



UNIVERSIDADE DE COIMBRA
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Departamento de Engenharia Electrotécnica e Computadores
Mestrado em Engenharia Electrotécnica e Computadores
2001-2002

TESE DE MESTRADO

TELE-3D

Um Scanner para Registo Tridimensional de Objectos

João Filipe Ferreira

Instituto de Sistemas e Robótica
Coimbra, Portugal

9 de Novembro de 2004



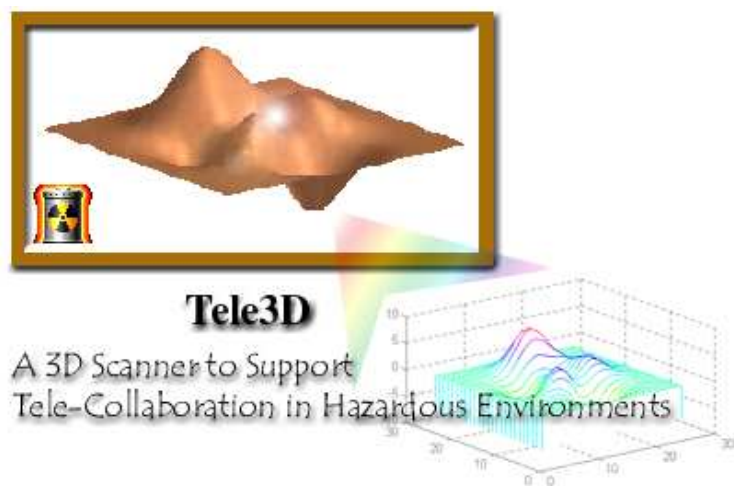
UNIVERSIDADE DE COIMBRA
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Departamento de Engenharia Electrotécnica e Computadores

Mestrado em Engenharia Electrotécnica e Computadores
2001-2002

TESE DE MESTRADO

TELE-3D

Um Scanner para Registo Tridimensional de Objectos



João Filipe Ferreira

Instituto de Sistemas e Robótica
Coimbra, Portugal

9 de Novembro de 2004

Agradecimentos

Queria começar por agradecer ao Professor Doutor Jorge Dias pela sua supervisão, pelo seu encorajamento nas horas difíceis que se nos apresentaram ao longo deste trabalho e pela sua paciência nos momentos em que a minha teimosia inata atrapalharam o seu decurso. Queria também agradecer ao Instituto de Sistemas e Robótica, Pólo de Coimbra, pelas instalações facilitadas e pelo apoio logístico, à Fundação para a Ciência e Tecnologia pelo financiamento do projecto na sua fase inicial e ao Departamento de Engenharia Electrotécnica e Computadores da Universidade de Coimbra pelo suporte vital ao meu trabalho.

Gostaria também de agradecer a todos os meus colegas e amigos de curso, de laboratório e de departamento, cujo apoio, tanto moral como científico, foram inestimáveis: ao Jorge Lobo, ao Nuno Cid, ao Luís Almeida, ao João Alves, ao Paulino, ao Jörg Rett, ao João Barreto, ao Gabriel, ao Peixoto, um enorme e comovido abraço!

A minha gratidão estende-se, obviamente, também a todos os meus amigos do coração pelo carinho que me deram. Gostaria de destacar particularmente a importância que teve o acompanhamento que tive de duas pessoas que, pelo amor e compreensão com que sempre me presentearam (talvez por estarem numa situação similar à minha), merece ser reconhecido em especial — à Cláudia e à Belinha, um “obrigado” impossível de descrever simplesmente por palavras.

Também sempre presente e importantíssima foi a minha família, à qual agradeço do fundo do coração. Particularmente, gostaria exprimir a minha gratidão pelo carinho incondicional da minha mãe, do meu pai, da minha irmã e dos meus avós, que tão ansiosa mas silenciosamente esperaram por este momento.

Por fim, queria agradecer em especial à Anita, minha mulher, pelo amor, carinho, enfim!, todo o apoio indescritível que me ofereceu e que me fez andar em frente em momentos dolorosos e entusiasmar-me nas ocasiões mais positivas. O meu sucesso não seria possível sem ti!

Resumo

A digitalização tridimensional tem operado uma revolução, às vezes bastante visível, mais vezes ainda silenciosa, nos últimos anos na nossa vida quotidiana. A necessidade imperiosa de adicionar uma dimensão à visualização do mundo que nos rodeia tem tornado essencial o papel das tecnologias que suportam esta revolução, nomeadamente a mais omnipresente visão por computador. Resumindo, a inspecção de objectos e cenas tridimensionais, seja na ciência, na indústria, ou qualquer outro campo de aplicação, deixou de se basear na detecção de apenas algumas das suas características particulares e passou a significar a recuperação completa da geometria visível desses objectos e dessas cenas.

Tendo em conta a importância desta técnica e à necessidade de a tornar cada vez mais eficiente, barata e fácil de utilizar, pretendeu-se neste trabalho efectuar-se o desenvolvimento e estudo extensivo de um sistema de digitalização tridimensional portátil movido “à mão”, baseado em visão por computador, projecção de luz estruturada e utilizando sensores de posicionamento e orientação.

Primeiramente, foi feito um estudo do “estado da arte”, a partir do qual foi delineado o rumo a seguir. Seguidamente, definiram-se os modelos do sistema e metodologias para efectuar a sua preparação para uso (ou seja, os procedimentos para a sua calibração), tendo sido lançadas as fundações teóricas (matemáticas e geométricas) de todo o processo, algumas das quais perspectivando soluções inovadoras, tanto de engenharia como a nível conceptual, para alguns problemas recorrentes em sistemas deste género.

Foi dedicada particular atenção à resolução do problema do registo tridimensional, visto ser a questão de mais difícil resolução em sistemas portáteis, arbitrariamente móveis, deste género; foi desenvolvida uma solução muito interessante que toma partido da geometria inerente ao método de triangulação utilizado.

Dissertação submetida ao Departamento de Engenharia Electrotécnica e de
Computadores da Universidade de Coimbra satisfazendo parcialmente os requisitos
para a obtenção do grau de Mestre

Trabalho desenvolvido sob orientação do Doutor Jorge Manuel Miranda Dias —
professor Associado do Departamento de Engenharia Electrotécnica e de
Computadores da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Esta tese é dedicada à parte mais importante do meu ser, a Anita.

SIMULTANEAMENTE O PROBLEMA E A UTILIDADE DA RECUPERAÇÃO
DE INFORMAÇÃO TRIDIMENSIONAL DO MUNDO:

A Natureza não revela os seus mistérios de uma só vez.

Sêneca, *Questões Naturais*, livro 7, século I

SERÁ QUE A SABEDORIA POPULAR ANTIGA JÁ SABIA FAZER A
COMPARAÇÃO ENTRE A DESCRIÇÃO VERBAL, A DESCRIÇÃO BIDIMEN-
SIONAL E A REPRESENTAÇÃO TRIDIMENSIONAL DO MUNDO?

Diz-me e eu esquecerei;

Mostra-me e eu lembrar-me-ei;

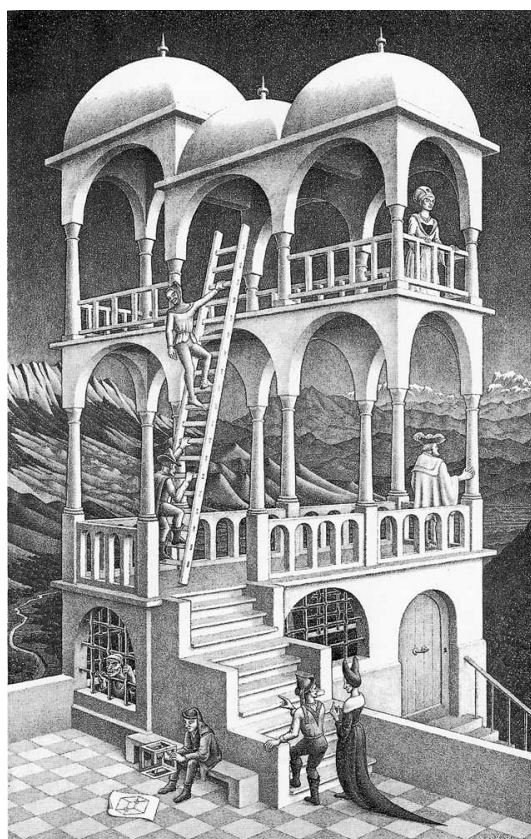
Envolve-me e eu compreenderei.

Antigo provérbio chinês

FALHAS NA PERCEPÇÃO TRIDIMENSIONAL A PARTIR DE REPRESENTAÇÕES BIDIMENSIONAIS DO MUNDO.

COMO PODEREMOS NÓS SER DIFERENTES DO RAPAZ DA FOLHA DE PAPEL?

Abaixo: figura tirada de *M. C. Escher, Gravura e Desenhos*, Taschen, GmbH, 2002.
Comentário do próprio Maurits Cornelis Escher, do mesmo livro.



M.C. Escher's 'Belvedere' (c) 2004 The M.C. Escher Company - The Netherlands. All rights reserved.
Used by permission.

Belvedere, litografia, 1958, 46 × 29,5 cm. *Em primeiro plano, em baixo à esquerda, está uma folha de papel, sobre a qual foram desenhadas as linhas de um dado. Dois círculos indicam os pontos onde as linhas se cruzam. Que linha está à frente, que linha está atrás? Atrás e à frente, ao mesmo tempo, não é possível num mundo tridimensional e não pode por isso ser representado. Mas pode ser desenhado um objecto que, visto de cima, representa uma realidade diferente da de quando visto de baixo. O rapaz, que está sentado no banco, tem nas mãos uma tal absurdidade, em forma de cubo. Ele observa pensativamente o objecto impossível e não parece ter consciência de que o belvedere, atrás das costas dele, é construído desta forma impossível. No piso inferior, no interior da casa, está encostada uma escada pela qual sobem duas pessoas. Mas chegados a um piso acima, estão de novo ao ar livre e têm de voltar a entrar no edifício. É então estranho que ninguém desta comunidade se preocupe com o destino do preso no subterrâneo que, queixoso, põe de fora a cabeça, através das grades?*

Conteúdo

Agradecimentos	iii
Resumo	v
1 Introdução	1
1.1 Motivações	1
1.2 Técnicas de Recuperação de Informação Tridimensional	10
1.2.1 A percepção da tridimensionalidade na Natureza	12
1.2.2 Projecção de luz estruturada	14
1.2.3 Sistemas multi-oculares	19
1.2.4 Outras técnicas	21
1.3 Técnicas para Recuperações Tridimensionais Completas	23
1.3.1 Sistemas fixos com partes móveis	23
1.3.2 Sistemas portáteis	24
1.3.3 Integração de informação tridimensional	27
1.4 Técnicas de Reconstrução de Cenas Tridimensionais	29
1.4.1 Computação gráfica 3D	29
1.4.2 Outras técnicas	31
1.5 Exemplos de Sistemas de Digitalização Tridimensional	33
1.5.1 Sistemas fixos	33
1.5.2 Sistemas portáteis	34
1.5.3 Comparações entre sistemas	39
1.6 Discussão: porque é que o <i>scanning</i> 3D é difícil?	42
1.7 Objectivos e estruturação do projecto Tele-3D	44
1.8 Estruturação da dissertação	45
2 O Sistema Tele-3D — Modelos e Calibração	47
2.1 Princípio de funcionamento do scanner Tele-3D	48
2.1.1 Modelo do sistema de triangulação tridimensional	49

2.1.2	Modelo do sistema de medição de “pose”	56
2.2	Estudo de sensibilidade ao erro do sistema de triangulação tridimensional	59
2.2.1	Conceitos de sensibilidade ao erro e sua aplicação genérica ao modelo do sistema de triangulação tridimensional	59
2.2.2	Estudo teórico da sua aplicação a uma configuração particular do sistema de triangulação	64
2.2.3	Simulação e análise de sensibilidade ao erro de uma instanciação aproximada do <i>scanner</i> Tele-3D	69
2.2.4	Análise dos resultados do estudo da sensibilidade ao erro do modelo de triangulação	75
2.3	Calibração do scanner Tele-3D	77
2.3.1	Calibração da câmara	77
2.3.2	Calibração do projector laser	79
2.3.3	Calibração do sistema de medição de “pose”	81
2.4	Resultados da simulação sobre a calibração do sistema de medição de “pose”	91
2.4.1	Inserção de ruído	92
2.4.2	Métrica para o erro	93
2.4.3	Primeiro conjunto de simulações — resultados para sensibilidade ao ruído em rotação e translação	93
2.4.4	Segundo conjunto de simulações — resultados para sensibilidade ao ruído com número variável de varrimentos/movimentos simulados	94
2.4.5	Terceiro conjunto de simulações — resultados para sensibilidade ao ruído com rotações de ângulos com amplitude variável e translações com comprimento variável	97
2.5	Calibração do protótipo laboratorial	103
2.5.1	Descrição dos procedimentos de calibração	103
2.5.2	Resultados da calibração	107
2.6	Conclusões	110
3	Reconstrução Tridimensional — Registo e Integração de Perfis	111
3.1	Varrimento e processamento de uma cena tridimensional	111
3.2	Registo e integração tridimensional de perfis de luz	113
3.2.1	Uma visão global sobre registo de perfis — correspondência entre perfis e correspondência entre pontos	115
3.2.2	Optimização local	116

3.2.3	Algoritmo de registo 3D de perfis	121
3.3	Reconstrução tridimensional — resultados	124
3.4	Conclusões	130
4	Conclusões e Trabalho Futuro	131
4.1	Apreciação dos resultados	131
4.2	Trabalho futuro	131
4.2.1	Melhoria das condições de funcionamento do sistema	132
4.2.2	Aposta em sensores inerciais	132
4.2.3	Utilização de visão estereoscópica activa	132
4.2.4	Refinação do processo de reconstrução tridimensional	132
A	Notação	137
B	Engenharia do Sistema Tele-3D	138
B.1	Descrição do hardware	138
B.2	Descrição do software	141
C	Quaterniões Duais e Transformações Rígidas	143
C.1	Rotações e Quaterniões Unitários	144
C.2	O Conceito de Número Dual	151
C.3	Movimento Helicoidal e Quaterniões Unitários Duais	153
D	Estimação Não Linear de Mínimos Quadrados	162
E	Estimação Linear de Mínimos Quadrados	164
E.1	Solução de sistemas lineares não homogéneos ($\mathbf{Ax} = \mathbf{b}$)	165
E.2	Solução de sistemas lineares homogéneos ($\mathbf{Ax} = \mathbf{0}$)	167

Lista de Tabelas

1.1	Exemplos de sistemas de sensores de <i>pose</i> (características fornecidas pelos fabricantes)	26
1.2	Exemplos de <i>scanners</i> 3D (características disponibilizadas pelos fabricantes)	41
A.1	Notação usada e descrições correspondentes	137

Lista de Figuras

1.1	Exemplo de uso de sistemas de recuperação tridimensional e sua integração com outros sistemas de aquisição de dados no campo da medicina	4
1.2	Exemplo de uso de recuperação e reconstrução tridimensionais na construção de moldes dentários	5
1.3	Reconstrução facial em ciências forenses	6
1.4	Imagens de uma digitalização 3D de uma estatueta de um buda, resultado do projecto para construir uma máquina de fax 3D	7
1.5	Imagens de modelos tridimensionais da estátua de David. Retirado do “ <i>Digital Michelangelo Project</i> ”	8
1.6	Exemplos de pós-produção cinematográfica recorrendo ao uso de reconstrução tridimensional	10
1.7	Cenário virtual encomendado pela BBC, criado a partir de uma recuperação estereoscópica	11
1.8	Espectro electromagnético com espectro visível, cores que compõem o espectro visível e respostas relativas de diversos dispositivos de emissão de luz comparativamente à resposta relativa do olho humano, em função do comprimento de onda	15
1.9	Exemplo de sombras de Moiré	16
1.10	Sistema de inspecção com projector de lista de luz, câmara e tapete rolante	18
1.11	Geometria epipolar de um sistema de estereovisão	20
1.12	Estéreo fotogramétrico	21
1.13	Exemplo de registo tridimensional	28
1.14	Exemplificação de um B-Spline	30
1.15	Exemplificação de uma reconstrução conseguida recorrendo ao algoritmo de Hugues Hoppe – <i>Surface Reconstruction from Unorganized Points</i> .	31
1.16	Imagem da reconstrução 3D por estereolitografia de uma estatueta de um buda	32
1.17	<i>Scanner</i> laser REPLICA 3D da 3D Scanners	35

1.18	<i>Scanner</i> laser 3030RGB/PS da Cyberware.	35
1.19	<i>Scanner</i> laser 3030RGB/MS da Cyberware	35
1.20	Cabine TorsoScan TriForm	36
1.21	O C3D-2030 (de C3D/Edinburgh Virtual Environment Centre)	36
1.22	O VIVID 700 (da Minolta Corporation)	37
1.23	OPTO 3D-Ranger (da Optonet)	38
1.24	O ModelMaker (da 3D Scanners)	38
1.25	O HLS/Polhemus FastSCAN (da ARANZ/Polhemus)	39
1.26	O Stradx (da Universidade de Cambridge)	40
1.27	1º problema de 3D <i>scanning</i> : materiais “não cooperantes” opticamente	43
1.28	2º e 6º problemas de 3D <i>scanning</i> : oclusões e lacunas	44
2.1	Esquema da arquitectura do scanner Tele-3D	49
2.2	Esquema para a geometria de uma cena com projecção da luz estruturada	51
2.3	Modelo da câmara “ <i>pinhole</i> ”	52
2.4	Projecção em perspectiva	53
2.5	Efeitos da amostragem de uma câmara CCD	54
2.6	Definição dos eixos/ângulos de rotação “Roll-Pitch-Yaw”	57
2.7	Elipsóide de incerteza: o caso tridimensional	63
2.8	Modelo simplificado do sistema Tele-3D	66
2.9	Função objectivo vs baseline para $\phi = 120^\circ$, $\gamma = 120^\circ$ e $h = 25$ cm . . .	71
2.10	Função objectivo vs baseline para $\phi = 150^\circ$, $\gamma = 60^\circ$ e $h = 25$ cm . . .	72
2.11	Função objectivo vs baseline para $\phi = 165^\circ$, $\gamma = 30^\circ$ e $h = 25$ cm . . .	73
2.12	Função objectivo vs posição vertical na imagem	74
2.13	Função objectivo vs ϕ e γ , para $B = 45$ cm e $h = 25$ cm	76
2.14	Esquema mostrando a configuração inerente simultaneamente à cali- bração da câmara e do laser	78
2.15	Calibração do sensor de “ <i>pose</i> ” — formulação estática	82
2.16	Esquematização para a formulação de movimento relativo para a cali- bração da transformação do referencial da câmara para o referencial do sensor de “ <i>pose</i> ”	84
2.17	Esquematização para a formulação de movimento relativo para a cali- bração da transformação do referencial do emissor do sensor de “ <i>pose</i> ” para o sistema de coordenadas absoluto (WCS)	85
2.18	Comparação da sensibilidade ao ruído presente apenas em rotações . .	95
2.19	Comparação da sensibilidade ao ruído presente apenas em translações .	96

2.20	Comparação da sensibilidade ao ruído para número variável de “ <i>poses</i> ”/movimentos simuladas/os	98
2.21	Comparação da sensibilidade ao ruído face a variações na amplitude dos ângulos de rotação	100
2.22	Comparação da sensibilidade ao ruído face a variações na norma das translações	101
2.23	Número de divergências usando a formulação de movimento relativo para os diversos conjuntos de simulações	102
2.24	Imagens exemplo da calibração da câmara através do software do OpenCV	104
2.25	Imagem tirada durante tentativa de calibração simultânea da câmara, do sensor de “ <i>pose</i> ” e do projector laser	105
2.26	Imagem da calibração do projector laser e do processamento de imagem envolvido	106
2.27	Transformada de Hough	107
3.1	Objectos digitalizados pelo sistema Tele-3D	112
3.2	Processamento e amostragem do perfil laser durante o varrimento de uma cena	112
3.3	Estimativas iniciais para a integração tridimensional	114
3.4	Registo de planos de luz	117
3.5	Exemplo prático do processo de registo de perfis 3D	123
3.6	Objectos com “clusters” registados	125
3.7	Exemplo de reconstrução tridimensional usando o software <i>Tele-3D Reconstructor</i>	126
3.8	Resultados da reconstrução tridimensional do manequim	127
3.9	Resultados da reconstrução tridimensional do manequim (vista de perfil)	128
3.10	Detalhe dos resultados da reconstrução tridimensional do manequim (face)	129
3.11	Resultados usando o software <i>Tele-3D Reconstructor</i> da reconstrução tridimensional da estátua de Nossa Senhora	129
B.1	“Deployment Diagram” (UML) para o sistema Tele-3D	139
B.2	Projecto CAD e fotografia do <i>scanner</i> Tele-3D	140
B.3	“Use Case Diagram” (UML) para o sistema Tele-3D	142
B.4	“Sequence Diagram” (UML) para o software <i>Tele-3D Grabber</i>	142
C.1	Quaterniões unitários e rotações — a fórmula de Rodrigues	148
C.2	A geometria do movimento helicoidal	154

Capítulo 1

Introdução

O Homem já há muito tempo que mede e representa informação sobre realidades tridimensionais com as mais variadas intenções, desde o planeamento de projectos de engenharia até à arte, passando mesmo pela necessidade de satisfação do desejo mais primário de comunicar, presente desde a Pré-História. Mas mesmo antes de saber comunicar de uma forma mais abstracta, já o ser humano estaria de forma inata equipado com sensores que lhe permitiriam analisar a tridimensionalidade do mundo que o rodeava. Tornou-se óbvio com o tempo, no entanto, que seria útil ter formas mais simples, automáticas e precisas de efectuar esta análise e a representação dos seus resultados.

Este capítulo apresenta um estudo de vários sistemas de recuperação e reconstrução de informação tridimensional, comumente conhecidos como *scanners* ou *digitalizadores 3D*¹, e as suas tecnologias, desde sistemas tradicionais, como os *sistemas multi-câmara* (também designados por *multi-oculares*) ou os *sistemas de varrimento com laser*, até a sistemas de vanguarda, como os *scanners portáteis e handheld* ².

O texto que se segue apresenta um resumo do estudo acima referido, que se encontra apresentado em maior detalhe no relatório técnico Ferreira [2002].

1.1 Motivações

A problemática da recuperação e reconstrução espacial tridimensional tem sido objecto de discussão durante as duas décadas passadas, para pessoas envolvidas em áreas tão diversas como, de um modo óbvio, as ciências informáticas, físicas, médicas,

¹Apesar da última designação ser formalmente mais correcta, a primeira é mais comum e, por esta razão, utilizar-se-á esta maioritariamente ao longo do resto do texto.

²de varrimento manual

ou como as áreas talvez menos óbvias respeitantes à produção cinematográfica e televisiva, estendendo-se mesmo essa discussão a todas as indústrias modernas em geral. Chegámos, portanto, a uma era em que este tipo de tecnologia já não faz mais apenas parte do imaginário da ficção científica. Sistemas usando este tipo de solução podem mostrar-se altamente eficazes, baratos e de boa qualidade para inspecção visual automática [Ferreira and Dias 2000].

As motivações mais óbvias para a realização de investigação nesta área são apresentadas de seguida:

- As potencialidades de uso prático destes sistemas confirmam a ideia de que a passagem do presente uso de descrições de realidades tridimensionais através das suas projecções bidimensionais para o uso directo de informação tridimensional virá a ser uma das mais importantes revoluções tecnológicas do futuro próximo — o advento da realidade virtual e do *standard* de codificação de multimedia *MPEG-4* implicará o uso recorrente de digitalizadores tridimensionais [Haskell et al. 1997].
- O aparecimento de sistemas de baixo custo devido à utilização racional de tecnologias mais baratas motivará um estudo ainda mais aprofundado para a sua optimização no sentido de se conseguirem obter a partir deles resultados ainda mais precisos [Ferreira et al. 2001, Ferreira and Dias 2000, Martins and Dias 1999].
- Dito isto, o espectro de tecnologias passíveis de ser utilizadas para versatilizar ainda mais estes sistemas está longe de estar esgotado.
- As técnicas de visualização e reconstrução tridimensional, incluindo a computação gráfica 3D, têm sofrido enormes evoluções, o que poderá galvanizar ainda mais esta área de investigação (ver Hoppe [1994], Nina Amenta and Kolluri [2001] e Farin [1999], M. Woo and Davis [1997], Edward [1990]).

Na verdade, várias aplicações práticas poderiam lucrar com a utilização deste tipo de sistemas [C3D/Edinburgh Virtual Environment Centre, Levoy et al. 2000, Weiss 1997, IMPROOFS project, 3D-MURALE Project, EPSRC 2000, Pérez-Fajado and Gutsens 2000]:

Engenharia e produção fabril (inspecção industrial, engenharia de metrologia dimensional, prototipagem rápida, “engenharia inversa”, etc.);

Arte (digitalização de artefactos de museu para arquivamento e reprodução em formato digital, incorporação em ferramentas artísticas digitais, animação por computador, pós-produção cinematográfica e vídeo, encenação virtual, etc.);

Simulação e realidade virtual (composição de descrições digitais para análise e simulação por computador, criação de ambientes virtuais simulando de forma realista o mundo físico, etc.);

“Faxing 3D”;

Modelos e Protótipos Industriais;

Factores humanos e antropometria (dimensionamento de roupa, *design* de veículos automóveis e de locais de trabalho, etc.);

Internet e telecomunicações (comércio virtual, etc.);

Medicina (examinação, diagnóstico, planeamento cirúrgico, simulação cirúrgica, projecto e fabrico de próteses, ortopedia, etc.);

Ciência forense (reconstrução facial, etc.).

É nas áreas onde é necessário ter descrições precisas do formato do corpo humano que esta tecnologia está a produzir resultados mais entusiasmantes. De seguida, ir-se-ão apresentar esses novos campos de aplicação com detalhe suficiente para dar uma ideia da revolução na vida do ser humano tornada possível através do seu uso.

Antropometria Os dados antropométricos são usados em muitas áreas da indústria para o design de produtos como calçado, equipamento de segurança, móveis, veículos e outros objectos com que o ser humano interage. Nas forças militarizadas e paramilitares, este tipo de dados é tão importante quanto a sobrevivência dos seus efectivos; como exemplo disso poder-se-á referir a necessidade de formatos precisos no equipamento justo de protecção, como capacetes, ou na sua importância na ergonomia de veículos de combate. Por outro lado, dados antropométricos precisos são também indispensáveis na gestão de gamas de tamanho e de stocks, tanto para roupa como para equipamento pessoal, e o crescente da Internet na venda a retalho on-line tem aberto horizontes novos no uso de dados antropométricos em aplicações de “venda virtual” de roupa.

A maioria dos dados antropométricos correntemente de uso generalizado deriva de informação adquirida por medição manual de uma amostra de população de

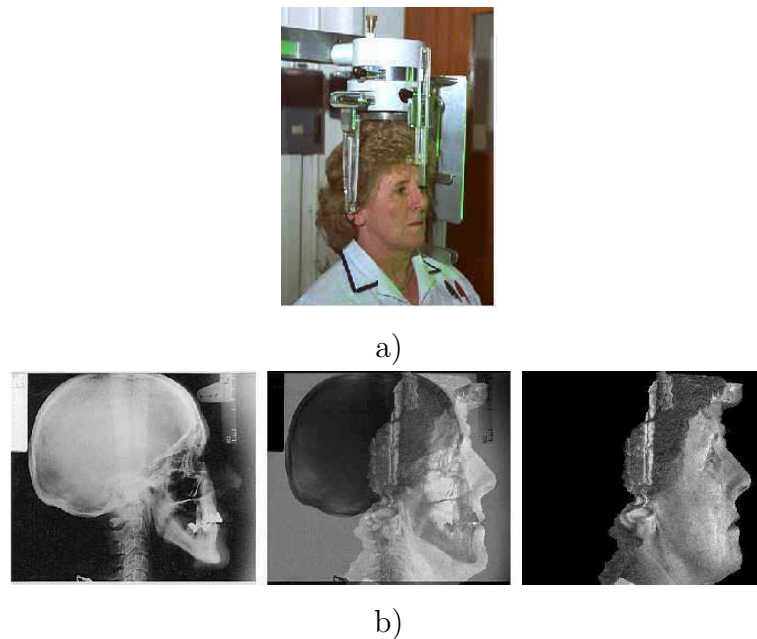


Figura 1.1: Exemplo de uso de sistemas de recuperação tridimensional e sua integração com outros sistemas de aquisição de dados no campo da medicina: a) a paciente e o sistema-exemplo; b) a integração de informação dos dois sistemas – tirado de EPRSC [2000], com a permissão gentil do Laboratório de Gráficos e Visão do Departamento de Ciências Computacionais da Universidade de Glasgow.

há cerca de 50 anos atrás. Melhorias nos cuidados com a alimentação e saúde humanas nos anos mais recentes resultaram no aumento do tamanho médio da população, tornando portanto obsoletos esses dados. Sondagens recentes conduzidas pelas Forças Armadas britânicas e norte-americanas entre outras foram um esforço consciente em direcção à resolução deste problema, mas saíram frustradas pela ausência de um método adequado para a aquisição rápida e eficaz de dados tridimensionais densos e precisos do corpo humano na sua totalidade. Técnicas manuais, que envolvem o uso, por exemplo, de fita métrica, são extremamente lentas e dispendiosas, produzem dados dispersos e são quase impossíveis de automatizar. Desde o advento das tecnologias de recuperação tridimensional, tem havido um crescente interesse na sua aplicação na medição do corpo humano e o objectivo neste momento é criar um sistema que possa fazer de um modo rotineiro essa recuperação de modo a obter dados densos e precisos, como requerido.

Medicina Um dos campos de maior disseminação desta tecnologia será a medicina, onde numerosos exemplos da sua utilização podem ser encontrados (examinação, diagnóstico, planeamento cirúrgico, simulação cirúrgica, projecto e fabrico de

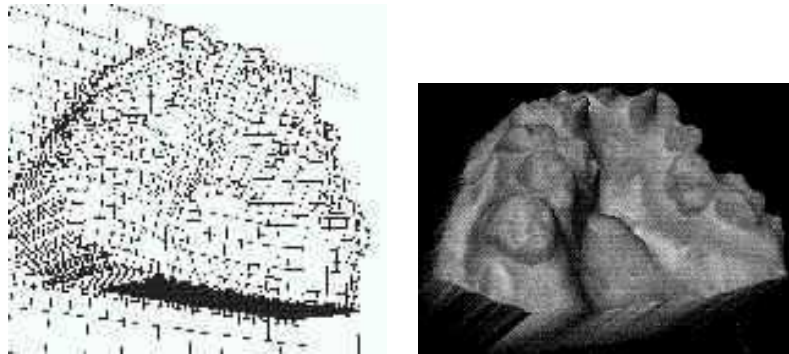


Figura 1.2: Exemplo de uso de recuperação e reconstrução tridimensionais na construção de moldes dentários — tirado de EPRSC [2000], com a permissão gentil do Laboratório de Gráficos e Visão do Departamento de Ciências Computacionais da Universidade de Glasgow.

próteses, ortopedia, etc.). Aliás, a integração de dados obtidos por sistemas de radiologia e similares com dados tridimensionais adquiridos pela recuperação tridimensional são exemplo dos avanços mais recentes deste campo; está previsto que a sua utilização venha a ser um passo incomensurável em frente na constituição de registos individuais dos pacientes nos serviços de saúde do futuro.

Nas figuras 1.1 e 1.2, apresentam-se imagens de alguns exemplos de casos de uso de sistemas de recuperação e reconstrução em diversas áreas e aplicações da medicina.

Ciência forense Outro campo de utilização desta tecnologia, talvez menos óbvio, será a ciência forense. Um dos exemplos do seu uso mais importantes é a reconstrução facial nos casos em que todos os outros métodos de identificação de cadáveres são impraticáveis — observe-se a figura 1.3 na página seguinte.

Realidade virtual nos *media*, Internet e telecomunicações Outras áreas que este tipo de tecnologia revolucionou foram os meios de comunicação social, a Internet (principalmente no comércio electrónico) e as telecomunicações, onde a realidade virtual, até aí apenas possível através da criatividade de alguns animadores, entrou numa era de realismo notável — se, até há bem pouco tempo, efeitos especiais deste tipo tinham um aspecto um pouco a lembrar bonecos animados, hoje em dia a sua imitação da realidade tornou-se quase perfeita.

Como exemplo de uso da digitalização tridimensional abrangendo a realidade virtual, a Internet, a arqueologia e a disponibilização de reproduções digitais 3D de artefactos de museu, uma equipa de 30 estudantes e professores da Universi-

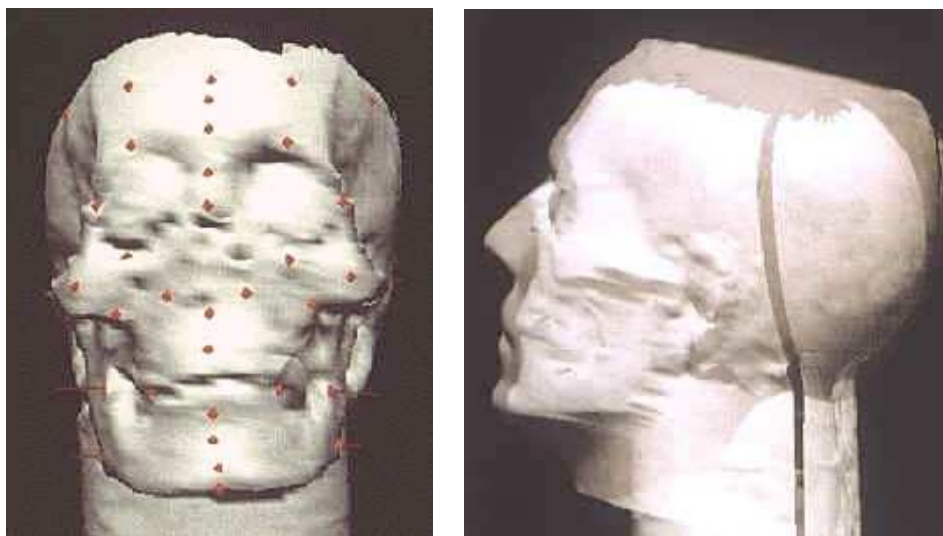


Figura 1.3: Reconstrução facial em ciências forenses – faz-se uma recuperação tridimensional do crânio (à esquerda), assinalando os pontos fulcrais na fisionomia, fazendo-se, seguidamente, o ajuste de um “rosto-tipo” a esses pontos (à direita) – tirado de EPRSC [2000], com a permissão gentil do Laboratório de Gráficos e Visão do Departamento de Ciências Computacionais da Universidade de Glasgow.

dade de Washington passaram o ano lectivo de 1998/99 em Itália recuperando as estruturas tridimensionais de esculturas e da arquitectura de Miguel Ângelo.

Desde 1992 que o líder desta equipa, o Professor Marc Levoy, e os seus estudantes têm estado a investigar métodos para digitalizar a forma de objectos tridimensionais usando scanners laser. Aos seus esforços iniciais foi dado optimisticamente o nome de “projecto para construir uma máquina de fax 3D”: em 1996 digitalizaram uma pequena estatueta de um buda e mandaram por correio electrónico o seu modelo estrutural de Palo Alto para uma unidade fabril em Los Angeles onde uma réplica precisa foi construída usando um processo denominado *stereolitografia* (ver figura 1.4 na próxima página), demonstrando, pelo menos, que tal sistema futurista era passível de ser funcional [Levoy et al. 2000].

O “*Digital Michelangelo Project*” começou oficialmente em Janeiro de 1997, implicando um período de planeamento de 2 anos. Durante esse período foi concebido, construído e testado equipamento constituído por um *laser rangefinder* montado numa peça mecânica móvel num suporte (ver figura 1.5 na página 8). Os requisitos deste sistema eram exigentes: era intenção da equipa do Professor Levoy recuperar marcas de buril menores que um milímetro, recuperação essa feita a uma distância de segurança, abarcando inclusivamente toda a altura da

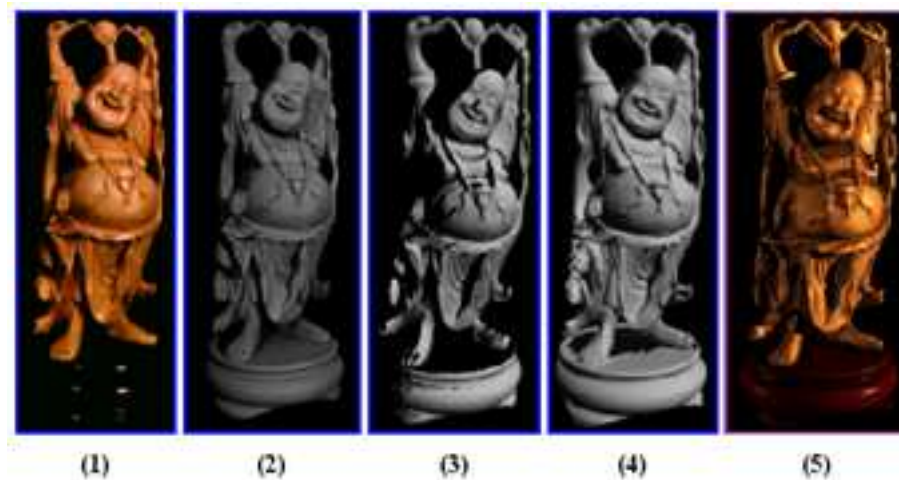


Figura 1.4: Imagens de uma digitalização 3D de uma estatueta de um buda, resultado do projecto para construir uma máquina de fax 3D – tirado de Curless and Levoy [1996-97] com a gentil permissão do Professor Marc Levoy. Da esquerda para a direita: estatueta original (de madeira e marfim falso), a mesma estatueta pintada de cinzento para simplificar a digitalização, recuperação 3D (parcial) feita a partir de um ponto de vista usando um scanner Cyberware 3030 (ver Cyberware) modificado para usar algoritmos de análise espaço-temporal, recuperação 3D completa feita a partir integração de recuperações de 48 pontos de vista, reconstrução 3D final usando algoritmos de preenchimento de lacunas.

estátua de David (10 metros!).

O sistema final, a que foi dado o nome de “*Scanner Stanford para Grandes Estátuas*”, foi fabricado pela empresa Cyberware (ver Cyberware). Para as zonas mais inacessíveis (as estátuas não foram movidas), foi utilizado um segundo *scanner*, um braço com juntas e um *laser rangefinder* por triangulação fabricado pela empresa 3D Scanners (ver 3D Scanners). Finalmente, para permitir a recuperação da arquitectura circundando as estátuas de Miguel Ângelo (principalmente a nova Sacristia das Capelas de Medicis, projectadas por ele) foi usado um protótipo de um *laser rangefinder* tipo *time-of-flight* feito por Cyra Technologies. Todos os três sistemas capturavam simultaneamente cor de alta-resolução [Levoy et al. 2000]. (Simultaneamente, por volta de 1996, um outro grupo no Concelho Nacional de Investigação Científica do Canadá — o NRC — estava a recuperar artefactos de museu a um ritmo regular. Porém, ninguém tinha conseguido resultados tão precisos e exaustivos quanto foram conseguidos no “*Digital Michelangelo Project*” [Levoy et al. 2000].)

Apesar dos modelos 3D deste projecto estarem ainda apenas parcialmente completos, devido a uma intensa procura foi disponibilizado um esqueleto de ar-



Figura 1.5: Imagens de modelos tridimensionais da estátua de David. Retirado do “*Digital Michelangelo Project*”, com a gentil permissão do Professor Levoy [Levoy et al. 2000].

quivo, tendo sido o seu catálogo colocado *online*. Apesar de as metodologias necessárias para criar e manter arquivos digitais de artefactos 2D terem amadurecido substancialmente nos últimos anos, o salto para a tridimensionalidade tem colocado vários novos problemas. Para os resolver, foi iniciado um novo projecto de investigação científica com o nome “Criando arquivos de obras de arte tridimensionais”, patrocinado pela National Science Foundation (Digital Libraries Initiative) [Levoy et al. 2000].

Um sub-projecto que se foi desenvolvendo paralelamente pela mesma equipa, o “*Digital Forma Urbis Romae Project*”, foi a digitalização de fragmentos da *Forma Urbis Romae*, um mapa gigante da Roma antiga esculpido em placas de mármore por volta do ano 200 A.D.. O mapa encontra-se actualmente fragmentado — em 1163 pedaços para se ser exacto. A reconstituição peça-a-peça deste mapa tem sido um dos grandes problemas por resolver do campo da arqueologia. Afortunadamente, os fragmentos têm vários centímetros de espessura, disponibilizando à equipa objectos de estrutura 3D facilmente recuperável, todos eles digitalizados numa maratona de 3 semanas a 24 horas por dia entre Maio e Junho de 1999. A equipa está neste momento a desenvolver algoritmos para tentar “reconstituir o puzzle” [Levoy et al. 2000].

As motivações por trás de qualquer destes projectos são o avanço da tecnologia de *scanning* 3D, a sua colocação ao serviço da humanidade e a criação de um arquivo a longo-prazo de artefactos culturais importantes, com o patrocínio da Universidade de Stanford, da Interval Research Corporation e da Fundação pelas Artes Paul G. Allen.

Outro exemplo de *scanning* 3D posto ao serviço da arqueologia é o *3D-MURALE*. Neste projecto, o objectivo da equipa supervisionada pelo Professor Marc Waelkens desde 1990 é a modelação do local arqueológico de Sagalassos, uma das três maiores cidades da antiga Pisidia, localizada *7km* a norte da aldeia de Aglasun na província turca de Burdur. Este projecto é suportado por um consórcio de universidades e empresas liderado pela Universidade de Brunel (Londres) e patrocinado pela União Europeia [3D-MURALE Project].

Finalmente, alguns exemplos de utilização de reconstrução tridimensional no campo da pós-produção cinematográfica podem ver-se na figura 1.6 na próxima página, e da pós-produção televisiva e vídeo na figura 1.7 na página 11.



Figura 1.6: Exemplos de pós-produção cinematográfica recorrendo ao uso de reconstrução tridimensional – tirado de EPRSC [2000], com a permissão gentil do Laboratório de Gráficos e Visão do Departamento de Ciências Computacionais da Universidade de Glasgow.

1.2 Técnicas de Recuperação de Informação Tridimensional

Vários tipos de técnicas de *scanning* 3D têm sido usadas na digitalização de realidades tridimensionais. A maior parte destes sistemas são baseados em *visão por computador*, podendo ser basicamente subdivididos nas seguintes categorias, dependendo da utilização ou não de uma fonte de energia para iluminar a cena:

- *Sistemas activos:*
 - Projecção de luz estruturada;
 - *Shadow scanning*.
- *Sistemas passivos:*
 - Monoculares estruturados;
 - Multi-oculares;
 - Câmaras não-planares;
 - *3D structure from motion*;
 - Fluxos de imagens;
 - Focagem selectiva.



Figura 1.7: Cenário virtual encomendado pela BBC (imagem em baixo), criado a partir de uma recuperação estereoscópica (duas imagens em cima) – tirado de EPRSC [2000], com a permissão gentil do Laboratório de Gráficos e Visão do Departamento de Ciências Computacionais da Universidade de Glasgow.

Além destas técnicas, temos ainda

- o *estéreo fotogramétrico*, e os “*3D structure from shading*” e “*3D structure from texture*”, técnicas geralmente classificadas como passivas na literatura, apesar de dependerem da variação activa de condições de iluminação para funcionar;
- sistemas mistos (como exemplo destes últimos, veja-se Dalmia and Trivedi [1996b]).

Porém, outras tecnologias além dos sensores de visão foram usadas na recuperação e reconstrução tridimensional; estas são usadas por outro grupo importante de sistemas de *scanning* 3D — as *sondas tridimensionais*.

Todas estas técnicas de recuperação de informação tridimensional, dando um ênfase especial aos sistemas de projecção de luz estruturada e aos sistemas multi-oculares, que tiveram mais sucesso até hoje que os seus pares, e incluindo também os sistemas encontrados na natureza que servem de inspiração a uma boa parte delas, serão apresentados nas secções que se seguem.

1.2.1 A percepção da tridimensionalidade na Natureza

A Biologia apresenta-nos variados exemplos de sensores naturais de que os animais dispõem para analisar o mundo tridimensional que os rodeia. É evidente que essa análise é a ferramenta mais importante a que estes recorrem para assegurar a sua sobrevivência. Tendo em conta todos estes exemplos de sucesso já existentes na natureza, como seria de esperar, a grande maioria dos sistemas de digitalização tridimensional inspiram-se claramente nestes sensores.

Boa parte do sistema sensorial animal é constituído por órgãos que complementam a informação geométrica que é possível medir no mundo: o olfacto e o paladar são exemplos óbvios deste tipo de órgãos, permitindo medir informação respectivamente sobre os cheiros e sabores que permeiam o mundo circundante.

Outro sensor menos óbvio mas não menos interessante é o sistema vestibular existente no ouvido interno de alguns animais e cujo objectivo é dotá-los de sentido de equilíbrio e de orientação relativa através da medição de acelerações lineares e velocidades angulares — não é, portanto, nada mais que um sensor inercial. Para saber mais sobre o sistema vestibular humano tendo em vista a sua utilidade como sensor, leia-se a discussão interessantíssima sobre este assunto que pode ser encontrada em Lobo [2002].

No entanto, são os sensores que permitem uma medição efectiva da tridimensionalidade do mundo que os rodeia que são fulcrais para a sua sobrevivência e à volta de cujo uso a sua vida revolve.

Existem animais, por exemplo, que para este efeito estão equipados com *sensores de ultrassons* inatos, constituídos por um sistema de audição apurada conjuntamente com um sistema vocal preparado para emitir sons de alta-frequência. Através do acto de emissão-recepção de ultrassons e da determinação subconsciente da distância percorrida pelas ondas sonoras através da medição instintiva do tempo gasto nesse percurso, esses animais conseguem informação complementar valiosa sobre a tridimensionalidade do ambiente circundante. Como exemplos, entre outros, temos os morcegos e os mamíferos marinhos, nomeadamente os golfinhos e as baleias.

Outro sistema sensorial activo usado por todos os animais é o *tacto*. Alguns usam mesmo os seus membros anteriores (no caso humano, as mãos) como *sondas tácteis*, através das quais medem intencionalmente as formas geométricas tridimensionais e a consistência do mundo que os rodeia.

Todavia, os mais importantes sensores naturais de percepção são sem sombra de dúvida os *sensores de visão*. Estes baseiam-se na medição passiva (ou seja, não intrusiva) através de *células fotorreceptoras* de emissões e reflexões de energia luminosa originadas na realidade circundante.

Na maior parte dos animais, estas células fazem geralmente parte dos *olhos*, órgãos de forma aproximadamente esférica alojados em cavidades ósseas denominadas órbitas e protegidos cada um por duas lâminas cutâneas, as pálpebras [Porto Editora 2002].

Este órgão, no ser humano, desempenha a sua função através de uma porção frontal transparente, a *córnea*, que se contrai ou expande, transmitindo e focando luz de forma a projectá-la numa camada foto-sensitiva na parte posterior, a *retina* [Grolier 1971]. Um estudo mais aprofundado sobre esta questão foi feito no decurso deste trabalho, encontrando-se apresentado com mais detalhe em Ferreira [2002].

Porém, parece bastante óbvio que um sistema de visão com um olho apenas não será uma solução muito eficaz para a sensação da tridimensionalidade do mundo — veja-se que, analogamente a uma câmara, sendo o olho um sensor que apenas capta imagens bidimensionais, existe informação de profundidade que é perdida no processo de projecção. Portanto, não será apenas pelas razões relativamente comuns de redundância de órgãos ou para o alargamento do campo de visão que é regra nos animais o sistema *binocular* ou de *estereovisão*.

A percepção da tridimensionalidade através da estereovisão funciona basicamente explorando a disparidade entre as imagens captadas pelos olhos para fornecer de forma “quantitativa” a informação de profundidade de que careceria cada uma das imagens por si só.

Por outro lado, torna-se evidente que a “terceira dimensão” não se desvanece se o mundo for visto por um só olho. A questão é que a percepção da tridimensionalidade não se esgota com a visão binocular — o cérebro humano utiliza várias estratégias para complementar a percepção 3D [Grolier 1971]:

- com a focagem selectiva, onde o ser humano infere distância a partir das diferenças de tensão dos músculos oculares na contracção ou distensão que sofrem para focar determinado objecto no campo de visão;
- com informação de cariz mais “qualitativo”, através da conjugação da experiência adquirida (isto é, a noção prévia dos tamanhos previsíveis de objectos conhecidos):
 - com inferências sobre a distância dos objectos a partir da noção de perspectiva — ou seja, um objecto mais próximo parece maior e vice-versa;
 - com inferências sobre distância dos objectos a partir de diferenças de cor e iluminação;
 - com o recurso a inferências sobre oclusões — isto é, um objecto mais próximo oculta um objecto mais longínquo, se se interpuser na linha de visão.

Outra forma de complementar a percepção tridimensional do sistema de visão consiste no seguimento de cenas em movimento (o processo designado por *3D structure from motion*); neste caso, o ser humano move-se de forma a ter várias perspectivas de uma cena e o seu sistema vestibular junta-se ao sistema de visão para formar um “super-sistema” sensorial que, seguindo a trajectória de um objecto no campo de visão, permite inferir a posição deste relativamente ao resto da cena — para compreender melhor este processo de colaboração próxima entre estes sistemas sensoriais, consulte-se Lobo [2002].

Além do sistema binocular, porém, existem outras configurações de sistemas de visão na natureza. Alguns insectos, por exemplo, têm sistemas multi-oculares constituídos por centenas de pequenos sensores visuais, funcionando como minúsculos microscópios, cada um por sua vez composto por dezenas de fotorreceptores. A vantagem destes sistemas reside na maior facilidade em analisar a tridimensionalidade com grande precisão, tendo como desvantagem a sua complexidade que geralmente tem como contrapartida a simplificação do sistema de processamento que lhe está associado.

1.2.2 Projectção de luz estruturada

A técnica mais utilizada como solução para a digitalização tridimensional será a projectção de luz estruturada, por razões relativamente fáceis de compreender, visto ser difícil obter características tridimensionais recorrendo apenas a imagens resultado de projectções do mundo, sem o recurso a outros artifícios — é, portanto, uma solução eficaz para digitalizações de precisão média-alta [Ferreira et al. 2002a].

Neste sentido, em seguida far-se-ão uma série de considerações práticas sobre a implementação destes sistemas e finalmente dar-se-ão exemplos dos diferentes sistemas que se têm construído baseados nesta técnica.

De entre as variadas considerações de ordem prática sobre a concepção de sistemas deste género, podem destacar-se como sendo as mais importantes as características espectrais da luz projectada e a influência da geometria de estruturação desta projectção [DePiero and Trivedi 1996].

Para compreender como o espectro de luz e os diversos tipos de emissão luminosa estão relacionados, atente-se a figura 1.8 na próxima página.

É possível verificar nessa figura que a maioria das lâmpadas de iluminação ambiente está concebida de modo a ter como zona de funcionamento a gama de comprimentos de onda que o olho humano consegue perceber (chamada espectro visível, indo aproximadamente desde os 350 nm até aos 750 nm), à excepção das lâmpadas de filamento de tungsténio (devido ao seu péssimo rendimento, que resulta em efeitos de Joule) e as

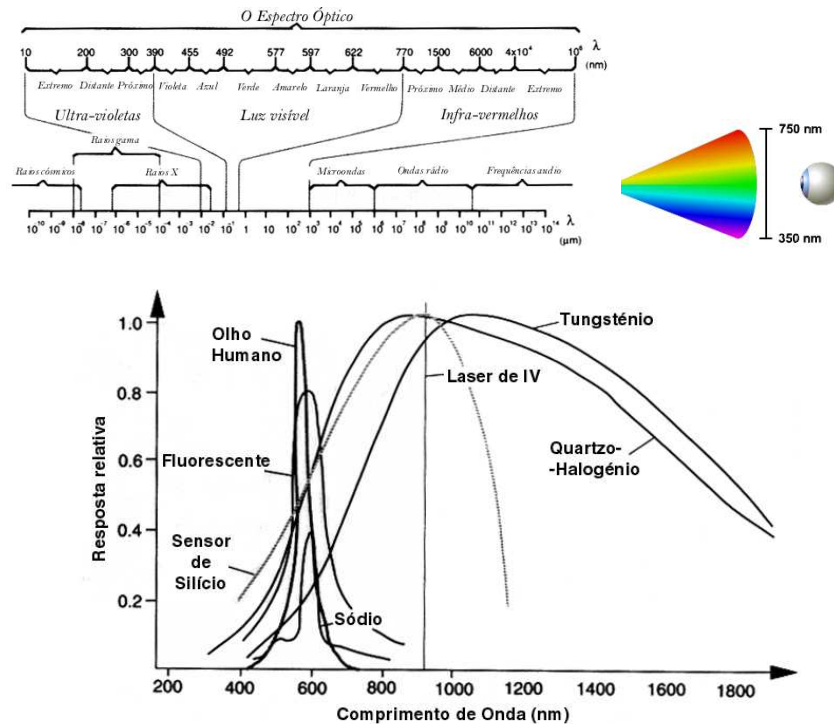


Figura 1.8: Espectro electromagnético com espectro visível, cores que compõem o espectro visível e respostas relativas de diversos dispositivos de emissão de luz comparativamente à resposta relativa do olho humano, em função do comprimento de onda – adaptado de [Awcock and Thomas 1996a] com a gentil permissão do Doutor Graeme Awcock.

lâmpadas de halogéneo.

Como a iluminação ambiente será possivelmente necessária para o operador do digitalizador e até mesmo, por exemplo, para a recuperação das texturas da cena, uma solução de engenharia interessante seria o uso de um tipo de luz cujos efeitos fossem relativamente fáceis de filtrar dos efeitos da luz não projectada. Como as lâmpadas de halogéneo e de filamento têm pouco uso em ambiente controlado, seja ele fabril ou laboratorial, pode verificar-se pela figura que uma excelente escolha para o projectador com esta ideia em mente seria o laser de infravermelhos, emissor de luz colimada a um comprimento de onda preciso (à volta dos 950 nm).

Veja-se que, como vantagem adicional, as câmaras de sensores de silício (incluindo as CCD), como se pode verificar pela figura, têm a resposta ideal aos infravermelhos, conquanto não venham equipadas com filtros internos. Usam-se filtros passa-banda para a gama de frequência de infravermelhos, os chamados *filtros de interferência*, e algoritmos de binarização de imagens para conseguir separar as zonas iluminadas de interesse do resto da imagem para processamento.

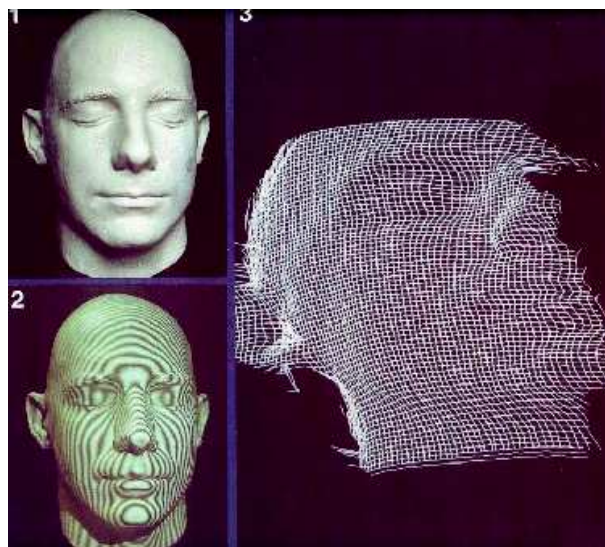


Figura 1.9: Exemplo de sombras de Moiré — à esquerda, a cena original e o efeito da projecção de luz; à direita o resultado da digitalização 3D.

O uso de laser relativamente a outros tipos de luz traz, porém, algumas implicações negativas, como a dificuldade em conseguir captar simultaneamente com uma câmara informação de cor a partir da cena a recuperar, seja devido aos filtros acima mencionados, seja devido à própria interferência do laser na luz reflectida pelo objecto. Além disto, é uma tecnologia ainda relativamente cara, comparando com outras fontes de energia luminosa.

Os sistemas de projecção de luz estruturada funcionam maioritariamente através da modelação dos percursos ópticos associados à emissão e detecção para efectuar a digitalização por triangulação [DePiero and Trivedi 1996]. As geometrias de estruturação da projecção de laser são geralmente simples, usando-se, por exemplo, projecção de uma *recta* (cujo resultado de intersecção com a cena a recuperar é um ponto) ou de um *plano* (cujo resultado de intersecção com a cena a recuperar é um perfil curvilíneo), de forma a simplificar também os algoritmos de triangulação.

Porém, seja para compensar as desvantagens inerentes quando se usa luz natural, seja por se querer acelerar o processo de digitalização quando se usa o laser, recorre-se muita vez à redundância inata de outras geometrias de projecção mais complexas.

Destas geometrias, a mais usada e mais óbvia será a grelha, seja ela de rectas (pontos) ou de planos (perfis). Outra solução interessante é a técnica das *sombras de Moiré*, onde se projecta luz usando uma fonte linear através de uma grelha óptica de espaçamento periódico — o resultado visível é uma imagem com uma espécie de curvas de nível (veja-se a figura 1.9).

Na prática, porém, torna-se claro que existe nestas geometrias mais complexas um problema difícil de correspondências para resolver. Por outras palavras, como ter a certeza de que um algoritmo automático distingue correctamente que recta/plano é que, numa determinada posição na imagem, está a ser observada/o?

Por esta razão, foram recentemente introduzidas técnicas de projecção diferentes — um exemplo é a *projecção de sequências de listas em padrões baseados em código de Gray*. Esta técnica sofre, apesar de tudo, de algumas deficiências: por exemplo, sendo um processo fundamentalmente discreto, a informação da orientação de projecção apenas pode ser obtida na fronteira das regiões, sendo necessário usar interpolação; é um processo muito sensível a erros devido a reflexões, já que estas comprometem seriamente a descodificação das tiras.

Os primeiros sistemas de recuperação de estruturas 3D não foram concebidos na prática para recuperar essas estruturas de forma explícita e quantitativa; a sua verdadeira missão era verificar/validar certas características tridimensionais de objectos para propósitos de inspecção visual industrial.

O mais popular entre estes seria *o sistema de iluminação através de uma lista de luz* [Dudek and Jenkin 2000a]. Este sistema consistiria habitualmente de um tapete rolante transportando os objectos a ser inspeccionados, uma câmara apontada perpendicularmente à superfície desse tapete e um projector de uma lista de luz, posicionado de forma a que o ângulo de incidência da luz fosse conhecido (geralmente de 45° de forma a facilitar os cálculos), de novo relativamente ao plano do tapete rolante. Uma esquematização da sua implementação e um exemplo do seu uso são mostrados na figura 1.10 na próxima página.

Mais recentemente têm-se concebido sistemas usando luz projectada com capacidades para recuperações tridimensionais mais intencionais e precisas.

Entre estes tem-se assistido a uma evolução muito grande nos que se baseiam na projecção de laser: os *laser rangefinders*.

Os sistemas *laser rangefinder*, a que por vezes também se dá o nome de *telémetros*, são sensores de varrimento com laser amplamente usados em aplicações de robótica móvel e sistemas autónomos para obter estimativas de distância a objectos do ambiente circundante. Três tipos deste sensor existem, baseados na mesma tecnologia mas com aproveitamentos diferentes desta [Dudek and Jenkin 2000b, Ferreira et al. 2002a]:

- **Triangulação.** Estes sistemas funcionam baseando-se nas relações geométricas entre o laser emitido, o laser reflectido e a sua posição num plano-imagem captado por um sistema de visão — o modelo para este sistema é equivalente ao do sistema mais genérico de projecção de luz estruturada;

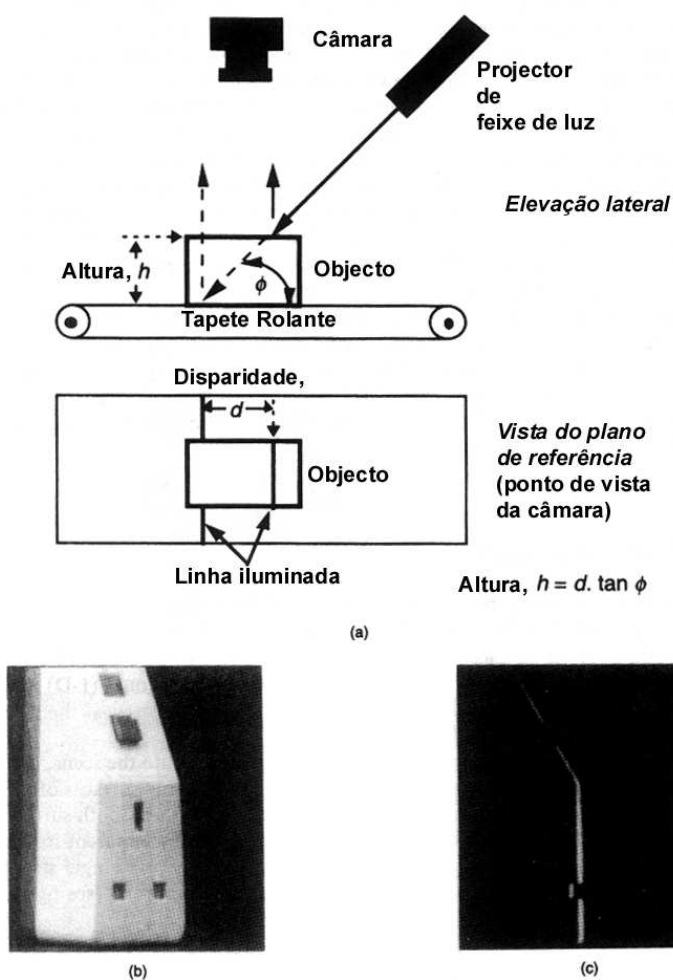


Figura 1.10: Sistema de inspeção com projector de lista de luz, câmara e tapete rolante (adaptado de [Awcock and Thomas 1996b] com a gentil permissão do Doutor Graeme Awcock).

- **“Time-of-flight”**. Estes baseiam-se na relação entre a distância ao objecto e o tempo que decorre entre o laser emitido e o laser reflectido;
- **“Phase-based”**. Estes utilizam a relação entre a distância ao objecto e a diferença de fase entre o laser emitido e o laser reflectido.

Qualquer dos dois últimos precisam de um uso da tecnologia laser de tal modo exigente para obter resultados precisos que, apesar de serem muito mais eficientes do que o primeiro, têm um preço proibitivo para a maioria das aplicações.

Porém, como foi dito na secção anterior, devido aos problemas práticos impostos pelo recurso ao laser, tem-se visto um recurso cada vez maior às técnicas de projecção de sequências de listas em código gray.

Tem-se também efectuado um estudo aprofundado na procura de técnicas diferentes que primem por ser de baixo custo, baixa complexidade e não sofrer dos problemas inerentes à análise directa da reflexão da luz. Neste sentido, desenvolveu-se recentemente a técnica de *“shadow scanning”*, cujo princípio de funcionamento se baseia na emissão de luz com uma fonte pontual e no movimento de um objecto que interfere nessa emissão, projectando uma sombra com a qual se faz a triangulação [Bouguet and Perona 1999].

1.2.3 Sistemas multi-oculares

Uma aproximação alternativa à solução do problema de recuperação de estruturas 3D é a exploração da geometria *multi-ocular*, nomeadamente a *estereovisão*. Esta parece, aliás, intuitivamente a configuração mais intuitiva, já que o ser humano obtém informação 3D do mundo circundante usando o seu próprio sistema estereovisual.

Qualquer sistema consistindo em duas ou mais fontes de imagem bidimensionais representando projecções do mundo tridimensional real pode ser descrito pela sua *geometria epipolar*, de que a figura 1.11 na página seguinte é o exemplo binocular.

Nela podem ver-se dois planos-imagem captados por duas câmaras, um ponto no “mundo real” 3D, as rectas de projecção desse ponto em cada um desses planos e os seus respectivos pontos de projecção — este exemplo é facilmente generalizável para n câmaras. A restrição imposta pela geometria epipolar implica que qualquer ponto P terá então que pertencer ao mesmo plano que os seus n pontos-imagem (e que os n pontos focais) pertencentes a cada uma das n câmaras do sistema, plano esse a que se dá o nome de *plano epipolar*. Este plano, por sua vez, intersecta os n planos-imagem formando as chamadas *rectas epipolares* [Hartley and Zisserman 2000, Faugeras 1993, Ferreira and Dias 2000, Ferreira et al. 2001, Dias 1994].

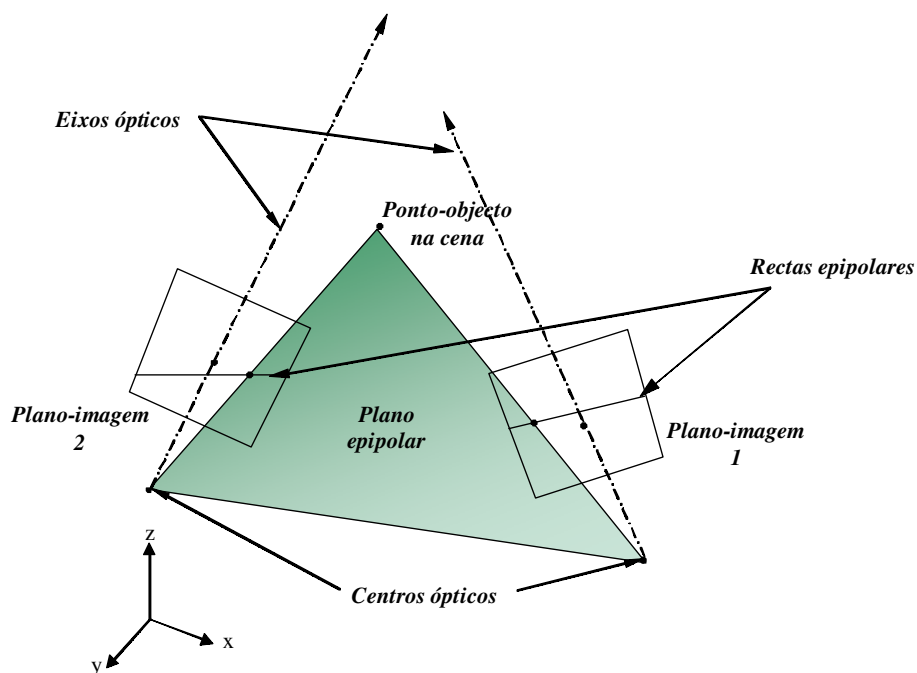


Figura 1.11: Geometria epipolar de um sistema de estereovisão [Ferreira and Dias 2000, Ferreira et al. 2001].

A conjugação das n relações para os n planos-imagem fornece o número de restrições necessárias para estabelecer uma relação unívoca entre um ponto 3D e os seus pontos de projecção nos planos-imagem, cuja correspondência é determinada a partir do uso das restrições impostas pela geometria epipolar [Hartley and Zisserman 2000, Faugeras 1993, Ferreira and Dias 2000, Ferreira et al. 2001, Dias 1994].

Devido aos problemas inerentes à aquisição de imagens por câmaras não-ideais, cada calibração feita para cada câmara implica uma degradação acumulativa na precisão da recuperação — esta é a desvantagem principal deste tipo de sistemas.

Outras desvantagens são: computações pesadas, algoritmos de correspondência entre pontos de projecção difíceis (geralmente usam-se algoritmos de correlação entre os pontos pertencentes a cada recta epipolar), e o facto de tais sistemas serem dispendiosos, mesmo sendo as câmaras CCD bastante mais baratas hoje em dia.

Como vantagens poder-se-ão referir o aproveitamento da redundância de rectas projectivas que cada câmara acrescenta ao sistema através de algoritmos de optimização, a ausência de necessidade de varrimento de uma perspectiva de uma cena tridimensional mais do que uma vez e a possibilidade de retirar informação de cor sem qualquer interferência.

Para retirar a desvantagem do preço de sistemas multi-oculares, têm-se substituído

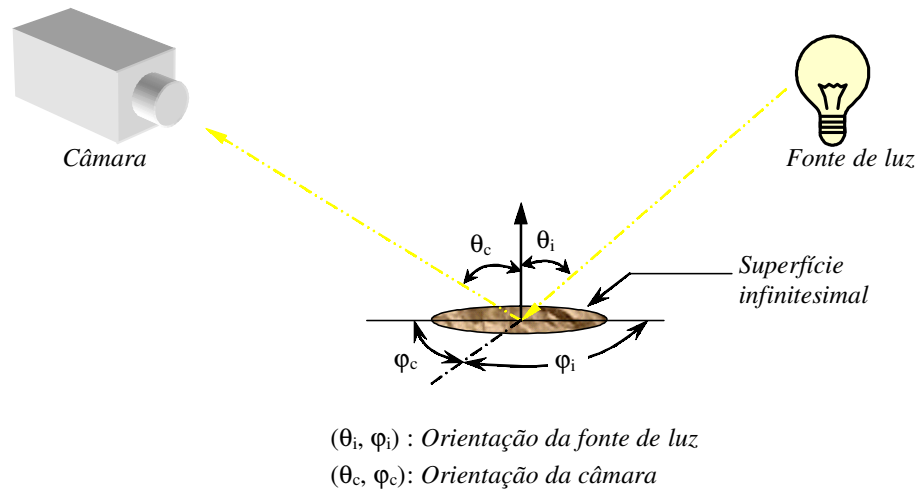


Figura 1.12: Estéreo fotogramétrico.

as câmaras redundantes para fornecer planos-imagem adicionais por espelhos planares — consulte-se, por exemplo, [Martins and Dias 1999]; as restantes desvantagens, porém, permanecem patententes.

1.2.4 Outras técnicas

Para além da projecção de luz estruturada, ainda dentro dos *scanners* 3D baseados em visão, existem várias tentativas para se usar apenas uma câmara para se obter informação tridimensional de uma cena. Como exemplo, temos o *estéreo fotogramétrico*. Este método permite utilizar apenas uma câmara, uma fonte de luz e a informação sobre o gradiente de irradiâncias, ultrapassando a ambiguidade da imagem única através da correlação de várias imagens de irradiância obtidas para geometrias idênticas de visualização mas segundo diferentes condições de iluminação [Woodham 1980] — veja-se a esquematização da figura 1.12.

Uma vez que não existe variação na localização da câmara, este processo tem a vantagem de não depender de algoritmos de correspondência; por outro lado, peca pela sua fraca viabilidade devido à maioria das superfícies de objectos estarem longe de ser lambertianas³, condição necessária para este método funcionar.

Outras técnicas existem que se aproximam muito desta, nomeadamente substituindo a luz simples por uma *textura periódica* [Blostein and Ahuja 1989].

Outro exemplo monocular será o recurso a marcas visuais de geometria e dimensões conhecidas colocadas propositadamente na cena que permitam inferir profundidades e

³Superfícies cujo material fornece uma difusão uniforme da radiação incidente.

efectuar triangulações, os *sistemas monoculares estruturados* — esta técnica, tirando algumas aplicações de inspecção automática industrial, é muito pouco interessante devido às suas limitações óbvias. Por fim, existe ainda o recurso a câmaras *não-planares* ou de *projectão ortogonal*, cujo interesse, pelas mesmas razões, é muito limitado.

Para além das soluções monoculares estáticas, existem ainda as técnicas de “*3D structure from motion*” [A. Chiuso and Soatto 2000] e os *fluxos de imagem* [Dalmia and Trivedi 1996a], que exploram a deslocação relativa da câmara relativamente à cena para retirar desta informação tridimensional — uma desvantagem que se lhes pode apontar será a sua falta de flexibilidade, nomeadamente face a cenas de pequena dimensão —, e a técnica da *focagem selectiva* [Favaro and Soatto 2002], que utiliza profundidades de campo pequenas e a focagem para distinguir profundidades. Esta última técnica tem óbvias limitações a nível de precisão, estando por isso confinada a aplicações de detecção de obstáculos e navegação em robótica móvel.

Porém, existe ainda outro importante grupo de sistemas de recuperação tridimensional que não recorre à visão por computadores — as *sondas tridimensionais*. Estes têm como vantagem relativamente aos sistemas baseados em visão o facto de não apresentarem problemas quanto à recuperação de estruturas tridimensionais de objectos transparentes ou translúcidos. Os sistemas baseados em câmaras necessitam de condições de iluminação especiais para se conseguir algum resultado com este tipo de objectos, e mesmo assim sem garantir precisão.

Estes sistemas podem subdividir-se em 3 tipos [Ferreira et al. 2001, Hindus 2001]:

- **Sondas de contacto**

Sondas de tacto, montadas em sistemas mecânicos de medição de coordenadas, são sistemas ainda muito utilizados nas indústrias automóvel e aeronáutica. Esta forma de aquisição de dados é muito precisa, mas morosa e dispendiosa, estando limitada aos materiais que podem resistir ao contacto mecânico, se bem que, como se poderá ver mais adiante, existam implementações recentes que já reduziram significativamente a extensão do problema.

- **Sondas de ultrasons**

Este tipo de sistemas usam um princípio físico semelhante ao do laser “*time-of-flight*” *rangefinder*, usando o tempo de propagação das ondas ultrassónicas para determinar profundidades.

- **Sondas de ressonância**

As sondas de ressonância electromagnética (IRM) são outra tecnologia de re-

cuperação de estruturas tridimensionais, muito usadas na medicina. Como as sondas ultrassónicas, este tipo de sistema é pouco prático devido à necessidade de uma significativa intervenção humana.

1.3 Técnicas para Recuperações Tridimensionais Completas

As técnicas de recuperação recorrendo a sensores de visão apresentadas na secção precedente não tomam em consideração o facto de que uma cena tridimensional apresenta invariavelmente oclusões, seja porque alguns objectos poderão ser opacos e tapar outros objectos na cena, seja porque é impossível recuperar cada objecto de forma completa a partir de um só ponto de vista. Para solucionar este problema foram concebidas várias técnicas que irão ser apresentadas nas secções que se seguem.

1.3.1 Sistemas fixos com partes móveis

Uma das soluções mais óbvias para este problema será acrescentar partes móveis a sistemas de recuperação tridimensional. Exemplos destes métodos serão:

- **Sistemas mecanizados — cena fixa e sensores móveis:**

Usando um suporte controlado por motores de passo de modo a permitir rotação lateral (*pan*) e vertical (*tilt*), ou uma qualquer estrutura com carris, torna-se possível executar tanto o varrimento da cena com o laser como o redireccionamento e movimentação de uma câmara. De qualquer dos modos, este método não resolve por completo o problema da oclusão a não ser, provavelmente, para objectos pequenos e, portanto, outras soluções tornam-se necessárias.

- **Sistemas mecanizados — cena móvel e sensores fixos:**

Colocando os objectos a recuperar num suporte móvel, sendo o uso de um disco rodando a uma velocidade angular conhecida o exemplo mais comum e óbvio, torna-se possível resolver o problema para cenas simples com poucos objectos.

- **Sistemas multi-oculares distribuídos:**

Outra solução é colocar planos-imagem em posições estratégicas em redor da cena a recuperar de modo a apanhar em todos os pontos de vista qualquer ponto que pudesse estar ocluído. A grande desvantagem deste método é inerente a todos os sistemas multi-oculares: a questão da determinação de correspondências entre

pontos, que piora devido ao facto de existirem pontos que poderão não ser visíveis simultaneamente para duas das câmaras, mesmo que pertençam ao mesmo plano epipolar.

É óbvio que a implementação deste tipo de solução permite o desenvolvimento de sistemas de recuperação tridimensional de âmbito limitado; para se conseguir algo mais versátil será necessário recorrer a outro tipo de técnicas.

1.3.2 Sistemas portáteis

Na procura da flexibilização de sistemas de *scanning* 3D, tem-se recentemente assistido à concentração de esforços das empresas que os comercializam no desenvolvimento de sistemas *portáteis*. A sua vantagem reside na possibilidade de calibrar o sistema apenas uma vez relativamente a um referencial privado e depois, através de várias técnicas de determinação da sua posição e orientação (informação a que se dá o nome de *pose*⁴ do sistema), compensar os movimentos executados para percorrer a cena a recuperar. Este tipo de solução resolve quase todos os problemas de oclusão.

Existem basicamente três formas possíveis de medir a *pose* do sistema:

- o uso de marcas visualmente contrastantes na cena (*landmarks*) colocadas estrategicamente de modo a poder recorrer-se a triangulações para medir as mudanças de *pose*;
- a abordagem emissor/receptor, onde um emissor é colocado numa parte da cena a ser recuperada enquanto um receptor no sistema vai registando as mudanças de posição e orientação relativas ao primeiro recorrendo à leitura das emissões daquele (o emissor e o receptor podem estar colocados de forma inversa, isto é, o primeiro no sistema e o segundo na cena);
- a abordagem “auto-contida”, onde o sistema obtém as suas leituras de *pose* através de uma configuração independente da cena a recuperar.

Qualquer sistema de emissor/receptor utilizando propriedades de tempo de voo ou de mudança de fase de realidades físicas ondulatórias como a luz ou como o som (mais

⁴Designação útil dada por muitos autores à informação de posição e orientação relativamente a um referencial global de um equipamento cuja direcção e sentido com que é orientado é essencial para descrever o seu princípio de funcionamento. Utiliza-se este termo frequentemente nas áreas da Robótica Móvel para modelar a cinemática de veículos autónomos e da Visão por Computador para descrever modelos de câmaras.

concretamente os ultrassons) pode ser utilizado para a segunda abordagem, assim como sistemas magnéticos explorando propriedades físicas como o efeito de Hall.

No entanto, para conceber sistemas usando a abordagem de “auto-contenção” torna-se necessário recorrer a sensores mais específicos.

Os chamados *sistemas de navegação inercial* obtêm informação de velocidade e posição a partir de medidas de sensores inerciais. Ao princípio básico empregue na navegação inercial dá-se o nome de “*deduced reckoning*”, já que a posição num dado instante é determinada a partir da posição num instante anterior, do caminho percorrido entre os dois instantes e da evolução temporal da velocidade ao longo desse percurso. Acelerómetros e/ou giroscópios, por exemplo, são utilizados para medir a aceleração ao longo de três eixos ortogonais, integrando duas vezes para determinar a posição, como pode ser visto em (1.1).

$$\Delta_{position} = \mathbf{x} = \int \dot{\mathbf{x}} dt = \int \int \ddot{\mathbf{x}} dt = \int \mathbf{a}_{sensed} dt \quad (1.1)$$

É possível, no entanto, substituir os giroscópios por sensores magnéticos ou *magnetómetros* para desempenhar a mesma função. Este sensor tem a vantagem de tornar redundante o uso de acelerómetros, visto que as suas medições não são afectadas por acelerações do sistema. A sua maior desvantagem é a sua sensibilidade a distorções no campo magnético terrestre resultantes da proximidade de materiais ferromagnéticos.

Uma apresentação mais aprofundada de um estudo sobre estas tecnologias pode ser encontrado em Ferreira [2002].

Para finalizar, é de referir que não é conhecido deste autor qualquer sistema de *scanning* 3D que use os métodos “auto-contidos” de medição de *pose* acima referidos.

Alguns sistemas de sensores de orientação comercializados actualmente estão organizados na tabela 1.1 para fins comparativos.

Tabela 1.1: Exemplos de sistemas de sensores de *pose* (características fornecidas pelos fabricantes) – X/Y/Z - Posição 3D; Y - Yaw; P - Pitch; R - Roll; H - Heading; Incl. - Inclination; E-R - Emissor-Receptor; Mag. - Magnético(s); Marc. Ópt. - Marcas Ópticas (Activas ou Passivas); IV - Infravermelhos; A - Acelerómetros; M - Magnetómetros (Genéricos); fM - Magnetómetros *Fluxgate*; G - Giroscópios; N/D - Informação Não Disponível.

Modelo	Fabricante	Valores Medidos	Tecnologia	Tipo de Saídas	Resolução	Precisão
<i>FASTRAK</i>	Polhemus	X/Y/Z e Y/P/R	E-R Mag.	dig. RS-232/422	$2 \mu\text{m}/\text{cm}$, $.025^\circ$	.8 mm, $.15^\circ$
<i>pciBIRD</i>	Ascension	X/Y/Z e Y/P/R	E-R Mag.	PCI-miniDIN	.5 mm, $.1^\circ @ 35 \text{ cm}$	1.8 mm, $.5^\circ$
<i>miniBIRD 500</i>	Ascension	X/Y/Z e Y/P/R	Mini E-R Mag.	dig. RS-232	.5 mm, $.1^\circ @ 35 \text{ cm}$	1.8 mm, $.5^\circ$
<i>OPTOTRAK</i>	NDI	X/Y/Z	Marc. Ópt. (IV)	N/D	.01 mm	.1 / .1 / .15 mm
<i>POLARIS</i>	NDI	X/Y/Z	Marc. Ópt.	dig. RS-232/422	N/D	.35 mm
<i>3DM</i>	Microstrain	$\pm 180^\circ / \pm 180^\circ / \pm 70^\circ$ (Y/P/R)	3M-3A	dig. RS-232	$< .1^\circ$	$\pm .7^\circ / \pm .7^\circ / 1.5^\circ$
<i>AHRS</i>	Crossbow	$\pm 180^\circ / \pm 180^\circ / \pm 90^\circ$ (H/P/R)	3G-3M-3A	anal. DB15	$\leq .1^\circ$	$\pm 3^\circ / \pm 2^\circ / 2^\circ$
<i>CXM539</i>	Crossbow	Campo Magn.	3fM	dig. e an. RS-232	N/D	$\pm 3 \text{ Gauss}$
<i>CXM547</i>	Crossbow	Azim./Incl./Toolface	3M-3A	dig. RS-232	N/D	$\pm 1.5^\circ / \pm .5^\circ / \pm .5^\circ$
<i>CXM543</i>	Crossbow	$\pm 360^\circ / \pm 90^\circ / \pm 180^\circ$ (Y/P/R)	3M-3A	dig. RS-232	N/D	$\pm 1.5^\circ / \pm .5^\circ / \pm .5^\circ$
<i>CXM544</i>	Crossbow	Yaw/Pitch/Roll	3fM-3A	dig. RS-232	N/D	$\pm 1.5^\circ / \pm .5^\circ / \pm .5^\circ$

1.3.3 Integração de informação tridimensional

Considere-se a seguinte descrição do uso de um sistema móvel de recuperação tridimensional (extensível evidentemente a sistemas fixos com partes móveis, como por exemplo *scanners* de disco rotativo) — câmaras ou projectores de luz estão fixados numa estrutura rígida que por sua vez será deslocada de forma a que toda a cena tridimensional seja percorrida. Imagens da cena serão capturadas em sequência enquanto o varrimento é efectuado e superfícies ou perfis curvilíneos tridimensionais serão medidos. Evidentemente, para que a informação tridimensional adquirida desta forma seja *integrada* num modelo global para a cena afectado a um único referencial, torna-se necessária a determinação de todas as translações e rotações intermédias do referencial privado do sistema de modo a realizar-se o processo a que se dá o nome de *registo 3D* (do inglês *3D registration*) da informação. Este processo foi definido de várias maneiras dependendo de cada implementação (consulte-se [Hilton 1997, Audette et al. 2000, Besl and McKay 1992, Huber 2000] para exemplos), mas pode ser formalmente definido de uma forma mais genérica como a *transformação de conjuntos de medidas tridimensionais num sistema de coordenadas comum*. Para se compreender melhor este processo, atente-se a figura 1.13 na página seguinte.

O conjunto de pontos recuperados $\mathbf{x}(t_n)$ para cada ponto de vista capturado em cada instante de amostragem $t = t_n$, é denominado $\mathbf{P}(t_n)$. O *scanner* e os seus referenciais locais, como já foi referido, são movidos e rodados ao longo da recuperação da cena. Como não estão envolvidas neste processo nem mudanças de escala nem reflexões, cada transformação de *pose* do sistema entre instantes de amostragem consecutivos pode ser modelada como uma *transformação rígida*, uma combinação de uma rotação e uma translação dada por [Ferreira et al. 2002a, Audette et al. 2000, Feldmar and Ayache 1994, Faugeras and Hebert 1986]

$$\mathbf{P}(t_n) = {}^n\mathbf{R}_{n-1}\mathbf{P}(t_{n-1}) + {}^n\vec{\mathbf{t}}_{n-1} \quad (1.2)$$

onde $\mathbf{x}(t_n)$ corresponde a qualquer ponto amostrado no instante $t = t_n$, com coordenadas afectadas a um qualquer referencial local do sistema.

A técnica de registo 3D é caracterizada principalmente pela maneira como resolve três problemas cruciais [Ferreira et al. 2002a, Audette et al. 2000, Besl and McKay 1992, Faugeras and Hebert 1986]:

1. Escolha do espaço de características (isto é, do tipo de medidas 3D usadas: pontos, curvas, planos, imagens voxel, etc.);

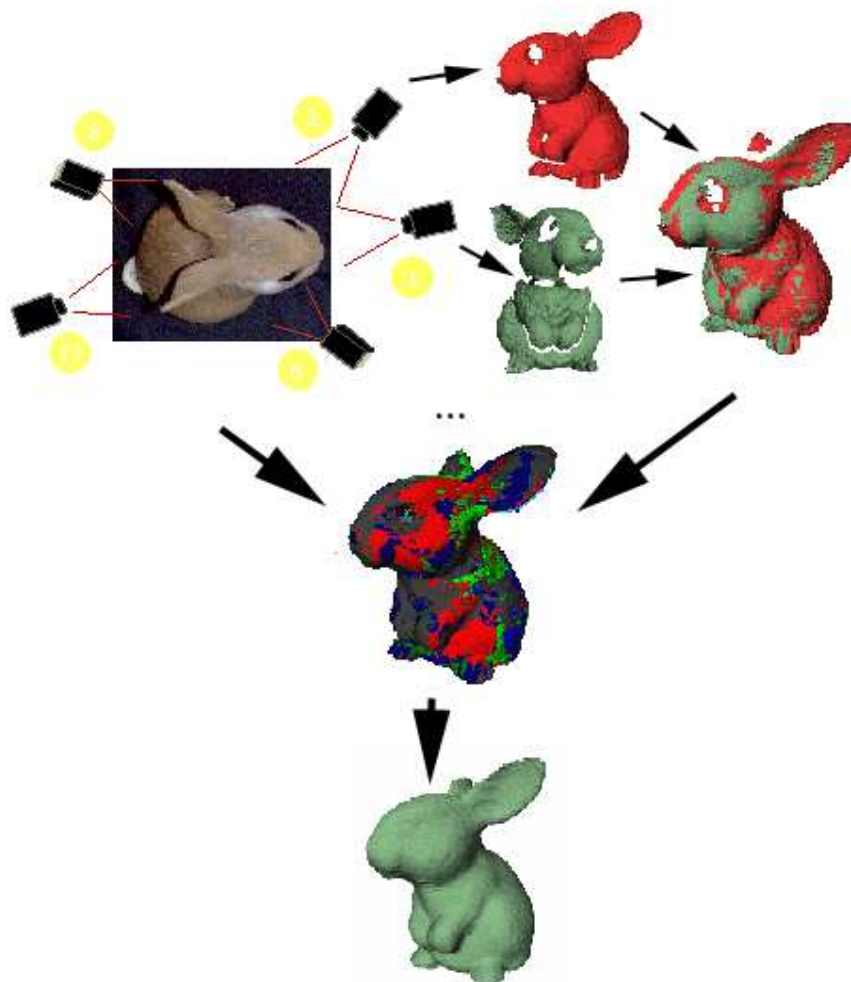


Figura 1.13: Exemplo de registo tridimensional – adaptado de [Huber 2000]. Veja-se a evolução desde os conjuntos de pontos retirados de cada ponto de vista de captura de imagens até à integração da informação num registo final afectado ao mesmo referencial.

2. Escolha da transformação (incluindo da representação da rotação — através de uma matriz ortonormal, de quatérnions, etc.) e, consequentemente, de uma função-objectivo a ser minimizada e do seu vector de parâmetros;
3. Correspondências entre características (este processo é obrigatório, já que a técnica de registo 3D é impossível sem ele) e a optimização global.

Repare-se que este método (confronte-se esta ideia com o que é dito em [Hébert and Rioux 1998]) poderá ser usado conjuntamente com sensores de *pose* para providenciar boas estimativas iniciais para a orientação e até para a posição do sistema de modo a melhorar o processo de integração [Ferreira et al. 2002a].

1.4 Técnicas de Reconstrução de Cenas Tridimensionais

Nesta secção descrever-se-ão de forma breve as ferramentas visuais usadas para a reconstrução das cenas tridimensionais a partir da informação recuperada pelo sistema sensorial. Apesar de não serem parte integrante dos *scanners* 3D discutidos neste capítulo, têm no entanto uma influência indirecta importante na forma como se armazena a informação recuperada.

1.4.1 Computação gráfica 3D

A forma mais comum, e durante muito tempo única, de realizar reconstruções tridimensionais é usando computação gráfica tridimensional. Uma razão muito importante para isto é a facilidade com que estas são executadas, visto que a informação é processada e visualizada no próprio computador onde foi armazenada; outra razão, não menos importante, é o facto de a área computação gráfica estar a sofrer constantes avanços devido à procura muito grande de soluções multimédia (como por exemplo o *MMX* em termos de hardware e o *DirectX* e o *OpenGL* em software) para galvanizar o mercado de videojogos, de produção video e de CAD.

Partindo do pressuposto que toda a informação adquirida pelo sistema sensorial está toda referida a *um só* sistema de eixos, existem várias formas básicas de reconstruir as cenas tridimensionais, sendo as duas mais comuns as *representações volumétricas* e as *representações superficiais*. O primeiro método utiliza uma unidade idealmente infinitesimal de volume, o *voxel* (reaproveitamento do pixel para volumes — *volume picture element*) [Kutulakos and Seitz 2000], para reconstruir os volumes tridimen-

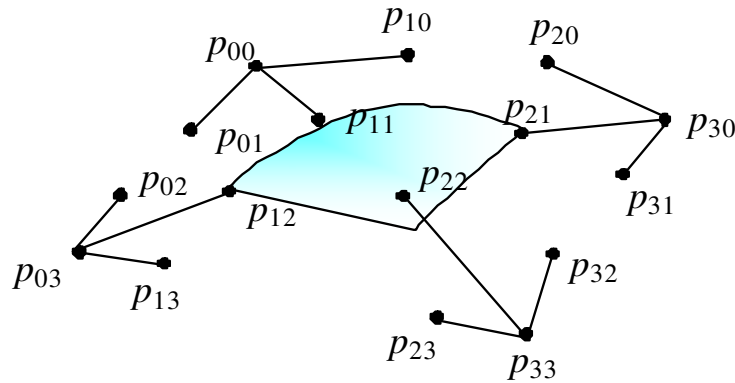


Figura 1.14: Exemplificação de um B-Spline – veja-se a necessidade de se usar 16 pontos de controlo para obter uma superfície suave, que se ajusta apenas aos 4 pontos centrais.

sionais que compõem a cena global. O segundo método recria os objectos da cena aproveitando pontos amostrados a partir das suas superfícies para as reconstruir.

O primeiro método não é tão conveniente quanto ao tipo mais comum de informação tridimensional recuperada por scanners 3D — a *nuvem de pontos* amostrados. Apesar de tudo, existem alguns estudos que reformulam todo o conceito de recuperação de estruturas tridimensionais de forma a utilizar os *voxels*: a técnica de *space carving* que, numa tradução livre, significa escavação ou escultura do espaço — para mais informação sobre este assunto, consulte-se [Kutulakos and Seitz 2000].

Porém, como ainda não foram apresentados resultados suficientemente convincentes neste campo, o método de longe mais utilizado é efectivamente a representação superficial, onde se ajustam superfícies às nuvens de pontos recuperadas. Este método implica a resolução de dois problemas — a reconstrução local, ou seja, a problemática do ajustamento da melhor superfície a pequenas quantidades localizadas de pontos, e a reconstrução global, isto é, a problemática do processamento da nuvem de pontos recuperados de forma a inferir correctamente, a partir desta, a topologia dos objectos da cena a reconstruir.

A primeira questão já foi resolvida com bastante sucesso recorrendo ao uso de “retalhos” de superfície idealmente infinitesimais com geometria variada — triângulos, polígonos, superfícies quádricas interpolantes, superfícies de Bezier, B-Splines (exemplificados pela figura 1.14), NURBS (Non-Uniform B-Splines), etc. Para um estudo mais aprofundado sobre estes assuntos, consulte-se [Edward 1990] ou [Farin 1999]. Existem variadas ferramentas de computação gráfica que implementam facilmente estas superfícies: como exemplos temos o *OpenGL*, o *DirectX*, ou ferramentas de mais alto nível como o *Visual ToolKit (VTK)*.

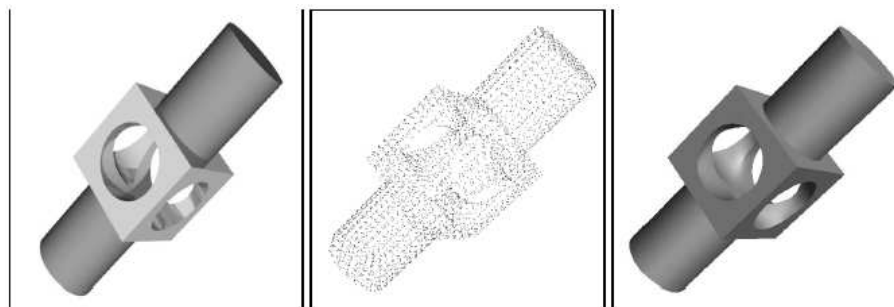


Figura 1.15: Exemplificação de uma reconstrução conseguida recorrendo ao algoritmo de Hugues Hoppe, em que da esquerda para direita se vêem consecutivamente o objecto original, os pontos amostrados e a sua reconstrução final – tirado de [Hoppe 1994] com a gentil permissão do Professor Hugues Hoppe. O algoritmo, como se pode verificar, inferiu correctamente a topologia complicada do objecto a partir dos dados de entrada.

Quanto ao segundo problema, existem várias soluções empolgantes a ser utilizadas com resultados bastante convincentes. Dois casos de sucesso serão o algoritmo concebido por Hugues Hoppe (*Surface Reconstruction from Unorganized Points* — consulte-se [Hoppe et al. 1994, Hoppe 1994] para mais informação) e um algoritmo recente que está a revolucionar a comunidade de computação gráfica concebido entre outros por Nina Amenta (o *PowerCrust* — consulte-se [Nina Amenta and Kolluri 2001] para mais informação), que tem como vantagem relativamente ao primeiro, além de uma maior eficiência, uma menor exigência relativamente à densidade espacial da nuvem de pontos.

1.4.2 Outras técnicas

Apesar da computação gráfica ser, sem dúvida, a principal ferramenta de reconstrução tridimensional utilizada em aplicações deste tipo, não é, de facto, a única. Novas tecnologias estão a surgir que irão revolucionar as técnicas de *fac-simile*: o *faxing 3D*.

Está-se a assistir ao começo de um estudo importantíssimo na reconstrução *física* de objectos tridimensionais, a técnica denominada *solid free-form fabrication*, que permitirá a re-instanciação remota de modelos de objectos, conceito inerente ao *fac-simile* [Curless and Levoy 1996-97, Hoppe et al. 1994, Hoppe 1994, Levoy et al. 2000]. Uma empresa que se notabilizou neste campo foi a 3D Systems Inc., usando um processo a que deram o nome de *estereolitografia*.

Como exemplo da sua utilização, pode-se referir que a empresa foi contratada pelo Professor Marc Levoy e a sua equipa para executar a reconstrução de uma estatueta de um buda (ver figura 1.4 na página 7). Converteram o modelo computacional que



Figura 1.16: Imagem da reconstrução 3D por estereolitografia de uma estatueta de um buda – tirado de [Curless and Levoy 1996-97] com a gentil permissão do Professor Marc Levoy.

lhês fora enviado por correio electrónico numa pilha de aproximadamente 500 contornos planares cheios espaçados de 150 microns. Para cada contorno, foi depositada uma camada de resina líquida, posteriormente endurecida por varrimento laser para reproduzir o contorno cheio. O buda foi fabricado deitado sobre as suas costas, começando pelas vestes e acabando no estômago, tendo o processo demorado cerca de 10 horas. Na pilha de contornos, foi incluído o suporte para as superfícies apontando para baixo, sendo este suporte preenchido após o processo de fabrico. O modelo final foi então lixado e sujeito a um tratamento com um jacto de partículas para suavizar as suas superfícies (ver figura 1.16 na página ao lado) [Curless and Levoy 1996-97].

Esta tecnologia está, no entanto, em fase embrionária e será difícil ter resultados convincentes ainda dentro dos próximos anos.

1.5 Exemplos de Sistemas de Digitalização Tridimensional

Nesta secção apresentar-se-ão exemplos representativos de sistemas usando as técnicas e tecnologias apresentadas anteriormente. Como a diferenciação mais simples entre sistemas de *scanning* 3D é a separação entre sistemas fixos e portáteis (e, mesmo assim, a fronteira não é claramente definida), será esse o critério de ordenamento das subsecções subsequentes.

1.5.1 Sistemas fixos

Vários sistemas de cariz tanto laboratorial como comercial foram desenvolvidos ao longo dos últimos anos. Até agora, a maioria destes sistemas foram projectados para resolver problemas específicos de recuperação de estruturas tridimensionais sendo, por motivos da disponibilidade tecnológica existente na altura da sua concepção, também na sua maioria sistemas fixos, possivelmente contendo partes móveis. Discutir-se-ão seguidamente precisamente os sistemas que se poderão classificar com mais ou menos segurança como pertencendo a esta categoria.

Os sistemas usando visão e projecção de luz estruturada, mais conhecidos por *laser rangefinders* por triangulação, são sem dúvida os *scanners* 3D mais comuns até à época.

Temos como primeiro exemplo o *scanner* apresentado na figura 1.17 na página 35, desenvolvido pela 3D Scanners. O sistema de recuperação consiste num laser e uma câmara montados numa estrutura que se move sobre carris de modo a varrer o objecto ao longo de um eixo. [3D Scanners].

Outros exemplos de *scanners* laser fixos seriam o 3030RGB/PS (figure 1.18) e o 3030RGB/MS (figure 1.19) da Cyberware — o primeiro também é conhecido como o “*Head & Face 3D Color Scanner*” e o último como o “*Model Shop Color 3D Scanner*”; as suas aplicações específicas transparecem claramente nas próprias designações [Cyberware]. Como se pode ver pelas figuras, ambos usam também carris para movimentar o sistema de recuperação 3D.

Um sistema revolucionário em termos de utilização do laser na recuperação tridimensional é o Arius3D *scanner*. Este sistema, que não é vendido como equipamento mas cujos serviços se podem alugar, foi concebido pelo CNRC (Canada’s National Research Council) e utiliza um método patenteado consistindo no uso de três comprimentos de onda — vermelho, verde e azul — de laser para capturar simultaneamente as coordenadas geométricas e de reflectância, ou seja, vectores de valores $[X, Y, Z, R, G, B]$. Os resultados afirmam-se como não sendo afectados pela iluminação ambiente, consistindo de dados de cor “de utilizador”, sem artifícios de luz, a uma resolução microscópica [Hindus 2001]. A Arius3D afirma estar a desenvolver um sistema novo, portátil, o Arius3D NOMAD, que assegurará aos seus clientes a possibilidade de fazer recuperações de cenas que não sejam passíveis de ser transportadas para as instalações da empresa.

Um sistema de *scanning 3D* diferente é o TriForm da 3D Scanning Systems, pertença da Wicks and Wilson Limited, mostrado na figura 1.20. Este sistema usa luz (branca) visível numa grelha linear para recuperar estruturas tridimensionais [3D Scanning Systems 2000]. Como se pode ver na figura, uma cabine de grande dimensão é utilizada para escudar o sistema de interferências luminosas externas. Como é evidente, cada estilo de cabine determinará o tamanho e o tipo do objecto a ser recuperado; também é óbvio, pelas alternativas apresentadas pela empresa, que este sistema foi pensado exclusivamente para a recuperação tridimensional do corpo humano.

Finalmente, usando uma tecnologia de visão diferente, o sistema multi-ocular C3D-2030 da C3D é apresentado na figura 1.21. O C3D-2030 consiste num suporte contendo um par estéreo de câmaras monocromáticas digitais de alta-resolução para recuperação das estruturas 3D e numa câmara policromática de alta-resolução para capturar a aparência da superfície conjuntamente com uma unidade de *flash* de iluminação de texturas superficiais [C3D/Edinburgh Virtual Environment Centre].

1.5.2 Sistemas portáteis

O primeiro exemplo de sistemas portáteis é o modelo VIVID 700 da Minolta, um *scanner* laser 3D (com varrimento por galvanómetro e espelhos), que é afirmado



Figura 1.17: *Scanner* laser REPLICA 3D da 3D Scanners [3D Scanners].



Figura 1.18: *Scanner* laser 3030RGB/PS da Cyberware [Cyberware]. Permissão para reprodução desta imagem gentilmente cedida pela Cyberware.



Figura 1.19: *Scanner* laser 3030RGB/MS da Cyberware [Cyberware]. Permissão para reprodução desta imagem gentilmente cedida pela Cyberware.



Figura 1.20: Cabine TorsoScan TriForm (da Wicks and Wilson Limited). Além desta cabine, estão disponíveis a BodyScanner e a HeadScan. [3D Scanning Systems 2000]



Figura 1.21: O C3D-2030 (de C3D/Edinburgh Virtual Environment Centre) – um sistema de *scanning* 3D multi-ocular [C3D/Edinburgh Virtual Environment Centre].



Figura 1.22: O VIVID 700 (da Minolta Corporation) – um sistema de *scanning* 3D “móvel-fixo” [Minolta Corporation 1997].

pela empresa como sendo “o primeiro digitalizador 3D verdadeiramente portátil do mercado”, tornando “fácil a digitalização em qualquer local” [Minolta Corporation 1997] (figura 1.22). Porém, este sistema é, na verdade, portátil até ao ponto em que é de pequenas dimensões e portanto fácil de transportar — como está calibrado para distâncias medidas relativamente ao referencial da câmara necessitará sempre, no entanto, de ser usado conjuntamente com uma mesa rotativa e técnicas de registo 3D, como as apresentadas em [Huber 2000].

Também apresentado como sendo portátil mas sofrendo exactamente do mesmo mal que o sistema anteriormente apresentado temos o OPTO 3D-Ranger da empresa italiana OPTONET, um *scanner* 3D usando projecção de luz estruturada em sequências de listas segundo códigos de Gray. O sistema consiste num tripé com uma câmara monocromática ou a cores e uma placa de captura de vídeo especializada para PC em termos de hardware, funcionando em conjunto com o software de registo 3D (que requer uma estimativa inicial dada manualmente pelo operador) PolyWorks/Modeler da InnovMetric Inc. (Quebec Canada) — veja-se a figura 1.23 [Optonet].

Um passo em direcção ao exemplo de um verdadeiro scanner 3D portátil será o ModelMaker da 3D Scanners (figura 1.24) — este sistema consiste num *scanner* laser montado num braço mecânico portátil, necessário para controlar melhor as variações de *pose* [3D Scanners].

Outro exemplo de um sistema portátil, é o HLS (“*Handheld Laser Scanner*”) da New Zealand’s Applied Research Associates, comercializado como sendo o Polhemus FastSCAN, representado na figura 1.25. Este sistema consiste em duas câmaras e

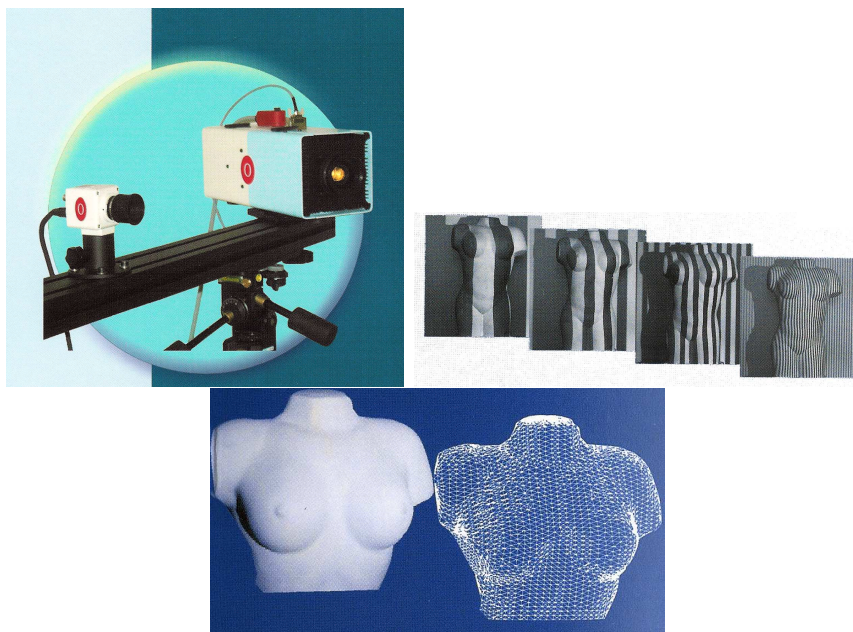


Figura 1.23: OPTO 3D-Ranger (da OPTONET) – outro sistema de *scanning* 3D “móvel-fixo” [Optonet]. Permissão para reprodução destas imagens gentilmente cedida pela Optonet Srl. De cima para baixo, da esquerda para a direita, pode ver-se uma perspectiva do sistema, um exemplo de uma sequência de projecções em código de Gray e os resultados obtidos para esse exemplo.



Figura 1.24: O ModelMaker (da 3D Scanners) – um sistema de recuperação 3D montado num braço mecânico portátil [3D Scanners].



Figura 1.25: O HLS/Polhemus FastSCAN (da ARANZ/Polhemus) – um sistema de scanning 3D portátil transportado manualmente (“*handheld*”) [Polhemus/ARANZ 2000].

um emissor magnético montados numa espécie de varinha, e um receptor que pode ser apenas a qualquer objecto a ser recuperado (o subsistema tipo emissor-receptor magnético é designado de Polhemus FASTRAK) — este digitalizador foi concebido para recuperar a estrutura tridimensional de objectos opacos e não-metálicos [Polhemus/ARANZ 2000].

Um último exemplo de um *scanner* portátil será o sistema de recuperação de estruturas tridimensionais ultrassónico da Universidade de Cambridge, o Stradx. Este sistema usa, para a recuperação 3D, um sistema emissor-receptor magnético (como por exemplo o Polhemus FASTRAK, o Ascension Bird ou o NDI POLARIS) e, para a reconstrução tridimensional, vários tipos de algoritmos volumétricos — veja-se a figura 1.26 [Gee and Prager 1999].

1.5.3 Comparações entre sistemas

Os sistemas de *scanning* 3D precedentes foram organizados na tabela 1.2 para comparação.

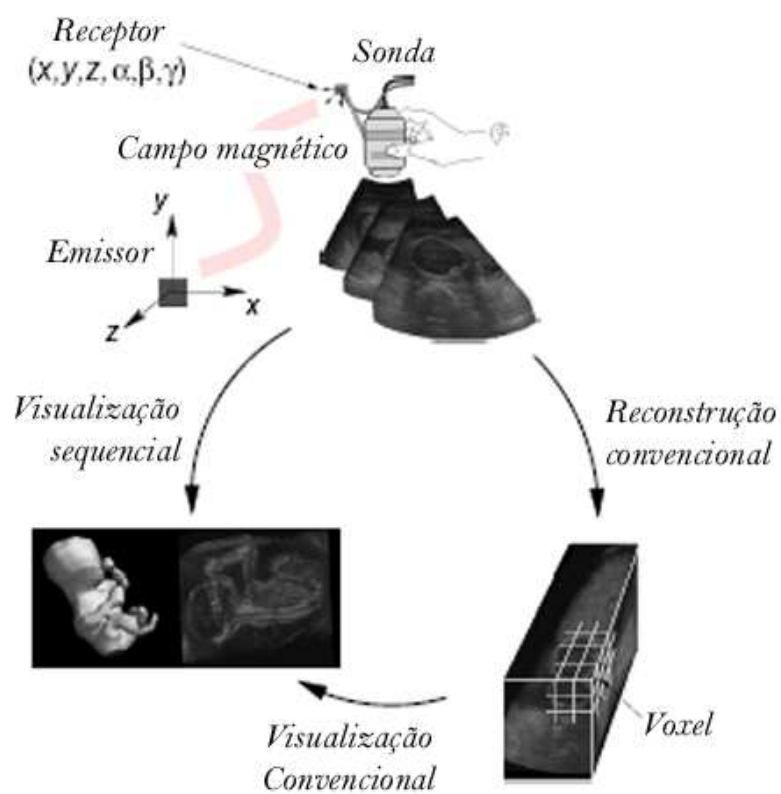


Figura 1.26: O Stradx (da Universidade de Cambridge) – um sistema ultrassónico de scanning 3D portátil. Adaptado de Gee and Prager [1999] com a gentil permissão do Doutor Richard Prager.

Tabela 1.2: Exemplos de *scanners* 3D – características disponibilizadas pelos fabricantes. Todas as medidas em milímetros excepto quando especificado. PLE - Projectão de Luz Estruturada; MO - Multi-Ocular; SP - Sensor de *Pose*; PL - Perfil (Plano) Laser; 3C - 3 Comprimentos de onda; LB - Luz Branca; CG - Sequência de padrões em Código Gray; E - Espelhos; SM - Sensor(es) Magnético(s); US - Ultrasons; N/D - Informação Não Disponível.

Modelo	Fabricante	Portabilidade	Tecnologia	Volume 3D Máx.	Resolução	Precisão
<i>REPLICA</i>	3D Scanners	Fixo	PLE, PL	N/D	.05	N/D
<i>Arius3D</i>	Arius3D	Fixo	PLE, PL (3C)	N/D	N/D	N/D
<i>TriForm</i>	3D Scan. Syst.	Fixo	PLE e MO, LB e E	Variável*	Variável*	Variável*
<i>C3D-2030</i>	C3D	Fixo	MO, Unid. Flash	200 × 260 × 150	.3	.5 RMS
<i>3030RGB/PS</i>	Cyberware	Fixo	PLE, PL	1000 × 340 × 300	De .1 a 1	N/D
<i>3030RGB/MS</i>	Cyberware	Fixo	PLE, PL	Modelos automóveis a 1/5	De .1 a 1	N/D
<i>VIVID 700</i>	Minolta	Móvel-Fixo	PLE, PL	1100 × 1100 × 2500	.35/.35/2.3	N/D
<i>OPTO 3D-Ranger</i>	OPTONET	Móvel-Fixo	PLE, LB e CG	300 × 225 × 200	.12/.12/.01	De .018 a .085
<i>ModelMaker</i>	3D Scanners	Braço Portátil	PLE, PL	Variável (Portátil)**	N/D	N/D
<i>FastSCAN</i>	Polhemus	Handheld	MO, PLE e SP, SM	Variável (Portátil)***	De .5 a 1	1 @ 200
<i>Stradx</i>	Univ. Cambr.	Handheld	US e SP, SM	Variável (Portátil)***	N/D	N/D

* Dependente da versão.

** Limitado pelo alcance do braço.

*** Limitados pelo alcance do sensor de *pose*, cerca de 760 mm.

1.6 Discussão: porque é que o *scanning* 3D é difícil?

No 1º Simpósio Internacional em Processamento, Visualização e Transmissão 3D, o Professor Marc Levoy apresentou uma “*keynote*” interessantíssima, sob o nome de “*Why Is 3D Scanning Hard?*”, em que abordava de um ponto de vista totalmente diferente o tema da digitalização tridimensional [Levoy 2002].

Olhando em retrospectiva para o projecto que orientou, apresentou as dificuldades que a sua equipa encontrou no Digital Michelangelo Project e no Digital *Forma Urbis Romae* Project, resumindo-as numa lista de 8 grandes problemas que, apesar de vistos sob a óptica dos projectos em causa, afectam qualquer sistema de *scanning* 3D baseado em visão [Levoy 2002]:

1. Materiais “não cooperantes” opticamente (figura 1.27 na página ao lado);
2. Digitalização na presença de oclusões (figura 1.28 na página 44);
3. Garantia da segurança de objectos delicados;
4. Digitalização de cenas grandes com grande resolução;
5. Digitalização *in loco* com boa precisão;
6. Preenchimento de lacunas (geradas, muito provavelmente, devido aos 2 primeiros problemas) em modelos poligonais densos (figura 1.28);
7. Tratamento de grandes conjuntos de dados de informação tridimensional;
8. Criação de arquivos digitais de informação tridimensional.

É previsível que o sistema que vier a resolver melhor o maior número destes problemas e outros já referidos ao longo deste texto será um caso de sucesso no mundo do *scanning* tridimensional, pelo menos no que se refere a sistemas baseados em visão por computador.

De qualquer das formas, torna-se também óbvio, tendo em conta tudo o que se disse até aqui, que os problemas mais directamente relacionados com a engenharia do sistema de digitalização em si serão os problemas 1, 2, 3 e 5. Serão estas, provavelmente, as questões-chave cuja resolução implicará uma melhoria substancial no desempenho final do sistema; seria, pois, desejável o desenvolvimento de um sistema que resolvesse estes 4 problemas, sendo simultaneamente simples e barato.

Por outro lado, também é desejável que seja fácil de preparar e manusear, e também que consiga resolver o problema da digitalização tridimensional de forma satisfatória para a gama mais alargada de volumes de tridimensionais.

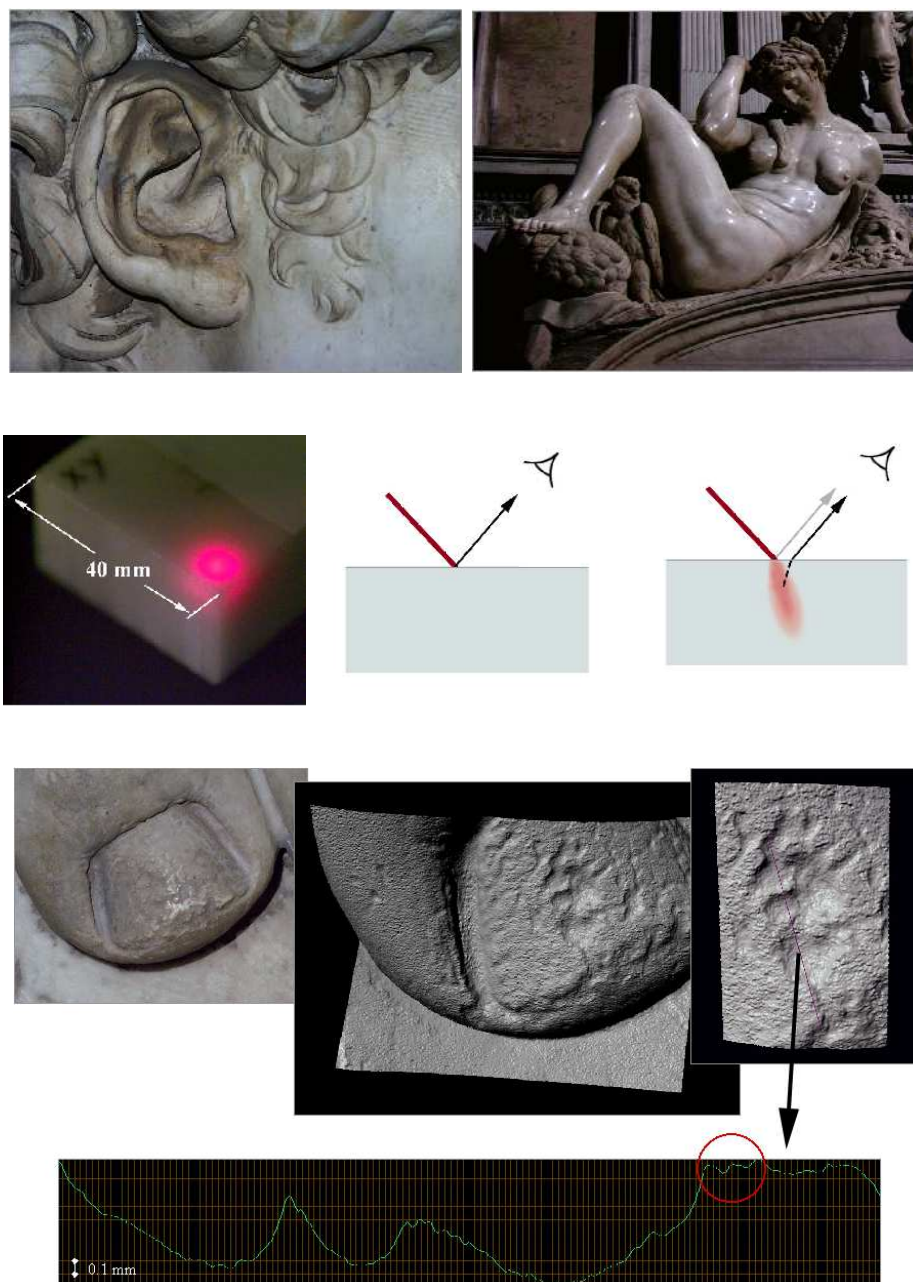


Figura 1.27: 1º problema de 3D *scanning*: materiais “não cooperantes” opticamente – tirado de Levoy [2002] com a gentil permissão do Professor Marc Levoy. Da esquerda para a direita, de cima para baixo: exemplo de mármore escuro na estátua de David, exemplo de mármore polido noutra estátua, imagem do efeito de espalhamento nalguns mármore e esquematizações comparativas de uma superfície lambertiana e do referido espalhamento, erros de reconstrução devido a reflexões especulares no dedo do pé da estátua de David.



Figura 1.28: 2º e 6º problemas de 3D *scanning*: oclusões e lacunas – tirado de Levoy [2002] com a gentil permissão do Professor Marc Levoy. Na figura apresenta-se um problema de oclusão no pé da estátua de David que por sua vez gerou uma lacuna.

É possível verificar no estudo efectuado neste texto que nenhum sistema conseguiu ainda atacar decisivamente estas questões. Os sistemas existentes resolvem apenas uma parte pequena dessas questões, sendo soluções concebidas “*à medida*” de cada *tipo de cena tridimensional*.

1.7 Objectivos e estruturação do projecto Tele-3D

Neste capítulo foram descritas várias técnicas e tecnologias para sistemas de recuperação e reconstrução de cenas tridimensionais. Vários exemplos destes sistemas foram também apresentados, tendo sido efectuadas análises comparativas tanto das técnicas e tecnologias como das suas implementações práticas.

Foi também feita uma breve discussão sobre alguns dos problemas que afectam os sistemas de digitalização tridimensional.

Uma primeira conclusão a que se pode chegar analisando o que foi dito até aqui será que a escolha de um *scanner* 3D depende das técnicas e das tecnologias que este usa no modo como estas abordam as seguintes questões [Hindus 2001]:

- **Tamanho** — Qual o tamanho da/dos cena/objectos a recuperar?
- **Precisão** — Que grau de precisão (baixo, alto ou de engenharia) é requerido?
- **Qualidade visual** — Será a cor necessária e querer-se-ão reconstruções fotorealistas?

Outra conclusão evidente será a de que o futuro dos digitalizadores 3D, representado por projectos como o “3D Faxing Machine” ou o “Digital Michelangelo”, está nos

sistemas portáteis multi-uso. Quem quer que ofereça a solução mais simples, barata e genérica, desde que não perca de vista a qualidade do sistema, terá vantagem sobre os restantes.

Porém, ainda muita investigação neste campo está para vir, juntamente com muitos mais avanços entusiasmantes — seja como for, a digitalização 3D *já é* uma revolução do presente.

Por todas estas razões, pretende-se com o projecto **Tele-3D** explorar os limites das problemáticas de *precisão vs flexibilidade* e de *qualidade e volume de trabalho vs eficiência* de um sistema de *scanning* 3D portátil *handheld*, baseado na técnica laser *rangefinder* por triangulação, com uma câmara, um projector de um plano de laser e um sensor de medição de *pose*, montados numa estrutura rígida.

Serão para tal estudados e testados numa primeira fase modelos e configurações para o sistema e algoritmos de calibração eficientes e flexíveis, sendo posteriormente estudada laboratorialmente a viabilidade de uma abordagem inovadora de uso de sensores de medição de *pose* conjuntamente com algoritmos eficientes de registo tridimensional de forma a explorar a geometria dos modelos adoptados, de forma a obter níveis de resolução e precisão satisfatórios.

1.8 Estruturação da dissertação

Começou-se neste capítulo por proporcionar uma introdução às técnicas e tecnologias correntemente utilizadas em sistemas de recuperação e reconstrução tridimensional. Estudaram-se sistemas práticos, usando tanto tecnologias mais comuns como mais modernas nesta área, de modo a demonstrar as suas vantagens e desvantagens. O seu propósito foi apontar caminhos futuros de investigação de modo a estabelecer também os objectivos e o contributo deste trabalho.

No Capítulo 2, far-se-á a apresentação dos modelos que regem o *scanner* Tele-3D, um estudo de sensibilidade ao erro desses modelos e a descrição dos algoritmos desenvolvidos para a calibração do sistema. Segue-se a apresentação dos resultados experimentais do trabalho laboratorial efectuado para a avaliação da qualidade do modelo e dos algoritmos utilizados.

No Capítulo 3, apresentar-se-á o processamento envolvido nos procedimentos de varrimento da cena de forma a obter-se uma primeira aproximação da reconstrução tridimensional. Será seguidamente apresentado um novo procedimento de registo de perfis tridimensionais que conjuga a utilização do sistema de medição de *pose* com

um algoritmo inovador que toma em consideração a geometria inerente ao sistema de triangulação usando a projecção de um plano de luz. Será nomeadamente testado o seu desempenho no sentido de verificar se são obtidos níveis de resolução e precisão satisfatórios.

Por fim, no Capítulo 4 tecer-se-ão as conclusões necessárias e propor-se-á trabalho futuro.

Capítulo 2

O Sistema Tele-3D — Modelos e Calibração

Tendo em vista a importância da última premissa definida no final do capítulo anterior para a comparação entre os desempenhos dos vários sistemas de digitalização tridimensional, o grau de *precisão* requerido, torna-se importante efectuar uma *análise ao erro* de forma a poder determinar a precisão mínima que se consegue nas digitalizações tridimensionais usando o sistema Tele-3D.

Por esta razão, é objectivo das primeiras secções deste capítulo definir os *modelos e princípios de funcionamento* do *scanner* Tele-3D e apresentar os fundamentos teóricos e os resultados práticos do *estudo de sensibilidade ao erro* do modelo de triangulação do sistema.

Porém, é evidente que, na prática, este estudo de sensibilidade ao erro fornece apenas um *limite inferior* para o erro médio previsto para o sistema, visto que não é possível determinar-se o valor exacto dos parâmetros que regem os modelos. Com efeito, os erros que corrompem inevitavelmente as medidas fornecidas pelos sensores influenciam também a qualidade da *calibração* do sistema. É necessário, pois, não só desenvolver procedimentos de calibração o mais robustos possível, como analisar essa robustez e a sua influência na precisão final do *scanner*.

Apesar de ser facto assente que a calibração, além de ser um desafio crucial nos sistemas de digitalização tridimensional, tem uma influência determinante no seu desempenho ao nível da precisão de medida (veja-se, por exemplo [DePiero and Trivedi 1996] e [Prager et al. 1997]), não tem sido feito, segundo a opinião deste autor, um trabalho suficiente abrangente para se afirmar que este assunto se encontra definitivamente arrumado. Para complicar significativamente esta questão, surgiu agora a necessidade de desenvolver *scanners* 3D portáteis, tecnologia na qual se enquadra o

scanner Tele-3D, que trouxe consigo abordagens e problemas completamente novos (consulte-se, por exemplo [McCallum et al. 1996]).

Por outro lado, a área da digitalização tridimensional baseada em imagens captadas por câmaras tenderá a ser impulsionada pelos algoritmos de calibração de câmaras robustos recentemente desenvolvidos, como por exemplo os procedimentos apresentados em [Zhang 1999] e [Heikkilä and Silven 1997].

A maioria dos estudos sobre calibração, extremamente dependentes de restrições severas nos seus modelos geométricos, como é o caso de [McIvor 1999], será brevemente substituída por soluções mais flexíveis e de mais simples execução, como por exemplo [Davis and Chen 2001], e soluções de grande grau de automatização, como é [Trucco et al. 1994].

Deste modo, no final deste capítulo apresentar-se-á uma solução que usa estas novas tendências no desenvolvimento de uma técnica de calibração que tenta satisfazer, tanto requisitos de flexibilidade e facilidade de utilização, como de precisão e desempenho final do sistema. Far-se-á primeiramente uma apresentação teórica dessa solução, seguindo-se a análise a nível experimental do seu desempenho.

2.1 Princípio de funcionamento do *scanner* Tele-3D

O *scanner* Tele-3D é um sistema de digitalização tridimensional portátil cujo princípio de funcionamento se baseia na fusão de medidas efectuadas por dois sistemas distintos: um sistema de digitalização tridimensional propriamente dito, que num dado instante efectua uma amostragem parcial de uma cena ou objecto cuja estrutura se quer recuperar tridimensionalmente, afectando as suas medidas a um referencial dependente da posição e orientação do *scanner*, e um sistema de medição de “*pose*” completa ou parcial, que torna possível a integração de todas as amostragens parciais num conjunto de medidas global afectado a um referencial único, independente das várias posições que o *scanner* tomou ao longo do varrimento da cena.

Na figura 2.1 está representada a arquitectura do *scanner*. Uma estrutura rígida com forma semelhante a um *boomerang* tem uma câmara fixada numa das pontas e um projector laser na outra, formando o sistema de digitalização tridimensional; o sensor de que consiste o sistema de medição de “*pose*” encontra-se acoplado à câmara. Uma pega está colocada no centro geométrico vertical da estrutura de modo a proporcionar a portabilidade manual requerida. Também na figura 2.1 estão representados todos os referenciais directa ou indirectamente envolvidos no funcionamento do scanner: o

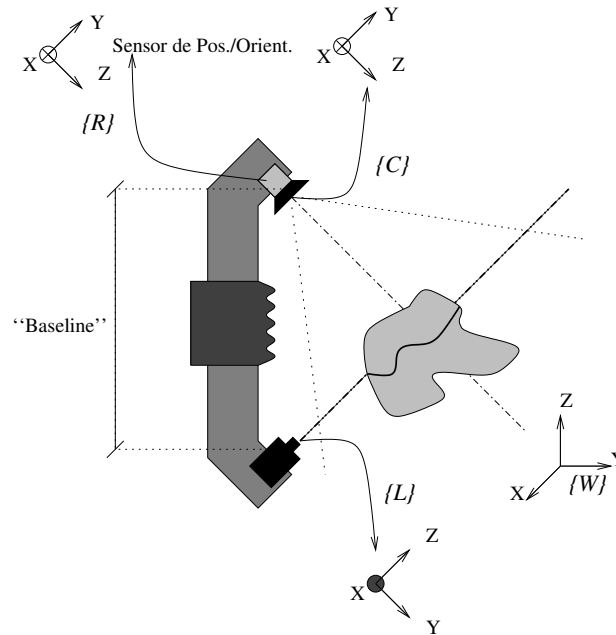


Figura 2.1: Esquema da arquitectura do scanner Tele-3D.

referencial da câmara, $\{C\}$, o referencial do projector laser, $\{L\}$, o referencial do sensor de medição de “pose”, $\{R\}$, e o referencial global final $\{W\}$.

Nas secções subsequentes far-se-á uma descrição exaustiva dos modelos geométricos inerentes a estes sistemas.

2.1.1 Modelo do sistema de triangulação tridimensional

O sistema de digitalização tridimensional do scanner Tele-3D é composto por dois dispositivos essenciais — um projector laser e uma câmara. A sua configuração geométrica, que não é mais do que a configuração fixa de um sistema laser *rangefinder* por triangulação, vai permitir a determinação das coordenadas tridimensionais de pontos na cena a recuperar. Nas subsecções que seguem, far-se-á uma exposição dessa configuração.

Modelo genérico da câmara

Qualquer ponto 3D genérico $\mathbf{P} = [X \ Y \ Z]^T$ e a sua projecção correspondente $\mathbf{p} = [u \ v]^T$ num plano-imagem podem ser relacionados matematicamente usando geometria projectiva e o conceito de coordenadas homogéneas através da equação seguinte, a *relação genérica da câmara projectiva*, onde s representa um factor de escala que afecta

globalmente essa relação [Dias 1994, Faugeras 1993, Hartley and Zisserman 2000],

$$\begin{aligned}
 \begin{bmatrix} su \\ sv \\ s \end{bmatrix} &= \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\
 &= \mathbf{A}_{3 \times 4} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{a}_1^T \mathbf{P} + a_{1,4} \\ \mathbf{a}_2^T \mathbf{P} + a_{2,4} \\ \mathbf{a}_3^T \mathbf{P} + a_{3,4} \end{bmatrix}
 \end{aligned} \tag{2.1}$$

sendo $\mathbf{a}_1 = [a_{1,1} \ a_{1,2} \ a_{1,3}]^T$, $\mathbf{a}_2 = [a_{2,1} \ a_{2,2} \ a_{2,3}]^T$ e $\mathbf{a}_3 = [a_{3,1} \ a_{3,2} \ a_{3,3}]^T$.

À matriz $\mathbf{A}$ dá-se o nome de *matriz de projecção*, e através do seu cálculo torna-se possível determinar a correspondência entre qualquer ponto 3D e a sua projecção no plano-imagem [Ferreira and Dias 2000].

A partir da equação da câmara projectiva podem deduzir-se duas equações homogéneas linearmente independentes relativamente a $\mathbf{P}' = [X \ Y \ Z \ 1]^T$ (ponto $\mathbf{P}$ em coordenadas homogéneas) [Ferreira et al. 2001, Dias 1994, Hartley and Zisserman 2000]. Para tal, elimina-se o factor de escala através do produto externo $\mathbf{p}' \times \mathbf{A}\mathbf{P}' = \mathbf{0}$ (com $\mathbf{p}' = [su \ sv \ s]^T$), obtendo-se então

$$\begin{cases} (\mathbf{a}_1 - u\mathbf{a}_3)^T \mathbf{P} + a_{1,4} - u = 0 \\ (\mathbf{a}_2 - v\mathbf{a}_3)^T \mathbf{P} + a_{2,4} - v = 0 \end{cases} \Leftrightarrow \begin{cases} \Pi_1 \mathbf{P}' = 0 \\ \Pi_2 \mathbf{P}' = 0 \end{cases} \tag{2.2}$$

Torna-se claro por esta expressão que estas equações representam, de facto, a solução da intersecção do plano Π_1 com Π_2 . A essa solução dá-se o nome de *recta de projecção* ou *recta projectante*, que pode ser apresentada de forma mais clara como uma única expressão de produto externo, dada pela equação (2.3) [Ferreira et al. 2001, Dias 1994].

$$\vec{\mathbf{n}} = (\mathbf{a}_1 - u\mathbf{a}_3) \times (\mathbf{a}_2 - v\mathbf{a}_3) \tag{2.3}$$

Esta equação indica que *todos os pontos 3D na recta projectante correspondem ao mesmo ponto projectado no plano-imagem*, o que significa que a *relação de projecção não é unívoca*.

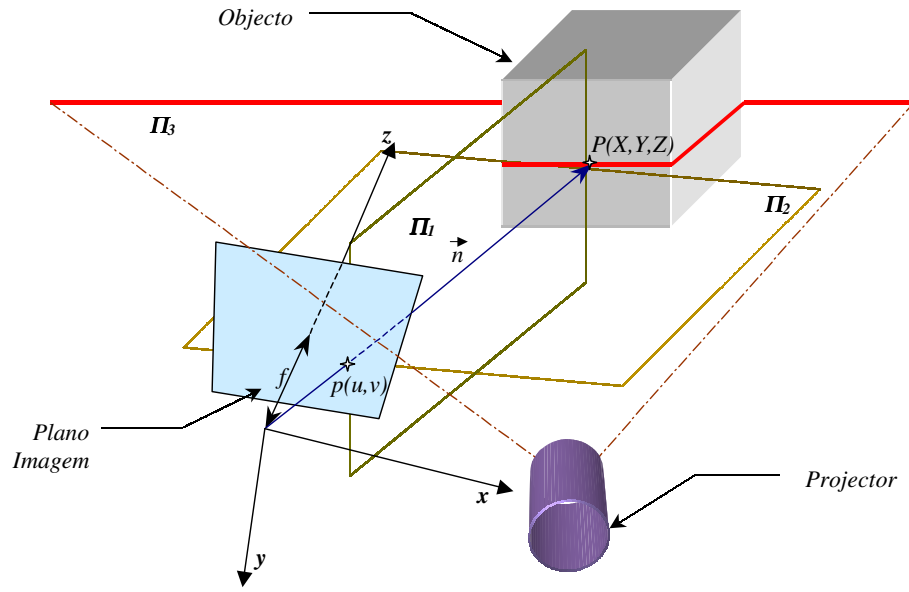


Figura 2.2: Esquema para a geometria de uma cena com projecção da luz estruturada capturada por uma câmara [Ferreira and Dias 2000].

Introdução da restrição geométrica adicional necessária — o plano laser

A geometria inerente a um sistema de recuperação tridimensional a partir da projecção luz estruturada, incluindo o modelo da câmara projectiva genérica, está representada na figura 2.2 na configuração mais habitualmente utilizada de *light-striping* [Dudek and Jenkin 2000a].

Torna-se claro nesta figura que a fonte de luz está de facto a projectar um plano de luz, que pode ser representado matematicamente pela equação (2.4).

$$\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = 0 \Leftrightarrow \Pi_3 \mathbf{P}' = 0 \quad (2.4)$$

Usando a equação (2.2) e a restrição do plano de luz dada pela equação (2.4), o ponto 3D na cena pode ser relacionado *de forma única* com o seu ponto de projecção bidimensional no plano-imagem da câmara por triangulação usando [Ferreira and Dias

2000, Dudek and Jenkin 2000c]:

$$\begin{cases} \Pi_1 \mathbf{P}' = 0 \\ \Pi_2 \mathbf{P}' = 0 \\ \Pi_3 \mathbf{P}' = 0 \end{cases} \quad (2.5)$$

Particularizando o modelo da câmara

Para controlar os efeitos que a câmara real utilizada no sistema poderá ter no processo de digitalização tridimensional, torna-se desejável refinar um pouco mais o seu modelo de forma a poder incluir parâmetros ajustáveis e até, se possível, melhorar o desempenho dos algoritmos de calibração do sistema. Para esse efeito, é importante notar que qualquer câmara é de facto um equipamento de funcionamento euclidiano sendo, portanto, possível basear-se o seu modelo em parâmetros de interpretação euclidiana e, consequentemente, com significado claro nesse tipo de geometria [Hartley and Zisserman 2000]. Nos parágrafos subsequentes far-se-á, então, uma decomposição analítica da matriz de projecção $\mathbf{A}$ através de um estudo mais aprofundado dos princípios de funcionamento de câmaras reais.

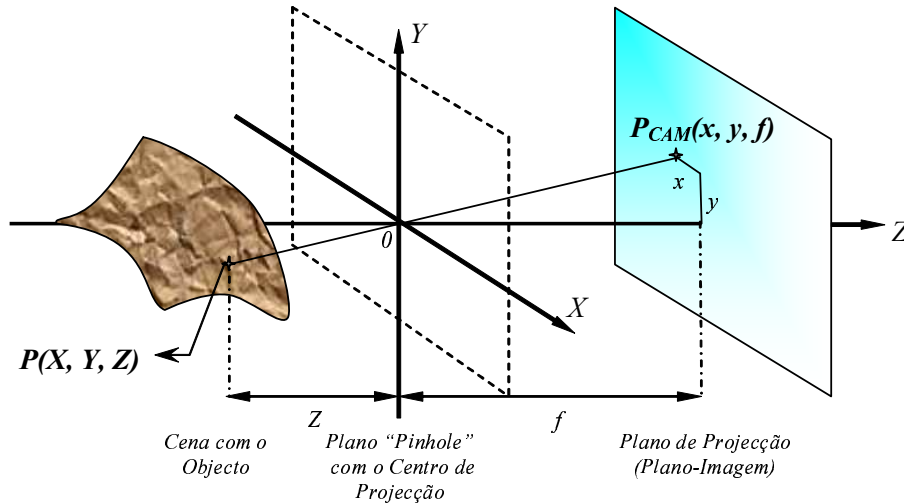


Figura 2.3: Modelo da câmara “pinhole”.

Na óptica e da fotogrametria¹ é frequentemente usada a simplificação das lentes finas para modelar um sistema de captura de imagens. Esta simplificação levou ao

¹Estudo da determinação das dimensões de objectos e de execução de levantamentos topográficos por meio de fotografias. Do grego *phōs*, *photós*, “luz” + *grama*, “inscrição” + *métrōn*, “medida” + *ia* [Porto Editora 2002].

desenvolvimento do chamado modelo “*pinhole*” da câmara, baseada no funcionamento da antecessora da câmara fotográfica actual, a *câmara estenotopaica*, que forçava os raios luminosos a passar por um orifício muito pequeno, o “*pinhole*” [Porto Editora 2002]. A geometria deste modelo encontra-se representada na figura 2.3 (f representa a distância focal).

Como se pode verificar na figura, este modelo implica que a imagem do mundo capturada pela câmara seja invertida; como, à semelhança do que ocorre no sistema de visão humana, esta imagem é re-invertida antes do seu processamento, habitualmente utiliza-se em sua substituição o modelo representado na figura 2.4.

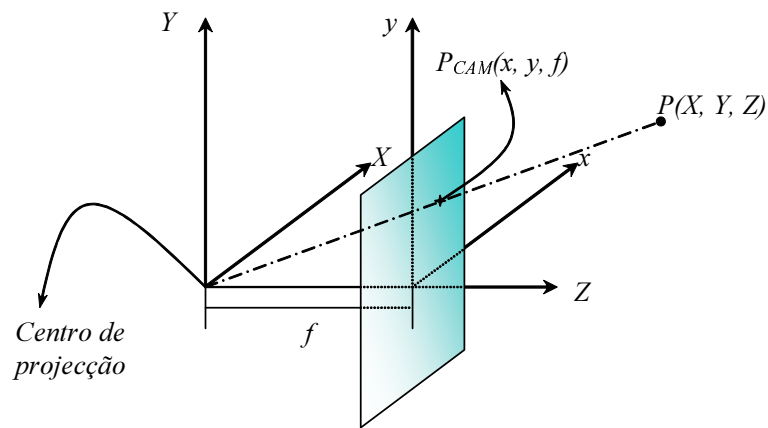


Figura 2.4: Projecção em perspectiva.

A geometria aí representada permite retirar facilmente a expressão para o fenómeno a que se dá o nome de *projecção em perspectiva*, dada pela equação (2.6) [Faugeras 1993, Hartley and Zisserman 2000].

$$\begin{cases} x = f \frac{X}{Z} \\ y = f \frac{Y}{Z} \end{cases} \quad (2.6)$$

Confirma-se neste raciocínio o facto de qualquer ponto na recta projectante ter o mesmo ponto de projecção no plano-imagem.

É de todo o interesse ter a projecção em perspectiva referida, não relativamente a coordenadas (x, y) no plano-imagem físico da câmara, mas relativamente a coordenadas (u, v) na imagem já digitalizada. É, por isso, essencial ter em conta os efeitos da amostragem, representados na figura 2.5.

Notando que poderá existir um deslocamento do centro do plano-imagem (o chamado *ponto principal*) O de forma a que este não coincida com a origem na imagem

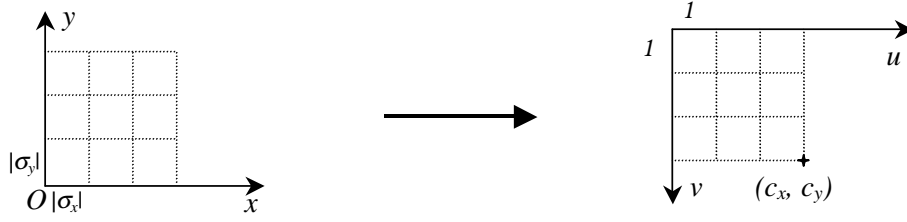


Figura 2.5: Efeitos da amostragem de uma câmara CCD.

digitalizada, tendo nesta coordenadas (c_x, c_y) , e denotando a razão entre factores de escala horizontais e verticais no plano-imagem físico e no plano-imagem amostrado por k_{CCD}/k_{im} e l_{CCD}/l_{im} , respectivamente, chegamos à seguinte expressão apresentada na equação (2.7) [Hartley and Zisserman 2000, Faugeras 1993].

$$\begin{cases} x = \frac{k_{CCD}}{k_{im}} (u - c_x) \\ y = -\frac{l_{CCD}}{l_{im}} (v - c_y) \end{cases} \Leftrightarrow \begin{cases} x = \sigma_x (u - c_x) \\ y = \sigma_y (v - c_y) \end{cases} \quad (2.7)$$

Substituindo, finalmente, a equação (2.7) na equação de projecção em perspectiva (equação (2.6)), obtém-se

$$\begin{cases} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sigma_x (u - c_x) \\ \sigma_y (v - c_y) \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \frac{X_C}{Z_C} \\ f \frac{Y_C}{Z_C} \end{bmatrix} \end{cases} \Leftrightarrow \begin{cases} u = \frac{f \frac{X_C}{Z_C} + \sigma_x c_x}{\sigma_x} \\ v = \frac{f \frac{Y_C}{Z_C} + \sigma_y c_y}{\sigma_y} \end{cases} \Leftrightarrow \begin{cases} u = \frac{f}{\sigma_x} \frac{X_C}{Z_C} + c_x \\ v = \frac{f}{\sigma_y} \frac{Y_C}{Z_C} + c_y \end{cases} \quad (2.8)$$

Normalizando f relativamente ao comprimento, denominando esse valor de $f_x = f/\sigma_x$, e largura, denominando esse valor de $f_y = f/\sigma_y$, dos pixel na imagem, temos, finalmente

$$\begin{cases} u = f_x \frac{X}{Z} + c_x \\ v = f_y \frac{Y}{Z} + c_y \end{cases} \Leftrightarrow \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2.9)$$

Este modelo é geralmente suficiente para qualquer câmara CCD, visto que estas geralmente têm células foto-receptoras perfeitamente rectangulares².

²Isto é, podem-se associar a sistemas de eixos ortogonais — consulte-se a interessante discussão sobre este assunto apresentada em Hartley and Zisserman [2000], página 151.

Se se quiser, no entanto, usar um sistema de coordenadas tridimensional genérico em vez do apresentado na figura 2.4, poder-se-á efectuar uma transformação rígida de forma a efectuar a mudança de sistema de eixos multiplicando previamente as coordenadas por uma matriz 4×4 de forma a obter a expressão final

$$\mathbf{A}_{3 \times 4} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_X \\ r_{2,1} & r_{2,2} & r_{2,3} & t_Y \\ r_{3,1} & r_{3,2} & r_{3,3} & t_Z \end{bmatrix} \quad (2.10)$$

Um caso particular de grande interesse de aplicação desta expressão é quando o referencial tridimensional escolhido está associado ao projector laser; neste caso, pode-se integrar nela o “*baseline*” entre a câmara e o projector (ver figura 2.1, página 49), cujo comprimento se poderá denotar como B , nas componentes da translação, recorrendo a coordenadas esféricas:

$$\mathbf{A}_{3 \times 4} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & B \cos \theta \sin \phi \\ r_{2,1} & r_{2,2} & r_{2,3} & B \sin \theta \sin \phi \\ r_{3,1} & r_{3,2} & r_{3,3} & B \cos \phi \end{bmatrix} \quad (2.11)$$

A expressão (2.10) dá-nos o modelo final para a câmara, denominada de *câmara “pinhole” completa* ou *câmara projectiva finita*, com 10 graus de liberdade [Hartley and Zisserman 2000]. Reformulando esta expressão, obtém-se

$$\mathbf{A}_{3 \times 4} = \mathbf{C}_{3 \times 3} \begin{bmatrix} {}^C \mathbf{R}_W & {}^C \vec{\mathbf{t}}_W \end{bmatrix} \quad (2.12)$$

onde $\mathbf{C}_{3 \times 3}$ é a *matriz de calibração* da câmara, também denominada de *matriz de parâmetros intrínsecos* visto ser independente da colocação e orientação da câmara no mundo, e a segunda matriz é uma matriz estendida contendo os *parâmetros extrínsecos* da câmara, que transforma o sistema de coordenadas do mundo no sistema de coordenadas da câmara [Ferreira and Dias 2000, Ferreira et al. 2001, Hartley and Zisserman 2000, Faugeras 1993, Zhang 2000].

O processo de triangulação

Como já foi dito na página 52, usa-se a triangulação através do recurso à equação (2.5) para determinar univocamente as coordenadas de um ponto tridimensional a partir da sua projecção, amostrada na imagem como pertencendo ao perfil iluminado pela lista *laser*.

A resolução da equação (2.5), ou, dito de outro modo, a determinação do ponto de intersecção entre os planos Π_1 , Π_2 e Π_3 , é conseguida através de álgebra vectorial, nomeadamente das propriedades dos determinantes, aplicada à geometria projectiva em $\mathbb{P}_3$ [Hartley and Zisserman 2000, Farin 1999, Vitória and de Lima 1998].

Seja $\mathbf{M} = [\Pi, \Pi_1, \Pi_2, \Pi_3]$, uma matriz composta por um plano tridimensional genérico Π e pelos três planos Π_i ($i = 1, \dots, 3$) que definem o ponto $\mathbf{P}'$. Verifica-se $\det \mathbf{M} = 0$ quando Π contém $\mathbf{P}'$, visto que nessas condições o plano Π pode ser expresso como combinação linear dos planos Π_i . Expandindo o determinante relativamente à coluna Π , obtém-se

$$\det \mathbf{M} = aD_{234} - bD_{134} + cD_{124} - dD_{123} \quad (2.13)$$

onde D_{jkl} é o determinante das matrizes formadas pelas linhas jkl da matriz 4×3 $[\Pi_1, \Pi_2, \Pi_3]$. Visto que $\det \mathbf{M} = 0$, é possível determinar a partir da equação anterior as coordenadas do ponto tridimensional amostrado como sendo

$$\mathbf{P}' = [D_{234}, -D_{134}, D_{124}, -D_{123}]^T \quad (2.14)$$

que é, efectivamente, o vector de solução (o espaço-nulo) da equação (2.5).

2.1.2 Modelo do sistema de medição de “pose”

O modelo geométrico para o sistema de medição de “pose” é simples e pode ser retirado directamente da figura 2.1 — usando as medições obtidas por esse sensor, é possível determinar a posição e orientação do referencial que lhe está associado, $\{R\}$, relativamente a $\{T\}$, sendo este último o sistema de coordenadas de referência do emissor, quando o sensor de medição de “pose” é do tipo emissor-receptor, ou o referencial de inicialização, quando o sensor é auto-contido [Ferreira et al. 2002b].

Mais concretamente, o sensor mede as seguintes relações entre os referenciais (pelo menos a primeira no caso do sensor ser auto-contido):

- 1º Os ângulos de três rotações no sentido inverso (sentido dos ponteiros do relógio) em cada eixo ortogonal que alinham o segundo referencial com o primeiro, pela seguinte ordem: primeiro uma rotação de α em Z , seguido de uma rotação de β em Y e por fim de uma rotação γ em X . Esta medição está de acordo com a convenção de ângulos/eixos “Roll-Pitch-Yaw” (ou “XYZ”), apresentada na figura 2.6 na próxima página [Collinson 1996].
- 2º A posição da origem do referencial do sensor medido relativamente a $\{T\}$, (x, y, z) .

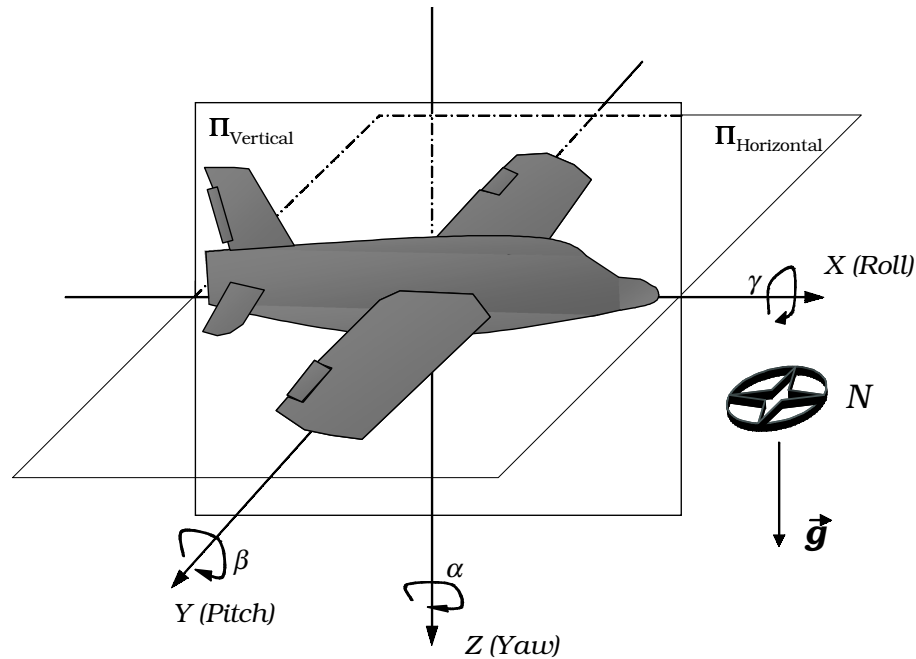


Figura 2.6: Definição dos eixos/ângulos de rotação “Roll-Pitch-Yaw”. Em aviação, geralmente a medição da inclinação longitudinal e latitudinal (“attitude”) e a orientação no plano horizontal (“heading”) é essencial, na medida em que permite a um piloto voar com o mínimo de segurança em quaisquer condições [Collinson 1996]. Esta convenção foi adoptada também pela maioria dos sensores de orientação em aplicações como as que são tratadas neste trabalho, visto que ela é baseada em tecnologia usada pelos seus congéneres de navegação aérea.

Alguns sensores estão apetrechados com *firmware* que processa estas medidas, calculando e disponibilizando inclusivamente a matriz de transformação correspondente ou até o quaternião correspondente à rotação composta como definido no ponto 1 acima descrito.

Desta forma, é possível determinar a transformação entre os referenciais através destas medidas, usando a seguinte convenção:

1º Uma rotação que alinha os eixos do primeiro referencial com o do segundo, composta por três rotações no sentido directo (contrário ao dos ponteiros do relógio) em cada eixo ortogonal, pela seguinte ordem: primeiro uma rotação de γ em X , seguido de uma rotação de β em Y e por fim de uma rotação α em Z .

2º Uma translação que leva a origem do primeiro referencial para o segundo.

As rotações em cada eixo referidas no primeiro ponto da convenção são dadas por

$$\mathbf{R}_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \cos(\gamma) \end{bmatrix} \quad (2.15a)$$

$$\mathbf{R}_Y = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \quad (2.15b)$$

$$\mathbf{R}_Z = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.15c)$$

e, portanto, como se pode constatar em [Prager et al. 1997], matematicamente a transformação em função das medidas do sensor é dada por

$${}^T\mathbf{T}_R = \begin{bmatrix} \mathbf{R}_Z \cdot \mathbf{R}_Y \cdot \mathbf{R}_X & \vec{\mathbf{r}} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & x \\ r_{2,1} & r_{2,2} & r_{2,3} & y \\ r_{3,1} & r_{3,2} & r_{3,3} & z \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\begin{aligned} r_{1,1} &= \cos(\alpha) \cos(\beta) \\ r_{1,2} &= \cos(\alpha) \sin(\beta) \sin(\gamma) - \sin(\alpha) \cos(\gamma) \\ r_{1,3} &= \cos(\alpha) \sin(\beta) \cos(\gamma) + \sin(\alpha) \sin(\gamma) \\ r_{2,1} &= \sin(\alpha) \cos(\beta) \\ r_{2,2} &= \sin(\alpha) \sin(\beta) \sin(\gamma) + \cos(\alpha) \cos(\gamma) \\ r_{2,3} &= \sin(\alpha) \sin(\beta) \cos(\gamma) - \cos(\alpha) \sin(\gamma) \\ r_{3,1} &= -\sin(\beta) \\ r_{3,2} &= \cos(\beta) \sin(\gamma) \\ r_{3,3} &= \cos(\beta) \cos(\gamma) \end{aligned} \quad (2.16)$$

2.2 Estudo de sensibilidade ao erro do sistema de triangulação tridimensional

Em seguida, discutir-se-á o método que permitirá estabelecer a relação causa-efeito entre mudanças nos parâmetros que regem o modelo do sistema de triangulação tridimensional e os erros cometidos na recuperação tridimensional, proceder-se-á à sua aplicação ao modelo genérico e finalmente far-se-á uma apresentação em *MATLAB* de resultados obtidos para a configuração particular do sistema Tele-3D.

2.2.1 Conceitos de sensibilidade ao erro e sua aplicação genérica ao modelo do sistema de triangulação tridimensional

Nesta secção, far-se-á uma apresentação aos conceitos de sensibilidade ao erro e aplicar-se-ão esses conceitos ao modelo geral de triangulação apresentado na secção 2.1.1.

Modelos sensoriais não lineares

Para se fazer um estudo de sensibilidade ao erro do modelo descrito na secção anterior para o sistema de triangulação tridimensional, é necessário em primeiro lugar determinar uma expressão que relacione as entradas do modelo com as suas saídas.

Para tal, defina-se $\boldsymbol{\theta} \in \mathbb{R}^m$ como o *vector de dados de medida*, com as m entradas do modelo, e $\mathbf{x} \in \mathbb{R}^n$ como o *vector de informação sensorial*, com as n saídas desse modelo. Desta forma, é possível representar $\mathbf{x}$ genericamente como uma função vectorial não linear como está representado na equação que se segue

$$\mathbf{x} = \mathbf{f}(\boldsymbol{\theta}) \quad (2.17)$$

onde se assume que $n \leq m$ [Nakamura and Xu 1989].

Mais concretamente, para o nosso modelo temos $\mathbf{x} \equiv \mathbf{P}$; recorrendo à equação (2.14) e ao conceito de coordenadas homogêneas, obtemos

$$\mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -\frac{D_{234}}{D_{123}} \\ \frac{D_{134}}{D_{123}} \\ -\frac{D_{124}}{D_{123}} \end{bmatrix} = \mathbf{f} \left(\begin{bmatrix} u \\ v \end{bmatrix} \right) \quad (2.18)$$

onde $\mathbf{f}$ é claramente não-linear.

Observando, no entanto, a geometria representada na figura 2.2 (página 51), torna-se óbvio que $\mathbf{x}$ corresponde a um ponto pertencente ao plano de luz de geometria fixa Π_3 . Desta forma, se se efectuar uma mudança de sistema de coordenadas de forma a que este corresponda ao referencial do projector laser (representado como $\{L\}$ na figura 2.1, página 49), obtemos para $\mathbf{A}_{3 \times 4}$ a expressão (2.11) e $\Pi_3 = 0$ transforma-se em $Y = 0$. Por esta razão, é possível simplificar a expressão anterior, obtendo-se

$$\mathbf{x} = \begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} -\frac{D_{234}}{D_{123}} \\ -\frac{D_{124}}{D_{123}} \end{bmatrix} = \mathbf{f} \left(\begin{bmatrix} u \\ v \end{bmatrix} \right) \quad (2.19)$$

Temos, deste modo, para o modelo do sistema de triangulação tridimensional, $n = 2$ e $m = 2$ como dimensões do vector de dados de medida e do vector de informação sensorial, respectivamente.

Noção de Diluição Geométrica de Precisão e matriz de covariância da informação sensorial

A expressão que permite relacionar a variação da informação sensorial com variações infinitesimais dos dados de medida é formalizada pela noção de *Diluição Geométrica de Precisão* (“*Geometric Dilution of Precision*”) que, como se pode ver na equação (2.20), é, por sua vez, igual ao Jacobiano de $\mathbf{f}$ [Dudek and Jenkin 2000d].

$$GDOP = \mathbf{J} = \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} \in \mathbb{R}^{n \times m} \quad (2.20)$$

Este, no caso do nosso modelo, é por sua vez dado por

$$\mathbf{J} = \begin{bmatrix} \frac{\partial X}{\partial u} & \frac{\partial X}{\partial v} \\ \frac{\partial Z}{\partial u} & \frac{\partial Z}{\partial v} \end{bmatrix} \quad (2.21)$$

A grande importância deste valor prende-se ao facto de ele ser uma medida da sensibilidade do sistema a erros nas medidas. Nos parágrafos que se seguem, demonstrar-se-á mais concretamente este facto.

O erro ou a incerteza incluídos nos dados de medida podem ser assumidos como sendo aditivos, sendo representados por

$$\boldsymbol{\theta} = \underline{\boldsymbol{\theta}} + \partial \boldsymbol{\theta} \quad (2.22)$$

onde $\underline{\theta} \in \mathbb{R}^m$ é o vector com as medidas sem erro e $\partial\theta \in \mathbb{R}^m$ representa a incerteza [Nakamura and Xu 1989]. Assume-se para $\partial\theta$ uma distribuição Gaussiana, ou seja

$$\begin{aligned} E[\partial\theta] &\triangleq \overline{\partial\theta} \\ &= \mathbf{0} \in \mathbb{R}^m \end{aligned} \quad (2.23)$$

$$\begin{aligned} V[\partial\theta] &\triangleq E[(\partial\theta - \overline{\partial\theta})(\partial\theta - \overline{\partial\theta})^T] \\ &= \mathbf{Q} \\ &= \text{diag}(\sigma_1^2, \dots, \sigma_m^2) \in \mathbb{R}^{m \times m} \end{aligned} \quad (2.24)$$

onde $E[*]$ representa a expectativa para $*$. É também assumido que $\partial\theta_j$ ($j = 1, \dots, m$), que é o j - ésimo elemento de $\partial\theta$, não está correlacionado e que σ_j^2 é a variância de $\partial\theta_j$. Por outro lado, $\mathbf{Q}$ é a matriz de covariância de $\partial\theta$ [Nakamura and Xu 1989].

Substituindo a equação (2.22) em (2.17) obtém-se

$$\mathbf{x} = \mathbf{f}(\underline{\theta} + \partial\theta) \quad (2.25)$$

Finalmente, se se assumir que o erro $\partial\theta$ é suficientemente pequeno, a equação anterior pode ser aproximada por

$$\mathbf{x} = \mathbf{f}(\underline{\theta}) + \mathbf{J}\partial\theta \quad (2.26)$$

onde $\mathbf{J}$ é o Jacobiano de $\mathbf{f}$, como definido na equação (2.20) [Nakamura and Xu 1989].

Das equações (2.23) e (2.26) obtém-se

$$\begin{aligned} E[\mathbf{x}] &\triangleq \bar{\mathbf{x}} \\ &= \mathbf{f}(\underline{\theta}) \end{aligned} \quad (2.27)$$

$$\begin{aligned} V[\mathbf{x}] &\triangleq E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] \\ &= E[\mathbf{J} \partial\theta \partial\theta^T \mathbf{J}^T] \\ &= \mathbf{J} \mathbf{Q} \mathbf{J}^T \end{aligned} \quad (2.28)$$

onde $E[\mathbf{x}]$ e $V[\mathbf{x}]$ são, respectivamente, a *vector-coluna de média* e a *matriz de covariância* da informação sensorial [Nakamura and Xu 1989].

Para o nosso caso, é possível simplificar a equação (2.28) se se considerar que o erro nas medidas das coordenadas na imagem tem a mesma representação estatística

(ou seja, $\sigma_u = \sigma_v = \sigma$) — neste caso, teremos $\mathbf{Q} = \sigma$, ou seja, a equação reduz-se a $V = \sigma \mathbf{J} \mathbf{J}^T$.

A equação (2.27) indica que, para uma calibração ideal, à medida que se repetem *ad infinitum* as medidas que permitem efectuar cálculos para determinar um dado $\mathbf{x}$, a média desses cálculos tenderá a convergir para o valor real desse $\mathbf{x}$ [Nakamura and Xu 1989]. Por esta razão, a importância da matriz de covariância da informação sensorial surge por esta ser *uma medida efectiva da incerteza introduzida pelo erro das medidas na estimativa final dada pelo modelo*.

Noção de elipsóide de incerteza

Como se pode verificar pela equação (2.28), a matriz de covariância de $\mathbf{x}$, ao contrário da matriz de covariância das medidas, não é diagonal, já que o Jacobiano, em geral, não o é. Por outras palavras, a correlação de x_i ($i = 1, \dots, m$), i -ésimo elemento de $\mathbf{x}$, está incluída no modelo, mesmo apesar de o conjunto de $\partial\theta_j$ ($j = 1, \dots, m$) ser assumido como não sendo correlacionado. Note-se, aliás, que para $\mathbf{J}$ de característica máxima, $\mathbf{J} \mathbf{Q} \mathbf{J}^T$ torna-se positivo definido, porque $\mathbf{Q}$ também o é, como se pode verificar pela equação (2.24) [Nakamura and Xu 1989].

Como $\mathbf{J} \mathbf{Q} \mathbf{J}^T$ é simétrico, a sua decomposição em valores singulares é representada por

$$\mathbf{J} \mathbf{Q} \mathbf{J}^T = \mathbf{U} \mathbf{D} \mathbf{U}^T \quad (2.29a)$$

$$\mathbf{U} = (\mathbf{e}_1 \dots \mathbf{e}_n) \in \mathbb{R}^{n \times n}$$

$$\mathbf{e}_j \in \mathbb{R}^n$$

$$\mathbf{e}_j^T \mathbf{e}_k = \begin{cases} 1 & , j = k \\ 0 & , j \neq k \end{cases} \quad (2.29b)$$

$$\mathbf{D} = \text{diag}(d_1, \dots, d_n)$$

$$d_1 \geq \dots \geq d_n \geq 0 \quad (2.29c)$$

onde $\mathbf{U}$ é uma matriz ortogonal e d_i ($i = 1, \dots, n$) são os valores singulares de $\mathbf{J} \mathbf{Q} \mathbf{J}^T$ [Nakamura and Xu 1989].

A partir deste resultado é possível retirar a variância escalar na direcção indicada

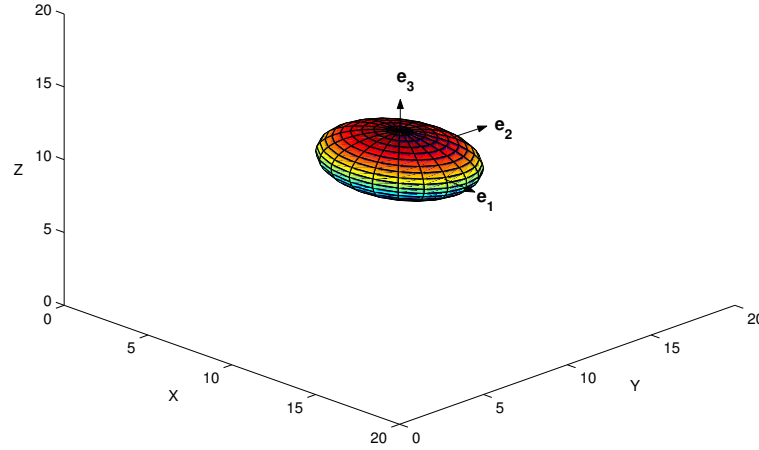


Figura 2.7: Elipsóide de incerteza: o caso tridimensional

pelo vector unitário $\mathbf{e}_j$, através da equação (2.30) [Nakamura and Xu 1989].

$$\begin{aligned}
 \mathbf{V} [\mathbf{e}_i^T \mathbf{x}] &= \mathbf{e}_i^T \mathbf{J} \mathbf{Q} \mathbf{J}^T \mathbf{e}_i \\
 &= \mathbf{e}_i^T \mathbf{U} \mathbf{D} \mathbf{U}^T \mathbf{e}_i \\
 &= d_i
 \end{aligned} \tag{2.30}$$

A incerteza de $\mathbf{x}$ na direcção de $\mathbf{e}_i$ é dada, portanto, por $\sqrt{d_i}$. Se se determinar as variâncias escalares para os vectores unitários correspondentes a todas as direcções cartesianas, o conjunto de vectores cuja direcção é dada pelos vectores unitários e cujas normas são dadas pelas incertezas correspondentes formam um elipsóide com eixos principais de direcção $\mathbf{e}_i$ e comprimento $2\sqrt{d_i}$. Este elipsóide, representado para o caso tridimensional na figura 2.7, é conhecido como *elipsóide de incerteza* — $\mathbf{e}_1$ e $\sqrt{d_1}$ correspondem à direcção de maior incerteza e $\mathbf{e}_n$ e $\sqrt{d_n}$ correspondem à direcção de menor incerteza [Nakamura and Xu 1989].

Volume do elipsóide de incerteza e sua minimização

O volume do elipsóide cujos eixos principais têm $2\sqrt{d_i}$ como comprimentos é calculado através de

$$volume = \frac{\pi^{n/2}}{\Gamma(1 + n/2)} \left(\prod_{i=1}^n d_i \right)^{1/2} \tag{2.31}$$

onde $\Gamma(*)$ é a função gama³ [Nakamura and Xu 1989].

Por outro lado, o determinante de uma matriz é igual ao produto dos seus valores singulares, e portanto, neste caso,

$$\begin{aligned}\det(\mathbf{JQJ}^T) &= \det(\mathbf{UDU}^T) \\ &= \prod_{i=1}^n d_i\end{aligned}\quad (2.32)$$

Substituindo (2.32) em (2.31), temos, finalmente,

$$volume = \frac{\pi^{n/2}}{\Gamma(1 + n/2)} \sqrt{\det(\mathbf{JQJ}^T)} \quad (2.33)$$

o que nos permite calcular o volume do elipsóide tendo o modelo do sistema de medida e os parâmetros (variâncias) das distribuições estatísticas dos dados de medida [Nakamura and Xu 1989].

Torna-se possível agora justificar a motivação para um estudo da sensibilidade ao erro do modelo do sistema de triangulação: é de todo o interesse estudar o comportamento do elipsóide de incerteza com mudanças nos parâmetros ajustáveis do modelo. Sobretudo, é de todo o interesse escolhê-los o mais possível com base na minimização do *volume* do elipsóide de incerteza; ou seja, minimizar através deles

$$\begin{aligned}objfun &\triangleq \det(\mathbf{V}) \\ &= \det(\mathbf{JQJ}^T) \\ &= \det(\sigma\mathbf{JJ}^T)\end{aligned}\quad (2.34)$$

2.2.2 Estudo teórico da sua aplicação a uma configuração particular do sistema de triangulação

Em seguida apresentar-se-á a dedução (feita recorrendo ao *Symbolic Toolbox* do *MATLAB*) das expressões para uma configuração de geometria simplificada para o sistema Tele-3D.

Modelo simplificado do sistema

Para aplicar a teoria apresentada nas secções anteriores, será de todo o interesse partir de alguns pressupostos que simplifiquem o modelo do sistema em análise, visto

³A designação de *volume* é, obviamente, um abuso de notação, dependendo da ordem da figura. Assim, no caso bidimensional, onde $n = 2$ (a figura é uma elipse), que mais nos interessa: $\Gamma(2) = (1 + \frac{2}{2} - 1)! = 1 \Rightarrow area = \pi \cdot \sqrt{d_1} \cdot \sqrt{d_2}$.

que, existindo uma estrutura rígida que suporta tanto o projector laser como a câmara, algumas das características geométricas do sistema podem ser sujeitas a aproximações inócuas para o estudo de sensibilidade ao erro que se pretende efectuar.

Para tal, definam-se os três ângulos de rotação dos eixos ortogonais do sistema referencial do projector laser $\{L\}$ que alinham estes eixos com os do referencial da câmara $\{C\}$: o ângulo de rotação em torno de Z , α , o ângulo de rotação em torno de Y , β , e o ângulo de rotação em torno de X , γ . Com estes três ângulos (veja-se a secção 2.1.2 na página 56) podemos redefinir a equação (2.11) de forma a obter uma expressão geral para a matriz de parâmetros extrínsecos (que, lembre-se, transformam o referencial do projector no referencial da câmara) dada por

$$\begin{bmatrix} \mathbf{R} & \vec{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \cos(\alpha) \cos(\beta) & \cos(\alpha) \sin(\beta) \sin(\gamma) - \sin(\alpha) \cos(\gamma) & \cos(\alpha) \sin(\beta) \cos(\gamma) + \sin(\alpha) \sin(\gamma) & B \cos(\theta) \sin(\phi) \\ \sin(\alpha) \cos(\beta) & \sin(\alpha) \sin(\beta) \sin(\gamma) + \cos(\alpha) \cos(\gamma) & \sin(\alpha) \sin(\beta) \cos(\gamma) - \cos(\alpha) \sin(\gamma) & B \sin(\theta) \sin(\phi) \\ -\sin(\beta) & \cos(\beta) \sin(\gamma) & \cos(\beta) \cos(\gamma) & B \cos(\phi) \end{bmatrix} \quad (2.35)$$

Uma configuração geométrica que, sendo uma aproximação suficientemente razoável da configuração apresentada na figura 2.1, possibilitaria uma simplificação considerável da expressão estabelecida em (2.35) está representada na figura 2.8. Aqui pode ver-se que a transformação de referenciais se dá toda sobre o plano YOZ do sistema de eixos do projector, sendo $\alpha = 180^\circ$ e $\beta = 0^\circ$, e $\theta = 90^\circ$. É possível verificar também pela figura que a lista de interferência no plano laser projectada no plano-imagem dará origem a um *perfil horizontal* e que os ângulos que os eixos principais⁴ fazem com a “baseline” deverão, por razões de ordem prática (campo de visão, etc.), estar compreendidos no intervalo $]0^\circ, 90^\circ[$, ou seja, $\phi \in]90^\circ, 180^\circ[$ e $\gamma \in]0^\circ, 180^\circ[$. Por outro lado, repare-se que nem todos os ângulos nestes intervalos mesmo assim terão utilidade na prática, visto que poderão implicar oclusões impossíveis de resolver ou disparidades nulas.

Substituindo estes valores na equação (2.35), a matriz de parâmetros extrínsecos simplificada torna-se

$$\begin{bmatrix} \mathbf{R} & \vec{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -\cos(\gamma) & \sin(\gamma) & B \sin(\phi) \\ 0 & \sin(\gamma) & \cos(\gamma) & B \cos(\phi) \end{bmatrix} \quad (2.36)$$

⁴O eixo óptico para a câmara e o eixo que direcciona o plano de laser, ou seja, ambos os eixos dos ZZ para cada referencial.

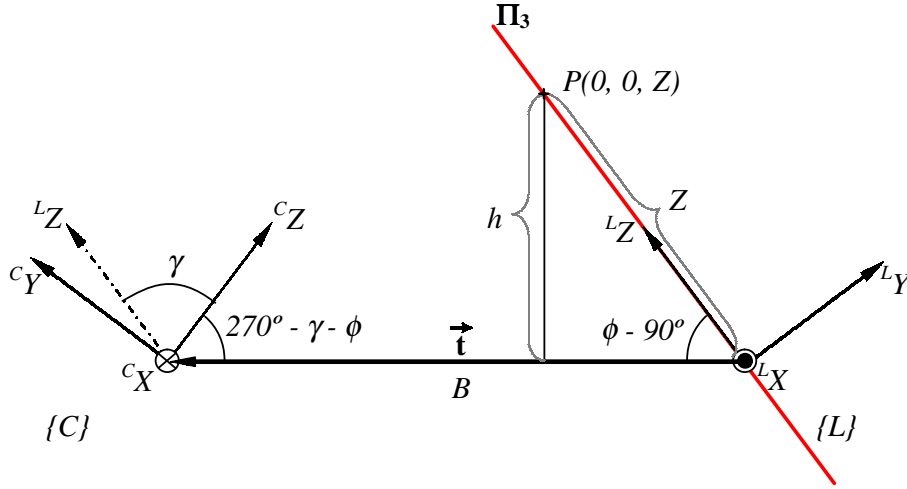


Figura 2.8: Modelo simplificado do sistema Tele-3D

o que substituindo em (2.11) e desenvolvendo a expressão produz

$$\mathbf{A}_{3 \times 4} = \begin{bmatrix} -f_x & c_x \sin(\gamma) & c_x \cos(\gamma) & c_x B \cos(\phi) \\ 0 & -f_y \cos(\gamma) + c_y \sin(\gamma) & f_y \sin(\gamma) + c_y \cos(\gamma) & f_y B \sin(\phi) + c_y B \cos(\phi) \\ 0 & \sin(\gamma) & \cos(\gamma) & B \cos(\phi) \end{bmatrix} \quad (2.37)$$

Usando a matriz de projecção dada pela equação anterior para efectuar a triangulação e determinar $\mathbf{f}$ (equação (2.19)), obtemos, finalmente,

$$\begin{aligned} X &= -\frac{(f_y \cos(\gamma) B \sin(\phi) + c_y \cos(\gamma) B \cos(\phi) - f_y \sin(\gamma) - c_y \cos(\gamma))}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))} u - \\ &\quad - \frac{(-c_x B \cos(\phi) \cos(\gamma) + c_x \cos(\gamma))}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))} v + \\ &\quad + \frac{c_x B \cos(\phi) f_y \sin(\gamma) - c_x f_y \cos(\gamma) B \sin(\phi)}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))} \end{aligned} \quad (2.38a)$$

$$Z = \frac{-v + f_y B \sin(\phi) + c_y B \cos(\phi)}{v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma)} \quad (2.38b)$$

Análise da GDOP

Como foi dito anteriormente, uma parte essencial do estudo de sensibilidade ao erro consiste na determinação do jacobiano de $\mathbf{f}$, ou seja, a *GDOP*, como apresentado nas equações (2.20) e (2.21). Para tal, apresentam-se seguidamente as equações para

as derivadas parciais dos elementos do vector de informação sensorial relativamente a cada elemento do vector de medidas, obtidas a partir de derivações consecutivas das expressões (2.38).

$$\frac{\partial X}{\partial u} = -\frac{c_y \cos(\gamma) B \cos(\phi) + \cos(\gamma) f_y B \sin(\phi) - c_y \cos(\gamma) - f_y \sin(\gamma)}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))} \quad (2.39a)$$

$$\begin{aligned} \frac{\partial X}{\partial v} &= \frac{\cos(\gamma) (c_y \cos(\gamma) B \cos(\phi) + \cos(\gamma) f_y B \sin(\phi) - c_y \cos(\gamma) - f_y \sin(\gamma))}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))^2} u - \\ &\quad - \frac{c_x \cos(\gamma) (c_y \cos(\gamma) B \cos(\phi) + \cos(\gamma) f_y B \sin(\phi) - c_y \cos(\gamma) - f_y \sin(\gamma))}{f_x (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))^2} \end{aligned} \quad (2.39b)$$

$$\frac{\partial Z}{\partial u} = 0 \quad (2.39c)$$

$$\frac{\partial Z}{\partial v} = -\frac{c_y \cos(\gamma) B \cos(\phi) + \cos(\gamma) f_y B \sin(\phi) - c_y \cos(\gamma) - f_y \sin(\gamma)}{(v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))^2} \quad (2.39d)$$

Pode-se concluir da *GDOP* que, tanto as variações de X face a diferenças infinitesimais em u , como as variações de Z face a diferenças infinitesimais em v não dependem da posição horizontal no plano-imagem da projecção do ponto amostrado. Mas, mais importante que isto, é possível verificar que *a estimativa para a coordenada Z do ponto não é sensível à posição horizontal do seu ponto de projecção.*

Análise da elipse de incerteza e sua minimização

Da *GDOP* e da equação (2.39) é possível determinar, como já foi visto, a função-objectivo a minimizar para reduzir ao máximo a área da elipse de incerteza do modelo do sistema Tele-3D. Aliás, como minimizar $\det(\sigma \mathbf{J} \mathbf{J}^T)$ é o mesmo que minimizar $\det(\mathbf{J} \mathbf{J}^T)$, visto que σ é um valor constante, podemos considerar $\sigma = 1 \text{ pixel}$ e teremos

$$\begin{aligned} objfun &= \det(\mathbf{J} \mathbf{J}^T) \\ &= \frac{(c_y \cos(\gamma) B \cos(\phi) + \cos(\gamma) f_y B \sin(\phi) - c_y \cos(\gamma) - f_y \sin(\gamma))^4}{f_x^2 (v \cos(\gamma) - c_y \cos(\gamma) - f_y \sin(\gamma))^6} \end{aligned} \quad (2.40)$$

É possível verificar de imediato desta equação que existem dois parâmetros cuja variação tem implicações evidentes na evolução da função-objectivo: *a “baseline” e a posição vertical da projecção no plano-imagem.*

Mantendo os restantes parâmetros constantes, a função-objectivo função da posição vertical é monótona ascendente em $] -\infty, v_{crit}[$ e monótona descendente em $] v_{crit}, +\infty[$, sendo

$$v_{crit} = \frac{c_y \cos(\gamma) + f_y \sin(\gamma)}{\cos(\gamma)} = c_y + f_y \tan(\gamma) \quad (2.41a)$$

$$\lim_{v \rightarrow v_{crit}} objfun = +\infty \quad (2.41b)$$

ou, por outras palavras, v_{crit} é o valor de posição vertical no ecrã do perfil projectado no plano-imagem para o qual não é possível obter *uma estimativa de posição 3D minimamente correlacionada com a posição real do ponto original*. No entanto, poucas vezes esta posição corresponde a uma posição dentro da imagem real (no caso PAL, por exemplo, dentro das linhas de 0 a 576) — os ângulos dos eixos principais com a “baseline”, que por sua vez condicionam como já se viu ϕ e γ , e os próprios parâmetros extrínsecos da câmara fazem com que a assíptota esteja na prática mais comumente na zona $v < 0$.

Da mesma forma, mantendo os restantes parâmetros constantes⁵, a função-objectivo função da “baseline” é monótona descendente em $] -\infty, B_{crit}[$ e monótona ascendente em $] B_{crit}, +\infty[$, sendo

$$B_{crit} = \frac{c_y \cos(\gamma) + f_y \sin(\gamma)}{c_y \cos(\gamma) \cos(\phi) + f_y \cos(\gamma) \sin(\phi)} \quad (2.42a)$$

$$\lim_{B \rightarrow B_{crit}} objfun = 0 \quad (2.42b)$$

ou, por outras palavras, B_{crit} é o valor de comprimento da “baseline” para o qual *o erro da estimativa de posição 3D se anula*. No entanto, mais uma vez e pelas mesmas razões que anteriormente, poucas vezes este valor corresponderá a um comprimento fisicamente possível (isto é, $B > 0$).

Por fim, note-se que se anulam os efeitos da “baseline” e de v quando $\gamma = 90^\circ$, tornando-se irrelevantes para o cálculo de $objfun$. Um caso de grande interesse prático em que este facto se dá é $\phi = 135^\circ, \gamma = 90^\circ$, para o qual ambos os eixos principais estão num ângulo de 45° com a “baseline”.

⁵Atente-se ao facto, porém, de que manter v e os restantes parâmetros constantes, variando B , implica que a posição (X, Z) que se quer medir varia (mais concretamente Z , e portanto a distância dessa posição ao *scanner*) para manter o mesmo fotosensor sensibilizado na câmara.

2.2.3 Simulação e análise de sensibilidade ao erro de uma instânciação aproximada do *scanner* Tele-3D

Far-se-á seguidamente uma apresentação em *MATLAB* de resultados obtidos para uma configuração particular do sistema Tele-3D — a maioria deste estudo de optimização foi inspirado em [Magrab et al. 2000].

No estudo dessa configuração usaram-se dados referentes a uma câmara PAL (768×576 pixels) que foi experimentada em laboratório no *scanner* Tele-3D, cujos parâmetros intrínsecos foram determinados por calibração como sendo $f_x = 1493.004150$, $f_y = 1438.083862$, $c_x = 402.759735$ e $c_y = 397.826294$ usando software baseado nos algoritmos do OpenCV da Intel [OpenCV Manual, Zhang 1999, Heikkilä and Silven 1997].

Estudo da evolução do volume da elipse de incerteza com a “*baseline*” e a posição vertical na imagem

Um parâmetro importante em termos de estudo da sensibilidade ao erro de um digitalizador tridimensional por triangulação é a “*baseline*”. Porém, como pode ser depreendido do que foi dito anteriormente, quando se varia a “*baseline*” torna-se necessário re-projectar no plano-imagem o ponto cujas coordenadas fixas (X, Z) se querem medir, isto é, *recalcular* u e v . Para isso usa-se a equação (2.43), obtida resolvendo as equações (2.38) (página 66) em ordem a u e v :

$$u = \frac{c_x B \cos(\phi) + c_x Z \cos(\gamma) - X f_x}{1 + Z \cos(\gamma)} \quad (2.43a)$$

$$v = \frac{f_y B \sin(\phi) + c_y B \cos(\phi) + Z c_y \cos(\gamma) + Z f_y \sin(\gamma)}{1 + Z \cos(\gamma)} \quad (2.43b)$$

Torna-se claro pelas equações anteriores e pela equação (2.40) (página 67) que, para efeitos de estudo da evolução do volume da elipse de incerteza, apenas é necessário determinar v e, por inerência, Z . Desta forma, pode-se evitar o cálculo de u e arbitrar o valor de X como sendo 0 — assim, estabelecendo uma distância fixa, h , medida desde a “*baseline*” até à posição que se quer determinar, é possível recorrer à trigonometria simples do triângulo representado na figura 2.8 na página 66 para estabelecer $Z = \frac{h}{\sin(\phi - 90^\circ)}$, e com o conhecimento deste valor determinar v , usando a equação (2.43b).

Nas figuras 2.9, 2.10 e 2.11 (páginas 71 a 73), mostram-se resultados da simulação da evolução de *objfun* face a B para vários valores de interesse prático de ϕ e γ e

para $h = 25$ cm. Veja-se que o erro apresenta um comportamento relativamente lógico: *é geralmente inversamente proporcional à “baseline”*. A única excepção a esta regra verifica-se para os gráficos da figura 2.10 na página 72, onde se pode ver a influência de v_{crit} — até esse valor ser atingido, como tinha sido previsto, o erro é ascendente, passando a ser inversamente proporcional à “baseline” desse valor em diante.

Também foi simulada a evolução de *objfun* face a B para $\phi = 135^\circ, \gamma = 90^\circ$ (com o mesmo h), onde se obteve, como seria de esperar, um valor constante para a função objectivo ($objfun = 2.1692 \times 10^{-13}$). Note-se, porém, que *estas condições são impossíveis de reproduzir em laboratório*, não só devido às diferenças inatas entre a simulação e a experimentação, como também devido aos erros derivados da calibração do sistema, dos quais estas simulações foram isoladas.

Na figura 2.12 na página 74, apresentam-se resultados da simulação da evolução de *objfun* face a v para $\phi = 120^\circ, \gamma = 120^\circ$. Veja-se que, mais uma vez, o erro tem o comportamento esperado: *é inversamente proporcional à posição vertical na imagem*.

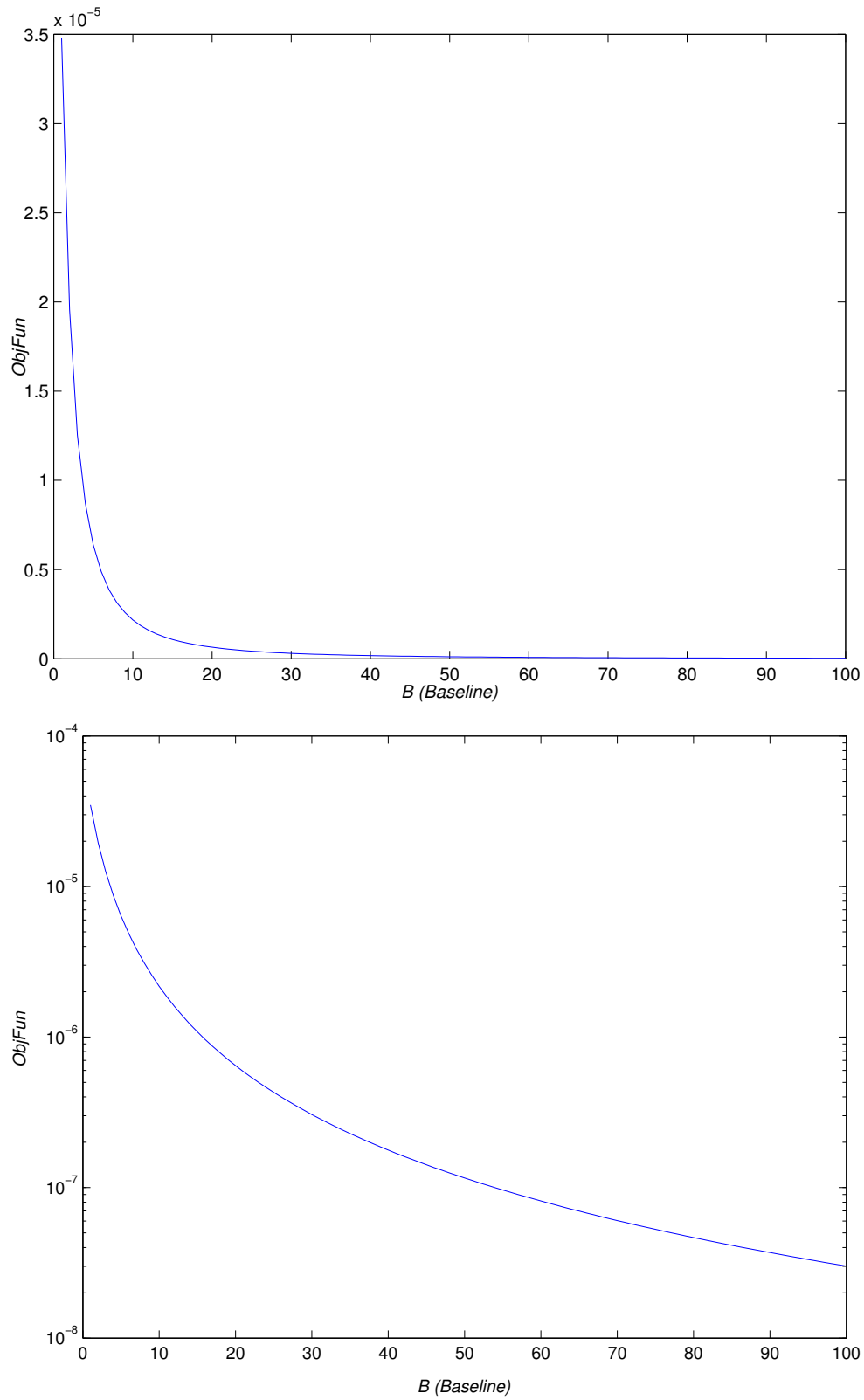


Figura 2.9: Função objectivo vs baseline (em centímetros) para $\phi = 120^\circ$, $\gamma = 120^\circ$ e $h = 25$ cm. Em cima, está representada num gráfico linear; em baixo, num gráfico semi-logarítmico.

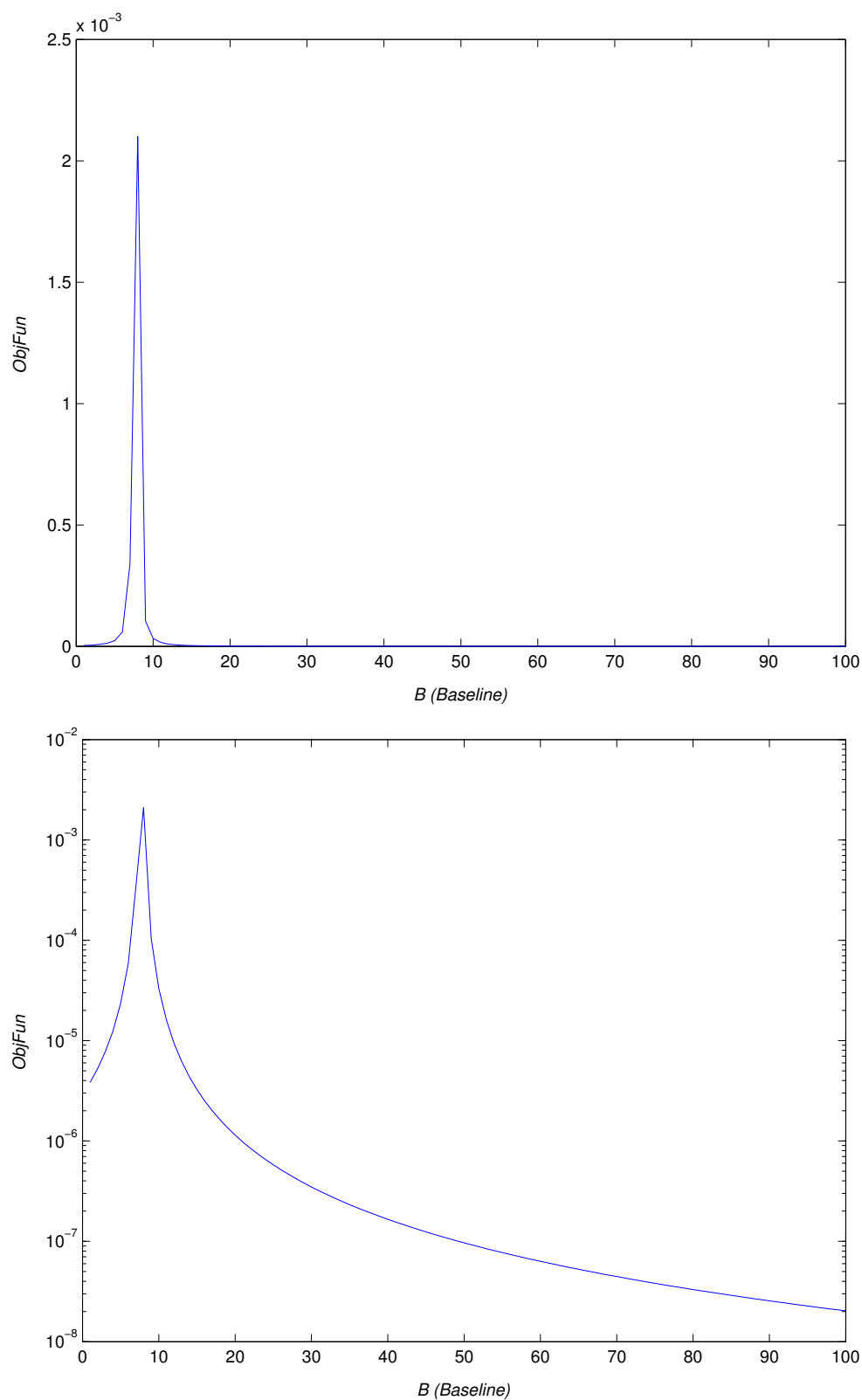


Figura 2.10: Função objectivo vs baseline (em centímetros) para $\phi = 150^\circ$, $\gamma = 60^\circ$ e $h = 25$ cm. Em cima, está representada num gráfico linear; em baixo, num gráfico semi-logarítmico.

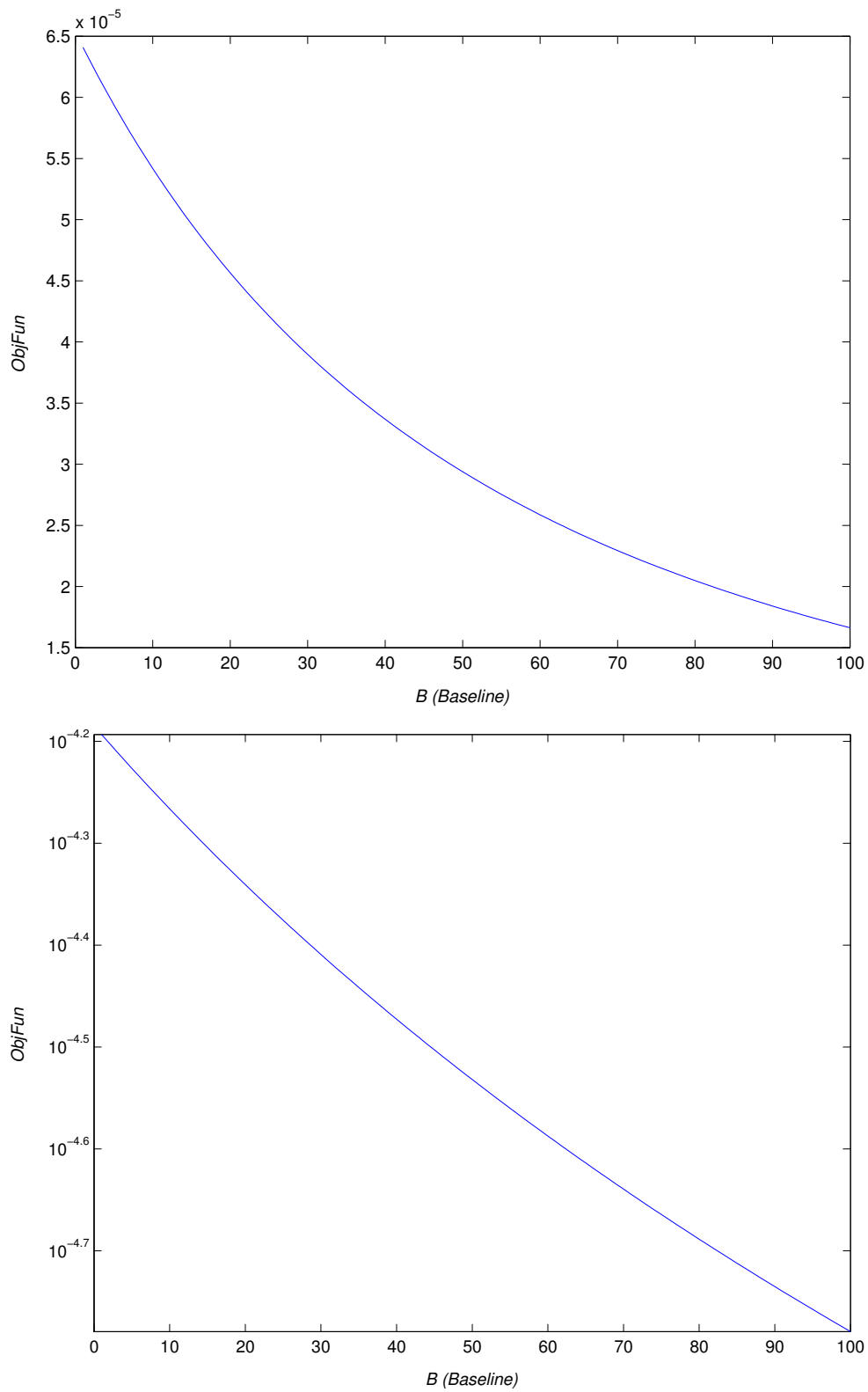


Figura 2.11: Função objectivo vs baseline (em centímetros) para $\phi = 165^\circ$, $\gamma = 30^\circ$ e $h = 25$ cm. Em cima, está representada num gráfico linear; em baixo, num gráfico semi-logarítmico.

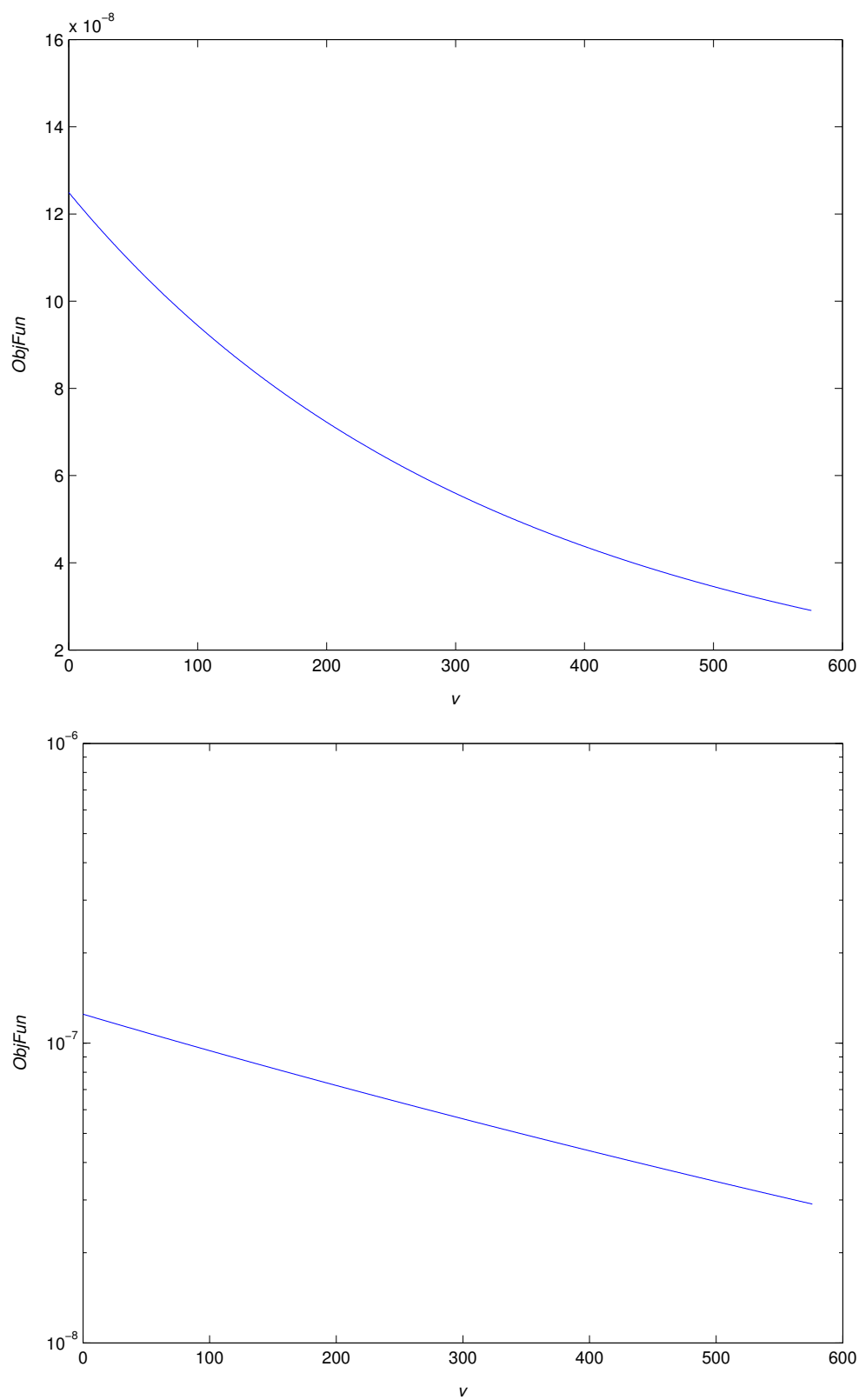


Figura 2.12: Função objectivo vs posição vertical na imagem (em pixels). Em cima, está representada num gráfico linear; em baixo, num gráfico semi-logarítmico.

Estudo da evolução do volume da elipse de incerteza com a vergência entre a câmara e o projector laser

Na figura 2.13 na página seguinte apresenta-se a evolução de $objfun$ face a variações em ϕ e γ , dos quais, por sua vez, depende a vergência da câmara com o projector laser, para $B = 45\text{ cm}$ e $h = 25\text{ cm}$. É possível verificar que, à parte das zonas em que ϕ e γ têm valores de pouca utilidade prática (veja-se os comentários relativos a este assunto na secção 2.2.2) e onde existem “variações quase assintóticas” de $objfun$, o valor desta vai diminuindo em direcção a $\gamma \approx 90^\circ$ — consulte-se a discussão sobre a importância deste último valor apresentada na secção 2.2.2.

2.2.4 Análise dos resultados do estudo da sensibilidade ao erro do modelo de triangulação

Analisando os resultados deste estudo, inclusivamente os gráficos obtidos a partir das simulações efectuadas, é possível concluir que os parâmetros decisivos para controlar a sensibilidade ao erro do modelo serão o comprimento da “baseline” e γ : torna-se claro que é possível diminuir drasticamente o erro cometido nas estimações de posição através do aumento de B e da utilização de $\gamma = 90^\circ$; regulando os outros parâmetros não terá tanta utilidade prática, como no caso de ϕ .

Por outro lado, será possível diminuir o erro se se forçar pontos a distâncias dentro da gama de operação a ter projecções mais perto da zona inferior do plano-imagem, se bem que se deve ter algum cuidado para não fazer ultrapassar os limites deste último. A vantagem que se obtém em termos de precisão procedendo-se desta forma é, porém, limitada.

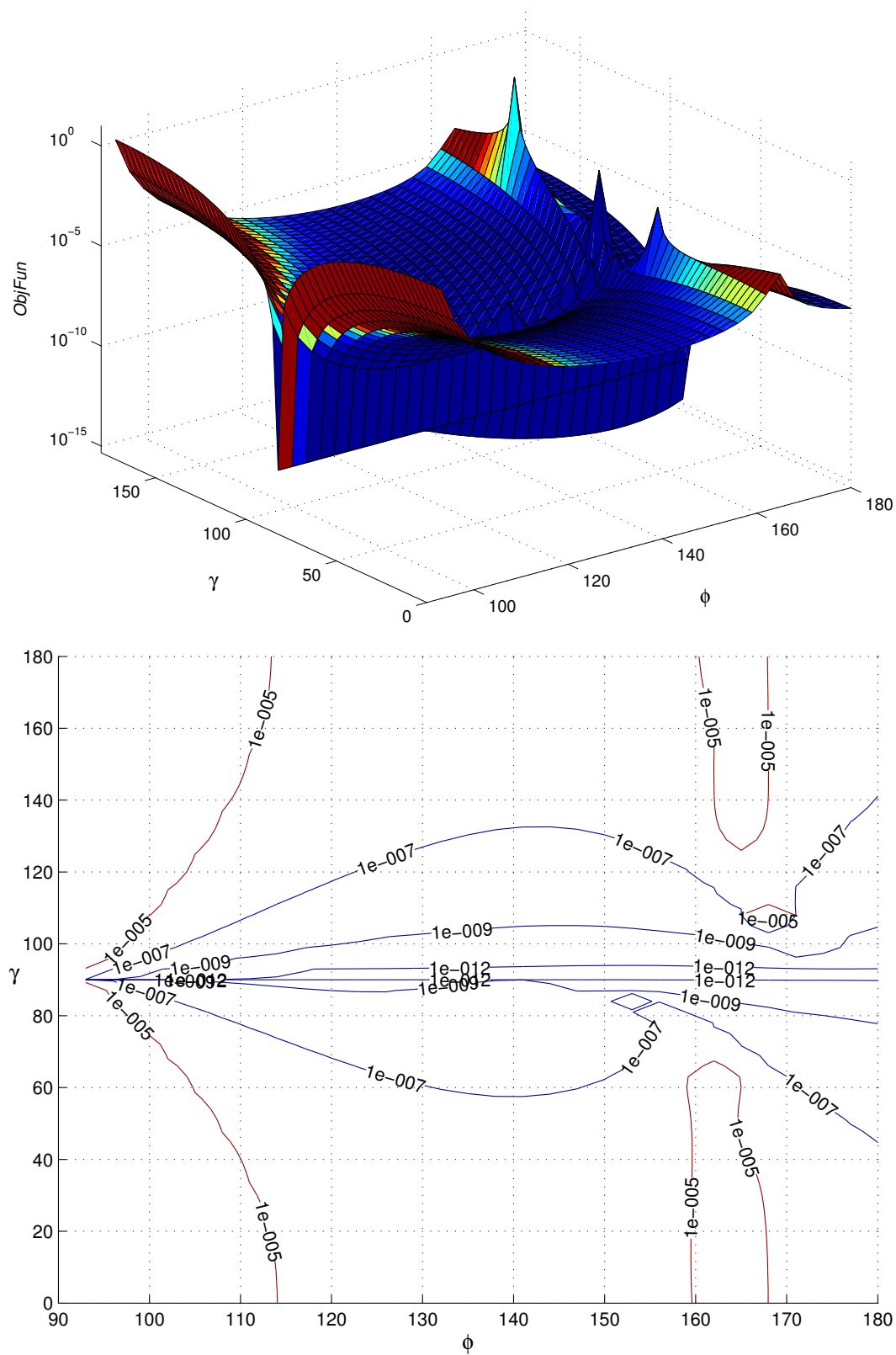


Figura 2.13: Função objetivo vs ϕ e γ , para $B = 45$ cm e $h = 25$ cm.

2.3 Calibração do *scanner* Tele-3D

Entende-se por calibração do *scanner* a determinação dos parâmetros definidos nos modelos geométricos anteriormente descritos, única e exclusivamente através de processos de estimação. O processo de calibração implica por inerência, portanto, a determinação da relação entre os sistemas de coordenadas representados na figura 2.1 (página 49).

Uma grande parte dos equipamentos de digitalização tridimensional estão desenhados de forma a que alguns dos parâmetros são já conhecidos de antemão; dito de outra forma, a sua geometria é concebida de maneira a que a maior parte das distâncias e ângulos envolvidos são já conhecidos de forma a simplificar a calibração ou mesmo eliminá-la por completo. Isto, porém, traz algumas desvantagens: qualquer defeito de fabrico poderá introduzir erros nas medições, prejudicando irremediavelmente a precisão do equipamento, e, ainda mais importante, perde-se por completo flexibilidade na concepção e construção desse equipamento.

Optou-se neste trabalho, por estas razões, por uma abordagem em que se usam modelos geométricos genéricos dependendo exclusivamente dos processos de calibração. Por outro lado, tentou-se simplificar estes processos o mais possível de modo a evitar grande complexidade na preparação do *scanner* para utilização, sem comprometer a sua precisão final nas medições.

Descrever-se-á nas subsecções que se seguem, portanto, a calibração de cada sistema envolvido.

2.3.1 Calibração da câmara

A calibração da câmara consiste na estimação dos seus parâmetros intrínsecos; a partir destes e do conhecimento da sua “*pose*” torna-se possível determinar a matriz de projecção que dá a relação entre pontos no mundo e as suas projecções no plano-imagem, como foi definido na equação (2.1) [Ferreira and Dias 2000, Ferreira et al. 2001, Dias 1994, Faugeras 1993, Hartley and Zisserman 2000].

Para o efeito, recorreu-se a um pacote de software de calibração robusta de câmaras usando o *Open Source Computer Vision Library* da Intel, ou em forma abreviada OpenCV [OpenCV Manual]. O método de calibração do OpenCV, baseado em Zhang [2000; 1999], Heikkilä and Silven [1997], é um algoritmo iterativo aplicado ao modelo de câmara descrito anteriormente, envolvendo também dois parâmetros que modelam a distorção radial e dois parâmetros que modelam a distorção tangencial. Estes parâmetros permitem relacionar coordenadas ideais no plano-imagem $[x \ y]^T$ e as co-

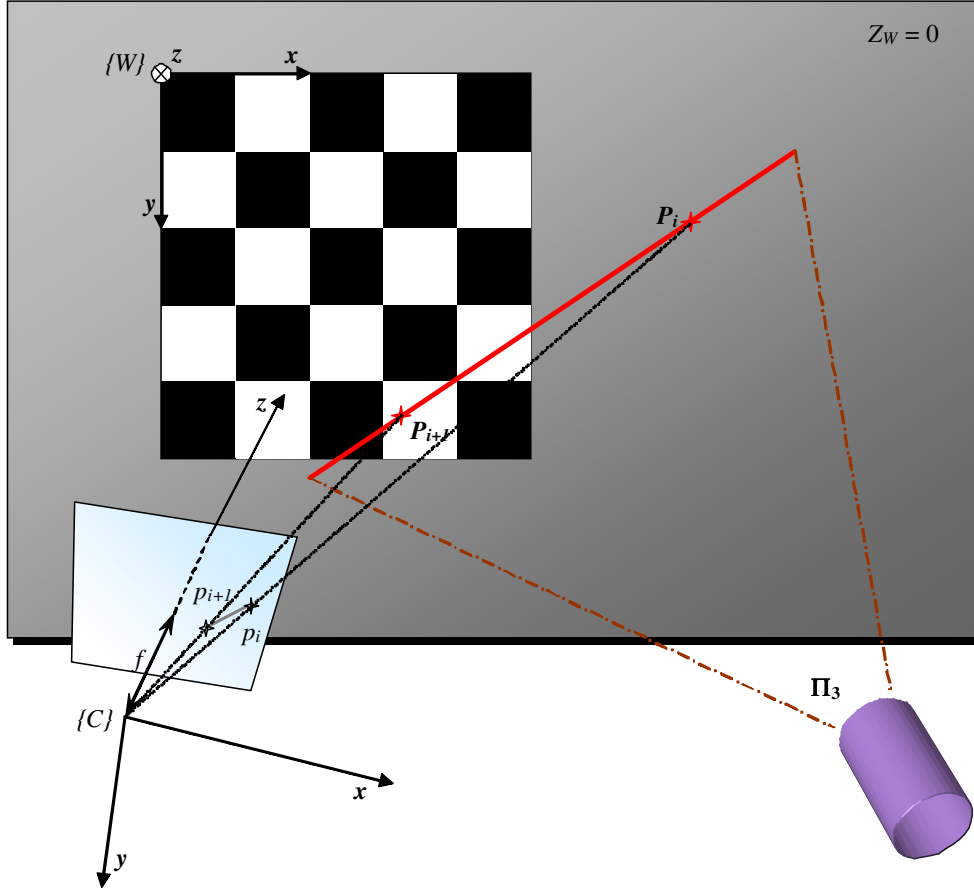


Figura 2.14: Esquema mostrando a configuração inerente simultaneamente à calibração da câmara e do laser.

ordenadas medidas correspondentes $[\tilde{x} \ \tilde{y}]^T$ antes do fenómeno de amostragem através de

$$\begin{cases} \tilde{x} = x + x(k_1 r^2 + k_2 r^4) + [2p_1 xy + p_2(r^2 + 2x^2)] \\ \tilde{y} = y + y(k_1 r^2 + k_2 r^4) + [2p_1 xy + p_2(r^2 + 2y^2)] \end{cases} \quad (2.44)$$

onde $r^2 = x^2 + y^2$.

Este algoritmo usa sequências de imagens de um padrão em xadrez que lhe fornece pontos tridimensionais cujas coordenadas são conhecidas com precisão, que se pode ver na figura 2.14, operando como descrito em seguida [OpenCV Manual]:

1. Calcula a homografia para todos os pontos entre cada imagem do padrão.
2. Inicializa os parâmetros intrínsecos; coloca os parâmetros de distorção a 0.

3. Descubra os parâmetros extrínsecos para cada imagem.
4. Executa um processo de optimização, minimizando o erro de projecção com todos os parâmetros.

Como foi referido acima, a calibração da câmara, em sentido estrito, consiste na determinação dos seus parâmetros intrínsecos; porém, torna-se facto adquirido que, na prática, durante o processo de calibração também se determinam os parâmetros extrínsecos para cada imagem do padrão.

Desta forma, para essas imagens obtém-se directamente uma estimativa para a transformação rígida correspondente entre o referencial do mundo e o referencial da câmara, ${}^C\mathbf{T}_W$, como está patente na equação (2.12), na página 55.

2.3.2 Calibração do projector laser

Durante a calibração da câmara descrita na subsecção anterior, é estimada a chamada *colinação* ou *homografia planar*⁶, relação que estabelece uma correspondência *unívoca* entre coordenadas homogéneas de pontos no plano de calibração e as coordenadas homogéneas das suas projecções correspondentes [OpenCV Manual, Dias 1994, Faugeras 1993, Hartley and Zisserman 2000]. De facto, no processo de calibração da câmara assume-se que o WCS⁷, $\{W\}$, está associado ao plano de calibração de forma a que a origem coincida com um dos cantos exteriores do xadrez e que esse plano coincida com o plano XOY com os eixos alinhados com o padrão (implicando, portanto, que $Z_W = 0$ para qualquer ponto pertencente ao plano de calibração — veja-se a figura 2.14); desta forma, a homografia pode ser retirada directamente das equações (2.1) e (2.10) como sendo

$$\begin{aligned} \begin{bmatrix} su \\ sv \\ s \end{bmatrix} &= \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ 1 \end{bmatrix} \\ &= \mathbf{H}_{3 \times 3} \begin{bmatrix} X_W \\ Y_W \\ 1 \end{bmatrix} \end{aligned} \quad (2.45)$$

⁶Geralmente abreviado para *homografia*.

⁷“World Coordinate System” — denominação dada ao referencial global em aplicações CAD (Desenho Assistido por Computador), bastante popular na comunidade de investigação 3D.

onde $\mathbf{r}_1$, $\mathbf{r}_2$ e $\mathbf{t}$ são obviamente os vectores coluna correspondentes na matriz de parâmetros extrínsecos da equação (2.10), e $s' = 1/s$ o factor de escala que afecta globalmente a homografia [OpenCV Manual, Dias 1994, Faugeras 1993, Hartley and Zisserman 2000]. Por questão de clareza, tornou-se explícito o facto de as coordenadas X e Y se referirem ao referencial $\{W\}$, subscrevendo-as em conformidade.

Como esta relação é unívoca, é sempre possível determinar a transformação inversa e portanto $\mathbf{H}_{3 \times 3}$ é sempre não-singular [Hartley and Zisserman 2000]. Logo, dado que a inversão de matrizes 3×3 é um processo trivial, é possível facilmente obter-se a relação de *homografia inversa*. Esta relação, dadas as coordenadas da projecção na imagem de pontos pertencentes ao plano de calibração, permite determinar as coordenadas desses pontos relativamente a $\{W\}$ usando

$$\begin{bmatrix} s'X_W \\ s'Y_W \\ s' \end{bmatrix} = \mathbf{H}_{3 \times 3}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (2.46)$$

onde $s' = 1/s$ é o factor de escala que afecta globalmente a homografia inversa.

Aplicando a equação (2.46) às projecções de todos os pontos correspondentes aos cantos dos quadrados do xadrez em cada imagem tirada deste último, torna-se possível determinar as coordenadas de cada ponto em $\{C\}$, o referencial privado da câmara, através de

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = {}^C\mathbf{T}_W \cdot \begin{bmatrix} X_W \\ Y_W \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ 0 \\ 1 \end{bmatrix} \quad (2.47)$$

Veja-se a figura 2.14 — durante a calibração da câmara, que se deverá então fazer tendo o projector *laser* ligado e movendo o sistema num procedimento parecido com o descrito em Prager et al. [1997] de forma a cobrir todos os graus de liberdade, é possível retirar sempre 2 pontos-imagem linearmente independentes, projecções de pontos 3D pertencentes à intersecção entre o plano-laser e o plano de calibração, e calcular as coordenadas relativamente a $\{C\}$ dos pontos que lhes deram origem usando as equações (2.46) e (2.47) [Ferreira et al. 2002b].

Reunindo três ou mais pontos obtidos dessa forma, estes podem por sua vez ser usados para determinar os parâmetros do plano-laser, dado pela equação (2.4), afectados ao referencial da câmara $\{C\}$ através da resolução da seguinte equação, usando

um método de mínimos quadrados apropriado:

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ X_3 & Y_3 & Z_3 & 1 \\ \dots & \dots & \dots & \dots \\ X_n & Y_n & Z_n & 1 \end{bmatrix} \begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (2.48)$$

Para se proceder à sua resolução, os parâmetros necessitam de ser normalizados. Escolheu-se a restrição $\|\hat{\mathbf{n}}\|^2 = 1$ com $\hat{\mathbf{n}} = [a' \ b' \ c']^T$, sendo a' , b' , c' as versões normalizadas requeridas e d' o parâmetro dependente — a intenção por detrás desta normalização específica é a obtenção da equação do plano na sua forma hermitiana, mais adequada para os fins em causa⁸. Neste forma, $\hat{\mathbf{n}}$ representa o versor normal a Π_3 e d' a distância de Π_3 a $\{C\}$.

Neste trabalho, o método de mínimos quadrados concreto utilizado na resolução deste problema encontra-se descrito no Apêndice E.2.

2.3.3 Calibração do sistema de medição de “pose”

Como foi dito no final da secção 2.1, o processo de integração de dados tridimensionais depende da computação das estimativas iniciais das transformações rígidas necessárias para o registo dos dados em cada “pose” do sistema. Torna-se, portanto, necessário estabelecer e calibrar a relação entre as leituras do sensor de “pose” e as suas transformações correspondentes ${}^T\mathbf{T}_R$, e as transformações inerentes ao registo 3D ${}^W\mathbf{T}_C$. Finalmente, para esta calibração ser possível, é essencial existir um meio de adquirir *valores de referência que se considerem as melhores estimativas possíveis para os valores exactos* (veja-se Chapra and Canale [2002a]) de ${}^W\mathbf{T}_C$. Os próximos parágrafos explicam como estes requisitos foram satisfeitos.

Uma matriz de rotação genérica é dada por uma matriz ortogonal 3×3 , e portanto $\mathbf{R}^{-1} = \mathbf{R}^T$ [Goldstein et al. 2002]; logo, se $\mathbf{P}_C$ for um vector de coordenadas correspondendo a um ponto referido ao referencial local da câmara e $\mathbf{P}_W$ um vector de coordenadas correspondendo ao mesmo ponto mas desta feita referido ao WCS, então

$$\begin{aligned} \mathbf{P}_C &= {}^C\mathbf{R}_W \cdot \mathbf{P}_W + {}^C\vec{\mathbf{t}}_W \quad \Leftrightarrow \\ \mathbf{P}_W &= {}^C\mathbf{R}_W^T \cdot \mathbf{P}_C - {}^C\mathbf{R}_W^T \cdot {}^C\vec{\mathbf{t}}_W \end{aligned}$$

⁸Para ver outras normalizações possíveis, consulte-se Zhang [1995].

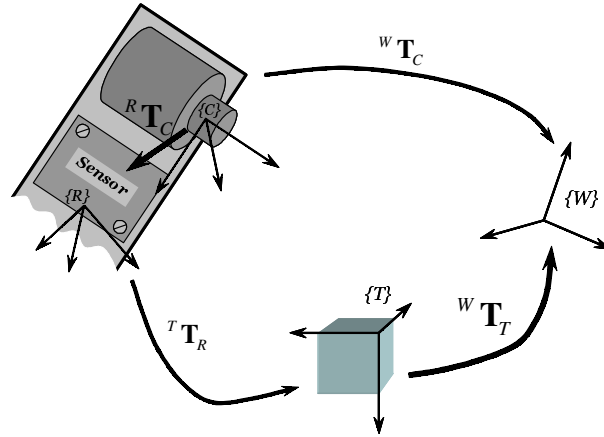


Figura 2.15: Calibração do sensor de “pose” — formulação estática. $\{T\}$ é o sistema de coordenadas do emissor quando o sensor usa a abordagem emissor/receptor, ou o referencial de inicialização quando o sensor é auto-contido.

Isto implica que a estimação dos parâmetros extrínsecos da câmara, executada implicitamente durante a calibração desta, pode ser usada para adquirir os valores de referência requeridos para ${}^W\mathbf{T}_C$, visto que a equação precedente torna trivial a determinação destes a partir dos componentes de rotação e translação da transformação inversa ${}^C\mathbf{T}_W$ como sendo

$${}^W\mathbf{T}_C = \begin{bmatrix} {}^C\mathbf{R}_W^T & -{}^C\mathbf{R}_W^T \cdot {}^C\vec{\mathbf{t}}_W \\ \mathbf{0} & 1 \end{bmatrix} \quad (2.49)$$

Isto também implica que os dados requeridos para as calibrações da câmara, do projector laser e do sensor de “pose” podem ser adquiridos simultaneamente e eficientemente com qualquer grau de automação desejado [Ferreira et al. 2002b].

No texto que se segue, irão ser derivadas duas formulações alternativas para a relação requerida entre ${}^T\mathbf{T}_R$ e ${}^W\mathbf{T}_C$, e as soluções matemáticas para os dois métodos de calibração correspondentes para cada relação vão ser descritas subsequentemente.

Formulação estática — uma solução iterativa, não linear, usando ângulos de Euler e matrizes de transformação homogêneas

A figura 2.15 mostra que, para qualquer momento no tempo, existe claramente uma relação entre ${}^W\mathbf{T}_C$ e ${}^T\mathbf{T}_R$ dada por

$${}^W\mathbf{T}_C = {}^W\mathbf{T}_T \cdot {}^T\mathbf{T}_R \cdot {}^R\mathbf{T}_C \quad (2.50)$$

onde ${}^R\mathbf{T}_C$ é a transformação entre o referencial da câmara e o referencial do sensor e ${}^W\mathbf{T}_T$ é a transformação entre $\{T\}$ e o WCS.

Se o padrão de calibração estiver fixo e o *scanner* for movido de modo a capturar várias perspectivas do xadrez, uma inspecção mais cuidadosa da figura 2.15 revela o facto importante de tanto ${}^R\mathbf{T}_C$ como ${}^W\mathbf{T}_T$ permanecem constantes durante o movimento (isto é, o referencial da câmara permanece estático relativamente a $\{R\}$ e $\{T\}$ permanece estático relativamente ao padrão). Portanto, se entretanto forem sendo recolhidos medidas do sensor e os resultados para os parâmetros extrínsecos da câmara correspondentes, um problema de estimação no sentido dos mínimos quadrados pode ser então formulado. Visto que esta acumulação de dados serve apenas o propósito de sobre-determinar o sistema de equações de estimação, sendo este dependente apenas da física estática do sistema, esta formulação e a sua solução correspondente, descrita nos parágrafos que se seguem, serão referidos como *formulação estática* para a calibração do sensor de “pose”.

${}^W\mathbf{T}_T$ e ${}^R\mathbf{T}_C$ são, pois, as matrizes de incógnitas a estimar (totalizando 24 incógnitas, 12 das quais linearmente independentes — 2×3 graus de liberdade em rotação + 2×3 graus de liberdade em translação).

Reformulando a relação (2.50), obtém-se a equação matricial homogénea

$${}^W\mathbf{T}_C - {}^W\mathbf{T}_T \cdot {}^T\mathbf{T}_R \cdot {}^R\mathbf{T}_C = \mathbf{0} \quad (2.51)$$

que implica $4 \times 3 = 12$ equações homogéneas não lineares, função dos parâmetros linearmente independentes a determinar, número de equações suficiente no limite para permitir o cálculo destas incógnitas. Assumindo que tanto as leituras do sensor de “pose” como da câmara serão afectadas por ruído, podem-se utilizar conjuntos destas leituras para acumular várias relações matriciais como (2.51) para obter-se um problema de mínimos quadrados caracterizado por um sistema não linear do tipo $\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \mathbf{0}$, se se formularem as transformações homogéneas em termos de ângulos de Euler. $\boldsymbol{\phi}$ é o vector de incógnitas a estimar e $\boldsymbol{\theta}$ é o vector de medidas, englobando os parâmetros extrínsecos da câmara e os ângulos de Euler e componentes de translação medidos pelo sensor de “pose”.

Este procedimento e as equações correspondentes são análogas às descritas no excelente trabalho em procedimentos de calibração automáticos para *scanners* de ultrassons “handheld” por R. Prager [Prager et al. 1997]. Este sistema tem de ser solucionado através de métodos iterativos de mínimos quadrados não lineares, como o algoritmo de Levenberg-Marquardt usado neste trabalho (consulte-se o Apêndice D); estes métodos têm como principais desvantagens a sua natureza iterativa e a sua sensibilidade à esco-

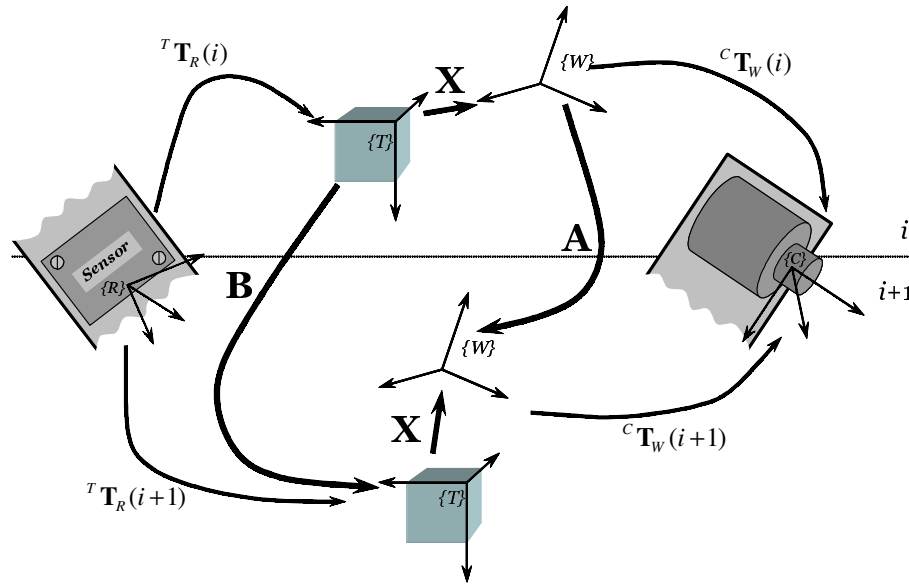


Figura 2.17: Esquematização para a formulação de movimento relativo para a calibração da transformação do referencial do emissor do sensor de “pose” para o sistema de coordenadas absoluto (WCS).

com

$$\mathbf{A} = {}^T\mathbf{T}_R^{-1}(i+1) {}^T\mathbf{T}_R(i) \quad (2.53a)$$

$$\mathbf{B} = {}^C\mathbf{T}_W(i+1) {}^C\mathbf{T}_W^{-1}(i) \quad (2.53b)$$

$$\mathbf{X} = {}^R\mathbf{T}_C \quad (2.53c)$$

para a figura 2.16 e

$$\mathbf{A} = {}^C\mathbf{T}_W^{-1}(i+1) {}^C\mathbf{T}_W(i) \quad (2.54a)$$

$$\mathbf{B} = {}^T\mathbf{T}_R(i+1) {}^T\mathbf{T}_R^{-1}(i) \quad (2.54b)$$

$$\mathbf{X} = {}^W\mathbf{T}_T \quad (2.54c)$$

para a figura 2.17.

Esta formulação foi estudada em profundidade no passado; foi exposta desta forma pela primeira vez por Y. Shiu em Shiu and Ahmad [1989] e tem sido habitualmente associada ao problema de calibração de sistemas sensoriais consistindo numa câmara utilizada para providenciar visão monocular montada num braço robotizado. Este é por vezes denominado problema de *calibração sensor-actuator* ou mais frequentemente

chamado de problema de *calibração “hand-eye”* (traduzido à letra, “mão-olho”) [Daniilidis 1998].

A justificação dada em Shiu and Ahmad [1989] para o uso desta formulação é muito similar aos objectivos do estudo apresentado neste texto na procura de processos de calibração automáticos; estes processos, pelo menos em teoria, ultrapassam claramente os processos mais comuns dependentes da geometria do sistema, de acordo com Shiu visto que “medidas directas são difíceis de obter, visto que podem existir obstáculos a obstruir o percurso de medida, os pontos de interesse podem estar dentro de um sólido e ser inatingíveis, e os sistemas de coordenadas podem diferir grandemente nas suas orientações. A obstrução do percurso de medida pode inclusivamente ser obstruído pela geometria do próprio sensor...”.

A maior parte das soluções automáticas para este problema têm sido de dois géneros:

- Estimação imediata das transformações desconhecidas recorrendo a algoritmos iterativos não lineares como o apresentado no Apêndice D;
- Divisão do processo num método com dois passos, estimando a rotação primeiro através de soluções lineares em forma fechada e posteriormente calculando a translação óptima *a posteriori* usando estes resultados.

As desvantagens do primeiro método são decorrentes da sua natureza (ver subsecção anterior e o Apêndice D — um exemplo deste tipo de solução pode ser encontrado em Horaud and Dornaika [1995]; o segundo método tem a desvantagem óbvia da propagação do erro do primeiro passo para o segundo passo — um exemplo deste tipo de solução, de acordo com Daniilidis [1998], pode ser encontrado em Chou and Kamel [1991].

Para evitar estas questões, a solução usada neste texto não é baseada em nenhum deste tipo de métodos; foi desenvolvida por Konstatinos Daniilidis como apresentado em Daniilidis [1998] e as suas vantagens comparando com outras soluções são óbvias: é um método unificado de mínimos quadrados em forma fechada não iterativo, usando quaterniões duais. Este método apresenta mesmo simplificações conseguidas através da utilização criteriosa das úteis propriedades da noção cinemática de *movimento helicoidal* (movimento em rosca, ou em inglês “*screw*”). Para uma compreensão completa deste processo, descrito no texto subsequente, recomenda-se a consulta preliminar do Apêndice C, no qual são apresentados os fundamentos analíticos e geométricos usados por este método, incluindo alguns teoremas e corolários essenciais.

A equação (2.52) pode ser formulada de forma alternativa, usando o Teorema C.3.1 e o Corolário C.3.4, em notação de quaterniões duais como

$$\check{\mathbf{a}} * \check{\mathbf{q}} = \check{\mathbf{q}} * \check{\mathbf{b}} \quad (2.55)$$

onde $\check{\mathbf{a}}$, $\check{\mathbf{q}}$ e $\check{\mathbf{b}}$ são os quaterniões duais unitários correspondendo às transformações rígidas dadas por $\mathbf{A}$, $\mathbf{X}$ e $\mathbf{B}$, respectivamente. Esta é a formulação mais compacta da relação “*hand-eye*”, visto que os coeficientes de um quaternião dual são 8, ao invés dos 12 das matrizes de transformação homogêneas da equação (2.52).

Esta equação pode ser manipulada usando o Corolário C.3.3 na página 160 de forma a obter

$$\check{\mathbf{a}} = \check{\mathbf{q}} * \check{\mathbf{b}} * \check{\mathbf{q}}^* \quad (2.56)$$

Daniilidis, e antes dele Chen em Chen [1991], introduziram e provaram o chamado Teorema de Congruência do Movimento Helicoidal. Este teorema e a sua demonstração como apresentada por Daniilidis, em notação de quaterniões duais, será apresentada de seguida.

Teorema 2.3.1. *O ângulo e o passo de um movimento helicoidal mantêm-se invariantes sob uma transformação de coordenadas.*

Demonstração. Facilmente se prova, usando as propriedades dos números duais sumariadas no Apêndice C, que a parte escalar de um quaternião dual $\check{\mathbf{a}}$ é dada por $(\check{\mathbf{a}} + \check{\mathbf{a}}^*)/2$, e portanto, usando as propriedades dos quaterniões duais e a equação (2.56), obtém-se

$$\begin{aligned} Sc(\check{\mathbf{a}}) &= \frac{\check{\mathbf{a}} + \check{\mathbf{a}}^*}{2} = \frac{\check{\mathbf{q}} * \check{\mathbf{b}} * \check{\mathbf{q}}^* + \check{\mathbf{q}} * \check{\mathbf{b}}^* * \check{\mathbf{q}}^*}{2} \\ &= \frac{\check{\mathbf{q}} * (\check{\mathbf{b}} + \check{\mathbf{b}}^*) * \check{\mathbf{q}}^*}{2} \\ &= \check{\mathbf{q}} * Sc(\check{\mathbf{b}}) * \check{\mathbf{q}}^* = Sc(\check{\mathbf{b}}) * \check{\mathbf{q}} * \check{\mathbf{q}}^* \\ &= Sc(\check{\mathbf{b}}) \end{aligned}$$

O Teorema C.3.1 implica que um quaternião dual genérico correspondendo a uma transformação rígida é dado por

$$\check{\mathbf{q}} = \begin{bmatrix} \cos\left(\frac{\theta + \varepsilon d}{2}\right) \\ \sin\left(\frac{\theta + \varepsilon d}{2}\right)(\mathbf{\hat{I}} + \varepsilon \mathbf{\hat{m}}) \end{bmatrix}$$

e, portanto, a combinação de ambos os resultados resulta em

$$\cos\left(\frac{\theta_a + \varepsilon d_a}{2}\right) = \cos\left(\frac{\theta_b + \varepsilon d_b}{2}\right)$$

o que, devido ao Corolário C.2.2, é equivalente a

$$\begin{aligned}\cos\left(\frac{\theta_a}{2}\right) &= \cos\left(\frac{\theta_b}{2}\right) \\ d_a \sin\left(\frac{\theta_a}{2}\right) &= d_b \sin\left(\frac{\theta_b}{2}\right)\end{aligned}$$

que, por sua vez, implica que $\theta_a = \theta_b$ e $d_a = d_b$.

□

Deste importante teorema resulta que a equação (2.56) pode ser reduzida, *se e só se* $\theta_{a,b} \neq 0$ e $\theta_{a,b} \neq 2\pi$, a [Daniilidis 1998]

$$\begin{bmatrix} 0 \\ \vec{\mathbf{a}} \end{bmatrix} = \check{\mathbf{q}} * \begin{bmatrix} 0 \\ \vec{\mathbf{b}} \end{bmatrix} * \check{\mathbf{q}}^* \quad (2.57)$$

Desta equação, e aplicando a definição de multiplicação de quaterniões (equação (C.9)) e a definição de multiplicação de números duais (equação (C.20c)), podem derivar-se os resultados seguintes para a parte real e para a parte dual de $\check{\mathbf{a}}$ [Daniilidis 1998]

$$\begin{aligned}\mathring{\mathbf{a}} &= \mathring{\mathbf{q}} * \mathring{\mathbf{b}} * \mathring{\mathbf{q}}^* \\ \mathring{\mathbf{a}}' &= \mathring{\mathbf{q}} * \mathring{\mathbf{b}} * \mathring{\mathbf{q}}'^* + \mathring{\mathbf{q}} * \mathring{\mathbf{b}}' * \mathring{\mathbf{q}}^* + \mathring{\mathbf{q}}' * \mathring{\mathbf{b}} * \mathring{\mathbf{q}}^*\end{aligned}$$

Aplicando a segunda condição para norma unitária do Teorema C.3.1, estas equações tornam-se [Daniilidis 1998]

$$\begin{aligned}\mathring{\mathbf{a}} * \mathring{\mathbf{q}} &= \mathring{\mathbf{q}} * \mathring{\mathbf{b}} \\ \mathring{\mathbf{a}}' * \mathring{\mathbf{q}} &= -\mathring{\mathbf{a}} * \mathring{\mathbf{q}}' + \mathring{\mathbf{q}} * \mathring{\mathbf{b}}' + \mathring{\mathbf{q}}' * \mathring{\mathbf{b}}\end{aligned}$$

Finalmente, todos os termos dos lados direitos podem ser transportados para as partes esquerdas respectivas, resultando nas equações homogêneas seguintes [Daniilidis 1998]

$$\mathring{\mathbf{a}} * \mathring{\mathbf{q}} - \mathring{\mathbf{q}} * \mathring{\mathbf{b}} = 0 \quad (2.58a)$$

$$(\mathring{\mathbf{a}}' * \mathring{\mathbf{q}} - \mathring{\mathbf{q}} * \mathring{\mathbf{b}}') + (\mathring{\mathbf{a}} * \mathring{\mathbf{q}}' - \mathring{\mathbf{q}}' * \mathring{\mathbf{b}}) = 0 \quad (2.58b)$$

Tendo em conta que as partes escalares de cada uma das equações acima é redundante devido a (2.57), temos um total de 6 equações homogéneas com 8 incógnitas que, aplicando a definição e as propriedades da multiplicação de quaterniões, podem ser escritas em formato matricial como apresentado de seguida [Daniilidis 1998]

$$\begin{bmatrix} \vec{a} - \vec{b} & [\vec{a} + \vec{b}]_{\times} & \mathbf{0}_{3 \times 1} & \mathbf{0}_{3 \times 3} \\ \vec{a}' - \vec{b}' & [\vec{a}' + \vec{b}']_{\times} & \vec{a} - \vec{b} & [\vec{a} + \vec{b}]_{\times} \end{bmatrix} \begin{bmatrix} \mathring{\mathbf{q}} \\ \mathring{\mathbf{q}}' \end{bmatrix} = \mathbf{0} \quad (2.59)$$

A matriz 6×8 vai ser denotada como $\mathbf{S}$; o vector de incógnitas tem obviamente 8 dimensões. Denota-se como $[\mathbf{a}]_{\times}$ a matriz anti-simétrica correspondendo ao produto vectorial com $\mathbf{a} = [a_1 \ a_2 \ a_3]^T$ [Hartley and Zisserman 2000, Daniilidis 1998]:

$$[\mathbf{a}]_{\times} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad (2.60)$$

A matrix $[\mathbf{a}]_{\times}$ é singular e $\mathbf{a}$ é o seu vector nulo (direito ou esquerdo). Logo, é definida a menos de um factor de escala [Hartley and Zisserman 2000].

Adicionalmente ao sistema de equações em (2.59), as restrições resultando da definição de quaternião dual unitário aplicam-se a $\mathring{\mathbf{q}}$, e portanto

$$\mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* = \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} = 1 \quad (2.61a)$$

$$\mathring{\mathbf{q}} * \mathring{\mathbf{q}}'^* + \mathring{\mathbf{q}}' * \mathring{\mathbf{q}}^* = \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}}' = 0 \quad (2.61b)$$

Infelizmente, apesar de parecer ter-se 8 equações para 8 incógnitas, os vectores da parte real $\vec{a}$ e $\vec{b}$ são vectores unitários e os vectores da parte dual $\vec{a}'$ e $\vec{b}'$ são perpendiculares aos vectores correspondentes da parte real — logo, 2 das equações são de facto redundantes. Geometricamente isto é compreensível, visto que pelo menos duas rectas são necessárias para se poder estimar movimento 3D a partir das suas correspondências [Daniilidis 1998].

Portanto pelo menos dois movimentos são necessários para conseguir dois eixos centrais das trajectórias helicoidais correspondentes, e cada uma destas trajectórias, como se pode inferir do que foi dito até aqui, têm de estar associadas a ângulos de rotação que impliquem que os seus eixos *não sejam paralelos* [Daniilidis 1998].

Finalmente, os passos necessários para o executar o método em forma fechada de mínimos quadrados para estimar o quaternião unitário dual $\check{\mathbf{q}}$ que representa a transformação $\mathbf{X}$ desconhecida podem ser descritos como se segue:

1. Dados $n \geq 2$ movimentos com eixos de rotação não paralelos, calcular $\check{\mathbf{a}}$ e $\check{\mathbf{b}}$ de leituras da câmara e do sensor de *pose* e verificar se as suas partes escalares são aproximadamente iguais — menores níveis de ruído tenderão a transformar a aproximação destes valores numa igualdade; em seguida, extrair as direcções e momentos dos eixos das trajectórias helicoidais de forma a construir a matriz $6n \times 8$

$$\mathbf{T} = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \dots \\ \mathbf{S}_n \end{bmatrix} \quad (2.62)$$

Visto que no caso de ausência de ruído estas equações surgem de restrições naturais, o espaço-nulo contém pelo menos a solução não-trivial desejada $(\check{\mathbf{q}}, \check{\mathbf{q}}')$. Porém, existe ainda uma solução ortogonal adicional: $(\mathbf{0}_{4 \times 1}, \check{\mathbf{q}})$. Logo, no caso de ausência de ruído, $\mathbf{T}$ tem, no máximo, *característica 6* [Daniilidis 1998].

2. Calcular a SVD de $\mathbf{T}$. Visto que a presença de ruído é quase certa, existirão dois valores singulares tendentes para — mas não iguais a — zero. Logo, tendo $\mathbf{T}$ teoricamente característica 6, a combinação linear dos dois vectores singulares direitos correspondentes $\mathbf{v}_7$ e $\mathbf{v}_8$

$$\begin{aligned} \begin{bmatrix} \check{\mathbf{q}} \\ \check{\mathbf{q}}' \end{bmatrix} &= \lambda_1 \mathbf{v}_7 + \lambda_2 \mathbf{v}_8 \\ &= \lambda_1 \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{y}_1 \end{bmatrix} + \lambda_2 \begin{bmatrix} \mathbf{x}_2 \\ \mathbf{y}_2 \end{bmatrix} \end{aligned} \quad (2.63)$$

será a solução para o problema no sentido dos mínimos quadrados⁹ (consulte-se o Apêndice E.2). Determinem-se, portanto, estes dois vectores a partir da computação da SVD.

⁹Os vectores $\mathbf{x}_1$ e $\mathbf{y}_1$, e $\mathbf{x}_2$ e $\mathbf{y}_2$ são obviamente os vectores coluna de dimensão 4 que compõem $\mathbf{v}_7$ e $\mathbf{v}_8$, respectivamente.

3. As duas restrições necessárias para a computação dos coeficientes λ_1 e λ_2 são obtidos através da aplicação das equações (2.61) à equação (2.63) resultando nas duas equações quadráticas seguintes [Daniilidis 1998]

$$\lambda_1^2 \mathbf{x}_1^T \mathbf{x}_1 + 2\lambda_1 \lambda_2 \mathbf{x}_1^T \mathbf{x}_2 + \lambda_2^2 \mathbf{x}_2^T \mathbf{x}_2 = 1 \quad (2.64a)$$

$$\lambda_1^2 \mathbf{x}_1^T \mathbf{y}_1 + \lambda_1 \lambda_2 (\mathbf{x}_1^T \mathbf{y}_2 + \mathbf{x}_2^T \mathbf{y}_1) + \lambda_2^2 \mathbf{x}_2^T \mathbf{y}_2 = 0 \quad (2.64b)$$

Obviamente λ_1 e λ_2 nunca se anulam e portanto $\mathbf{x}_1^T \mathbf{x}_1 \neq 0$ é assumido de forma a que $\lambda_2 \neq 0$. Introduzindo $s = \lambda_1/\lambda_2$, calcular 2 soluções para s usando as equações (2.64).

4. Substituindo $\lambda_1 = s\lambda_2$ na equação (2.64a) resulta em

$$\lambda_2^2 (s^2 \mathbf{x}_1^T \mathbf{x}_1 + 2s \mathbf{x}_1^T \mathbf{x}_2 + \mathbf{x}_2^T \mathbf{x}_2) = 1 \quad (2.65)$$

Um dos valores para s produz o menor valor para o trinómio $s^2 \mathbf{x}_1^T \mathbf{x}_1 + 2s \mathbf{x}_1^T \mathbf{x}_2 + \mathbf{x}_2^T \mathbf{x}_2$ e portanto corresponde à solução ortogonal ($\mathbf{0}_{4 \times 1}, \hat{\mathbf{q}}$) (a razão para isto é explicada em Daniilidis [1998]). Logo, o valor para s que maximiza o trinómio deve ser usado para calcular λ_2 da equação (2.65) e depois para calcular $\lambda_1 = s\lambda_2$, para se obter a solução desejada ($\hat{\mathbf{q}}, \hat{\mathbf{q}}'$).

5. Finalmente, $\mathbf{v}_7, \mathbf{v}_8, \lambda_1$ e λ_2 podem ser utilizados conjuntamente na equação (2.63) de forma a calcular as soluções real e dual para a estimativa requerida do quaternião unitário dual $\check{\mathbf{q}}$.

2.4 Resultados da simulação sobre a calibração do sistema de medição de “pose”

Uma comparação estatística extensiva foi empreendida entre os dois métodos propostos para a calibração do sensor de “pose”. Este estudo foi realizado usando o MATLAB R12.1, recorrendo às funções de origem `svd` e `lsqnonlin`¹⁰, correndo num PC equipado com um processador AMD Athlon™ XP 2500+ (1.83 MHz), 512 MBytes de RAM e Windows XP Professional.

Antes de apresentar os resultados, algumas considerações respeitando os procedimentos desta simulação serão apresentadas.

¹⁰Usando o algoritmo de Levenberg-Marquardt — veja-se o Apêndice D.

2.4.1 Inserção de ruído

Visto que o processo de simulação tem como intenção estudar o erro de resposta do sistema face à presença de ruído nas leituras simuladas, algum cuidado é necessário na definição das condições em que o ruído é inserido nessas leituras.

Como explicado anteriormente, as leituras do sensor de “pose” e a estimação dos parâmetros extrínsecos da câmara resultam em duas matrizes de leituras, ${}^T\mathbf{T}_R$ e ${}^W\mathbf{T}_C$, respectivamente. Sendo que em condições realistas estas leituras são corrompidas por ruído, este facto precisa de ser levado em conta no processo de simulação.

Seja $\nu = \frac{1}{SNR} \in (0, 1)$ o recíproco do valor estabelecido para o parâmetro razão sinal-ruído. Como metodologia generalizada, adiciona-se ruído aos vectores de translação definindo

$$\vec{\tilde{\mathbf{t}}} = \vec{\mathbf{t}} + \nu \|\vec{\mathbf{t}}\| \mathbf{n}, \quad \mathbf{n} = \begin{bmatrix} n_1 \equiv rand \cap \mathcal{N}(0, 1) \\ n_2 \equiv rand \cap \mathcal{N}(0, 1) \\ n_3 \equiv rand \cap \mathcal{N}(0, 1) \end{bmatrix} \quad (2.66)$$

e aos ângulos de Euler de rotação definindo

$$\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} + \nu \boldsymbol{\theta} w, \quad \begin{cases} w \equiv rand \cap \mathcal{N}(0, 1) \\ \boldsymbol{\theta} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \end{cases} \quad (2.67)$$

onde $rand \cap \mathcal{N}(0, 1)$ denota um número aleatório simulando ruído gaussiano branco com $\mu = 0, \sigma = 1$.

Usando estas definições, o ruído é adicionado directamente a ${}^T\mathbf{T}_R$ — a matriz resultante é denotada por ${}^T\tilde{\mathbf{T}}_R$.

Por outro lado, os parâmetros extrínsecos resultando da calibração da câmara corrompidos por ruído são muito mais difíceis de descrever através de um modelo concreto; muitos autores argumentam inclusivamente que estes são mesmo muito mais sensíveis ao ruído. Para simular esta sensibilidade acrescida, foi adicionado ruído simultaneamente a ${}^W\mathbf{T}_T$ e a ${}^R\mathbf{T}_C$, obtendo-se ${}^W\tilde{\mathbf{T}}_T$ e ${}^R\tilde{\mathbf{T}}_C$, respectivamente, e calculada a matriz final através da reformulação da equação (2.50) como

$${}^W\tilde{\mathbf{T}}_C = {}^W\tilde{\mathbf{T}}_T \cdot {}^T\mathbf{T}_R \cdot {}^R\tilde{\mathbf{T}}_C \quad (2.68)$$

redobrando portanto desta forma os efeitos do ruído.

2.4.2 Métrica para o erro

Primeiramente, reafirme-se que o principal interesse por trás da calibração do sensor de “pose” é estabelecer um modo de relacionar as suas leituras, isto é, a matriz ${}^T\mathbf{T}_R$, com a transformação ${}^W\mathbf{T}_C$ que permite a integração de diferentes varrimentos do *scanner* num WCS comum, $\{W\}$. Com isto em mente, após adicionar ruído e efectuar as simulações, algum tipo de métrica para o erro que relacione o valor exacto ${}^W\mathbf{T}_C$ com a estimativa para esta transformação, denotada por ${}^W\hat{\mathbf{T}}_C$, é necessária para proceder-se à comparação entre procedimentos de calibração.

Para a componente de translação da matriz de transformação de integração tridimensional ${}^W\mathbf{T}_C$, decidiu-se usar o erro absoluto em $\mathbb{R}^3$, $e_{trans} = \left\| {}^W\vec{\mathbf{t}}_C - {}^W\vec{\hat{\mathbf{t}}}_C \right\|$, dado que este também dá uma noção intuitiva da distância entre a posição esperada e a posição estimada para o sistema de coordenadas.

Quanto à orientação, não está definida qualquer medida canónica [Andreff et al. 2000]. Apesar deste facto, para cumprir o objectivo de procura de uma métrica intuitiva para o erro, foi decidido usar-se a definição apresentada abaixo.

Considere-se a igualdade que se segue:

$$\hat{\mathbf{q}} = \mathbf{e} * \hat{\mathbf{q}}, \quad \mathbf{e} = \begin{bmatrix} \cos\left(\frac{\theta_e}{2}\right) \\ \sin\left(\frac{\theta_e}{2}\right) \hat{\mathbf{w}}_e \end{bmatrix} \quad (2.69)$$

Se $\hat{\mathbf{q}}$ representa o quaternião unitário correspondendo à componente de rotação ${}^W\mathbf{R}_C$ e $\hat{\mathbf{q}}$ for o quaternião unitário correspondendo à estimativa respectiva ${}^W\hat{\mathbf{R}}_C$, então $\mathbf{e}$ representa o quaternião unitário correspondendo à rotação definida pelo par ângulo/eixo $(\theta_e, \hat{\mathbf{w}}_e)$ que transforma a rotação estimada no valor exacto. Isto significa que $e_{rot} = |\theta_e|$ pode ser usado como uma métrica de erro para a rotação.

2.4.3 Primeiro conjunto de simulações — resultados para sensibilidade ao ruído em rotação e translação

Os desempenhos de ambos os métodos foram comparados em termos de sensibilidade ao ruído para 10 níveis diferentes de razão sinal-ruído (ν variando de 1% a 10%), com o ruído adicionado primeiro apenas aos ângulos de rotação, como definido na equação (2.67), e depois apenas aos componentes de translação, como definido na equação (2.66).

Para este efeito, foram efectuados 500 ensaios com 20 “poses” simuladas (ou seja, 19 movimentos simulados) para cada um dos 10 níveis em ambas as simulações correspondendo a cada método.

Em cada um dos ensaios, as condições necessárias para a calibração foram simuladas através da geração de valores aleatórios para os ângulos de rotação, variando de 0° a 360° , e para os componentes de translação, indo de -500 mm a 500 mm , para as transformações a estimar, ${}^W\mathbf{T}_T$ e ${}^R\mathbf{T}_C$.

Adicionalmente, nas simulações para a formulação estática, foram usados ângulos e componentes de translação aleatórios tomados da mesma forma para gerar valores iniciais para as estimativas.

Além disto, os movimentos entre cada “pose” foram simulados através da geração de valores aleatórios para os ângulos de rotação, variando de -90° a 90° , e para os componentes de translação, indo de -500 mm a 500 mm , para ${}^T\mathbf{T}_R$ correspondentes a cada instante amostrado. Isto simula movimentos do *scanner* executados dentro de um volume cúbico de 1 m^3 centrado em $\{T\}$. Então, ${}^W\mathbf{T}_C$ é calculado usando a equação (2.50). Finalmente para cada instante é adicionado ruído para simular os valores para ${}^T\tilde{\mathbf{T}}_R$ e ${}^W\tilde{\mathbf{T}}_C$ resultantes das leituras, usando as equações (2.67), (2.66) e (2.68) conforme o necessário.

Foi constatado que a formulação de movimento relativo *nem sempre é estável* — inclusive, a resolução das equações (2.64) levou por vezes não só à obtenção de erros de estimação elevados como a valores complexos para as soluções para $s = \lambda_1/\lambda_2$. Isto significa que, nestes casos, a formulação de movimento relativo *chegou mesmo a divergir*. Logo, para ultrapassar este problema, cada ensaio foi repetido para ambas formulações com novos valores aleatórios até se atingir a convergência; todas as divergências e repetições respectivas foram contabilizadas e foram incorporadas nos resultados para o método do movimento relativo, tanto para este conjunto de simulações como para os conjuntos de simulações seguintes.

Os resultados para este primeiro conjunto de simulações estão apresentados nas figuras 2.18 e 2.19. Neste conjunto em particular, a formulação de movimento relativo divergiu apenas quando se procedeu à adição de ruído aos ângulos de rotação (ver a figura 2.23(a) na página 102); em contrapartida, esta formulação convergiu *sempre* quando o ruído era adicionado apenas à translação.

2.4.4 Segundo conjunto de simulações — resultados para sensibilidade ao ruído com número variável de varrimentos/movimentos simulados

Os desempenhos de ambos os métodos foram comparados em termos de sensibilidade ao ruído simultaneamente presente nos parâmetros de rotação e de translação

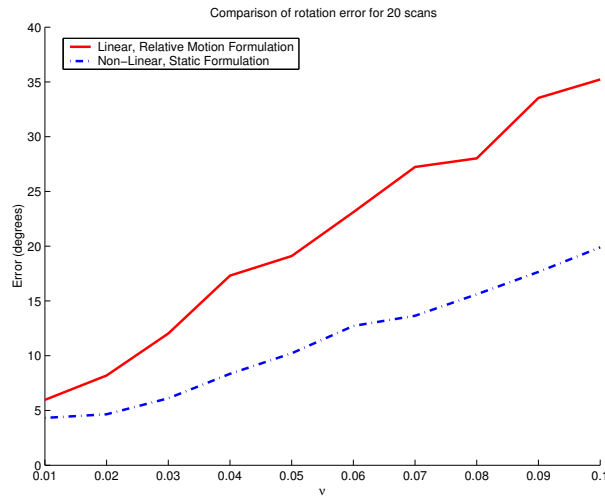
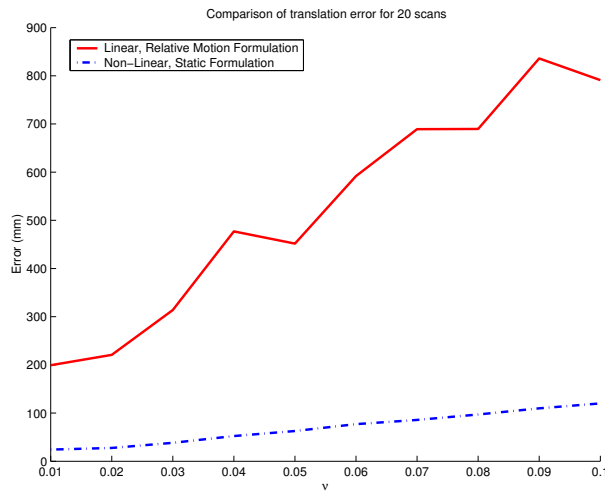
(a) ν para a rotação vs e_{rot} (em graus)(b) ν para a rotação vs e_{trans} (em milímetros)

Figura 2.18: Comparação da sensibilidade ao ruído presente apenas em rotações para 20 “poses” simuladas (ou seja, 19 movimentos simulados). Os movimentos entre “poses” foram simulados através da geração de valores aleatórios para os ângulos de rotação, variando entre -90° e 90° , e para os componentes de translação, indo de -500 mm a 500 mm, para ${}^T\mathbf{T}_R$ para cada “pose”.

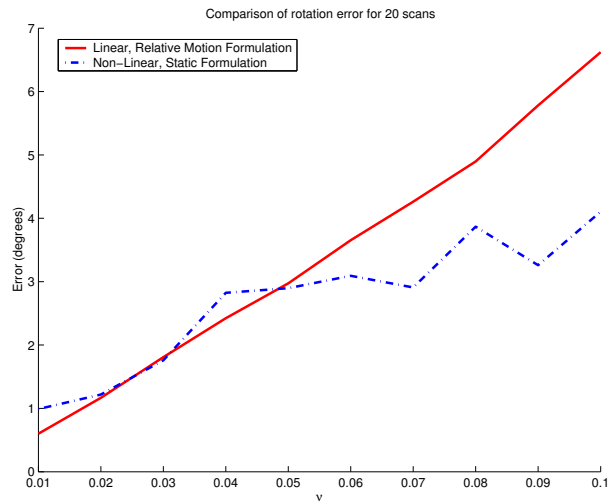
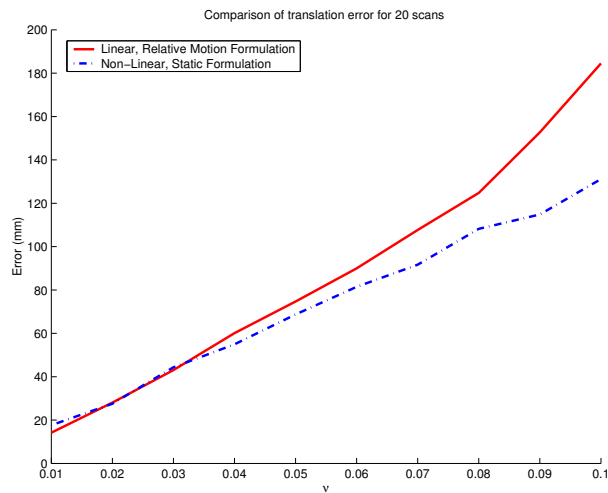
(a) ν para a translação vs e_{rot} (em graus)(b) ν para a translação vs e_{trans} (em milímetros)

Figura 2.19: Comparação da sensibilidade ao ruído presente apenas em translações para 20 “poses” simuladas (ou seja, 19 movimentos simulados). Os movimentos entre “poses” foram simulados através da geração de valores aleatórios para os ângulos de rotação, variando entre -90° e 90° , e para os componentes de translação, indo de -500 mm a 500 mm, para ${}^T\mathbf{T}_R$ para cada “pose”.

(fixando $\nu = 10\%$ para ambos desta vez), para 5 números diferentes de “poses” simuladas, variando entre 10 e 30 em intervalos regulares.

Para este efeito, foram realizados 500 ensaios, mais uma vez, para cada uma das simulações. Em cada um destes ensaios, as condições necessárias para a calibração foram reunidas como no conjunto de simulações anterior.

Os resultados para este segundo conjunto de simulações estão apresentados na figura 2.20, incluindo uma comparação em termos de tempo dispendido. Os resultados para o número de divergências para a formulação em movimento relativo estão apresentados na figura 2.23(b) na página 102.

2.4.5 Terceiro conjunto de simulações — resultados para sensibilidade ao ruído com rotações de ângulos com amplitude variável e translações com comprimento variável

Os desempenhos de ambos os métodos foram comparados em termos de sensibilidade ao ruído em rotação e translação (de novo $\nu = 10\%$ para ambas), para 20 “poses” simuladas.

Primeiramente, a simulação foi realizada com os movimentos tendo rotações $\mathbf{R}_m$ referidas ao sistema de coordenadas local do receptor do sensor de “pose” com eixos orientados aleatoriamente e 19 amplitudes de ângulo Φ diferentes, indo de 0 a 180° em intervalos regulares. A orientação para as leituras do sensor de “pose” são dadas portanto por

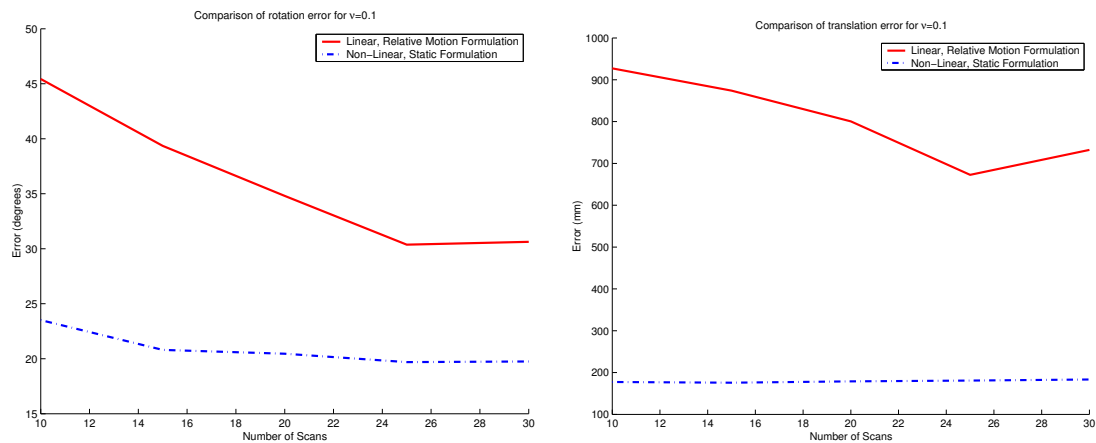
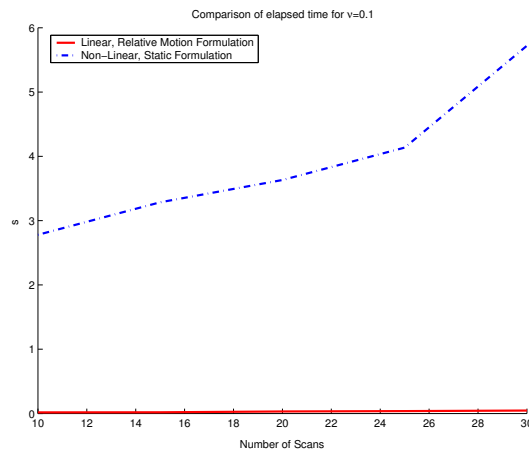
$${}^T\mathbf{R}_R(n+1) = \mathbf{R}_m(\Phi(n)) {}^T\mathbf{R}_R(n), \quad n = 1..19 \quad (2.70)$$

com n representando o número do instante de uma determinada “pose”.

As translações foram simuladas como anteriormente, de modo a ter de novo movimentos dentro de um volume cúbico de 1 m^3 centrado em $\{T\}$.

Seguidamente, outra simulação foi realizada para ambos os métodos de calibração, mas desta vez para movimentos simulados com rotações $\mathbf{R}_m$ com ângulos de amplitude fixa (90°) e eixo orientado aleatoriamente. Mais uma vez, os efeitos destas rotações são calculados usando a equação (2.70). Nesta circunstância, 11 vectores de translação diferentes $\vec{\mathbf{t}}_m$ com direcções e normas l variáveis, indo de 0 a 100 mm em intervalos regulares, foram usados, de novo referidos ao sistema de coordenadas local do receptor do sensor de “pose”. Os efeitos destas translações são dadas por

$${}^T\vec{\mathbf{t}}_R(n+1) = \mathbf{R}_m(\Phi(n)) {}^T\vec{\mathbf{t}}_R(n) + \vec{\mathbf{t}}_m(l(n)), \quad n = 1..19 \quad (2.71)$$

(a) Número de “poses” vs e_{rot} (em graus)(b) Número de “poses” vs e_{trans} (em milímetros)

(c) Número de “poses” vs tempo dispendido (em segundos)

Figura 2.20: Comparação da sensibilidade ao ruído ($\nu = 10\%$ tanto para rotações como para translações) para número variável de “poses”/movimentos simulados/os. Os movimentos entre “poses” foram simulados através da geração de valores aleatórios para os ângulos de rotação, variando entre -90° e 90° , e para os componentes de translação, indo de -500 mm a 500 mm, para ${}^T\mathbf{T}_R$ para cada “pose”.

Para este efeito, para cada uma destas simulações, mais uma vez foram realizados 500 ensaios. Para tudo o que não tenha sido explicado nos parágrafos anteriores, em cada um destes ensaios as condições necessárias para a simulação da calibração foram criadas como nos conjuntos de simulações antecedentes.

Os resultados para este terceiro conjunto de simulações estão apresentados nas figuras 2.21 e 2.22. Os resultados para o número de divergências para a formulação de movimento relativo estão apresentados nas figuras 2.23(c) e 2.23(d) na página 102.

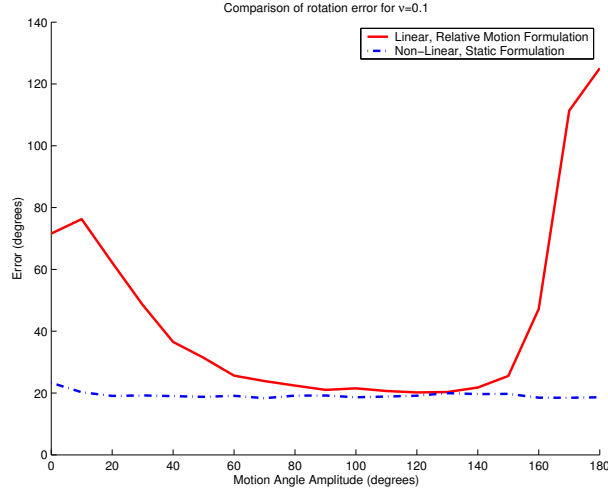
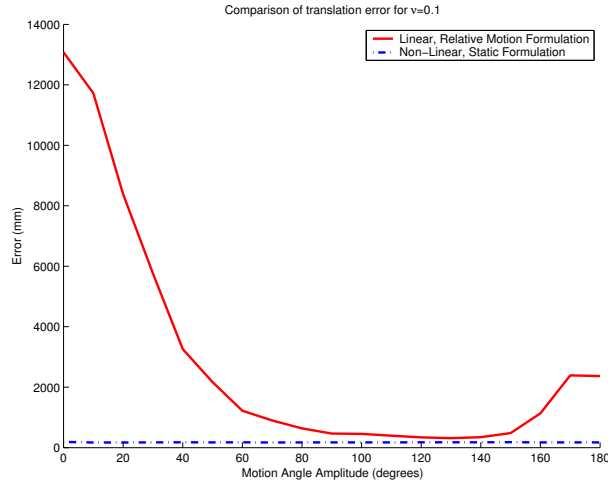
(a) ν para a rotação vs e_{rot} (em graus)(b) ν para a rotação vs e_{trans} (em milímetros)

Figura 2.21: Comparação da sensibilidade ao ruído face a variações na amplitude dos ângulos de rotação para 20 “poses” simuladas (ou seja, 19 movimentos simulados). Movimentos entre “poses” foram simulados através da geração de valores aleatórios para o eixo e 19 diferentes amplitudes para o ângulos de cada rotação, variando entre 0° e 180° . As translações foram simuladas de forma a ter movimentos dentro de um volume cúbico de 1 m^3 , centrado em $\{T\}$.

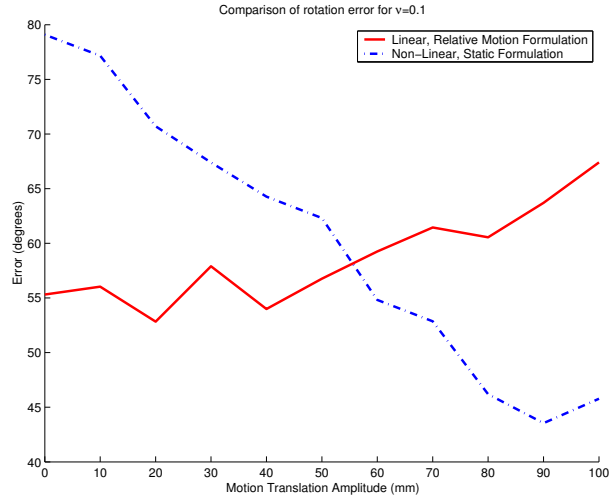
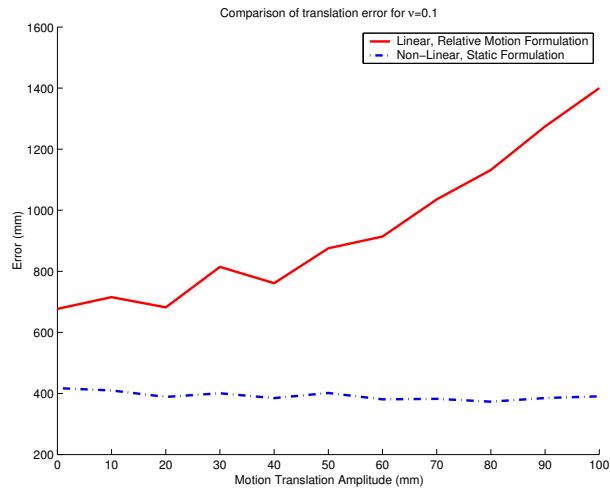
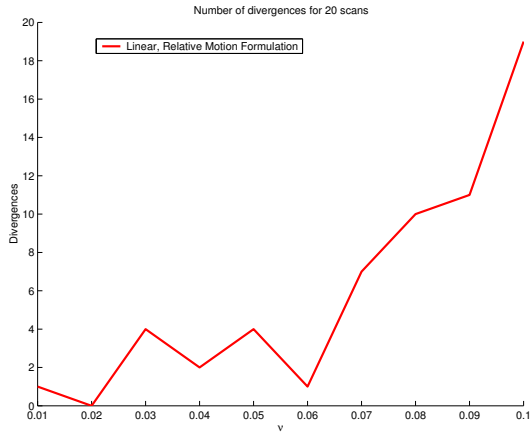
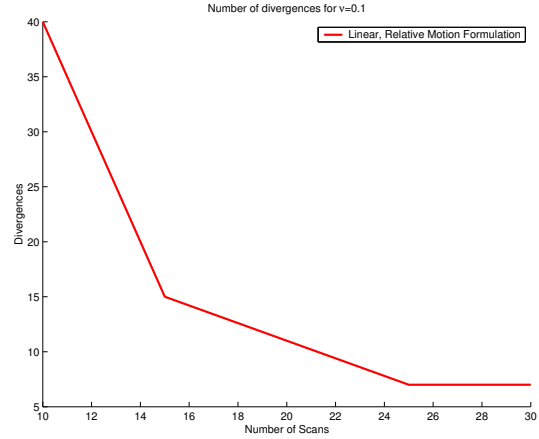
(a) ν para a translação vs e_{rot} (em graus)(b) ν para a translação vs e_{trans} (em milímetros)

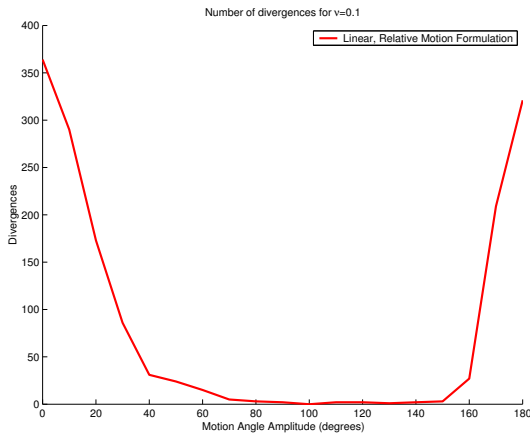
Figura 2.22: Comparação da sensibilidade ao ruído face a variações na norma das translações para 20 “poses” simuladas (ou seja, 19 movimentos simulados). Movimentos entre “poses” foram simulados através da geração de valores aleatórios para o eixo de cada rotação, com ângulo respectivo de amplitude fixa (-90°), e 11 vectores de translação com direcção aleatória e normas variantes, indo de 0 a 100 mm, sendo todos estes valores referidos ao sistema de coordenadas local do receptor do sensor de “pose”.



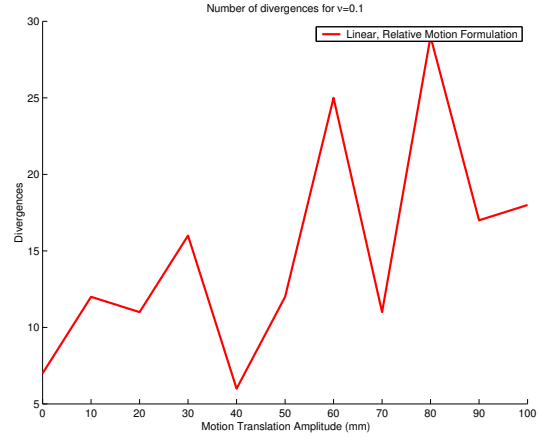
(a) Número de divergências usando a formulação de movimento relativo para ruído em rotação apenas (20 “poses”/19 movimentos).



(b) Número de divergências usando a formulação de movimento relativo para ruído com $\nu = 10\%$, tanto em rotação como em translação, e número variável de “poses”/movimentos simuladas/os.



(c) Número de divergências usando a formulação de movimento relativo para ruído com $\nu = 10\%$, tanto em rotação como em translação, e número variável de amplitudes de ângulos de rotação (20 “poses”/19 movimentos).



(d) Número de divergências usando a formulação de movimento relativo para ruído com $\nu = 10\%$, tanto em rotação como em translação, e número variável de normas para o vector de translação (20 “poses”/19 movimentos).

Figura 2.23: Número de divergências usando a formulação de movimento relativo para os diversos conjuntos de simulações. As condições de simulação de movimento são como as descritas nas figuras 2.18, 2.20, 2.21 e 2.22, respectivamente.

2.5 Calibração do protótipo laboratorial

De seguida, serão apresentados os procedimentos de calibração na sua vertente prática e os resultados da sua aplicação na calibração do protótipo criado em laboratório descrito no Apêndice B.

2.5.1 Descrição dos procedimentos de calibração

No texto que se segue, serão apresentados os pormenores práticos dos procedimentos de calibração do sistema Tele-3D, nomeadamente no que respeita aos passos a seguir e ao processamento de imagem envolvido.

Calibração da câmara e do sensor de pose

Para todo o processo de calibração foi utilizado o pacote de software em Visual C++ *Tele-3D Grabber* para a captura das imagens requeridas e das leituras do sensor de “pose”. Como já foi referido, foi usado o software de calibração de câmaras do Intel OpenCV para a calibração da câmara a partir das imagens capturadas — veja-se a figura 2.24 na página seguinte.

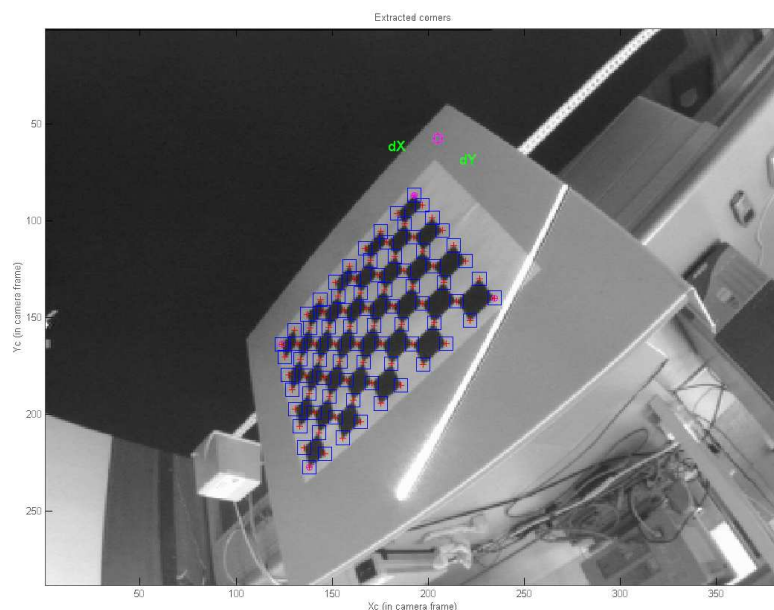
Para a calibração do sensor de “pose” foi produzido software em MATLAB usando a formulação estática descrita anteriormente.

Os passos para a calibração em laboratório destes dois componentes do sistema pode resumir-se aos seguintes passos:

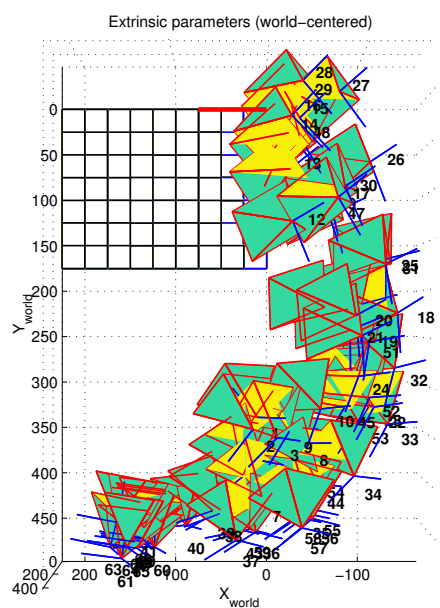
1. capturar imagens do padrão em xadrez em diferentes “poses” e respectivas leituras do sensor;
2. calibrar a câmara usando o software OpenCV;
3. calibrar o sensor de “pose” usando os resultados da calibração da câmara em termos de parâmetros extrínsecos e as leituras do sensor obtidas no ponto 1.

Calibração do laser

A calibração do plano de luz projectado do laser foi originalmente pensada para ser feita com as mesmas imagens (entretanto corrigidas em termos de distorção) e dados recolhidos na calibração da câmara e sensor de “pose”, deixando o projector ligado ao longo das capturas de imagem, como tinha sido apresentado na figura 2.14 na página 78 e é mostrado na prática na figura 2.25 na página 105.



(a) Imagem da detecção de cantos do padrão em xadrez.



(b) Figura mostrando o sistema de coordenadas privado da câmara para várias “poses” apresentadas relativamente ao plano de calibração.

Figura 2.24: Imagens exemplo de capturas efectuadas com o *Tele-3D Grabber* sendo usadas na calibração da câmara através do software do OpenCV.

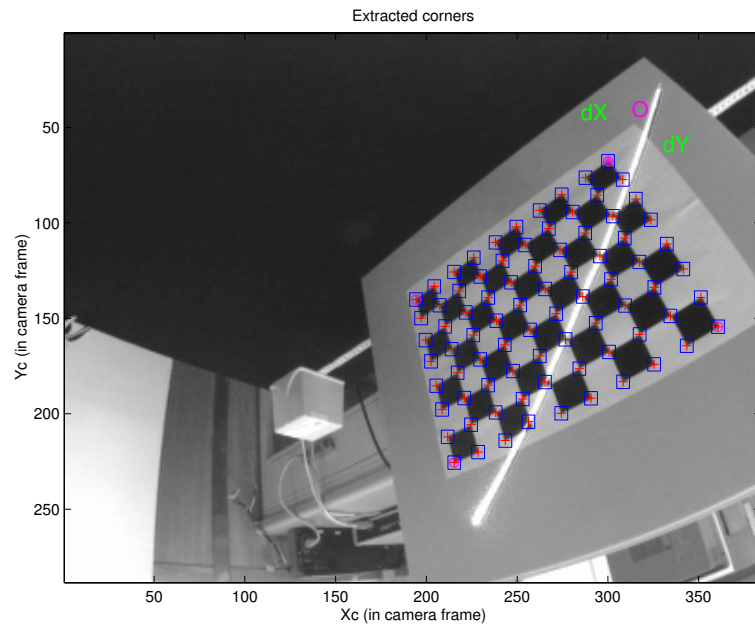


Figura 2.25: Imagem tirada durante tentativa de calibração simultânea da câmara, do sensor de “pose” e do projector laser, mostrada durante o processo de calibração da câmara. Veja-se o erro cometido pelo algoritmo de detecção dos cantos do padrão em xadrez nas zonas atravessadas pelo perfil laser.

Como se pode ver pela figura 2.25, porém, o projector laser estar ligado durante esta operação diminui drasticamente a qualidade da calibração da câmara¹¹ e, consequentemente, do sensor de “pose” e do próprio projector laser, visto que os parâmetros extrínsecos estimados entretanto para cada imagem estariam por si só bastante afectados por este erro acrescido na calibração da câmara! Como desvantagem adicional, pode-se referir a maior dificuldade no processamento da imagem na segmentação do perfil laser, visto que este processo impossibilita o uso do filtro de interferência.

Por estas razões, foi decidido separar na prática a calibração do laser do resto dos procedimentos de calibração — desta forma, o primeiro passo de calibração seria recolher dados iniciais para calibrar a câmara e o sensor de “pose” e, posteriormente, recolher dados adicionais usando o filtro para a calibração do laser.

Para este último processo, existiriam duas alternativas para o calibrador:

- Tirar-se duas imagens por pose, uma com o projector laser ligado e outra desligado, visualizando sempre o padrão de calibração, de forma a poder determinar os parâmetros extrínsecos da câmara e consequentemente obter para cada pose

¹¹Mais concretamente, uma calibração que normalmente daria $.2 \text{ pixel}$ de erro de reprojeção é prejudicada de tal forma pela presença do perfil de luz que passa a dar um erro de reprojeção de $.7 \text{ pixel}$!

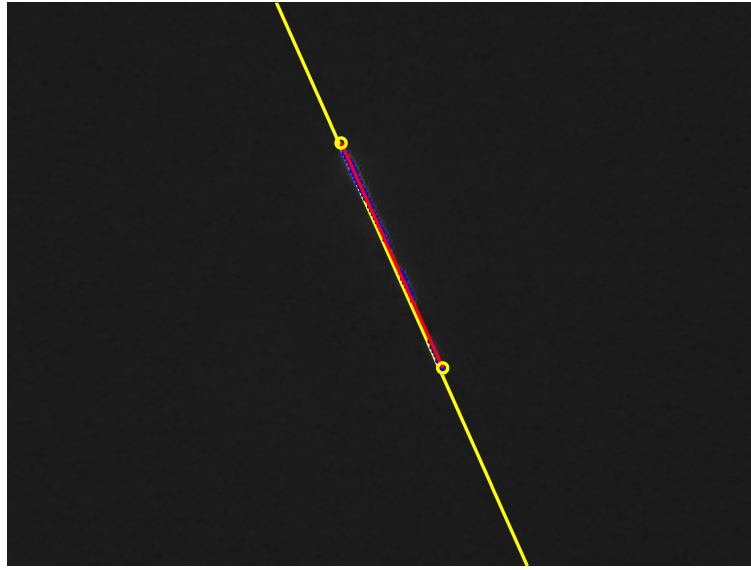


Figura 2.26: Imagem da calibração do projector laser e do processamento de imagem envolvido. A linha a amarelo representa a primeira estimativa para a recta dada pela Transformada de Hough, os pontos a azul são os pontos usados para a interpolação linear, a linha a vermelho representa a estimativa final para o segmento de recta e as circunferências a amarelo salientam os pontos extremos do segmento usados para formar a matriz da equação (2.48).

as transformações ${}^C\mathbf{T}_W$ requeridas para resolver a equação (2.47), e processar o perfil laser, separadamente.

- Usar-se as leituras do sensor de “*pose*” e a sua calibração prévia para fornecer as mesmas transformações ${}^C\mathbf{T}_W$. Esta última alternativa tem a desvantagem de o erro da calibração do sensor de “*pose*” se propagar no erro de calibração do projector laser; porém, tem duas grandes vantagens — permite um exercício muito mais amplo dos graus de liberdade do “scanner”, visto não ser necessário manter o padrão na imagem, permite o uso do filtro de interferência durante a calibração, e além disso facilita muito a redundância, visto ser um processo muito mais automático.

Na figura 2.26, encontra-se apresentada uma imagem de um perfil de laser já segmentado no decurso da calibração do projector. A segmentação do perfil é executada procedendo aos passos descritos seguidamente:

1. criação de uma máscara resultando da binarização da imagem original através da aplicação de um limiar adaptativo [Gonzalez and Woods 2002, Shapiro and Stockman 2001], usando-a em seguida para isolar os pontos de interesse através de uma operação de produto lógico;

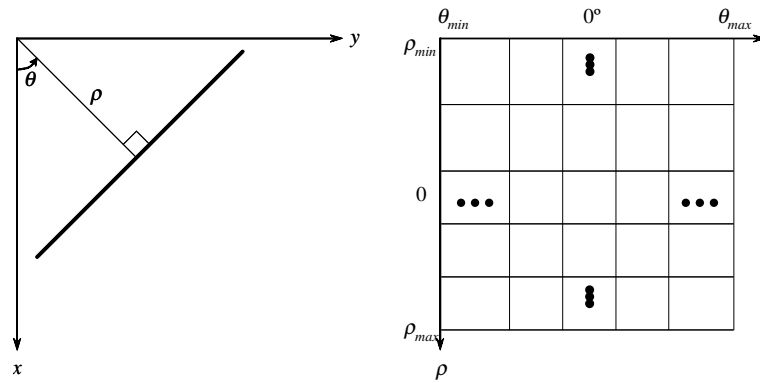


Figura 2.27: Transformada de Hough. À esquerda pode ver-se a representação normal de uma recta e à direita a subdivisão do plano $\rho\theta$ em células, formando a tabela de acumulação.

2. estimação da recta de perfil através do recurso à Transformada de Hough da imagem para fornecer uma primeira estimativa livre de *outliers*, seguida de uma interpolação linear usando os pontos a uma distância de 5 *pixel* dessa primeira estimativa;
3. determinação dos pontos extremos do segmento de recta para formar a matriz da equação (2.48).

Veja-se a figura 2.27: a Transformada de Hough estabelece uma transformação do domínio do plano euclidiano XY para o domínio discreto do plano $\rho\theta$ através do uso equação normal da recta $x \cos(\theta) + y \sin(\theta) = \rho$. Esta transformação dá-se de tal forma que Q pontos pertencentes à recta geram Q curvas sinusoidais intersectando-se em (ρ_i, θ_j) no espaço de parâmetros; através da incrementação de θ e resolvendo para ρ geram-se Q entradas numa tabela de acumulação $A(i, j)$ associadas com a célula determinada por (ρ_i, θ_j) . A determinação de células de interesse na tabela de acumulação — máximos locais ou globais — permite uma segmentação virtualmente livre de *outliers* de rectas importantes na imagem (para mais informação sobre a Transformada de Hough consulte-se [Gonzalez and Woods 2002, Shapiro and Stockman 2001]).

2.5.2 Resultados da calibração

Foi efectuada uma calibração do protótipo (Apêndice B) e chegou-se aos resultados apresentados em seguida.

Calibração da câmara

Parâmetros: Listagem dos parâmetros fixos da câmara:

- Parâmetros intrínsecos, todos medidos em pixel:

$$f_x = 508,667466138224260$$

$$f_y = 484,771697394041610$$

$$c_x = 350,071619468620950$$

$$c_y = 293,610609020277820$$

- Factores de distorção radial e tangencial:

$$k_1 = -0,320572115748301$$

$$k_2 = 0,099703917567058$$

$$p_1 = -0,009317538174552$$

$$p_2 = 0,019130748099614$$

Listagem de resultados (erro, etc.): *Output* do calibrador OpenCV:

`%-- Focal length uncertainty:`

`fc_error = [ 4.809369789589532 ; 5.021308770483099 ];`

`%-- Principal point uncertainty:`

`cc_error = [ 8.603408250523470 ; 5.332534187183376 ];`

`%-- Distortion coefficients uncertainty:`

`kc_error = [ 0.010858513491872 ; 0.010914146381690 ;`

`0.001280482120130 ; 0.003554243775997 ];`

`%-- Image size:`

`nx = 768;`

`ny = 576;`

`% Number of calibration images`

`n_ima = 15;`

Calibração do sensor de “pose”**Parâmetros:** Matrizes de transformação (componentes de translação em milímetros):

$${}^W\mathbf{T}_T = \begin{bmatrix} -0,0786 & -0,9911 & 0,1075 & 116,2456 \\ 0,0348 & -0,1105 & -0,9933 & 185,7426 \\ 0,9963 & -0,0743 & 0,0432 & -315,9158 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$${}^R\mathbf{T}_C = \begin{bmatrix} 0,6887 & -0,1123 & -0,7163 & 104,4547 \\ -0,2452 & -0,9658 & -0,0844 & -57,9762 \\ -0,6823 & 0,2338 & -0,6926 & -130,7045 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Listagem de resultados (erro, etc.): *Output* do calibrador Tele-3D:

Mean rotation error (degrees):

MeanError_theta = 2.719212

Mean translation error (mm):

MeanError_t = 12.60196

Calibração do projector laser**Parâmetros:** Listagem dos parâmetros do plano-luz: $a = -0,0069103$ $b = -0,084969$ $c = 0,99636$ $d = -353,32$ **Listagem de resultados (erro, etc.):** *Output* do calibrador Tele-3D:

Mean triangulation error (mm):

Error = 12.6258

2.6 Conclusões

Tentou-se neste capítulo apresentar uma solução que se acredita satisfazer os requisitos apresentados na secção introdutória, e que além disso permite, devido à sua flexibilidade, que o seu grau de automatização seja apenas dependente da sua realização algorítmica. Esta solução representa a fundação para o funcionamento do *scanner* Tele-3D.

Os modelos e os procedimentos de calibração, foram validados com sucesso, tanto durante as simulações, como no decorrer dos testes laboratoriais. Em consequência das dificuldades inerentes ao tipo de *scanner* 3D em estudo neste texto, porém, os resultados laboratoriais apresentados para as calibrações não são, só por si, suficientes para garantir uma boa digitalização tridimensional. Esta importante questão, que é o fulcro do estudo apresentado nesta tese, irá ser tratada no capítulo que se segue.

Capítulo 3

Reconstrução Tridimensional — Registo e Integração de Perfis

A primeira questão que surge na reconstrução tridimensional é o varrimento da cena de forma a obter uma reconstrução inicial aproximada. Esta primeira parte da reconstrução, bastante dependente de metodologias de processamento de imagem, será apresentada na primeira parte do capítulo. O registo tridimensional e subsequente integração dos dados serão tratados seguidamente.

Por fim, os resultados da reconstrução tridimensional a partir dos pontos integrados (sem ordenação espacial) serão apresentados e as devidas conclusões serão tiradas.

Para efectuar os testes laboratoriais apresentados neste capítulo, foram utilizados os dois objectos apresentados na figura 3.1 na página seguinte.

3.1 Varrimento e processamento de uma cena tridimensional

O processamento da cena tridimensional parte da ideia simples de amostrar essa cena a cada instante através da captura de uma imagem e uma leitura de “*pose*” em n instantes de um varrimento da cena com o *scanner*; todos os pontos para cada perfil amostrado são sujeitos a uma integração inicial aproximada e armazenados numa estrutura de dados conveniente. O software usado neste trabalho está sumariamente apresentado no Apêndice B.2; nesta secção, descrever-se-ão essencialmente os procedimentos de processamento de imagem, essenciais neste estágio de reconstrução 3D.

Num instante i , uma imagem (usando o filtro de interferência e corrigida em termos de distorção) e a leitura de “*pose*” correspondente são capturadas. A imagem passa



Figura 3.1: Objectos digitalizados pelo sistema Tele-3D — um manequim e uma pequena estátua de Nossa Senhora.

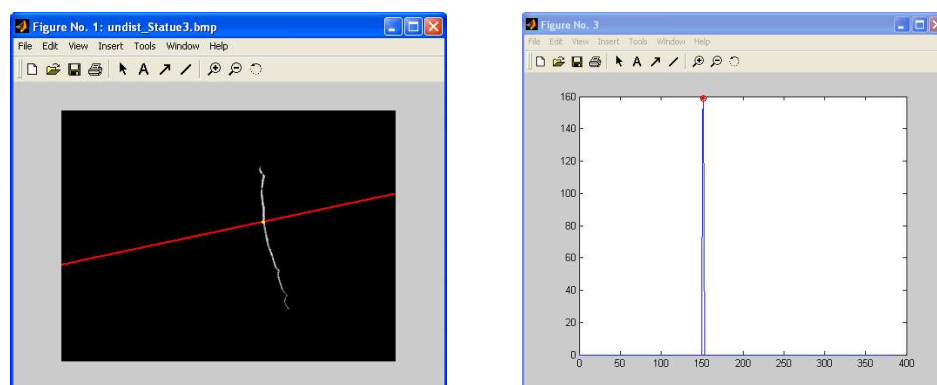


Figura 3.2: Processamento e amostragem do perfil laser durante o varrimento de uma cena. À esquerda pode ver-se a secção de amostragem para o histograma, apresentado à direita. No histograma pode ver-se uma circunferência a vermelho indicando o seu valor máximo e um asterisco da mesma cor indicando a estimativa com precisão sub-pixel determinada para o pico do perfil — apesar de próximos, não são coincidentes. Também se pode ver, na imagem à esquerda, a localização estimada para o ponto amostrado a partir da determinação deste pico. A recta de amostragem percorrerá progressivamente o perfil de uma ponta a outra de modo a amostrar automaticamente todos os pontos que for possível.

então por vários estágios de processamento (veja-se a figura 3.2 na página anterior):

1. binarização através da aplicação de um limiar adaptativo (consulte-se Gonzalez and Woods [2002], Shapiro and Stockman [2001]) para a determinação de uma máscara para separar os pontos de interesse através da aplicação de um produto lógico na imagem original;
2. determinação do eixo principal da zona de interesse para determinar uma estimativa aproximada da direcção global do perfil;
3. varrimento da imagem através da função interna do MATLAB `improfile`, que produz um histograma bidimensional amostrado da imagem ao longo de uma recta, na direcção perpendicular à que foi determinada no passo anterior;
4. determinação para cada histograma do pico do perfil de luz com precisão sub-pixel considerando-o aproximadamente gaussiano e aplicando o procedimento descrito em Trucco et al. [1994].
5. para todos os pontos do perfil amostrados no ponto anterior, aplicação da equação de triangulação (2.14) para determinar os pontos no mundo tridimensional.

Podem ver-se as estimativas iniciais para a integração tridimensional dos objectos conseguida desta forma representadas na figura 3.3 na página seguinte. No cálculo das triangulações escolheu-se, por conveniência de versatilidade, o sistema de coordenadas associado ao emissor do sensor de “pose” como coincidente com o WCS (ou seja, ${}^W\mathbf{T}_T = \mathbf{I}_{4 \times 4}$, com $\mathbf{I}$ representando a matriz identidade 4×4).

3.2 Registo e integração tridimensional de perfis de luz

Vários problemas foram identificados na secção introdutória desta dissertação para o registo e integração tridimensionais. As soluções concebidas ao longo deste trabalho no contexto da reconstrução tridimensional recorrendo ao uso de um *scanner handheld* serão apresentados de seguida.

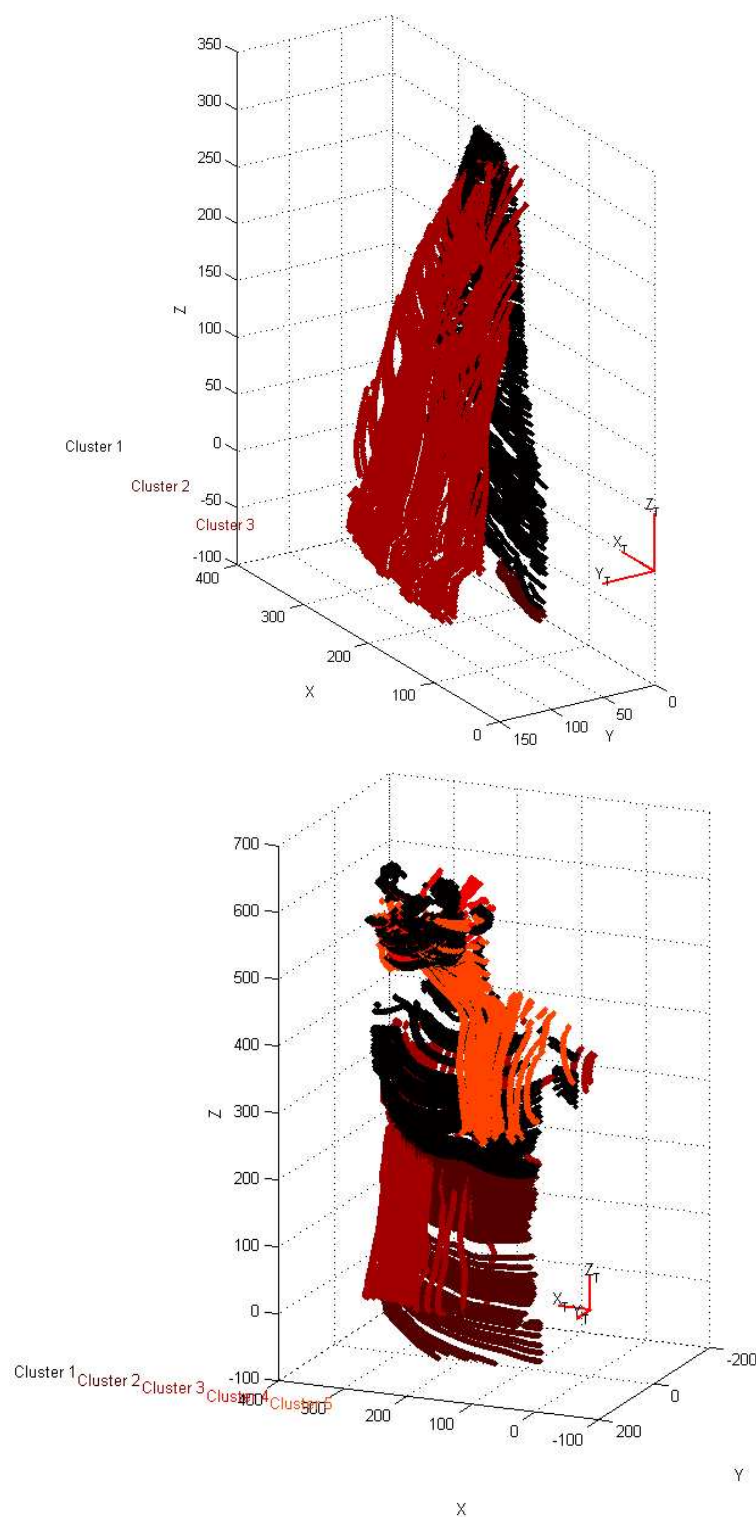


Figura 3.3: Estimativas iniciais para a integração tridimensional. Podem ver-se: na origem, a vermelho e etiquetado com $\{T\}$ o WCS, fixo para todos os instantes, e os perfis associados em vários grupos (denominados de “clusters”) de cores diversas, conforme a orientação do plano de luz correspondente. Note-se também o erro presente na orientação dos perfis.

3.2.1 Uma visão global sobre registo de perfis — correspondência entre perfis e correspondência entre pontos

Uma das diferenças principais no processo de registo entre *scanners* 3D normais e *scanners handheld* é que, quando realizado para os primeiros, conjuntos de amostras de vários perfis tirados de *cada ponto de vista* podem ser formados como estando *localmente registados* dentro desse conjunto. Isto significa que na realidade é possível assumir que sub-conjuntos da superfície no seu todo foram amostrados em cada ponto de vista e que, como tal, a integração pode ser conseguida através do **registo de várias superfícies sobrepostas** obtidas através de interpolação, usando os pontos de controlo fornecidos pelos perfis amostrados. Esta é a base da maior parte da investigação em registo 3D (consulte-se Chen and Medioni [1992], Rusinkiewicz and Levoy [2001]).

Em contrapartida, existem duas grandes dificuldades respeitando o processo de registo tridimensional que implicam uma grande perda de redundância útil quando se considera o *handheld 3D scanning*:

1. Nunca mais que um perfil será amostrado de um mesmo ponto de vista, logo nenhum perfil *pode ser considerado registado de antemão* com qualquer outro perfil;
2. Cada perfil pode apenas ser registado com perfis que o *intersectem* dentro dos limites da superfície visível do objecto da cena a ser amostrada [Hébert and Rioux 1998].

Isto significa que é apenas possível registar **perfis intersectantes** em oposição a *retalhos* de superfície sobrepostos e que o espaço de características de características é obviamente composto dos pontos de intersecção desses perfis.

Assuma-se por agora que a segunda questão levantada acima foi resolvida de forma a que foram determinadas as correspondências entre os perfis intersectantes. Nesse caso, uma solução preliminar para o problema de registo 3D para um perfil em particular com o seu conjunto de perfis intersectantes, na linha de algoritmos bem conhecidos como o ICP (*Iterative Closest Point*), poderia ser dada pela equação seguinte (*cf.* Chen and Medioni [1992], Hébert and Rioux [1998])

$$e = \sum_{i=1}^N \left\| {}^W\mathbf{T}_{C_n} \mathbf{P}_{ni} - {}^W\mathbf{T}_{C_i} \mathbf{P}_{in}'' \right\|^2,$$

$$\mathbf{P}_{in}'' = \mathbf{P} \mid \min_{\mathbf{P} \in p_i} \left\| {}^W\mathbf{T}_{C_n} \mathbf{P}_{ni} - {}^W\mathbf{T}_{C_i} \mathbf{P} \right\| \quad (3.1)$$

onde e é a função objectivo a ser minimizada no sentido dos mínimos quadrados. Esta função representa a soma dos quadrados das distâncias euclidianas entre elementos de cada par, índice i , dos pontos de intersecção correspondentes. Estes pares consistem nos pontos de intersecção $\mathbf{P}_{ni}$, tirados do perfil p_n considerado a partir do ponto de vista n , para o qual a transformação ${}^W\mathbf{T}_{C_n}$ terá de ser estimada para atingir o registo, e nos pontos $\mathbf{P}_{in}''$ dos perfis intersectantes p_i considerados no ponto de vista i (emparelhados com as transformações ${}^W\mathbf{T}_{C_i}$, assumidas como dadas), tomados como correspondentes através de emparelhamento baseado na determinação das distâncias euclidianas mínimas.

Apesar de, como se poderia suspeitar através da análise da equação (3.1), o registo local de um perfil com o conjunto dos seus perfis intersectantes levar a uma solução possivelmente linear, para se conseguir um registo global de todos os perfis uns com os outros não existe outra opção que não seja uma solução iterativa [Hébert and Rioux 1998]. Poder-se-á, todavia, escolher entre usar uma função objectivo *generalizada* a minimizar e efectuar optimização global ou executar optimização local até que algum limiar requerido seja atingido.

Esta escolha está, porém, limitada pela solução do registo local, e a equação (3.1) poderá não ser a mais indicada, visto que as particularidades da geometria de geração de perfis não está a ser levada em conta.

Alguns estudos foram conduzidos por Hébert e Rioux, documentados em Hébert and Rioux [1998], que tentam solucionar o problema de registo local através do aproveitamento da geometria de geração de perfis indirectamente através do uso das propriedades do plano tangente à superfície amostrada de forma a achar correspondências entre os pontos de perfis intersectantes.

A nossa solução, descrita no texto que se segue, também tenta explorar os benefícios da geometria de geração de perfis para conseguir melhores correspondências entre pontos, mas desta feita usando o conhecimento de que pontos de intersecção entre perfis pertencem também às linhas de intersecção entre planos de luz, bastante mais fáceis de determinar.

3.2.2 Optimização local

Seguidamente, descrever-se-ão os princípios e a metodologia desenvolvida neste trabalho para a optimização local no processo de registo tridimensional de perfis.

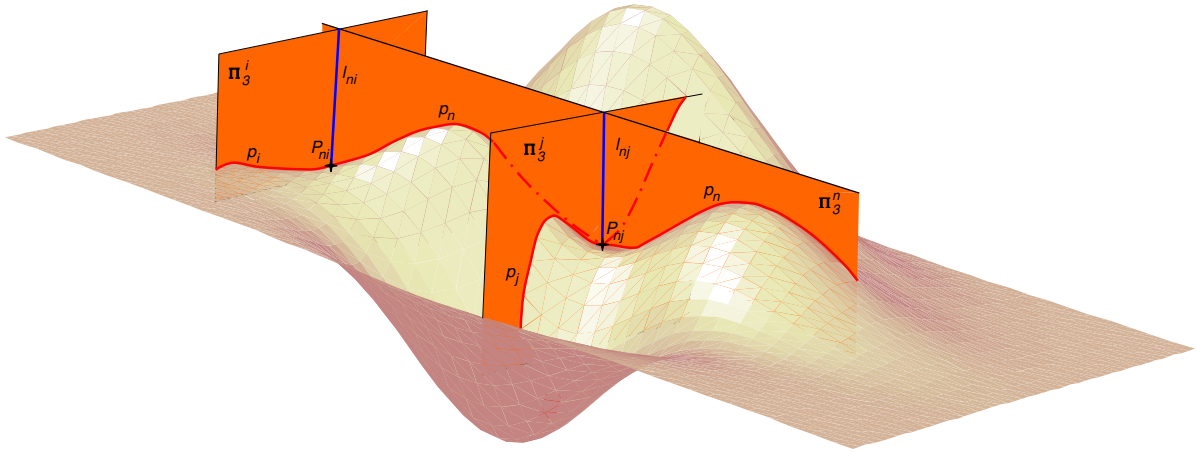


Figura 3.4: Registo de planos de luz. Um plano pode ser visto intersectando dois outros planos; as suas rectas de intersecção, os perfis correspondentes e os pontos de intersecção entre esses perfis estão apresentados nesta figura.

Escolhendo a função objectivo para optimização local — tomando partido da recta de procura dada pela intersecção entre planos de luz

Na figura 3.4, apresenta-se um exemplo de um varrimento hipotético (visto que todas os valores exactos para as “poses” de todas as entidades assumem-se como sendo conhecidas), consistindo em um plano de luz intersectando dois outros planos, todos estes intersectando uma superfície ondulante gerando perfis. Nesta figura é perfeitamente visível a redundância útil que se pode conseguir levando em conta a geometria do varrimento, visto que se estão a considerar várias entidades diferentes que intersectam nos mesmos pontos de registo: dois planos de luz, uma recta e dois perfis por cada um destes pontos.

Em algoritmos como o ICP e as suas variantes (comparadas num excelente estudo apresentado em Rusinkiewicz and Levoy [2001]), as correspondências entres pontos são encontradas, normalmente, seja determinando o ponto amostrado mais próximo através do cálculo da distância ao longo de um percurso directo até ao ponto de controlo considerado (veja-se, por exemplo, Besl and McKay [1992]), seja encontrando o ponto amostrado mais próximo através do cálculo da distância ao longo de um percurso perpendicular ao plano tangente à superfície (veja-se, por exemplo, Chen and Medioni [1992]). Note-se que a expressão “ponto amostrado mais próximo” implica que uma sub-amostragem, geralmente recorrendo a uma grelha regular (consulte-se Rusin-

kiewicz and Levoy [2001], Chen and Medioni [1992]), dos sub-conjuntos de superfícies ou perfis para recolher os pontos de controlo é necessária previamente a encontrar qual destes é “o mais próximo”; isto porque não existe nenhuma maneira de se conhecer correspondências de antemão. Isto leva geralmente a um dos maiores desafios do registo 3D: evitar um grande número de *outliers* resultando de “más” escolhas locais para a sub-amostragem.

Propõe-se neste trabalho que cada recta de intersecção referida ao ponto de vista i denotada como l_{in} pode ser usada como um percurso de procura para determinar que pontos pertencendo ao perfil p_i correspondem a um outro ponto pertencendo ao perfil p_n num ponto de vista de referência n .

A formalização matemática do registo tridimensional de um perfil realizado desta maneira seria então formulada como

$$e = \sum_{i=1}^N \left\| {}^W\mathbf{T}_{C_n} \mathbf{P}_{ni} - {}^W\mathbf{T}_{C_i} \mathbf{P}'_{in} \right\|^2, \quad \begin{cases} \mathbf{P}'_{in} = l_{in} \cap p_i \\ l_{in} = {}^{C_i}\Pi_3^n \cap {}^{C_i}\Pi_3^i \end{cases} \quad (3.2)$$

onde todas as entidades têm o mesmo significado que na equação (3.1), com a excepção de $\mathbf{P}'_{in}$, que representa o ponto correspondente no perfil p_i tomado usando a recta de procura l_{in} , que é o resultado da intersecção entre ${}^{C_i}\Pi_3^n$, o plano de luz tomado do ponto de vista n , e ${}^{C_i}\Pi_3^i$, o plano de luz tomado do ponto de vista i , ambos referidos a $\{C_i\}$.

Metodologia para a optimização local

O processo de minimização da função objectivo dada pela equação (3.2) é denominada de *optimização local*. Esta parte particular do registo tridimensional foi estudada em profundidade no passado — encontra-se intimamente ligada ao popular *problema de estimação da orientação absoluta*, que por sua vez tem também aplicações em fotogrametria, na análise de movimento de objectos, na determinação da transformação sensor-actuator e em estimação de “*pose*” [Eggert et al. 1997].

Um estudo generalizado dos desempenhos dos principais métodos não-iterativos foi feito em Eggert et al. [1997]. Quatro algoritmos mais importantes foram comparados neste trabalho:

1. Minimização linear directa de (3.2) através da estimação prévia da matriz orto-

normal de rotação usando SVD, usando de seguida esta estimativa para determinar o vector de translação — apresentado em Arun et al. [1987].

2. Um método similar recorrendo ao cálculo de valores/vectores próprios — apresentado em Horn et al. [1988].
3. Um método que, como os dois algoritmos anteriormente descritos, estima primeiro a rotação e depois a translação, como no algoritmo anterior recorre também ao cálculo de valores/vectores próprios, mas desta vez usando a notação de quaterniões mais compacta para representar a rotação — também concebido por Horn e apresentado em Horn [1987].
4. Um algoritmo que lida com a rotação e a translação em simultâneo de forma a providenciar uma abordagem linear de mínimos quadrados, usando a notação de quaterniões duais e recorrendo ao cálculo de valores/vectores próprios — apresentado em Walker et al. [1991].

Neste estudo foi tornado claro que, em geral, estes algoritmos obtêm resultados similares em condições realistas. Com isto em mente, foi decidido usar-se neste trabalho a solução apresentada por Michael Walker, que usa uma formulação elegante e evita também os efeitos da propagação de erro da estimação da rotação para a estimação da translação.

De acordo com este método, a equação que representa a transformação $\mathbf{T}$ que leva o ponto $\mathbf{P}_{ni}$ a coincidir com o seu correspondente $\mathbf{P}_{in}$ é dado pela equação (C.31) (na página 160), onde $\mathbf{\hat{q}}$ (representando a rotação) e $\mathbf{\hat{q}}'$ (representando a translação) são as partes real e dual, respectivamente, do quaternião dual unitário $\mathbf{\check{q}}$ correspondendo ao movimento helicoidal relacionado com $\mathbf{T}$, e $\mathbf{\hat{p}}_{ni}$ e $\mathbf{\hat{p}}_{in}$ são quaterniões puramente imaginários correspondendo a $\mathbf{P}_{ni}$ e $\mathbf{P}_{in}$, respectivamente — propõe-se nova consulta ao Apêndice C para mais informação sobre esta equação e os princípios geométricos e matemáticos subjacentes.

Algumas diferenças serão visíveis entre a equação (C.31) e a expressão apresentada em Walker et al. [1991] — estas são devidas à notação vectorial diferente usada por Walker et al., que trocam as posições relativas da partes real e imaginária de forma a que $\mathbf{\hat{q}}_{walker} = [q_x \ q_y \ q_z \ q_0]^T$. Isto resulta na troca natural de células das matrizes $\mathbf{Q}(\mathbf{\hat{q}})$ e $\mathbf{W}(\mathbf{\hat{q}})$ e numa expressão ligeiramente diferente para (C.31); para saber as expressões exactas para estas entidades, recomenda-se a consulta do texto original de Walker et al. [1991].

Nas linhas que se seguem, optar-se-á por um abuso de notação e denotar-se-ão as entidades segundo estes autores da mesma forma que os seus correspondentes apre-

sentados no Apêndice C — pede-se que o leitor se recorde, porém, que o formato dos resultados que se seguem não seria o mesmo usando a notação apresentada no Apêndice referido; apesar disto, note-se que a solução para o problema difere apenas algebricamente e não em termos dos princípios subjacentes.

Dito isto, rescrevendo a equação (3.2) de acordo com Walker et al. [1991] e considerando $\tilde{\mathbf{p}}_{in}$, correspondendo a $\tilde{\mathbf{P}}_{in} = {}^W\mathbf{T}_{C_i} \mathbf{P}'_{in}$, and $\mathring{\mathbf{p}}_{in}$ resultante da equação (C.31), referidos respectivamente como ponto “medido” e “transformado”, respectivamente, obtém-se

$$e = \sum_{i=1}^N \beta_i \left\| \mathring{\mathbf{p}}_{in} - \tilde{\mathbf{p}}_{in} \right\|^2, \quad \begin{cases} \tilde{\mathbf{P}}_{in} = {}^W\mathbf{T}_{C_i} \mathbf{P}'_{in} \mapsto \tilde{\mathbf{p}}_{in} \\ \mathbf{P}'_{in} = l_{in} \cap p_i \\ l_{in} = {}^{C_i}\Pi_3^n \cap {}^{C_i}\Pi_3^i \end{cases} \quad (3.3)$$

onde β_i são coeficientes de peso constantes e positivos que podem ser utilizados para reflectir a fiabilidade dos dados fornecidos [Walker et al. 1991, Eggert et al. 1997].

Usando as propriedades dos quaterniões duais e a equação (C.31), poder-se-á expandir o quadrado da norma de forma a obter [Walker et al. 1991]

$$\begin{aligned} \left\| \mathring{\mathbf{p}}_{in} - \tilde{\mathbf{p}}_{in} \right\|^2 &= \mathring{\mathbf{q}}^T \mathring{\mathbf{q}}' + 2\mathring{\mathbf{q}}^T \left(\mathbf{W}(\mathring{\mathbf{p}}_{ni}) - \mathbf{Q}(\tilde{\mathbf{p}}_{in}) \right) \mathring{\mathbf{q}} \\ &\quad - 2\mathring{\mathbf{q}}^T \mathbf{Q}(\tilde{\mathbf{p}}_{in})^T \mathbf{W}(\mathring{\mathbf{p}}_{ni}) \mathring{\mathbf{q}} \\ &\quad + \left((\mathring{\mathbf{p}}_{ni})^T \mathring{\mathbf{p}}_{ni} + (\tilde{\mathbf{p}}_{in})^T \tilde{\mathbf{p}}_{in} \right) \end{aligned} \quad (3.4)$$

Logo, a função objectivo pode ser escrita como uma função quadrática de $\mathring{\mathbf{q}}$ e $\mathring{\mathbf{q}}'$ [Walker et al. 1991]

$$e = \mathring{\mathbf{q}}^T \mathbf{C}_1 \mathring{\mathbf{q}} + \mathring{\mathbf{q}}'^T \mathbf{C}_2 \mathring{\mathbf{q}}' + \mathring{\mathbf{q}}'^T \mathbf{C}_2 \mathring{\mathbf{q}} + K \quad (3.5)$$

onde

$$\begin{aligned}
\mathbf{C}_1 &= -2 \sum_{i=1}^N \beta_i \mathbf{Q}(\tilde{\mathbf{p}}_{in})^T \mathbf{W}(\mathring{\mathbf{p}}_{ni}) \\
\mathbf{C}_2 &= \left(\sum_{i=1}^N \beta_i \right) \mathbf{I} \\
\mathbf{C}_3 &= 2 \sum_{i=1}^N \beta_i \left(\mathbf{W}(\mathring{\mathbf{p}}_{ni}) - \mathbf{Q}(\tilde{\mathbf{p}}_{in}) \right) \\
K &= \sum_{i=1}^N \beta_i \left((\mathring{\mathbf{p}}_{ni})^T \mathring{\mathbf{p}}_{ni} + (\tilde{\mathbf{p}}_{in})^T \tilde{\mathbf{p}}_{in} \right)
\end{aligned} \tag{3.6}$$

Usando a condição de norma unitária para o quaternião dual $\mathring{\mathbf{q}}$, que em termos das suas partes real e dual implica $\mathring{\mathbf{q}}^T \mathring{\mathbf{q}} = 1$ e $\mathring{\mathbf{q}}'^T \mathring{\mathbf{q}}' = 0$, Walker et al. recorrem aos multiplicadores de Lagrange (consulte-se Rao [2002a]) para provar que a solução para $\mathring{\mathbf{q}}$ é dada pelo vector próprio de

$$\mathbf{A} = \frac{1}{2} (\mathbf{C}_3^T (\mathbf{C}_2 + \mathbf{C}_2^T)^{-1} \mathbf{C}_3 - \mathbf{C}_1 - \mathbf{C}_1^T) \tag{3.7}$$

correspondente ao seu valor próprio positivo de maior valor absoluto. A solução para a parte dual é facilmente demonstrada como sendo dada por $\mathring{\mathbf{q}}' = -(\mathbf{C}_2 + \mathbf{C}_2^T)^{-1} \mathbf{C}_3 \mathring{\mathbf{q}}$ [Walker et al. 1991].

3.2.3 Algoritmo de registo 3D de perfis

O algoritmo global de registo tridimensional de perfis pode ser descrito como a iteração consecutiva dos passos que se seguem, usando as leituras do sensor de “*pose*” para determinar os valores iniciais para as estimativas das transformações.

Na iteração t estimam-se todos os $\mathbf{T}_n(t)$, $n = 1..m$ para um total de m perfis:

1. Primeiramente, criam-se listas de todos os planos de luz que intersectam cada plano de luz que tenha sido gerado. Para que uma intersecção seja válida é necessário, como foi dito anteriormente, que esta aconteça na superfície do objecto a ser reconstruído. Para proceder a esta validação de intersecções, determina-se o paralelepípedo de menor volume (acima de determinado limiar) que englobe completamente o perfil de referência: consideram-se planos intersectantes válidos aqueles cujas rectas de intersecção correspondentes intersectem também este paralelepípedo. Planos sem intersectantes são classificados como sendo deficientes de informação, sendo consequentemente eliminados.

2. Em seguida, percorrem-se as listas de planos intersectantes por plano de referência, determinando-se os pontos a menor distância das rectas de intersecção, tanto do plano de referência, como de cada plano da lista. Os pares ordenados de pontos resultantes são consideradas candidatos a correspondências. A validade das candidaturas verifica-se usando um limiar de distância escolhido tomando em conta uma avaliação prévia da precisão máxima da estimativa inicial. Se não existirem candidaturas válidas, considera-se o perfil como registado, passando-se ao perfil seguinte.
3. Resolve-se a equação (3.3) usando o método descrito na secção antecedente para determinar $\mathbf{T}_n(t)$. Nesta equação são usadas transformações ${}^W\mathbf{T}_{C_i}$ correspondendo à iteração $t - 1$. Os pesos β_i são calculados levando em conta a ortogonalidade dos planos de luz considerados (quanto “mais ortogonais” os planos, mais incisivo será o resultado da sua intersecção) e a distância entre os pontos correspondentes (quanto “menos distantes” os pontos, mais fiáveis são as correspondências). Para se conseguir independência das unidades e escalas utilizadas, optou-se por se usar a chamada *medida ou distância de Tanimoto* — pertencente ao que se dá o nome de medidas de similaridade — dada entre dois pontos 3D $\mathbf{x}$ e $\mathbf{y}$ por [Tanimoto 1959, Scott and Cao 1999]

$$S_T(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \mathbf{x}^T \mathbf{y}} = \frac{1}{1 + \frac{(\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})}{\mathbf{x}^T \mathbf{y}}} \quad (3.8)$$

que, como se vê, é inversamente proporcional à distância euclidiana ao quadrado entre esses pontos dividida pela sua correlação — isto é, esta medida, que está normalizada de forma a variar entre 0 e 1, é tanto maior quanto maior for a correlação, e logo a proximidade/similaridade, entre os pontos $\mathbf{x}$ e $\mathbf{y}$; usou-se também

$$O(\hat{\mathbf{n}}, \hat{\mathbf{m}}) = 1 - \hat{\mathbf{n}} \cdot \hat{\mathbf{m}} \quad (3.9)$$

como medida de ortogonalidade (o produto interno das normais aos planos dará mais próximo de 0 quanto mais próximo se estiver da ortogonalidade).

Desta forma, considerou-se a formula geral $\beta_i = S_T^j \cdot O^k$, onde j e k podem ser escolhidos empiricamente, conforme as características e desempenhos médios do *scanner*.

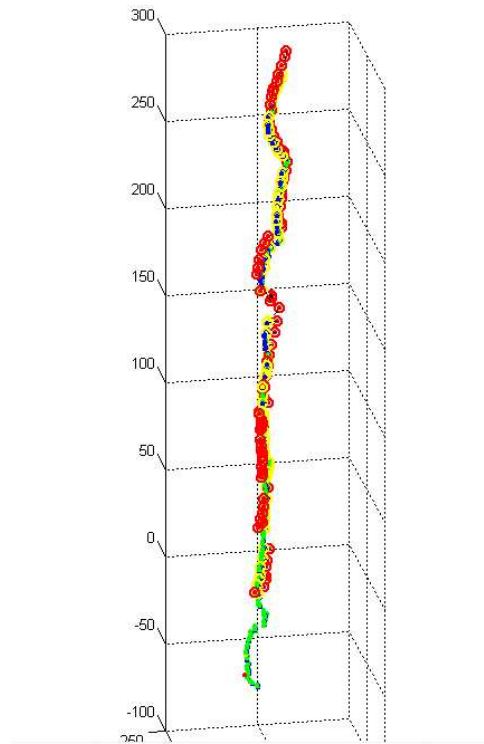


Figura 3.5: Exemplo prático do processo de registo de perfis 3D. Podem ver-se: a azul, os pontos do perfil de referência na iteração $t - 1$; a amarelo e vermelho os pontos candidatos a correspondência no perfil de referência e nos perfis intersectantes, respectivamente; os candidatos válidos envolvidos por uma circunferência; finalmente, a verde os pontos do perfil de referência depois da iteração t .

4. Finalmente, calcula-se o erro global de correspondência dado pela soma das distâncias entre pontos — se for menor que um limiar mínimo para o erro, pára-se o algoritmo; senão, passa-se à iteração seguinte. Também serve como condição de paragem o facto de todos os $\mathbf{T}_n(t)$ se aproximarem tanto da matriz identidade que se pode considerar que o algoritmo atingiu um mínimo, que se espera global.

Como exemplo prático de um registo de um perfil, veja-se a figura 3.5.

Foram feitas algumas experiências com este algoritmo a partir das integrações provisórias apresentadas na figura 3.3 na página 114; provou-se, porém, muito problemático obter registos convincentes a partir destes dados — os níveis de ruído, perfeitamente visíveis na figura, devidos à dificuldade em obter uma calibração que permita alta precisão, foram determinantes.

Uma leitura mais atenta de Hébert and Rioux [1998] revela um conceito bastante importante para a intuição da resolução deste problema: Hébert and Rioux defendem que, quando as condições iniciais não são suficientemente próximas da solução final,

uma boa política será agrupar os perfis em grupos.

Foi então decidido proceder-se a um pré-registo usando o mesmo algoritmo, mas desta vez aplicando-o não a perfis, e sim a grupos de perfis — a diferença no procedimento encontrar-se-ia no terceiro passo, onde a transformação rígida seria estimada para registar *cada grupo de perfis de uma só vez*. Os grupos, denominados de “*clusters*”, foram automaticamente agrupados de acordo com a orientação do seu plano-luz correspondente. No caso do manequim, foram identificados 5 “*clusters*”; no caso da estátua, foram identificados 3 “*clusters*”. Podem ver-se os resultados de pré-registo efectuado desta forma na figura 3.6 na próxima página.

Os resultados finais do registo efectivo de perfis foram usados para as reconstruções tridimensionais apresentadas no texto que se segue.

3.3 Reconstrução tridimensional — resultados

Para a reconstrução tridimensional final, a efectuar após a integração dos dados, foi concebido um pacote de software em Visual C++, como está apresentado no Apêndice B.2, com o nome de *Tele-3D Reconstructor*, utilizando a ferramenta “*open source*” Visualization Toolkit (VTK) [Schroeder et al. 2004]. O software desenvolvido utiliza o algoritmo *PowerCrust*, apresentado em Nina Amenta and Kolluri [2001], para efectuar a reconstrução 3D, gerando uma iso-superfície¹ a partir de pontos de controlo sem ordenação específica — exemplos de resultados do uso deste software podem ser vistos na figura 3.7.

Nas figuras 3.8, 3.9, 3.10 e 3.11 apresentam-se os resultados de reconstrução dos dois objectos digitalizados pelo sistema Tele-3D. É visível o efeito nefasto do erro (maioritariamente devido à calibração menos conseguida do projector laser, cujo erro introduzido foi estimado como sendo de à volta de 1,2 cm — ver secção 2.5) nalgumas partes dessas reconstruções, particularmente na da estátua, visto ter ao mesmo tempo uma geometria mais redundante e também uma superfície com detalhes mais finos; porém, também se torna muito claro os bons resultados conseguidos pelo processo de registo, mesmo em condições tão adversas.

¹Envolvendo ou não um volume — o algoritmo tenta resolver esta questão inferindo a partir dos dados fornecidos.

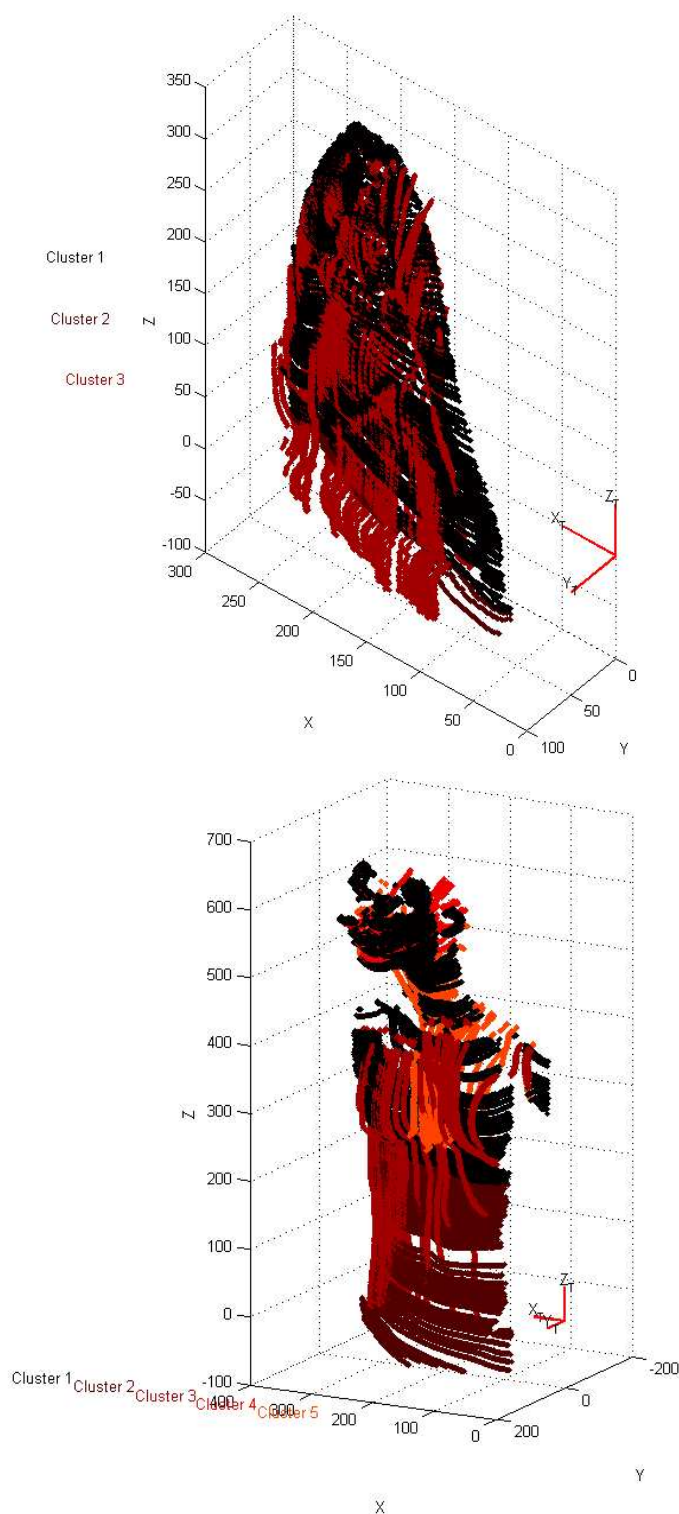


Figura 3.6: Objectos com “clusters” registados — pode ver-se a melhoria substancial obtida na integração dos dados depois do registo comparando esta figura com a figura 3.3 na página 114.

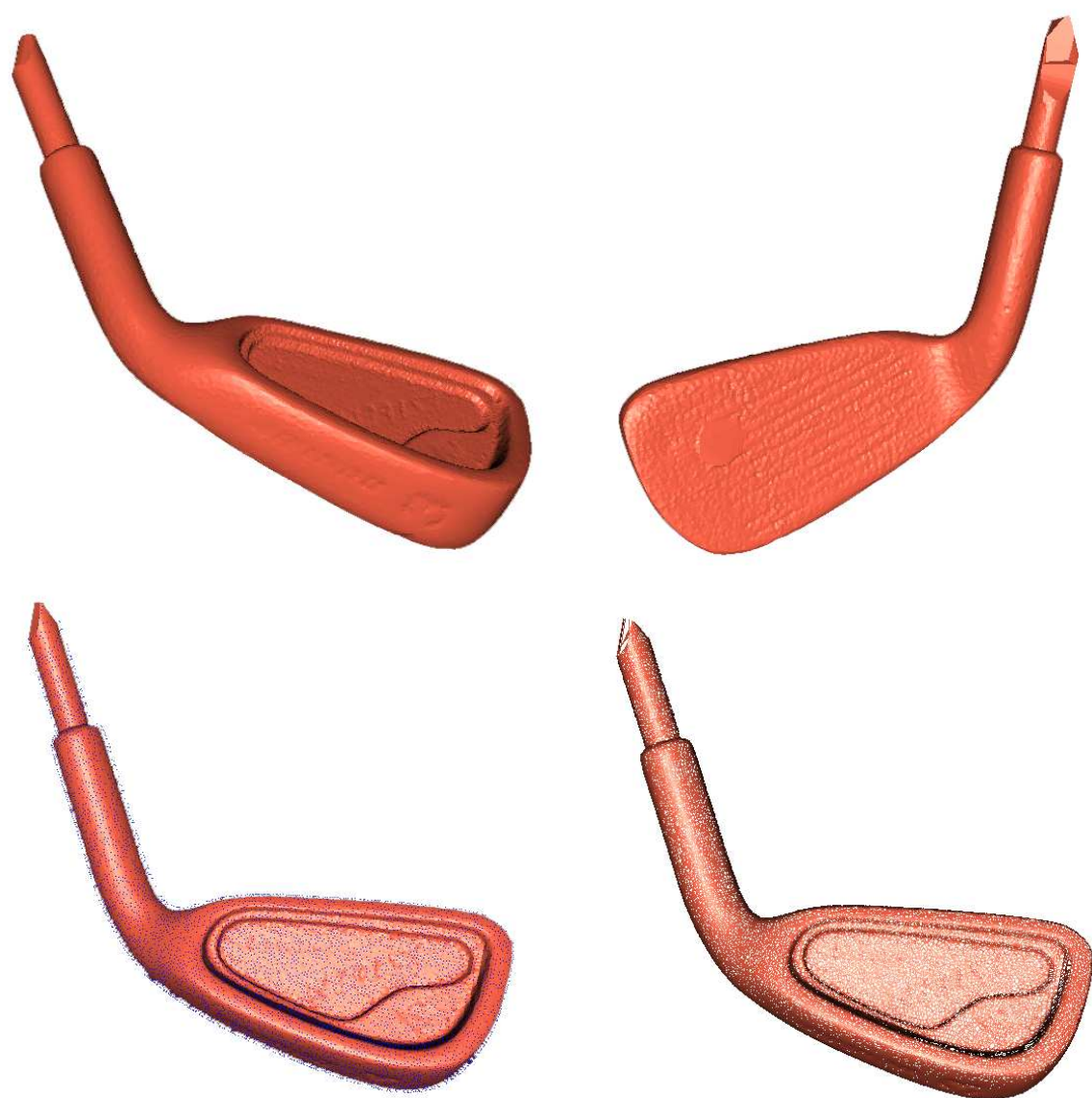


Figura 3.7: Exemplo de reconstrução tridimensional usando o software *Tele-3D Reconstructor*. Na figura podem ver-se duas perspectivas diferentes da reconstrução 3D parcial de um taco de golfe, os pontos de controlo a partir dos quais se fez a reconstrução (retirados, com a gentil permissão do Doutor Hugues Hoppe, do site relativo a Hoppe et al. [1994], e representados a azul na imagem em baixo, à esquerda) e a sua representação em “mesh” tridimensional (em baixo, à direita).

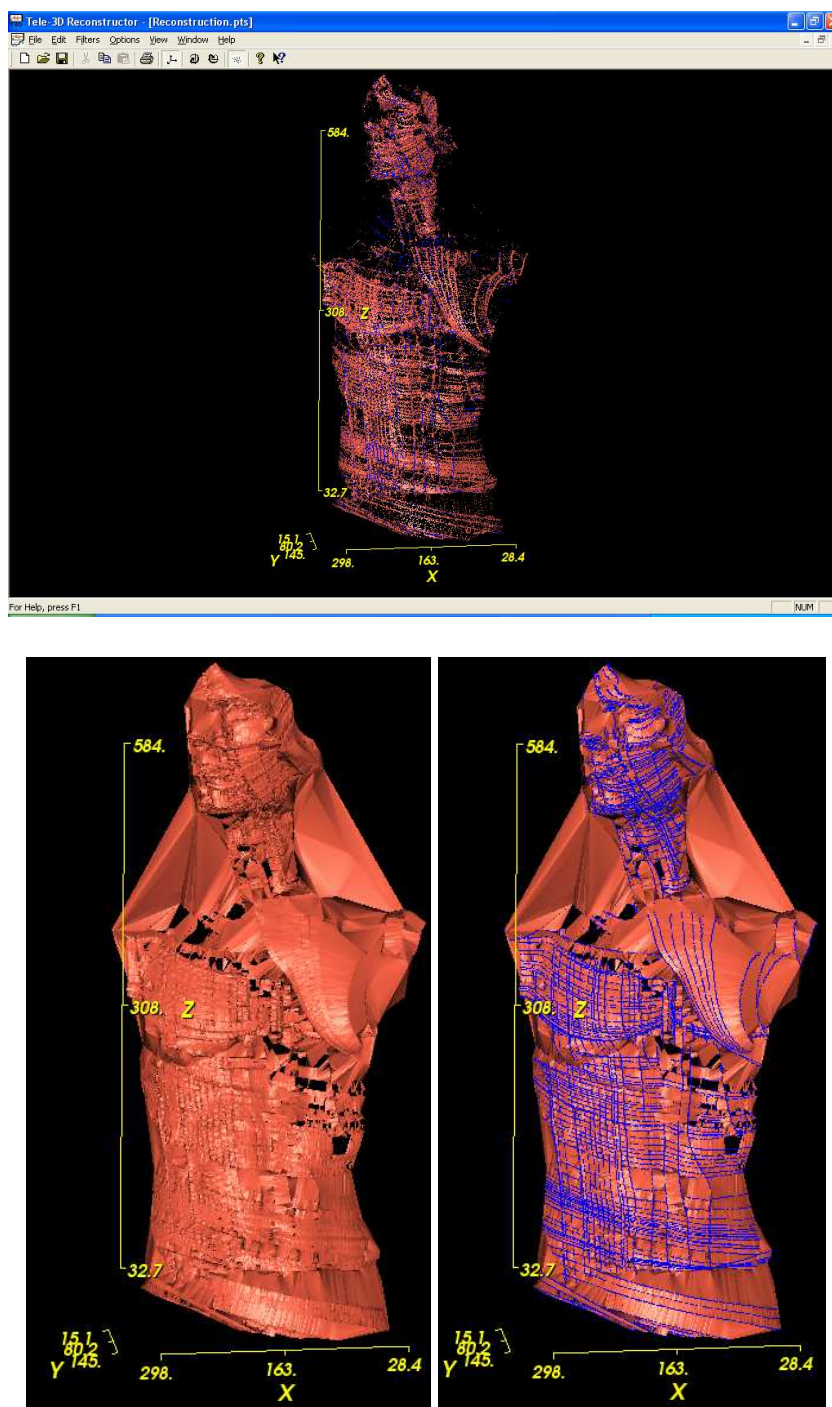


Figura 3.8: Resultados da reconstrução tridimensional do manequim, usando o software *Tele-3D Reconstructor*. É possível verificar-se que o registo para o lado direito do torso (lado esquerdo na imagem) e para a cara funcionou bastante bem. Em cima podem ver-se os pontos de controlo (digitalizados pelo *scanner*) a azul e os pontos adicionados pelo *PowerCrust* a vermelho — como só foi digitalizada a superfície frontal do manequim, o algoritmo de reconstrução ligou os ombros do manequim à cara, o que não aconteceria se a digitalização tivesse sido mais extensiva.

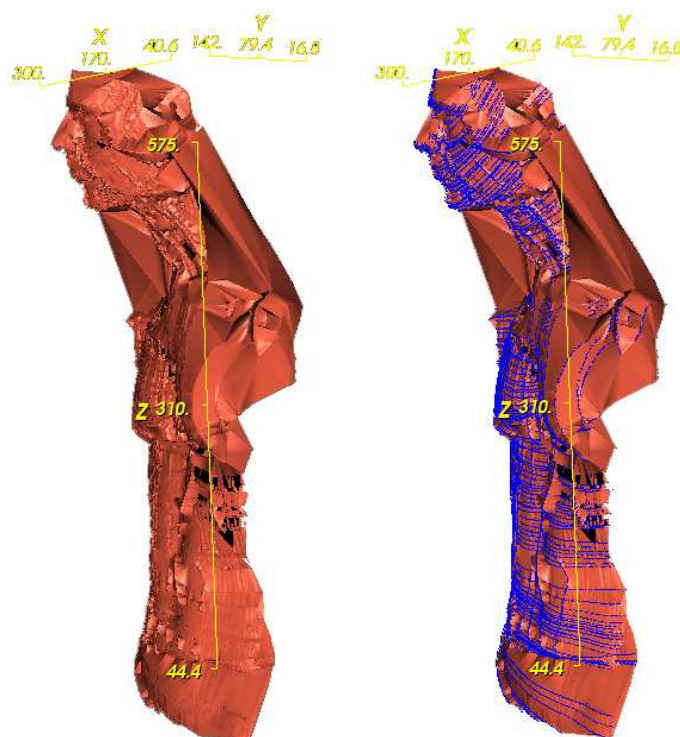
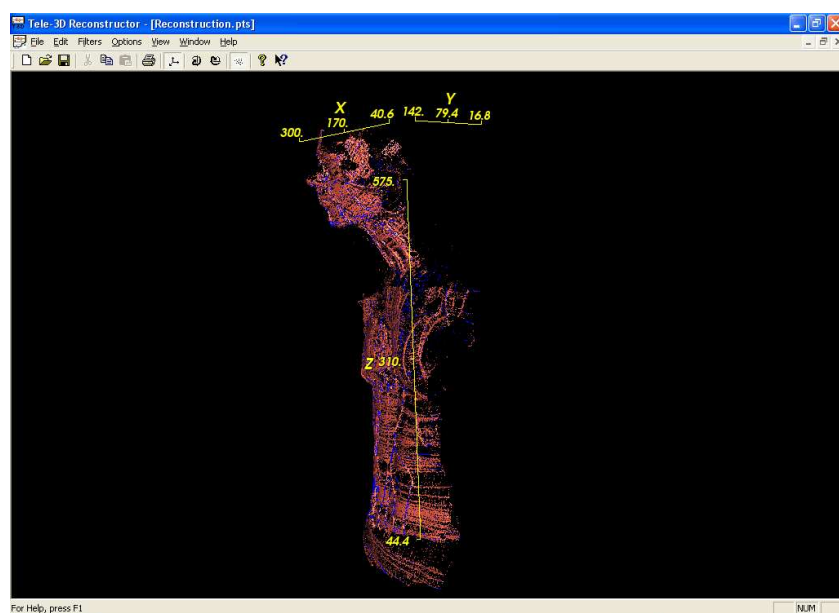


Figura 3.9: Resultados da reconstrução tridimensional do manequim (vista de perfil), usando o software *Tele-3D Reconstructor*. Mais uma vez, em cima podem ver-se os pontos de controlo (digitalizados pelo *scanner*) a azul e os pontos adicionados pelo *PowerCrust* a vermelho.

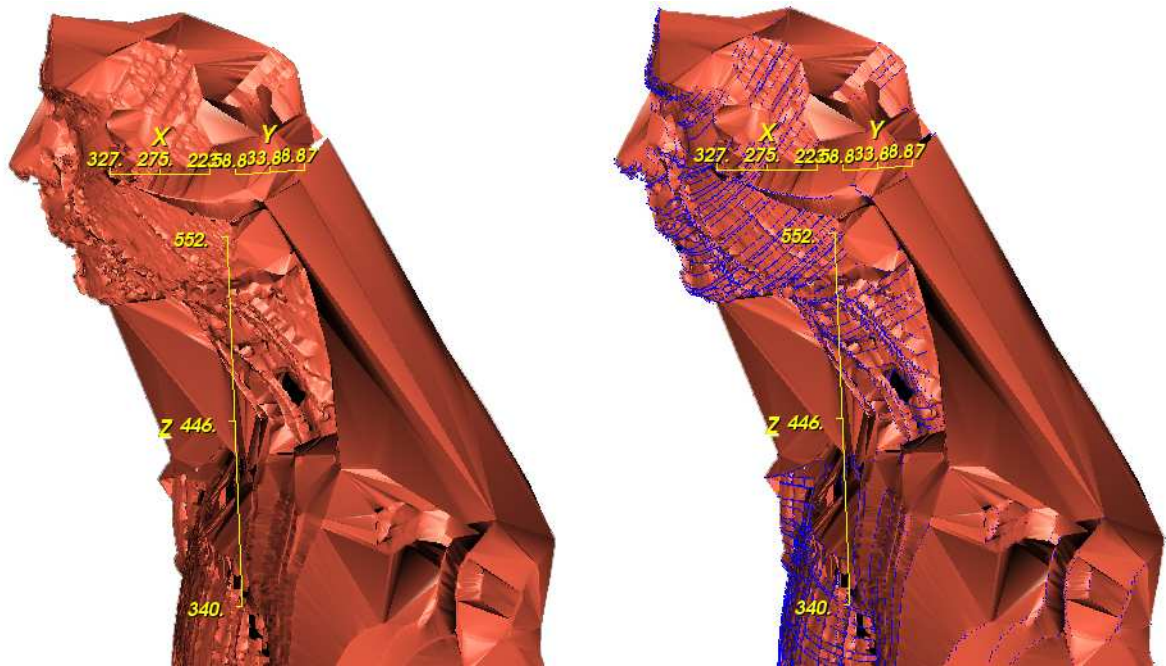


Figura 3.10: Detalhe dos resultados usando o software *Tele-3D Reconstructor* da reconstrução tridimensional do manequim (face). O único problema surgido nesta reconstrução pode ver-se na orelha, local de detalhe mais fino e, portanto, mais sensível a ruído.

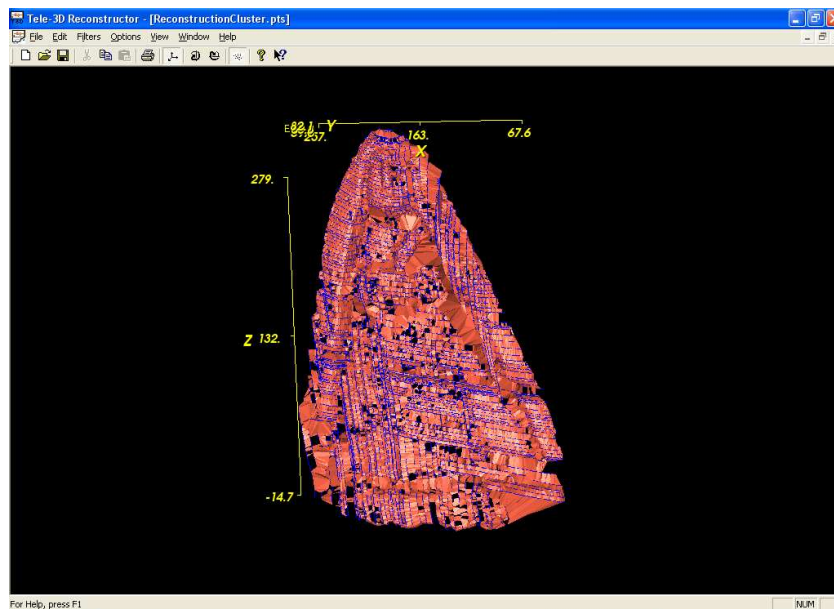


Figura 3.11: Resultados usando o software *Tele-3D Reconstructor* da reconstrução tridimensional da estátua de Nossa Senhora. Repare-se que muito do detalhe mais fino perdeu-se no ruído que permeia os perfis, em média maior que o tamanho desse detalhe, e nos erros de registo.

3.4 Conclusões

Tentou-se neste capítulo conceber um bom método automático de processamento para o varrimento das cenas tridimensionais a recuperar, formular uma solução matemática e algorítmica coerente e interessante para o registo tridimensional subsequente e um pacote de software de reconstrução tridimensional a partir de pontos desordenados.

Apesar dos problemas inerentes à natureza do scanner *handheld* Tele-3D e também relacionados com algumas questões que se prendem com a engenharia do sistema (sensibilidade dos componentes, condições de trabalho, etc.), os resultados conseguidos após o processamento dos dados reais obtidos mostram ter-se atingido o objectivo de apresentar uma abordagem inovadora e consequente às questões apresentadas neste capítulo.

Capítulo 4

Conclusões e Trabalho Futuro

Ao longo deste texto foi efectuado um estudo exaustivo e apresentada uma solução teórica e prática para um sistema de digitalização tridimensional amovível, que culminou na construção do protótipo Tele-3D.

Várias conclusões em diversas áreas foram atingidas e foram-se também delineando várias ideias para trabalho futuro, concretamente neste protótipo e mais genericamente nesta área de investigação. Estas conclusões e ideias vão ser desenvolvidas seguidamente neste capítulo, com que se encerra esta dissertação.

4.1 Apreciação dos resultados

Como foi sendo dito ao longo deste texto, na sua maioria, as questões apresentadas (essencialmente os modelos, calibração e registo tridimensional) foram sendo resolvidas com uma razoável taxa de sucesso.

O maior problema foi, sem dúvida, conseguir-se um sistema que tivesse um nível de precisão e robustez satisfatório, o que se acredita ser possível facilmente com um pouco mais de trabalho de refinação da engenharia e das condições de funcionamento do *scanner*.

4.2 Trabalho futuro

Como se disse acima, algum trabalho futuro no sentido de melhorar o desempenho do sistema Tele-3D torna-se necessário e está apresentado nas linhas que se seguem.

4.2.1 Melhoria das condições de funcionamento do sistema

Ao longo dos testes laboratoriais foi-se tornando claro que algumas questões que muitas vezes não se consideram com tanto cuidado em teoria se tornam críticas na prática durante um funcionamento de sistemas deste género.

Nomeadamente no caso do sistema Tele-3D, isto notou-se particularmente no desempenho do sensor de “*pose*” utilizado, que sofria bastante na presença de campos electromagnéticos externos, na câmara, que devido às suas características de focagem próxima apresentava uma distorção radial severa, e no projector laser, que não estaria muito provavelmente a projectar um plano-luz equilibrado (foi notória muitas vezes a diferença de brilho e grossura nos perfis de intersecção).

Estas difíceis condições de funcionamento apelam para uma melhor afinação futura do sistema, que, como já foi dito, foi construído a partir de componentes “*off-the-shelf*”, o que obriga a maiores cuidados na sua utilização, nomeadamente no processo de calibração.

4.2.2 Aposta em sensores inerciais

Depois de efectuadas as afinações acima referidas, de forma a tornar o sistema completamente robusto e fiável, torna-se possível dar o passo seguinte em direcção a uma significativa melhoria na sua autonomia — a aposta eventualmente mais arriscada em sensores de “*pose*” auto-contidos, como os sensores inerciais, para conseguir aumentar ainda mais o grau de portabilidade do sistema.

4.2.3 Utilização de visão estereoscópica activa

Outra aposta de futuro seria o recurso à visão estereoscópica activa, como já é feito em *scanners* similares ao Tele-3D (vide Polhemus/ARANZ [2000]), de forma a evitar oclusões e providenciar imagens de textura das cenas tridimensionais digitalizadas.

4.2.4 Refinação do processo de reconstrução tridimensional

Outra área em que seria importante fazer-se uma melhoria substancial é na parte final do processo de digitalização 3D — a reconstrução. Seria importante tratar de questões como:

- a sub-amostragem dos pontos que constituem a superfície (neste trabalho efectuada através de uma simples decimação) de forma a atenuar o peso de cada reconstrução (que em formato VRML chegam a ocupar 50 MBytes!);

- uma passagem prévia de um filtro de mediana pelos pontos de controlo usados para a reconstrução, de forma a controlar os efeitos do ruído e suavizar a superfície (este efeito foi simulado ao longo deste trabalho em pós-processamento através de um filtro passa-alto aplicado à iso-superfície após esta ser gerada);
- a aplicação de textura.

APÊNDICES

Apêndice A

Notação

Na tabela A.1 apresenta-se a notação usada ao longo do texto para maior clareza de exposição.

Tabela A.1: Notação usada e descrições correspondentes.

Símbolo	Descrição
a	escalares
a^{-1}	inversos (o mesmo para todos os seguintes)
$\hat{\mathbf{n}}$	vectores unitários (ex.: vector normal)
$\vec{\mathbf{t}}$	vectores aplicados (ex.: vector de translação)
$\mathbf{p}, \mathbf{P}$	restantes vectores (ex.: vector de coordenadas)
$\mathbf{R}$	matrizes
$\mathbf{R}^T$	matrizes transpostas
$\mathbf{R}^{-T}$	inversos de matrizes transpostas
$\check{a}$	números duais genéricos
$\check{a}^*$	conjugados de números duais genéricos
$\check{\mathbf{a}}$	vectores duais
$\check{\mathbf{a}}^*$	conjugados de vectores duais
$\overset{\circ}{\mathbf{q}}$	quaterniões (notação vectorial)
$\overset{\circ}{\mathbf{q}}^*$	conjugados de quaterniões (notação vectorial)
$\overset{\circ}{\mathbf{q}}_{\circ}$	quaterniões duais
$\overset{\circ}{\mathbf{q}}_{\circ}^*$	conjugados de quaterniões duais
$\{W\}$	sistemas referenciais de coordenadas
X_W	valores referidos a um referencial (ex.: coordenada)
${}^W\mathbf{T}_C$	transformações entre referenciais
$\vec{\mathbf{t}}, {}^W\tilde{\mathbf{T}}_C$	valores medidos
$\hat{\mathbf{t}}, {}^W\hat{\mathbf{T}}_C$	valores estimados (Diferem de valores unitários pelo contexto. Ex.: vector unitário estimado $\hat{\mathbf{n}}$.)

Apêndice B

Engenharia do Sistema Tele-3D

B.1 Descrição do hardware

O sistema Tele-3D (veja-se a figura B.1 na página ao lado) é constituído por:

- Um suporte obedecendo ao esboço apresentado na figura 2.1 na página 49, sustentando o *scanner* Tele-3D propriamente dito:
 - Uma câmara Pulnix (c-mount) equipada com uma lente de 3,7 mm (1:1,6) sobreposta por um filtro de interferência quando necessário;
 - Um projector de laser (que projecta um perfil linear, ou seja, um plano);
 - O lado receptor de um sensor de “*pose*” miniBIRD 500 da Ascension (cujas características se podem encontrar na tabela 1.1 na página 26).
- O lado emissor do miniBIRD;
- Um PC de controlo, cujas características foram variando ao longo do tempo; os principais detalhes a reter são uma placa de captura gráfica Matrox Meteor e com porta série reservada para a ligação do sensor de “*pose*”.

O suporte do *scanner* Tele-3D é uma estrutura leve em acrílico com uma pega esponjosa ao centro, cuja geometria e dimensões precisas foram resultados de um projecto CAD reproduzido na figura B.2 na página 140.

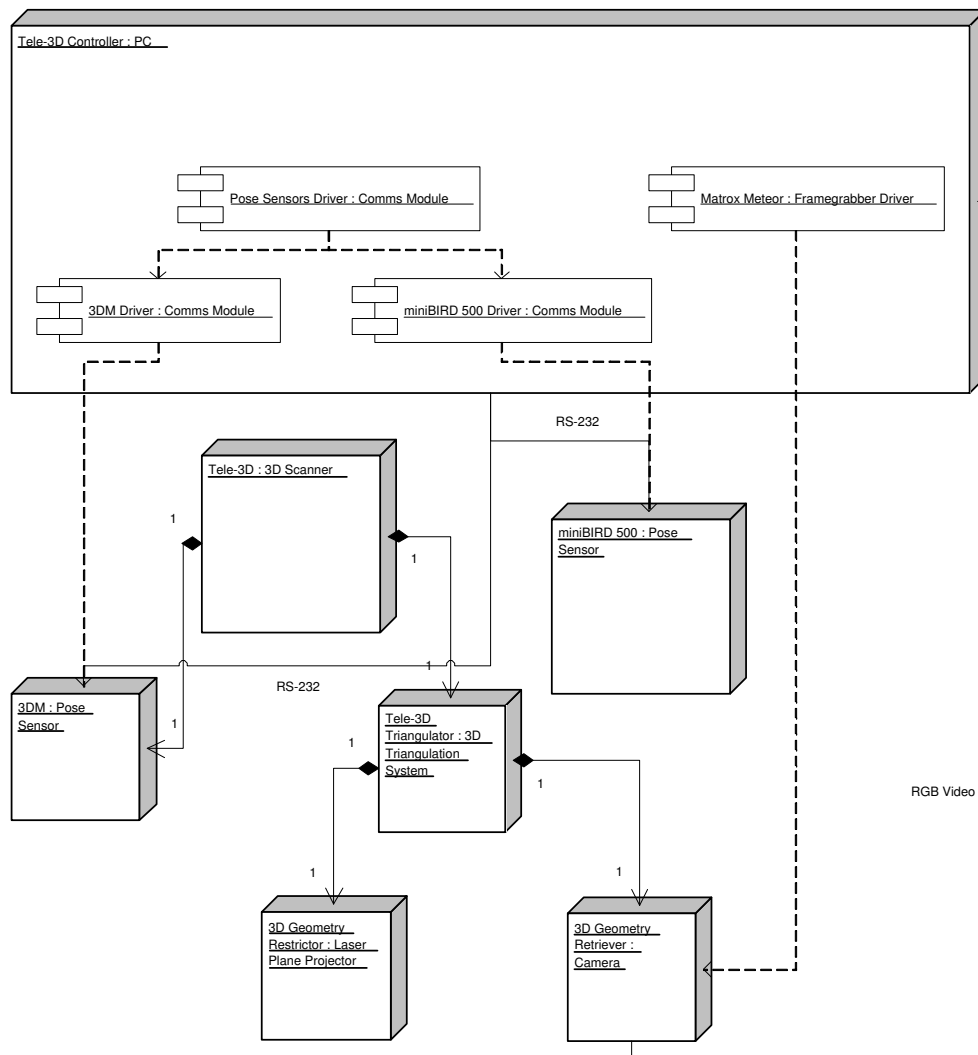
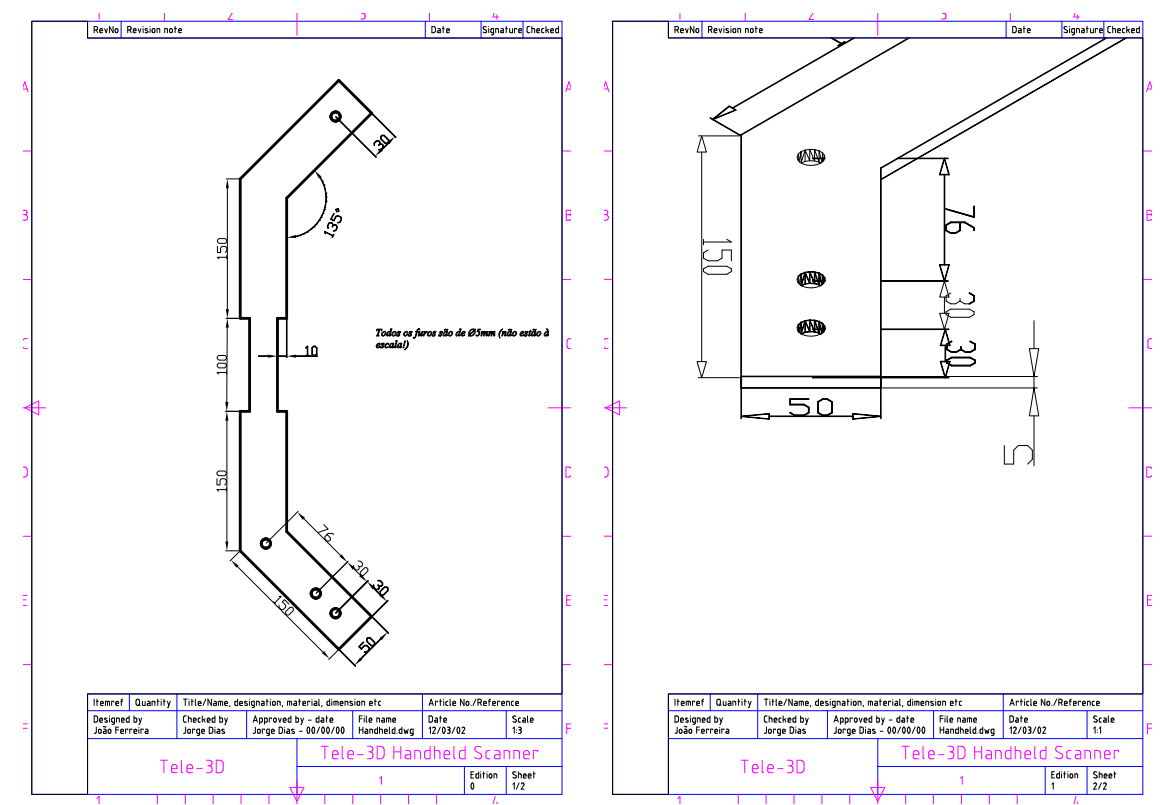


Figura B.1: “Deployment Diagram” (UML) para o sistema Tele-3D [Fowler and Kendall 1997, UML 1997]. O sensor de “pose” que acabou por ser usado exclusivamente foi o Ascension miniBIRD 500.



(a) Vista de topo

(b) Vista isométrica



(c) Fotografia em uso

Figura B.2: Projecto CAD do *scanner* Tele-3D em duas vistas e fotografia em uso. **Atenção:** (a) e (b) são reproduções do projecto original e portanto não estão coerentes com as escalas indicadas.

B.2 Descrição do software

Em seguida, apresentam-se algumas esquematizações que sumariam o sistema Tele-3D em termos de software. A figura B.3 mostra os casos de uso e a inter-relação entre os diversos módulos e a figura B.4 mostra alguns pormenores importantes de um desses módulos.

Os módulos *Tele-3D Calibrator* e *Tele-3D Scanner* foram programados em MATLAB, o primeiro recorrendo em parte ao calibrador de câmaras do OpenCV [OpenCV Manual]; os módulos *Tele-3D Grabber* e *Tele-3D Reconstructor* foram programados em Visual C++, tendo-se recorrido no primeiro aos drivers da Meteor e no segundo ao Visual Toolkit (VTK) da Kitware e ao algoritmo de reconstrução de superfícies a partir de pontos não ordenados concebido por Nina Amenta [Nina Amenta and Kolluri 2001] e adaptado para o VTK por Tim Hutton.

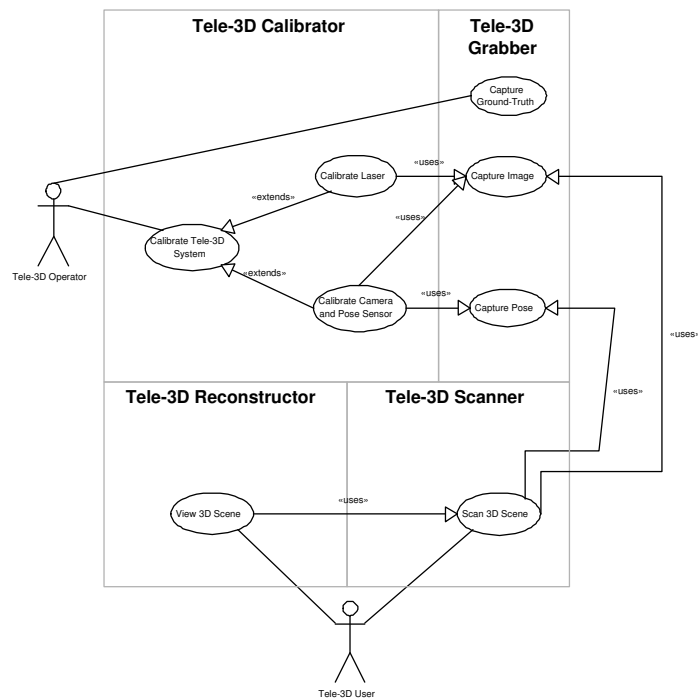


Figura B.3: “Use Case Diagram” (UML) para o sistema Tele-3D [Fowler and Kendall 1997, UML 1997].

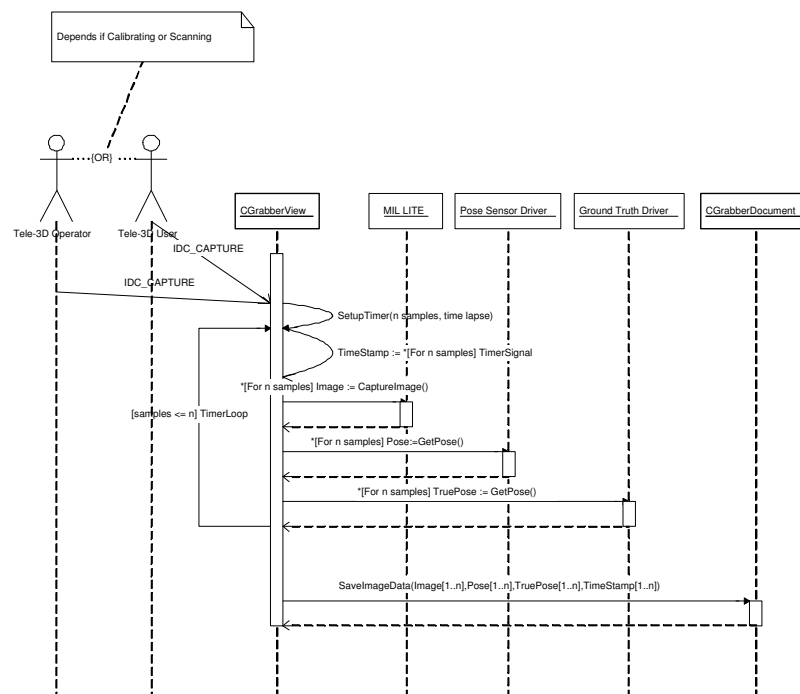


Figura B.4: “Sequence Diagram” (UML) para o software *Tele-3D Grabber* [Fowler and Kendall 1997, UML 1997].

Apêndice C

Quaterniões Duais e Transformações de Corpos Rígidos

O desenvolvimento da física e tecnologia no século XX gerou um interesse especial nalguns campos da matemática: teoria dos grupos, álgebra geral e álgebra de Clifford e, em particular, álgebra de quaterniões [Meister 1997].

A álgebra de quaterniões foi inventada no século XIX por W. R. Hamilton [Hamilton 1850]. Desde os primórdios, os quaterniões foram criados para desempenhar um papel especial no estudo de rotações em espaços vectoriais tridimensionais; a linguagem dos quaterniões era mais simples e descritiva do que outras, mas sendo uma linguagem nova não era fácil de usar. Logo, até meados do século XX, o uso prático de quaterniões tem sido mínima em comparação com outros métodos mais tradicionais, como por exemplo a amplamente divulgada representação de rotações através de matrizes ortogonais [Meister 1997].

Correntemente, a situação modificou-se essencialmente devido ao progresso nos campos da tecnologia espacial, de navegação terrestre e por satélite, etc. A estimação de transformações de corpos rígidos e estimação de movimento são precisamente o tipo de novos problemas onde o uso de quaterniões é de grande importância [Meister 1997].

Visto que a maior parte dos algoritmos usados ao longo deste texto são, de facto, tentativas para resolver problemas de optimização envolvendo a estimação de transformações de corpos rígidos, foi decidido que os quaterniões, devido às razões apresentadas acima, iriam desempenhar um papel essencial. Consequentemente, um estudo metuculoso de trabalhos prévios respeitando o uso de quaterniões na representação de tais transformações foi empreendido e irá ser sumariado no texto que se segue.

Primeiramente, a definição para os quaterniões, incluindo uma notação vectorial alternativa muito útil que vai ser usada ao longo do texto, será apresentada, seguida

da descrição de algumas das suas propriedades e um teorema que introduz o uso de quaterniões unitários na representação de rotações. Finalmente, uma introdução aos conceito de número dual será apresentado e a definição de quaterniões duais irá ser fornecida, incluindo um teorema que introduz o uso de quaterniões duais unitários para a representação completa de transformações de corpos rígidos através da noção cinemática de movimento helicoidal.

C.1 Rotações e Quaterniões Unitários

Apresentar-se-á agora a definição do quaternião, será discutida a relação entre quaterniões unitários e matrizes 3×3 ortogonais de rotação e referir algumas das suas propriedades matemáticas e geométricas mais úteis.

Um quaternião pode ser visto como um caso especial dos números complexos, neste caso com uma real e três partes imaginárias, e é denotado na sua forma polinomial expandida como

$$\mathring{q} = q_0 + iq_x + jq_y + kq_z \quad (\text{C.1})$$

com as propriedades que se seguem, respeitando os números imaginários envolvidos

$$i^2 = j^2 = k^2 = ijk = -1 \quad (\text{C.2a})$$

$$ij = k, \quad jk = i, \quad ki = j \quad (\text{C.2b})$$

$$ji = -k, \quad kj = -i, \quad ik = -j \quad (\text{C.2c})$$

Devido a estas propriedades, a multiplicação de quaterniões pode ser definida como

$$\begin{aligned} \mathring{r}\mathring{q} = & (r_0q_0 + r_xq_x + r_yq_y + r_zq_z) \\ & + i(r_0q_x + r_xq_0 + r_yq_z - r_zq_y) \\ & + j(r_0q_y - r_xq_z + r_yq_0 + r_zq_x) \\ & + k(r_0q_z + r_xq_y - r_yq_x + r_zq_0) \end{aligned} \quad (\text{C.3})$$

O produto $\mathring{q}\mathring{r}$ tem uma forma similar, mas seis dos sinais são invertidos, como pode ser prontamente verificado através da aplicação das propriedades apresentadas na equação (C.2) na equação (C.3). Logo, em geral $\mathring{r}\mathring{q} \neq \mathring{q}\mathring{r}$.

Outra entidade importante na álgebra de quaterniões é o quaternião conjugado. Considerando o quaternião dado pela equação (C.1), o seu conjugado é definido como

$$\hat{q}^* = q_0 - iq_x - jq_y - kq_z \quad (\text{C.4})$$

Não obstante, como será visto mais adiante, a maneira mais útil de denotar um quaternião é como um vector-coluna com quatro componentes, ou um vector coluna com um valor escalar e um componente vectorial tridimensional:

$$\hat{\mathbf{q}} = \begin{bmatrix} q_0 \\ q_x \\ q_y \\ q_z \end{bmatrix} = \begin{bmatrix} q_0 \\ \mathbf{q} \end{bmatrix} \quad (\text{C.5})$$

Usando esta notação, torna-se então possível definir o produto escalar de dois quaterniões como

$$\hat{\mathbf{p}} \cdot \hat{\mathbf{q}} = p_0q_0 + p_xq_x + p_yq_y + p_zq_z \quad (\text{C.6})$$

e, conseqüentemente, o quadrado da magnitude de um quaternião como o produto escalar desse quaternião consigo próprio:

$$\|\hat{\mathbf{q}}\|^2 = \hat{\mathbf{q}} \cdot \hat{\mathbf{q}} = q_0^2 + q_x^2 + q_y^2 + q_z^2 \quad (\text{C.7})$$

Considerem-se as matrizes auxiliares seguintes:

$$\mathbf{Q}(\hat{\mathbf{q}}) = \begin{bmatrix} q_0 & -q_x & -q_y & -q_z \\ q_x & q_0 & -q_z & q_y \\ q_y & q_z & q_0 & -q_x \\ q_z & -q_y & q_x & q_0 \end{bmatrix} \quad (\text{C.8a})$$

$$\mathbf{W}(\hat{\mathbf{q}}) = \begin{bmatrix} q_0 & -q_x & -q_y & -q_z \\ q_x & q_0 & q_z & -q_y \\ q_y & -q_z & q_0 & q_x \\ q_z & q_y & -q_x & q_0 \end{bmatrix} \quad (\text{C.8b})$$

A definição homóloga da operação de multiplicação de quaterniões dada pela equação (C.3) mas agora em notação vectorial, que se denotará de $*$ para distinguir da multiplicação vectorial normal, é dada por:

$$\begin{aligned}\mathring{\mathbf{r}} * \mathring{\mathbf{q}} &= \begin{bmatrix} r_0 q_0 - \mathbf{r}^T \mathbf{q} \\ r_0 \mathbf{q} + q_0 \mathbf{r} + \mathbf{r} \times \mathbf{q} \end{bmatrix} \\ &= \mathbf{Q}(\mathring{\mathbf{r}}) \mathring{\mathbf{q}} = \mathbf{W}(\mathring{\mathbf{q}}) \mathring{\mathbf{r}}\end{aligned}\quad (\text{C.9})$$

Qualquer vector tridimensional pode ser escrito alternativamente como um quaterniões puramente imaginário e a sua multiplicação possui uma propriedade muito útil, que pode ser derivada directamente da equação (C.9):

$$\begin{bmatrix} 0 \\ \mathbf{r} \end{bmatrix} * \begin{bmatrix} 0 \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} -\mathbf{r}^T \mathbf{q} \\ \mathbf{r} \times \mathbf{q} \end{bmatrix} \quad (\text{C.10})$$

Três propriedades interessantes, facilmente demonstráveis, que surgem da definição de multiplicação de quaterniões são

$$\|\mathring{\mathbf{q}}\|^2 = \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* = \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} \quad (\text{C.11})$$

$$\|\mathring{\mathbf{r}} * \mathring{\mathbf{q}}\|^2 = \|\mathring{\mathbf{r}}\|^2 \|\mathring{\mathbf{q}}\|^2 \quad (\text{C.12})$$

e

$$(\mathring{\mathbf{q}}^* * \mathring{\mathbf{p}}^*) = (\mathring{\mathbf{p}} * \mathring{\mathbf{q}})^* \quad (\text{C.13})$$

Como consequência da primeira destas três propriedades, pode-se concluir que o inverso de um quaterniões não nulo é dado por

$$\mathring{\mathbf{q}}^{-1} = \frac{\mathring{\mathbf{q}}^*}{\mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}}} \quad (\text{C.14})$$

Outras importantes, e óbvias, propriedades respeitando o uso das matrizes auxili-

ares dadas nas equações (C.8) são

$$\mathbf{Q}(\mathring{\mathbf{q}})^T \mathbf{Q}(\mathring{\mathbf{q}}) = \mathbf{Q}(\mathring{\mathbf{q}}) \mathbf{Q}(\mathring{\mathbf{q}})^T = \mathring{\mathbf{q}}^T \mathring{\mathbf{q}} I_{4 \times 4} \quad (\text{C.15a})$$

$$\mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{W}(\mathring{\mathbf{q}}) = \mathbf{W}(\mathring{\mathbf{q}}) \mathbf{W}(\mathring{\mathbf{q}})^T = \mathring{\mathbf{q}}^T \mathring{\mathbf{q}} I_{4 \times 4} \quad (\text{C.15b})$$

$$\mathbf{Q}(\mathring{\mathbf{q}})^T \mathring{\mathbf{q}} = \mathbf{W}(\mathring{\mathbf{q}})^T \mathring{\mathbf{q}} = \mathring{\mathbf{q}}^T \mathring{\mathbf{q}} \mathring{\mathbf{e}} \quad (\text{C.15c})$$

$$\mathbf{Q}(\mathring{\mathbf{r}}) \mathbf{Q}(\mathring{\mathbf{q}}) = \mathbf{Q}(\mathbf{Q}(\mathring{\mathbf{r}}) \mathring{\mathbf{q}}) \quad (\text{C.15d})$$

$$\mathbf{W}(\mathring{\mathbf{r}}) \mathbf{W}(\mathring{\mathbf{q}}) = \mathbf{W}(\mathbf{W}(\mathring{\mathbf{r}}) \mathring{\mathbf{q}}) \quad (\text{C.15e})$$

$$\mathbf{Q}(\mathring{\mathbf{r}}) \mathbf{W}(\mathring{\mathbf{q}})^T = \mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{W}(\mathring{\mathbf{r}}) \quad (\text{C.15f})$$

onde $\mathring{\mathbf{e}} = [1 \ 0 \ 0 \ 0]^T$ é o quaternião unidade (a parte real é 1 e as partes imaginárias são 0, o que implica $\mathring{\mathbf{e}} * \mathring{\mathbf{q}} = \mathring{\mathbf{q}} * \mathring{\mathbf{e}} = \mathring{\mathbf{q}}$).

Seguidamente, o teorema que ilustra a utilidade geométrica do quaternião como forma alternativa de representação de rotações face a outras representações puramente analíticas será apresentado. Este teorema e os seus corolários são um sumário, usando notação própria (consulte-se o Apêndice A), do trabalho apresentado em diversos artigos sobre o tema dos quaterniões e o seu uso na estimação de transformações de corpos rígidos, nomeadamente [Horaud and Dornaika 1995] e, mais com mais importância ainda, [Horn 1987]. A notação concebida para acomodar a teoria apresentada neste Apêndice enquadra-se no esforço empreendido na apresentação de uma descrição matemática uniforme e intuitiva dos conceitos rodeando os quaterniões, visto que cada autor nesta área usa a sua própria notação, tornando ainda mais difícil a assimilação desta linguagem matemática.

Teorema C.1.1. *Seja $\vec{\mathbf{r}}_i = [x \ y \ z]^T$ um vector tridimensional arbitrário que é transformado em $\vec{\mathbf{r}}_{i+1} = [x' \ y' \ z']^T$ através de uma rotação de ângulo θ sobre um eixo (passando através da origem) com direcção definida pelo versor $\hat{\mathbf{w}} = [w_x \ w_y \ w_z]^T$. Se*

$$\left\{ \begin{array}{l} \mathring{\mathbf{r}}_i = \begin{bmatrix} 0 \\ \vec{\mathbf{r}}_i \end{bmatrix} \\ \mathring{\mathbf{r}}_{i+1} = \begin{bmatrix} 0 \\ \vec{\mathbf{r}}_{i+1} \end{bmatrix} \\ \mathring{\mathbf{q}} = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \hat{\mathbf{w}} \end{bmatrix} = \begin{bmatrix} q_0 \\ \mathbf{q} \end{bmatrix} \end{array} \right.$$

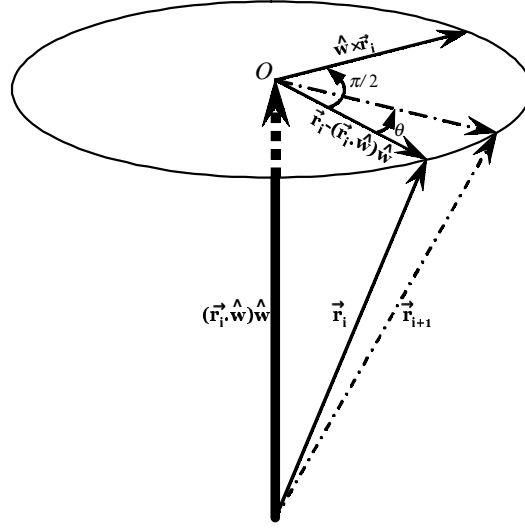


Figura C.1: A rotação de $\vec{r}_i$ para $\vec{r}_{i+1}$ (com o eixo de rotação representado pelo segmento de recta grosso apontando para o centro de rotação) pode ser entendida, como afirmado na fórmula de Rodrigues, em termos das entidades apresentadas neste diagrama [Horn 1987].

então a fórmula de Rodrigues (veja-se a figura C.1) tem uma representação equivalente e alternativa na forma de multiplicação de quaterniões, ou seja,

$$\begin{aligned}
 \vec{r}_{i+1} &= \mathbf{R} \vec{r}_i \\
 &= \cos \theta \vec{r}_i + \sin \theta \hat{\mathbf{w}} \times \vec{r}_i + (1 - \cos \theta) (\hat{\mathbf{w}} \cdot \vec{r}_i) \hat{\mathbf{w}} \\
 &\quad \Downarrow \\
 \hat{\mathbf{r}}_{i+1} &= \hat{\mathbf{q}} * \hat{\mathbf{r}}_i * \hat{\mathbf{q}}^*
 \end{aligned} \tag{C.16}$$

Demonstração. Através de manipulação simples mostra-se que, se a parte escalar de $\hat{\mathbf{r}}_i$ é nula e se $\hat{\mathbf{q}}$ for definido como anteriormente, então o produto composto $\hat{\mathbf{q}} * \hat{\mathbf{r}}_i * \hat{\mathbf{q}}^*$ tem parte escalar nula e parte vectorial dada por

$$\vec{r}_{i+1} = (q_0^2 - \mathbf{q} \cdot \mathbf{q}) \vec{r}_i + 2q_0 \mathbf{q} \times \vec{r}_i + 2(\mathbf{q} \cdot \vec{r}_i) \mathbf{q}$$

Expandindo as partes escalar e vectorial, substituindo e usando as identidades

$$\begin{aligned}
 2 \sin(\theta/2) \cos(\theta/2) &= \sin \theta \\
 \cos^2(\theta/2) - \sin^2(\theta/2) &= \cos \theta
 \end{aligned}$$

obtem-se a fórmula de Rodrigues

$$\vec{\mathbf{r}}_{i+1} = \cos \theta \vec{\mathbf{r}}_i + \sin \theta \hat{\mathbf{w}} \times \vec{\mathbf{r}}_i + (1 - \cos \theta) (\hat{\mathbf{w}} \cdot \vec{\mathbf{r}}_i) \hat{\mathbf{w}}$$

como requerido.

□

Corolário C.1.1. *O quaternião de rotação $\mathring{\mathbf{q}}$ é um **quaternião unitário**¹, ou seja*

$$\|\mathring{\mathbf{q}}\|^2 = \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} = \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* = 1 \quad (\text{C.17})$$

Demonstração. Sabe-se que $\hat{\mathbf{w}}$ é um vector unitário e que $\cos^2 \left(\frac{\theta}{2}\right) + \sin^2 \left(\frac{\theta}{2}\right) = 1$.

Logo, temos

$$\begin{aligned} \|\mathring{\mathbf{q}}\|^2 &= \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} \\ &= q^2 + \|\mathbf{q}\|^2 \\ &= \cos^2 \left(\frac{\theta}{2}\right) + \sin^2 \left(\frac{\theta}{2}\right) (\hat{\mathbf{w}}_x^2 + \hat{\mathbf{w}}_y^2 + \hat{\mathbf{w}}_z^2) \\ &= \cos^2 \left(\frac{\theta}{2}\right) + \sin^2 \left(\frac{\theta}{2}\right) \\ &= 1 \end{aligned}$$

□

Corolário C.1.2. *A rotação dada por $\mathring{\mathbf{q}}$ é a mesma rotação dada por $-\mathring{\mathbf{q}}$.*

Corolário C.1.3. *A rotação inversa de $\mathring{\mathbf{q}}$ (ou seja, a rotação de ângulo $-\theta$ sobre o mesmo eixo) é dada pelo seu conjugado, $\mathring{\mathbf{q}}^*$.*

Demonstração. Esta é uma propriedade muito útil do quaternião unitário, frequentemente usada ao longo deste texto. Pode ser derivada da sua definição e através da equação (C.14):

$$\mathring{\mathbf{q}}^{-1} = \frac{\mathring{\mathbf{q}}^*}{\mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}}} = \frac{\mathring{\mathbf{q}}^*}{1} = \mathring{\mathbf{q}}^* \quad (\text{C.18})$$

□

Corolário C.1.4. *O quaternião unitário correspondendo à rotação resultante da composição de duas rotações arbitrárias é dada pelo produto dos quaterniões de rotação componentes.*

¹Não confundir *quaternião unitário*, que é um quaternião de norma unitária, com o *quaternião unidade*, definido anteriormente.

Demonstração. Considere-se mais uma vez a rotação dada por $\mathring{\mathbf{r}}_{i+1} = \mathring{\mathbf{q}} * \mathring{\mathbf{r}}_i * \mathring{\mathbf{q}}^*$. Suponha-se que uma segunda rotação é aplicada, representada pelo quaternião unitário $\mathring{\mathbf{p}}$. Temos

$$\begin{aligned}\mathring{\mathbf{r}}_{i+2} &= \mathring{\mathbf{p}} * \mathring{\mathbf{r}}_{i+1} * \mathring{\mathbf{p}}^* \\ &= \mathring{\mathbf{p}} * (\mathring{\mathbf{q}} * \mathring{\mathbf{r}}_i * \mathring{\mathbf{q}}^*) * \mathring{\mathbf{p}}^*\end{aligned}$$

Usando a propriedade da multiplicação de quaterniões dada pela equação (C.13), tem-se, como era esperado,

$$\begin{aligned}\mathring{\mathbf{r}}_{i+2} &= (\mathring{\mathbf{p}} * \mathring{\mathbf{q}}) * \mathring{\mathbf{r}}_i * (\mathring{\mathbf{p}} * \mathring{\mathbf{q}})^* \\ &= \mathring{\mathbf{s}} * \mathring{\mathbf{r}}_i * \mathring{\mathbf{s}}^*\end{aligned}$$

e, portanto, a rotação global é dada por $\mathring{\mathbf{s}} = \mathring{\mathbf{p}} * \mathring{\mathbf{q}}$. □

Corolário C.1.5. *A representação ortonormal para a rotação $\mathbf{R}$ está relacionada com a correspondente representação em quaternião dual $\mathring{\mathbf{q}}$ através de*

$$\begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{R} \end{bmatrix} = \mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{Q}(\mathring{\mathbf{q}}) \quad (\text{C.19})$$

Demonstração. Pode ser facilmente demonstrado a partir da definição e propriedades da multiplicação de quaterniões que

$$\begin{aligned}\mathring{\mathbf{r}}_{i+1} &= \mathring{\mathbf{q}} * \mathring{\mathbf{r}}_i * \mathring{\mathbf{q}}^* \\ &= \mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{Q}(\mathring{\mathbf{q}}) \mathring{\mathbf{r}}\end{aligned}$$

O produto de matrizes auxiliares mostrado acima é dado por

$$\mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{Q}(\mathring{\mathbf{q}}) \mathring{\mathbf{r}} = \begin{bmatrix} \mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} & 0^T \\ 0 & \mathbf{R} \end{bmatrix}$$

onde $\mathbf{R}$ é uma matriz 3×3 .

Visto que $\mathring{\mathbf{q}}$ é um quaternião unitário, por definição $\mathring{\mathbf{q}} \cdot \mathring{\mathbf{q}} = 1$ e demonstra-se facilmente que $\mathbf{W}(\mathring{\mathbf{q}})^T \mathbf{Q}(\mathring{\mathbf{q}})$, e consequentemente $\mathbf{R}$, são ortonormais. Adicionalmente, sendo $\mathring{\mathbf{r}}_i$ e $\mathring{\mathbf{r}}_{i+1}$ puramente imaginários, através de simples manipulação algébrica mostra-se que

$$\begin{aligned} \mathring{\mathbf{r}}_{i+1} &= \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{R} \end{bmatrix} \mathring{\mathbf{r}}_i \\ &\Downarrow \\ \vec{\mathbf{r}}_{i+1} &= \mathbf{R} \vec{\mathbf{r}}_i \end{aligned}$$

□

C.2 O Conceito de Número Dual

Os números duais foram inventados por W. Clifford (veja-se [Clifford 1873]) e foram posteriormente aprofundados por E. Study (consulte-se [Study 1891] e [Study 1903]), nos finais do século XIX, inícios do século XX [Walker et al. 1991, Daniilidis 1998].

Um número dual genérico $\check{a} = a + \varepsilon b$ pode ser definido como uma combinação de um par de entidades reais ordenadas com o mesmo número de dimensões (por exemplo, números reais, vectores de números reais, matrizes de números reais, etc.) a e b com uma regra de multiplicação especial para ε dada por $\varepsilon^2 = 0$. As duas entidades a e b podem ser descritas como pertencendo à parte real e à parte dual do número dual, respectivamente [Walker et al. 1991].

A adição, subtracção e multiplicação de números duais são definidas através das fórmulas que se seguem [Walker et al. 1991]

$$(a + \varepsilon b) + (c + \varepsilon d) = (a + c) + \varepsilon (b + d) \quad (\text{C.20a})$$

$$(a + \varepsilon b) - (c + \varepsilon d) = (a - c) + \varepsilon (b - d) \quad (\text{C.20b})$$

$$(a + \varepsilon b)(c + \varepsilon d) = ac + \varepsilon(ad + bc) \quad (\text{C.20c})$$

As operações de adição e de multiplicação fazem com que o caso especial de números duais com componentes real e dual unidimensionais seja um anel abeliano chamado Δ contendo elementos singulares, visto que apenas números duais com uma parte real não nula possuem um elemento inverso [Daniilidis 1998].

Vectores duais tridimensionais, por outro lado, estão definidos em Δ^3 , e com a adição e o produto externo compõem um módulo sobre o anel Δ . Vectores duais tridimensionais com partes real e dual ortogonais são uma representação de rectas em $\mathbb{R}^3$ conhecida como *coordenadas de Plücker*. A parte real é a direcção da recta e a

parte dual é o seu momento. O produto interno entre dois vectores duais deste tipo é igual ao co-seno do *ângulo dual* $\check{\theta} = \theta + \varepsilon d$, introduzido por Study na sua investigação (veja-se [Study 1903]), que tem uma interpretação geométrica muito útil: θ é o ângulo e d a distância entre duas rectas no espaço [Daniilidis 1998, Walker et al. 1991]. Estas duas rectas podem por exemplo definir movimentos helicoidais [Walker et al. 1991] — este facto será utilizado mais adiante neste texto.

Uma propriedade essencial de funções com números duais como argumentos e o seu corolário envolvendo o ângulo dual, que será usado em secções posteriores, será apresentada de seguida [Walker et al. 1991, Daniilidis 1998].

Lema C.2.1. *Seja $\check{a} = a + \varepsilon b$ um número dual number com os números (unidimensionais) reais a e b como as suas partes real e dual, respectivamente. A sua expansão em série de Taylor no ponto a é dada por*

$$f(\check{a}) = f(a + \varepsilon b) = f(a) + \varepsilon b f'(a) \quad (\text{C.21})$$

Demonstração. De acordo com o teorema de expansão em série de Taylor, se uma função $f(\check{a})$ é analítica dentro do círculo $|\check{a} - c| < R$ (com $R > 0$), onde c é o centro desse círculo, então $f(\check{a})$ pode ser expandida numa série de Taylor única nesse círculo,

$$f(\check{a}) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!} (\check{a} - c)^n$$

Se se expandir $f(\check{a})$ no ponto a , a série de Taylor será

$$f(\check{a}) = f(a + \varepsilon b) = f(a) + \varepsilon b f'(a) + \varepsilon^2 \frac{b^2}{2} f''(a) + \dots$$

Contudo, visto que a definição de números duais afirma que $\varepsilon^2 = 0$, todos os termos da série de Taylor com expoentes de ε maiores do que um anulam-se.

Portanto,

$$f(\check{a}) = f(a + \varepsilon b) = f(a) + \varepsilon b f'(a)$$

□

Corolário C.2.2. *O seno e o co-seno do ângulo dual $\check{\theta} = \theta + \varepsilon d$ são dados por*

$$\cos(\check{\theta}) = \cos(\theta + \varepsilon d) = \cos(\theta) - \varepsilon d \sin(\theta) \quad (\text{C.22a})$$

$$\sin(\check{\theta}) = \sin(\theta + \varepsilon d) = \sin(\theta) + \varepsilon d \cos(\theta) \quad (\text{C.22b})$$

Demonstração. A demonstração deste corolário é obtida directamente através da expansão consecutiva em série de Taylor de $\sin(\check{\theta})$ e $\cos(\check{\theta})$ em θ , através do Lema C.2.1. $\square$

C.3 Movimento Helicoidal e Quaterniões Unitários Duais

Os quaterniões duais são definidos de forma similar aos quaterniões reais. Um quaternião dual é dado genericamente por

$$\check{\mathbf{q}} = \begin{bmatrix} \check{q}_0 \\ \check{\mathbf{q}} \end{bmatrix} \quad (\text{C.23})$$

onde $\check{q}_0$ é um número dual unidimensional e $\check{\mathbf{q}}$ é um vector 3D dual.

Dado que as operações são definidas da mesma forma que com os quaterniões reais, os quaterniões duais são similarmente um módulo- Δ , um anel não-abeliano com o quaternião unidade como elemento identidade, e uma álgebra não-comutativa e associativa [Daniilidis 1998].

Qualquer vector dual tridimensional, como foi visto anteriormente com vectores 3D e quaterniões reais, pode ser escrito como um quaternião dual puramente imaginário e a sua multiplicação possui a mesma útil propriedade, dada por

$$\begin{bmatrix} 0 \\ \check{\mathbf{r}} \end{bmatrix} * \begin{bmatrix} 0 \\ \check{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} -\check{\mathbf{r}}^T \check{\mathbf{q}} \\ \check{\mathbf{r}} \times \check{\mathbf{q}} \end{bmatrix} \quad (\text{C.24})$$

A quadrado da norma de um quaternião dual é definida como

$$\|\check{\mathbf{q}}\|^2 = \check{\mathbf{q}} * \check{\mathbf{q}}^* = \check{\mathbf{q}} \cdot \check{\mathbf{q}} \quad (\text{C.25})$$

e é um número dual com parte real positiva.

Se a norma tem *uma parte real não-nula*, então o quaternião dual tem um inverso, que tem uma forma similar ao inverso de um quaternião real [Daniilidis 1998].

Um caso especial importante do quaternião dual é o *quaternião dual unitário*, definido de forma análoga como anteriormente como um quaternião dual com norma unitária — neste caso, o seu inverso é também analogamente igual ao seu conjugado.

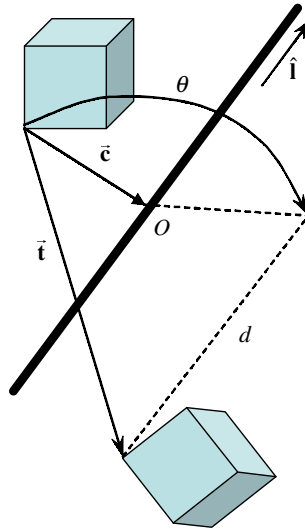


Figura C.2: A geometria do movimento helicoidal (“em rosca”): qualquer transformação de corpo rígido pode ser modelada através de uma rotação com ângulo θ em volta do eixo passando por $\vec{c}$ com direcção $\hat{i}$ e uma translação subsequente d paralela à direcção desse eixo [Daniilidis 1998].

Visto que um quaternião dual pode obviamente ser definido de forma alternativa como $\check{\mathbf{q}} = \mathbf{\hat{q}} + \varepsilon \mathbf{\hat{q}}'$, as suas condições de norma unitária podem ser derivadas usando a operação multiplicação de números duais da pela equação (C.20c) como sendo [Daniilidis 1998]

$$\begin{aligned} \|\check{\mathbf{q}}\|^2 &= \mathbf{\hat{q}} * \mathbf{\hat{q}}^* + \varepsilon (\mathbf{\hat{q}} * \mathbf{\hat{q}}'^* + \mathbf{\hat{q}}' * \mathbf{\hat{q}}^*) \\ &= 1 \Rightarrow \begin{cases} \mathbf{\hat{q}} * \mathbf{\hat{q}}^* = 1 \\ \mathbf{\hat{q}} * \mathbf{\hat{q}}'^* + \mathbf{\hat{q}}' * \mathbf{\hat{q}}^* = 0 \end{cases} \end{aligned} \quad (\text{C.26})$$

Seguidamente, o papel do quaternião dual como uma representação matemática alternativa importante para transformações de corpos rígidos será demonstrado.

De acordo com o teorema de Chasles’ (consulte-se Chen [1991] para mais informação), uma transformação rígida, dada por uma rotação em volta de um eixo passando pela origem seguida de uma translação, pode ser alternativamente modelada como uma rotação *com o mesmo ângulo* em volta de um eixo, denominado *eixo central*, que *não passe* pela origem (cf. Teorema C.1.1) e uma translação simultânea ao longo desse eixo, descrevendo uma trajectória “em rosca” — esta é a conhecida noção cinemática de *movimento helicoidal* [Daniilidis 1998].

A geometria do movimento helicoidal (consulte-se a figura C.2) vai ser agora examinada com mais detalhe de modo a tirar conclusões sobre a sua relação com quaterniões

duais.

Uma recta no espaço com direcção dada pelo vector unitário $\hat{\mathbf{l}}$ passando pelo ponto $\vec{\mathbf{c}}$ pode ser representada pelo par ordenado tridimensional (à primeira vista com 6 graus de liberdade) $(\hat{\mathbf{l}}, \vec{\mathbf{m}})$ a que é dado o nome de par *direcção e momento*, onde o momento é dado por $\vec{\mathbf{m}} = \vec{\mathbf{c}} \times \hat{\mathbf{l}}$. Note-se que o momento da recta é perpendicular ao plano contendo a recta e a origem com magnitude igual à distância dessa recta à origem. Logo, duas restrições, $\|\hat{\mathbf{l}}\|^2 = 1$ e $\vec{\mathbf{c}}^T \hat{\mathbf{l}} = 0$, garantem que os graus de liberdade de uma recta arbitrária no espaço são efectivamente quatro [Daniilidis 1998].

Dado que o eixo central do movimento helicoidal é uma recta no espaço, este depende dos quatro parâmetros implícitos no seu par direcção e momento correspondendo a cada um dos seus graus de liberdade; estes parâmetros, juntamente com o ângulo de rotação θ e a distância de translação ao longo do eixo central d (a que também se dá o nome de *passo da “rosca”* ou helicóide, ou simplesmente *passo*), constituem os seis graus de liberdade de uma transformação de corpo rígido [Daniilidis 1998].

Como pode ser visto na figura C.2, o versor de direcção $\hat{\mathbf{l}}$ é paralelo ao eixo de rotação. O passo d é a projecção do vector de translação sobre eixo central, e portanto $d = \vec{\mathbf{t}}^T \hat{\mathbf{l}}$. Como foi mencionado acima, o ângulo θ é o mesmo tanto no convencional $(\mathbf{R}, \vec{\mathbf{t}})$ como na representação helicoidal [Daniilidis 1998].

Primeiramente, o sistema de coordenadas, que na figura C.2 coincide originalmente com um dos vértices e as arestas correspondentes do cubo tridimensional representando o corpo rígido, é deslocado para o ponto $\vec{\mathbf{c}}$ — este ponto é a projecção da origem do sistema de coordenadas no eixo central — e seguidamente a transformação é realizada [Daniilidis 1998].

Como no Teorema C.1.1, usa-se a fórmula de Rodrigues (notando o facto de que o eixo de rotação, depois da deslocação do sistema de coordenadas, passa agora pela origem) de forma a obter [Daniilidis 1998]

$$\mathbf{R} \vec{\mathbf{c}} = \vec{\mathbf{c}} + \sin \theta \hat{\mathbf{l}} \times \vec{\mathbf{c}} + (1 - \cos \theta) \hat{\mathbf{l}} \times (\hat{\mathbf{l}} \times \vec{\mathbf{c}})$$

Visto que $\vec{\mathbf{c}}^T \hat{\mathbf{l}} = 0$, e que da geometria inerente ao movimento helicoidal $\vec{\mathbf{t}} = d\hat{\mathbf{l}} + (\mathbf{I} - \mathbf{R}) \vec{\mathbf{c}}$, segue-se que [Daniilidis 1998]

$$\vec{\mathbf{c}} = \frac{1}{2} \left(\vec{\mathbf{t}} - (\vec{\mathbf{t}}^T \hat{\mathbf{l}}) \hat{\mathbf{l}} + \hat{\mathbf{l}} \times \vec{\mathbf{t}} \cot \frac{\theta}{2} \right)$$

Deste resultado torna-se óbvio que $\vec{\mathbf{c}}$, e portanto o eixo central, *não estão definidos se o ângulo θ for 0 ou π* . Em todos os outros casos, o vector de momento torna-

-se [Daniilidis 1998]

$$\vec{\mathbf{m}} = \vec{\mathbf{c}} \times \hat{\mathbf{l}} = \frac{1}{2} \left(\vec{\mathbf{t}} \times \hat{\mathbf{l}} + \hat{\mathbf{l}} \times (\vec{\mathbf{t}} \times \hat{\mathbf{l}}) \cot \frac{\theta}{2} \right) \quad (\text{C.27})$$

Seguidamente, mostrar-se-á que transformações rígidas de rectas no espaço podem ser representadas muito a propósito como um produto de quaterniões duais unitários, de uma forma muito similar a como as rotações no Teorema C.1.1 eram representadas por um produto de quaterniões reais unitários [Daniilidis 1998].

Lema C.3.1. *Seja $(\hat{\mathbf{l}}_i, \vec{\mathbf{m}}_i)$ o par direcção e momento de uma recta no espaço. Transforme-se esta recta com $(\mathbf{R}, \vec{\mathbf{t}})$ numa recta com par direcção e momento $(\hat{\mathbf{l}}_{i+1}, \vec{\mathbf{m}}_{i+1})$. Se ambas as rectas forem representadas como quaterniões duais puramente imaginários dados por*

$$\check{\mathbf{l}}_i = \mathring{\mathbf{l}}_i + \varepsilon \mathring{\mathbf{m}}_i$$

e

$$\check{\mathbf{l}}_{i+1} = \mathring{\mathbf{l}}_{i+1} + \varepsilon \mathring{\mathbf{m}}_{i+1}$$

com

$$\mathring{\mathbf{l}}_i = \begin{bmatrix} 0 \\ \hat{\mathbf{l}}_i \end{bmatrix}, \quad \mathring{\mathbf{m}}_i = \begin{bmatrix} 0 \\ \vec{\mathbf{m}}_i \end{bmatrix}$$

e

$$\mathring{\mathbf{l}}_{i+1} = \begin{bmatrix} 0 \\ \hat{\mathbf{l}}_{i+1} \end{bmatrix}, \quad \mathring{\mathbf{m}}_{i+1} = \begin{bmatrix} 0 \\ \vec{\mathbf{m}}_{i+1} \end{bmatrix}$$

respectivamente, então existe um quaternião dual unitário $\check{\mathbf{q}} = \mathring{\mathbf{q}} + \varepsilon \mathring{\mathbf{q}}'$ que permite a representação dessa transformação como

$$\check{\mathbf{l}}_{i+1} = \check{\mathbf{q}} * \mathring{\mathbf{l}}_i * \check{\mathbf{q}}^* \quad (\text{C.28})$$

O quaternião unitário $\mathring{\mathbf{q}} = [\cos \frac{\theta}{2} \sin \frac{\theta}{2} \hat{\mathbf{w}}]^T$ é o quaternião de rotação como apresentado no Teorema C.1.1 e o quaternião unitário $\mathring{\mathbf{q}}' = \frac{1}{2} \mathring{\mathbf{t}} * \mathring{\mathbf{q}}$ é metade do produto do quaternião de rotação com o quaternião puramente imaginário correspondendo ao vector de translação $\vec{\mathbf{t}}$.

Demonstração. Aplicando uma rotação $\mathbf{R}$ e uma translação $\vec{\mathbf{t}}$ a uma recta dada por $(\hat{\mathbf{l}}_i, \vec{\mathbf{m}}_i)$ obtém-se a recta transformada $(\hat{\mathbf{l}}_{i+1}, \vec{\mathbf{m}}_{i+1})$ com

$$\hat{\mathbf{l}}_{i+1} = \mathbf{R}\hat{\mathbf{l}}_i$$

e

$$\begin{aligned}\vec{\mathbf{m}}_{i+1} &= \vec{\mathbf{c}}_{i+1} \times \hat{\mathbf{l}}_{i+1} = (\mathbf{R}\vec{\mathbf{c}}_i + \vec{\mathbf{t}}) \times (\mathbf{R}\hat{\mathbf{l}}_i) \\ &= \mathbf{R}(\vec{\mathbf{c}}_i \times \hat{\mathbf{l}}_i) + \vec{\mathbf{t}} \times (\mathbf{R}\hat{\mathbf{l}}_i) \\ &= \mathbf{R}\vec{\mathbf{m}}_i + \vec{\mathbf{t}} \times (\mathbf{R}\hat{\mathbf{l}}_i)\end{aligned}$$

Se se mudar de notação puramente vectorial para notação vectorial de quaterniões, os termos com rotações são fáceis de representar, o único problema sendo o termo de produto externo restante. Todavia, considerando a equação (C.10) e que $\hat{\mathbf{l}}_{i+1} = \mathbf{R}\hat{\mathbf{l}}_i$ pode-se representar facilmente este termo verificando que

$$\begin{aligned}\begin{bmatrix} 0 \\ \vec{\mathbf{t}} \times \hat{\mathbf{l}}_{i+1} \end{bmatrix} &= \frac{1}{2} \left(\begin{bmatrix} \vec{\mathbf{t}}^T \hat{\mathbf{l}}_{i+1} - \hat{\mathbf{l}}_{i+1}^T \vec{\mathbf{t}} \\ -\hat{\mathbf{l}}_{i+1} \times \vec{\mathbf{t}} + \vec{\mathbf{t}} \times \hat{\mathbf{l}}_{i+1} \end{bmatrix} \right) \\ &= \frac{1}{2} (\hat{\mathbf{l}}_{i+1} * \vec{\mathbf{t}}^* + \vec{\mathbf{t}} * \hat{\mathbf{l}}_{i+1})\end{aligned}$$

Portanto, a transformação completa em termos de quaterniões torna-se, de acordo com o Teorema C.1.1,

$$\hat{\mathbf{l}}_{i+1} = \hat{\mathbf{q}} * \hat{\mathbf{l}}_i * \hat{\mathbf{q}}^*$$

e

$$\vec{\mathbf{m}}_{i+1} = \hat{\mathbf{q}} * \vec{\mathbf{m}}_i * \hat{\mathbf{q}}^* + \frac{1}{2} (\hat{\mathbf{q}} * \hat{\mathbf{l}}_i * \hat{\mathbf{q}}^* * \vec{\mathbf{t}}^* + \vec{\mathbf{t}} * \hat{\mathbf{q}} * \hat{\mathbf{l}}_i * \hat{\mathbf{q}}^*)$$

Se se usarem as definições para os quaterniões duais unitários dados por este Lema correspondendo a cada recta no espaço, mostra-se facilmente que as equações precedentes equivalem à equação única

$$\hat{\mathbf{l}}_{i+1} + \varepsilon \vec{\mathbf{m}}_{i+1} = (\hat{\mathbf{q}} + \varepsilon \hat{\mathbf{q}}')(\hat{\mathbf{l}}_i + \varepsilon \vec{\mathbf{m}}_i)(\hat{\mathbf{q}}^* + \varepsilon \vec{\mathbf{q}}'^*)$$

Isto é o mesmo que

$$\check{\hat{\mathbf{l}}}_{i+1} = \check{\hat{\mathbf{q}}} * \check{\hat{\mathbf{l}}}_i * \check{\hat{\mathbf{q}}}^*$$

O único facto que resta demonstrar é que $\check{\mathbf{q}}$ é, sem margem para dúvidas, um quaternião dual *unitário*. Isto pode ser feito verificando a sua norma, sabendo que $\mathring{\mathbf{t}}$ é um quaternião puramente imaginário, que $\mathring{\mathbf{q}}' = \frac{1}{2} \mathring{\mathbf{t}} * \mathring{\mathbf{q}}$, que $\mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* = 1$ do Corolário C.1.1, e usando a propriedade (C.13) e a operação de multiplicação de números duais definida na equação (C.20c) de forma a obter

$$\begin{aligned} \|\check{\mathbf{q}}\|^2 &= \check{\mathbf{q}} * \check{\mathbf{q}}^* = \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* + \varepsilon(\mathring{\mathbf{q}} * \mathring{\mathbf{q}}'^* + \mathring{\mathbf{q}}' * \mathring{\mathbf{q}}^*) \\ &= \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* + \frac{\varepsilon}{2}(\mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* * \mathring{\mathbf{t}}^* + \mathring{\mathbf{t}} * \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^*) \\ &= 1 + \frac{\varepsilon}{2}(\mathring{\mathbf{t}}^* + \mathring{\mathbf{t}}) \\ &= 1 + \frac{\varepsilon}{2} \left(\begin{bmatrix} 0 \\ -\vec{\mathbf{t}} \end{bmatrix} + \begin{bmatrix} 0 \\ \vec{\mathbf{t}} \end{bmatrix} \right) \\ &= 1 \end{aligned}$$

□

Este Lema vai ser usado em seguida para formular um teorema que mostra que quaterniões duais unitários podem ser usados para representar o movimento helicoidal associado a uma transformação de corpo rígido, como provado no decurso investigação de Konstantinos Daniilidis em [Daniilidis 1998], e um corolário que mostra a relação matemática entre os componentes desse quaternião unitário dual e os parâmetros da transformação de corpo rígido correspondente, como foi apresentado no trabalho de Michael Walker *et al's* em [Walker et al. 1991].

Teorema C.3.1. *Seja $(\hat{\mathbf{l}}, \vec{\mathbf{m}})$ o par direcção e momento do eixo central do movimento helicoidal com passo d e ângulo de rotação θ correspondendo à transformação rígida de parâmetros $(\mathbf{R}, \vec{\mathbf{t}})$, como afirmado no teorema de Chasles'. Se*

$$\check{\mathbf{l}} = \vec{\mathbf{l}} + \varepsilon \vec{\mathbf{m}}$$

então, em concordância com o Lema C.3.1, existe um quaternião dual unitário $\check{\mathbf{q}} = \mathring{\mathbf{q}} + \varepsilon \mathring{\mathbf{q}}'$ que corresponde à transformação aplicada a um ponto arbitrário com vector de posição $\vec{\mathbf{r}}_i$ para o deslocar até $\vec{\mathbf{r}}_{i+1}$. Este quaternião dual unitário é dado por

$$\check{\mathbf{q}} = \begin{bmatrix} \cos\left(\frac{\check{\theta}}{2}\right) \\ \sin\left(\frac{\check{\theta}}{2}\right)\check{\mathbf{l}} \end{bmatrix} \quad (\text{C.29})$$

onde $\check{\theta} = \theta + \varepsilon d$ é denominado de ângulo dual, como definido por Study em [Study 1903]. Dado que $\check{\mathbf{q}}$ tem norma unitária, esta correspondência apenas se mantém se as restrições seguintes, as chamadas **condições ou restrições de norma unitária** para quaterniões duais, forem satisfeitas:

$$\begin{aligned} \|\check{\mathbf{q}}\|^2 &= \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* + \varepsilon(\mathring{\mathbf{q}} * \mathring{\mathbf{q}}'^* + \mathring{\mathbf{q}}' * \mathring{\mathbf{q}}^*) \\ &= 1 \Rightarrow \begin{cases} \mathring{\mathbf{q}} * \mathring{\mathbf{q}}^* = 1 \\ \mathring{\mathbf{q}} * \mathring{\mathbf{q}}'^* + \mathring{\mathbf{q}}' * \mathring{\mathbf{q}}^* = 0 \end{cases} \end{aligned} \quad (\text{C.30})$$

Demonstração. Visto que o eixo central de uma trajectória helicoidal é uma recta no espaço, então, de acordo com o Lema C.3.1 e o Teorema C.1.1, a parte real do quaterniões dual unitário requerido correspondendo à matriz de rotação $\mathbf{R}$ é dado por

$$\mathring{\mathbf{q}} = \begin{bmatrix} q_0 \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \hat{\mathbf{l}} \end{bmatrix}$$

Logo, pode-se rescrever a equação do momento da recta (C.27) como

$$\sin\left(\frac{\theta}{2}\right) \vec{\mathbf{m}} = \frac{1}{2} \left(\vec{\mathbf{t}} \times \mathbf{q} + q_0 \vec{\mathbf{t}} - (\hat{\mathbf{l}}^T \vec{\mathbf{t}}) \hat{\mathbf{l}} \cos\left(\frac{\theta}{2}\right) \right)$$

Usando $d = \vec{\mathbf{t}}^T \hat{\mathbf{l}}$, tem-se

$$\sin\left(\frac{\theta}{2}\right) \vec{\mathbf{m}} + \frac{d}{2} \hat{\mathbf{l}} \cos\left(\frac{\theta}{2}\right) = \frac{1}{2} \left(q_0 \vec{\mathbf{t}} + \vec{\mathbf{t}} \times \mathbf{q} \right)$$

Aplicando de novo o Lema C.3.1 e a definição de multiplicação de quaterniões dada por (C.9), a parte dual do quaterniões dual unitário desejado é dada por

$$\mathring{\mathbf{q}}' = \frac{1}{2} \mathring{\mathbf{t}} * \mathring{\mathbf{q}} = \begin{bmatrix} -\frac{1}{2} \mathbf{q}^T \vec{\mathbf{t}} \\ \frac{1}{2} \left(q_0 \vec{\mathbf{t}} + \vec{\mathbf{t}} \times \mathbf{q} \right) \end{bmatrix}$$

Reunindo estes resultados, tem-se a expressão completa para o quaterniões dual unitário:

$$\begin{aligned} \check{\mathbf{q}} &= \begin{bmatrix} q_0 \\ \mathbf{q} \end{bmatrix} + \varepsilon \begin{bmatrix} -\frac{1}{2} \mathbf{q}^T \vec{\mathbf{t}} \\ \frac{1}{2} \left(q_0 \vec{\mathbf{t}} + \vec{\mathbf{t}} \times \mathbf{q} \right) \end{bmatrix} \\ &= \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \hat{\mathbf{l}} \end{bmatrix} + \varepsilon \begin{bmatrix} -d \sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \vec{\mathbf{m}} + \frac{d}{2} \hat{\mathbf{l}} \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \end{aligned}$$

Usando o Lema C.2.1 e o Corolário C.2.2, é fácil provar que

$$\check{\mathbf{q}} = \begin{bmatrix} \cos\left(\frac{\theta+\varepsilon d}{2}\right) \\ \sin\left(\frac{\theta+\varepsilon d}{2}\right)(\check{\mathbf{I}} + \varepsilon\check{\mathbf{m}}) \end{bmatrix}$$

Efectuando as últimas substituições considerando as definições de ângulo dual $\check{\theta}$ e de $\check{\mathbf{I}}$ tem-se

$$\check{\mathbf{q}} = \begin{bmatrix} \cos\left(\frac{\check{\theta}}{2}\right) \\ \sin\left(\frac{\check{\theta}}{2}\right)\check{\mathbf{I}} \end{bmatrix}$$

que já tinha sido provado no Lema C.3.1 como tendo norma unitária, como requerido. $\square$

Corolário C.3.2. *A transformação rígida com parâmetros $(\mathbf{R}, \vec{\mathbf{t}})$ que transforma o ponto $\mathbf{p}_i$ em $\mathbf{p}_{i+1}$ pode ser derivada do quaterniã dual unitário $\check{\mathbf{q}}$ correspondendo ao movimento helicoidal relacionado com essa transformação usando*

$$\check{\mathbf{p}}_{i+1} = \mathbf{W}(\check{\mathbf{q}})^T \mathbf{Q}(\check{\mathbf{q}}) \check{\mathbf{p}}_i + 2\mathbf{W}(\check{\mathbf{q}})^T \check{\mathbf{q}}' \quad (\text{C.31})$$

onde $\check{\mathbf{q}}$ e $\check{\mathbf{q}}'$ são as partes real e dual do quaterniã dual, respectivamente, e $\check{\mathbf{p}}_i$ e $\check{\mathbf{p}}_{i+1}$ são os quaterniões puramente imaginários correspondendo a $\mathbf{p}_i$ e $\mathbf{p}_{i+1}$, respectivamente.

Demonstração. A transformação de corpo rígido envolvida é tradicionalmente dada por

$$\mathbf{p}_{i+1} = \mathbf{R}\mathbf{p}_i + \vec{\mathbf{t}}$$

Usando as propriedades de multiplicação de quaterniões, do Teorema C.1.1 tem-se que a rotação de $\mathbf{p}_i$ dada por $\mathbf{R}$ pode ser representada em notação de quaterniões como

$$\check{\mathbf{q}} * \check{\mathbf{p}}_i * \check{\mathbf{q}}^* = \mathbf{W}(\check{\mathbf{q}})^T \mathbf{Q}(\check{\mathbf{q}}) \check{\mathbf{p}}_i$$

e do Teorema C.3.1, pode-se derivar

$$\check{\mathbf{q}}' = \frac{1}{2} \check{\mathbf{t}} * \check{\mathbf{q}} \Leftrightarrow \check{\mathbf{t}} = 2 \check{\mathbf{q}}' * \check{\mathbf{q}}^* = 2\mathbf{W}(\check{\mathbf{q}})^T \check{\mathbf{q}}'$$

$\square$

Corolário C.3.3. *O inverso de uma transformação de corpo rígido pode ser representada pelo conjugado do quaterniã dual unitário que corresponde a essa transformação.*

Demonstração. Este corolário é provado analogamente ao Corolário C.1.3. $\square$

A expressão (C.31) e o Corolário C.3.2 levaram ao desenvolvimento de várias soluções em forma fechada para problemas de estimação no sentido dos mínimos quadrados como a que é apresentada no trabalho de investigação de *et al* (consulte-se [Walker et al. 1991]), mostrando as vantagens no uso desta notação para a representação de transformações rígidas genéricas. Todavia, como pode ser constatado no corolário seguinte, ainda existe uma outra vantagem no seu uso que será frequentemente explorada ao longo do texto principal desta tese, que assume protagonismo quando se lida com composição de transformações [Daniilidis 1998].

Corolário C.3.4. *Uma transformação de corpo rígido composta por duas ou mais transformações é derivada da concatenação de todos os quaterniões duais unitários correspondentes através da operação de multiplicação de quaterniões.*

Demonstração. Este facto é facilmente demonstrável através da verificação do factos seguintes:

- O Lema C.3.1 e o Teorema C.3.1 mostram que a maneira como os quaterniões duais unitários se relacionam com transformações rígidas depende da transformação do eixo central do movimento helicoidal correspondente através de $\check{\mathbf{q}} * \check{\mathbf{l}} * \check{\mathbf{q}}^*$.
- Visto que esta expressão tem a mesma forma que a relação que permite a representação de rotações através de quaterniões unitários reais, isto implica que o Corolário C.1.4 também se aplica a quaterniões duais unitários nesta situação.

$\square$

Apêndice D

Resolução de Problemas de Mínimos Quadrados Não Lineares Sem Restrições — o Método de Levenberg-Marquardt

Quando a relação mais simples possível estabelecida entre vários valores conhecidos (geralmente medidas de algum tipo) e as incógnitas que se quer estimar é *não linear*, estamos perante o problema de estimação de parâmetros mais genérico que existe envolvendo a minimização de uma função-objectivo. Neste caso, apenas métodos iterativos não lineares, como os descritos em Chapra and Canale [2002b] e dos quais os mais facilmente reconhecíveis serão o método de descida máxima e o método Newton, podem ser usados para a estimação de parâmetros.

Qualquer método não linear necessita de valores iniciais para as estimativas dos parâmetros para a primeira iteração, e consequentemente, estes métodos requerem mais do que uma mera formulação teórica do problema — precisam também de uma boa análise das suas condições práticas. De outra forma, qualquer destes métodos terá uma alta probabilidade em convergir, não para um mínimo global que representaria a solução desejada, e sim para um qualquer mínimo local indesejado mais próximo [Chapra and Canale 2002b].

Neste trabalho, o robusto algoritmo de Levenberg-Marquardt foi escolhido para resolver este tipo de problemas — este algoritmo vai ser sumariado nos parágrafos seguintes.

Considere-se a seguinte função-objectivo não linear genérica a ser minimizada

$$\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \mathbf{0} \quad (\text{D.1})$$

onde $\boldsymbol{\theta}$ e $\boldsymbol{\phi}$ são genericamente denominados de *vector de medidas* e *vector de incógnitas*, respectivamente. O objectivo é portanto estimar $\boldsymbol{\phi}$ de forma a minimizar $\mathbf{f}$.

Durante a iteração j , a actualização $\Delta\boldsymbol{\phi}$ para a estimativa presente $\boldsymbol{\phi}_j$ é derivada da expansão de Taylor de primeira ordem de (D.1) [Prager et al. 1997, Hartley and Zisserman 2000]

$$\begin{aligned} \mathbf{0} = \mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}) &\approx \mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_j) + \frac{\partial \mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_j)}{\partial \boldsymbol{\phi}} (\boldsymbol{\phi} - \boldsymbol{\phi}_j) \\ \Rightarrow \Delta\mathbf{f} &= \mathbf{J}(\boldsymbol{\phi} - \boldsymbol{\phi}_j) = \mathbf{J}\Delta\boldsymbol{\phi} \end{aligned} \quad (\text{D.2})$$

onde $\Delta\mathbf{f} = -\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_j)$ é o vector de erro e $\mathbf{J} = \frac{\partial \mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_j)}{\partial \boldsymbol{\phi}}$ é a matriz de gradientes, também chamada de matriz Jacobiana.

O algoritmo de Levenberg-Marquardt devolve os valores actualizados para os parâmetros $\boldsymbol{\phi}_{j+1}$ através das *equações normais aumentadas* $\mathbf{N}'\Delta\mathbf{f} = (\mathbf{J}^T\mathbf{J} + \varepsilon\mathbf{I})^{-1}\mathbf{J}^T\Delta\mathbf{f}$ como sendo [Prager et al. 1997, Hartley and Zisserman 2000]

$$\boldsymbol{\phi}_{j+1} = \boldsymbol{\phi}_j + \mathbf{N}'\Delta\mathbf{f} \quad (\text{D.3})$$

onde ε é o factor de amortecimento escolhido em cada iteração de forma a estabilizar a convergência do algoritmo. A cada passo, $\Delta\mathbf{f}$ e $\mathbf{J}$ são avaliados em função da estimativa presente $\boldsymbol{\phi}_j$. Este processo é iterado até que as correcções $\Delta\boldsymbol{\phi}$ sejam suficientemente pequenas [Prager et al. 1997, Hartley and Zisserman 2000].

Note-se que a popularidade deste algoritmo reside no facto de que este se aproxima do método de descida máxima quando ε é grande, efectivamente acelerando a sua convergência em direcção ao mínimo mais próximo, e do método de Newton quando a convergência está próxima de se consumir e $\varepsilon \approx 0$ (visto que as equações normais aumentadas se tornam nas equações normais regulares $\mathbf{N}\Delta\mathbf{f} = \mathbf{J}^T\mathbf{J}\Delta\mathbf{f}$ do método de Newton), ganhando assim precisão nessa mesma convergência [Hartley and Zisserman 2000, Prager et al. 1997, Chapra and Canale 2002b].

Apêndice E

Resolução de Problemas de Mínimos Quadrados Lineares

Qualquer matriz genérica $\mathbf{A}_{m \times n}$ pode ser decomposta como sendo

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (\text{E.1})$$

onde $\mathbf{U}$ é uma matriz ortogonal $m \times m$, $\mathbf{V}$ é uma matriz ortogonal $n \times n$ e $\mathbf{\Sigma}$ é uma matriz $m \times n$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix} \quad (\text{E.2})$$

cujos elementos não diagonais são todos nulos e cujos elementos diagonais satisfazem $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. Esta é a denominada Decomposição em Valores Singulares (*Singular Value Decomposition*) ou SVD [Leon 1998].

A prova deste facto é apresentada, por exemplo, em Leon [1998], e tem sido reconhecida (por exemplo, em Hartley and Zisserman [2000]) como sendo um dos resultados mais úteis no que respeita a soluções de problemas do campo de Álgebra Linear Numérica. Provavelmente, a sua aplicação mais vantajosa é no seu uso na solução de uma grande variedade problemas de mínimos quadrados.

Outra ferramenta indispensável na resolução de problemas lineares de mínimos quadrados é a teoria de valores/vectores próprios e os multiplicadores de Lagrange [Rao 2002b;a].

Dois destes problemas e as suas soluções correspondentes através de SVD e va-

lores/vectores próprios, usadas neste trabalho, vão ser apresentadas ao longo deste apêndice. Os algoritmos de SVD e de determinação de valores/vectores próprios foram aplicados neste trabalho usando a implementação do MATLAB das funções `svd` e `eig`.

E.1 Solução de sistemas lineares sobre-determinados não homogêneos ($\mathbf{Ax} = \mathbf{b}$)

Se $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ for uma matriz $m \times n$ de característica r , a sua matriz $\mathbf{\Sigma}$ será uma matriz $m \times n$ [Leon 1998],

$$\mathbf{\Sigma} = \left[\begin{array}{c|c} \mathbf{\Sigma}_1 & \mathbf{0} \\ \hline \mathbf{0}^T & 0 \end{array} \right] = \left[\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ \hline & & & & \mathbf{0} \\ & & & & 0 \end{array} \right] \quad (\text{E.3})$$

e pode-se definir [Leon 1998]

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T \quad (\text{E.4})$$

onde $\mathbf{\Sigma}^+$ é matriz $n \times m$ dada por

$$\mathbf{\Sigma}^+ = \left[\begin{array}{c|c} \mathbf{\Sigma}_1^{-1} & \mathbf{0} \\ \hline \mathbf{0}^T & 0 \end{array} \right] = \left[\begin{array}{cccc|c} \frac{1}{\sigma_1} & & & & \\ & \frac{1}{\sigma_2} & & & \\ & & \ddots & & \\ & & & \frac{1}{\sigma_r} & \\ \hline & & & & \mathbf{0} \\ & & & & 0 \end{array} \right] \quad (\text{E.5})$$

A equação (E.4) apresenta uma generalização natural do inverso de uma matriz. Desta forma, a matriz $\mathbf{A}^+$ é denominada a *pseudo-inversa* de $\mathbf{A}$, visto que também satisfaz as quatro propriedades algébricas conhecidas como as *Condições de Penrose*: $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+$, $(\mathbf{AA}^+)^T = \mathbf{AA}^+$ and $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$ (é favor consultar Leon [1998] para mais informação).

Pode-se provar (veja-se Leon [1998]) que, se uma estimativa $\mathbf{x}$ satisfaz o conjunto de $n \times n$ equações lineares denominadas *equações normais*

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$$

para $\mathbf{A}_{m \times n}$, então $\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$ é a solução única para o problema linear de mínimos quadrados do tipo $\mathbf{Ax} = \mathbf{b}$ [Leon 1998, Hartley and Zisserman 2000]. Obviamente, para isto ser verdade, $\mathbf{A}^T \mathbf{A}$ é também provada ser não-singular.

O conceito de matriz pseudo-inversa pode ser usada para resolver problemas deste tipo [Leon 1998, Hartley and Zisserman 2000]. Para ver como isto é feito, consideremos primeiro o caso onde $\mathbf{A}$ é uma matriz $m \times n$ de característica n . Logo Σ é dada por

$$\Sigma = \begin{bmatrix} \Sigma_1 \\ \mathbf{0}^T \end{bmatrix}$$

onde Σ_1 é uma matriz diagonal $n \times n$ não-singular. A matriz $\mathbf{A}^T \mathbf{A}$ é também não singular e

$$\mathbf{A}^T = \mathbf{V} \Sigma^T \mathbf{U}^T \Rightarrow (\mathbf{A}^T \mathbf{A})^{-1} = \mathbf{V} (\Sigma^T \Sigma)^{-1} \mathbf{V}^T$$

Portanto, a solução para as equações normais é dada por

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \Rightarrow \\ \mathbf{x} &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \\ &= \mathbf{V} (\Sigma^T \Sigma)^{-1} \mathbf{V}^T \mathbf{V} \Sigma^T \mathbf{U}^T \mathbf{b} \\ &= \mathbf{V} (\Sigma^T \Sigma)^{-1} \Sigma^T \mathbf{U}^T \mathbf{b} \\ &= \mathbf{V} \Sigma^+ \mathbf{U}^T \mathbf{b} \\ &= \mathbf{A}^+ \mathbf{b} \end{aligned}$$

que efectivamente mostra que, como foi referido anteriormente, $\mathbf{A}^+$ é a generalização do inverso de $\mathbf{A}$. Desta forma, $\mathbf{A}^+ \mathbf{b}$ é a solução do problema linear de mínimos quadrados quando $\mathbf{A}$ tem característica máxima. No caso onde $\mathbf{A}$ tem característica $r < n$ existe uma infinidade de soluções para o problema de mínimos quadrados; porém, pode-se provar (mais uma vez, consulte-se Leon [1998]) que, não só $\mathbf{A}^+ \mathbf{b}$ é uma dessas soluções, visto que $\mathbf{A}^T \mathbf{b}$ pertence ao espaço-coluna de $\mathbf{A}^T \mathbf{A}$, como também que é a *solução mínima* quanto a $\|\mathbf{b} - \mathbf{Ax}\|_2^2$, como requerido [Leon 1998, Hartley and Zisserman 2000].

Finalmente, um último apontamento deve ser tomado: quando a característica de $\mathbf{A}$ é desconhecido, normalmente os valores singulares σ_j abaixo de um certo limiar infinitesimal são postos a zero.

A pseudo-inversa é calculada através da implementação em MATLAB da função `pinv`, que por sua vez usa implicitamente a função `svd`.

E.2 Solução de sistemas lineares sobre-determinados homogêneos ($\mathbf{Ax} = \mathbf{0}$)

Durante a prova da existência universal de uma decomposição em valores singulares de uma matriz $\mathbf{A}_{m \times n}$ apresentada em Leon [1998], várias observações importantes são feitas. Duas destas afirmam que os vectores coluna de $\mathbf{V}$, denotados por $\mathbf{v}_j$ e denominados *vectores singulares direitos* de $\mathbf{A}$, formam uma *base ortonormal* para $N(\mathbf{A})$ para $r + 1 \leq j \leq n$ se $\mathbf{A}$ tiver uma característica $r < n$ [Leon 1998]. Torna-se portanto óbvio que um sistema de equações homogêneas do tipo $\mathbf{Ax} = \mathbf{0}$ tem uma solução não-trivial *se e só se* $\mathbf{A}$ tiver uma característica $r < n$ [Hartley and Zisserman 2000].

O facto de que $\mathbf{v}_j, r + 1 \leq j \leq n$ formam uma base ortonormal para $N(\mathbf{A})$ significa que existe uma infinidade de soluções para $\mathbf{Ax} = \mathbf{0}$ que podem ser formadas por qualquer combinação linear da forma genérica

$$\mathbf{x} = \lambda_1 \mathbf{v}_{r+1} + \dots + \lambda_{n-(r+1)} \mathbf{v}_n \quad (\text{E.6})$$

Por essa razão, essas combinações lineares de soluções para problemas do tipo $\mathbf{Ax} = \mathbf{0}$ são habitualmente restringidas por alguma condição relevante, que normalmente implica a normalização dos elementos de $\mathbf{x}$.

Uma restrição comum em estimação de parâmetros, segundo Zhang [1995], é $\|\mathbf{x}\| = 1$. Problemas do tipo $\mathbf{Ax} = \mathbf{0}$ sujeitos a esta restrição são resolvidos como apresentado em seguida [Hartley and Zisserman 2000].

Considere-se a SVD de $\mathbf{A}$. O problema requer a minimização de $\|\mathbf{U}\Sigma\mathbf{V}^T\mathbf{x}\|$. Porém, visto que $\mathbf{U}$ e $\mathbf{V}$ são ortogonais, as distâncias euclidianas são preservadas (veja-se Leon [1998] ou Hartley and Zisserman [2000]); logo $\|\mathbf{U}\Sigma\mathbf{V}^T\mathbf{x}\| = \|\Sigma\mathbf{V}^T\mathbf{x}\|$ e $\|\mathbf{x}\| = \|\mathbf{V}^T\mathbf{x}\|$. Portanto, torna-se necessário minimizar $\|\Sigma\mathbf{V}^T\mathbf{x}\|$ sujeito à restrição $\|\mathbf{V}^T\mathbf{x}\| = 1$. Efectuando substituições usando $\mathbf{y} = \mathbf{V}^T\mathbf{x}$, o problema torna-se: minimizar $\|\Sigma\mathbf{y}\|$ sujeito a $\|\mathbf{y}\| = 1$. Visto que Σ está na forma apresentada em (E.2), com os seus elementos diagonais em ordem descendente, consequentemente a solução para este problema auxiliar é $\mathbf{y} = [0 \ 0 \ \dots \ 0 \ 1]^T$. Voltando a substituir, torna-se evidente que a solução para o problema original é $\mathbf{v}_n$, ou seja, a última coluna de $\mathbf{V}$.

Por vezes, porém, torna-se conveniente normalizar apenas um subconjunto de elementos de $\mathbf{x}$. Seja $\mathbf{x}'$ um sub-vector de $\mathbf{x}$ contendo os seus n primeiros elementos e $\mathbf{x}''$ um sub-vector contendo os restantes elementos de $\mathbf{x}$. Se se impuser a restrição $\|\mathbf{x}'\|^2 = 1$ de forma a obrigar este vector a ser unitário, pode-se resolver $\mathbf{Ax} = \mathbf{0}$ sujeita a esta restrição, calculando $\mathbf{x}''$ em função de $\mathbf{x}'$.

Para este efeito, torna-se necessário rescrever a equação $\mathbf{Ax} = \mathbf{0}$ como sendo $\mathbf{Cx}' + \mathbf{Dx}'' = \mathbf{0}$. O problema original passa agora a ser formulado como

$$\min_{w', w''} \|\mathbf{O}\|^2, \quad \mathbf{O} = \mathbf{Cx}' + \mathbf{Dx}'' \quad (\text{E.7})$$

sujeito à restrição $\|\mathbf{x}'\|^2 = 1$. Usando multiplicadores de Lagrange (consulte-se Rao [2002a]), isto é equivalente a [Shapiro and Stockman 2001]

$$\min_{w', w''} \left[\|\mathbf{Cx}' + \mathbf{Dx}''\|^2 + \lambda(1 - \|\mathbf{x}'\|^2) \right] \quad (\text{E.8})$$

Este problema pode ser resolvido tomando as derivadas parciais da função objectivo quanto a $\mathbf{x}'$ e $\mathbf{x}''$ e igualando-as a zero [Shapiro and Stockman 2001]:

$$\frac{\partial \mathbf{O}}{\partial \mathbf{x}'} = 2\mathbf{C}^T(\mathbf{Cx}' + \mathbf{Dx}'') - 2\lambda\mathbf{x}' = 0 \quad (\text{E.9a})$$

$$\frac{\partial \mathbf{O}}{\partial \mathbf{x}''} = 2\mathbf{D}^T(\mathbf{Cx}' + \mathbf{Dx}'') = 0 \quad (\text{E.9b})$$

Esta última equação é equivalente a

$$\mathbf{x}'' = -(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{Cx}' \quad (\text{E.10})$$

Substituindo (E.10) na primeira equação, obtém-se

$$\lambda\mathbf{x}' = [\mathbf{C}^T\mathbf{C} - \mathbf{C}^T\mathbf{D}(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{C}] \mathbf{x}' \quad (\text{E.11})$$

É evidente portanto que λ é vector próprio de

$$\mathbf{M} = \mathbf{C}^T\mathbf{C} - \mathbf{C}^T\mathbf{D}(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{C} \quad (\text{E.12})$$

e portanto a solução para $\mathbf{x}'$ é dada pelo vector próprio associado ao valor próprio mais pequeno de $\mathbf{M}$. A equação (E.10) permite calcular consequentemente o vector de coeficientes dependentes $\mathbf{x}''$ dado $\mathbf{x}'$ [Shapiro and Stockman 2001].

Bibliografia

3D-MURALE PROJECT. Web site – <http://www.prip.tuwien.ac.at/Research/Murale>.

3D SCANNERS. 3D Scanners Home Page. Web site – www.3Dscanners.com.

3D SCANNING SYSTEMS. 3D Scanning Systems Home Page. Web site – <http://www.wvl.co.uk>, July 2000.

ANDREFF, N., HORAUD, R., AND ESPIAU, B. Robot Hand-Bye Calibration using Structure-from-Motion. Technical report, INRIA Rhône-Alpes and GRAVIR-IMAG, February 2000. URL citeseer.nj.nec.com/andreff00robot.html.

ARUN, K. S., HUANG, T. S., AND BLOSTEIN, S. D. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-9 (5):698–700, September 1987.

AUDETTE, M. A., FERRIE, F. P., AND PETERS, T. M. An Algorithmic Overview of Surface Registration Techniques for Medical Imaging. *Medical Image Analysis*, 4 (3):201–217, September 2000.

AWCOCK, G. W. AND THOMAS, R. *Applied Image Processing*, chapter 3, pages 85–86. McGraw-Hill, international edition, 1996a. ISBN 0-07-001470-1.

AWCOCK, G. W. AND THOMAS, R. *Applied Image Processing*, chapter 2, pages 39–40. McGraw-Hill, international edition, 1996b. ISBN 0-07-001470-1.

A. CHIUSO, R. B. AND SOATTO, S. Optimal Structure From Motion: Local Ambiguities and Global Estimates. *Intl. J. of Comp. Vision*, 39(3):195–228, September 2000.

BESL, P. J. AND MCKAY, N. D. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

- BLOSTEIN, D. AND AHUJA, N. Shape from texture: Integrating texture element extraction and surface estimation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, pages 1233–1251, December 1989.
- BOUGUET, J.-Y. AND PERONA, P. 3D Photography using shadow in dual-space geometry. In *IJCV*, November/December 1999.
- C3D/EDINBURGH VIRTUAL ENVIRONMENT CENTRE. C3D/Edinburgh Virtual Environment Centre: 3D imaging in a flash. Web site – <http://www.c3d.com/products/singlePod.htm>.
- CHAPRA, S. C. AND CANALE, R. P. *Numerical Methods for Engineers: With Software and Programming Applications*, chapter 3, pages 54–57. McGraw-Hill, fourth edition, 2002a. Part One— Modeling, Computers and Error Analysis, Section 3.3 — Error Definitions, International Edition.
- CHAPRA, S. C. AND CANALE, R. P. *Numerical Methods for Engineers: With Software and Programming Applications*, chapter 14 and 17 and and Ep. Part V, pages 375–397, 468–470, 563–568. McGraw-Hill, fourth edition, 2002b. Part Four — Nonlinear Multidimensional Unconstrained Optimization, Part Five — Nonlinear Least-Squares Regression, International Edition.
- CHEN, H. A screw motion approach to uniqueness analysis of head-eye geometry. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 145–151, Maui, Hawaii, June 3–6 1991.
- CHEN, Y. AND MEDIONI, G. Object Modelling by Registration of Multiple Range Images. *Image and Vision Computing*, 10(3):145–155, April 1992.
- CHOU, J. AND KAMEL, M. Finding the Position and Orientation of a Sensor on a Robot Manipulator Using Quaterions. *International Journal of Robotics Research*, 3 (10):240–254, 1991. As cited in Daniilidis [1998].
- CLIFFORD, W. Preliminary sketch of bi-quaternions. In *Proc. London Mathematical Society*, number 4, pages 381–395, 1873. As cited in Daniilidis [1998].
- COLLINSON, R. *Introduction to Avionics*. Chapman & Hall, 1996. ISBN 0-412-48250-9.
- CURLESS, B. AND LEVOY, M. Project to Build a 3D Fax Machine. Website – <http://graphics.stanford.edu/projects/faxing>, 1996-97. Last update: February 10, 1997.
- CYBERWARE. Cyberware Home Page. Web site – www.cyberware.com.

- DALMIA, A. K. AND TRIVEDI, M. M. 3-d structure extraction from image streams: a review. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1996a.
- DALMIA, A. K. AND TRIVEDI, M. M. Integrating depth from motion and binocular stereo approaches. *Machine Vision and Applications*, 1996b.
- DANIILIDIS, K. Hand-Eye Calibration Using Dual Quaternions. *International Journal of Robotics Research*, 1998.
- DAVIS, J. AND CHEN, X. A Laser Range Scanner Designed for Minimum Calibration Complexity. In *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*. IEEE, May 2001.
- DEPIERO, F. W. AND TRIVEDI, M. M. 3-D Computer Vision Using Structured Light: Design, Calibration, and Implementation Issues. *Advances in Computers*, 43: 243–278, 1996.
- DIAS, J. *Reconstrução Tridimensional Utilizando Visão Dinâmica*. PhD thesis, University of Coimbra, Portugal, November 1994.
- DUDEK, G. AND JENKIN, M. *Computational Principles of Mobile Robotics*, chapter 4, page 112. Cambridge University Press, 2000a.
- DUDEK, G. AND JENKIN, M. *Computational Principles of Mobile Robotics*, chapter 3, page 67. Cambridge University Press, 2000b.
- DUDEK, G. AND JENKIN, M. *Computational Principles of Mobile Robotics*, chapter 4, page 104. Cambridge University Press, 2000c.
- DUDEK, G. AND JENKIN, M. *Computational Principles of Mobile Robotics*, chapter 7, pages 181–183. Cambridge University Press, 2000d.
- EDWARD, A. *Computer Graphics*. Addison Wesley, 1990.
- EGGERT, D. W., LORUSSO, A., AND FISHER, R. B. Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, (9):272–290, 1997.
- EPRSC. Faraday Home Frameset. Web site – <http://www.esat.kuleuven.ac.be/~konijn/improofs.htm>, August 2000. EPRSC (Sponsored Faraday Partnership Administered by the Department of Computing Science), University of Glasgow.
- FARIN, G. *NURBS — from Projective Geometry to Practical Use*, chapter 1. A. K. Peters, 2nd. edition, 1999.

- FAUGERAS, O. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.
- FAUGERAS, O. D. AND HEBERT, M. The Representation, Recognition and Locating of 3-D Objects. *The International Journal of Robotics Research*, 5(3):27–52, Fall 1986. Massachusetts Institute of Technology.
- FAVARO, P. AND SOATTO, S. Learning shape from defocus. In *Proceedings of the ECCV*. LNCS, Springer Verlag, May 2002.
- FELDMAR, J. AND AYACHE, N. Rigid, Affine and Locally Affine Registration of Free-Form Surfaces. Research Report 2220, Institut National De Recherche En Informatique Et En Automatique (INRIA), March 1994. Project Epidaure.
- FERREIRA, J. F. Scanners 3D — State-Of-The-Art. Technical report, Instituto de Sistemas e Robótica, Universidade de Coimbra, 2002.
- FERREIRA, J. AND DIAS, J. A 3D Scanner — Three-Dimensional Reconstruction From Multiple Images. In *Proc. Controlo 2000 Conf. on Automatic Control*, pages 690–695. University of Minho, Portugal, 2000. Student Forum.
- FERREIRA, J., LOBO, J., AND DIAS, J. Tele-3D — Developing a Handheld Scanner Using Structured Light Projection. In CORTELAZZO, G. AND GUERRA, C., editors, *3DPVT Proceedings — 1st International Symposium on 3D Processing, Visualization and Transmission*, pages 788–791. University of Padua, IEEE Computer Society, June 19–21 2002a.
- FERREIRA, J., LOBO, J., AND DIAS, J. Tele-3D — Developing a Handheld Scanner Using Structured Light Projection. In *RecPad Proceedings — 12th Portuguese Conference On Pattern Recognition*. University of Aveiro, Portugal, June 27–28 2002b. CD Proceedings.
- FERREIRA, J., MARTINS, N., AGNELO, L., AND DIAS, J. Imagiologia Tridimensional Digital para Construção de Protótipos Industriais. *Robótica*, 4th Trimester 2001(45): 28–34, November 2001.
- FOWLER, M. AND KENDALL, S. *UML Distilled - Applying the Standard Object Modelling Language*. Addison Wesley, 2nd. edition, 1997.
- GEE, A. AND PRAGER, R. Sequential 3D diagnostic ultrasound using the Stradx system. Medical Image Computing and Computer-Assisted Intervention — MICCAI'99, September 1999. Revised and extended, June 2001.

- GOLDSTEIN, H., POOLE, C., AND SAFKO, J. *Classical Mechanics*. Addison-Wesley, 3rd edition, 2002. ISBN 0-201-65702-3.
- GONZALEZ, R. C. AND WOODS, R. E. *Digital Image Processing*. Prentice Hall, second edition, 2002. International Edition.
- GROLIER, editor. *Encyclopedia International*. Number 7. Grolier Incorporated, 1971. ISBN 0-7172-0702-1.
- HAMILTON, W. R. On quaternions, or on a new system of imaginaries in algebra. *Phil. Mag.*, pages 489–495, 1850. As cited in Daniilidis [1998].
- HARTLEY, R. AND ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- HASKELL, B. G., PURI, A., AND NETRAVALI, A. N. *Digital Video: An Introduction to MPEG-2*, chapter 15 and 17, pages 322–358, 382–387. Chapman and Hall, New York, 1997.
- HÉBERT, P. AND RIOUX, M. Toward a hand-held laser range scanner: Integrating observation-based motion compensation. In *Proceedings of IS&T/SPIE's 10th Annual Symposium on Electronic Imaging (Photonics West); Conference 3313: Three-Dimensional Image Capture and Applications*, San Jose, Ca, USA, January 1998.
- HEIKKILÄ AND SILVEN. A Four-step Camera Calibration Procedure with Implicit Image Correction. In *CVPR97 Proceedings*, 1997.
- HILTON, A. Model Building from 3D Surface Measurements. Web page – http://www.ee.surrey.ac.uk/Research/VSSP/3DVision/model_building/model.html, December 1997. Tutorial.
- HINDUS, L. A. 3D Scanners: Scanning Takes Shape. *Advanced Imaging*, 16(7): 42,43,62, July 2001.
- HOPPE, H., DEROSE, T., DUCHAMP, T., McDONALD, J., AND STUETZLE, W. Surface Reconstruction from Unorganized Points. In *SIGGRAPH '94 Proceedings*. Washington University, 1994.
- HOPPE, H. *Surface Reconstruction from Unorganized Points*. PhD thesis, Washington University, 1994.
- HORAUD, R. AND DORNAIKA, F. Hand-Eye Calibration. *Int. Journal of Robotics Research*, 14(3):195–210, June 1995.

- HORN, B. K. P., HILDEN, H. M., AND NEGAHDARIPOUR, S. Closed-form solution of absolute orientation using orthonormal matrices. *J. Optical Society of America Ser. A*, (5):1127–1135, 1988.
- HORN, B. K. P. Closed-form solution of absolute orientation using unit quaternions. *J. Optical Society of America Ser. A*, 4(4):629–642, April 1987.
- HUBER, D. *Automatic 3D Modeling of Unstructured Scenes*. PhD thesis, The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, June 2000. Thesis Proposal.
- IMPROOFS PROJECT. Web site – <http://www.esat.kuleuven.ac.be/~konijn/improofs.htm>.
- OpenCV Manual. *Open Source Computer Vision Library*. Intel Corporation, USA, 1999-2001. Reference Manual.
- KUTULAKOS, K. N. AND SEITZ, S. M. A Theory of Shape by Space Carving. In *International Journal of Computer Vision*, 2000.
- LEON, S. J. *Linear Algebra With Applications*. Prentice Hall, 5th edition, 1998.
- LEVOY, M. Why Is 3D Scanning Hard? In CORTELAZZO, G. AND GUERRA, C., editors, *3DPVT Proceedings — 1st International Symposium on 3D Processing, Visualization and Transmission*. University of Padua, IEEE Computer Society, 2002. Key-note — only slides are publicly available at <http://graphics.stanford.edu/talks>.
- LEVOY, M. ET AL. The Digital Michelangelo Project: 3D scanning of large statues. In *SIGGRAPH 2000 Proceedings*. New Orleans, LA, July 2000.
- LOBO, J. Inertial Sensor Data Integration In Computer Vision Systems. Master's thesis, Faculty of Science and Technology, Electrical and Computer Engineering Department, University of Coimbra, 2002.
- MAGRAB, E. B. ET AL. *An Engineer's Guide to MATLAB*, chapter 5, 6, 7 and 13, pages 101–238, 571–618. Prentice-Hall, New Jersey, 2000. ISBN 0-13-0011335-2.
- MARTINS, N. AND DIAS, J. Visual Inspection based on mirror images - The 3D Reconstruction Case. In *SIRS '99 Proceedings*. University of Coimbra, Portugal, 1999.
- MCCALLUM, B. C., FRIGHT, W. R., NIXON, M. A., AND PRICE, N. B. A Feasibility Study of Hand-held Laser Surface Scanning. In *Proc. Image and Vision Computing New Zealand*, pages 103–108, Lower Hutt, August 1996.

- McIVOR, A. M. Calibration of a Laser Stripe Profiler. In *Second International Conference on 3-D Digital Imaging and Modeling*, Los Alamitos, CA, USA, 1999. IEEE Comput. Soc.
- MEISTER, L. *Quaternions and Their Application In Photogrammetry and Navigation*. PhD thesis, Fakultät für Mathematik und Informatik der TU Bergakademie Freiberg, December 1997.
- MINOLTA CORPORATION. Minolta 3D Home Page. Web site – <http://www.minolta3d.com>, 1997.
- M. WOO, J. N. AND DAVIS, T. *Open GL Programming Guide*. Addison Wesley, 1997.
- NAKAMURA, Y. AND XU, Y. Geometrical Fusion Method for Multi-Sensor Robotic Systems. In *Proc. 1989 IEEE International Conference on Robotics and Automation*, pages 668–673, Scottsdale, Arizona, 1989.
- NINA AMENTA, S. C. AND KOLLURI, R. K. The power crust, unions of balls, and the medial axis transform. In *SIGGRAPH2001*, 2001.
- OPTONET. Optonet 3D Home Page. Web site – <http://www.optonet.it>.
- PÉREZ-FAJADO, Y. AND GUTSENS, O. As imagens em 3D ajudam os cirurgiões - Manter o que está bem. *Super Interessante*, (29):84–90, September 2000. ISSN 0874-0682.
- POLHEMUS/ARANZ. Polhemus Home Page. Web site – <http://www.polhemus.com>, 2000.
- PORTO EDITORA. Diciopédia 2002. DVD-ROM by Porto Editora Ltd, 2002.
- PRAGER, R., ROHLING, R., GEE, A., AND BERMAN, L. Automatic calibration for 3-D free-hand ultrasound. Technical Report CUED/F-INFENG/TR 303, Cambridge University Department of Engineering, September 1997.
- RAO, S. S. *Applied Numerical Methods for Engineers and Scientists*, chapter 12, pages 899–902. Prentice Hall, 2002a. Sections 12.4.3/12.4.4 — Function of Several Variables with Equality Constraints/Mixed Equality and Inequality Constraints.
- RAO, S. S. *Applied Numerical Methods for Engineers and Scientists*, chapter 4, pages 270–363. Prentice Hall, 2002b. Solution of Matrix Eigenvalue Problem.
- RUSINKIEWICZ, S. AND LEVOY, M. Efficient variants of the ICP algorithm. In *Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling*, pages 145–152, 2001. URL citeseer.nj.nec.com/rusinkiewicz01efficient.html.

- SCHROEDER, W., MARTIN, K., AND LORENSSEN, B. *The Visualization Toolkit — An Object-Oriented Approach To 3D Graphics*. Kitware Inc., third edition, 2004. CD-ROM (VTK version 4.2), ISBN 1-930934-07-6.
- SCOTT, S. AND CAO, L. Topic clustering basic, 1999. URL citeseer.ist.psu.edu/438032.html. CSCE 970: Pattern Recognition Spring Semester.
- SHAPIRO, L. G. AND STOCKMAN, G. C. *Computer Vision*. Prentice-Hall, 2001. International Edition, ISBN 0-13-030796-3.
- SHIU, Y. C. AND AHMAD, S. Calibration of Wrist-Mounted Robotic Sensors by Solving Homogeneous Transform Equations of the Form $\mathbf{AX} = \mathbf{XB}$. *IEEE Transactions on Robotics and Automation*, 5(1):16–27, February 1989.
- STUDY, E. Von den Bewegungen und Umlegungen. *Mathematische Annalen*, 39: 441–566, 1891. As cited in Daniilidis [1998].
- STUDY, E. Geometrie der Dynamen, 1903. Leipzig. As cited in Daniilidis [1998].
- TANIMOTO, T. T. An elementary mathematical. Technical report, IBM, 1959. As cited by Scott and Cao [1999].
- TRUCCO, E., FISHER, R., AND FITZGIBBON, A. Direct calibration and data consistency in 3-D laser scanning. In HANCOCK, E., editor, *BMVC94: Proceedings of the 5th British Machine Vision Conference*, pages 489–498. BMVA Press, September 1994.
- VITÓRIA, J. AND DE LIMA, T. P. *Álgebra Linear*. Universidade Aberta, Lisboa, February 1998.
- WALKER, M. W., SHAO, L., AND VOLZ, R. A. Estimating 3-D Location Parameters Using Dual Number Quaternions. *CVGIP: Image Understanding*, 54(3):358–367, 1991.
- WEISS, L. SFF Processes. Technical report, JTEC/WTEC Panel Report on Rapid Prototyping in Europe and Japan, March 1997.
- WOODHAM, R. J. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, pages 139–144, January/February 1980.
- ZHANG, Z. Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting. Research Report 2676, Institut National De Recherche En Informatique Et En Automatique (INRIA), October 1995. ISSN 0249-6399.

- ZHANG, Z. Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In *International Conference on Computer Vision (ICCV'99)*, pages 666–673, Corfu, Greece, September 1999.
- ZHANG, Z. A Flexible New Technique For Camera Calibration. Technical report, Microsoft Research, 2000.
- UML Notation Guide - version 1.1. Technical Report 1.1, Rational Software Corporation, September 1997.