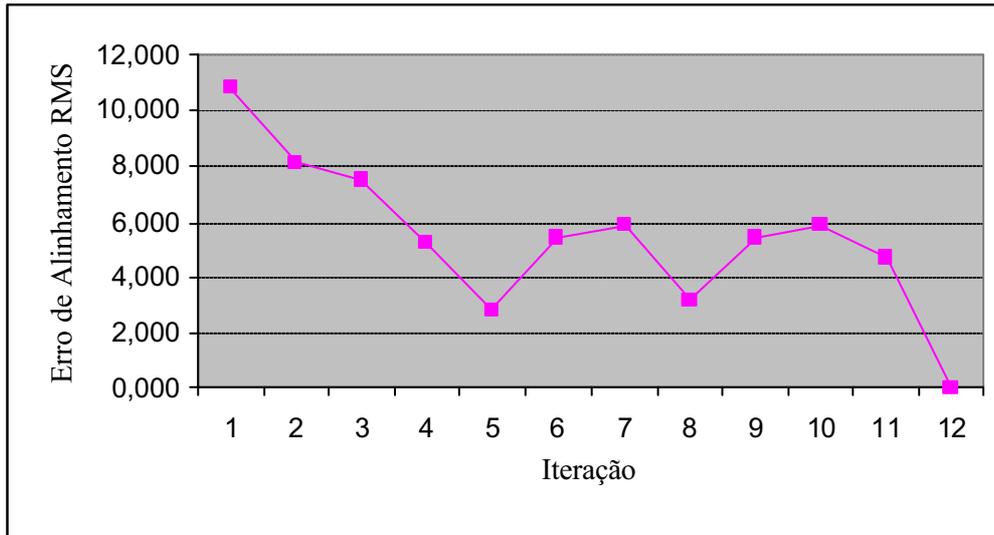
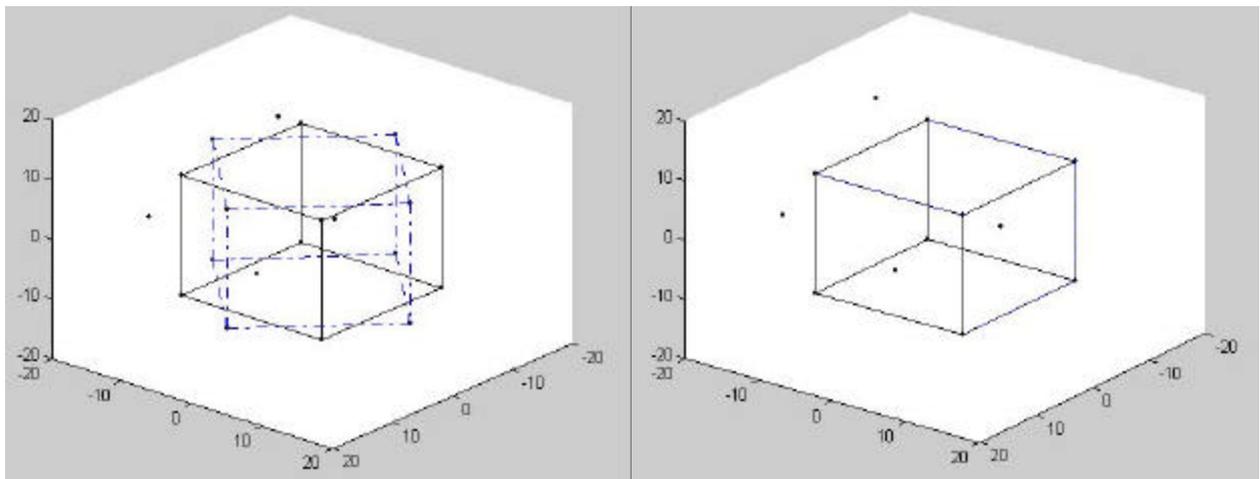


Figura 6.8 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 25% de ruído.



(a)

(b)

Figura 6.9 - Renderização do alinhamento dos cubos com 25% de ruído. (a) Cubos na primeira iteração e (b) cubos alinhados.

#### 6.2.4 - Quarto teste: 50% de ruído

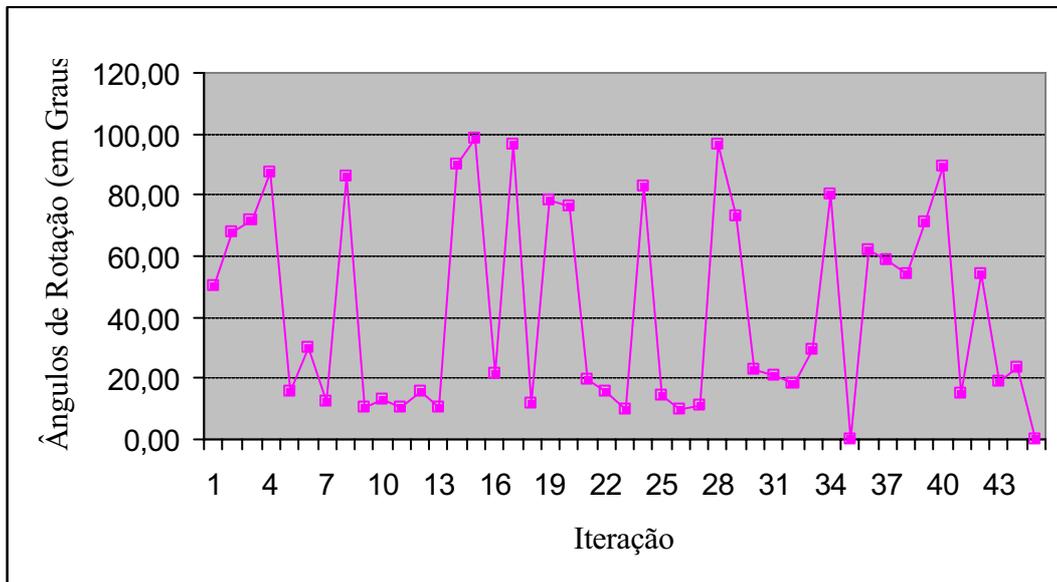


Figura 6.10 - Ângulos rotação estimados em função da iteração no experimento dos cubos com 50% de ruído.

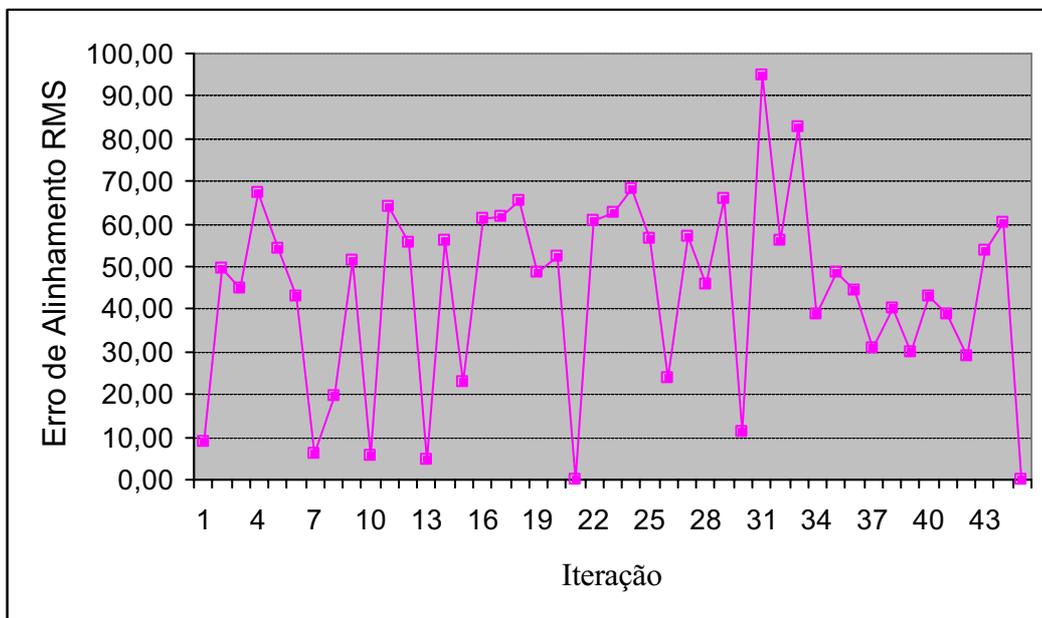
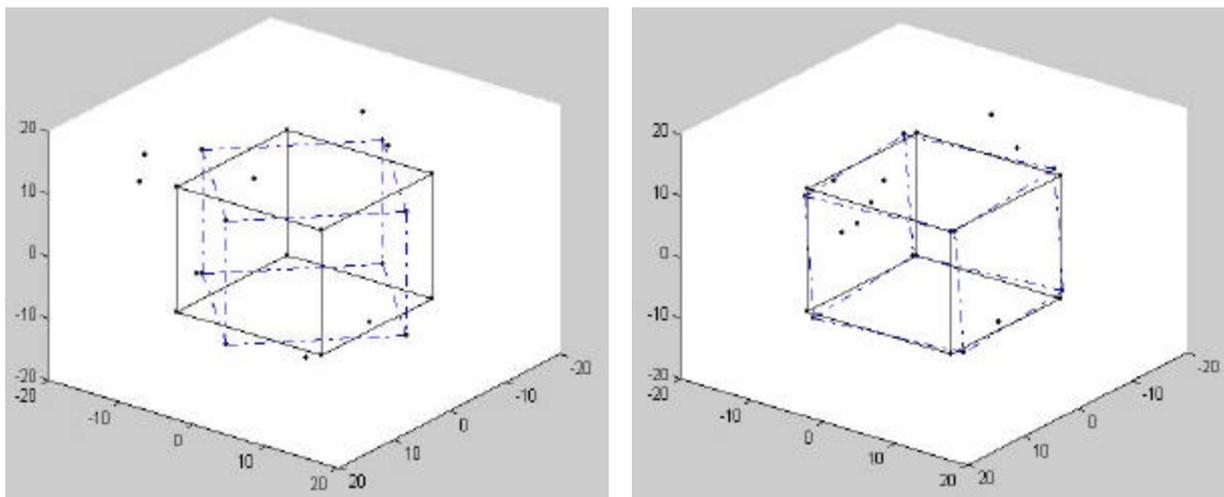


Figura 6.11 - Erro de alinhamento RMS em função da iteração no experimento dos cubos com 50% de ruído.



(a)

(b)

Figura 6.12 - Renderização do alinhamento dos cubos com 50% de ruído. (a) Cubos na primeira iteração e (b) cubos alinhados.

Tabela 6.1 - Dados obtidos nos testes de execução do algoritmo do *ICP* com dados de dois cubos rotacionados 45° em relação ao eixo *Z*.

	1° Teste	2° Teste	3° Teste	4° Teste
<b>Incertezas</b>	Sem Ruído	12,5% de ruído de ruído branco gaussiano por imagem	25% de ruído branco gaussiano por imagem	50% de ruído de ruído branco gaussiano por imagem
<b>Estimativa da Rotação Inicial</b>	50°	50°	50°	50°
<b>Número de Pontos de Controle</b>	5	5	5	5
<b>Critério de Parada</b>	Limiar de 1° para a rotação	Limiar de 1° para a rotação	Limiar de 1° grau para a rotação	Limiar de 1° para a rotação
<b>Tempo de execução (em segundos)</b>	1,15 s	1,54 s	1,67 s	5,53 s
<b>Eixo de Rotação Estimado</b>	(0; 0; 1)	(0,155; -0,614; 0,774)	(-0,191; -0,461; -0,867)	(0,04; 0,02; -1)

Tabela 6.2 - Ângulos de rotação estimados e o erro RMS correspondente a cada iteração.

Iterações	1º Teste		2º Teste		3º Teste		4º Teste	
	Ângulo	RMS	Ângulo	RMS	Ângulo	RMS	Ângulo	RMS
1	50,00°	10,823	50,00°	10,431	50,00°	10,823	50,00°	8,68
2	19,53°	4,911	26,95°	8,023	16,53°	8,145	67,92°	49,41
3	0,00°	0,002	14,23°	5,029	27,46°	7,475	71,77°	44,79
4			14,23°	5,608	20,23°	5,206	87,12°	67,2
5			4,84°	1,137	10,95°	2,827	15,88°	54,1
6			0,00°	0,002	11,41°	5,432	29,70°	42,96
7					4,46°	5,849	12,19°	6,29
8					12,13°	3,164	86,34°	19,66
9					15,56°	5,432	10,45°	51,46
10					16,75°	5,843	13,31°	5,55
11					18,88°	4,65	10,22°	63,83
12					0,00°	0,002	15,63°	55,61
13							10,72°	4,69
14							90,29°	56,27
15							98,77°	23,02
16							21,68°	61,34
17							96,53°	61,87
18							11,88°	65,29
19							78,38°	48,66
20							76,40°	52,37
21							19,33°	7,64
22							15,85°	60,8
23							10,03°	62,64
24							83,13°	68,26
25							14,03°	56,41
26							10,07°	23,84
27							11,23°	57
28							96,33°	45,73
29							72,97°	66,09
30							22,61°	11,16
31							20,62°	94,82
32							18,50°	56,08
33							29,61°	82,94
34							80,37°	38,76
35							10,9	48,54
36							62,06°	44,19
37							58,93°	30,71
38							53,92°	40,34
39							71,00°	29,7
40							89,35°	42,97
41							15,15°	38,8
42							54,31°	28,74
43							18,62°	53,66
44							23,80°	60,3
45							0,00°	0

## 6.3 - RESULTADOS EXPERIMENTAIS COM IMAGENS DE PROFUNDIDADE REAIS

### 6.3.1 - Primeiro teste

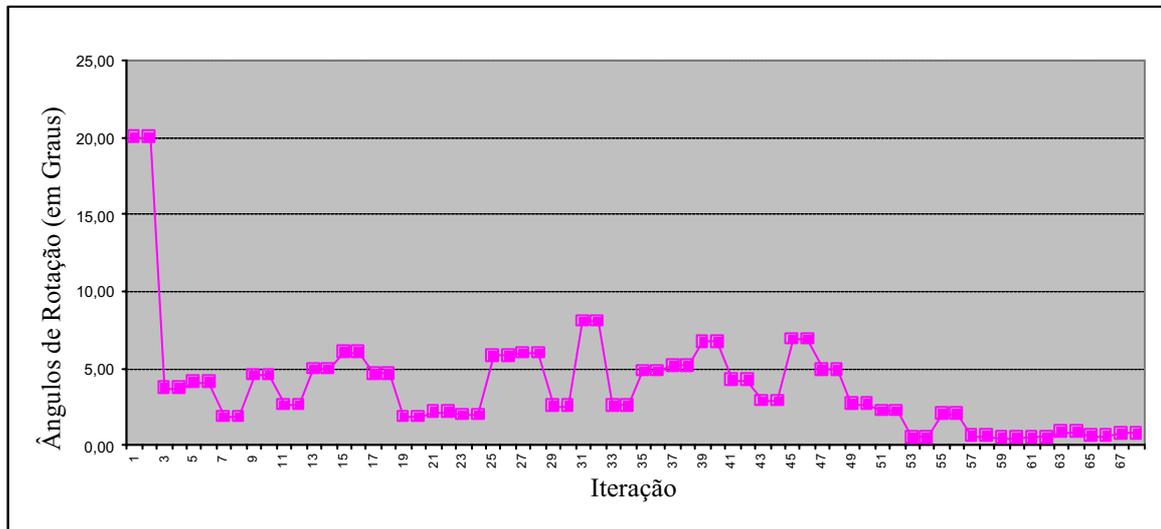


Figura 6.13 - Ângulos rotação estimados em função da iteração no 1º teste de alinhamento do par de imagens de profundidade.

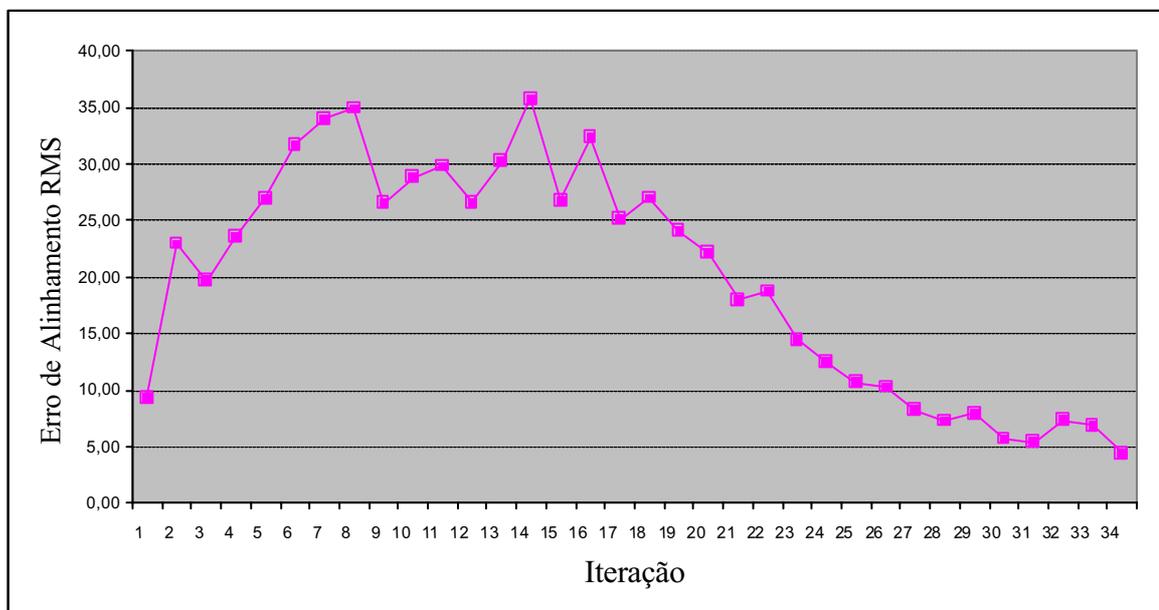
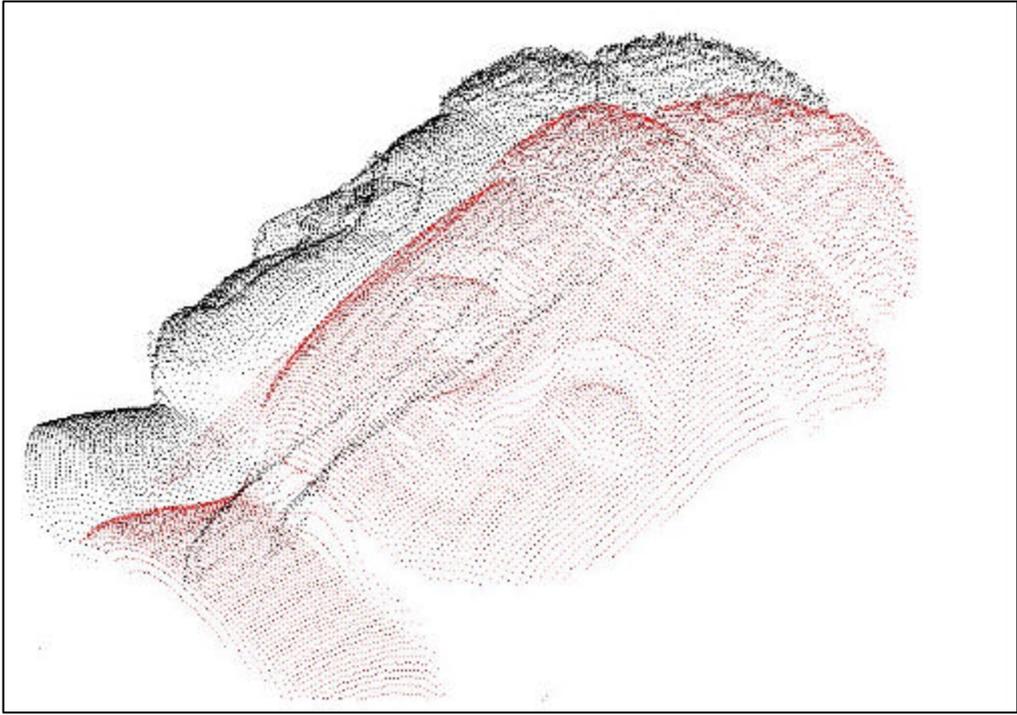
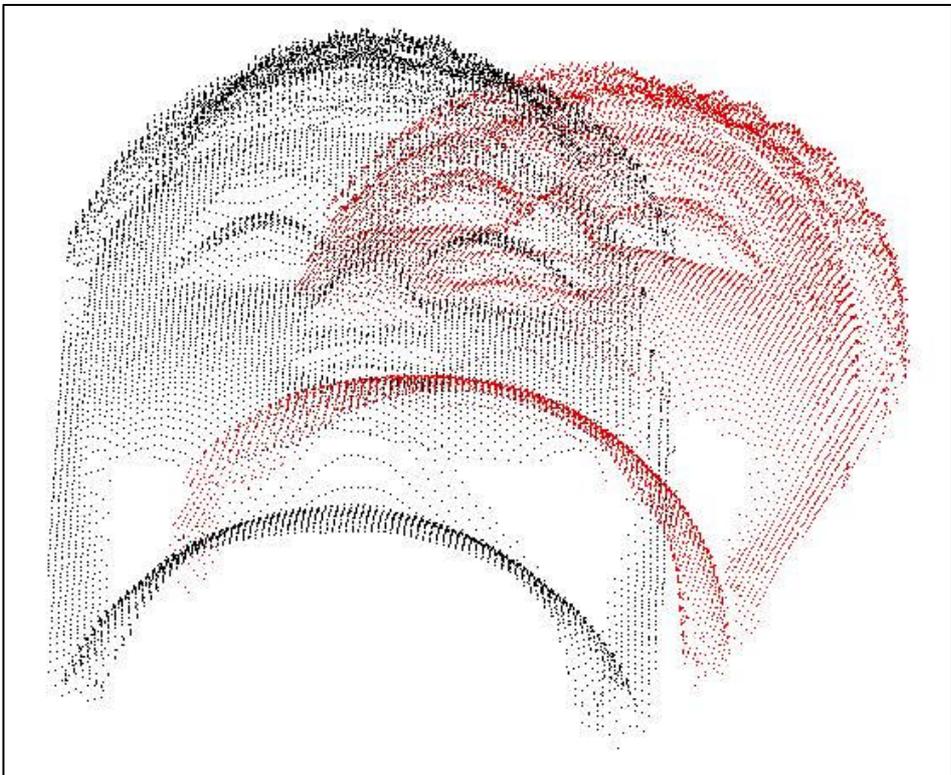


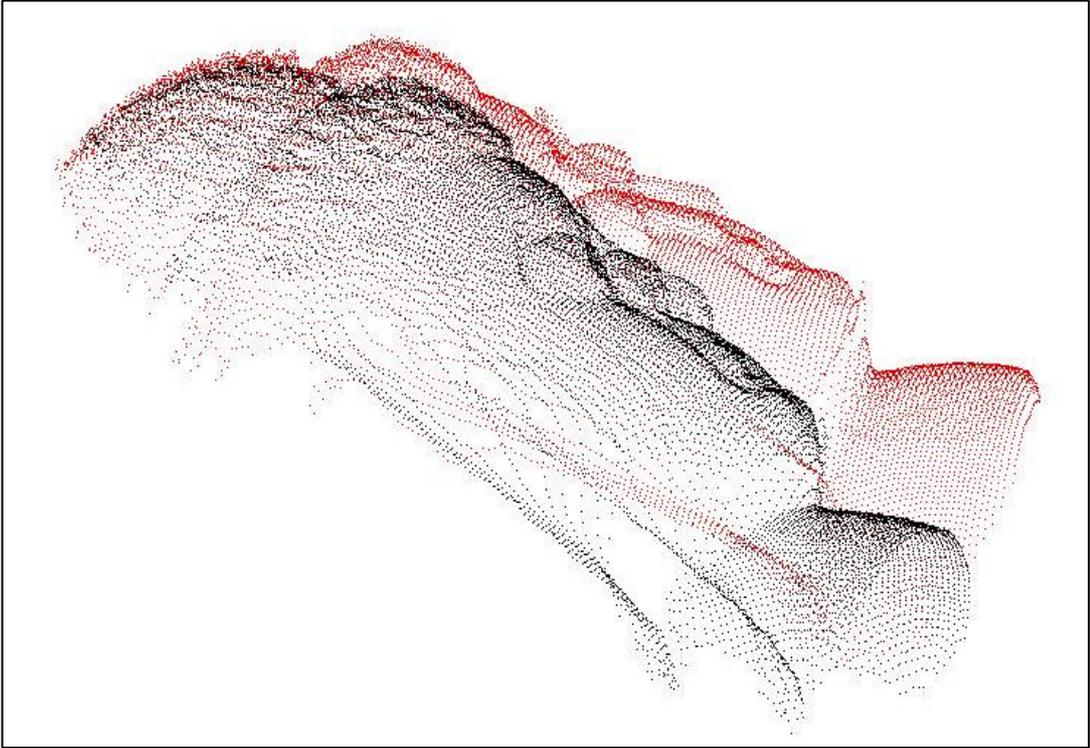
Figura 6.14 - Erro de alinhamento RMS em função da iteração no 1º teste de alinhamento do par de imagens de profundidade.



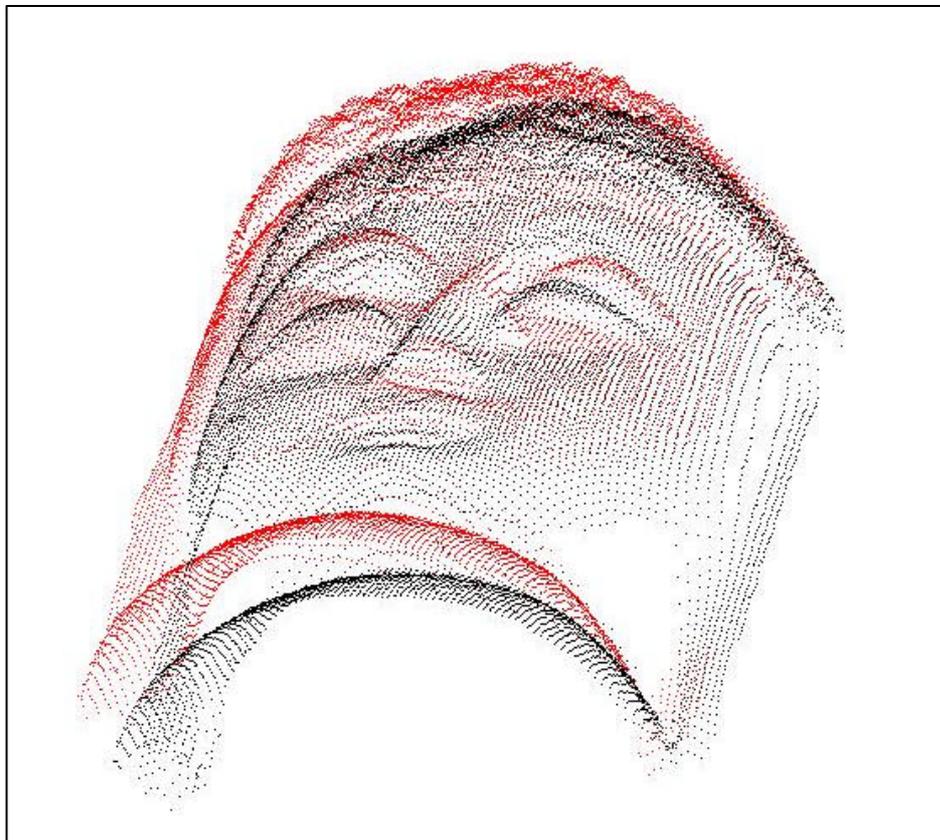
(a)



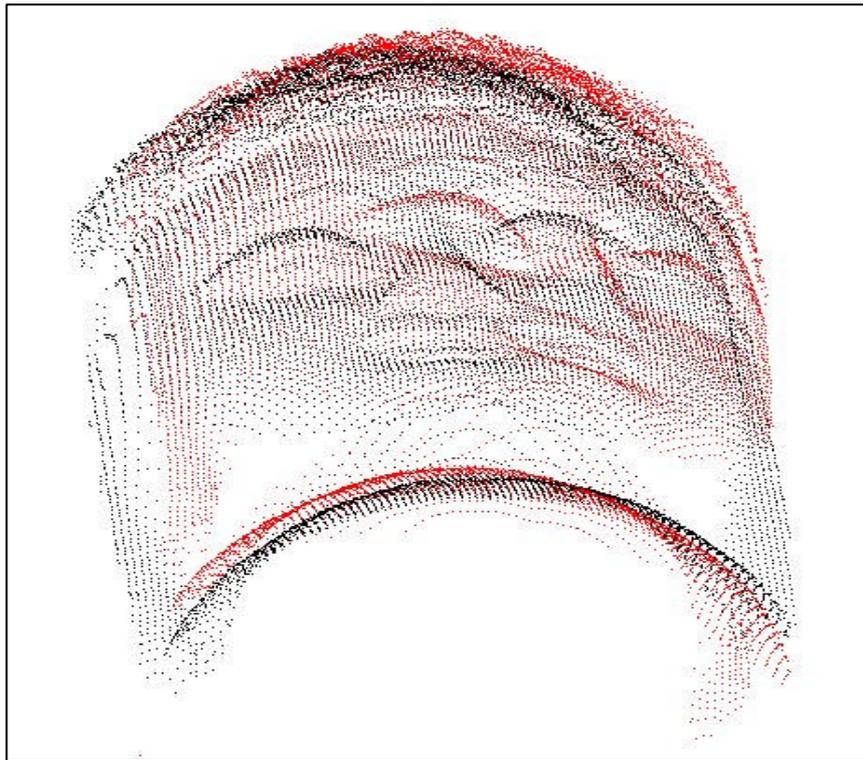
(b)



(c)



(d)



(e)

Figura 6.15 - De (a) a (e): renderização como pontos do alinhamento do par de imagens de profundidade no 1º teste. A primeira imagem (a) é o alinhamento inicial, e a última imagem (e), é do alinhamento final.

### 6.3.2 - Segundo teste

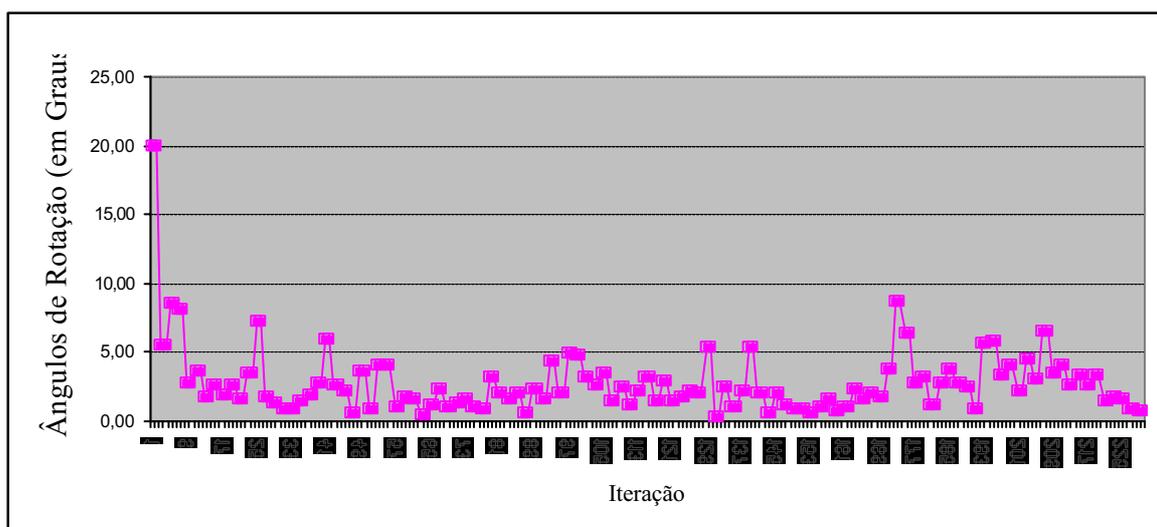


Figura 6.16 - Ângulos rotação estimados em função da iteração no 2º teste de alinhamento do par de imagens de profundidade.

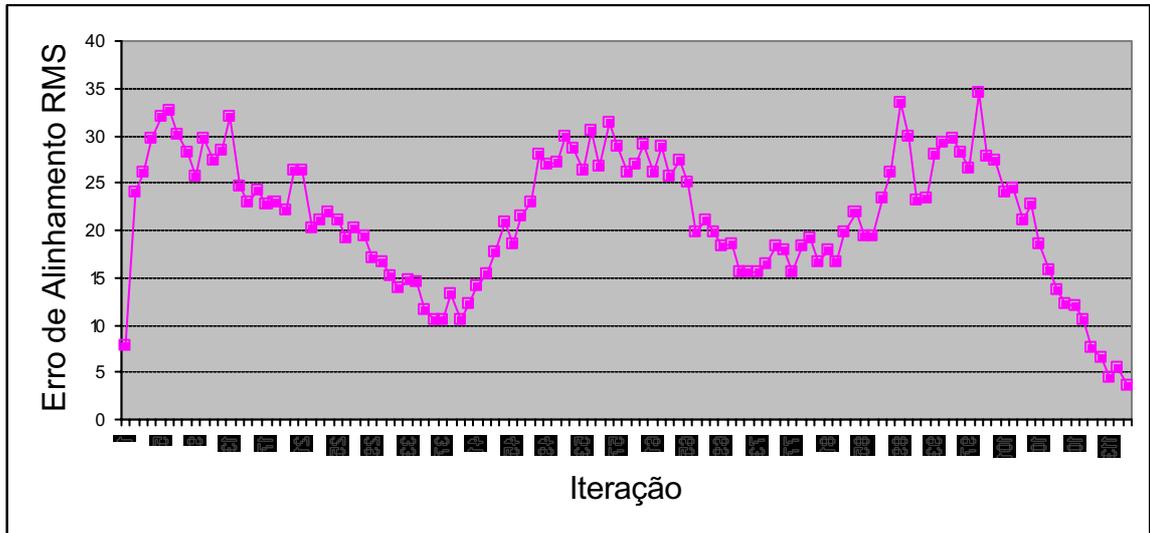
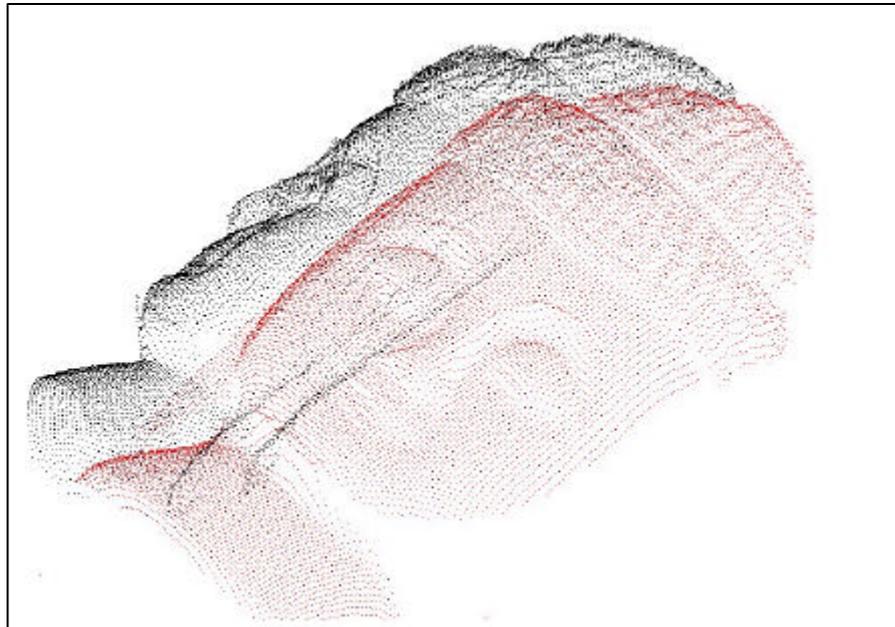
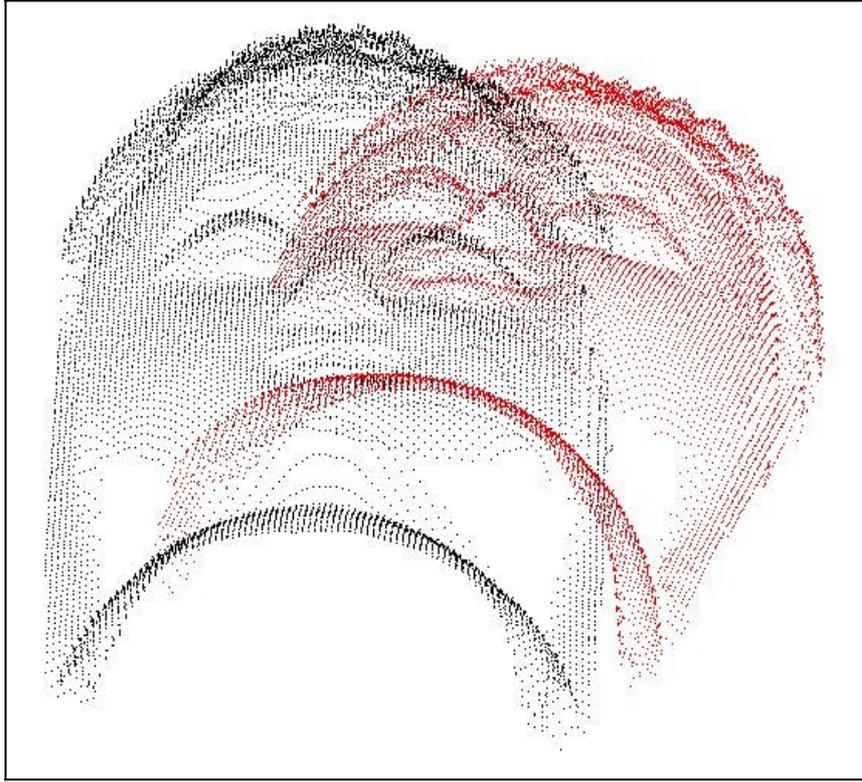


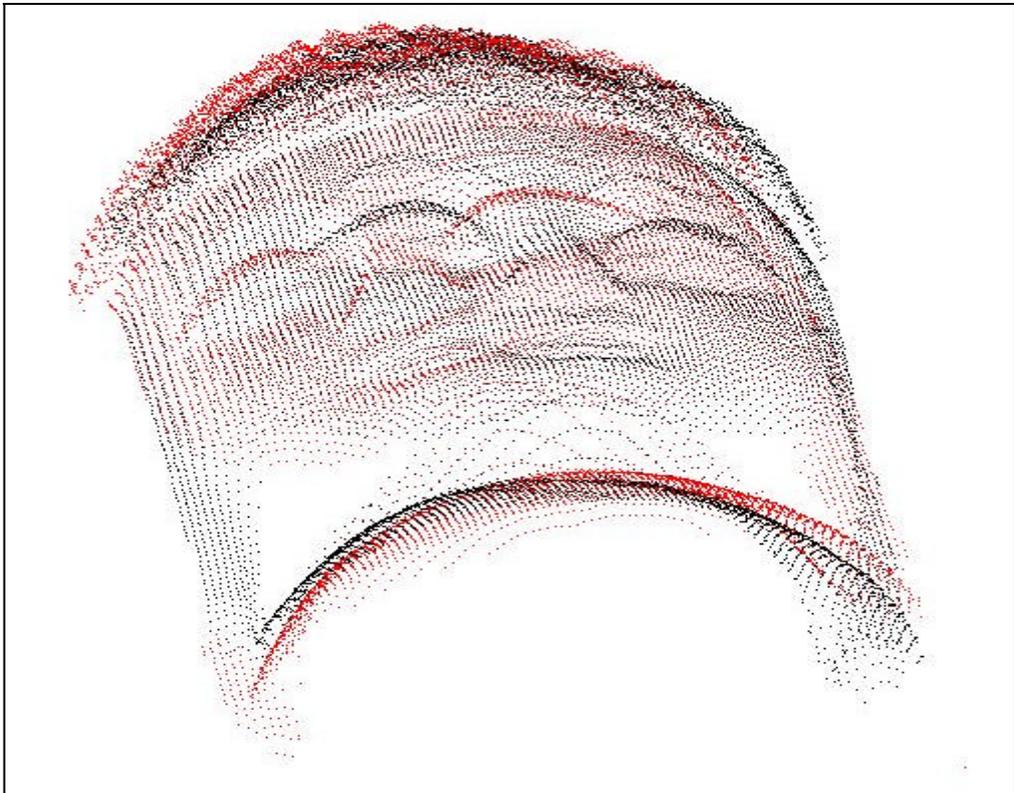
Figura 6.17 - Erro de alinhamento RMS em função da iteração no 2º teste de alinhamento do par de imagens de profundidade.



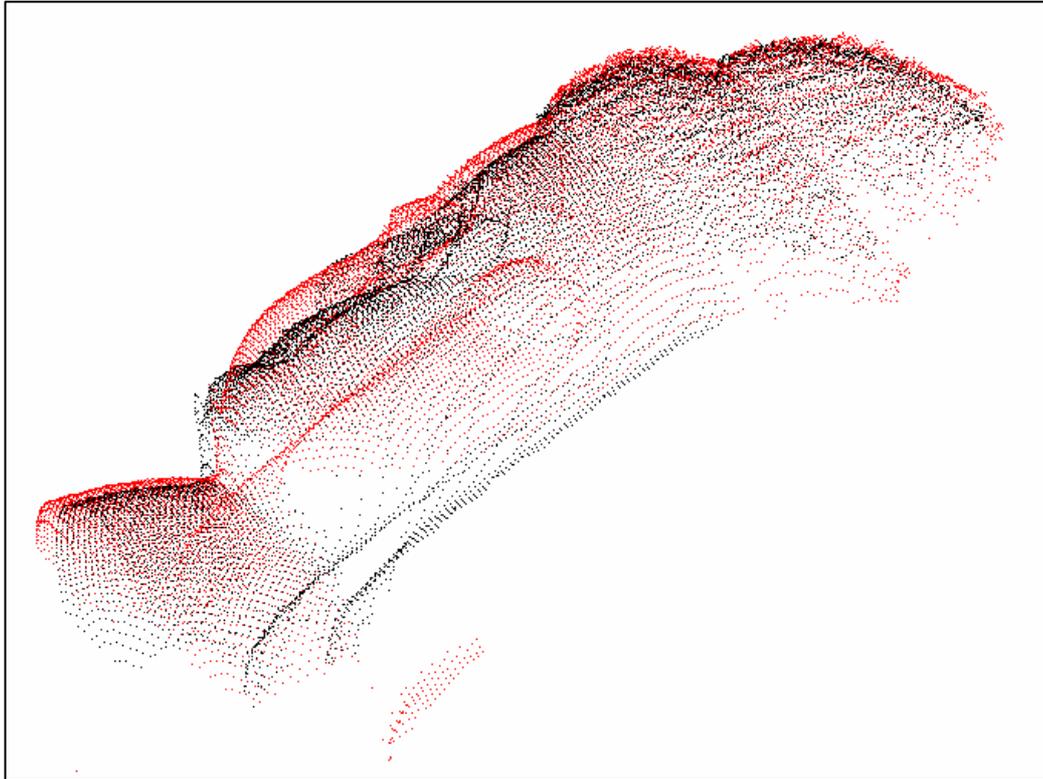
(a)



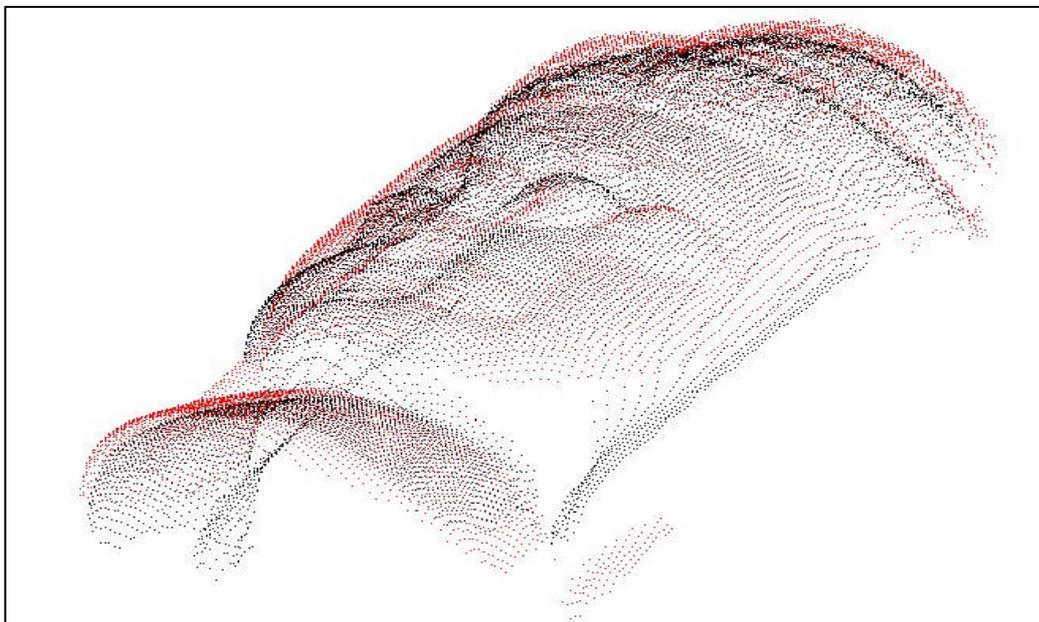
(b)



(c)



(d)



(e)

Figura 6.18 - De (a) a (e): renderização como pontos do alinhamento do par de imagens de profundidade no 2º teste. A primeira imagem (a) é do alinhamento inicial, e a última imagem (e) é do alinhamento final.

## **CAPÍTULO 7 – ANÁLISE DOS RESULTADOS**

### **7.1 - CONSIDERAÇÕES COM RELAÇÃO AOS TESTES COM O CUBO**

No primeiro teste, com dados 3D sem ruído, pode ser visualizado no gráfico da Figura 6.2 que os valores dos erros RMS obtidos nas iterações formam uma seqüência monotonamente decrescente. O mesmo também pode ser verificado com a seqüência de rotações estimadas na Figura 6.1. Esse comportamento ocorre porque, numa situação ideal, ou seja, com pares correspondentes corretos, o método dos mínimos quadrados (núcleo do *ICP*) cumpre o seu papel adequadamente, possibilitando a estimação dos parâmetros de movimento corretos (neste caso são somente rotações) que minimizam o erro médio quadrático em cada iteração.

No segundo, terceiro e quarto testes pode-se perceber que os erros RMS calculados formam uma seqüência com aspecto decrescente, mas não monótona. A explicação para esse comportamento vem da análise da seqüência de rotações estimadas. Como há a presença de ruído, a seleção aleatória de pontos de controle pode incluir pontos espúrios, conseqüentemente, a busca de pontos correspondentes para esses dados espúrios dentro da outra imagem monta falsos pares correspondentes. As falsas correspondências por sua vez comprometem a matriz de covariância. E por fim, a matriz de rotação estimada a partir dessa matriz de covariância pode estar degenerada, ou seja, ela não representa uma matriz de rotação verdadeiramente. Pode ser visualizado no gráfico da rotação que o valor de uma rotação estimada nem sempre é menor que a rotação estimada anteriormente. Desse fato pode-se concluir que o estimador de mínimos quadrados ficou enviesado pela presença de pareamentos errôneos, e por isso proporcionou a estimação de valores de movimento rígido incorretos.

### **7.2 - CONSIDERAÇÕES COM RELAÇÃO AOS TESTES COM IMAGENS DE PROFUNDIDADE REAIS**

Quanto aos testes com as imagens de profundidade reais, a percepção do alinhamento não-ótimo é evidente a partir da análise visual. Os gráficos das rotações estimadas e erros RMS obtidos revelam um comportamento pior da convergência das imagens em relação aos

testes do cubo com ruído, devido à quantidade de incertezas presentes no conjunto de dados.

De um modo geral, pode-se concluir que o algoritmo *ICP* clássico, implementado sem dispositivos que efetuem uma filtragem ou rejeição de dados ou pares discrepantes, e utilizando otimização não-robusta, não apresenta garantias mínimas de convergência. Tecnicamente, a função de custo do *ICP* é do tipo não-convexa, significando que existem muitos pontos de mínimo local. Na busca pelo mínimo global, o algoritmo pode enterrar em um mínimo não-global e parar, nunca atingindo o desejado mínimo global. Claramente, o método dos mínimos quadrados do *ICP* não se comporta bem em situações que violam as hipóteses de máxima verossimilhança sob as quais ele foi projetado.

### **7.3 - ASPECTOS QUE INFLUENCIAM A CONVERGÊNCIA DO ALGORITMO**

Os resultados experimentais sugerem ainda as seguintes conclusões com relação à convergência do algoritmo:

- **Qualidade do alinhamento inicial**

Uma boa estimativa do alinhamento inicial é condição imprescindível para se encontrar o alinhamento ótimo, pois o *ICP* apenas refina essa estimativa inicial. Mesmo em condições onde os dados não se apresentarem contaminados com ruído, o algoritmo pode não convergir para o mínimo global se uma boa estimativa inicial não for empregada.

- **Escolha dos pontos de controle**

A rejeição de dados discrepantes pode começar na fase de escolha de pontos de controle. Uma escolha criteriosa de uma quantidade pequena de pontos de controle de modo a rejeitar dados discrepantes pode ser mais eficiente do que uso de uma quantidade grande de pontos a esmo, que podem conter ruído ou estarem fora da região de sobreposição. O número de pontos de controle também influenciam a complexidade, visto que esta ela cresce proporcionalmente ao número de pontos de controle.

- **Os métodos de combinação e rejeição de falsos pares**

Estratégias de combinação na etapa de estabelecimento de correspondências afetam decisivamente a convergência e velocidade do *ICP*. Os pareamentos errados perturbam função de custo e são ocasionados em pela violação das hipóteses anteriores.

- **Métrica de erro**

O núcleo da função de custo a ser montada é uma média. Cada par contribui com o quadrado da distância entre os seus pontos para formar a média de quadrados de distâncias, de acordo com a métrica de erro de mínimos quadrados. Ressalta-se que a média é uma medida de tendência central não-robusta a dados discrepantes, e que apenas um dado discrepante é suficiente para perturbá-la. Mínimos quadrados, a partir de uma perspectiva de máxima verossimilhança, só apresenta resultados não-tendenciosos quando a hipóteses de normalidade da distribuição e homogeneidade da variância são válidas. Imagens de profundidade muitas vezes violam essas hipóteses.

- **Métodos de recuperação da transformação 3D**

Os métodos disponíveis para a representação e recuperação do movimento rígido são bem condicionados e precisos. A única preocupação é quanto ao reforço da ortonormalidade para evitar a obtenção de rotações degeneradas, quando não se utiliza a representação de quaternions. Conclui-se que todas as etapas influenciam a convergência do *ICP*. A convergência depende de fatores como: a qualidade do alinhamento inicial, dos métodos de escolha, dos métodos de combinação de pontos, dos métodos de minimização da função de custo e dos métodos de recuperação da transformação 3D.

## **CAPÍTULO 8 – CONCLUSÕES E SUGESTÕES PARA PESQUISAS FUTURAS**

### **8.1 - CONCLUSÕES**

A tecnologia da reconstrução 3D de modelos digitais a partir de imagens de profundidade obtidas por varredura laser é o estado-da-arte para aquisição de modelos. A abordagem consiste de três etapas: a primeira é a aquisição de dados, a segunda é o registro de imagens, e a terceira é a integração de imagens e a reconstrução de superfície.

Esta dissertação faz uma revisão dos estágios e métodos envolvidos em cada etapa da reconstrução 3D de modelos digitais e apresenta uma implementação do algoritmo *ICP* para o alinhamento de múltiplas imagens de profundidade obtidas através de varredura laser, com o propósito de construir modelos geométricos digitais para aplicações da engenharia que tem por foco a exatidão do modelo.

A etapa da aquisição de dados é de extrema importância para a reconstrução 3D. No contexto da engenharia moderna, o paradigma de desenvolvimento rápido de protótipo, eficiente e a baixo custo podem tornar até mesmo proibitivo o uso de instrumentos lentos e que não podem captar com riqueza de detalhes geometrias de superfícies complexas, como é a característica das superfícies de objetos reais de forma livre. Os tradicionais digitalizadores de contato são altamente exatos, mas eles também são lentos e fornecem baixa taxa de amostragem. O progresso dos equipamentos de iluminação e o aumento da capacidade de processamento de computadores combinada com a redução de custo desses aparatos vêm propiciando o avanço da tecnologia de obtenção de modelos por digitalização e estabelecendo como tendência a digitalização óptica sem contato. Mesmo estando em constante aprimoramento, pode-se dizer que a abordagem óptica a laser já atingiu um estágio de desenvolvimento tecnológico que a qualifica para ser usada para a construção de modelos com alto grau de precisão. Além de serem rápidos, compactos e gerarem uma grande e densa quantidade de dados, outras razões para utilização dessa categoria de digitalização são a segurança, devido a utilização de laser de baixa intensidade, e a precisão, que é suficiente para ser usada na engenharia reversa de modelos digitais de alto nível.

O problema do alinhamento ou registro, foco deste trabalho, é de caráter essencialmente particular, não existindo nenhuma solução conclusiva e generalista. Testes experimentais com o algoritmo *ICP* demonstraram que, em contrapartida a rapidez, facilidade de implementação e generalidade, o seu ponto fraco é o mesmo mal que aflige métodos numéricos em geral: hipóteses precisam ser satisfeitas para que haja garantia mínima de convergência. Sem a imposição de restrições o algoritmo pode convergir para um mínimo não global, visto que a função de custo é não-convexa, resultando em um alinhamento incorreto e conseqüentemente em um modelo distorcido.

## 8.2 - SUGESTÕES PARA PESQUISAS FUTURAS

Ficou evidenciado neste estudo que a construção de modelos geométricos é um tema atual e um elemento chave para diversos campos de conhecimento. Nesse sentido, todas etapas da construção de modelos digitais pela reconstrução 3D, bem como as etapas do algoritmo *ICP*, representam férteis campos de estudo. Com respeito especificamente ao algoritmo apresentado, duas direções de pesquisa tem ficado mais evidentes. A primeira é voltada para a obtenção de pontos de controle e pares correspondentes mais confiáveis, que caracteriza uma abordagem mais voltada para o tratamento dos dados diretamente, como a filtragem para a rejeição de ruído e a utilização de dispositivos para descartar ou evitar falsos pares correspondentes. A segunda direção tem por foco a função de custo do *ICP*, que utiliza ferramentas oriundas de áreas como a otimização, a estatística e a visão computacional robusta.

Alguns tópicos sugestivos de estudo são:

- O alinhamento inicial, pois trata-se de um tema ainda pouco explorado. Nesse aspecto o termo “estimativa inicial boa” é relativo e precisa de melhor análise e conceituação.
- A recuperação de movimento é uma etapa que já foi provada apresentar boas soluções, especialmente quando se utiliza quaternions, apresentando poucas diferenças entre os métodos disponíveis, além de apresentar atrativas soluções de forma fechada quando se utiliza correspondência ponto a ponto. Entretanto, um estudo da estimação da matriz de rotação em situações de grande quantidade de ruído é salutar.

- Considerando que o componente ruído deve ser modelado como variável aleatória, pesquisas com abordagens probabilísticas podem apresentar bons resultados.
- Existem poucas pesquisas na literatura no sentido de se testar estimadores robustos para a função de custo do *ICP*.
- Temas como a validação de registro e o registro global são campos que geram uma desconfiança de que há ainda muitas idéias a serem propostas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Aloimonos, Y. e Rosenfeld, A. (1993). "Principles of Computer Vision." In: Young, T., (ed.) *Handbook of Pattern Recognition na Image Processing: Computer Vision*, Academic Press, New York.
- Arun, K. S., Huang, T. S. e Blostein, S. D. (1987). "Least-Squares Fitting of Two 3-D Point Sets." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 5, pp. 698-700.
- Banta, J.E., Wong, L.M., Dumont, C. e Abidi, M.A. (2000). "A Next-Best-View System for Autonomous 3-D Object Reconstruction." In: *IEEE Transactions Systems, Man, and Cybernetics*, Vol. 30, No. 5, pp. 589-597.
- Batchelor, B. G. e Whelan, P. F. (1997). "Intelligent Vision Systems for Industry." Springer-Verlag, New York
- Besl, P. J. e McKay, N. D. (1992). "A Method for Registration of 3-D Shapes." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239-256.
- Bhanu, B. (1984). "Representation and shape matching of 3-D objects." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 6, No. 3, pp. 340-351.
- Bischoff, R. e Graefe, V. (1998). "Vision-Guided Intelligent Robots for Automating Manufacturing, Materials Handling and Services." In: *Workshop on European Scientific and Industrial Collaboration on Promoting Advanced Technologies in Manufacturing*, Girona, pp. 105-109.
- Blais, G. e Levine, M. D. (1995). "Registering multiview range data to create 3D computer objects." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, pp. 820-824.
- Borenstein, J., Everett, H. R. e Feng, L. (1996). "Where am I ? Systems and Methods for Mobile Robot Positioning", Technical Report, University of Michigan
- Borges, D. L., Bispo, E. M. e Fisher, R. B. (1994). "Técnicas de aquisição automática de Modelos Geométricos para Reconhecimento e Inspeção." In: *Revista Brasileira de Computação (RBC)*, Vol. 7, No 2, pp. 49-60.

- Brown, L. G. (1992). "A survey of image registration techniques." In: *ACM Computing Surveys*, Vol. 24, No. 4, pp. 325-376.
- Campbell, R. e Flynn, P. (1998). "A WWW-Accessible 3D Image and Model Database for Computer Vision Research." In: Bowyer, K.W. e Phillips, P.J. (eds.) *Empirical Evaluation Methods in Computer Vision*. IEEE Computer Society Press, pp. 148-154.
- Cerrada, C., Ikeuchi, K., Weiss, L., e Reddy, R. (1990). "A 3D-Object Reconstruction System Integrating Range-Image Processing and Rapid Prototyping", Technical Report CMU-RI-TR-90-32, Robotics Institute, Carnegie Mellon University.
- Chen, C., Hung, Y e Cheng, J. (1999). "RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping rangeimages." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 11, pp. 1229-34.
- Chen, Y. e Medioni, G. (1992). "Object modelling by registration of multiple range images." In: *Image and Vision Computing*, Vol. 10, N0. 3, pp. 145-155.
- Chen, Y. e Medioni, G. (1995). "Description of Complex Objects from Multiple Range Images Using an Inflating Balloon Model", *CVIU(61)*, No. 3, May 1995, pp. 325-334.
- Coelho, M C. P. e Tavares, J. (2003). "Técnicas Base para Aquisição de Informação Tridimensional sem Contacto: Uma Descrição." In: *RESI - Revista Eletrônica de Sistemas de Informação*, Vol. 2, No. 1.
- Connolly, C. (1985). "The Determination of Next Best Views." In: Proc. *IEEE Int'l Conf. Robotics and Automation*, pp. 432-435.
- Conrad. D.J. e Sampson, R.E. (1990). "3D Range Imaging Sensors." In: Henderson, T.C., (ed.) *Traditional and Non-Traditional Robotic Sensors*. NATO ASI Series, Springer-Verlag, Vol. 63, pp. 35-47.
- Curless, B. L. (1997). *New Methods for Surface Reconstruction from Range Images*, PhD Thesis, Stanford University.
- Curless, B. L. (2000). *Overview of Active Vision Techniques*, SIGGRAPH 2000 Course on 3D Photography, University of Washington.
- D'eharbe, D. (2005). Elementos de complexidade computacional DIMAp/UFRN.
- Dorai, C, Wang, G., Jain, K. J. e Mercer, C. (1998). "Registration and Integration of Multiple Object Views for 3D Model Construction." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No 1.
- Dorai, C., Weng, J. e Anil, K. J. (1997). "Optimal Registration of Object views using range data." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 10, pp. 1131-1138.

- Dorai, C., Weng, J. e Jain, A., K. (1994). "Optimal registration of multiple range views." In: *12th Int. Conference on Pattern Recognition*, pp. A569-571, Jerusalem, Israel.
- Eggert, D. W., Fitzgibbon, A. W. e Fisher, R. B. (1998). "Simultaneous registration of multiple range views for use in reverse engineering of CAD models." In: *Computer Vision and Image Understanding*, Vol. 69, pp. 253-272.
- Eggert, D. W., Lorusso, A., e Fisher, R. B. (1997). "Estimating 3-D Rigid Body Transformations: A Comparison of Four Major Algorithms." In: *Machine Vision and Applications*, Vol.9, pp. 272-290.
- Elsner, D. L. Whitaker, R.T. e Abidi, M. A. (1999). "Volumetric Reconstruction of Objects and Scenes using Range Images." *Digital Signal Processing*, Vol. 9, No. 2, pp. 120-135.
- Faugeras, O. (1993). "*Three-dimensional Computer Vision*" MIT Press, Cambridge, MA.
- Faugeras, O. D. e Herbert, M. (1983). "A 3-D Recognition and Positioning Algorithm Using Geometrical Matching Between Primitive Surfaces." In: *Proceeding of Seventh International Joint Conference on Artificial Intelligence*, pp. 996-1002.
- Faugeras, O. D. e Herbert, M. (1986). "The representation, recognition and locating of 3-D objects." In: *International Journal of Robotic Research*, Vol. 5, pp. 21-52.
- Faugeras, O. e Luong, Q. T. (2004). *Geometry of Multiple Images* MIT Press, 644 pp.
- Favaro, P. (2000). "Shape from Focus/Defocus". Washington University.
- Florczyk, S. (2005). "*Robot Vision: Video-based Indoor Exploration with Autonomous and Mobile Robots*" Weinheim: Wiley-VCH.
- Forsyth, D.A. e Ponce, J. (2003). *Computer Vision: A Modern Approach*, Prentice Hall.
- Gibson, S. Fyock, C. Grimson, E., Kanade, T. Kikinis, R. Lauer, H. McKenzie, N. Mor, A. Nakajima, S. Ohkami, H. Osborne, R. Samosky, J. e Sawada, A. (1998). "Volumetric Object Modeling for Surgical Simulation," *Medical Image Analysis*, Vol. 2, No. 2, , pp. 121-132.
- Gonzalez, R. C. e Woods, R. E. (2000). *Processamento de Imagens Digitais*, Editora Edgard Blücher Ltda.
- Goshtasby, A. (1998). "Three-Dimensional Model Construction from Multiview Range Images: Survey With New Results." In: *Pattern Recognition*, Vol. 31, No. 11, pp. 705-1714.
- Guda, P., Cao, J., Gailey, J. H. e Hall, E. L. (2000). "*Machine Vision Fundamentals*." In: *Handbook of Industrial Automation*, Marcel Dekker, New York.

- Hartley, R. e Zisserman, A. (2000). “*Multiple View Geometry in Computer Vision*” Cambridge.
- Horn, B. K. P. (1987). “Closed-Form Solution of Absolute Orientation Using Unit Quaternions.” In: *Journal of the Optical Society of America A*, Vol. 4, No. 4, pp. 629-642.
- Horn, B., Hilden, H., e Negahdaripour, S. (1988). “Closed-Form Solution of Absolute Orientation Using Orthonormal Matrices,” *JOSAA*, Vol. 5, No. 7.
- Horowitz, E., Sahni, S., e Rajasekaran, S. (1998). “*Computer Algorithms*”. W.H. Freeman Press.
- Huber, D. (2002). *Automatic Three-dimensional Modeling from Reality*, Doctoral Dissertation, Carnegie Mellon University.
- ISR/Coimbra, (2001). “Alinhamento de Mapas 3D , ICP - Iterative Closest Point Algorithm, Baseado em Trabalhos de P. Besl, N. McKay, Szymon Rusinkiewicz, Marc Levoy, Zhengyou Zhang, Y. Chen, Medioni”, Disponível online - <http://paloma.isr.uc.pt/Tele3DWeb/downloads/3dlaa1.ppt>.
- Jähne, B. (2002). “*Digital Image Processing*”, 5th revised and extended edition, Springer.
- Jain, R., Kasturi, R. e Schunck, B. G. (1995 ). “*Machine Vision*”. McGraw Hill.
- Kanatani. K. (1994). “Analysis of 3D rotation fitting.” In: *IEEE Trans. Pattern Analysis and Machine Intell.*, 16(5):543-549.
- Kreyszig, E. (1998). *Advanced Engineering Mathematics*, 8th Edition, John Wiley and Sons, New York.
- Levoy, M. (2003). “The Digital Michelangelo Project”. Disponível online - <http://graphics.stanford.edu/projects/mich/>.
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., e Fulk, D. (2000). “The Digital Michelangelo Project: 3D Scanning of Large Statues.” In: *Proceedings of ACM SIGGRAPH 2000*, pp. 131-144.
- Li., S. Z. (1999). “Shape Matching Based on Invariants.” In: Omid Omidvar (ed.), *Shape Analysis*, pp.203-228, Volume 6 of Progress in Neural Networks. Ablex.
- Medioni, G. e Kang, S. B. (2004). *Emerging Topics in Computer Vision*, Prentice Hall.
- Meer, P. (2004). “Robust techniques for computer vision.” In: Medioni, G. e Kang, S. B. (Eds.) *Emerging Topics in Computer Vision*, Prentice Hall, pp. 107-190.
- Nishino, K. e Ikeuchi, K. (2002). “Robust Simultaneous registration of Multiple Range Images”, In: *The 5th Asian Conference on Computer Vision (ACCV2002)*, pp. 23-25.

- Page, D. L. Koschan, A. Sun, Y. e Abidi, M. (2003). "Laser-based imaging for reverse engineering." *Sensor Review, Special Issue on Machine Vision and Laser Scanners*, Vol. 23, No. 3, pp. 223-229.
- Paiva, A. C., Seixas, R.B. e Gattass, M. (1999). "Introdução a Visualização Volumétrica", Relatório Técnico. Departamento de Informática, PUC-Rio, Rio de Janeiro.
- Persiano, R.C.M. e Oliveira, A.A.F. (1989). *Introdução à Computação Gráfica: Livros Técnicos e Científicos Editora Ltda.*
- Piegl, L. e Tiller, W. (1997). "The NURBS Book", Springer-Verlag.
- Pito, R. e Bajcsy, R. (1995). "A Solution to the Next Best View Problem for Automated CAD Model Acquisition of Free-form Objects using Range Cameras." In: *Proceedings SPIE Symposium on Intelligent Systems and Advanced Manufacturing*, Philadelphia, PA.
- Potmesil, M. (1983). "Generating Models of Solid Objects by matching 3D surface segments. In: *VIII International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, pp. 1089-1093.
- Press, W. H., Teukolsky, S. A., Vetterling, W.T. e Flannery, B. P. (1992). "Numerical Recipes in C: The Art of Scientific Computing", Second Edition, Cambridge University Press, New York, NY.
- Pulli, K. (1997). *Surface Reconstruction and Display from Range and Color Data*, PhD Thesis, University of Washington.
- Pulli, K. (1999). "Multiview registration for large data sets." In: *Proceedings of Second International Conference on 3D Digital Imaging and Modeling (3DIM)*, pp. 160-168.
- Reed, M. K. (1998). *Solid Model Acquisition from range imagery*, PhD Thesis, Columbia University.
- Requicha, A. A. G. (1999 ). *Geometric Modeling: A First Course*, University of Southern California, (available on-line).
- Rosenhahn, B. (2003). *Pose Estimation Revisited*, PhD thesis, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik.
- Rusinkiewicz, S. e Levoy, M. (2001). "Efficient Variants of the ICP Algorithm." In: *Third International Conference on 3D Digital Imaging and Modeling*, pp. 145-152.
- Schut, G. H. (1960). "On exact linear equations for the computation of rotational element of absolute orientation" In: *Photogrammetria*, 15(1):34:37.

- Schwarte, R. (1990). "Multiresolutional laser radar." In: Henderson, T.C., (ed.) *Traditional and Non-Traditional Robotic Sensors*, NATO ASI Series, Springer-Verlag, Vol. F63, pp. 127-142.
- Shah, M. (1997). "Fundamentals of Computer Vision." Computer Science Department, University of Central Florida, Orlando, FL. Disponível em: <http://www.cs.ucf.edu/courses/cap6411/book.pdf>.
- Shum, H., Hebert, M., Ikeuchi, K. e Reddy, R. (1997). "An integral approach to free-form object modeling." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 12, pp. 1366 - 1370.
- Silva, L. e Bellon, O. R. P. (1995). "Robust range image registration using genetic algorithms and the surface interpenetration measure." Series in machine perception and artificial intelligence, Vol. 60.
- Simon, D. (1996). *Fast and Accurate Shape-Based Registration*, Ph.D Dissertation, Carnegie Mellon University.
- Simon, D. (1997). "What is "Registration" and Why is it so Important in CAOS?" In : *Proceedings of the First Joint CVRMed / MRCAS Conference*, pp. 57-60.
- Sinha, S. S. e Jain, R. (1994). "Range Image Analysis." In: Young, T., (ed.) *Handbook of Pattern Recognition na Image Processing: Computer Vision*, Academic Press, New York.
- Souza, M.A., Ricetti, F., Centeno, T. M., Pedrini, H., Erthal, J.L. e Mehl, A. (2001). "Reconstrução de Imagens Tomográficas Aplicada à Fabricação de Próteses por Prototipagem Rápida usando Técnicas de Triangulação." In: *Proceedings of II Latin American Congress on Biomedical Engineering*, Havana, Cuba.
- Spitz, E., N, (1999). *Dimensional Inspection Planning for Coordinate Measuring Machines*, PhD Thesis, University of Shouthern California.
- Spitz, S. N. e Requicha, A. A. G. (2000). "Multiple-Goals Path Planning for Coordinate Measuring Machines." In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2000)*, San Francisco, California, pp. 2322-2327.
- Stockeryale (2006). What is Structured Light ? - [http://www.stockeryale.com/i/lasers/structured\\_light.htm](http://www.stockeryale.com/i/lasers/structured_light.htm).
- Stockmann, M. e Naumann, J. (2005). *Moiré interferometry - Technique and Application*. Hungary, Anyagvizsgálók Lapja.
- Stoddart, A. J. e Hilton. A. (1996). "Registration of multiple point sets." In: *13th Int. Conference on Pattern Recognition*, pp. B40-44, Vienna, Austria.

- Thompson, E. H. (1958). An exact linear orientation of the problem of absolute orientation. *Photogrammetria*, 15(4):163:178.
- Tienstra, M.. (1969). Calculation of Orthogonal Matrices. ITC, Delft.
- Toscani, L. V. e Veloso, P. A. S. (2001). “Complexidade de algoritmos”, Editora Sagra Luzzatto.
- Trucco, E. e Verri, A. (1998). “Introductory Techniques for 3-D Computer Vision”, Prentice Hall.
- Turk, G. e Levoy, M. (1994). “Zippered Polygon Meshes from Range Images.” In: *Proceedings of ACM Computer Graphics SIGGRAPH94*, Orlando, Florida, USA, pp. 311-318.
- Umeyama, S. (1991). “Least-Square Estimation of transformation parameters between two point patterns.” In: *IEEE transactions on pattern analysis and machine intelligence*, Vol 13, No 4, pp 376-380.
- Varady, T., Martin, R. R. e Cox, J. (1997). “Reverse Engineering Of Geometric Models - An Introduction” In: *Computer-Aided Design*, Vol. 29, No 4, pp. :255-268.
- Viana, V. (1998). “Meta-Heurísticas e Programação Paralela em Otimização Combinatória.”, UFC Edições, Fortaleza, CE.
- Vuytsteke, P., Price, C.B. e Oosterlinck, A. (1990). “Image Sensors for Real-Time 3D Acquisition, Part 1.” In: Henderson, T.C., (ed.) *Traditional and Non-Traditional Robotic Sensors*, NATO ASI Series, Springer-Verlag, Vol. F63, pp. 187-210.
- Weisstein, E. W. (2006). “Covariance Matrix.” From MathWorld--A Wolfram Web Resource. Disponível online - <http://mathworld.wolfram.com/CovarianceMatrix.html>
- Weschler, M. (2006). “Como funciona o laser”. Disponível online - <http://ciencia.hsw.uol.com.br/laser.htm>
- Wikipédia (2006). “Modulação em frequência”. Disponível online - [http://pt.wikipedia.org/wiki/Modula%C3%A7%C3%A3o\\_em\\_freq%C3%BC%C3%A](http://pt.wikipedia.org/wiki/Modula%C3%A7%C3%A3o_em_freq%C3%BC%C3%A) Ancia.
- Williams, J. e Bennamoun., M. (2001). “Simultaneous Registration of Multiple Corresponding Point Sets.” In: *Computer Vision and Image Understanding*, Vol. 81, No. 1, pp. 117-142.
- Wong, L. M., Dumont, C. e Abidi, M. A. (1999). “Next best view system in a 3-d object modeling task.” In: *Proc. International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, pp. 306-311.

- Zhang, R., Tsai, P. -S., Cryer, J. E. e Shah, M. (1999). "Shape from Shading: A Survey." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21 No. 8, pp. 690-706.
- Zhang, Z. (1992). Iterative Point Matching for Registration of Free-Form Curves, Technical Report RR-1658, INRIA.
- Zhang, Z. (1995). Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting, Technical Report, INRIA.

## **APÊNDICES**

## APÊNDICE A – DIAGRAMA DA IMPLEMENTAÇÃO

