

Mercedes Rocío Gonzales Márquez

Pré-alinhamento de Imagens de Profundidade via Malhas Simplificadas

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Engenharia de Computação.

Orientadora: Profa. Dr.-Ing. Wu, Shin-Ting

Campinas, SP
2009

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

G589p	Gonzales Márquez, Mercedes Rocío Pré-alinhamento de imagens de profundidade via malhas simplificadas / Mercedes Rocío Gonzales Márquez. – Campinas, SP: [s.n.], 2009. Orientador: Wu Shin-Ting. Tese de Doutorado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação. 1. Imagem tridimensional. 2. Processamento de imagens. 3. Visão por computador. 4. Malhas. 5. Modelagem de dados. I. Wu Shin-Ting. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título
-------	---

Título em Inglês:	Crude registration of range images through simplified meshes
Palavras-chave em Inglês:	Three-dimensional imaging, Image processing, Vision computer, Meshes, Modeling data
Área de concentração:	Engenharia de Computação
Titulação:	Doutor em Engenharia Elétrica
Banca Examinadora:	Olga Regina Pereira Bellon, Antônio Castelo Filho, Clésio Luis Tozzi, Alice Maria Bastos Hubinger Tokarnia
Data da defesa:	22/01/2009
Programa de Pós Graduação:	Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidata: Mercedes Rocío Gonzales Márquez

Data da Defesa: 22 de janeiro de 2009

Título da Tese: "Pré-alinhamento de Imagens de Profundidade via Malhas Simplificadas"

Profa. Dra. Wu Shin-Ting (Presidente):

Wu Shin-Ting

Profa. Dra. Olga Regina Pereira Bellon:

Olga Regina Pereira Bellon

Prof. Dr. Antônio Castelo Filho:

Antônio Castelo Filho

Profa. Dra. Alice Maria Bastos Hubinger Tokarnia:

Alice Maria Bastos Hubinger Tokarnia

Prof. Dr. Clésio Luis Tozzi:

Clésio Luis Tozzi

Resumo

O pré-alinhamento de duas imagens de profundidade parcialmente sobrepostas de objetos rígidos consiste na determinação automática de uma transformação rígida T que coloca ambas imagens em um mesmo referencial. Ele deve ser independente da geometria da imagem e deve ser eficiente, no sentido de reduzir a complexidade temporal da tarefa de determinação e casamento de pares de correspondências. A solução mais popular consiste na extração de n pares de pontos correspondentes nas regiões comuns de ambas imagens para, a partir desses pares, estimar a transformação T . As técnicas existentes na literatura ainda apresentam diversas limitações quanto à restrição geométrica das imagens e à eficiência. Este trabalho propõe uma solução alternativa eficiente. Tendo como hipótese de que a forma de um objeto pode ser essencialmente capturada por um subconjunto finito dos seus pontos, esta tese propõe o uso de uma malha triangular simplificada, em lugar de uma imagem de profundidade densa, para reduzir o espaço de busca de correspondências. Propõe-se ainda a construção de um descritor, denominado triedro, com alta capacidade discriminante para possibilitar uma busca mais eficiente de correspondências no espaço reduzido de amostras. O algoritmo proposto foi validado através de testes com diversas imagens reais.

Palavras-chave: Descritor Geométrico, Correspondência, Pré-alinhamento, Reconstrução 3D.

Abstract

The crude registration of two partially overlapping range images from rigid objects consists of the determination of a rigid transformation T which brings both images into the same reference system. A crude registration should be independent of data geometry, and be efficient, in the sense that has low time complexity in finding the correspondence pairs. The most popular solution consists of extracting n pairs of correspondences in the overlapping regions and estimating from them the transformation T . Current works on crude registration still present drawbacks in relation to data geometry and efficiency. This work proposes an alternative solution for crude registration. Based on the hypothesis that the shape of an object is essentially captured by a finite subset of its points, this thesis proposes the use of a simplified triangular mesh, instead of a dense range image, to reduce the search space of correspondences. Besides, the construction of a descriptor, called trihedron is proposed. It has higher discriminating capacity which makes correspondence search still more efficient. The proposed algorithm has been validated through experiment tests with several real images.

Keywords: Geometrical Descriptor, Correspondence, Crude Registration, 3D Reconstruction.

Agradecimentos

Ao Deus maravilhoso, por ser constantemente minha luz, minha força e meu refúgio. Luz, que ilumina meu caminho; força, que me impele para frente e me renova a cada passo, e refúgio, quando preciso de calma e segurança. Obrigada Senhor!.

À minha amada família: Ao meu esposo Cosme, por seu amor, companhia e encorajamento durante o desenvolvimento deste trabalho, e à minha filha Paulinha, por alegrar minha vida com sua doce presença.

Aos meus pais Marco e Maria, por seu exemplo de honestidade, perseverança e fé inabaláveis. Também à minha irmã Milagro e ao meu cunhado Alex, pela casa sempre aberta para me acolher durante minhas estadas em Campinas, depois que mudei para Dourados/MS. Obrigada a vocês quatro por seu amor, incentivo e orações.

À minha orientadora Wu Shin-Ting, pelo ensino e valiosa orientação.

Aos membros da banca examinadora desta tese, professores Olga R.P.Bellon, Antônio Castelo Filho, Clésio Luis Tozzi e Alice M.B.H.Tokarnia, pelas valiosas sugestões e contribuições para o aprimoramento deste trabalho.

À minha igreja Nova Terra em Dourados representada pelos Pastores Marcos, Raquel, Gilmar e Rose. Agradeço o carinho, encorajamento e o precioso suporte espiritual durante a reta final desta jornada. Vocês são muito especiais para mim.

Aos professores do DCA-FEEC, pelos conhecimentos compartilhados.

Aos meus colegas de trabalho da Universidade Estadual de Mato Grosso do Sul, pelo apoio e amizade.

À minha amiga Regina, pela correção do português deste texto.

À CAPES e FAPESP, pelo apoio financeiro.

A todos aqueles que, de alguma forma, tornaram possível a realização deste trabalho.

*Dedico esta tese ao Deus Todo Poderoso.
“Porque dEle e por Ele, e para Ele, são todas as coisas; glória, pois, a Ele
eternamente. Amém”.
Romanos 11:36.*

Sumário

1	Introdução	1
1.1	Imagens de profundidade	2
1.2	Reconstrução 3D em Visão Computacional	4
1.3	O problema do pré-alinhamento	8
1.4	Conteúdo do Trabalho	10
1.4.1	Escopo	11
1.4.2	Hipótese	11
1.4.3	Contribuições	11
1.5	Organização do trabalho	12
2	Trabalhos Correlatos	15
2.1	Propriedades Geométricas dos Pontos Amostrais de uma Superfície	16
2.2	Uso do conjunto completo de dados	19
2.2.1	DARCES baseado em RANSAC (DBR)	20
2.2.2	Casamento de Pontos Orientados (Imagens <i>Spin</i> - CPO)	21
2.2.3	Casamento de Histogramas Geométricos (CHG)	23
2.2.4	Limitações	24
2.3	Uso de um conjunto de dados com características especiais	25
2.3.1	Casamento de Curvas Bitangentes (CCB)	25
2.3.2	Casamento de Curvas Intrínsecas (CCI)	27
2.3.3	Casamento baseado no gradiente de intensidade (CGI)	29
2.3.4	Limitações	30
2.4	Considerações Finais	31
3	Algoritmo de Busca	35
3.1	Construção de um Descritor-Referência	36
3.2	Casamento de triedros	39
3.2.1	Procedimento de Escolha de um Descritor-Referência	39

3.2.2	Procedimento de Busca por um Descritor-Alvo	39
3.3	Remoção de Correspondências Falsas	41
3.4	Análise	44
3.4.1	Existência de um par de triedros correspondentes	44
3.4.2	Análise da resolução da malha	45
3.5	Considerações Finais	46
4	Redução do Espaço de Busca	49
4.1	Redução por Decimação da Malha	50
4.2	QSLIM: Um Algoritmo de Decimação	51
4.3	Adaptações	54
4.3.1	Controle na Densidade Amostral	54
4.3.2	Controle no Limiar para Espaço de Busca	61
4.4	Pré-alinhamento com uso de malhas simplificadas	66
4.5	Análise	67
4.5.1	Existência de um par de triedros correspondentes	67
4.5.2	Redução de Potenciais Candidatos	69
4.6	Considerações Finais	71
5	Validação	75
5.1	Algoritmo	75
5.2	Ajuste Fino (aplicação de ICP)	76
5.3	Resultados	76
5.3.1	Qualidade do Pré-alinhamento	78
5.3.2	Registro de n vistas para reconstrução completa	79
5.3.3	Comparações	93
5.4	Considerações Finais	96
6	Conclusões e Trabalhos Futuros	97
	Referências bibliográficas	100
A	Outros Resultados	107
B	Pseudo-código	121

Lista de Figuras

1.1	Aquisição de uma imagem de profundidade (Fonte: [3])	4
1.2	Múltiplas vistas de um objeto 3D (Fonte: [50])	5
1.3	(a)-(b) Duas vistas do modelo <i>bunny</i> com pares de correspondências conhecidos e (c) o resultado do registro das duas vistas.	6
1.4	(a) Imagem angel; (b) Malha densa; (c) Malha simplificada.	12
2.1	Curvatura de uma superfície	17
2.2	Classificação dos pontos em (a) parabólico, (b) hiperbólico e (c) elíptico	18
2.3	Vetores normais e arestas da face	19
2.4	(a-c) Busca de um triângulo correspondente (d) Triângulo grande (15 pontos)	21
2.5	Imagem <i>spin</i>	22
2.6	Histograma geométrico	23
2.7	Curvas de máxima curvatura	26
2.8	(a) Pontos bitangentes (b) Plano bitangente rolando sobre a superfície	26
2.9	Distância entre um par de pontos bitangentes	27
2.10	Curva interpolante	28
2.11	(a) Curvas Intrínsecas (b) Representações de <i>features</i>	28
3.1	<i>Shape context</i>	36
3.2	Triedro Espacial.	37
3.3	Triedro Espacial.	38
3.4	(a) $f_1 = 0.1, f_2 = 0.2$ (b) $f_1 = 0.5, f_2 = 0.7$ (c) $f_1 = 0.3, f_2 = 0.4$	40
3.5	Filtragem por curvatura média	42
3.6	Obtenção do quarto ponto correspondente.	43
3.7	Nível de Redução vs. Tempo de Processamento.	46
4.1	(a) Superfície do objeto (b) Pontos distintos	51
4.2	Rosto de Chopin	52
4.3	Imagem Esfcub	52

4.4	Contração de aresta	52
4.5	Face da malha e pixels ocupados na imagem	55
4.6	Imagens de Teste: PITT-PLANE [51]	56
4.7	Imagens de Teste: PITBULL [51]	56
4.8	Malhas Simplificadas com nível $\epsilon = 0.2 * s$ (a) PITT-PLANE, (b) PITBULL	57
4.9	Classificação de vértices Imagem PITT-PLANE	59
4.10	Classificação de vértices Imagem PITBULL	60
4.11	Imagens de Teste: FROG [35]	61
4.12	Malhas Simplificadas com nível ϵ (a) PITT-PLANE, (b) PITBULL	62
4.13	Malhas Simplificadas com nível ϵ (a) FROG, (b) BUNNY	63
4.14	Pares de malhas simplificadas (PITT-PLANE e PITBULL) no mesmo referencial	64
4.15	Pares de malhas simplificadas (FROG e BUNNY) no mesmo referencial	65
5.1	Pré-alinhamento de um par de imagens de profundidade.	77
5.2	Pré-alinhamento e Registro Fino (PITT-PLANE).	80
5.3	Pré-alinhamento e Registro Fino (PITBULL).	81
5.4	Pré-alinhamento e Registro Fino (FROG).	82
5.5	Pré-alinhamento e Registro Fino (BUNNY).	83
5.6	Registro de pares de vistas: BUNNY.	85
5.7	Propagação de erros no registro de múltiplas imagens.	86
5.8	Registro global de vistas: BUNNY.	86
5.9	Registro de pares de vistas: PITT-PLANE.	87
5.10	Registro global de vistas: PITT-PLANE.	88
5.11	Registro de pares de vistas: PITBULL.	89
5.12	Registro global de vistas: PITBULL.	90
5.13	Registro de pares de vistas: FROG.	91
5.14	Registro global de vistas: FROG.	92
5.15	(a) Imagem $\mathcal{R}_1$ (b) Imagem $\mathcal{R}_2$ (c) Pré-alinhamento.	95
A.1	Imagens de Teste: TELETUBBY [35] e BIRD [35].	108
A.2	Imagens de Teste: DINO [51]	109
A.3	Pré-alinhamento e Registro Fino (TELETUBBY).	111
A.4	Pré-alinhamento e Registro Fino (BIRD).	112
A.5	Pré-alinhamento e Registro Fino (DINO).	113
A.6	Registro de pares de vistas: TELETUBBY.	114
A.7	Registro global de vistas: TELETUBBY.	115
A.8	Registro de pares de vistas: BIRD.	116

A.9 Registro global de vistas: BIRD.	117
A.10 Registro de pares de vistas: DINO.	118
A.11 Registro global de vistas: DINO.	119

Lista de Tabelas

2.1	Tabela comparativa dos principais trabalhos presentes na literatura	33
3.1	Parâmetros da transformação entre pares de imagens do modelo BUNNY	45
3.2	Número de Vértices vs. Tempo de Processamento	46
4.1	Redução do volume de dados PITT-PLANE e PITBULL.	58
4.2	Classificação de vértices de PITT-PLANE e PITBULL.	59
4.3	Redução do volume de dados PITT-PLANE e PITBULL.	59
4.4	Redução do volume de dados.	66
4.5	Porcentagem de Vértices Correspondentes/Vértices em A_{S1}	67
4.6	Parâmetros da transformação entre pares de imagens do modelo BUNNY	69
4.7	Redução de falsas correspondências.	70
4.8	Tabela comparativa dos principais trabalhos presentes na literatura	73
5.1	Parâmetros da transformação entre pares de imagens do modelo BUNNY	79
5.2	Percentual da Região de sobreposição	84
5.3	Quadro Comparativo de Algoritmos (NT = Não trabalha bem).	94
A.1	Redução do volume de dados.	107

Trabalhos Publicados Pela Autora

1. S.T. Wu, Mercedes R.G.Márquez. "A non-self-intersection Douglas-Peucker Algorithm". *Proceedings of the 16th Brazilian Symposium on Computer Graphics and Image Processing* (São Carlos, SP, Brasil, Outubro de 2003), IEEE CS Press, pg. 60-66.

A idéia inicial para se desenvolver um novo algoritmo de registro foi criar um algoritmo de simplificação próprio que pudesse ser integrado no processo de registro. Ao analisar o algoritmo Douglas-Peucker de simplificação de polilinhas conjecturou-se que se podia estendê-lo para 3D após solucionar o problema de auto-interseções. Concebeu-se, então, uma solução para este problema, que teve como fruto a elaboração deste artigo. Este artigo foi selecionado para publicação de uma versão estendida na revista da Sociedade Brasileira de Computação.

2. S.T. Wu, A.C.G. da Silva, Mercedes R.G.Márquez. "The Douglas-Peucker Algorithm: Sufficient Conditions for Non-Self-Intersecting". *Journal of the Brazilian Computer Society*, Volume 9, Numero 3, pp.67-79, Setembro 2004.

É uma versão estendida do artigo anterior. Neste trabalho foi adicionada uma formalização dos resultados obtidos. Apesar da melhoria no desempenho do algoritmo, através de uma nova implementação feita pelo engenheiro Adler Cardoso da Silva, constatou-se que a extensão deste algoritmo para 3-D não seria trivial. Optou-se, portanto, usar o algoritmo de simplificação QSLIM e tentar adaptá-lo para satisfazer os requisitos de pré-alinhamento proposto neste trabalho.

3. Mercedes R.G.Márquez, S.T. Wu "Using simplified meshes for crude registration of two partially overlapping range images". *Short Communications Proceedings of the 15th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2008 - WSCG'2007, Plzen, REPUBLICA TCHECA, 2007*, Volume 15, Numero 2, pp.183-190, Janeiro 2007.

Este trabalho apresenta uma versão preliminar do algoritmo de pré-alinhamento com uso de malhas simplificadas. Ela sintetiza as idéias apresentadas no Capítulo 4, relacionadas com a redução do espaço de busca.

4. Mercedes R.G.Márquez, S.T. Wu. "An automatic crude registration of two partially overlapping range images". *Proceedings of the 21th Brazilian Symposium on Computer Graphics and Image Processing* (Campo Grande, MS, Brasil, Outubro de 2008), IEEE CS Press, pp.245-252.

Neste trabalho o foco é a construção de um descritor a partir de uma malha simplificada. O objetivo é aumentar a eficiência na busca de correspondências, necessárias para a estimativa de um pré-alinhamento. Capítulo 3 é uma versão estendida deste artigo.

Capítulo 1

Introdução

Reconstrução 3D de objetos do mundo real destaca-se como uma área muito importante na Visão Computacional e consiste em modelar objetos a partir de imagens. No contexto deste trabalho estas imagens são imagens de profundidade para as quais o processo de reconstrução compreende três fases principais: aquisição de imagens de profundidade parcialmente sobrepostas, registro e integração destas imagens. Na aquisição, a superfície do objeto é escaneada a partir de vários pontos de vista, gerando um conjunto de n imagens de profundidade em forma de nuvem de amostras. Na fase de registro, uma transformação rígida (matriz de rotação e translação) deve ser estimada para alinhar as imagens, par a par, em um único referencial. Finalmente, uma vez que todas as vistas estejam registradas em um único sistema de coordenadas, elas são integradas para construir um modelo 3D completo. A complexidade dessa tarefa é uma função do número n de vistas adquiridas para amostrar a geometria total do objeto, portanto procura-se utilizar um mínimo número de imagens para aliviar os custos de memória e processamento da tarefa de integração [33, 46]. Esta tese foca o problema de registro de imagens parcialmente sobrepostas.

O paradigma clássico para estimar a transformação rígida que registra duas imagens de profundidade é a minimização do erro da distância entre pares de pontos correspondentes extraídos de ambas imagens. Portanto, a tarefa principal no ato de registro é a determinação automática de pares de correspondências entre vistas. A solução deste problema torna-se um desafio pois a correspondência entre as regiões de sobreposição, das quais precisa-se extrair os pares de pontos correspondentes, é desconhecida. Besl apresentou o algoritmo iterativo ICP [5], propondo que os pontos mais próximos entre as duas imagens sejam os pontos correspondentes. Para encontrar tais correspondências, o algoritmo necessita de uma transformação inicial T_0 que aproxime suficientemente bem ambas imagens. Assim sendo, surge a necessidade de uma aproximação inicial do registro o que é conhecido como pré-alinhamento de imagens de profundidade.

A literatura apresenta algumas técnicas para resolver esta questão, as quais se diferenciam basicamente pela procedência dos pares de correspondências. Há duas abordagens: na primeira, as

correspondências são extraídas a partir do conjunto completo de amostras originais e, na segunda, as correspondências são obtidas a partir de um subconjunto de dados originais composto por amostras que possuem características especiais. Estas técnicas, no entanto, ainda apresentam certas limitações. No caso da primeira, o espaço de busca gerado é muito grande o que torna a técnica ineficiente; e no caso da segunda, alguma informação importante para as correspondências pode ser descartada, ao restringir-se somente a um conjunto específico de dados, como pontos bitangentes [53]. Esta restrição faz com que estas técnicas sejam dependentes da geometria e não asseguram a extração de pares de correspondências para qualquer forma da superfície.

Neste trabalho propõe-se um algoritmo de registro que procura contornar estas limitações com base em duas inovações. Propõe-se, em lugar de usar o conjunto completo de amostras de uma imagem de profundidade, o uso de uma malha triangular simplificada que preserve as características globais da forma original da imagem e que os pares de correspondências sejam extraídos a partir dos espaços reduzidos, compostos pelos vértices da malha. Além disso, propõe-se um novo descritor, denominado “triedro”, para buscar as correspondências entre os espaços reduzidos das amostras. Diferentemente das propostas existentes, este descritor integra informações geométricas locais e globais das amostras, o que o leva a ser mais discriminante na busca.

Este capítulo está organizado da seguinte maneira. A Seção 1.1 apresenta o conceito de imagens de profundidade, sensor de profundidade e formato de dados. A Seção 1.2 contextualiza o problema do registro de duas imagens de profundidade no domínio da Reconstrução 3D e serve de embasamento ao trabalho desenvolvido. A Seção 1.3 apresenta uma introdução ao problema de registro de duas imagens de profundidade, expondo sucintamente seus principais desafios. Mais adiante, a Seção 1.4 discute brevemente o conteúdo deste trabalho, a sua hipótese e principais contribuições. Finalmente, o capítulo é concluído na Seção 1.5 com a apresentação da organização da tese.

1.1 Imagens de profundidade

Reconstrução 3D enquadra-se na área de Análise de Imagens ou Visão Computacional. Esta área tem como finalidade obter informações geométricas, topológicas ou físicas de uma cena a partir de uma ou mais imagens desta cena [42, 23]. Em particular, a obtenção da informação de profundidade (terceira dimensão) foi tradicionalmente dependente de imagens (bidimensionais) de intensidade e das técnicas conhecidas por *shape from stereo* [34], *motion* [10], *texture* [30], *shading* [19] ou combinação destas técnicas [14]. Este processo é complexo já que a extração de informações 3D depende das projeções, que contêm essencialmente dados 2D.

A aparição de novos sensores que possibilitam a obtenção direta de informação tridimensional tem atraído a atenção para a nova representação de dados tridimensionais em forma de imagem, adquirida por estes dispositivos. Uma vantagem da imagem 3D é que suas amostras podem ser

trabalhadas como um conjunto de pontos no espaço tridimensional. Portanto, além de poder empregar as técnicas tradicionais do processamento de imagens (2D), pode-se também fazer uso dos conceitos e ferramentas matemáticas existentes na teoria de superfícies, como os conceitos clássicos de geometria diferencial e topologia [16].

Uma imagem digital pode ser representada por uma matriz bidimensional, onde cada elemento representa a característica de um ponto da imagem. Estes elementos são conhecidos como *pixels*. O número de linhas e o número de colunas definem a sua resolução. Um modelo matemático natural para descrever uma imagem digital é uma função $i = f(x, y)$, conhecida como função-imagem definida em uma região bidimensional. Dependendo da natureza da imagem, esta função representa a intensidade ou a profundidade de cada pixel.

- Uma imagem de intensidade (imagem colorida) é uma matriz $m \times n$, onde o valor da função em cada *pixel* representa a cor do *pixel* que, por sua vez, é normalmente representada por três valores reais, R, G e B (vermelho (red), verde (green) e azul (blue) em sistemas de cores aditivo).
- Uma imagem de profundidade [46, 4] é representada por uma matriz $m \times n$, onde $f(x, y)$ é um valor real que representa a profundidade da amostra localizada em coordenadas (x, y) . Essa profundidade é calculada com base na distância entre o sensor tridimensional e o objeto amostrado. Assim, os pontos $(x, y, f(x, y))$ encontram-se posicionados no sistema de referência determinado pelo sensor que fez a aquisição da imagem.

Hoje em dia existem vários dispositivos de aquisição de imagens de profundidade como, por exemplo, escaneadores que usam métodos como a triangulação e métodos baseados no tempo de reflexão, através de raios laser. Na abordagem de triangulação, por exemplo, a configuração utilizada pelo sensor consiste em uma fonte pontual emissora de energia e em um receptor; onde este tem a localização conhecida em relação ao emissor, como é mostrado na Figura 1.1(a). O emissor projeta um plano de luz sobre o objeto, fazendo com que sua imagem seja capturada à medida em que a fonte de luz faz a varredura de toda a superfície do objeto. O receptor captura a reflexão do feixe e, então, calcula a informação de profundidade [3]. Conectando as amostras vizinhas da imagem de profundidade (Figura 1.1(b)) através de elementos triangulares, pode-se construir uma malha triangular como é mostrado na Figura 1.1(c).

Mais precisamente, a partir de quatro amostras vizinhas, localizadas em linhas e colunas adjacentes, pode-se criar uma faceta triangular, a partir da qual possa-se obter dois triângulos da malha. Há duas possibilidades para escolher os triângulos. O critério de escolha é baseado na avaliação dos comprimentos das duas diagonais. A configuração que corresponde à menor diagonal é escolhida. Para evitar arestas excessivamente grandes, um limiar máximo para o comprimento das arestas é pré-definido. Uma sugestão de limiar é usar-se um múltiplo da distância de amostragem s , a saber,

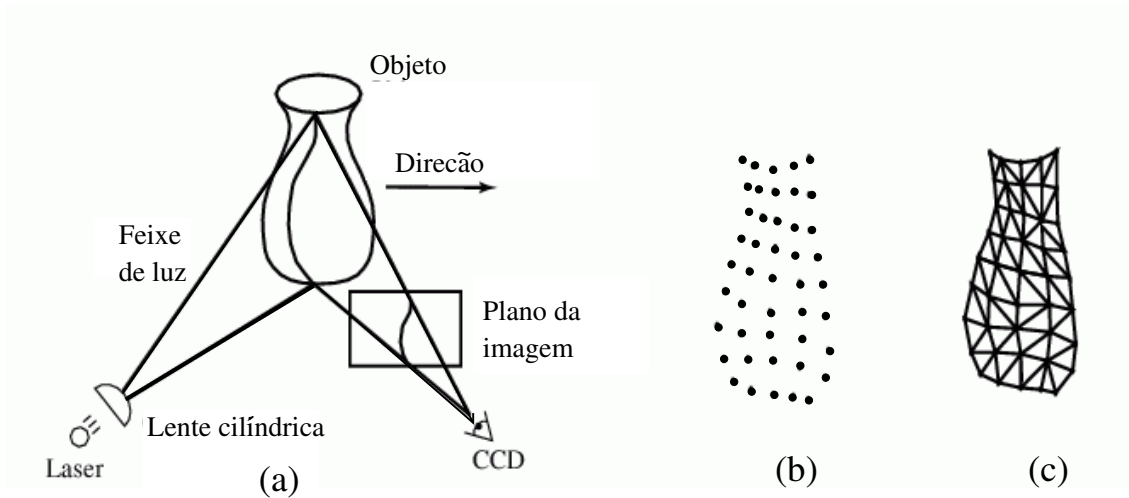


Fig. 1.1: Aquisição de uma imagem de profundidade (Fonte: [3])

$6 * s$ [52]. Embora esse limiar possa impedir a união de amostras que deveriam, de fato, estar conectadas, espera-se que outras imagens de profundidade (aquelas com ângulo de vista mais favorável), possam prover a informação correta de adjacência.

1.2 Reconstrução 3D em Visão Computacional

A facilidade e eficiência com que um sensor 3D captura informação tridimensional tem tornado imagens de profundidade populares na área de Reconstrução 3D em Visão Computacional, com numerosas aplicações que vão desde entretenimento à automação industrial. O processo de Reconstrução 3D envolve as etapas de aquisição, registro e integração [52, 43, 53, 26], que são descritas brevemente a seguir:

1. *Aquisição*: uma única imagem de profundidade não é suficiente para modelar completamente um objeto de forma livre (*free-form*) em razão às auto-oclusões. Assim sendo, múltiplas imagens devem ser adquiridas para cobrir inteiramente a superfície do objeto. Imagens capturadas podem ter em comum uma região geométrica chamada de “região de sobreposição”. Segundo [33, 46], é desejável que o número de vistas necessárias para capturar a geometria completa seja mínimo para aliviar a complexidade da sua integração. No entanto trabalhar com um número pequeno de imagens implica em se ter pares de imagens com pouca região de sobreposição e isso pode ser um empecilho para uma reconstrução confiável com a tecnologia disponível. A Figura 1.2 ilustra as imagens de profundidade do modelo *bunny* em seis distintos pontos de vista. Observe a diferença entre o tamanho da região de sobreposição do par BUNNY0-BUNNY45 e do par BUNNY0-BUNNY90.



Fig. 1.2: Múltiplas vistas de um objeto 3D (Fonte: [50])

2. *Registro*: já que cada imagem se encontra no seu próprio sistema local de coordenadas, para representar um único modelo 3D a partir delas, é preciso que sejam colocadas em um sistema de referência comum. Nesta fase distinguem-se ainda duas etapas: pré-alinhamento e registro fino. Na primeira etapa uma aproximação inicial do registro é estimada e na segunda etapa, iterações são aplicadas sobre ela para que os erros de alinhamento sejam minimizados.

- *Pré-alinhamento*: Esta etapa consiste de duas fases: Correspondência e Posicionamento. Na fase de correspondência procura-se por amostras nas duas imagens, que ocupem a mesma posição no espaço do objeto. Essas amostras são conhecidas como pontos correspondentes. As Figura 1.3(a) e (b) apresentam duas vistas, $vista_1$ e $vista_2$, do modelo *bunny* em diferentes cores. Observe que três pares de correspondências são marcados nas vistas. Na fase de Posicionamento uma transformação rígida T , que alinhe ambas as imagens, sob certo erro, é obtida a partir das correspondências estimadas. T consiste de

uma matriz de rotação R e um vetor-translação t , obtidos tal que minimizem a distância quadrada entre as correspondências (p_i, q_i) , onde p_i e q_i são os pontos correspondentes da $vista_1$ e $vista_2$, respectivamente (Equação 1.1). Um conjunto mínimo de três pontos correspondentes não colineares entre as duas vistas (incluindo seu baricentro) é suficiente para determinar o R e o t , embora muitas vezes um número maior de $N > 3$ pontos seja usado [25].

$$f = \frac{1}{N} \sum_{i=1}^N \|Rp_i + t - q_i\|^2. \quad (1.1)$$

Na Figura 1.3(c) visualiza-se o pré-alinhamento destas vistas a partir dos 3 pares de pontos correspondentes. Note-se a contribuição de cada vista em diferentes cores na vista registrada. A qualidade do registro pode ser avaliada pelo “intercalamento” de cores atribuídas às duas vistas [15, 27].

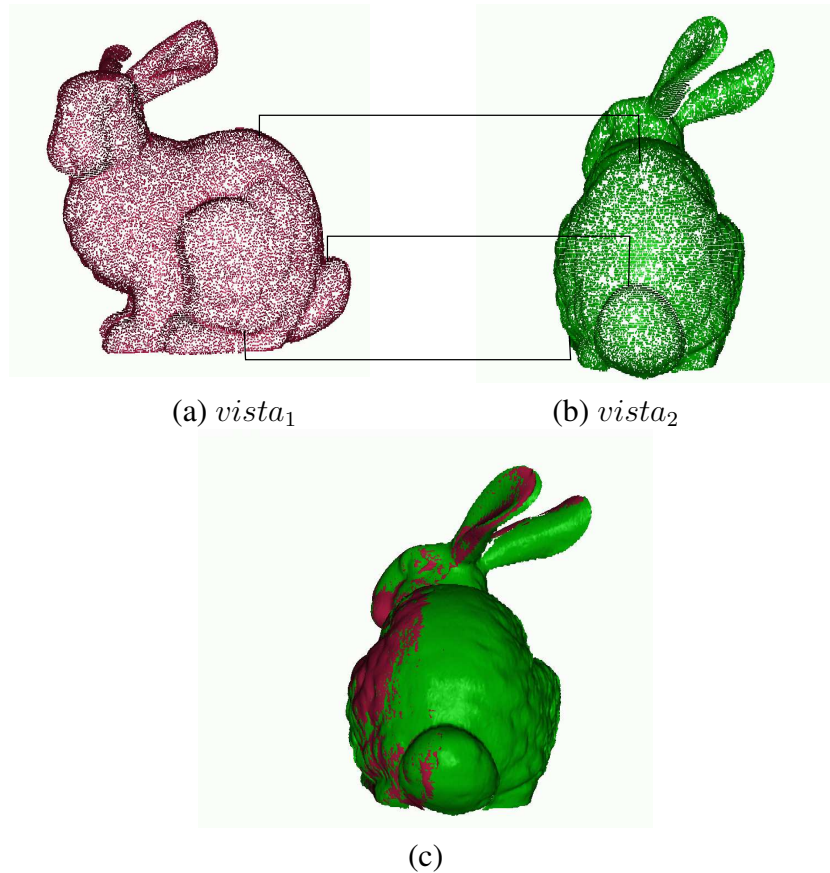


Fig. 1.3: (a)-(b) Duas vistas do modelo *bunny* com pares de correspondências conhecidos e (c) o resultado do registro das duas vistas.

O pré-alinhamento pode ser manual ou automático dependendo da escolha manual ou automática dos pontos correspondentes. O foco deste trabalho é o pré-alinhamento auto-

mático de duas imagens de profundidade.

Nesta altura da discussão, é conveniente introduzir algumas notações. Denominam-se de $\mathcal{R}_1$ e $\mathcal{R}_2$ à primeira e segunda imagem de profundidade, respectivamente. $\mathcal{R}_1$ deverá ser transformada às coordenadas da imagem $\mathcal{R}_2$, após o pré-alinhamento. Denota-se o conjunto de amostras de $\mathcal{R}_1$ como $A_1 = A_{I1} \cup A_{S1}$, e o conjunto de amostras de $\mathcal{R}_2$ como $A_2 = A_{I2} \cup A_{S2}$, onde os subconjuntos de n_0 pontos pertencentes à região de sobreposição de A_1 e A_2 são denotados por A_{S1} e A_{S2} [11], isto é, $A_{S1} = \{p_i = (x_i, y_i, z_i), i = 1, 2, \dots, n_0\}$ e $A_{S2} = \{Rp_i + t, i = 1, 2, \dots, n_0\}$.

Denomina-se de descritor-referência $\mathcal{D}_1$ o conjunto de N pontos p_i em $\mathcal{R}_1$, e descritor-alvo $\mathcal{D}_2$ o conjunto de N pontos q_i em $\mathcal{R}_2$, onde (p_i, q_i) são pares de pontos correspondentes estabelecidos para determinar a transformação T .

- *Registro fino*: consiste de um ajuste fino da matriz T de pré-alinhamento, reduzindo seus erros. O método tradicional de registro fino é o algoritmo “Ponto mais próximo por iteração” (ICP - *Iterative Closest Point*) [5, 12] e suas variantes [20, 52, 6, 29, 32, 54, 38, 17, 44]. Esse método precisa, para sua convergência, de uma transformação inicial boa (ou seja, razoavelmente próxima da solução), a qual é tipicamente a transformação obtida no pré-alinhamento. Com esta transformação inicial, estabelece-se pares de pontos correspondentes associando-se as amostras de $\mathcal{R}_1$ com suas amostras mais próximas na imagem $\mathcal{R}_2$, a partir dos quais uma nova transformação rígida T é calculada pela equação 1.1. Espera-se que se minimize ainda mais a distância quadrada entre os pares correspondentes. Após isso, as amostras de $\mathcal{R}_1$ são transformadas pela nova T e o processo é repetido para continuar ajustando os elementos da transformação, até que a distância entre os pontos correspondentes seja igual a um limiar pré-definido ou até que a diferença entre os resultados de duas iterações sucessivas seja menor que um limiar.

3. *Integração*: depois que todas as vistas, par a par, são registradas usando-se o processo descrito, elas são também integradas para obter uma representação única (não redundante) do objeto. E, finalmente, os dados integrados são aproximados por uma única superfície suave, tendo como resultado final a construção de um modelo 3D completo.

Todas as fases da Reconstrução 3D têm sido automatizadas com relativo sucesso, exceto a etapa de correspondência dentro da fase de pré-alinhamento, a qual ainda continua sendo um desafio. Entre algumas técnicas de aquisição de imagens, tem-se [41, 56]. Para registro fino de pares de vistas, tem-se [5, 12, 20, 52, 6, 29, 32, 54, 38, 17, 44] e para técnicas de integração automática, [49, 3, 17].

1.3 O problema do pré-alinhamento

O problema do pré-alinhamento, como comentado na Seção 1.2, pode ser dividido em dois sub-problemas: o problema de correspondência que consiste na identificação automática de um conjunto de N pares de pontos correspondentes (p_i, q_i) , $p_i \in \mathcal{R}_1$ e $q_i \in \mathcal{R}_2$, e o problema de posicionamento que consiste na obtenção da transformação rígida T que alinha ambas imagens. T deve ser obtida pela solução do problema de mínimos quadrados (equação 1.1), que minimiza os erros entre os N pares de correspondentes.

Para entender a dificuldade inerente ao problema de correspondência, é preciso desmembrá-lo em duas etapas:

1. A escolha dos N pontos p_i em $\mathcal{R}_1$ (ou determinação do descritor-referência $\mathcal{D}_1$)
2. A busca em $\mathcal{R}_2$ dos N pontos correspondentes q_i para os p_i (busca do descritor-alvo $\mathcal{D}_2$).

Para garantir pares válidos de pontos correspondentes devem-se escolher os N pontos p_i dentro da região de sobreposição A_{S1} . De fato, se a escolha for fora da região de sobreposição, não se poderá obter um T correto, pois estar-se-ia tentando alinhar informação que não é comum a ambas imagens. Portanto, esta etapa se constitui em um desafio: como se escolher estes pontos dentro da região de sobreposição, quando a tal região é completamente desconhecida? Uma solução (de força bruta) seria considerar todas as possíveis combinações $C_N^{m_1}$ onde n_1 é o número de amostras de $\mathcal{R}_1$ ¹. No entanto, a óbvia limitação desta opção é a ineficiência. A literatura consultada apresenta duas alternativas de solução :

- A redução do espaço de escolha, limitando-o ao uso de amostras com características especiais ou feições [53, 26, 43];
- A escolha aleatória [11, 1, 2].

Apesar da redução ser considerável, o que redundava a favor do quesito “eficiência”, as amostras com características especiais podem estar concentradas fora da região de sobreposição ou, simplesmente, dependendo da geometria particular da superfície, podem ser raras, ou, ainda, podem não existir; dificultando e até impedindo a extração dos N pares de correspondências. Mais sobre as desvantagens desta abordagem pode ser visto na Seção 2.3.4. Por outro lado, a maior dificuldade da escolha aleatória é a determinação do número de tentativas suficientes para escolher os p_i . De fato, é difícil determinar o número certo de tentativas que garanta que os p_i estejam contidos na região de sobreposição. Novamente, a extração de N pares de correspondências poderá não ser garantida.

¹Uma combinação C_s^n sem repetição, em análise combinatória, indica quantas variedades de subconjuntos diferentes com s elementos existem em um conjunto com n elementos[55].

Quanto à segunda etapa, encontrar os q_i correspondentes para os p_i leva a considerar um problema de busca. De fato, enquanto o problema anterior limitou-se apenas à escolha dos p_i , agora se tem em mãos um problema de busca com dois consequentes sub-problemas : “aonde se busca?”, se referindo ao espaço de busca, e “como se busca?”, se referindo ao procedimento de busca.

A alternativa mais trivial, e também a menos eficiente, para o espaço de busca é considerar o conjunto de todas as possíveis combinações $C_N^{n_2}$ onde n_2 é o número de amostras de $\mathcal{R}_2$. Esta é denominada de busca exaustiva. A literatura apresenta, como uma alternativa para ganho de eficiência, reduzir o espaço de busca, limitando-o ao uso de $\eta \ll n_2$ amostras com características especiais. Assim, um espaço menor de C_N^η combinações é considerado.

Quanto ao procedimento de busca, existe um consenso na literatura para estabelecer correspondências pela identificação de similaridade geométrica entre o descritor-referência e os descritores-alvo no espaço de busca de $\mathcal{R}_2$. Os descritores devem ser caracterizados por propriedades geométricas invariantes sob transformações rígidas, tais como ângulos, curvaturas e distâncias. Essas características são conhecidas também como propriedades intrínsecas da superfície, pois são independentes das posições e orientações dos descritores [9, 13]. Então, os potenciais candidatos para descritores-alvo são aqueles que possuem medidas invariantes, similares com as medidas do descritor-referência. É fortemente desejável que os descritores possuam medidas invariantes não triviais e assim, possam reduzir a ambiguidade no processo de busca. Um descritor com baixa discriminância resulta em múltiplas correspondências ambíguas.

Ainda com descritores discriminantes, o processo de comparação pode ser afetado por variações causadas por ruído gerado ou distorções próprias do processo de aquisição, o que pode ocasionar que as amostras ou feições, pertencentes à região de sobreposição, não se correspondam exatamente. Para tanto, o procedimento de busca de correspondentes deve considerar certa tolerância, o que aumenta ainda mais o número de candidatos. O desafio é então, um esquema de busca que seja capaz de manejar variações pequenas nos pares de pontos correspondentes na região de sobreposição, mantendo ainda a capacidade discriminante do descritor.

Mesmo com um descritor discriminante, a etapa de busca de um descritor-alvo para um descritor referência conduz a uma relação $1:m$ (onde m é um número inteiro positivo), fornecendo correspondências incorretas que levam à estimativa de transformações T_i falsas positivas. Então, um procedimento de filtragem se faz necessário para eliminar $m - 1$ transformações e ficar com uma única T . Para isso, as transformações candidatas devem passar por uma avaliação global para eliminar as falsas positivas. Usualmente esta avaliação consiste em contar o número de *ajustes*, ou seja, o número de amostras em $\mathcal{A}_1$ transformadas por T_i que têm distância até seu ponto mais próximo em $\mathcal{A}_2$, abaixo de um limiar pré-estabelecido. A T_i que possui o maior número de ajustes deverá ser a escolhida, resultando em um único T que resolve o sub-problema de posicionamento. Especial cuidado deve ser tido para que a transformação T_i que fornece a solução desejada não seja descartada neste processo.

Para evitar isso, a solução mais fácil porém não a mais eficiente consiste em que fazer com que todas as amostras de $\mathcal{A}_1$ e todas as transformações candidatas T_i participem do processo de avaliação.

À luz das considerações sobre os problemas do pré-alinhamento podem-se vislumbrar dois requisitos: existência e eficiência. Existência é referido a assegurar que sempre se consiga extrair N pares de correspondências, independente da geometria da imagem. Eficiência é referido à redução dos dados processados e a complexidade das operações realizadas. Para maior clareza, estes requisitos podem ser desmembrados nos seguintes itens:

Existência

- deve-se assegurar que haja pelo menos um descritor-referência na região de sobreposição A_{S1} : já que é o alinhamento de descritores que dá origem à transformação rígida, o número de pontos ou características na região A_{S1} deve ser suficiente para definir um descritor.
- deve-se assegurar que haja pelo menos um casamento válido de descritores: pelo menos um descritor-referência encontrado na região de sobreposição A_{S1} deve encontrar um descritor-alvo em A_{S2} . A condição necessária para isso é que os pontos pertencentes a A_{S1} e A_{S2} se correspondam, ou seja, representem aproximadamente a mesma localização no espaço da superfície.

Eficiência

- eficiência na determinação de descritores: deve-se determinar uma estratégia simples de construção de descritores a partir de um conjunto denso de dados para evitar a explosão combinatória de $C_N^{m_1}$ descritores.
- eficiência no casamento de descritores: este quesito tem a ver com as propriedades geométricas invariantes dos descritores construídos. Quanto maior o número de medidas invariantes caracterizando o descritor, maior será a sua capacidade discriminante.
- eficiência no processo de filtragem de falsas correspondências: a validação das correspondências deve passar por uma avaliação global. Esta avaliação não deve envolver o conjunto denso de dados.

1.4 Conteúdo do Trabalho

Após se estudar o problema de pré-alinhamento automático, consegue-se elaborar um conjunto de condições que garantem sua aplicação em quaisquer duas imagens de profundidade com a única

restrição de que possuam uma geometria não trivial, conforme explicado na Seção 1.4.1. Com base nestas condições, apresenta-se um algoritmo de pré-alinhamento automático de duas imagens de profundidade.

1.4.1 Escopo

No escopo deste trabalho, assume-se, como entrada do algoritmo proposto, duas imagens de profundidade parcialmente sobrepostas. Não há restrição para a geometria das superfícies, a não ser que elas não possuam somente pontos com mesmas curvaturas, como a esfera ou plano. Estas superfícies representam casos triviais que não são resolvidos, até o momento da elaboração deste trabalho, por nenhuma estratégia de registro.

1.4.2 Hipótese

O sistema visual humano consegue fazer registro de objetos através da essência da sua forma. Adicionalmente, Belongie e Malik [47] afirmam que um objeto pode ser modelado como um possivelmente infinito conjunto de pontos, mas sua forma é essencialmente capturada por um subconjunto finito dos seus pontos. Portanto, no lugar de um conjunto de dados denso, pode-se usar um conjunto esparso de dados que aproxime a forma representada pelos dados originais. Assim, a hipótese deste trabalho é que, para realizar um pré-alinhamento, é suficiente considerar uma representação que, com poucas amostras, consiga exprimir a essência da forma geométrica pertencente à superfície do objeto. Para isso, considera-se a versão simplificada (Figura 1.4(c)) da malha densa (Figura 1.4(b)), construída pela interpolação das amostras da imagem de profundidade (Figura 1.4(a)). É importante que a simplificação remova detalhes irrelevantes da superfície e gere uma malha que preserve vértices nas regiões de considerável variação de curvatura e, também, conserve uma certa densidade de vértices nas regiões com variações menores de curvatura. Isto garante que sobre a região de sobreposição, independente da sua geometria, sempre exista uma quantidade de pontos suficiente para construir descritores.

1.4.3 Contribuições

Os itens seguintes sintetizam as principais contribuições deste trabalho:

1. Uma proposta da redução do volume de dados de uma imagem de profundidade densa em uma malha triangular simplificada. Esta redução resulta em um ganho de eficiência nos procedimentos do algoritmo de determinação do descritor-referência, de busca pelo descritor-alvo e de eliminação de falsas correspondências, sem acarretar perdas geométricas significativas e, portanto, sem comprometer a precisão dos resultados.

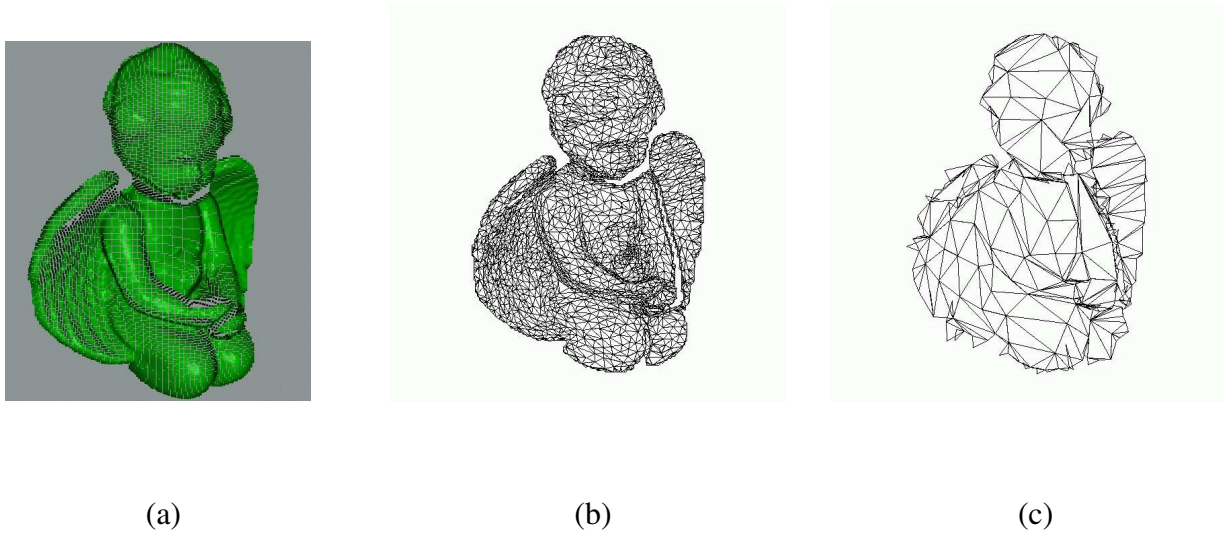


Fig. 1.4: (a) Imagem angel; (b) Malha densa; (c) Malha simplificada.

2. Determinação fácil de um descritor espacial que cubra uma região maior que uma vizinhança local e, ao mesmo tempo, esteja dentro da região de sobreposição para garantir o alinhamento. Tendo uma cobertura consideravelmente grande aumenta-se a capacidade discriminante do descritor, reduzindo, assim, o número de correspondências ambíguas.
3. Como consequência dos dois itens anteriores obtém-se um procedimento simples para eliminação de falsas correspondências. A qualidade do descritor proposto reduz drasticamente o número de transformações falsas positivas. Elas são facilmente eliminadas através da avaliação de erros de um conjunto reduzido de dados, composto pelos vértices da malha triangular simplificada.

1.5 Organização do trabalho

No restante, a dissertação está organizada em três partes, que estão distribuídas em cinco capítulos. A primeira parte, composta apenas pelo Capítulo 2, descreve o estado-da-arte, dificuldades e alternativas de solução para o problema de pré-alinhamento de duas imagens de profundidade. A segunda parte, composta pelos Capítulos 3 e 4 validam a hipótese deste trabalho, apresentando um algoritmo de pré-alinhamento eficiente de duas imagens de profundidade. A terceira e última parte, composta pelos Capítulos 5 e 6, apresenta alguns resultados do algoritmo proposto, as conclusões e os trabalhos futuros. Os parágrafos seguintes apresentam uma breve descrição do conteúdo de cada capítulo.

Capítulo 2 introduz conceitos relacionados ao pré-alinhamento de duas imagens de profundidade. Em seguida, descreve as principais abordagens de pré-alinhamento, apresentando, para cada uma, os principais trabalhos presentes na literatura com suas correspondentes limitações. Analisam-se, também, com certo detalhe, os principais problemas do pré-alinhamento e citam-se suas possíveis soluções.

Capítulo 3 apresenta um descritor tridimensional que não apenas incorpora informação local (dos próprios vértices da malha) mas também leva em consideração a informação global expressa pela posição relativa de vértices da malha “suficientemente afastados” entre si. Verifica-se como o fato de se considerar uma região mais abrangente, ao invés de uma simples região local, torna o descritor mais discriminante, reduzindo a quantidade de potenciais candidatos.

Capítulo 4 apresenta a proposta de reduzir o tamanho do conjunto de dados simplificando as imagens em malhas triangulares com poucos, porém suficientes, vértices. Cada etapa do procedimento é explicada em detalhes, validando a idéia de que a representação proposta descreve, de forma suficientemente fiel, o conjunto denso de dados. Apresenta-se, ainda neste capítulo, o algoritmo de pré-alinhamento baseado na proposta de uma representação malha triangular simplificada.

Capítulo 5 apresenta alguns resultados do algoritmo de pré-alinhamento proposto. Esses resultados validam a hipótese de que uma representação com número reduzido de amostras, porém fiel à essência da forma geométrica dos dados originais, é suficiente para realizar um pré-alinhamento.

Capítulo 6 resume as principais conclusões deste trabalho, destacando as contribuições e limitações. Em seguida, apresenta algumas melhorias que podem ser contempladas em trabalhos futuros.

Capítulo 2

Trabalhos Correlatos

Sejam $\mathcal{R}_1$ e $\mathcal{R}_2$ duas imagens de profundidade parcialmente sobrepostas, adquiridas a partir de diferentes pontos de vista e posicionadas em seu próprio sistema de referência local, o objetivo do pré-alinhamento é obter a transformação rígida T que coloque estas imagens em um sistema de coordenadas comum. Essa transformação pode ser obtida pela minimização do erro da distância entre pontos de descritores de ambas imagens. No entanto, estabelecer correspondências consistentes em forma automática é ainda um desafio devido ao desconhecimento da região de sobreposição, da qual os descritores correspondentes devem ser extraídos. Frente a isto, técnicas baseadas em tentativas e erros têm sido propostas, nas quais extrai-se um descritor-referência a partir de $\mathcal{R}_1$ e, sob comparações de medidas invariantes, tenta-se achar um descritor-alvo em $\mathcal{R}_2$. Visto que nenhuma medida invariante caracteriza de forma única um descritor, mais de um candidato em $\mathcal{R}_2$ será associado como potencial correspondente para um descritor-referência em $\mathcal{R}_1$, precisando de um procedimento adicional de filtragem de falsas correspondências.

As técnicas de pré-alinhamento automática podem ser classificadas em duas abordagens. A primeira considera a amostragem densa completa de ambas imagens para extrair os descritores correspondentes. Já a segunda considera apenas um subconjunto dos dados originais, restringido-se a usar apenas dados com certas características especiais como, por exemplo, pontos de curvatura média igual a zero, para reduzir o volume de dados processados. Este capítulo descreve os principais trabalhos que representam o estado-da-arte do pré-alinhamento de duas imagens de profundidade parcialmente sobrepostas, com especial ênfase nas suas limitações. Essas técnicas são analisadas à luz da análise feita no Capítulo 1, na qual foram determinados alguns critérios para a existência e eficiência de um pré-alinhamento, a saber, extração de pelo menos um descritor na região de sobreposição e, de pelo menos um casamento válido de descritores; e ainda, eficiência na determinação, no casamento dos descritores e no descarte de falsas correspondências.

O restante deste capítulo está organizado da seguinte forma. A Seção 2.1 apresenta algumas propriedades geométricas locais de superfícies as quais podem ser extraídas de uma imagem de profun-

didade, quando ela é considerada uma superfície parametrizada. Em seguida, a Seção 2.2 apresenta três técnicas pertencentes à abordagem baseada na totalidade dos dados. Mais adiante, na Seção 2.3 três técnicas correspondentes à abordagem baseada em apenas um subconjunto de pontos característicos são descritas. Na descrição de cada técnica inclui-se uma análise de acordo com os critérios mencionados e coloca-se em destaque as principais limitações. Por último, na Seção 2.4, conclui-se o capítulo, apresentando-se alguns requisitos para a existência e eficiência do pré-alinhamento de duas imagens de profundidade.

2.1 Propriedades Geométricas dos Pontos Amostrais de uma Superfície

Quando as amostras de uma imagem de profundidade são devidamente ligadas, constrói-se uma malha triangular, também denominada uma superfície de profundidade (*range surface*) [3]. Esta malha triangular é uma boa aproximação de uma superfície e os conceitos da Geometria Diferencial podem ser aplicados para estimar suas propriedades geométricas diferenciais, tais como curvaturas.

Considera-se uma superfície parametrizada S definida como o mapeamento suave (infinitamente diferenciável) $X : U \subset \mathcal{R}^2 \rightarrow \mathcal{R}^3$ que associa um par (u, v) no subconjunto aberto U de $\mathcal{R}^2$ ao vetor coordenado $X(u, v)$. Para garantir que o plano tangente esteja bem definido, assume-se que as derivadas parciais $X_u = \partial X / \partial u$ e $X_v = \partial X / \partial v$ existam. Seja $\alpha : I \subset \mathcal{R} \rightarrow U$ uma curva suave com $\alpha(t) = (u(t), v(t))$, então $\beta = X \circ \alpha$ é uma curva espacial parametrizada sobre a superfície. De acordo com a regra da cadeia, o vetor tangente β no ponto $\beta(t)$ é $u'(t)X_u + v'(t)X_v$. Deduz-se que o plano tangente da superfície em $X(u, v)$ é paralelo ao plano gerado pelos vetores X_u e X_v . O vetor normal à superfície é então

$$N = \frac{X_u \times X_v}{|X_u \times X_v|}.$$

A norma ao quadrado do vetor tangente $\beta'(t) = u'(t)X_u + v'(t)X_v$ à curva $\beta(t)$ no ponto $X(u, v)$ é:

$$|\beta'(t)|^2 = (u'X_u + v'X_v) \cdot (u'X_u + v'X_v)$$

ou

$$|\beta'(t)|^2 = Eu'^2 + 2F'u'v' + Gv'^2,$$

onde $E = X_u \cdot X_u$, $F = X_u \cdot X_v$, $G = X_v \cdot X_v$ são conhecidos como os coeficientes da primeira forma fundamental. Assim, a primeira forma fundamental é definida como a forma bilinear que associa a dois vetores do plano tangente o seu produto interno

$$I(\mathbf{t}_1, \mathbf{t}_2) = \mathbf{t}_1 \cdot \mathbf{t}_2.$$

Para $\mathbf{t}_1 = \mathbf{t}_2$

$$I(\mathbf{t}, \mathbf{t}) = |\mathbf{t}|^2 = Eu'^2 + 2Fu'v' + Gv'^2.$$

Seja p um ponto sobre a superfície S e seja N o vetor normal unitário de S em p . Ao se construir um plano Π_p que contém p e N , tem-se que a interseção de Π_p com S forma uma curva β em S com $\beta(0) = p$ e $\beta'(0) = \mathbf{t}$ (Figura 2.1(a)). Esta curva tem uma curvatura k que também será a curvatura normal $k_p(\mathbf{t})$ de S em p na direção $\mathbf{t}$. No entanto Π_p não é o único plano. Ao rodar Π_p ao redor de N , forma-se uma nova curva sobre S com sua própria curvatura normal.

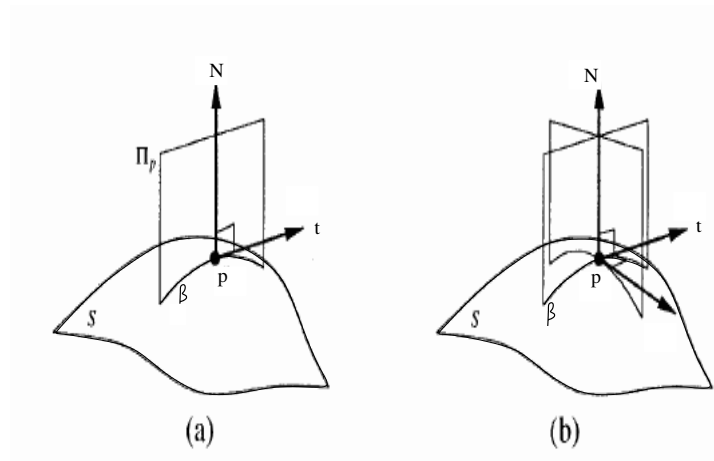


Fig. 2.1: Curvatura de uma superfície

De fato, pode-se ver na Figura 2.1(b) que se tem um conjunto infinito dessas curvaturas normais ao redor de p , uma em cada direção. As direções onde k_n alcança o maior valor e o menor valor são chamadas de direções principais e os valores são chamados de curvaturas principais k_1 e k_2 , respectivamente.

Seja $N : X \rightarrow S^2$ o mapa de Gauss que associa a cada ponto sobre a superfície seu vetor normal unitário $N_p = \frac{X_u \times X_v}{|X_u \times X_v|}$ na esfera unitária S^2 . A derivada dN_p do mapa de Gauss em p mede a variação do vetor normal numa vizinhança de p , ou seja, mede quanto a superfície se “curva” próxima de p .

A segunda forma fundamental é definida por:

$$II(\mathbf{t}, \mathbf{t}) = \langle dN_p(\mathbf{t}), \mathbf{t} \rangle = eu'^2 + 2fu'v' + gv'^2, \quad (2.1)$$

onde $e = N_u \cdot X_u$, $f = N_u \cdot X_v = N_v \cdot X_u$, $g = N_v \cdot X_v$ são seus coeficientes.

Os coeficientes da primeira e segunda forma fundamental são relacionados pelo mapa de Gauss-Weingarten [16]:

$$dN_p(\mathbf{t}) = \begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} = \frac{1}{EG - F^2} \begin{pmatrix} fF - gG & gF - fG \\ eF - fE & fF - eE \end{pmatrix}.$$

Se forem considerados vetores base (X_u, X_v) ortonormais, tem-se que $E = G = 1, F = 0$ e portanto o tensor segunda forma fundamental é

$$II = dN_p = \begin{pmatrix} e & f \\ f & g \end{pmatrix} = \begin{pmatrix} N_u \cdot X_u & N_v \cdot X_u \\ N_u \cdot X_v & N_v \cdot X_v \end{pmatrix} \quad (2.2)$$

Os autovalores dessa matriz correspondem às curvaturas normais mínima e máxima k_1 e k_2 , respectivamente, e os autovetores correspondentes são denominados as direções principais. A grandeza H é chamada de curvatura média e é definida como a média das curvaturas principais $(k_1 + k_2)/2$. E a grandeza K é chamada de curvatura gaussiana e é definida como o produto das curvaturas principais $k_1 k_2$.

De acordo com o valor dessas curvaturas um ponto $p \in S$ é denominado:

1. elíptico, se $K > 0$ (k_1 e k_2 com o mesmo sinal); (Figura 2.2(c))
2. hiperbólico, se $K < 0$ (k_1 e k_2 com sinais opostos); (Figura 2.2(b))
3. parabólico, se $K = 0$ e $H \neq 0$ (k_1 ou k_2 nula, mas não ambas); (Figura 2.2(a))
4. planar, se $K = H = 0$ ($k_1 = k_2 = 0$);
5. umbílico, se $K \geq 0$ ($k_1 = k_2$).

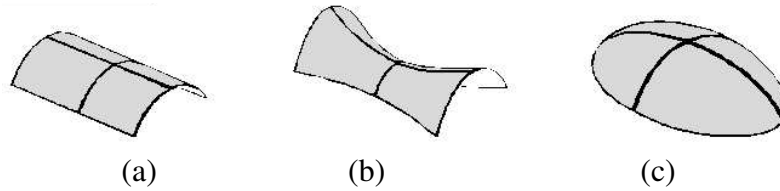


Fig. 2.2: Classificação dos pontos em (a) parabólico, (b) hiperbólico e (c) elíptico

Neste trabalho trata-se de malhas triangulares, ao invés de superfícies contínuas, precisamos, portanto, uma versão discreta do algoritmo de determinação de curvaturas. O algoritmo de Rusinkiewicz [45] é aplicado. Para obter as curvaturas principais, calcula-se, em forma discreta, o tensor II e então extrai-se os seus auto-valores. Para calcular o tensor II em um vértice p , Rusinkiewicz propõe que II seja a média dos tensores calculados individualmente em cada uma das faces F da vizinhança de p (faces adjacentes). II é estimado em cada face usando as três direções bem definidas dadas por suas três arestas e_i e diferenças finitas para aproximar a derivada direcional de N na direção de e_i . Sejam n_1, n_2 e n_3 os vetores normais associados aos três vértices da face F e seja u, v dois vetores ortonormais no plano tangente, conforme mostrado na Figura 2.3 [45].

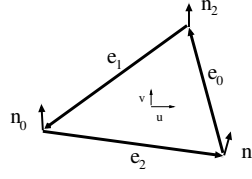


Fig. 2.3: Vetores normais e arestas da face

Da equação 2.1 pode-se deduzir as seguintes três equações.

$$II \begin{pmatrix} e_0.u \\ e_0.v \end{pmatrix} = \begin{pmatrix} (n_2 - n_1).u \\ (n_2 - n_1).v \end{pmatrix}$$

$$II \begin{pmatrix} e_1.u \\ e_1.v \end{pmatrix} = \begin{pmatrix} (n_0 - n_2).u \\ (n_0 - n_2).v \end{pmatrix}$$

$$II \begin{pmatrix} e_2.u \\ e_2.v \end{pmatrix} = \begin{pmatrix} (n_1 - n_0).u \\ (n_1 - n_0).v \end{pmatrix}$$

O sistema de equações obtido pode ser resolvido para II usando mínimos quadrados.

2.2 Uso do conjunto completo de dados

O uso do conjunto completo de dados visa a garantir a melhor solução, por considerar a avaliação de todas as possíveis correspondências, porém o processo não é eficiente como todo algoritmo de força bruta. Enquanto um grande esforço é dispensado para garantir a extração de correspondências, o quesito eficiência é prejudicado. Consequentemente, esta abordagem se torna impraticável pela explosão combinatorial resultante do uso de uma grande amostragem. De fato, já que um mínimo de três pares de pontos não colineares são necessários para encontrar uma transformação rígida, os descritores devem ter no mínimo três amostras. Supondo que tanto $\mathcal{R}_1$ como $\mathcal{R}_2$ têm n amostras cada uma, então, todas as possíveis combinações de três amostras para formar descritores em $\mathcal{R}_1$ seria de C_3^n (espaço de escolha), e em $\mathcal{R}_2$ seria, também, de C_3^n (espaço de busca); então, o espaço de busca total por pares de correspondentes seria de $C_3^n C_3^n$ o que dá uma complexidade de $O(n^6)$. É por tal razão que os métodos desta primeira abordagem fazem uma escolha aleatória dos descritores em $\mathcal{R}_1$ com o intuito de reduzir o espaço de escolha de C_3^n a k tentativas. Desta forma, o espaço de busca por pares de correspondentes fica reduzido de $C_3^n C_3^n$ a $k C_3^n$. Este número é ainda grande, pois as imagens são compostas de milhares de amostras. Detalhar-se-ão, a seguir, três técnicas pertencentes a esta abordagem.

2.2.1 DARCES baseado em RANSAC (DBR)

DARCES baseado em RANSAC é um algoritmo iterativo [11] que encaixa o algoritmo DARCES dentro do processo iterativo do algoritmo RANSAC cujas iniciais correspondem a *Random Sample Consensus* ou “Consenso de amostragem aleatória” ([28, 31]). O princípio básico de RANSAC é tomar um número n de amostras aleatórias, durante um processo iterativo de k passos. Após as k iterações é esperado que as amostras sejam pontos válidos para a aplicação, na qual RANSAC está sendo usado. O número k depende da probabilidade de se escolher uma boa amostra (*inlier*), dentre um conjunto completo de amostras. DARCES corresponde às iniciais de *Data-Aligned Rigidity Constrained Exhaustive Search* ou “busca exaustiva restrita à rigidez dos dados alinhados”, e consiste em construir um triângulo a partir de três amostras ($n = 3$) s_1 , s_2 e s_3 , na imagem $\mathcal{R}_1$ e tentar encontrar o mesmo triângulo na imagem $\mathcal{R}_2$, realizando uma varredura na sua amostragem. Para escolher esses pontos, primeiro uma amostra s_1 é escolhida aleatoriamente e os outros pontos de controle s_2 e s_3 são determinados de forma que um triângulo equilátero seja formado.

Para encontrar um triângulo correspondente em $\mathcal{R}_2$ é feito o seguinte procedimento: cada amostra de $\mathcal{R}_2$ é assumida como uma correspondente provável m_1 para s_1 e, já que a distância é invariante sob transformações rígidas, a busca por um correspondente para o ponto s_2 é restrita a uma esfera de raio $d_{12} = \|s_1 - s_2\|$, com seu centro em m_1 (Figura 2.4(b)); a busca pelo ponto correspondente para s_3 é realizada na interseção de duas esferas com raios $d_{23} = \|s_2 - s_3\|$ e $d_{13} = \|s_1 - s_3\|$ com centros em m_2 e m_1 , respectivamente (Figura 2.4(c)).

Cada par de triângulos potenciais correspondentes gerará uma transformação T_i , a qual será considerada uma transformação candidata para o registro se passar pela avaliação do ajuste de um conjunto de 15 amostras s_j as quais são selecionadas de forma que formem um triângulo grande em volta do triângulo $s_1 s_2 s_3$ (Figura 2.4(d)). Quando o erro das amostras transformadas $T_i s_j$ e o ponto mais próximo na imagem $\mathcal{R}_2$ for maior que um pré-determinado limiar, a transformação T_i é rejeitada. Finalmente as transformações T_i remanescentes passam por uma avaliação global para descartar as falsas positivas. O ajuste é avaliado considerando apenas um subconjunto $\mathcal{R}_1$ de $\mathcal{A}_1$. A transformação com o maior número de ajustes é finalmente escolhida.

A seguir apresenta-se uma breve análise dos itens correspondentes ao critério de existência e eficiência.

Existência

- Extração de pelo menos um descritor-referência $\mathcal{D}_1$ em $\mathcal{A}_{S1}$: O princípio do método RANSAC não é determinístico. Deve-se conhecer a probabilidade de escolha de boas amostras que conduzam a um alinhamento razoável.
- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: se for encontrado um D_1 em $\mathcal{A}_{S1}$, é garan-

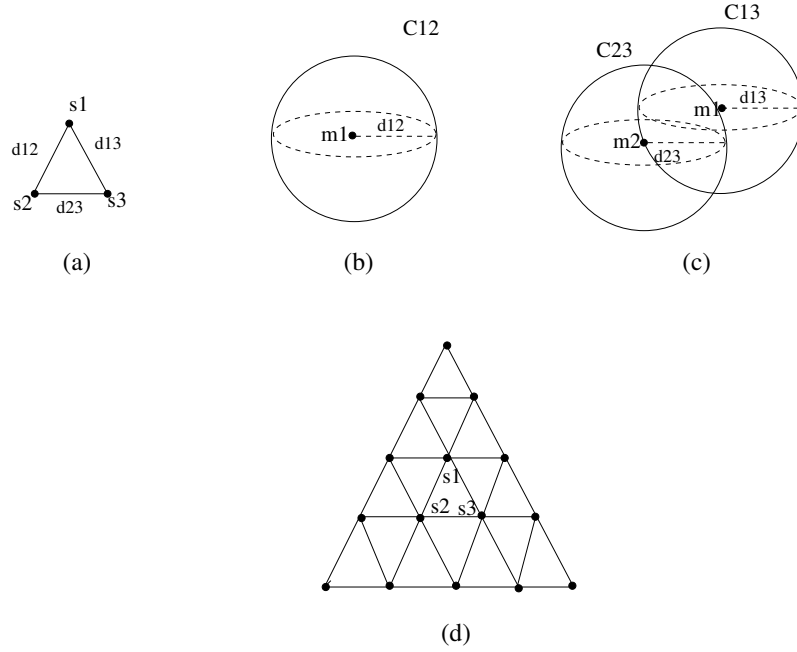


Fig. 2.4: (a-c) Busca de um triângulo correspondente (d) Triângulo grande (15 pontos)

tida pela busca exaustiva realizada em $\mathcal{R}_2$ (amostragem completa).

Eficiência

- Eficiência na construção do descritor-referência $\mathcal{D}_1$: é simples; são apenas três amostras, uma das quais é escolhida aleatoriamente.
- Eficiência na busca de um descritor-alvo $\mathcal{D}_2$: embora o descritor $\mathcal{D}_2$ seja procurado em regiões esféricas, estas regiões são subconjuntos densos de $\mathcal{R}_2$. O descritor é formado por amostras sem nenhuma característica distinta e com a configuração triangular. Portanto, a orientação espacial não é suficientemente discriminada, resultando em muitos casamentos ambíguos.
- Eficiência no processo de filtragem : é realizada através da avaliação de erros das amostras de um triângulo grande e de uma subamostragem de $\mathcal{R}_1$.

2.2.2 Casamento de Pontos Orientados (Imagens *Spin* - CPO)

CPO procura pares de correspondentes individualmente (p_i, q_i) , $p_i \in \mathcal{R}_1$ e $q_i \in \mathcal{R}_2$, para depois determinar pares de descritores [1]. Os descritores são construídos formando grupos consistentes de $N > 3$ pontos escolhidos aleatoriamente em $\mathcal{R}_1$ (descritor $\mathcal{D}_1$), com seus correspondentes em $\mathcal{R}_2$ ($\mathcal{D}_2$). A busca de um ponto correspondente q_i para um ponto p_i em $\mathcal{R}_1$ é realizada pela comparação de imagens *spin* associadas aos pontos. Uma imagem *spin* consiste de um histograma 2D construído pela

distância horizontal (até o vetor normal da superfície) e a distância perpendicular (ao plano tangente no ponto orientado) de pontos de uma região vizinha, conforme mostrado na Figura 2.5.

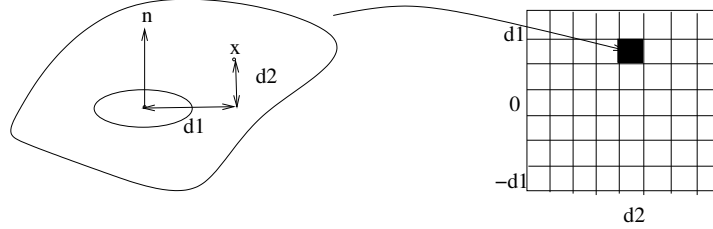


Fig. 2.5: Imagem *spin*

Duas etapas de busca de correspondências são consideradas: busca de um ponto correspondente q_i para p_i , feito pela comparação das suas imagens *spin*, e busca de agrupamentos consistentes de $N > 3$ pontos com seus correspondentes (casamento de descritores). Os grupos de pares são considerados consistentes quando, baseados no princípio de rigidez, a distância entre seus pontos (p_i, p_j) é aproximadamente igual à distância entre os seus pontos correspondentes (q_i, q_j) . Finalmente, para cada grupo, uma transformação candidata T_i é determinada, a qual é validada escolhendo alguns pontos na imagem e contando as distâncias, entre seus pontos transformados e os pontos mais próximos na imagem $\mathcal{R}_2$, que são menores que um limiar pré-estabelecido. O T_i que fornecer o maior número de ajustes será o T escolhido.

Veja-se, a seguir, uma breve análise dos itens relacionados aos quesitos de existência e eficiência.

Existência

- Extração de pelo menos um descritor-referência $\mathcal{D}_1$ em A_{S1} : não está garantida pois, em razão da aleatoriedade, as $N > 3$ amostras podem não estar contidas na região de sobreposição.
- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: se for encontrado um $\mathcal{D}_1$ em A_{S1} , é garantida pois a busca de pontos correspondentes é realizada, exaustivamente, usando-se a amostragem completa da imagem $\mathcal{R}_2$.

Eficiência

- Eficiência na construção do descritor-referência $\mathcal{D}_1$: a construção de um descritor consiste em escolher aleatoriamente $N > 3$ amostras na imagem $\mathcal{R}_1$. O custo para extrair as imagens *spin* não é alto, simplesmente calcula-se o vetor normal e determinam-se os pontos da vizinhança que estão a certas distâncias do ponto orientado.
- Eficiência na busca de um descritor-alvo $\mathcal{D}_2$: não é eficiente porque usa a amostragem completa como espaço de busca e porque a medida invariante associada ao descritor é somente a distân-

cia. A configuração espacial entre as amostras não é considerada redundando na obtenção de muitas correspondências ambíguas.

- Eficiência no processo de filtragem: o método não passa pela avaliação global pois escolhe apenas uma certa quantidade de amostras para medir os erros de aproximação até a imagem $\mathcal{R}_1$. Este procedimento procura ser eficiente ao usar pouca informação, porém pode redundar em escolher erroneamente uma transformação falsa positiva como solução. Isto foi verificado por Planitz [7, 8] em casos de geometrias que apresentam simetria (Seção 5.3.3).

2.2.3 Casamento de Histogramas Geométricos (CHG)

CHG usa RANSAC para, durante um processo iterativo de k passos, escolher aleatoriamente uma face t_{1i} da malha triangular que representa $\mathcal{R}_1$ e procurar por uma face correspondente t_{2i} na malha triangular que representa a imagem $\mathcal{R}_2$ [2]. A busca por correspondentes é feita pela comparação de histogramas geométricos associados às faces. Um histograma geométrico [2] descreve a relação de uma face t_i com as faces t_{ij} dentro de uma vizinhança pré-definida. Esta relação é caracterizada por duas medidas: o ângulo relativo entre os vetores normais das faces e a distância d do centróide da face t_i até os centróides das faces t_{ij} (Figura 2.6).

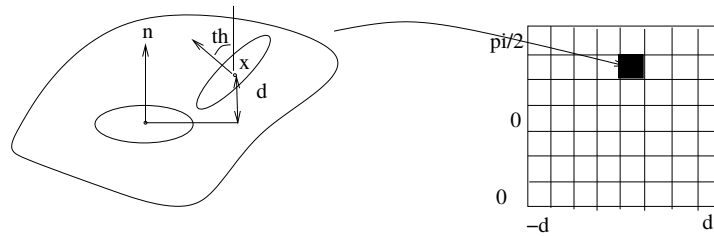


Fig. 2.6: Histograma geométrico

Potenciais correspondentes para uma face t_i na malha associada a $\mathcal{R}_1$ são determinados pelo uso de uma métrica de similaridade para obter uma transformação T_i . Diferente de DBR, a filtragem de falsas transformações é feita pela avaliação global de algumas transformações candidatas, escolhidas aleatoriamente, e não de todas. Na avaliação global todas as faces da malha transformada são usadas.

A seguir apresenta-se uma breve análise dos itens correspondentes ao critério de existência e eficiência:

Existência

- Extração de pelo menos um descritor-referência $\mathcal{D}_1$ em A_{S1} : O princípio do método RANSAC não é determinístico. Deve-se conhecer a probabilidade de escolha de boas amostras que

conduzam a um alinhamento razoável.

- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: se for encontrado um $\mathcal{D}_1$ em A_{S1} , é garantida a extração pois é realizada exaustivamente a procura usando-se a malha que representa a imagem $\mathcal{R}_2$.

Eficiência

- Na construção do descritor-referência $\mathcal{D}_1$: a construção de descritores é simples pois se usa apenas a informação geométrica da faceta t_i . O custo para construir os histogramas geométricos não é alto, simplesmente calcula-se o vetor normal e determinam-se as faces da vizinhança que estão a uma certa distância e a um certo ângulo da face que está sendo avaliada.
- Na busca de um descritor-alvo $\mathcal{D}_2$: não é eficiente pois usa a amostragem completa como espaço de busca e a configuração do descritor é planar. O método tende também a gerar muitas correspondências ambíguas; em especial, quando houver regiões suaves na superfície. Os pontos dentro destas regiões possuem mesma variação angular (quase mesma curvatura) e portanto resultam em um mesmo histograma geométrico.
- No processo de filtragem: diferente dos métodos anteriores, não todas as transformações candidatas passam pela avaliação global. Apenas algumas que são escolhidas aleatoriamente. Aparentemente este procedimento é mais eficiente, pois não avalia exaustivamente o grande número de transformações candidatas, porém pode conduzir à escolha errada de um falso positivo como analisado por Planitz e apresentado na seção 5.3.3.

2.2.4 Limitações

O ponto forte da abordagem baseada na amostragem completa é o princípio básico de busca exaustiva. De fato, considerando exaustivamente todos os prováveis descritores em $\mathcal{R}_1$ e buscando um correspondente dentre todas as possibilidades em $\mathcal{R}_2$, assegura-se que nenhum potencial candidato ao pré-alinhamento seja perdido. Isto satisfaz o quesito “existência”. Porém há duas limitações:

- os espaços de escolha e busca se tornam enormes devido à explosão combinatorial. Para tentar aliviar a ineficiência do procedimento, acaba-se optando (a) pela escolha aleatória de alguns descritores-referência dentre o grande número de combinações e (b) pela simplificação do processo de avaliação global das transformações candidatas. Os resultados obtidos pela escolha aleatória, serão, no entanto, razoáveis somente se a probabilidade de obtenção de boas amostras, for conhecida. Quanto ao procedimento de filtragem, acredita-se que se não se levar em consideração todas as transformações candidatas e, todos os pontos da imagem ou um conjunto representativo deles, corre-se o risco de escolher um falso positivo como solução.

- os descritores propostos são geralmente pouco expressivos no sentido de serem pouco representativos da forma geométrica global da superfície.

2.3 Uso de um conjunto de dados com características especiais

Visando uma maior eficiência, as técnicas pertencentes a esta abordagem reduzem o volume de dados, restringido-o somente a amostras que possuam certas características especiais. Se n é o número total de amostras das imagens $\mathcal{R}_1$ e $\mathcal{R}_2$, espera-se que o subconjunto tenha η amostras, onde $\eta \ll n$ (η muito menor que n).

Assim sendo, o número de todas as possíveis combinações de três amostras para formar descritores, na imagem $\mathcal{R}_1$, seria de C_3^η , e o espaço de busca total para encontrar pares de correspondentes, seria de $C_3^\eta C_3^\eta$ o que dá uma complexidade de $O(\eta^6)$. Como $\eta \ll n$, pode-se perceber um ganho em eficiência proporcional à redução do conjunto de amostras. No entanto, o principal problema desta abordagem surge quando as características não são muito discriminantes, de forma que podem não estar presentes em todas as imagens, fazendo que o quesito existência não seja satisfeito.

Há um grande número de características que podem ser extraídas a partir de uma imagem de profundidade, de modo a estabelecer critérios de similaridade entre amostras (da imagem de profundidade) ou de vértices (da malha que aproxima a imagem de profundidade). As características mais utilizadas são medidas de curvatura (gaussiana e média), as quais são geralmente extraídas da malha triangular, usando-se uma versão discreta da geometria diferencial apresentada na Seção 2.1. Ou então, podem ser estimadas diretamente das imagens de profundidade através da aplicação de máscaras usadas no processamento de imagens bidimensionais. As seguintes máscaras podem ser usadas nos pontos para estimar derivadas (diferenças finitas) no ponto $(u, v, i(u, v))$ [18, 39, 24, 4]:

- Diferenças de um único lado com respeito a u ou v : $\partial_u i(u, v) = i(u, v) - i(u - 1, v)$
- Diferenças centrais simétricas: $\partial_u i(u, v) = i(u + 1, v) - i(u - 1, v)$
- Filtro de Prewitt : média das diferenças centrais em três linhas ou colunas.

Com uso destes filtros, outras características, que dependem do cálculo de derivadas, tais como quinas e bordas sobre a imagem podem também ser detectadas. Na Figura 2.7 alguns pontos característicos das imagens são apresentadas.

A seguir, três técnicas que exploram feições características de uma imagem.

2.3.1 Casamento de Curvas Bitangentes (CCB)

Este método é baseado na extração e casamento de descritores formados por pontos de curvas bitangentes [53]. Pontos bitangentes são pares de pontos na superfície que têm o mesmo plano tan-

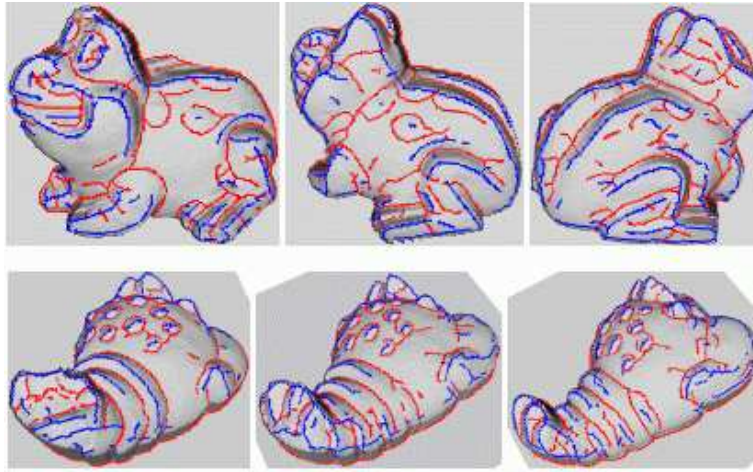


Fig. 2.7: Curvas de máxima curvatura

gente ou, também chamado, plano bitangente (Fig 2.8a.). Quando o plano bitangente rola sobre a superfície, os pontos bitangentes descrevem curvas bitangentes (Fig 2.8b.). Um descriptor é definido por quatro pontos, isto é, os pontos extremos de dois segmentos, um segmento em cada curva do par de curvas bitangentes. O espaço de escolha é, então, composto pelos pontos das curvas bitangentes. E a distância entre os pontos bitangentes, ao longo das suas respectivas curvas bitangentes, é usada como uma medida invariante, conforme ilustrado na Figura 2.9. Esta medida é usada para buscar descritores-alvo no espaço de busca formado pelo conjunto de pontos das curvas bitangentes na imagem $\mathcal{R}_2$. As transformações candidatas são aplicadas e verificadas, calculando-se a distância de uma subamostragem da imagem $\mathcal{R}_1$ até os pontos mais próximos na imagem $\mathcal{R}_2$.

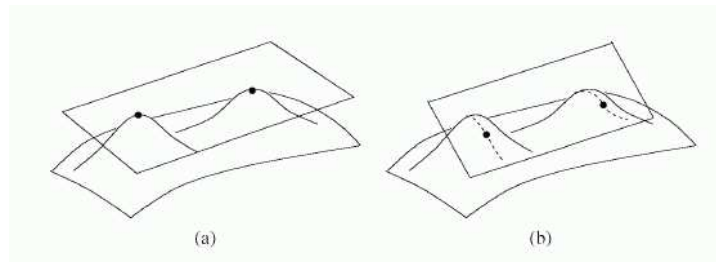


Fig. 2.8: (a) Pontos bitangentes (b) Plano bitangente rolando sobre a superfície

Veja-se, a seguir, uma breve análise dos itens relacionados aos quesitos de existência e eficiência.

Existência

- Extração de pelo menos um descriptor-referência $\mathcal{D}_1$ em A_{S1} : não está garantida, pois os pontos bitangentes são pontos característicos muito especiais que podem não estar presentes em todas as superfícies.

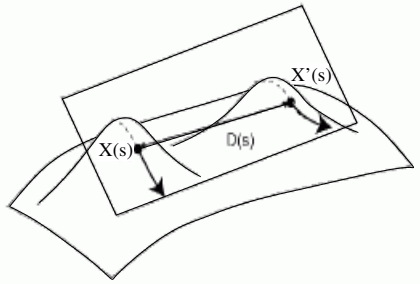


Fig. 2.9: Distância entre um par de pontos bitangentes

- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: se houver pontos bitangentes na região de sobreposição A_{S1} e o método de detecção de características for preciso, estas também serão detectadas em A_{S2} pois as características especiais são invariantes sob transformações rígidas. Além disso a busca é realizada exaustivamente usando-se como espaço de busca, o conjunto de pontos das curvas bitangentes em $\mathcal{R}_2$.

Eficiência

- Eficiência na construção do descritor-referência $\mathcal{D}_1$: a determinação das curvas bitangentes segue um processo relativamente complexo já que precisa-se determinar o plano tangente em cada ponto, os pontos que compartilham o mesmo plano, e as curvas bitangentes contínuas. Os descritores, par de segmentos são parametrizados pelo comprimento de arco. Este procedimento é usualmente eficiente em vista do número reduzido de dados.
- Eficiência na busca de um descritor-alvo $\mathcal{D}_2$: é eficiente pois usa como espaço de busca uma subamostragem composta por pontos bitangentes. O descritor é composto por um par de segmentos cuja configuração é distinta porque representa a ligação entre duas curvas bitangentes que são curvas distintas da superfície.
- Eficiência no processo de filtragem: esta técnica, similar à DBR, usa a avaliação de erros de uma sub-amostragem de $\mathcal{R}_1$, sendo portanto mais eficiente que as técnicas que usam uma amostragem densa.

2.3.2 Casamento de Curvas Intrínsecas (CCI)

Malhas ligando todas as amostras das imagens de profundidade são construídas neste trabalho [26]. Em cada vértice da malha, a curvatura média é determinada e um processo de interpolação é feito para determinar as posições sobre a malha as quais definem uma curva com curvatura média igual a zero (Figura 2.10). Essas curvas são agrupadas de duas em duas e, em cada par, o segmento mais curto (conector) que as une é determinado. Os pontos extremos de um par de conectores definem

um descritor e o comprimento dos conectores é usado como propriedade invariante para estabelecer correspondências (Figura 2.11). A qualidade da transformação T é verificada em duas etapas: a primeira avalia o ajuste dos pontos característicos e, a segunda, o ajuste de todas as amostras. Na etapa final, a transformação que ajusta o maior número de amostras é escolhida.

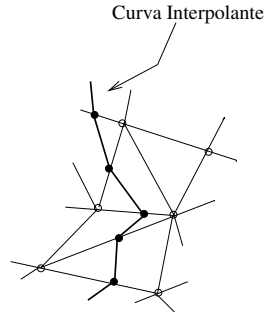


Fig. 2.10: Curva interpolante

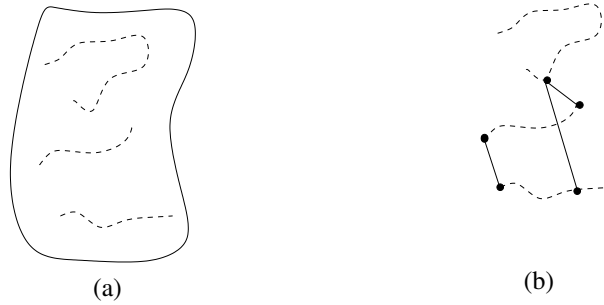


Fig. 2.11: (a) Curvas Intrínsecas (b) Representações de *features*

A seguir apresenta-se uma breve análise dos itens correspondentes ao critério de existência e eficiência.

Existência

- Extração de pelo menos um descritor-referência $\mathcal{D}_1$ em A_{S1} : não está garantida, pois os pontos com curvatura média igual a zero, assim como os pontos bitangentes, são pontos característicos muito especiais que podem não estar presentes em todas as superfícies. Além disso, a extração de curvaturas é através do cálculo de derivadas, portanto é condicionada a não ter quinas ou arestas com descontinuidade de orientação onde não é possível obter derivadas.
- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: se houver pontos com curvatura média igual a zero na região de sobreposição A_{S1} e o método de detecção de características for preciso, estas

também serão detectadas em A_{S2} . A busca é realizada exaustivamente usando-se como espaço de busca o conjunto de pontos das curvas de curvatura média igual a zero em $\mathcal{R}_2$.

Eficiência

- Eficiência na construção do descritor-referência $\mathcal{D}_1$: a extração de curvas segue um processo relativamente complexo. A determinação da curvatura média é feita em cada amostra de ambas imagens e a extração da curva contínua, através de um processo de interpolação. As curvas são agrupadas, duas a duas, para obter o menor segmento que une cada par de curvas, cujos extremos compõem os descritores. No entanto o processo se torna eficiente por trabalhar em cima de um conjunto reduzido de dados.
- Eficiência na busca de um descritor-alvo $\mathcal{D}_2$: é eficiente, pois usa como espaço de busca uma subamostragem composta por pontos característicos. O descritor é composto por pares de segmentos cuja configuração é distinta já que representam a menor ligação entre curvas de curvatura média igual a zero, as quais são curvas distintas da superfície.
- Eficiência no processo de filtragem: a filtragem é conduzida em duas etapas, após avaliar o ajuste da subamostragem de pontos especiais, avalia-se também o ajuste da amostragem completa de $\mathcal{R}_1$. Esta dupla avaliação incrementa o custo deste processo, portanto não é eficiente.

2.3.3 Casamento baseado no gradiente de intensidade (CGI)

CGI usa as imagens de intensidade $\mathcal{I}_1$ e $\mathcal{I}_2$ associadas a $\mathcal{R}_1$ e $\mathcal{R}_2$, respectivamente [43]. A partir destas imagens, os máximos locais do gradiente de intensidade são extraídos como pontos característicos. O número de pontos especiais encontrados está em função do limiar considerado e da textura do objeto; esta técnica, portanto, é restrita a objetos texturizados (cantos resultam em bons candidatos pois neles o gradiente de intensidade é máximo em ambas direções).

Com os pontos 3D associados a esses pontos 2D é feita uma tetraedrização Delaunay a qual é invariante a transformações geométricas. Portanto, podem-se usar os seus triângulos como descritores. Invariantes, como distância e ângulos, são medidas usadas para comparação e obtenção de triângulos correspondentes. Somente os triângulos que têm aproximadamente os mesmos comprimentos de aresta e diferença angular, entre os vetores normais do seus vértices, serão considerados triângulos correspondentes. Finalmente, a eliminação de transformações candidatas é feita pela escolha da transformação que ajusta o maior número de pontos especiais.

Veja-se, a seguir, uma breve análise dos itens relacionados aos quesitos de existência e eficiência.

Existência:

- Extração de pelo menos um descritor-referência $\mathcal{D}_1$ em A_{S1} : o sucesso do algoritmo é altamente dependente do procedimento de extração de pontos especiais em objetos texturizados, portanto, este algoritmo é limitado quanto ao tipo de imagens.
- Extração de pelo menos um descritor-alvo para $\mathcal{D}_1$: Mesmo que triângulos formados por pontos especiais forem detectados na região de sobreposição A_{S1} , o método não pode garantir que os seus triângulos correspondentes sejam também detectados. De fato, medidas fotométricas são mais sensíveis às interações luz-superfície-observador, assim, a intensidade pode variar de um ponto de vista a outro.

Eficiência

- Eficiência na construção do descritor-referência $\mathcal{D}_1$: o custo está em função do processamento de imagens bidimensionais para se extrair os pontos máximos de gradiente e da determinação da tetraedrização para se construírem os descritores. Considera-se eficiente uma vez que trabalha-se com um conjunto reduzido de dados.
- Eficiência na busca de um descritor-alvo $\mathcal{D}_2$: é eficiente, pois usa um espaço de busca limitado a uma subamostragem composta por um conjunto de pontos com características especiais. As amostras do descritor são distintas, porém a sua configuração é planar.
- Eficiência no processo de filtragem: usa-se a subamostragem composta pelos pontos especiais para avaliar o erro global.

2.3.4 Limitações

A abordagem baseada em pontos característicos surge como uma alternativa para a redução do espaço de busca de pares de correspondentes. A detecção de feições diferenciadas fornece uma representação concisa da geometria global, gerando um número de pontos significativamente menor que o número original de amostras. Além disso, essa representação é independente do ponto de vista, pois os pontos característicos se mantêm inalteráveis ao mudar o ângulo de visão. Portanto, esta representação é, além de reduzida, também invariante sob transformações rígidas.

No entanto, após uma análise, pode-se perceber algumas desvantagens nesta representação, as quais, a seguir, apresentamos em separado :

- Restrição à classe de superfície: uma escolha particular de pontos característicos pode limitar o escopo do algoritmo para classes particulares de superfícies. Isto impede assegurar a extração de pontos na região de sobreposição.

- Tendência de concentração de pontos em um local: a concentração de pontos especiais em pequenas regiões e falta deles em outras coloca em risco a extração de pontos na região de sobreposição, quando esta não apresenta esses detalhes. Além disso, a obtenção de curvas muito densas ou regiões de conglomerado de pontos conduz à necessidade de filtrar pontos demasiadamente próximos para construir descritores, o que onera o processo de construção [37].

Quanto aos descritores propostos pelas técnicas descritas pode-se dizer que por um lado, são expressivos porque são construídos a partir de amostras salientes (pontos característicos), mas por outro lado, a configuração não captura bem a forma tridimensional global da superfície.

2.4 Considerações Finais

Neste capítulo apresentaram-se as duas abordagens nas quais as técnicas de pré-alinhamento se enquadram: baseada na amostragem completa e restrita a um subconjunto de pontos com características especiais. Também são apresentadas algumas das mais importantes técnicas de pré-alinhamento pertencentes a cada categoria. Essas técnicas foram analisadas de acordo ao conjunto de critérios de existência e eficiência estabelecidos no Capítulo 1. A Tabela 2.1 resume os aspectos mais relevantes de cada técnica. Esta análise mostra que, embora muito se tenha evoluído, as técnicas ainda apresentam diversas limitações. Nenhum dos algoritmos reúne todos os requisitos expressos nos critérios considerados, o que limita a sua aplicabilidade.

Neste trabalho propõe-se um algoritmo que visa contornar as limitações encontradas, reunindo as vantagens das duas abordagens:

- O uso de um conjunto reduzido de dados, para satisfazer o quesito de eficiência. Este conjunto reduzido é, na verdade, uma representação que, com poucos pontos, consegue extrair as características globais essenciais da forma do objeto.
- O uso de todos os pontos dessa representação reduzida, porém global, da superfície, para contribuir com o quesito de existência. O uso exaustivo de todos os pontos desta representação assegurará que informação essencial para o registro não seja perdida.

No entanto, deve-se assegurar que a representação reduzida não venha a descartar informação relevante para o registro como os N pares de pontos correspondentes, para qualquer tipo de geometria. A partir da referida análise conseguiram-se identificar dois requisitos desejáveis para que uma representação reduzida satisfaça o quesito existência.

1. independência à classe de superfície: a representação não deve ser limitada a um certo tipo de superfície para garantir a extração dos pares de correspondências.

2. adequabilidade na densidade: a redução deve reter um número mínimo de amostras e deve conter amostras espalhadas em todas as regiões da superfície e, não somente, concentradas em certos locais. Uma densidade adequada, portanto, possibilitará que haja amostras na região de sobreposição para garantir a extração de pares de correspondências.

Adicionalmente, para satisfazer o quesito de eficiência, identificaram-se quatro características que podem aumentar a capacidade discriminante de um descritor. Estas são:

- distinguibilidade geométrica das suas amostras.
- medidas invariantes associadas às suas amostras individualmente: informação geométrica local como curvatura.
- as medidas invariantes associadas à relação entre suas amostras: relação espacial entre as amostras, como distância.

Analisou-se que quanto mais características o descritor conseguisse expressar, maior seria a sua capacidade informativa e, portanto, discriminante.

Neste trabalho, propõe-se um novo descritor de configuração tridimensional que possui amostras distintas, medidas invariantes locais e relação espacial distinta entre suas amostras. Com estas características, acredita-se que este descritor possui uma maior capacidade discriminante que os apresentados anteriormente.

Ao mesmo tempo de garantir o quesito de existência, o algoritmo de pré-alinhamento deve também assegurar que a transformação T resultante seja correta e não um falso positivo, como acontece com alguns casos particulares de aplicação do algoritmo CPO e CHG (Seção 2.2). Conjetura-se, com base na análise dos métodos, que o processo de avaliação global deve possuir as seguintes duas características para conduzir a uma correta eliminação de falsos positivos:

1. deve-se avaliar exaustivamente todas as transformações candidatas.
2. deve-se avaliar os ajustes de todas as amostras da imagem.

Para conciliar corretude e eficiência, neste trabalho propõe-se um procedimento de eliminação de falsas positivas que possui as características 1. e 2. Pode-se avaliar todas as transformações candidatas as quais são em número reduzido em relação às abordagens baseadas em amostragem completa. E, adicionalmente, para a avaliação de erros pode-se considerar o conjunto reduzido de dados proposto, em representação (substituição) do conjunto denso de dados originais.

Técnica	Espaço de escolha do descritor-referência	Configuração dos descritores	Medidas invariantes dos descritores		Espaço de busca do descritor-alvo	Procedimento de Filtragem
			das amostras	da relação entre as amostras		
DBR	todas as amostras (A_1)	triangular (três amostras)	não possui	distâncias	todas as amostras (A_2)	Avaliação de ajuste de vizinhança e subamostragem de $\mathcal{R}_1$
CPO	todas as amostras (A_1)	$N > 3$ amostras sem estrutura definida	histogramas <i>spin</i> com posto por distâncias acumuladas	distâncias	todas as amostras (A_2)	Avaliação de ajuste de alguns pontos de A_1 .
CHB	malha associada a A_1	triangular (três vértices de um triângulo da malha)	histograma geométrico (distância e ângulo)	distâncias	malha associada a A_2	Avaliação de ajuste da malha associada a A_1 , porém só de algumas transformações candidatas.
CCB	pontos bitangentes em $\mathcal{R}_1$	quadrilateral (dada pelos pontos extremos de dois segmentos)	não possui	distância	pontos bitangentes em $\mathcal{R}_2$	Avaliação de ajuste de uma subamostragem de A_1
CCI	pontos de curvatura média igual a zero em A_1	linear (dada pelos pontos extremos de N segmentos)	não possui	distância	pontos de curvatura média igual a zero em A_2	Avaliação de ajuste de pontos especiais e de A_1
CGI	pontos máximos locais de gradiente de intensidade em I_1	triangular (dada pelos três vértices do triângulo da tetraedrização)	não possui	distância e ângulos	pontos máximos locais de gradiente de intensidade em I_2	Avaliação de ajuste de uma subamostragem de pontos especiais

Tab. 2.1: Tabela comparativa dos principais trabalhos presentes na literatura

Capítulo 3

Algoritmo de Busca

O pré-alinhamento automático de imagens de profundidade é um problema fundamental na área de Visão Computacional. Vários algoritmos têm sido propostos para este propósito, porém ainda sofrem de várias limitações principalmente relacionadas à garantia da existência de descritores para casamento e à eficiência na busca de descritores correspondentes como vimos no Capítulo 2. Neste capítulo apresenta-se um novo algoritmo de busca que contorna o problema de existência e alivia a complexidade combinatorial na busca de correspondências. Visto que a determinação de um par de descritores correspondentes na região de sobreposição conduz à obtenção da transformação responsável pelo pré-alinhamento, temos que o algoritmo de busca é crucial dentro deste processo.

Propõe-se substituir a imagem de profundidade por uma malha triangular a qual é caracterizada por não ser restrita a tipos especiais de geometria e ser dotada de topologia (relação espacial entre as amostras). A existência é assegurada ao construir um descritor em cada vértice da malha $\mathcal{M}_1$, que representa $\mathcal{R}_1$ e ao buscar exaustivamente seus descritores-alvo na malha $\mathcal{M}_2$, que representa a imagem $\mathcal{R}_2$. Uma redução significativa na complexidade combinatorial na busca de descritores correspondentes é obtida pela proposta de um descritor discriminante com configuração espacial e informações locais e globais associadas. Tal descritor é um triedro espacial que consiste de uma base composta por três vetores linearmente independentes quase ortogonais, cuja origem e pontos extremos dos eixos, são vértices da malha triangular. A esse triedro será associada informação local (dada pela curvatura média em cada vértice) e, também, informação global, fornecida pela posição relativa de três vértices “suficientemente afastados” do vértice na origem. O fato de se considerar uma região mais abrangente permite representar uma região com maiores aspectos da forma geométrica da superfície, ao invés de uma simples região local. Isto torna o descritor altamente discriminante diminuindo drasticamente o número de correspondências inválidas; e a busca exaustiva garante que, pelo menos, um par de triedros correspondentes seja sempre encontrado.

O restante deste capítulo está organizado da seguinte forma. A Seção 3.1 descreve como os triedros espaciais são gerados a partir da malha triangular $\mathcal{M}_1$. Mais adiante, na Seção 3.2 discute-se o

procedimento detalhado de casamento de pares de triedros. Na Seção 3.3 apresenta-se o procedimento de descarte de falsas transformações para obter a transformação de registro T . E na Seção 3.4 são avaliadas as condições para a existência de triedros correspondentes e é feita uma análise de tempo de execução do algoritmo. A Seção 3.5 apresenta as considerações finais deste capítulo.

3.1 Construção de um Descritor-Referência

A comparação das propriedades geométricas invariantes dos descritores definidos em cada imagem de profundidade permite encontrar os pares de correspondências entre imagens. Portanto, estas propriedades são de importância crucial para o estabelecimento correto e eficiente de correspondências. De fato, um descritor com propriedades geométricas invariantes pouco significantes resultará em múltiplos casamentos ambíguos e tornará computacionalmente caro o procedimento de identificação de um casamento correto (processo de filtragem). No Capítulo 2, pode-se analisar que um descritor será mais expressivo se conseguir incorporar maior número possível de propriedades geométricas locais e globais nas amostras que o constituem. Além do mais, acredita-se que se essa configuração for tridimensional e, representar uma “região suficientemente grande” da superfície, a capacidade discriminatória do descritor, poderá ser aumentada.

A junção de propriedades locais e globais em um único descritor tem sido uma importante contribuição na área de reconhecimento de objetos [47, 57]. O *shape context* é um descritor que descreve a relação de um dado ponto do objeto com todos os outros pontos do mesmo. Ou seja, é composto por um conjunto de vetores partindo de um ponto p até todos os outros pontos da imagem (Figura 3.1). Isto faz com que o descritor, através dos seus vetores, expresse a configuração da forma completa relativa ao ponto de referência p .

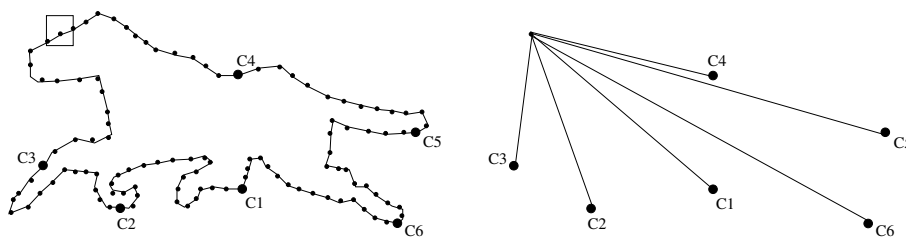


Fig. 3.1: *Shape context*

Essa idéia se aplica no contexto de reconhecimento do objetos, onde naturalmente os objetos em análise devem ser similares ao de referência. Já no caso de imagens de profundidade parcialmente sobrepostas, este conceito não pode ser aplicado diretamente.

No entanto, usando o *shape context* como inspiração, propõe-se um descritor que incorpore medidas locais e globais significativas de forma que contribuam com a sua capacidade discriminante. Tal

descritor é um triedro espacial composto por três vetores, linearmente independentes, que ligam um vértice $\mathcal{V}$ da malha, chamado de origem, a outros três vértices da mesma $v_i, i = 1, 3$ (Figura 3.2). As medidas invariantes que são incorporadas são:

- a curvatura média $H = (k_1 + k_2)/2$, a qual é obtida a partir das curvaturas principais, k_1 e k_2 , estimadas pelo algoritmo de Rusinkiewicz (Seção 2.1), e
- as magnitudes d_i dos vetores $\overrightarrow{v_i\mathcal{V}}$. Estas distâncias d_i devem estar no intervalo $[R_{min}, R_{max}]$, onde R_{min} e R_{max} representam os raios mínimo e máximo de esferas com centro em $\mathcal{V}$. Esses valores devem ser adequadamente escolhidos de forma a prover ao triedro uma configuração global, porém sem ultrapassar a região de sobreposição. Conjeturamos que assim o descritor terá maior capacidade para discriminar distintas feições na busca.

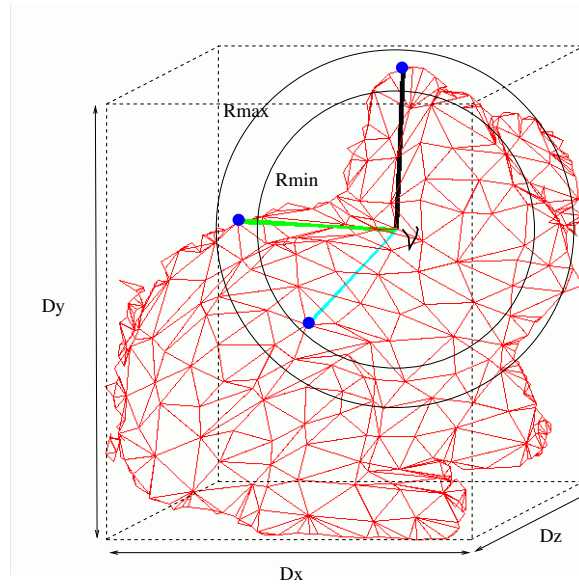


Fig. 3.2: Triedro Espacial.

A atribuição de valores para d deve considerar as limitações impostas por dois requisitos conflitantes: maior abrangência e, por outro lado, a sua pertinência à região de sobreposição pois um descritor excessivamente grande pode não ficar inteiramente contido dentro desta região. Idealmente, dever-se-ia utilizar um valor d_i maximamente grande e que caiba ainda na região de sobreposição. Como a região de sobreposição é desconhecida, pensou-se que d_i poderia depender de um valor conhecido relacionado com o tamanho do objeto, como a média M das dimensões da caixa limitante da malha que representa o objeto, onde $M = (D_x + D_y + D_z)/3$ com D_x denotando a largura, D_y a altura e D_z a profundidade da caixa limitante (Figura 3.2). Outra alternativa seria calcular as dimensões exatas do objeto alinhando-o com seus eixos principais. No entanto, por simplicidade, optou-se por usar M .

Assim, estimar o valor de d_i se reduz a procurar por uma fração f de M ($d_i = fM$). Para dar uma maior liberdade para o triedro se acomodar à geometria da superfície, uma flexibilidade maior pode ser dada fazendo $f_1M \leq d \leq f_2M$, onde $f_1 < f_2$, resultando em $f_1M = R_{min}$ e $f_2M = R_{max}$ da Figura 3.2. Os valores f_1 e f_2 podem ser facilmente fornecidos pelo usuário como a menor e a maior estimativa da porcentagem, em pixels efetivos, que a região de sobreposição representa em relação à imagem de profundidade.

É desejável evitar triedros com seus três eixos muito próximos, no sentido de formar ângulos muito pequenos como o da Figura 3.3, que não abrangem muito bem a região geométrica em volta do vértice origem $\mathcal{V}$. De fato, comparando a Figura 3.2 com a Figura 3.3, na primeira se tem um triedro que abrange uma região maior do que na segunda. A primeira configuração foi obtida ao impor que os três eixos sejam, maximamente, ortogonais entre si, ou seja, eles devem ser preferencialmente perpendiculares entre si.

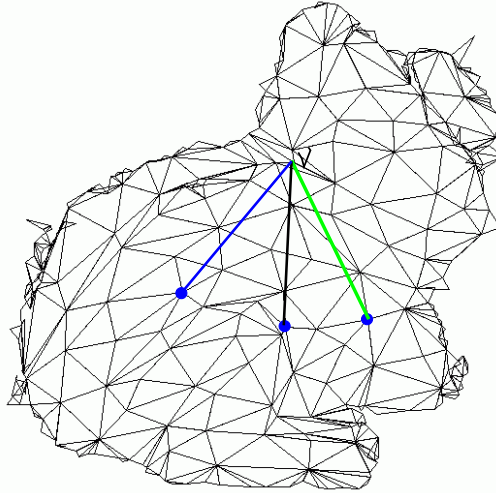


Fig. 3.3: Triedro Espacial.

Antes de prosseguir é necessário introduzir mais duas notações que serão usadas no decorrer desta dissertação. Chamar-se-ão de $\mathcal{M}_1$ e $\mathcal{M}_2$ às malhas densas construídas a partir de todas as amostras de $\mathcal{R}_1$ e $\mathcal{R}_2$, respectivamente, e de $\mathcal{MS}_1$ e $\mathcal{MS}_2$ às malhas obtidas da simplificação de $\mathcal{M}_1$ e $\mathcal{M}_2$, respectivamente.

Para construir um triedro com origem no vértice $\mathcal{V}$ da malha $\mathcal{M}_1$ e vértices extremos v_i , com as características mencionadas deve-se seguir o seguinte procedimento: O vértice v_1 pode ser qualquer vértice que pertença à região radial $\mathcal{R}$, limitada pelas esferas de raio mínimo $R_{min} = f_1M$ e o raio máximo $R_{max} = f_2M$. E o vértice v_2 deve ser escolhido entre os contidos na região $\mathcal{R}$, como aquele que gera o vetor $\overrightarrow{v_2\mathcal{V}}$ que forma o ângulo mais próximo a noventa graus com o vetor $\overrightarrow{v_1\mathcal{V}}$. E o vértice

v_3 , além de estar em $\mathcal{R}$, deve ser o vértice mais afastado do plano gerado pelos vetores $\overrightarrow{v_1\mathcal{V}}$ e $\overrightarrow{v_2\mathcal{V}}$. Os eixos assim definidos podem não ser rigorosamente ortogonais, mas são uma boa opção para satisfazer a propriedade almejada de abrangência máxima.

Na Figura 3.4 mostram-se triedros espaciais gerados para a imagem FROG180 (Figura 4.7), considerando diferentes tamanhos de sobreposição entre duas imagens. O triedro gerado por $f_1 = 0.1$ e $f_2 = 0.2$ é construído quando se presume que a região de sobreposição representa entre 10% e 20% da região total da imagem (Figura 3.4(a)). Usa-se $f_1 = 0.3$ e $f_2 = 0.4$ quando estimasse que tal porcentagem está entre 30% e 40% (Figura 3.4(b)). E, finalmente, quando a porcentagem estiver entre 50% e 70%, usa-se $f_1 = 0.5$ e $f_2 = 0.7$ (Figura 3.4(c)).

Na Seção 3.4 é apresentada uma análise do quesito de existência considerando os valores f_1 e f_2 sugeridos.

3.2 Casamento de triedros

Nesta seção será descrito o procedimento de casamento de triedros, que inclui a escolha de descritores-referência na malha $\mathcal{M}_1$ e busca de descritores-alvo na malha $\mathcal{M}_2$.

3.2.1 Procedimento de Escolha de um Descritor-Referência

Para garantir que pelo menos um descritor-referência seja determinado na região de sobreposição ter-se-ia que considerar todas as combinações $C_n^{m_1}$ de descritores. Porém essa opção é ineficiente. Neste trabalho propõe-se uma alternativa exaustiva mais eficiente: a construção de apenas um triedro em cada vértice da malha $\mathcal{M}_1$. O uso de um único triedro, ao invés de todas as possíveis combinações do vértice em relação a outro, tem se mostrado suficiente para se obter resultados satisfatórios. Acredita-se que o triedro definido na Seção 3.1 atinge uma região espacial suficientemente grande em volta de cada vértice, tornando-se um bom representante dessa região, mesmo que não a cubra. O procedimento de escolha de descritores-referência é então baseado nos seguintes dois estágios:

- Obtenção da malha $\mathcal{M}_1$ conforme explicado na Seção 1.1.
- Construção de um triedro espacial em cada vértice da malha $\mathcal{M}_1$.

3.2.2 Procedimento de Busca por um Descritor-Alvo

Propõe-se substituir a busca das três amostras constituintes do triângulo usado no algoritmo DAR-CES (Seção 2.2.1) pela busca de quatro amostras que descrevem o descritor apresentado na Seção 3.1.

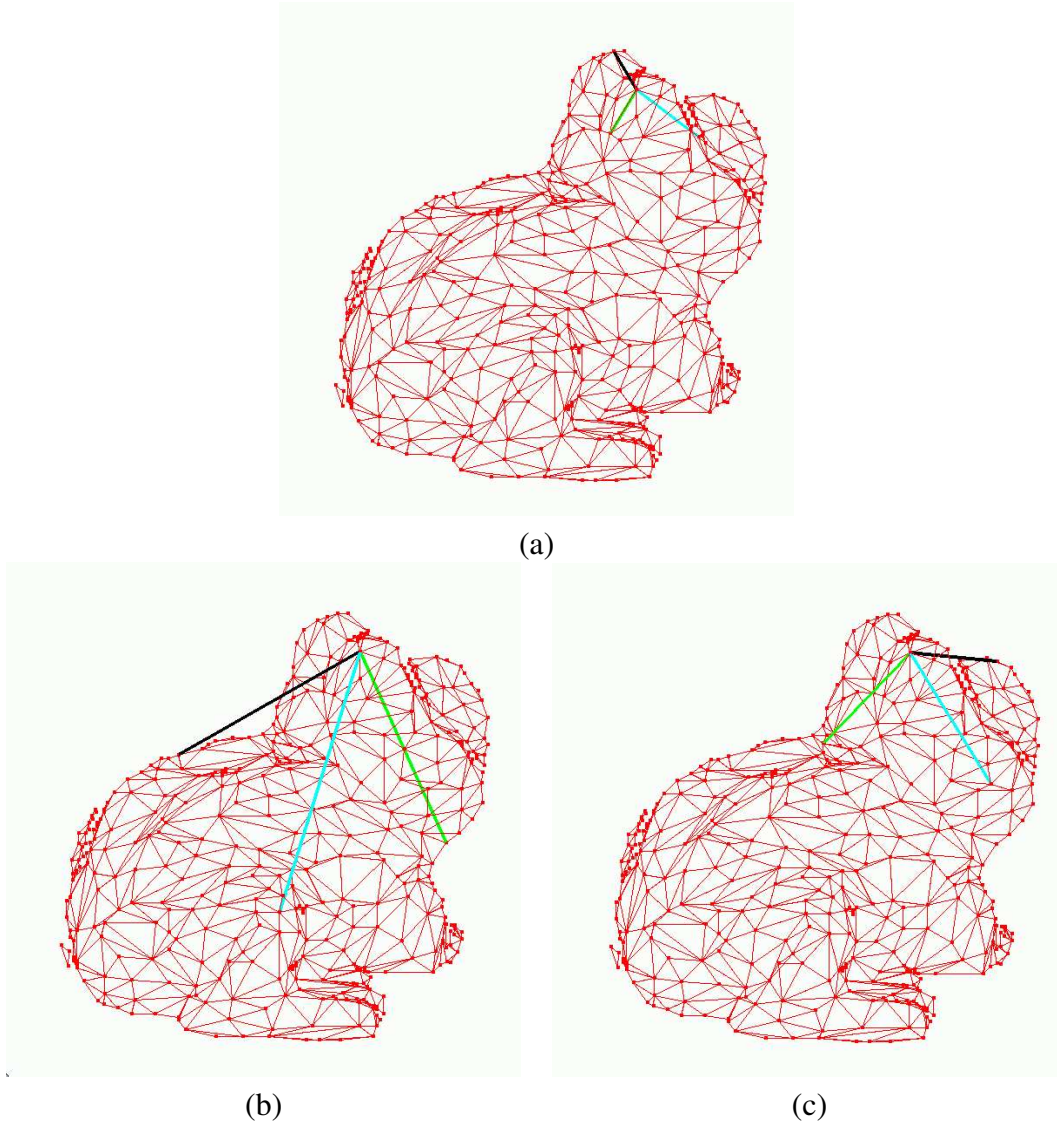


Fig. 3.4: (a) $f_1 = 0.1, f_2 = 0.2$ (b) $f_1 = 0.5, f_2 = 0.7$ (c) $f_1 = 0.3, f_2 = 0.4$

O princípio básico da busca é percorrer vértices na malha $\mathcal{M}_2$ que sejam candidatos para se corresponderem com o vértice origem $\mathcal{V}$ do triedro da malha $\mathcal{M}_1$, logo então, deve-se procurar os vértices correspondentes aos restantes três vértices do triedro. Segue-se o procedimento detalhado.

Busca da Origem

Seja $\mathcal{T}_k$, um triedro definido no vértice $\mathcal{V}$ da malha $\mathcal{M}_1$, com vértices extremos $v_i, i = 1, 2, 3$. O primeiro estágio restringe-se à procura de candidatos $\mathcal{C}_i$ para origem do descritor-alvo em $\mathcal{M}_2$. Estes devem ter curvatura média H_2 aproximadamente igual que a curvatura média H_1 de $\mathcal{V}$. Considera-se uma tolerância Δ a qual é dependente de s (tamanho do pixel, para ser independente da escala da ima-

gem de profundidade). Esse primeiro casamento reduz significativamente, o número de candidatos como se pode visualizar na Figura 3.5. Foi aplicado o procedimento em três pares de malhas (a) bird, (b) teletubby e (c) dino. Os pontos em preto na malha $\mathcal{M}_2$ são os vértices que foram identificados como semelhantes em curvatura média com o vértice origem do triedro $\mathcal{T}$ na malha $\mathcal{M}_1$.

Busca das Três Amostras Adicionais

Os potenciais correspondentes c_{1l} e c_{2l} de v_1 e v_2 , respectivamente, são obtidos na mesma forma como no algoritmo DARCES (Seção 2.2.1), com uma restrição adicional: possuir similar curvatura média. A região de busca dos pontos c_{1l} é reduzida à superfície da esfera com raio $\|\mathcal{V} - v_1\|$ e centro em $\mathcal{C}_l$ e a região de busca dos pontos c_{2l} é restrita à interseção das esferas de raios $\|\mathcal{V} - v_2\|$ e $\|v_1 - v_2\|$, com centros em $\mathcal{C}_l$ e c_{1l} , respectivamente, conforme ilustrado na Figura 2.4. Uma tolerância $\delta = s$ dos raios das esferas de procura é considerada. Finalmente, uma busca dos candidatos c_{3l} que correspondem a v_3 , poderia ser restrita à interseção de três esferas, porém uma região mais reduzida é obtida ao se seguir o seguinte procedimento: calcula-se a distância ponto-plano d de v_3 até o plano P , formado pelos vetores $\overrightarrow{\mathcal{V}v_1}$ e $\overrightarrow{\mathcal{V}v_2}$, e calcula-se também as coordenadas baricêntricas, da projeção ortogonal x_1 do vértice v_3 sobre o plano P em relação ao triângulo $\mathcal{V}v_1v_2$. Logo, os pontos candidatos c_{3l} estarão em um raio $\delta = s$ em volta do ponto p , sendo $p = x_2 + n_2d$, onde x_2 é a posição correspondente a x_1 no triângulo $\mathcal{C}_lc_{1l}c_{2l}$ e n_2 é o vetor normal do plano gerado pelos vetores $\overrightarrow{\mathcal{C}_lc_{1l}}$ e $\overrightarrow{\mathcal{C}_lc_{2l}}$ (Figura 3.6). Resumindo, o procedimento de busca do descritor-alvo é realizado através dos seguintes passos:

- Obtenção da malha triangular $\mathcal{M}_2$.
- Comparação da curvatura média H para filtrar os candidatos a vértice origem do descritor-alvo.
- Aplicação do DARCES estendido para 3D para encontrar descritores-alvo que se aproximem do descritor-referência tanto em curvatura como em distâncias.

3.3 Remoção de Correspondências Falsas

O casamento dos triedros reduz drasticamente o número de possíveis correspondências, porém ainda se tem um conjunto considerável de remanescentes. Considerando que as duas malhas só diferem por uma transformação rígida, a transformação candidata $T_k(R_k, t_k)$ pode ser determinada a partir de um par de triedros correspondentes $\mathcal{T}_1$ e $\mathcal{T}_2$ constituídos pelos vértices $(\mathcal{V}, v_1, v_2, v_3)$ e $(\mathcal{C}, c_1, c_2, c_3)$ respectivamente.

A transformação que alinha ambos triedros é calculada da seguinte forma simples

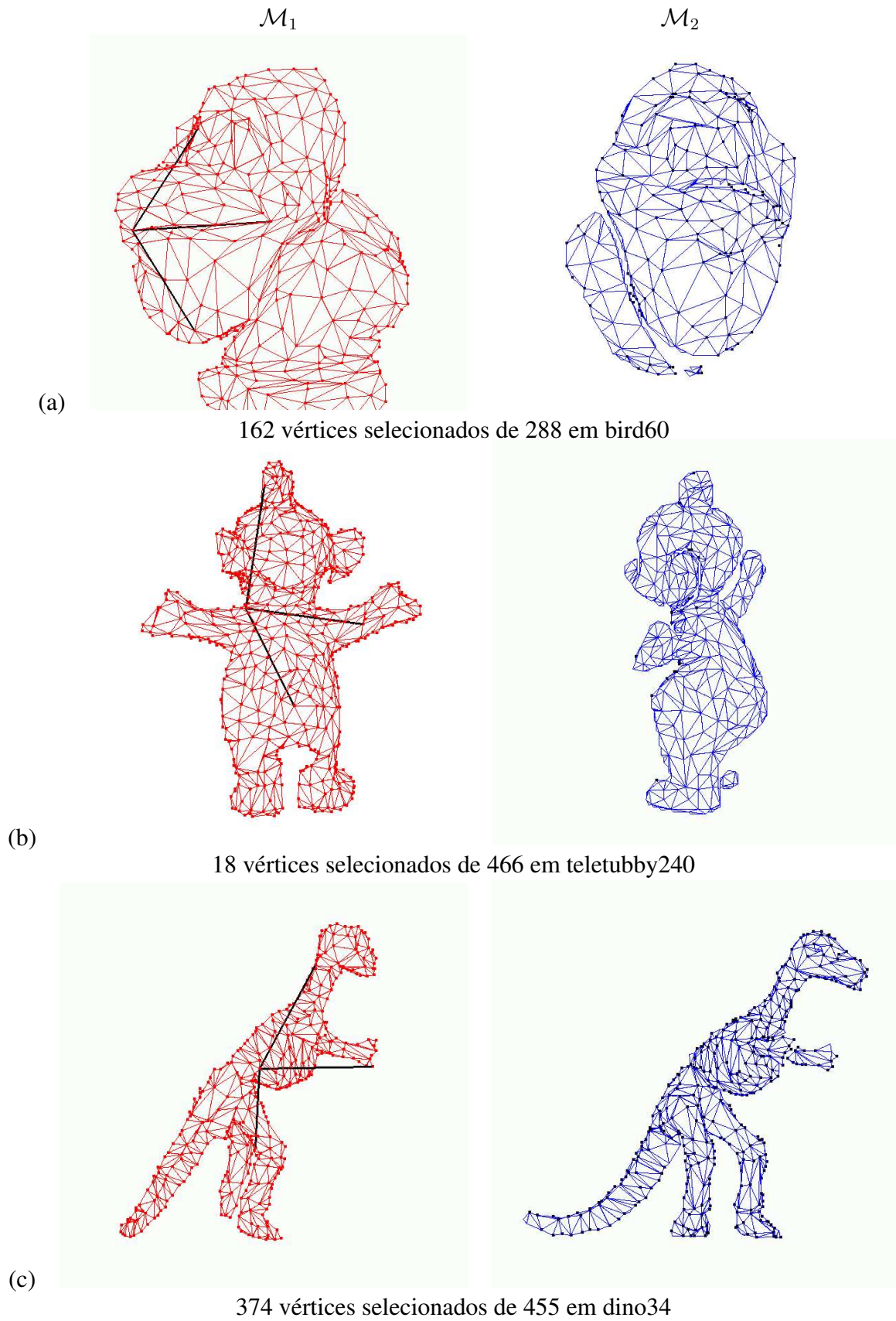


Fig. 3.5: Filtragem por curvatura média

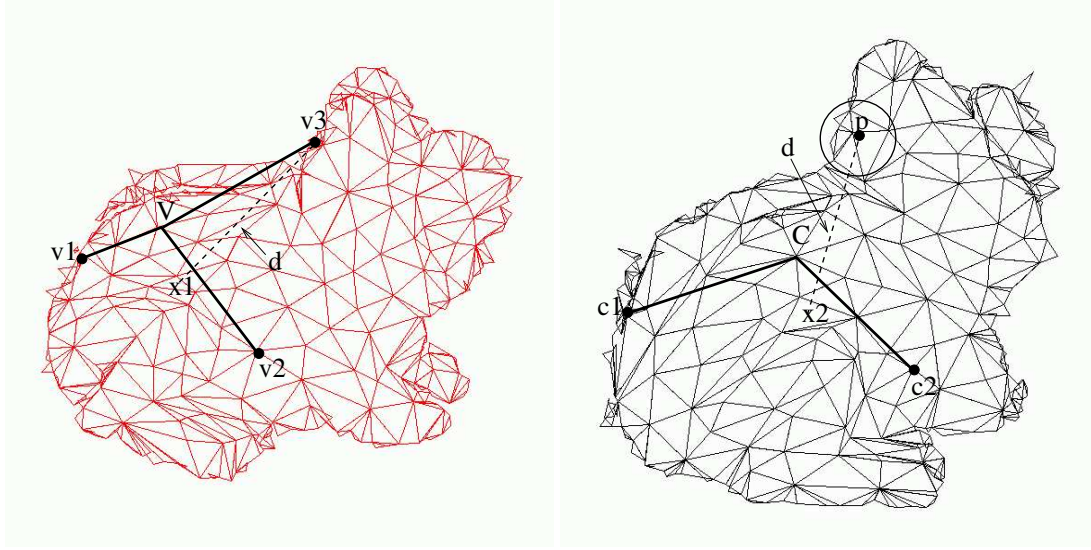


Fig. 3.6: Obtenção do quarto ponto correspondente.

$$T_k = \begin{bmatrix} \mathcal{C}_x & c_{1,x} & c_{2,x} & c_{3,x} \\ \mathcal{C}_y & c_{1,y} & c_{2,y} & c_{3,y} \\ \mathcal{C}_z & c_{1,z} & c_{2,z} & c_{3,z} \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathcal{V}_x & v_{1,x} & v_{2,x} & v_{3,x} \\ \mathcal{V}_y & v_{1,y} & v_{2,y} & v_{3,y} \\ \mathcal{V}_z & v_{1,z} & v_{2,z} & v_{3,z} \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}. \quad (3.1)$$

Cada transformação candidata T_k determinada passa por uma avaliação de ajuste global através da contagem de vértices na malha transformada $\mathcal{M}_1$, que alinharam suficientemente bem com vértices na malha $\mathcal{M}_2$. Ressalta-se aqui que a avaliação de ajuste se dá, de fato, entre os vértices de $\mathcal{M}_1$ e as faces em $\mathcal{M}_2$, já que as amostras em duas malhas não são exatamente idênticas. São computadas as distâncias entre os vértices de $\mathcal{M}_1$ e suas faces mais próximas em $\mathcal{M}_2$ e contabilizada a quantidade de vértices que tiverem distâncias menores que um limiar pré-especificado. Mais especificamente o procedimento segue os seguintes passos:

- a malha $\mathcal{M}_1$ é transformada com uso de T_k .
- mede-se a distância entre o baricentro de cada face $\mathcal{F}$ de $\mathcal{M}_1$ transformada e a face em $\mathcal{M}_2$ mais próxima a $\mathcal{F}$.
- Contabiliza-se o número *cont* de vértices cuja distância for menor que um limiar pré-determinado.

Finalmente, dentre todas as transformações candidatas, escolhe-se aquela que teve o maior número *cont*, baseado no fato que a transformação que registra a região de sobreposição de ambas imagens deve ter a maior quantidade de vértices alinhados. No apêndice B encontra-se o pseudo-código desta etapa de pré-alinhamento.

3.4 Análise

Nesta seção é feita análise, em termos práticos, para avaliar o desempenho do algoritmo de busca proposto.

3.4.1 Existência de um par de triedros correspondentes

Foi aplicado o algoritmo proposto sobre pares de imagens oriundas de conjuntos de imagens encontradas nos repositórios [35, 51, 50] para validar os valores sugeridos para f_1 e f_2 como sendo a estimativa de menor e maior porcentagem da região de sobreposição em relação à imagem original. Para ilustração foi escolhido o conjunto de seis imagens de profundidade visualizado na Figura 1.2 tomadas do site [50]: BUNNY0, BUNNY45, BUNNY90, BUNNY180, BUNNY270 e BUNNY315 e testou-se o desempenho do algoritmo nos seguintes pares de imagens:

- BUNNY0-BUNNY90,
- BUNNY90-BUNNY180,
- BUNNY180-BUNNY270,
- BUNNY270-BUNNY0,
- BUNNY270-BUNNY315,
- BUNNY315-BUNNY0.

As seis imagens foram obtidas rotacionando o sensor em torno do eixo y . Para cada par que diferencia por um ângulo de 45° (exemplo BUNNY270-BUNNY315), usou-se $f_1 = 0.5$ e $f_2 = 0.7$, pois percebeu-se que a menor e a maior estimativa da porcentagem, em pixels efetivos, de sobreposição em relação à imagem de profundidade é 50% e 70%, respectivamente. Por mesmo motivo, para cada par que tem diferença angular em torno de 90° (exemplo BUNNY0-BUNNY90), usou-se $f_1 = 0.0$ e $f_2 = 0.3$. A Tabela 3.1 mostra comparativamente o ângulo de rotação real, realizado no sensor, e o ângulo de rotação obtido pela proposta de pré-alinhamento. Na coluna Par de imagens são especificados os pares de imagens usadas, e os parâmetros da transformação de registro real¹ são especificados nas colunas R_y (R) (Rotação em torno ao eixo y), x (R), y (R) e z (R). Estas três últimas colunas correspondem aos três valores do vetor translação². Os parâmetros da transformação obtidos pelo algoritmo de pré-alinhamento apresentado são, por sua vez, registrados nas colunas R_y (M),

¹Esses parâmetros foram obtidos a partir do ângulo fornecido pelo próprio nome da imagem, por exemplo para BUNNY0-90 se tem $R_y = 90^\circ$ [50].

²Estes deslocamentos foram obtidos com uso do aplicativo *Scanalyze*, após realizar manualmente a rotação R_y correspondente

x (M), y (M) e z (M). A qualidade da transformação do pré-alinhamento foi avaliada pela diferença de ambos valores (R e M) que deu uma diferença angular, em média, de 1.61° e uma diferença de deslocamento, em média, de 0.01.

Par de Imagens	R_y (R)	x (R)	y (R)	z (R)	R_y (M)	x (M)	y (M)	z (M)
0-90	90	-0.0001	0.0002	-0.0001	90.0834	-0.0015	0.0000	-0.0025
90-180	90	-0.0001	0.0001	-0.0001	87.9930	-0.0010	-0.0015	0.0006
180-270	90	-0.0000	0.0003	0.0000	88.2712	-0.0014	0.0021	-0.0022
270-0	90	-0.0001	0.0000	0.0004	91.9029	-0.0022	-0.0012	0.0013
270-315	45	0.0137	-0.0003	0.0045	43.7342	-0.0100	-0.0012	0.0040
315-0	45	-0.0065	-0.0000	-0.0128	42.3506	-0.0040	-0.0001	-0.0060

Tab. 3.1: Parâmetros da transformação entre pares de imagens do modelo BUNNY

Observe que mesmo com uma pequena área de sobreposição em pares 0-90, 90-180, 180-270 e 270-0, o algoritmo foi capaz de identificar pares de correspondências que conduziram a um registro bem sucedido com uso de valores f_1 e f_2 propostos.

3.4.2 Análise da resolução da malha

Para validar a conjectura de que malhas simplificadas, que retêm a essência da forma geométrica, são suficientes para registrar duas imagens, serão apresentados nesta seção resultados coletados do pré-alinhamento do par de imagens BUNNY0-BUNNY90 para cinco níveis de resolução das malhas. O algoritmo de redução de amostras ilustrado aqui será detalhado no Capítulo 4. A Tabela 3.2 apresenta o número de vértices da malha $\mathcal{M}_1$ na coluna **Vértices em $\mathcal{M}_1$** e o número de vértices da malha $\mathcal{M}_2$ na coluna **Vértices em $\mathcal{M}_2$** e os parâmetros da transformação obtida são mostrados nas colunas R_y e x , y , z . A qualidade das transformações do pré-alinhamento foi avaliada pela diferença dos valores R e M (da Tabela 3.1), tanto na Rotação, apresentados na coluna ΔR , como no deslocamento, apresentados na coluna ΔD . Foram também medidos os tempos de execução do procedimento em um computador Pentium IV 3.2 GHz com 512 MB de RAM (plataforma Linux), para avaliar o compromisso entre a qualidade de pré-alinhamento e o tempo de processamento. Este tempo (em segundos) é apresentado na coluna **Tempo**. O gráfico da Figura 3.7 mostra a diminuição drástica do tempo de processamento com o aumento do nível de decimação. Foi utilizado o pacote QSLIM para decimar as malhas e o nível de decimação corresponde ao nível considerado neste pacote [21]. Observe-se que os valores ΔR e ΔD variam em um intervalo de aproximadamente um grau, isto é, os valores dos parâmetros da transformação são similares a despeito do número de amostras em diminuição. Isto sugere que uma redução dos dados seria adequada para garantir uma maior eficiência no algoritmo sem sacrificar a precisão nos resultados do mesmo. Cabe mencionar que os valores ΔR e ΔD não aumentam à medida que a resolução da malha diminui, pois o alinhamento

depende do limiar estabelecido para o casamento dos descritores. Observe-se que no nível 4 se tem um erro menor do que no nível 1. Este fato será explorado no Capítulo 4 para propor uma nova forma de redução do espaço de busca.

Nível	Vértices em $\mathcal{M}_1$	Vértices em $\mathcal{M}_2$	R_y	x	y	z	ΔR	ΔD	Tempo
1	3314	4152	89.8519	-0.0019	0.0004	-0.0019	0.15	0.00	67992
2	1487	1829	90.0639	-0.0006	0.0003	0.0000	0.06	0.00	13612
3	959	1135	91.1277	-0.0035	-0.0000	-0.0049	1.13	0.01	1628
4	688	855	90.0834	-0.0015	0.0000	-0.0025	0.08	0.00	831
5	426	553	91.5342	-0.0007	0.0016	-0.0019	1.53	0.00	138
6	299	395	90.2239	-0.0001	0.0002	-0.0001	0.22	0.00	45

Tab. 3.2: Número de Vértices vs. Tempo de Processamento

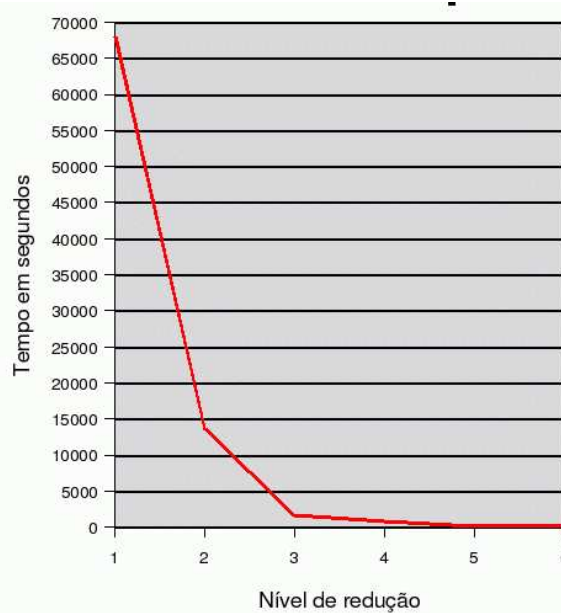


Fig. 3.7: Nível de Redução vs. Tempo de Processamento.

3.5 Considerações Finais

Neste capítulo apresenta-se o algoritmo de busca baseado na construção de triedros espaciais com a finalidade de serem usados no pré-alinhamento de duas imagens parcialmente sobrepostas. Entre as características relevantes do triedro destacam-se:

- são de simples construção: uma vez que a malha é obtida, a construção é feita pela escolha de vértices a uma distância d de um vértice fixo arbitrário, de forma que resulte em vetores com direções quase ortogonais.
- podem ser construídos em superfícies de objetos com arbitrárias formas: qualquer superfície permite sua representação por uma malha triangular.
- possuem medidas invariantes locais: são usadas as curvaturas médias para comparação.
- possuem medidas invariantes globais: são usadas as relações métricas entre as amostras.
- possuem configuração espacial: o formato tridimensional agrega aos valores escalares informação espacial das amostras selecionadas para pré-alinhamento.

A junção do uso de uma malha triangular para representar a imagem de profundidade, a definição de triedros na malha $\mathcal{M}_1$ e propriamente o algoritmo de busca na malha $\mathcal{M}_2$ dá origem a um algoritmo de pré-alinhamento com as seguintes características:

- Eficiente filtragem de falsos candidatos (Seção 3.2.1), sem comprometer o resultado da busca.
- Simplicidade na avaliação de erros para eficiente filtragem de falsos candidatos (Seção 3.3) sem comprometer o resultado da busca.

Enquanto o quesito existência é assegurado sob ajuste adequado dos parâmetros f_1 e f_2 ainda deve-se assegurar o quesito eficiência. Acredita-se que isso será conseguido com a redução de volume de dados que é proposta no próximo capítulo, com base nos resultados apresentados na Tabela 3.2.

Capítulo 4

Redução do Espaço de Busca

Os dados adquiridos a partir de um *hardware* 3D de aquisição moderno, como escaneadores de profundidade podem resultar em milhões de amostras ou mais, fazendo com que a construção de procedimentos eficientes para o processamento de dados de tão grande tamanho seja uma tarefa desafiadora.

De fato, o ponto crucial para aumentar a eficiência, nas diferentes etapas do registro, é a redução do espaços de escolha dos descritores em $\mathcal{R}_1$ e o espaço de busca de descritores correspondentes em $\mathcal{R}_2$. Essa redução torna-se possível ao diminuir o volume de dados das imagens de profundidade, substituindo-se o conjunto denso das suas amostras por um conjunto menor que consiga capturar a essência da forma da superfície original.

Uma forma de reduzir o volume de dados é limitar o pré-alinhamento com base nas características especiais. Este paradigma é conciso e independente do ponto de vista, porém apresenta algumas limitações como : (1) restrição a classes particulares de superfícies, e (2) concentração de pontos em certas regiões da superfície e, ainda, ausência dos mesmos em outras regiões.

Visto que qualquer superfície pode ser aproximada por uma malha triangular, sugere-se utilizar uma “malha triangular simplificada” para reduzir o volume de dados da imagem de profundidade sem deixar de manter a conectividade entre as amostras que é necessária para avaliar erros (Seção 3.3). Os resultados mostram que com esta representação pode-se conseguir uma redução do volume de dados de até 96%, o que garante o quesito eficiência, e, sob ajuste adequado de parâmetros de controle, pode-se ter uma densidade adequada para preservar o quesito existência.

O restante deste capítulo está organizado da seguinte forma. Na Seção 4.1 descreve-se a proposta de redução de volume de dados, através do uso de malha triangular, apresentando uma primeira vantagem: a irrestrrição à geometria da superfície. A seguir, na Seção 4.2, descreve-se o algoritmo de simplificação QSLIM, proposto por Garland e Heckbert [21], que é capaz de produzir de forma rápida, aproximações de alta qualidade onde a forma geral da superfície é preservada. Na mesma seção discorre-se sobre algumas características que fazem o método QSLIM conflitante com os requisitos

apresentados na Seção 2.4. Na Seção 4.3 apresentam-se as adaptações que foram introduzidas para resolver os pontos conflitantes do algoritmo QSLIM. Desta forma, podem-se obter as características de densidade adequada de vértices. Com base na malha simplificada proposta, na Seção 4.4 apresenta-se o algoritmo de pré-alinhamento via malhas simplificadas. Mais adiante, na Seção 4.5 mostra-se a capacidade do algoritmo proposto de satisfazer os quesitos eficiência e existência, ainda após uma drástica simplificação. A Seção 4.6 apresenta as considerações finais do capítulo.

4.1 Redução por Decimação da Malha

Uma imagem de profundidade descreve parcialmente a superfície de um objeto através de N amostras $P = \{p_1, p_2, \dots, p_N\}$. Este conjunto de amostras pode ser usado para construir uma malha triangular M que é uma aproximação à superfície original, sendo $M = \{t_1, t_2, \dots, t_N\}$ onde t_i é uma face triangular da malha, com três pontos pertencentes a P . Um algoritmo para construção desta malha foi explicado na Seção 1.1, porém os atuais escaneadores de profundidade produzem milhões de amostras (N muito grande), o que resulta na geração de malhas muito densas.

No contexto deste trabalho, conjectura-se que a malha não precisa usar a amostragem completa, pois a obtenção de uma transformação rígida que realiza um pré-alinhamento de duas imagens de profundidade não necessita de modelos excessivamente detalhados. Modelos que preservem as características salientes da forma da superfície são suficientes. Para tanto, a proposta é usar uma malha triangular simplificada que com um número reduzido de amostras consiga preservar as principais características globais da superfície.

Além do ganho em eficiência com sua representação, a malha triangular simplificada apresenta uma vantagem adicional: não impõe limitações a tipos de superfícies que podem ser processadas. Isto é, qualquer superfície pode ser aproximada por uma malha triangular, sobrepujando a representação por pontos característicos a qual pode limitar seu escopo a classes particulares de superfícies. Mais uma vantagem da malha simplificada que foi aproveitada neste trabalho é a preservação da conectividade entre as amostras do conjunto reduzido. De fato, para aplicar o algoritmo de Rusinkiewicz, descrito na Seção 2.1, para extrair as curvaturas, precisa-se o conhecimento dos pontos vizinhos. A conectividade é essencial também para avaliar erros do pré-alinhamento (Seção 3.3).

Segundo alguns autores, uma malha triangular simplificada é uma malha triangular que consegue reter as partes mais salientes da geometria da malha original, onde partes salientes referem-se a um pequeno número de partes que capturam a forma do objeto [36, 37]. Outros autores têm tentado especificar melhor este conceito: uma malha simplificada preserva os pontos distintos da superfície do objeto, onde estes pontos podem ser entendidos como as características que fazem com que o objeto seja facilmente identificável. Podem ser pontos ou traços (curvas) sobre a superfície do objeto que capturam a essência ou abstração de alto nível da sua forma [40]. A Figura 4.1 mostra dois

objetos e seu correspondente conjunto de pontos distintos.

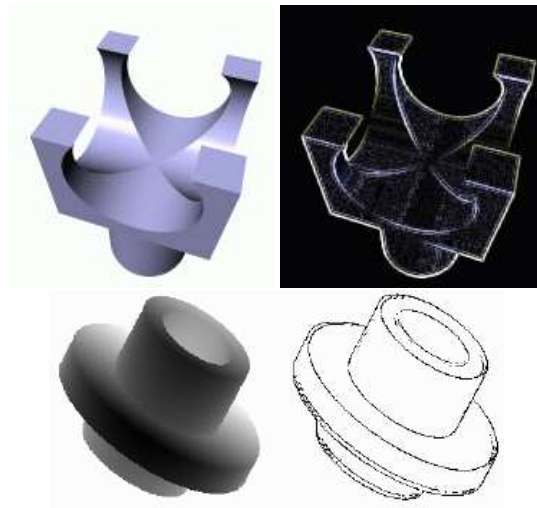


Fig. 4.1: (a) Superfície do objeto (b) Pontos distintos

Com base nessas definições, elaborou-se o seguinte conceito equivalente: a simplificação de uma malha triangular consiste na remoção de detalhes desnecessários para identificá-la prontamente, sem ambiguidade.

Frente a uma grande variedade de algoritmos de decimação disponíveis [36, 37] optou-se por utilizar um deles como base para validar a proposta deste trabalho. Em continuidade, descreve-se o método de simplificação QSLIM [22, 21], o qual, a partir de uma malha triangular, realiza a decimação das suas arestas, iterativamente, de forma que resulta-se um novo nível de simplificação. Este método se destaca como um método que preserva as características salientes da superfície, em cada iteração.

4.2 QSLIM: Um Algoritmo de Decimação

No trabalho [22] demonstra-se, através de exaustivos experimentos e análise comparativa com outros métodos de simplificação, que o algoritmo de simplificação QSLIM fornece um bom compromisso entre eficiência e qualidade da aproximação resultante, no sentido de conseguir reter as características mais distinguíveis do objeto. Na Figura 4.2(a) pode-se observar uma malha densa representando o rosto de Chopin [39] e à direita (Figura 4.2(b)) a malha simplificada em 95%. Note-se como o nariz, o queixo e outras características mais notáveis ainda são mantidos, mesmo, após a drástica simplificação.

Na Figura 4.3 também pode-se apreciar uma malha densa e sua correspondente simplificada. Apesar da porcentagem de simplificação ser grande, maior que 90%, as bordas de descontinuidade e bordas de orientação da superfície ainda são preservadas.

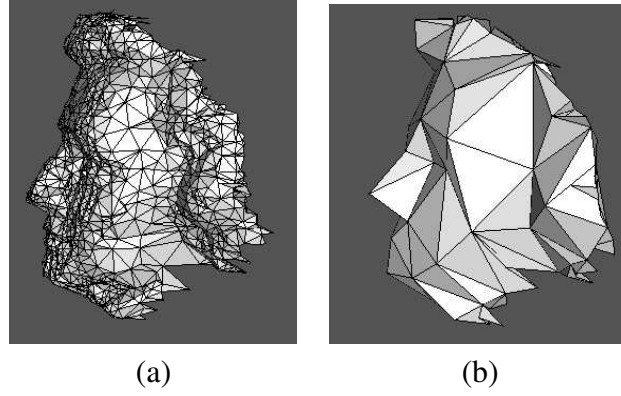


Fig. 4.2: Rosto de Chopin

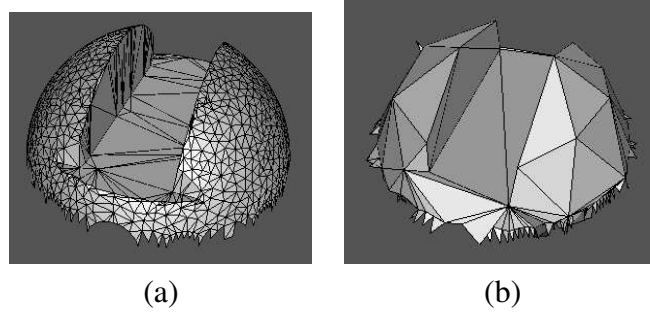


Fig. 4.3: Imagem Esfcub

O algoritmo de simplificação QSLIM [22, 21] usa o princípio de decimação por contração de arestas. Este princípio estabelece a eliminação de uma aresta $v_i v_j$ e substituição dela por um ponto $\bar{v}$, conforme ilustrado pela Figura 4.4. QSLIM busca preservar maximamente a geometria da superfície, pela minimização dos erros introduzidos na simplificação. Isto é, em cada iteração do algoritmo, uma aresta é substituída pelo ponto que introduz o menor erro em relação à malha da iteração anterior. O erro quádrico $E(v)$ é utilizado para avaliar as discrepâncias.

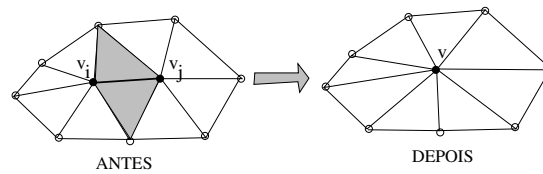


Fig. 4.4: Contração de aresta

O erro quádrico E associado a um vértice v é definido como a soma das distâncias quadradas do vértice a todas suas faces adjacentes

$$E(v) = \sum_k D_k^2(v),$$

onde a distância quadrada de v a uma face adjacente k , representada por $\langle n_k, v \rangle = d$, pode ser expressa por

$$\begin{aligned} D_k^2(v) &= (n_k^t v + d)^2 = (n_k^t v + d)(n_k^t v + d) \\ &= v^t (n_k n_k^t) v + 2d n_k^t v + d^2, \end{aligned}$$

sendo n_k^t o vetor normal ao plano (face k) e d é um escalar constante. Esta distância pode também ser escrita em forma quadrática

$$Q(v) = v^t A v + 2b^t v + c = D^2(v).$$

Fazendo $A = n_k n_k^t$, $b = d n_k$, $c = d^2$, Q , então, pode ser definido como o triplo $Q = (A, b, c)$ onde A é uma matriz 3×3 , b , um vetor de 3 escalares, e c , um escalar. Por propriedade de quádricas, a adição de quádricas é realizada componente a componente, assim: $Q_1(v) + Q_2(v) = (Q_1 + Q_2)(v)$ onde $(Q_1 + Q_2) = (A_1 + A_2, b_1 + b_2, c_1 + c_2)$.

Com isso, a soma das distâncias quadradas de um vértice v até suas faces adjacentes é igual a uma quádrica Q

$$E(v) = \sum_k D_k^2(v) = \sum_k Q_k(v) = Q(v). \quad (4.1)$$

Para determinar o ponto $\bar{v}$ que introduz o menor erro na contração de uma aresta $v_i v_j$, precisa-se modelar o erro da aresta e minimizá-lo. Pela propriedade de soma de quádricas, o erro quádrico da aresta $v_i v_j$ é dado pela função quadrática:

$$Q_1(\bar{v}) + Q_2(\bar{v}) = Q(\bar{v})$$

onde Q_1 e Q_2 são erros quádricos associados aos vértices v_i e v_j , respectivamente. A minimização desta função levará a se obter o vértice $\bar{v}$ que substitui a aresta $v_i v_j$. Já que $Q(\bar{v})$ é quadrático, encontrar seu mínimo é um problema linear; o mínimo ocorre quando $\partial Q / \partial x = \partial Q / \partial y = \partial Q / \partial z = 0$. Calculando as derivadas parciais de $Q = \bar{v}^t A \bar{v} + 2b^t \bar{v} + c$, tem-se o gradiente de Q

$$\Delta Q(\bar{v}) = 2A\bar{v} + 2b$$

Resolvendo para $\Delta Q(v) = 0$ encontra-se:

$$\bar{v} = -A^{-1}b. \quad (4.2)$$

Assim, $\bar{v}$ é o vértice que introduz o menor erro o que representa a contração ótima, para a aresta $v_i v_j$.

Para determinar, dentre todas as arestas da malha triangular, a aresta cuja eliminação introduzirá o menor erro em relação à malha triangular da iteração anterior, deve-se (1) calcular a contração ótima $\bar{v}$ de cada aresta e seu erro $E(\bar{v})$, e (2) escolher para contração, a aresta que apresenta o menor erro E . O processo é repetido, até que uma determinada quantidade de triângulos seja alcançada.

4.3 Adaptações

Da descrição do método e dos exemplos, pode-se verificar que o processo de simplificação QSLIM cumpre o objetivo de um método de simplificação conforme apresentado na Seção 4.1: remover detalhes menores da imagem e preservar as características mais proeminentes. No entanto deve-se lembrar que o processo de redução deve ainda preservar pelo menos um par de triedros correspondentes para garantir o registro. Portanto, as seguintes questões precisam ser respondidas e tratadas para que QSLIM seja efetivo para o propósito deste trabalho:

1. Qual é o nível aceitável de perda de detalhes? Como limitar o nível de simplificação de modo que a região de sobreposição não seja excessivamente decimada e possa se assegurar o quesito densidade adequada mencionado na Seção 2.4?
2. Os vértices que amostram a região de sobreposição A_{S1} não coincidem com os vértices que amostram a região de sobreposição A_{S2} . Como estabelecer um limiar para o espaço de busca destas correspondências?

A resposta da primeira questão será dada na Subseção 4.3.1 e a da segunda questão, na Subseção 4.3.2.

4.3.1 Controle na Densidade Amostral

No intuito de se obter uma redução do espaço de busca, que vise o aumento da eficiência, podem-se gerar malhas excessivamente simplificadas, de modo que se perca o mínimo de amostras necessárias para a correspondência. Para evitar isso precisa-se determinar um teto máximo para o nível de simplificação, de forma que uma densidade aceitável de vértices na malha simplificada seja garantida.

Para quantificar o nível de simplificação máximo requerido teve-se que modificar o parâmetro usado pelo algoritmo QSLIM (Seção 4.2) o qual não era adequado para os objetivos deste trabalho. QSLIM usa um critério de número mínimo de faces como critério de parada da simplificação. Como era preciso ter duas malhas simplificadas, no mesmo nível de simplificação, para substituir as imagens de profundidade, essa medida não era adequada. De fato, precisa-se de um nível de simplificação relacionado com fidelidade geométrica, ou seja, duas malhas com mesmo nível de perda ou preservação geométricas. O critério de número de faces não se relaciona com a fidelidade geométrica, pois

uma malha triangular composta por mil triângulos, representando uma região planar, fornece uma representação geométrica tão fiel quanto uma malha formada por apenas dois triângulos.

Para medir o grau de fidelidade da malha simplificada em relação à malha densa original foi introduzido um erro ϵ . Este erro é definido como a média das distâncias ponto-plano de todas as faces em relação a suas amostras associadas. As amostras associadas da face F são aquelas que correspondem aos pixels ocupados por F , como mostrado pela Figura 4.5.

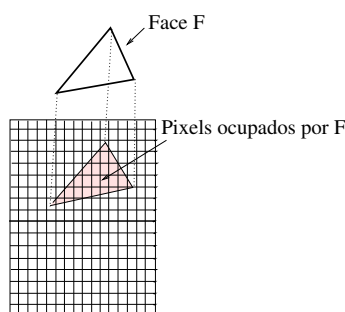


Fig. 4.5: Face da malha e pixels ocupados na imagem

Com este erro, o critério de parada adotado originalmente em QSLIM foi substituído pelo erro ser menor ou igual a ϵ .

Para definir o teto máximo T foram usados diversos conjuntos de imagens do repositório [51], as quais, na média, possuem 22000 amostras. Verificou-se que para um valor de $\epsilon = 0.2 * s$, onde s é o espaçamento horizontal e vertical entre as amostras¹, todos os pares obtiveram um pré-alinhamento que fez o algoritmo ICP convergir à solução correta². Para ilustração apresentam-se nas Figuras 4.6 e 4.7 dois conjuntos do mencionado repositório: PITT-PLANE e PITBULL, respectivamente. E na Figura 4.8 mostram-se o primeiro par de imagens de cada conjunto após sua simplificação no nível determinado por $\epsilon = 0.2 * s$.

¹O valor s é introduzido na fórmula para fazer ϵ independente de escala

²A solução correta do registro de pares de imagens deste repositório não está disponível, no entanto para efeitos de comparação assumiu-se que o registro correto é aquele que resulta de um pré-alinhamento manual mais um registro fino, ambos realizados com uso do software Scanalyze [48].

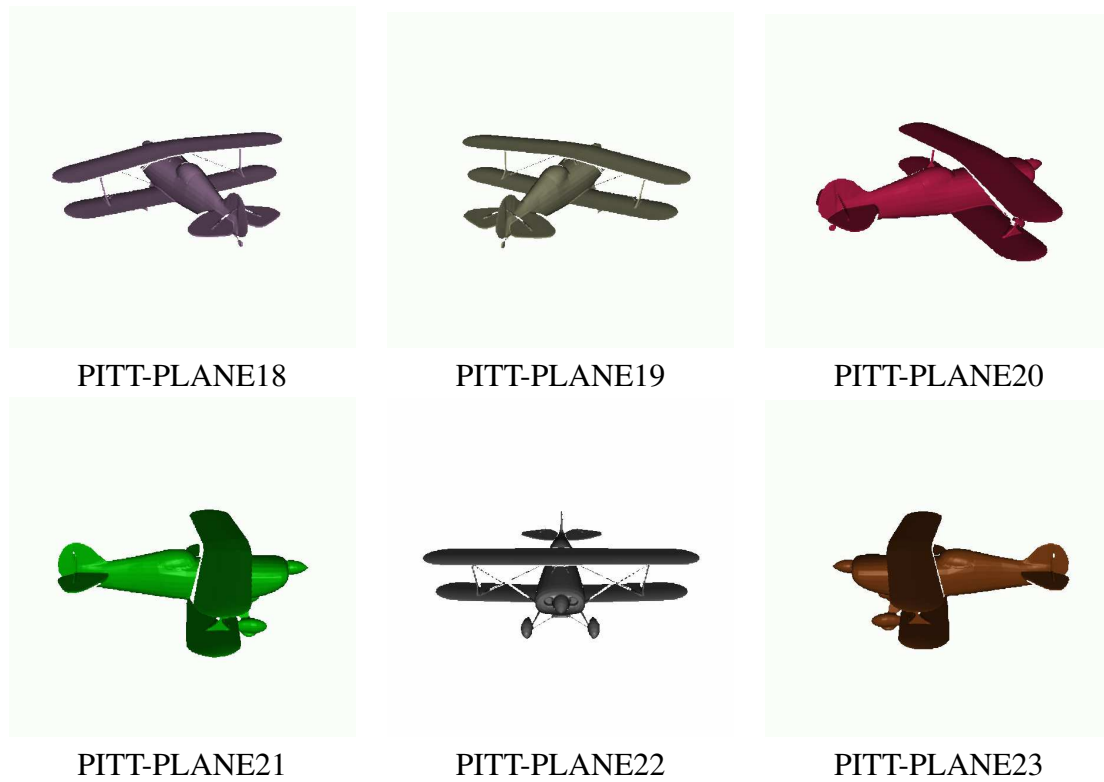


Fig. 4.6: Imagens de Teste: PITT-PLANE [51]

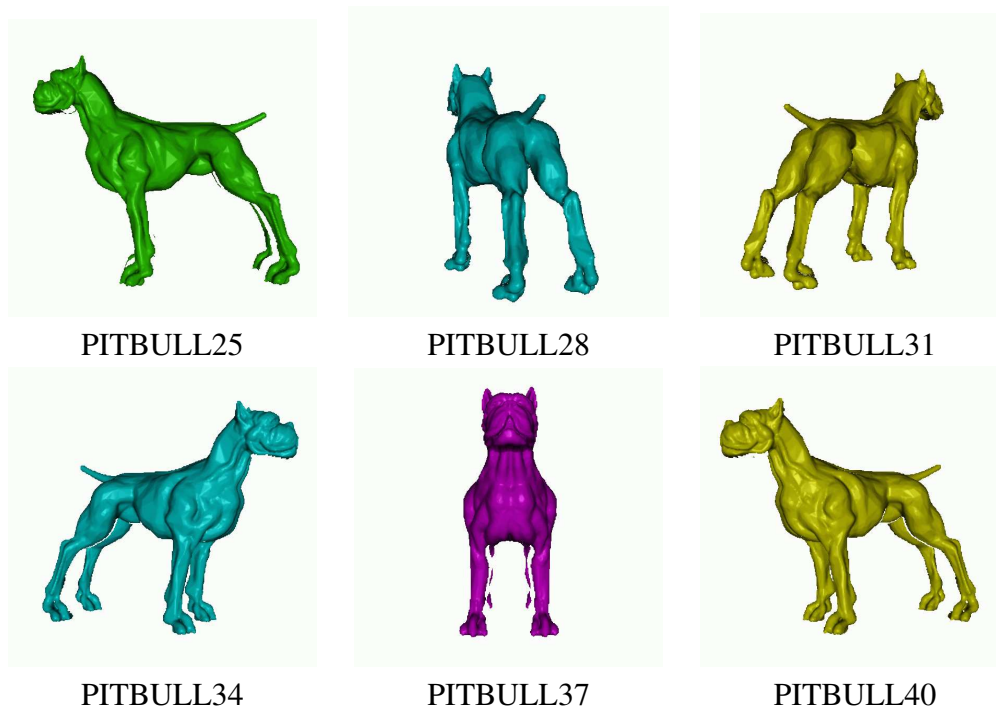


Fig. 4.7: Imagens de Teste: PITBULL [51]

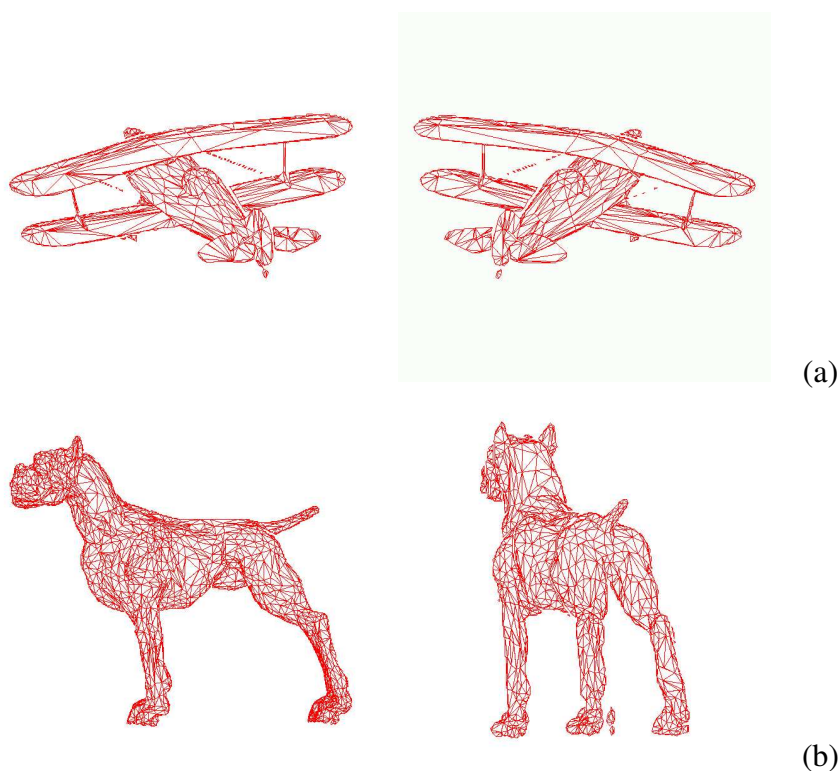


Fig. 4.8: Malhas Simplificadas com nível $\epsilon = 0.2 * s$ (a) PITT-PLANE, (b) PITBULL

Observa-se que as malhas simplificadas dos pares de imagens apresentam diferentes densidades embora tenha sido considerado o mesmo nível de simplificação ϵ . Isto quer dizer que a quantidade de vértices retidos depende do tipo de geometria que localmente a superfície apresenta:

- em regiões suaves, um número relativamente pequeno de vértices é preservado (menor densidade)
- em regiões menos suaves, um número maior de vértices são encontrados (maior densidade).

Na verdade, pode-se pensar em QSLIM como um decimador orientado a faces (regiões que são quase planares de acordo com um erro), pois substitui esse grupo por apenas uma única face representante.

A tabela 4.1 exibe estes resultados mostrando que no caso de PITT-PLANE foram preservados, na média, 44,8% das amostras, enquanto que no caso de PITBULL foram retidos, na média, 91,8%. Este resultado é esperado em razão do princípio de substituição de uma aresta por um ponto apenas quando a inclusão deste introduz o menor erro dentre todas as possibilidades. Isto significa que os pontos preservados são aqueles cuja exclusão introduzirá um grande erro em relação à geometria original da malha. Estes pontos são então pontos localmente não planares com características proeminentes em relação à sua vizinhança. Já os pontos eliminados são planares ou “quase planares”. Como PITT-PLANE contém muitas amostras em áreas localmente planas (por exemplo nas asas), consegue-se eliminar mais amostras para o mesmo nível de erro.

Imagens	Amostras (N)	Vértices (N)	% Preservação
PITT-PLANE 18	21261	1005	47,3
PITT-PLANE 19	21278	858	40,3
PITBULL25	23363	2133	90,9
PITBULL28	20088	1491	92,6

Tab. 4.1: Redução do volume de dados PITT-PLANE e PITBULL.

Por um lado, a característica de preservação dos pontos mais salientes faz com que o método retenha as características essenciais da sua forma, o qual é vantajoso para o propósito deste trabalho. Por outro lado, a definição do limiar $\epsilon = 0.2 * s$ permite que as regiões suaves não sejam excessivamente decimadas. No entanto, este limiar pode resultar em excessiva amostragem das regiões com considerável curvatura. De fato, estas regiões podem ser mais simplificadas que as regiões suaves. Em frente disso, propõe-se o uso de um índice de suavidade $\%_{suav}$ da geometria da superfície o qual possa introduzir uma certa ponderação para o limiar $\epsilon = \frac{0.2*s}{\%_{suav}}$. Esta ponderação faz com que a densidade de vértices diminua em função da suavidade da malha.

O índice de suavidade de uma imagem de profundidade pode ser facilmente definido estimando a curvatura mínima e máxima dos vértice de uma malha fina com uso do algoritmo de Rusinkiewicz [45]. Seguindo a formulação descrita na seção 2.1 acerca de classificação de pontos categorizamos os pontos planares (p), parabólicos (pa), elípticos (e) e hiperbólicos (h). Assim sendo, definiu-se o índice de suavidade $\%_{suav}$ da seguinte maneira:

$$\%_{suav} = \frac{p + pa + e}{p + pa + e + h}.$$

As Figuras 4.9 e 4.10 (segunda linha) mostram a classificação dos vértices dos pares de imagens PITT-PLANE e PITBULL, respectivamente. Na cor amarela, preta, azul e verde apresentam-se, respectivamente, os vértices parabólicos (pa), planares (p), elípticos (e) e hiperbólicos (h) e o número total de cada uma das categorias é apresentado na Tabela 4.2. Os vértices em vermelho são vértices da borda da malha os quais não foram usados na classificação por terem suas vizinhanças incompletas.

Com base no índice calculado espera-se que o índice de simplificação para PITBULL e PITT-PLANE seja aproximadamente o mesmo. A Tabela 4.3 apresenta os novos valores obtidos após a simplificação baseada na redefinição de ϵ .

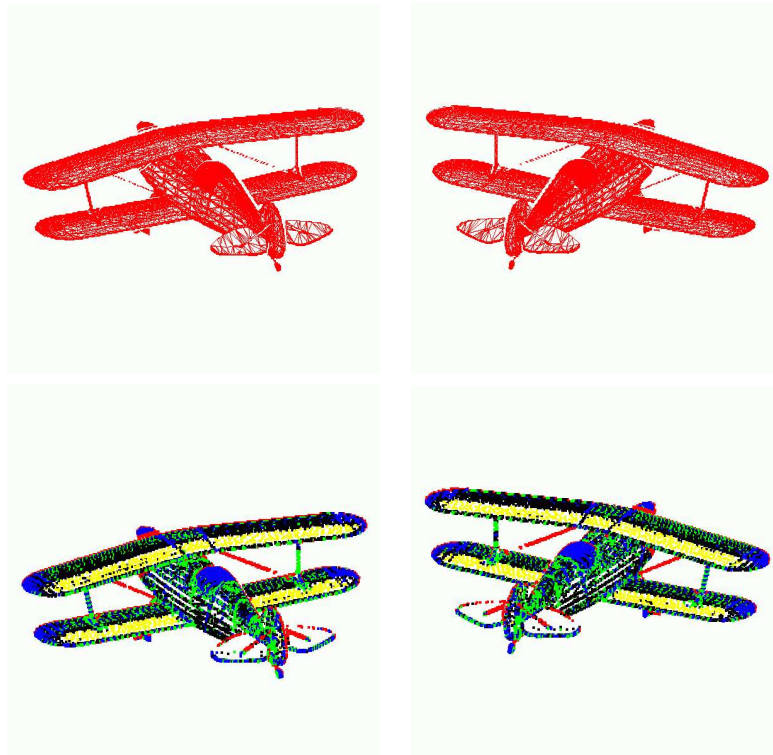


Fig. 4.9: Classificação de vértices Imagem PITT-PLANE

Imagens	Planares (p)	Parabólicos (pa)	Elípticos (e)	Hiperbólicos (h)	%Suav
PITT-PLANE 18	694	3976	2225	2116	0.765
PITT-PLANE 19	635	3956	2283	2144	0.762
PITBULL25	3	792	4335	5184	0.497
PITBULL28	1	698	3430	4250	0.493

Tab. 4.2: Classificação de vértices de PITT-PLANE e PITBULL.

Imagens	Amostras ($\hat{N}$)	Vértices (N)	% Preservação
PITT-PLANE 18	21261	506	97,6
PITT-PLANE 19	21278	468	97,8
PITBULL25	23363	995	95,7
PITBULL28	20088	756	96,2

Tab. 4.3: Redução do volume de dados PITT-PLANE e PITBULL.

Para validar a nova formulação de ϵ foram usadas imagens de outros repositórios e percebeu-se uma outra situação a ser tratada. Considerou-se o conjunto FROG e o conjunto BUNNY, obtidos dos

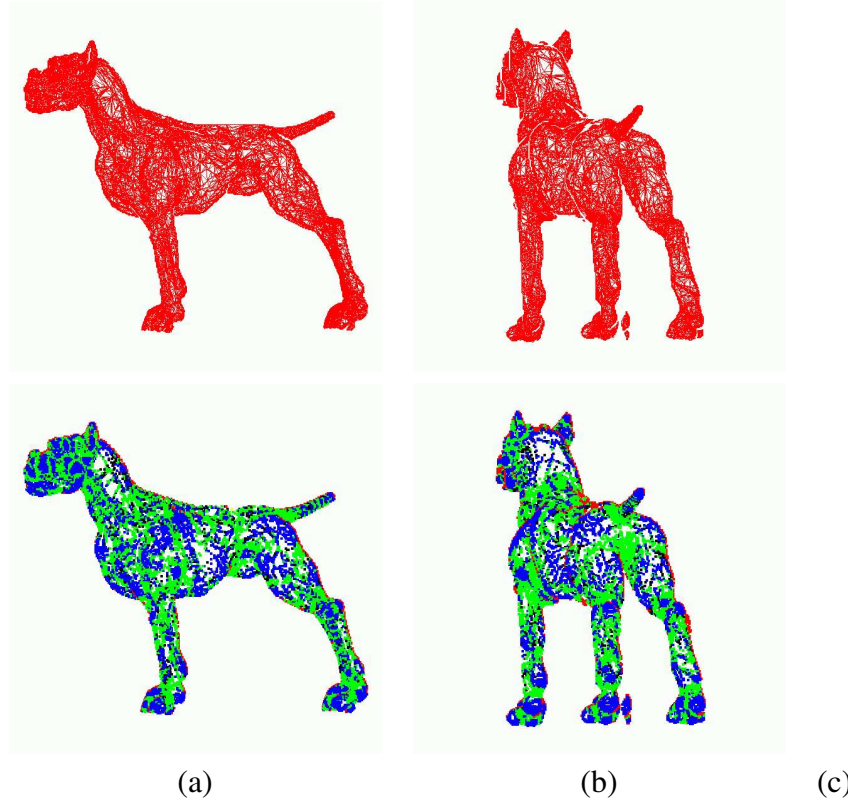


Fig. 4.10: Classificação de vértices Imagem PITBULL

repositórios [35, 50] e apresentados nas Figuras 4.11 e 1.2. Foi notada uma diferença muito grande no número de amostras por área nos dois conjuntos, por exemplo enquanto que BUNNY0 apresenta um $s = 0.7$ e uma área de 77.88, FROG0 possui um $s = 0.48$ e uma área de 29.31. Para estabelecer um erro justo deve-se então levar em consideração uma estimativa da área da imagem, além do espaçamento s . Fazendo isso, estar-se-á relacionando convenientemente as seguintes situações: maior frequência de amostragem com menor erro de reconstrução. Considera-se como frequência de amostragem s/A onde $A = \max(D_x, D_y)$. Os valores D_x e D_y obtidos na Seção 3.1, correspondem ao maior valor horizontal e vertical da imagem, respectivamente.

Assim sendo, já que o valor de ϵ foi baseado nos testes realizados com as imagens do repositório [51], as quais possuem na média uma área de 70, considerou-se uma redefinição de ϵ levando em consideração a proporção da frequência de amostragem à de valor 70. Consequentemente, tem-se

$$\epsilon = \frac{0.2 * s * A}{70 * \%suav}$$

As Figuras 4.12 e 4.13 mostram o primeiro par de imagens de cada conjunto após sua simplificação no nível determinado pelo valor definido para ϵ .

Na Tabela 4.4 apresenta-se a porcentagem de redução alcançada quando as imagens das Figu-

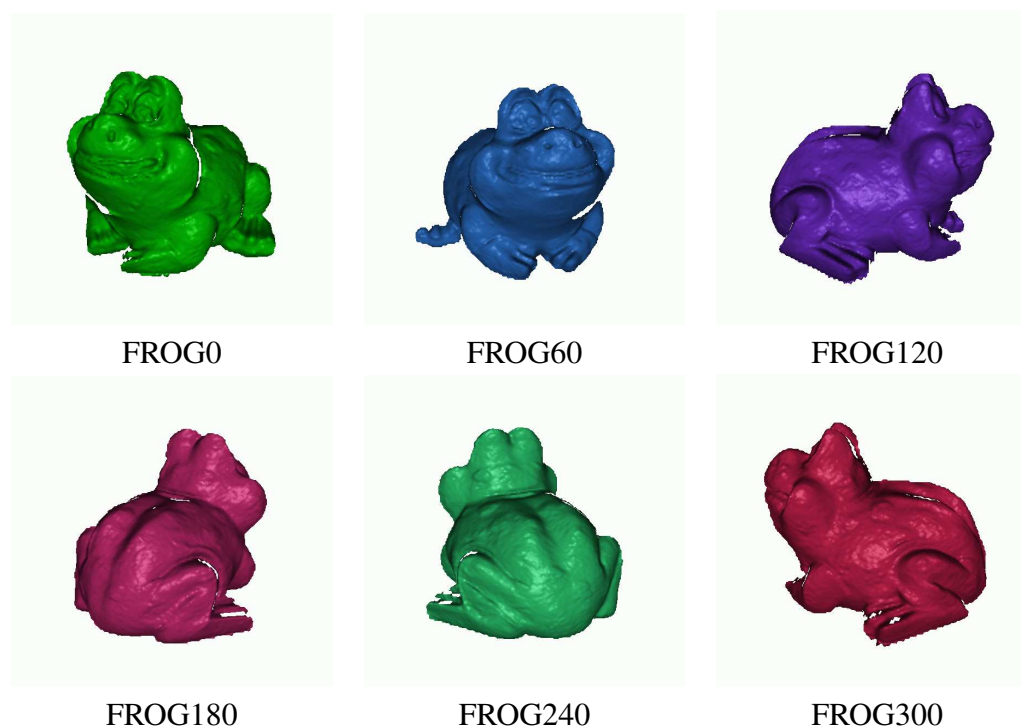


Fig. 4.11: Imagens de Teste: FROG [35]

ras 4.6, 4.7, 4.11 e 1.2 são reduzidas de acordo com o limiar ϵ definido. A redução do volume de dados foi na média de 96%, em relação à amostragem original. Com as malhas assim reduzidas pode-se obter bons resultados como será apresentado nas próximas seções.

4.3.2 Controle no Limiar para Espaço de Busca

Quando duas vistas do mesmo objeto são obtidas a partir do sensor, este não captura superfícies exatamente iguais. Como já foi visto, haverá regiões de sobreposição e também haverá regiões não comuns (dependendo da grau de variação da direção de vista). Nas regiões de sobreposição, as amostras dificilmente ocupam exatamente a mesma posição espacial. Nesta seção propõe-se introduzir um parâmetro de controle para contornar tais discrepâncias durante o estágio de busca de um descritor-alvo.

Essas discrepâncias podem ser claramente visualizadas nas Figuras 4.14 e 4.15, onde os pares de malhas simplificadas apresentadas nas Figuras 4.12 e 4.13 têm sido colocadas no mesmo referencial³. A explicação das discrepâncias pode ser por dois aspectos: a natureza de aquisição das imagens de profundidade e o princípio de decimação do método de simplificação.

- processo de aquisição das imagens de profundidade: a imagem de profundidade já incorpora

³As transformações usadas foram obtidas por pré-alinhamento manual mais um registro fino, ambos realizados com uso do software Scanalyze [48], exceto as transformações dos pares de conjunto BUNNY as quais são conhecidas

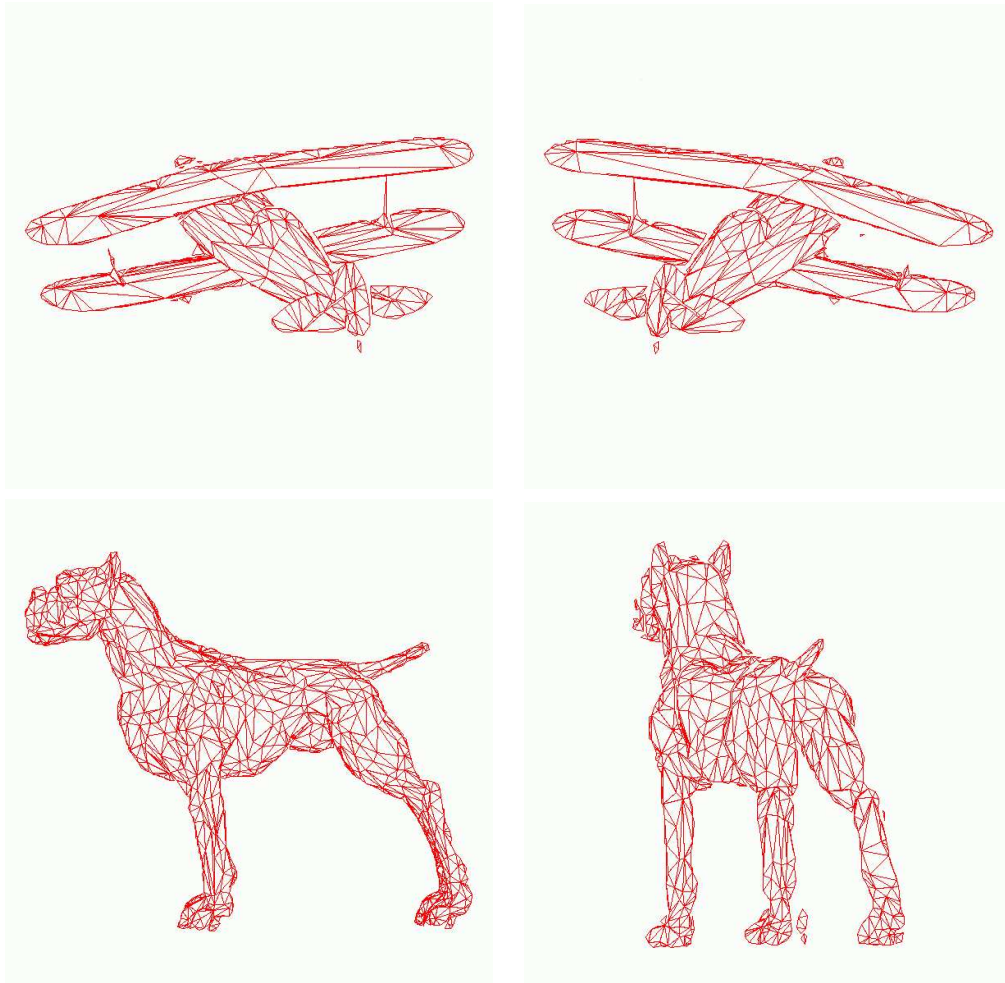


Fig. 4.12: Malhas Simplificadas com nível ϵ (a) PITT-PLANE, (b) PITBULL

discrepâncias próprias do processo de aquisição ao ponto que imagens de profundidade de uma mesma cena adquiridas em dois momentos diferentes dará como resultado duas imagens também diferentes.

- comportamento de QSLIM nas regiões “quase-planares”: quando os vértices de uma aresta e tem faces incidentes coplanares ou quase coplanares, a matriz A pode se tornar singular. Logo, a inversa não existe e não se pode encontrar $\bar{v}$ de forma única, usando a equação 4.2. Nesta situação o conjunto de pontos, para os quais $Q(v)$ é mínimo, formam ou uma reta ou um plano, o que traz muitas possibilidades para a seleção da posição $\bar{v}$.

Em razão dos erros de aquisição e da própria estratégia de decimação, os conjuntos de vértices obtidos pela simplificação de malhas parcialmente sobrepostas podem não coincidir; haverá uma disparidade entre os “vértices correspondentes”, a qual deve ser levada em consideração quando for realizado o procedimento de busca de correspondentes.

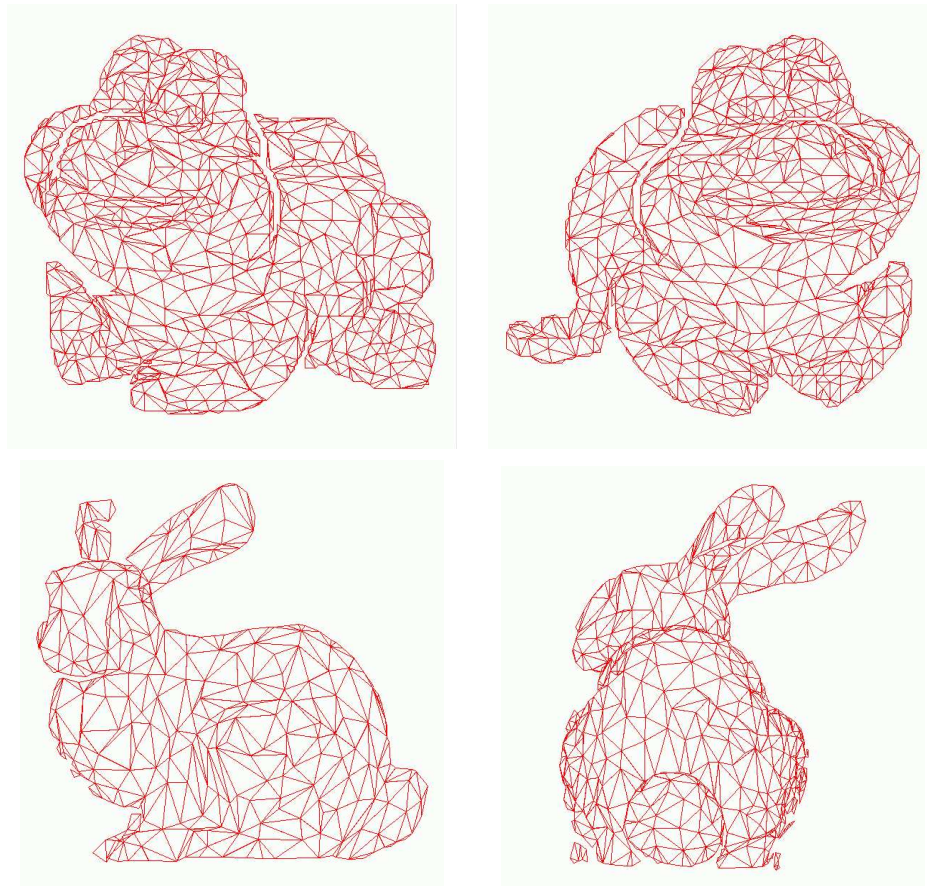


Fig. 4.13: Malhas Simplificadas com nível ϵ (a) FROG, (b) BUNNY

A solução proposta neste trabalho foi estabelecer uma tolerância de busca $\delta = 4 * s * A/70$, a qual foi adequada para os pares de imagens testados dos três repositórios [35, 51, 50]. Para ilustração, usaram-se um par de imagens dos conjuntos de imagens teste apresentados nas Figuras 4.6, 4.7, 4.11 e 1.2 para verificar a adequabilidade dos valores ϵ e δ propostos. Para isso colocaram-se cada par de malhas no mesmo referencial ⁴, como mostrado nas Figuras 4.14 e 4.15 e calculou-se o percentual de vértices na região de sobreposição da malha $\mathcal{M}_1$ que estariam a uma distância δ de vértices da malha $\mathcal{M}_2$ em relação ao número total de vértices na região de sobreposição. A Tabela 4.5 contém na primeira coluna o nome do par de imagens, na segunda coluna o número de vértices de $\mathcal{M}_1$, e na terceira a porcentagem da simplificação obtida em relação ao conjunto completo de amostras. Na quarta coluna apresenta-se o número de vértices encontrados na região de sobreposição A_{S1} (VS) e na quinta coluna, o número dos vértices (VC) que estão a uma distância δ de um vértice da malha $\mathcal{M}_2$. E, finalmente, na última coluna, é apresentado o percentual de vértices que obtiveram um correspondente em relação ao número total de vértices na região de sobreposição ($\%VC/VS$). Note-se que no mínimo

⁴As transformações usadas foram obtidas por pré-alinhamento manual mais um registro fino, ambos realizados com uso do aplicativo *Scanalyze*. Como esta análise foi feita a-posteriori, esta suposição faz sentido

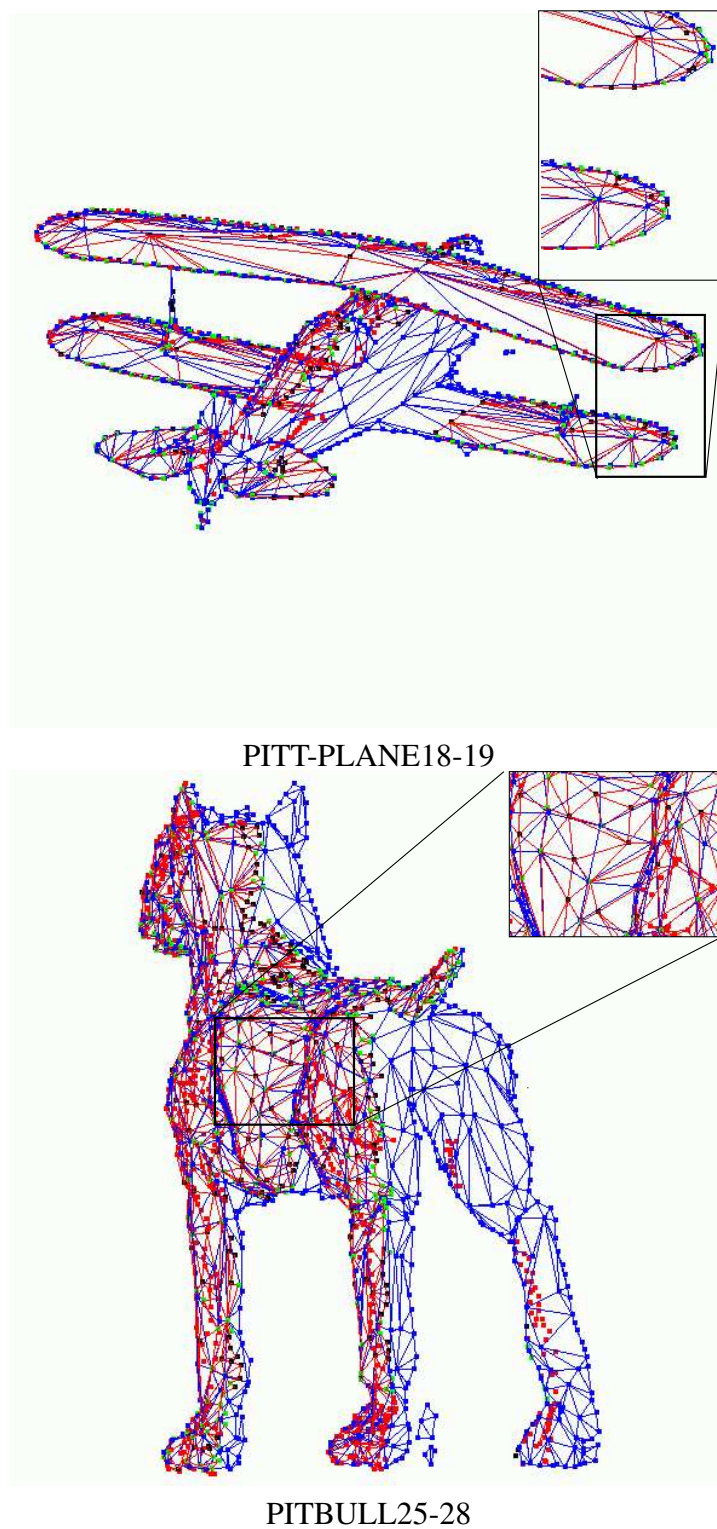
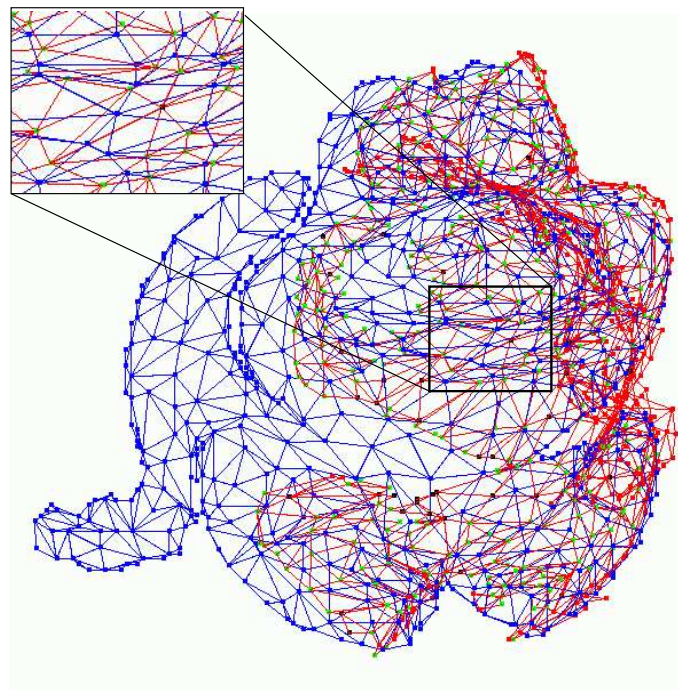
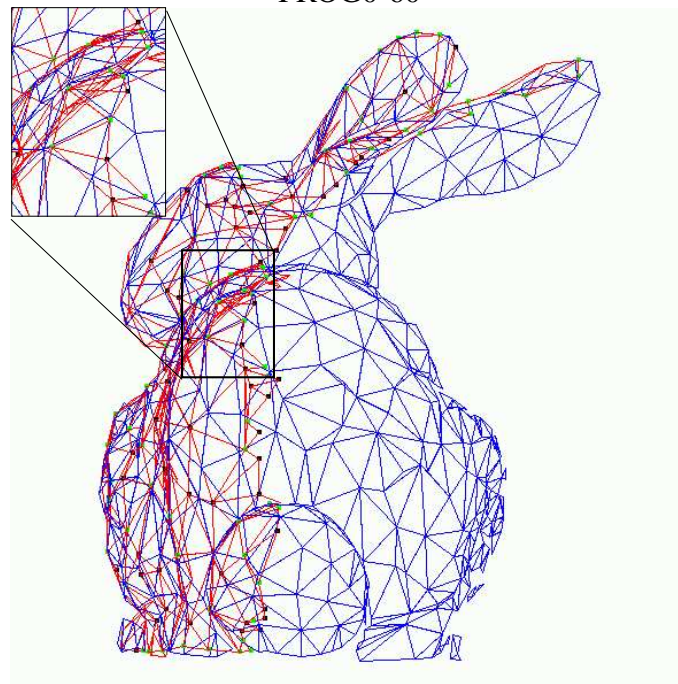


Fig. 4.14: Pares de malhas simplificadas (PITT-PLANE e PITBULL) no mesmo referencial



FROG0-60



BUNNY0-90

Fig. 4.15: Pares de malhas simplificadas (FROG e BUNNY) no mesmo referencial

Imagens	Amostras ($\mathcal{N}$)	Vértices (N)	% Redução
PITT-PLANE 18	21261	506	97,6
PITT-PLANE 19	21278	468	97,8
PITT-PLANE 20	22027	722	96,7
PITT-PLANE 21	21017	363	98,3
PITT-PLANE 22	18868	647	96,6
PITT-PLANE 23	21019	371	98,2
PITBULL25	23363	995	95,7
PITBULL28	20088	756	96,2
PITBULL31	25303	736	97,1
PITBULL34	26116	794	97,0
PITBULL37	14807	2139	85,6
PITBULL40	26007	799	96,9
FROG0	9640	744	92,3
FROG60	8618	704	91,8
FROG120	10767	609	94,3
FROG180	10779	560	94,8
FROG240	10993	499	95,5
FROG300	10983	691	93,7
BUNNY0	40256	341	99,2
BUNNY90	30379	554	98,2
BUNNY180	40251	247	99,4
BUNNY270	31701	458	98,6

Tab. 4.4: Redução do volume de dados.

57,4% dos vértices na região de sobreposição, possuem um correspondente.

4.4 Pré-alinhamento com uso de malhas simplificadas

Após a explanação sobre a obtenção de uma malha simplificada adequada para o pré-alinhamento, pode-se trocar os dois primeiros passos do algoritmo apresentado no Apêndice B.

1. Construir a malha-referência
2. Construir a malha-alvo

por

1. Construir a malha simplificada $\mathcal{MS}_1$ com erro ϵ .
2. Construir a malha simplificada $\mathcal{MS}_2$ com erro ϵ igual ao considerado para a redução da malha $\mathcal{MS}_1$.

	Vértices	% Simplif.	Vértices em A_{S1} (VS)	VS c/corresp.(VC)	% VC/VS
PITT-PLANE18-19	506	97,6%	326	226	69,3%
PITBULL25-28	995	95,7%	463	276	60,0%
FROG0-60	744	92,3%	372	336	90,3%
BUNNY0-90	341	99,2%	129	74	57,4%

Tab. 4.5: Porcentagem de Vértices Correspondentes/Vértices em A_{S1}

Com essa substituição

- os espaços de escolha e de busca são reduzidos drasticamente, a razão de 4% do total de amostras do conjunto original. Apesar disso a malha simplificada preserva as principais características da forma global da imagem de profundidade, provendo um conjunto de vértices dispersos e representativos, ainda em regiões com considerável baixa curvatura.
- o sucesso na determinação de pelo menos um triedro correspondente é assegurado pelo controle da densidade ϵ da malha (Seção 4.3.1) e pelo limiar de busca δ estabelecido na Seção 4.3.2.
- o processo de filtragem baseado nas malhas simplificadas, embora seja análogo aos já propostos (Capítulo 2), apresenta duas vantagens:
 - Não há necessidade de uso da amostragem completa para uma avaliação global confiável. A malha simplificada representa em forma concisa a forma global da superfície, portanto considerar seus vértices é equivalente a usar a amostragem completa para efeito de obter uma aproximação grosseira do alinhamento.
 - A malha simplificada contém um conjunto reduzido de amostras, na ordem de 4%, o que permite realizar, com eficiência, o cálculo dos erros para todos os vértices da malha $\mathcal{MS}_1$.

4.5 Análise

Nesta seção é feita uma análise, em termos práticos, sobre o procedimento de registro com base na malha simplificada.

4.5.1 Existência de um par de triedros correspondentes

O uso da malha simplificada reduz de fato drasticamente o volume dos dados usados no pré-alinhamento, contribuindo obviamente no aumento da eficiência do algoritmo. No entanto, fica ainda

por avaliar se este ganho em eficiência não prejudica a existência do par de triedros correspondentes a qual é necessário para que o algoritmo obtenha o resultado desejado.

Na seção 3.4.1 foi apresentada uma ilustração na qual mostrou-se o resultado do algoritmo proposto nos pares de imagens pertencentes ao conjunto BUNNY. Conferiram-se os resultados obtidos com os parâmetros reais do registro e verificou-se uma grande semelhança com um erro mínimo de 1.6° no ângulo de rotação e 0.01 no vetor translação. Porém esses resultados foram obtidos quando o algoritmo foi aplicado em malhas muito refinadas e não malhas simplificadas.

Nesta seção mostrar-se-á que apesar da redução drástica dos dados evidenciada nas malhas simplificadas, na média, em 96%, ainda conseguem-se resultados muito próximos dos reais. Isto significa que a malha-referência e malha-alvo simplificadas preservaram uma adequada densidade através do parâmetro ϵ sugerido a qual foi suficiente para se encontrar pelo menos um triedro na região de sobreposição e assegurou a correspondência deste triedro na malha-alvo simplificada através do parâmetro δ sugerido. De fato, foi mostrado na Seção 4.3.2 que com a redução sugerida neste trabalho assegura-se, na média, que mais do 50% dos vértices que se encontram na região de sobreposição das malhas simplificadas obtenham um correspondente.

Foi aplicado o algoritmo proposto sobre pares de imagens oriundas de conjuntos de imagens encontradas nos repositórios [35, 51, 50] para validar os valores sugeridos para f_1, f_2, ϵ e δ . Porém para ilustração foi escolhido o conjunto de seis imagens de profundidade visualizado na Figura 1.2 [50], cujos valores dos parâmetros de alinhamento são conhecidos, o que facilita a avaliação da qualidade dos resultados do algoritmo proposto. Estas imagens são BUNNY0, BUNNY45, BUNNY90, BUNNY180, BUNNY270 e BUNNY315 com os quais formou-se os seguintes pares de imagens:

- BUNNY0-BUNNY90,
- BUNNY90-BUNNY180,
- BUNNY180-BUNNY270,
- BUNNY270-BUNNY0,
- BUNNY270-BUNNY315,
- BUNNY315-BUNNY0.

As seis imagens foram obtidas rotacionando o sensor em torno do eixo y . Para cada par que foi obtido por uma rotação de 45° (exemplo BUNNY270-BUNNY315), usamos $f_1 = 0.5$ e $f_2 = 0.7$, e para cada par que tem diferença angular de 90° (exemplo BUNNY0-BUNNY90), usamos $f_1 = 0.0$ e $f_2 = 0.3$. Como na Tabela 3.2, a Tabela 4.6 especifica o par de imagens usados na coluna Par de imagens e os parâmetros da transformação de registro real, especificados nas colunas R_y (R) (Rotação em torno ao eixo y), x (R), y (R) e z (R), e os parâmetros da transformação obtidos pelo

algoritmo de pré-alinhamento nas colunas R_y (M), x (M), y (M) e z (M). Enquanto que na Tabela 3.1 foi escolhido um número arbitrário de faces como valor de entrada para a simplificação, nesta tabela usou-se o erro de simplificação ϵ , discutido e sugerido na Seção 4.3.1. A qualidade da transformação do pré-alinhamento foi verificada pela diferença de ambos valores (R e M) que deu uma diferença angular, na média, de 1.4° e uma diferença entre os valores do vetor translação, na média, de 0.002. Comparando estes valores com os obtidos a partir da Tabela 3.1, verifica-se que, em média, houve melhora.

Par de Imagens	Rot _y (R)	x (R)	y (R)	z (R)	Rot _y (M)	x (M)	y (M)	z (M)
0-90	90	-0.0001	0.0002	-0.0001	89.9975	-0.0001	0.0004	-0.0007
90-180	90	-0.0001	0.0001	-0.0001	91.3913	-0.0007	0.0008	-0.0022
180-270	90	-0.0000	0.0003	0.0000	90.1022	-0.0004	0.0001	0.0009
270-0	90	-0.0001	0.0000	0.0004	93.696	0.0009	0.0035	0.0001
270-315	45	0.0137	-0.0003	0.0045	44.4341	0.0116	-0.0013	0.0074
315-0	45	-0.0065	-0.0000	-0.0128	47.6569	-0.0092	-0.0000	-0.0131

Tab. 4.6: Parâmetros da transformação entre pares de imagens do modelo BUNNY

4.5.2 Redução de Potenciais Candidatos

Nesta seção mostra-se a capacidade do algoritmo proposto para reduzir o número de potenciais candidatos. Na Tabela 4.7 apresenta-se o número de candidatos obtidos após a busca exaustiva de descritores-alvo em $\mathcal{MS}_2$, para descritores-referência construídos em $\mathcal{MS}_1$. O número de amostras $\mathcal{N}$ da imagem de profundidade é mostrada na coluna **Amostras** e o número de vértices N da malha simplificada é mostrado na coluna **Vértices** da Tabela 4.4. Usando o conjunto completo de dados, dever-se-ia avaliar exaustivamente $(\mathcal{N}_1 \times C_3^{\mathcal{N}_1-1}) \times (\mathcal{N}_2 \times C_3^{\mathcal{N}_2-1})$ possibilidades para assegurar que o melhor candidato fosse escolhido. Esse valor é apresentado na coluna **Candidatos 1**. Após a redução pela simplificação das malhas, o número de possibilidades, reduziria a $(N \times C_3^{(N-1)}) \times (N \times C_3^{(N-1)})$, que é apresentado na coluna **Candidatos 2**. No entanto, ao invés de encontrar a melhor correspondência dentre esse número grande de possibilidades para cada par de imagens, o algoritmo proposto reduz, ainda mais, o número de possibilidades. O número de potenciais transformações, que devem ser avaliadas, é dado na coluna **Candidatos 3**. Note-se que o percentual de redução é, em média, mais que 99% no espaço de busca. Por outro lado, em média, um tempo de 86 segundos mais um *overhead* de, em média, 30 segundos de tempo de simplificação, foram requeridos para obter o pré-alinhamento de cada par de imagens com uso de um computador Pentium IV 3.2GHz com 512 MB de RAM (plataforma Linux). Ao se comparar este tempo com os apresentados na coluna **Tempo** da Tabela 3.2 percebeu-se, um ganho em eficiência quando é realizada uma redução criteriosa priorizando a preservação da informação essencial da geometria global da superfície.

Pares de imagens	f_1	f_2	Candidatos 1	Candidatos 2	Candidatos 3
PITT-PLANE18-PITT-PLANE19	0.5	0.7	1.16e33	8.52e19	3257
PITT-PLANE19-PITT-PLANE20	0.5	0.7	1.34e33	3.55e20	4210
PITT-PLANE20-PITT-PLANE21	0.5	0.7	1.28e33	1.28e20	455
PITT-PLANE21-PITT-PLANE22	0.1	0.3	6.86e32	8.246e19	3752
PITT-PLANE22-PITT-PLANE23	0.1	0.3	6.87e32	8.99e19	63
PITT-PLANE23-PITT-PLANE18	0.5	0.7	1.11e33	3.35e19	993
PITBULL25-PITBULL28	0.5	0.7	1.35e33	8.77e21	645
PITBULL28-PITBULL31	0.5	0.7	1.85e33	2.62e21	1068
PITBULL31-PITBULL34	0.5	0.7	5.29e33	3.19e21	771
PITBULL34-PITBULL37	0.01	0.3	6.21e32	2.29e23	19767
PITBULL37-PITBULL40	0.01	0.3	6.10e32	2.35e23	6960
PITBULL40-PITBULL25	0.5	0.7	3.78e33	1.09e22	519
FROG0-FROG60	0.3	0.6	1.32e30	2.06e21	6887
FROG60-FROG120	0.3	0.6	2.06e30	9.21e20	3493
FROG120-FROG180	0.3	0.6	5.03e30	3.68e20	3347
FROG180-FROG240	0.3	0.6	5.47e30	1.66e20	4521
FROG240-FROG300	0.3	0.6	5.90e30	3.85e20	3995
FROG300-FROG0	0.3	0.6	3.49e30	1.91e21	5167
BUNNY0-BUNNY90	0.01	0.3	1.32e30	3.44e19	12797
BUNNY90-BUNNY180	0.01	0.3	2.06e30	9.40e18	4058
BUNNY180-BUNNY270	0.01	0.3	5.03e30	4.38e18	3389
BUNNY270-BUNNY0	0.01	0.3	5.47e30	1.60e19	7407

Tab. 4.7: Redução de falsas correspondências.

4.6 Considerações Finais

Em analogia ao sistema de visão humano, para pré-alinhamento somente os principais aspectos da forma da imagem são de relevância. Modelos excessivamente detalhados não necessariamente fornecem resultados mais precisos. De fato, o objetivo do alinhamento é de casar a região de sobreposição das duas imagens. Seguindo este raciocínio, propõe-se a idéia de usar as amostras de uma malha simplificada para encontrar a transformação rígida que alinhe as duas imagens de um objeto rígido.

No entanto, o objetivo da simplificação de malhas tem sido, tradicionalmente, renderização mais eficiente. Portanto, preocupa-se somente em reduzir o número de vértices que descrevem o objeto, enquanto a sua forma geral é preservada. Porém, no contexto de alinhamento, não apenas a redução do número de faces e preservação da forma do objeto, mas também outra característica é fundamental: a densidade adequada, pois precisa-se de uma amostragem mínima para busca de correspondências. Esse tipo de redução garante que, mesmo sem se conhecer a região de sobreposição, sempre se tenham amostras presentes nessa região e, conseqüentemente, descritores gerados dentro da região. Por outro lado, é também necessário que as representações das duas malhas, que se pretendem registrar, possuam pontos nas suas regiões de sobreposição que se correspondam mutuamente, sob certa tolerância. Isso assegurará a determinação de descritores correspondentes.

Face à disposição de uma grande variedade de implementação de algoritmos de decimação [36, 37], houve preocupação de se mostrar neste capítulo como reutilizá-los, após devidas adequações, para o propósito específico deste trabalho. Mais especificamente, apresentou-se uma forma de adaptar a técnica de simplificação QSLIM, muito popular e de fácil acesso [21]. Para esta técnica foi suficiente introduzir dois parâmetros de controle: erro de simplificação ϵ para ajustar o nível de decimação e intervalo de tolerância δ para aumentar a folga na busca pelas amostras correspondentes com uso destas malhas simplificadas. O parâmetro ϵ é dependente de um índice de suavidade da superfície o qual é estimado a partir da classificação de vértices da malha fina.

A inserção do uso de uma malha triangular simplificada no algoritmo de busca apresentado no Capítulo 3 dá origem a um algoritmo de pré-alinhamento, com as seguintes características:

- Espaço de escolha reduzido: os vértices de uma malha simplificada adaptada $\mathcal{M}_1$ são usados ao invés do conjunto denso das amostras de $\mathcal{R}_1$.
- Espaço de busca reduzido: os vértices de uma malha simplificada adaptada $\mathcal{M}_2$ são usados ao invés do conjunto denso das amostras de $\mathcal{R}_2$.
- Definição de um descritor-referência em cada vértice de $\mathcal{M}_1$ em lugar de selecionar aleatoriamente um descritor em $\mathcal{R}_1$: com a redução do espaço de escolha, pode-se analisar todos os vértices da malha simplificada adaptada (Seção 3.2.1).

- Busca exaustiva eficiente: a busca exaustiva é ainda uma técnica confiável para garantir que um descritor-alvo, se existir, seja encontrado. Os resultados experimentais mostram que a percentagem de redução dos dados, em relação aos dados originais, é maior de 96%, portanto, é factível realizar uma busca exaustiva entre todos os vértices. Além disso, a capacidade discriminante do triedro aumenta a eficiência na busca por descritores-alvo.
- Filtragem eficiente e precisa de transformações hipotéticas: não é necessário considerar a amostragem completa para medição de erros; os vértices da malha fornecem um subconjunto de pontos representativos da superfície, o que permite avaliar com confiabilidade o erro das transformações.

A Tabela 4.8 é uma extensão da Tabela 2.1, incluindo adicionalmente as características do procedimento proposto. Através desta tabela, tem-se uma visão geral comparativa deste trabalho em relação aos existentes na literatura.

Técnica	Espaço de escolha do descritor-referência	Configuração dos descritores	Medidas invariantes dos descritores		Espaço de busca do descritor-alvo	Procedimento de Filtragem
			das amostras	da relação entre as amostras		
DBR	todas as amostras (A_1)	triangular (três amostras)	não possui	distâncias	todas as amostras (A_2)	Avaliação de ajuste de vizinhança e subamostragem de $\mathcal{R}_1$
CPO	todas as amostras (A_1)	$N > 3$ amostras sem estrutura definida	histogramas com posto por distâncias acumuladas	distâncias	todas as amostras (A_2)	Avaliação de ajuste de alguns pontos de A_1
CHB	malha associada a A_1	triangular (três vértices de um triângulo da malha)	histograma geométrico (distância e ângulo)	distâncias	malha associada a A_2	Avaliação de ajuste da malha associada a A_1 , porém só de algumas transformações.
CCB	pontos bitangentes em $\mathcal{R}_1$	quadrilateral (dada pelos pontos extremos de dois segmentos)	não possui	distância	pontos bitangentes em $\mathcal{R}_2$	Avaliação de ajuste de uma subamostragem de A_1
CCI	pontos de curvatura média igual a zero em A_1	linear (dada pelos pontos extremos de N segmentos)	não possui	distância	pontos de curvatura média igual a zero em A_2	Avaliação de ajuste de pontos especiais e de A_1
CGI	pontos máximos locais de gradiente de intensidade em I_1	triangular (dada pelos três vértices do triângulo da tetraedrização)	não possui	distância e ângulos	pontos máximos locais de gradiente de intensidade em I_2	Avaliação de ajuste de uma subamostragem de pontos especiais
A proposta	vértices da malha simplificada associada a $\mathcal{R}_1$	tetraedral (dada por quatro vértices de um triedro espacial)	curvatura média	distâncias e ângulos	vértices da malha simplificada associada a $\mathcal{R}_2$	Avaliação de ajuste dos vértices da malha simplificada

Tab. 4.8: Tabela comparativa dos principais trabalhos presentes na literatura

Capítulo 5

Validação

Um registro totalmente automático de duas imagens de profundidade, parcialmente sobrepostas de um objeto rígido, em um sistema de referência comum, é ainda um problema não totalmente resolvido. O procedimento mais conhecido é dividir o problema em dois sub-problemas: estimar um pré-alinhamento (alinhamento inicial) e refinar este alinhamento até que o erro quadrático entre os pontos correspondentes das duas imagens seja mínimo.

Este trabalho contribui com uma nova proposta para estimação do alinhamento inicial, cujos detalhes foram descritos nos capítulos 3 e 4. Para validar o algoritmo proposto, foi implementado um protótipo de registro de pares de imagens em duas etapas. Na primeira etapa é estimado o pré-alinhamento, com uso do referido algoritmo e, na segunda etapa, este alinhamento é refinado por meio de um algoritmo ICP (*Iterative Closest Point*). Diferentes geometrias da superfície do objeto foram testadas [35, 51, 50]. Em todos os testes, o algoritmo proposto gerou uma estimativa inicial que permitiu ao algoritmo ICP convergir para uma solução visualmente correta. Para os testes realizados neste trabalho usou-se o aplicativo **Scanalyze** [48] para ajuste fino.

O restante deste capítulo está organizado da seguinte forma. Na Seção 5.1 apresenta-se o fluxograma do novo algoritmo de pré-alinhamento. Na Seção 5.2 apresenta-se o método ICP que é usado tradicionalmente para realizar registro fino. Na Seção 5.3 são apresentados os resultados obtidos após aplicar este algoritmo a um conjunto de pares de imagens de profundidade e após refinar o alinhamento, com o aplicativo **Scanalyze**. Também são apresentados experimentos feitos usando imagens similares às usadas por Planitz et al. [7, 8], para analisar e comparar alguns dos mais conhecidos algoritmos de pré-alinhamento. E a Seção 5.4 apresenta as considerações finais do capítulo.

5.1 Algoritmo

O pré-alinhamento de duas imagens de profundidade $\mathcal{R}_1$ e $\mathcal{R}_2$ é conseguido através do casamento das malhas simplificadas $\mathcal{MS}_1$ e $\mathcal{MS}_2$, o qual, por sua vez, é realizado pelo casamento de triedros

espaciais propostos no Capítulo 3. A partir de um triedro espacial definido em cada vértice da $\mathcal{MS}_1$ procuram-se prováveis triedros correspondentes na malha $\mathcal{MS}_2$ (Correspondência), e determina-se uma transformação rígida para cada par (Posicionamento). Como vários são os vértices da malha e vários são os candidatos correspondentes para cada triedro em $\mathcal{MS}_1$, então, várias transformações serão determinadas. Para verificar a validade das transformações candidatas, cada uma delas passa por uma avaliação de erros de ajuste entre as duas malhas (Filtragem). A transformação que alinha o maior número de faces das malhas é considerada a solução.

A Figura 5.1(a) apresenta o par $\mathcal{P}_{120-160}$ do conjunto FROG (ângulos de vista, 120° e 160°). Convertem-se essas nuvens de pontos a malhas triangulares densas e simplificam-se as malhas com erro de aproximação ϵ como explicado no Capítulo 4. A Figura 5.1(b) apresenta o resultado da simplificação para $\mathcal{P}_{120-160}$. Descritores são construídos em cada vértice da $\mathcal{MS}_1$ e um procedimento de busca é realizado para encontrar potenciais candidatos na malha $\mathcal{MS}_2$, com os quais determinam-se potenciais transformações. Na Figura 5.1(c) podem-se ver os descritores de um par de vértices correspondentes. Todas as transformações passam por um processo de filtragem para, finalmente, obter-se a melhor solução T (Figura 5.1(d)). Esta transformação poderá ser refinada ao se usar como transformação inicial para o algoritmo ICP.

5.2 Ajuste Fino (aplicação de ICP)

O algoritmo ICP proposto por Besl & McKay [5] é o método clássico para resolver o problema do registro fino. O algoritmo, iterativamente, determina uma transformação T_k com base na suposição de que pares de correspondências podem ser estabelecidos tomando um ponto p_i em $\mathcal{R}_1$ e o seu mais próximo na imagem $\mathcal{R}_2$. Então, os pontos de $\mathcal{R}_1$ são transformados pela aplicação de T_k , sendo o processo, de estabelecimento de correspondências e determinação de uma T_{k+1} , repetido até que o quadrado da soma das distâncias entre os pares de pontos seja menor que um limiar pré-definido. O algoritmo ICP trabalha bem, quando uma transformação inicial T_0 próxima de solução correta esteja disponível. Caso contrário, o algoritmo converge a um mínimo local no espaço das transformações, dando uma solução incorreta. Com base nessa característica do ICP, na próxima seção utilizou-se o número de iterações necessárias para convergência como uma medida de qualidade do algoritmo de pré-alinhamento proposto.

5.3 Resultados

Na aquisição dos resultados desta seção foram utilizados os quatro diferentes conjuntos de imagens mostradas na Figura 4.6, 4.7 e 1.2. Os conjuntos PITT-PLANE, PITBULL, FROG e BUNNY. Os dois primeiros provêm do site referenciado em [51], o terceiro do site citado em [35] e o quarto

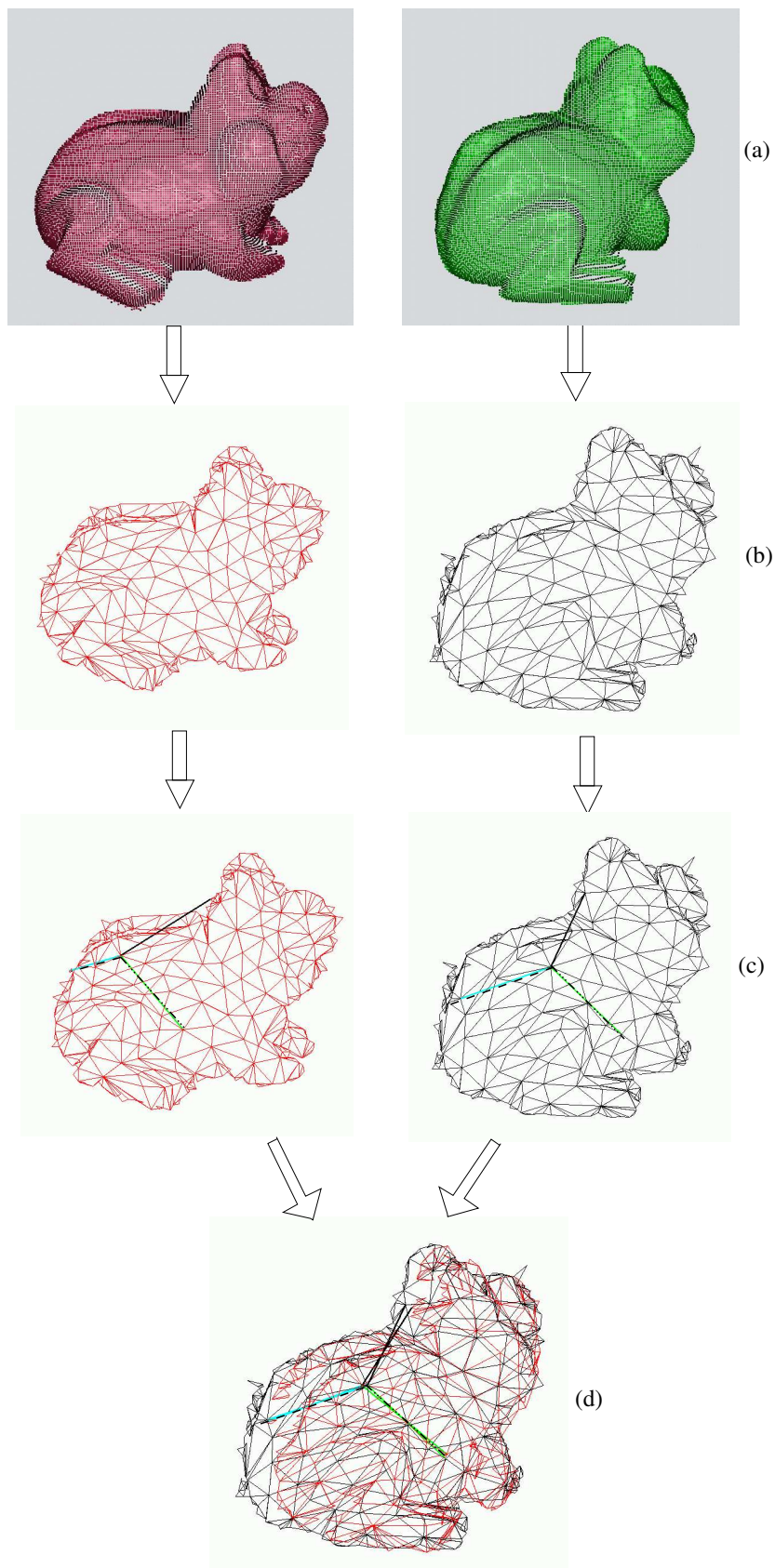


Fig. 5.1: Pré-alinhamento de um par de imagens de profundidade.

do site referenciado em [50]. Esses quatro conjuntos de objetos foram escolhidos com o intuito de mostrar o comportamento do algoritmo proposto quando aplicado a imagens de diferentes repositórios. O algoritmo proposto funcionou satisfatoriamente para os quatro casos e mostra que o seu bom desempenho independe da forma geométrica e tamanho dos objetos. Todos os resultados desta tese foram obtidos com uso de computador Pentium IV 3.2GHz com 512 MB de RAM (plataforma Linux). Os valores usados para ϵ e δ , foram os sugeridos no Capítulo 4 e os valores atribuídos a f_1 e f_2 foram mostrados na Tabela 4.7.

5.3.1 Qualidade do Pré-alinhamento

Nesta seção são apresentados pares de imagens de cada conjunto, os quais foram selecionados para mostrar o pré-alinhamento obtido pelo algoritmo proposto e o resultado final de registro após a aplicação do algoritmo ICP. Nas Figuras 5.2 e 5.3 mostram-se os pares PITT-PLANE20-PITT-PLANE21 e PITBULL31-PITBULL34. Em seguida, nas Figuras 5.4 e 5.5, os pares FROG0-FROG60 e BUNNY0-BIRD90. Em todas as figuras, as imagens de profundidade originais são mostradas separadamente nos itens (a) e (b). Em (c) exibe-se o par de malhas triangulares simplificadas, correspondentes às imagens, registrado pela transformação rígida T_0 , obtida pelo algoritmo proposto. Em (d), a transformação T_0 é aplicada em malhas finas para efeito de comparação com o registro fino. Em (e), o registro final das imagens.

A qualidade do registro fino pode ser avaliado subjetivamente pelo intercalamento das superfícies como sugerido por Flynn [15] ou quantitativamente pela sua proximidade da transformação real de registro. Ou seja, a qualidade do pré-alinhamento será julgada boa se fizer o método ICP convergir ao resultado correto ou se após a aplicação do ICP obtiver um aumento no grau de intercalamento entre as superfícies. Pode-se observar mediante uma avaliação visual que os quatro pares tiveram o grau de intercalamento aumentado após a aplicação do registro fino. Resta apenas apresentar uma avaliação quantitativa a qual é possível fazer com os pares do conjunto BUNNY cujas transformações reais são conhecidas. A Tabela 4.6 especifica os pares de imagens usados na coluna **Par de imagens** e os parâmetros da transformação de registro real, especificados nas colunas Rot_y (R) (Rotação em torno ao eixo y) e $(x$ (R), y (R), z (R)) especificando os três valores do vetor translação. Da mesma forma os parâmetros da transformação obtidos pelo algoritmo de pré-alinhamento apresentado são apresentados nas colunas Rot_y (M) e $(x$ (M), y (M), z (M)). Já na Tabela 5.1 são apresentados os parâmetros da transformação obtido após aplicar o registro fino. A qualidade do pré-alinhamento pode ser verificada tanto pela diferença de ambos valores (R e M), isto é, uma diferença angular, na média, de 1.4° e uma diferença entre os valores do vetor translação, na média, de 0.002, e também pela diferença dos valores (R e I) que deu uma diferença angular, na média, de 0.11° e uma diferença entre os valores do vetor translação de 0^1

¹Somente os quatro pares que usamos para realizar o registro global de *bunny* (Figura 5.6), foram considerados na

Par de Imagens	Rot _y (l)	x (l)	y (l)	z (l)
0-90	90.2239	-0.0001	0.0002	-0.0001
90-180	90.1095	-0.0001	0.0001	-0.0001
180-270	90.0138	-0.0000	0.0003	0.0000
270-0	90.0944	-0.0001	0.0000	0.0004

Tab. 5.1: Parâmetros da transformação entre pares de imagens do modelo BUNNY

Vale ainda a pena mencionar que muitos dos pares de imagens considerados neste trabalho tiveram uma região de sobreposição consideravelmente pequena, como foi estimado com o seguinte procedimento realizado a-posteriori dos resultados: (a) Colocaram-se as imagens de profundidade no mesmo referencial. (b) Contabilizam-se o número num de amostras da imagem $\mathcal{R}_1$ cujas amostras mais próximas na imagem $\mathcal{R}_2$ estivessem abaixo do limiar s . (c) Calculou-se a percentagem $\%_{sobrepos}$ dividindo num pelo número total de amostras da imagem $\mathcal{R}_1$. A Tabela 5.2 apresenta o percentual de sobreposição que cada par de imagens tem em relação ao número de amostras da primeira imagem. Pode-se concluir com isso, que ajustando adequadamente os parâmetros f_1 , f_2 conseguiu-se obter um pré-alinhamento de qualidade mesmo para pares com pouca região de sobreposição como por exemplo FROG300-0 que apresenta apenas 24.5%.

5.3.2 Registro de n vistas para reconstrução completa

Nesta seção, várias vistas do mesmo objeto são registradas, par a par, pelo algoritmo proposto. Como os conjuntos PITT-PLANE, PITBULL e FROG têm 6 vistas e o conjunto BUNNY é composto só de 4, denotaremos j o número de vistas do conjunto. As j imagens $\mathcal{R}_i$ de cada conjunto apresentam vistas consecutivas do objeto, portanto, os pares $\mathcal{R}_i \mathcal{R}_{mod(i+1,j)}$ são parcialmente sobrepostos. As j imagens foram escolhidas para representar a rotação panorâmica do objeto (360 graus).

Para cada conjunto, é apresentado o pré-alinhamento de cada um dos j pares $\mathcal{R}_i \mathcal{R}_{mod(i+1,j)}$. E para cada par registrado é anotado o número de iterações que foram necessárias para o algoritmo ICP convergir, usando a transformação T_0 , obtida pelo algoritmo de pré-alinhamento proposto, como entrada. As Figuras 5.9, 5.11 e 5.13 mostram os pares de imagens pertencentes aos conjuntos PITT-PLANE, PITBULL e FROG, respectivamente, e por último, a Figuras 5.6 exibe os 4 pares de imagens correspondentes ao conjunto BUNNY.

Para construir um modelo completo, os pares registrados devem ser integrados em um único sistema de referência. A forma mais comum de realizar o registro de múltiplas vistas, quando se tem o registro par a par, é a seguinte: Seja 4 vistas $\mathcal{R}_i$ e seja T_i a transformação que alinha o par $\mathcal{R}_i \mathcal{R}_{mod(i+1,4)}$, e assuma que o sistema global de referência é o da vista $\mathcal{R}_1$, então para obter a trans-



(a) PITT-PLANE20



(b) PITT-PLANE21

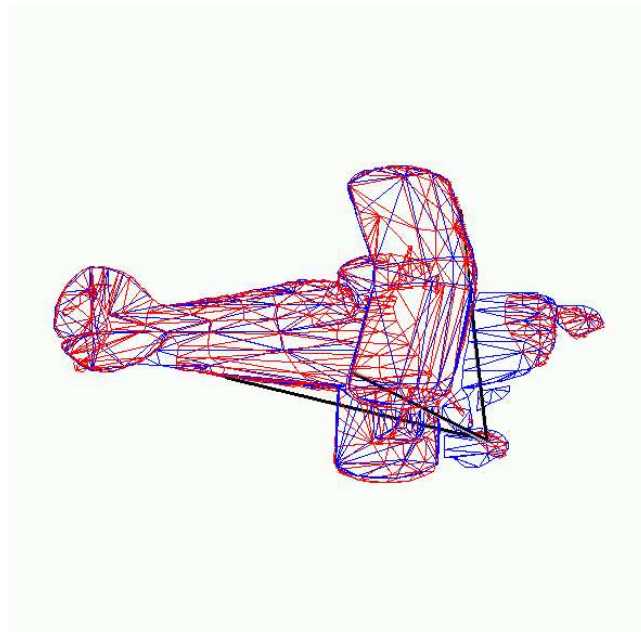
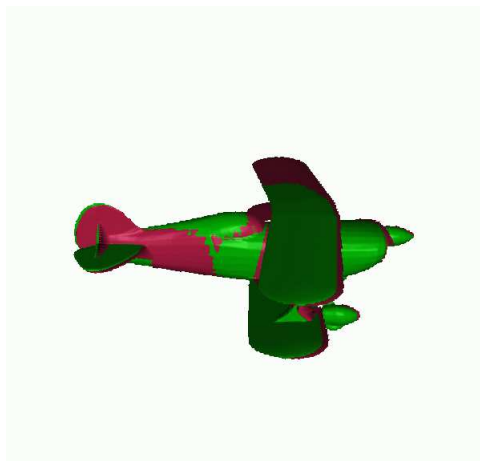
(c) Pré-alinhamento de malhas simplificadas (T_0)Pré-alinhamento de malhas finas (T_0)(e) Registro após refinamento de T_0 (ICP - 7 iterações)

Fig. 5.2: Pré-alinhamento e Registro Fino (PITT-PLANE).

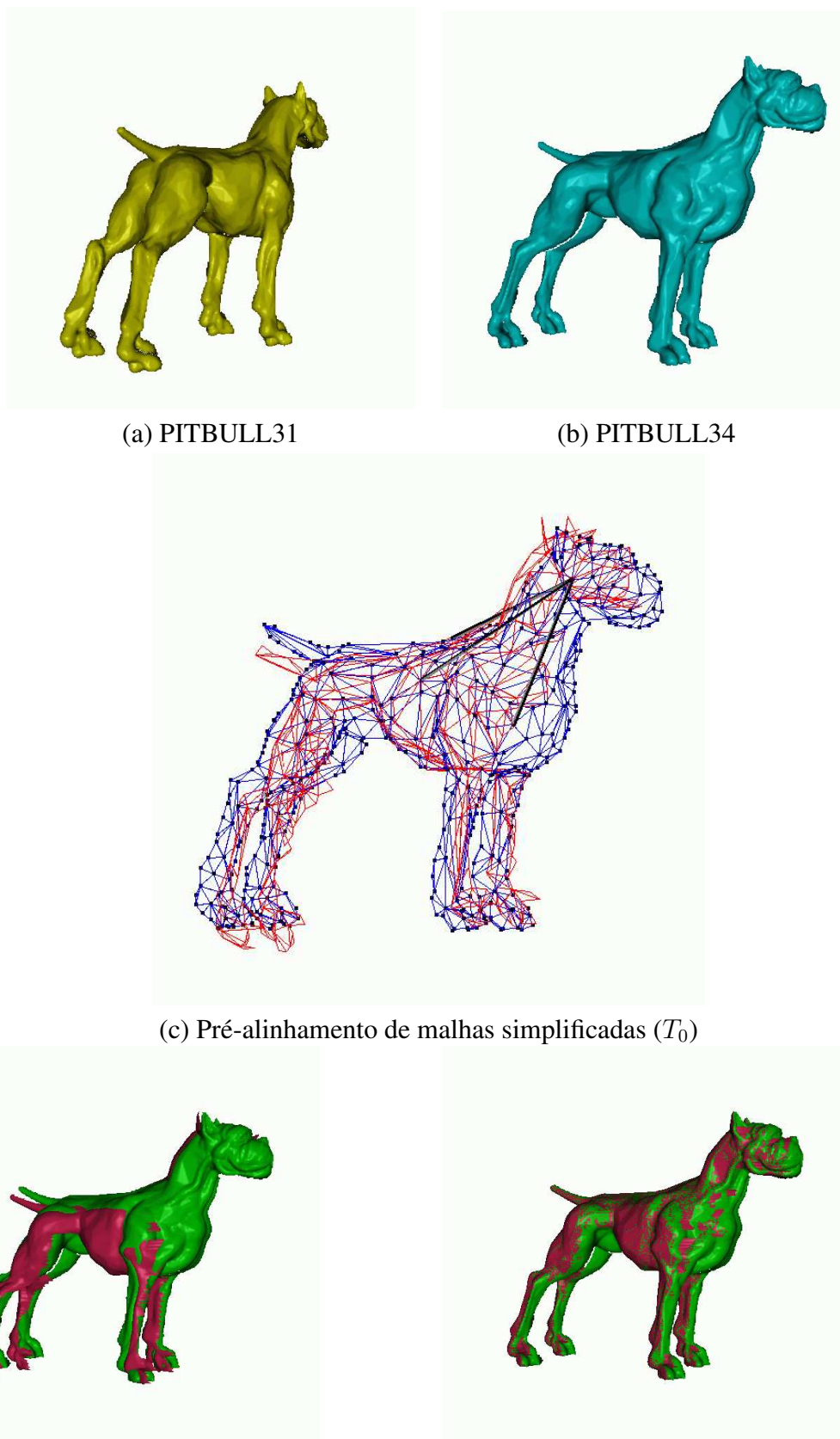
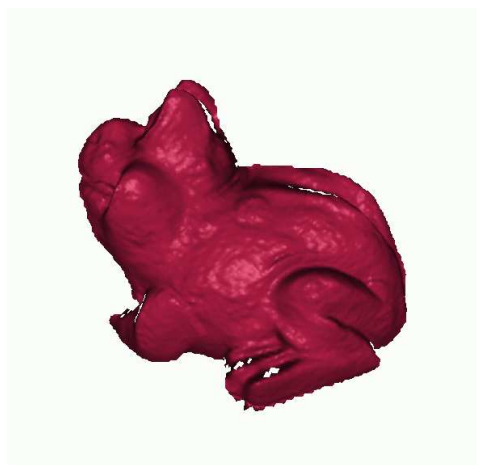


Fig. 5.3: Pré-alinhamento e Registro Fino (PITBULL).



(a) FROG300



(b) FROG0

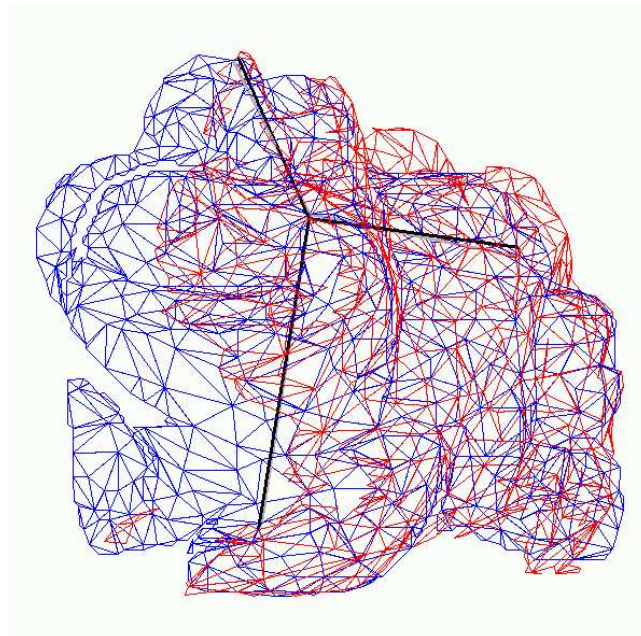
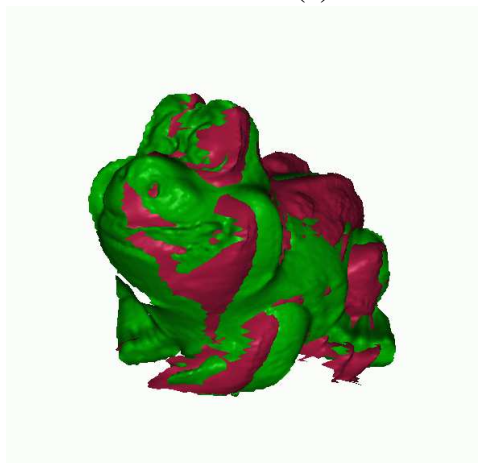
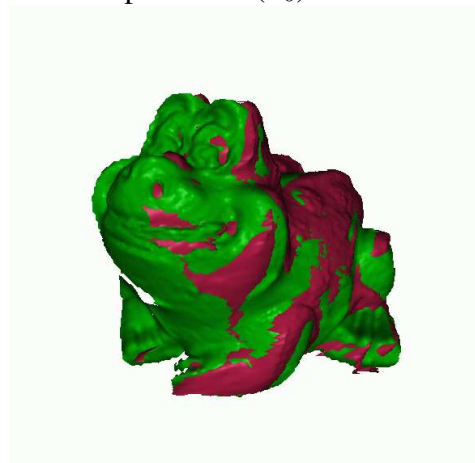
(c) Pré-alinhamento de malhas simplificadas (T_0)Pré-alinhamento de malhas finas (T_0)(e) Registro após refinamento de T_0 (ICP - 9 iterações)

Fig. 5.4: Pré-alinhamento e Registro Fino (FROG).

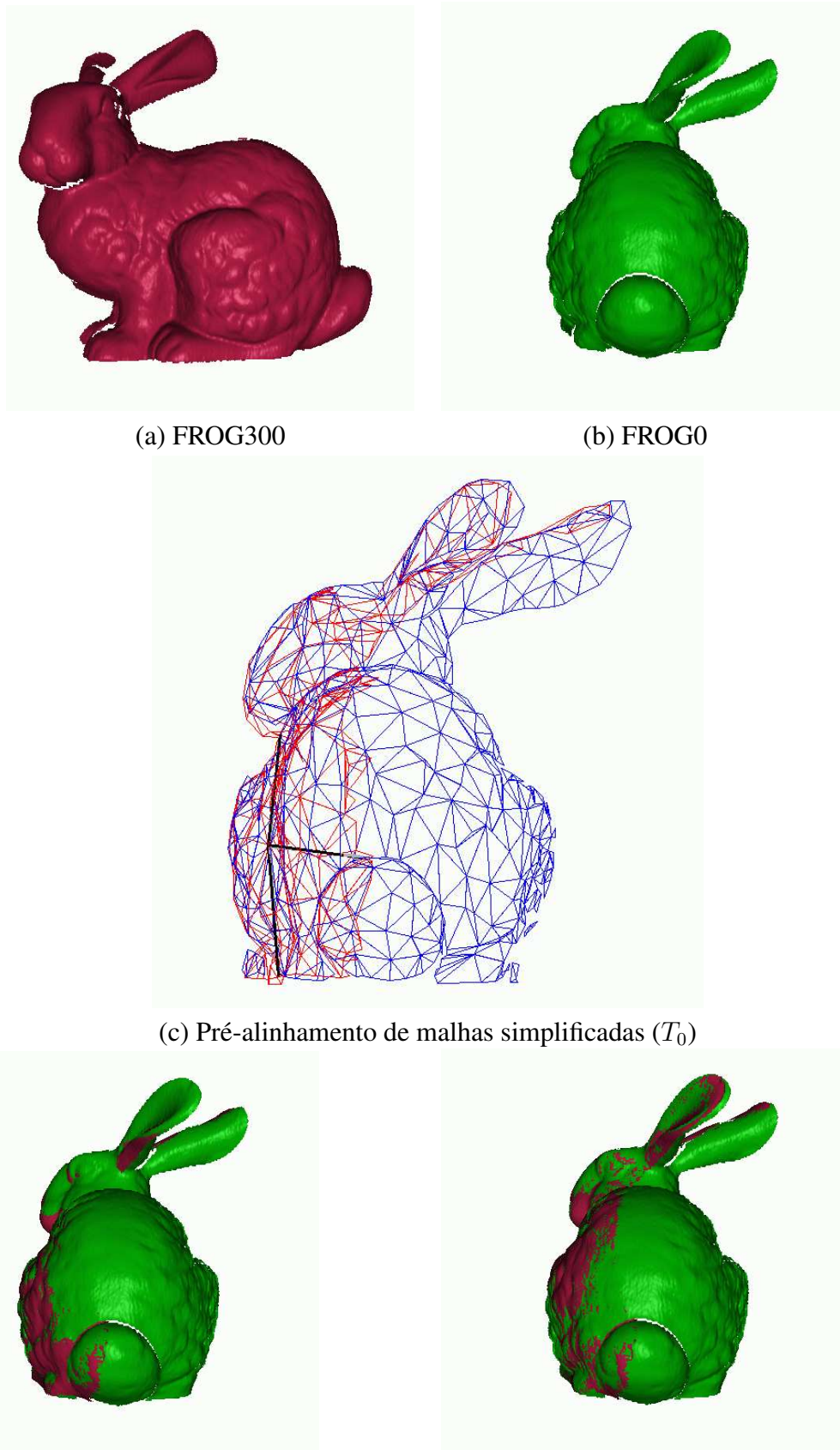


Fig. 5.5: Pré-alinhamento e Registro Fino (BUNNY).

Par de Imagens	%sobreposição
PITT-PLANE18-PITT-PLANE19	51.2
PITT-PLANE19-PITT-PLANE20	65.2
PITT-PLANE20-PITT-PLANE21	68.1
PITT-PLANE21-PITT-PLANE22	41.4
PITT-PLANE22-PITT-PLANE23	41.5
PITT-PLANE23-PITT-PLANE18	64.2
PITBULL25-PITBULL28	42.0
PITBULL28-PITBULL31	28.0
PITBULL31-PITBULL34	55.7
PITBULL34-PITBULL37	28.7
PITBULL37-PITBULL40	47.2
PITBULL40-PITBULL25	75.7
FROG0-FROG60	23.3
FROG60-FROG120	25.4
FROG120-FROG180	30.0
FROG180-FROG240	36.6
FROG240-FROG300	34.2
FROG300-FROG0	24.5
BUNNY0-BUNNY90	25.4
BUNNY90-BUNNY180	28.7
BUNNY180-BUNNY270	25.2
BUNNY270-BUNNY0	25.1

Tab. 5.2: Percentual da Região de sobreposição

formação entre a vista $\mathcal{R}_4$ e vista $\mathcal{R}_1$ deve, em princípio, se fazer:

$$T_{4,1} = (T_3 \circ T_2 \circ T_1)^{-1}$$

O problema típico nesta abordagem é que os erros dos registros de pares são propagados ao longo da composição de transformações. Observe nas Figuras 5.7 o resultado dessa propagação. A cabeça do coelho em duas imagens está nitidamente desalinhada.

Para minimizar o efeito desta propagação foi usado o aplicativo **Scanalyze** para realizar simultaneamente o registro fino do conjunto completo de todas as imagens de forma que os erros individuais sejam minimizados. Com isso conseguiu-se um melhor registro global como mostrado na Figura 5.8. Com uso do mesmo procedimento foram integradas as 6 imagens dos conjuntos PITT-PLANE, PITBULL e FROG (Figuras 5.12, 5.12 e 5.14). Quatro vistas foram capturadas para mostrar a junção dos pares. Para avaliar visualmente o registro, cada imagem é exibida em cor diferente. O intercalamento das cores mostra a qualidade do registro obtido. Os pares registrados podem ser usados para passar

por um processo final de integração para obter uma única malha integrada. Os resultados da aplicação do algoritmo proposto em outros conjuntos de imagens são exibidos no Apêndice A.

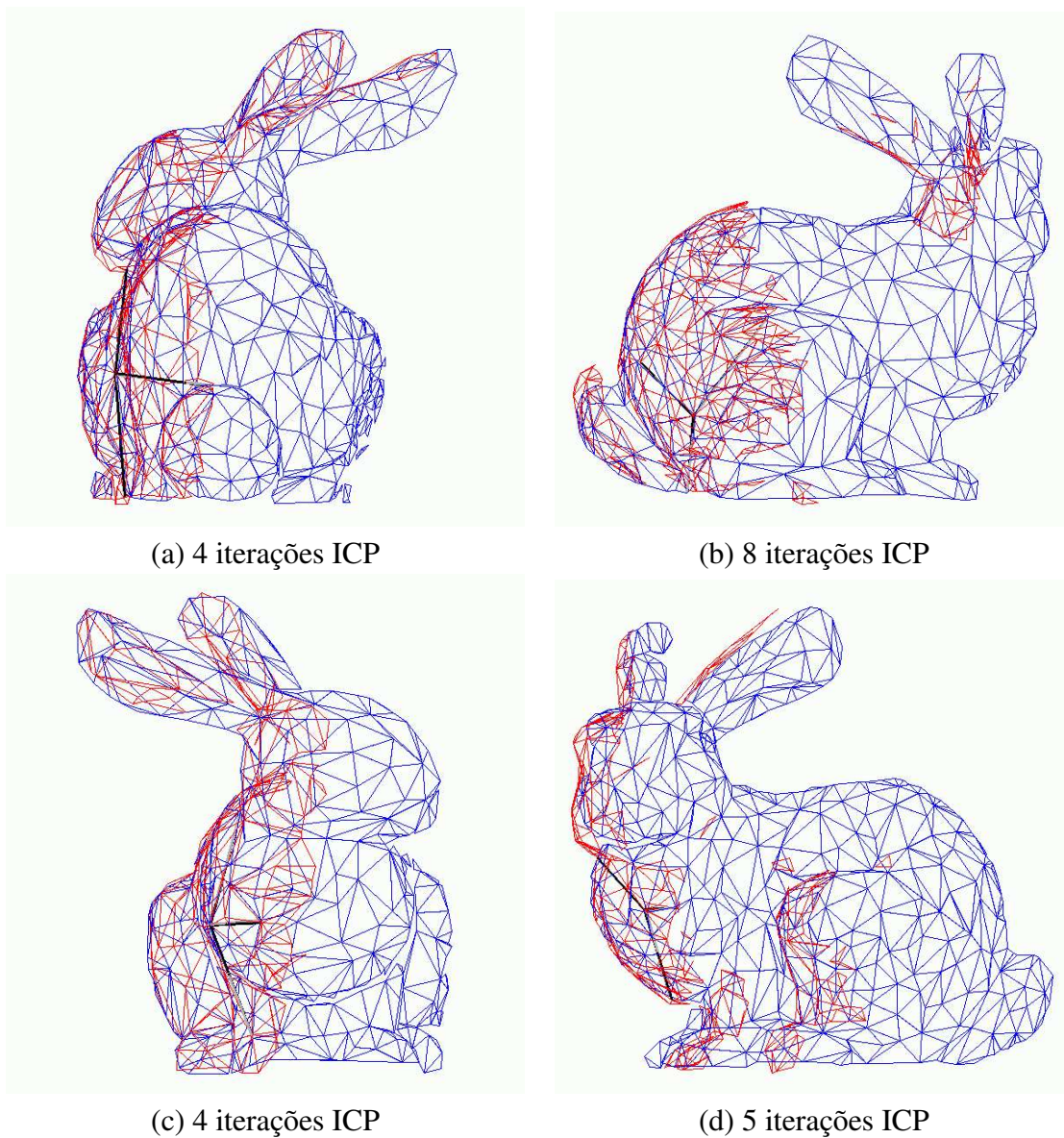


Fig. 5.6: Registro de pares de vistas: BUNNY.

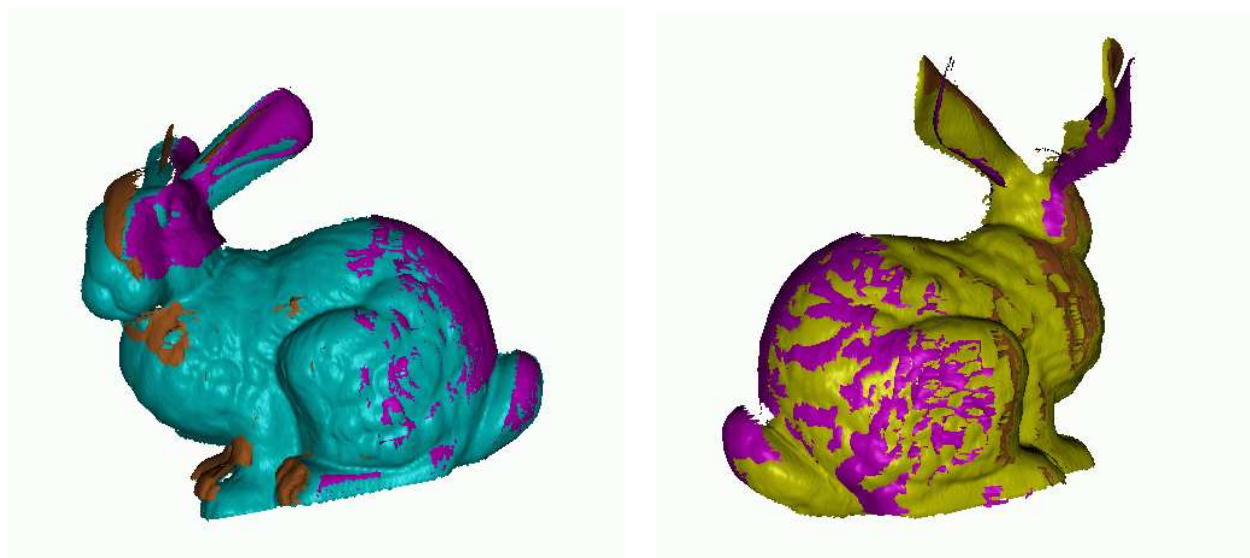


Fig. 5.7: Propagação de erros no registro de múltiplas imagens.

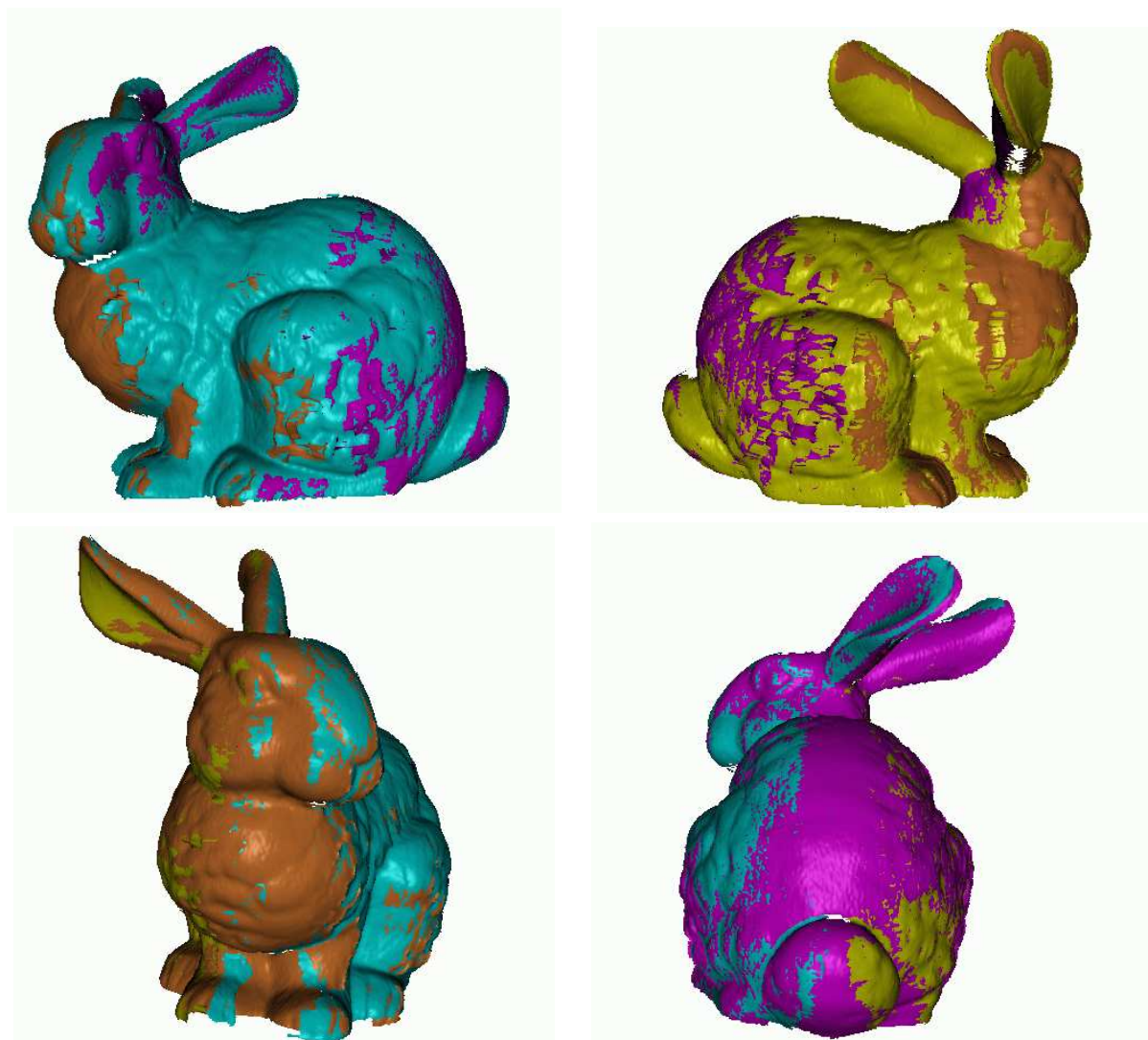
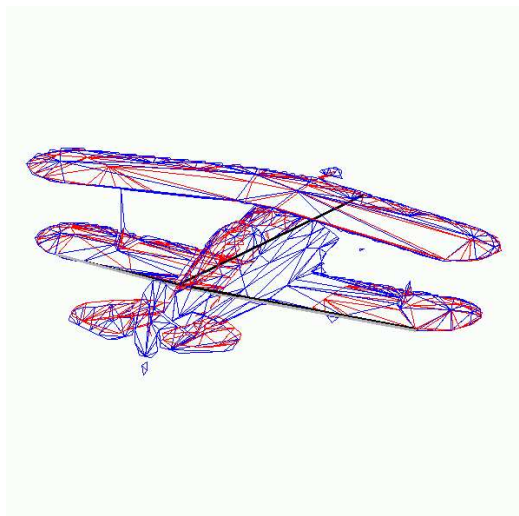
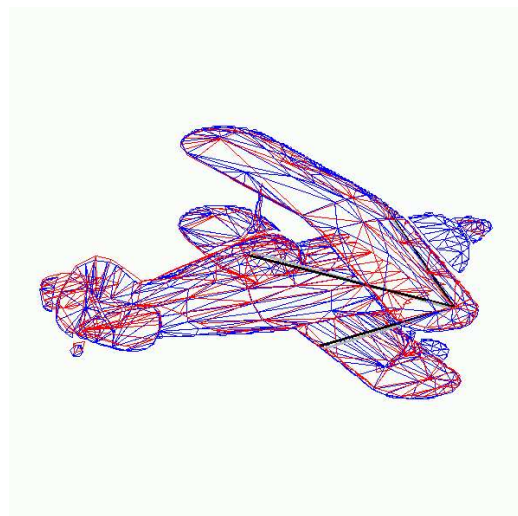


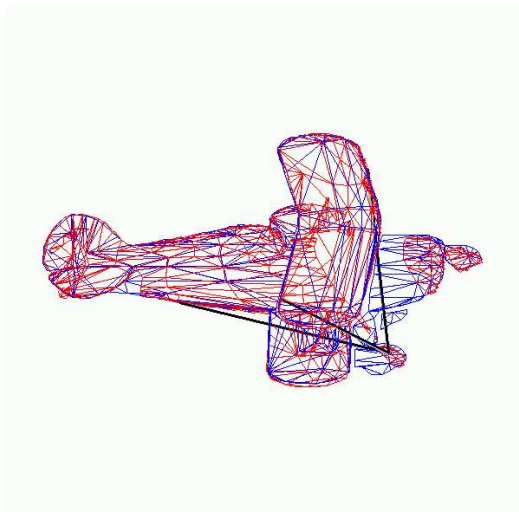
Fig. 5.8: Registro global de vistas: BUNNY.



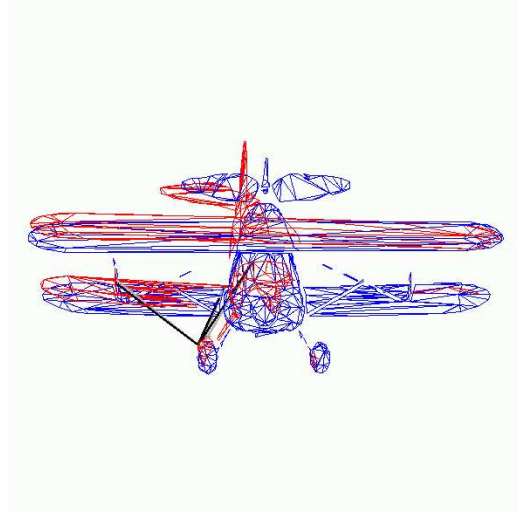
(a) 6 iterações ICP



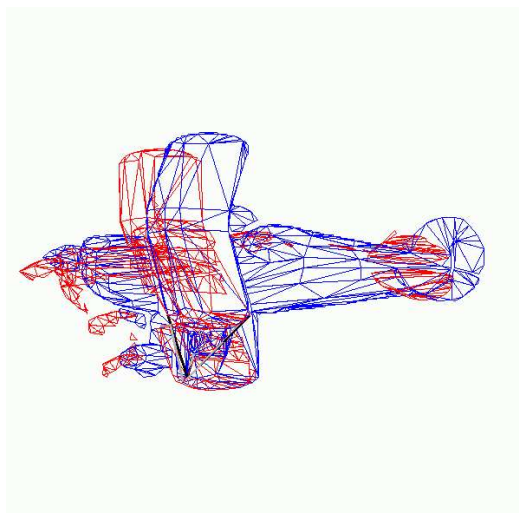
(b) 6 iterações ICP



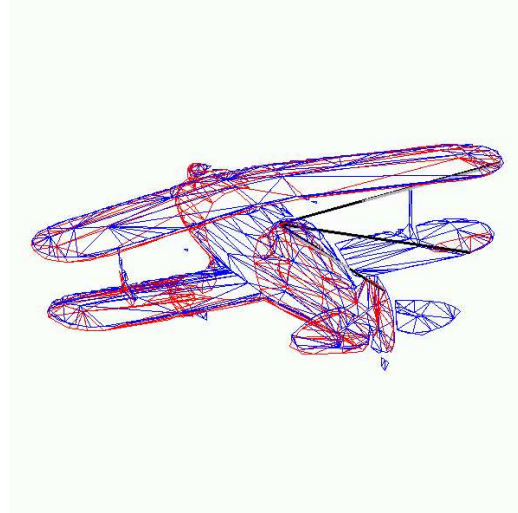
(c) 6 iterações ICP



(d) 8 iterações ICP



(e) 5 iterações ICP



(f) 5 iterações ICP

Fig. 5.9: Registro de pares de vistas: PITT-PLANE.



Fig. 5.10: Registro global de vistas: PITT-PLANE.

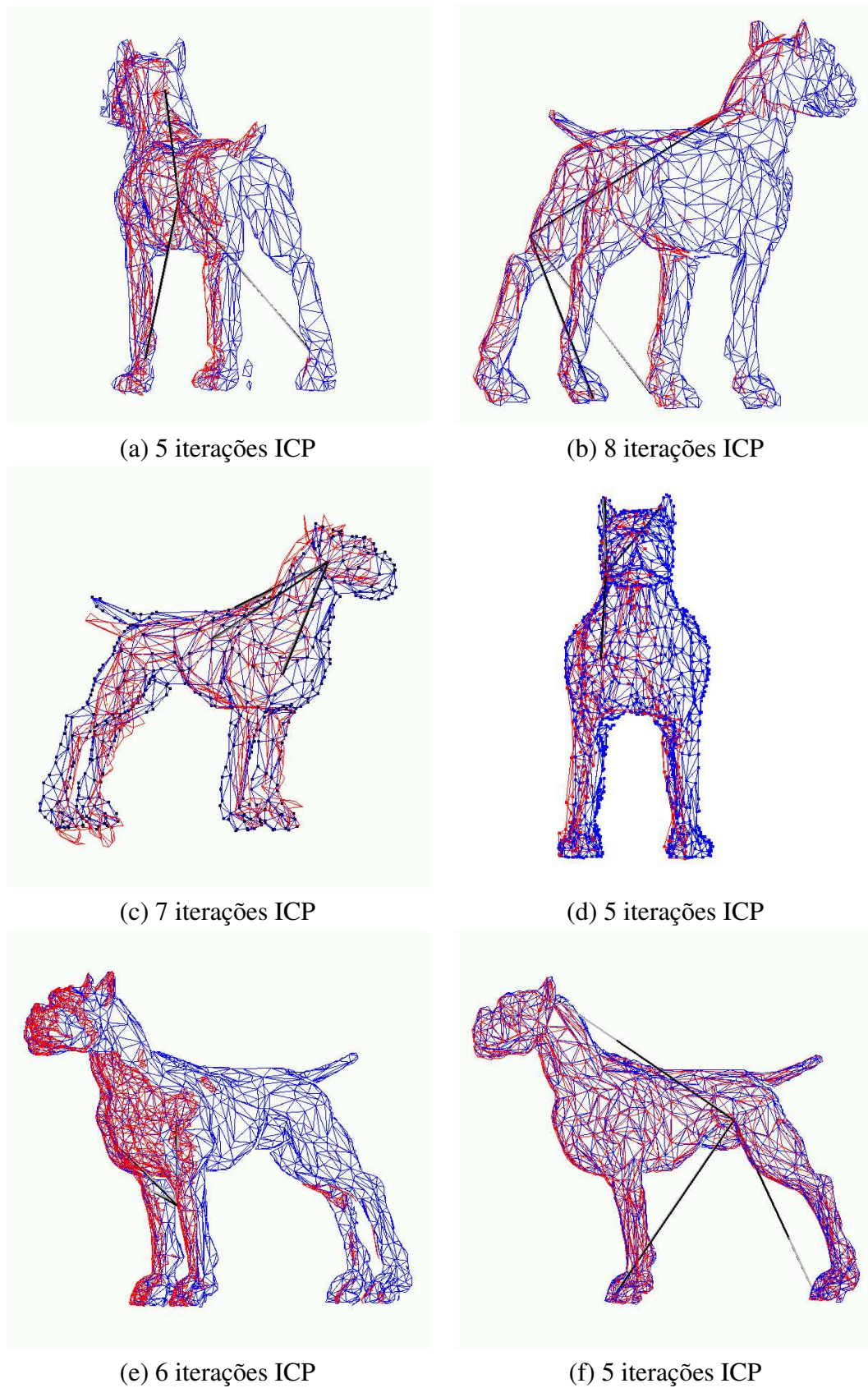


Fig. 5.11: Registro de pares de vistas: PITBULL.

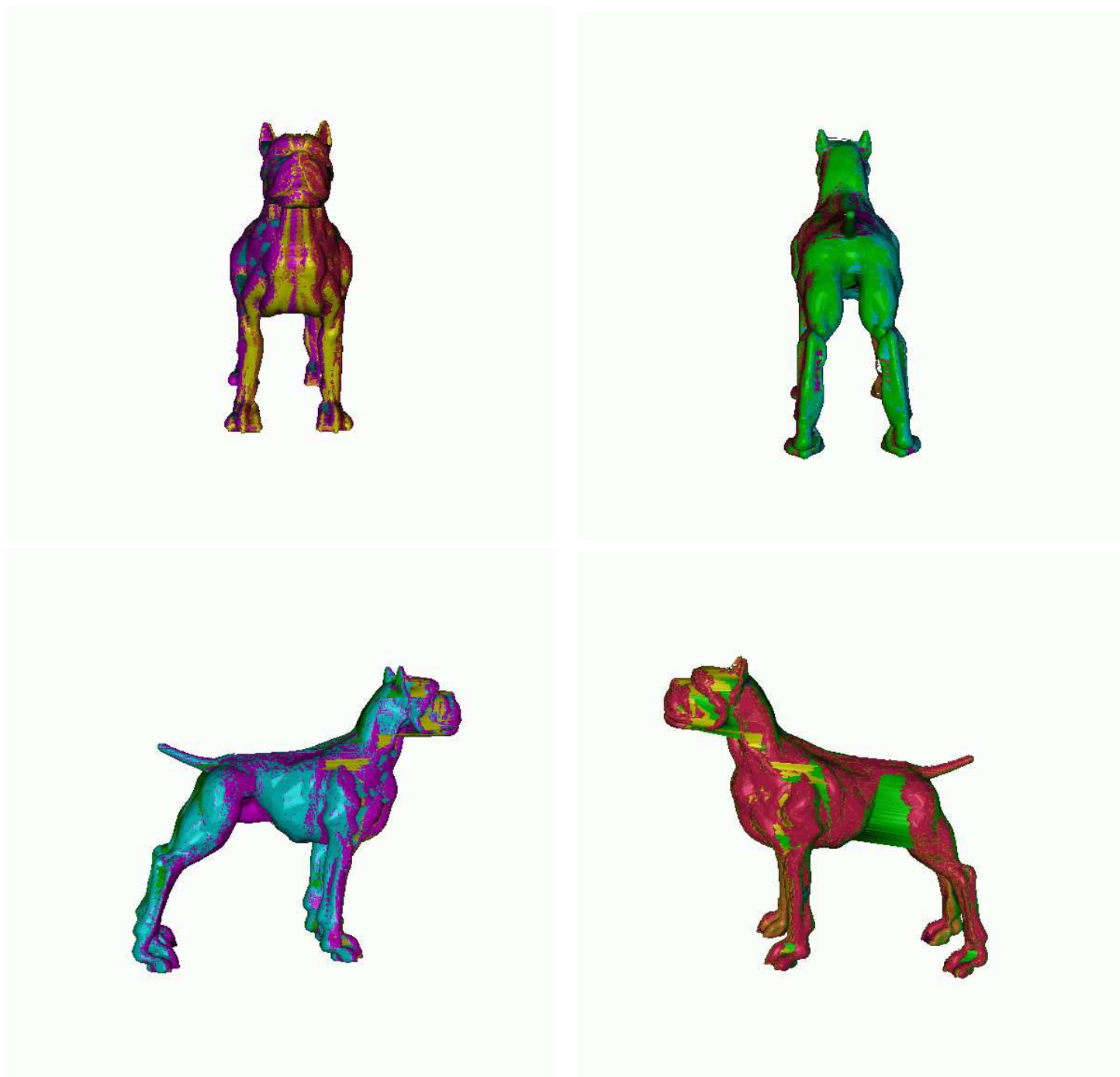


Fig. 5.12: Registro global de vistas: PITBULL.

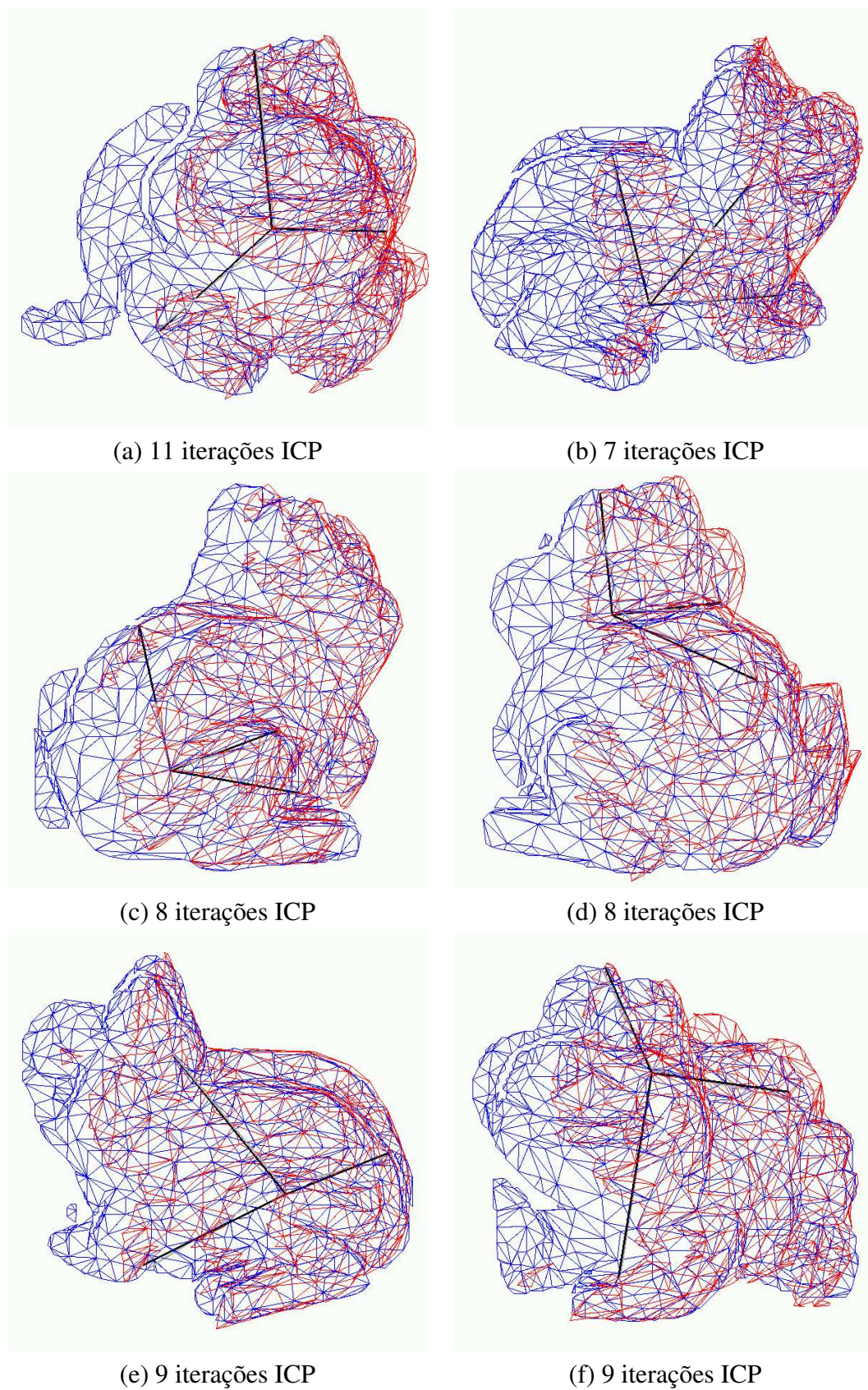


Fig. 5.13: Registro de pares de vistas: FROG.

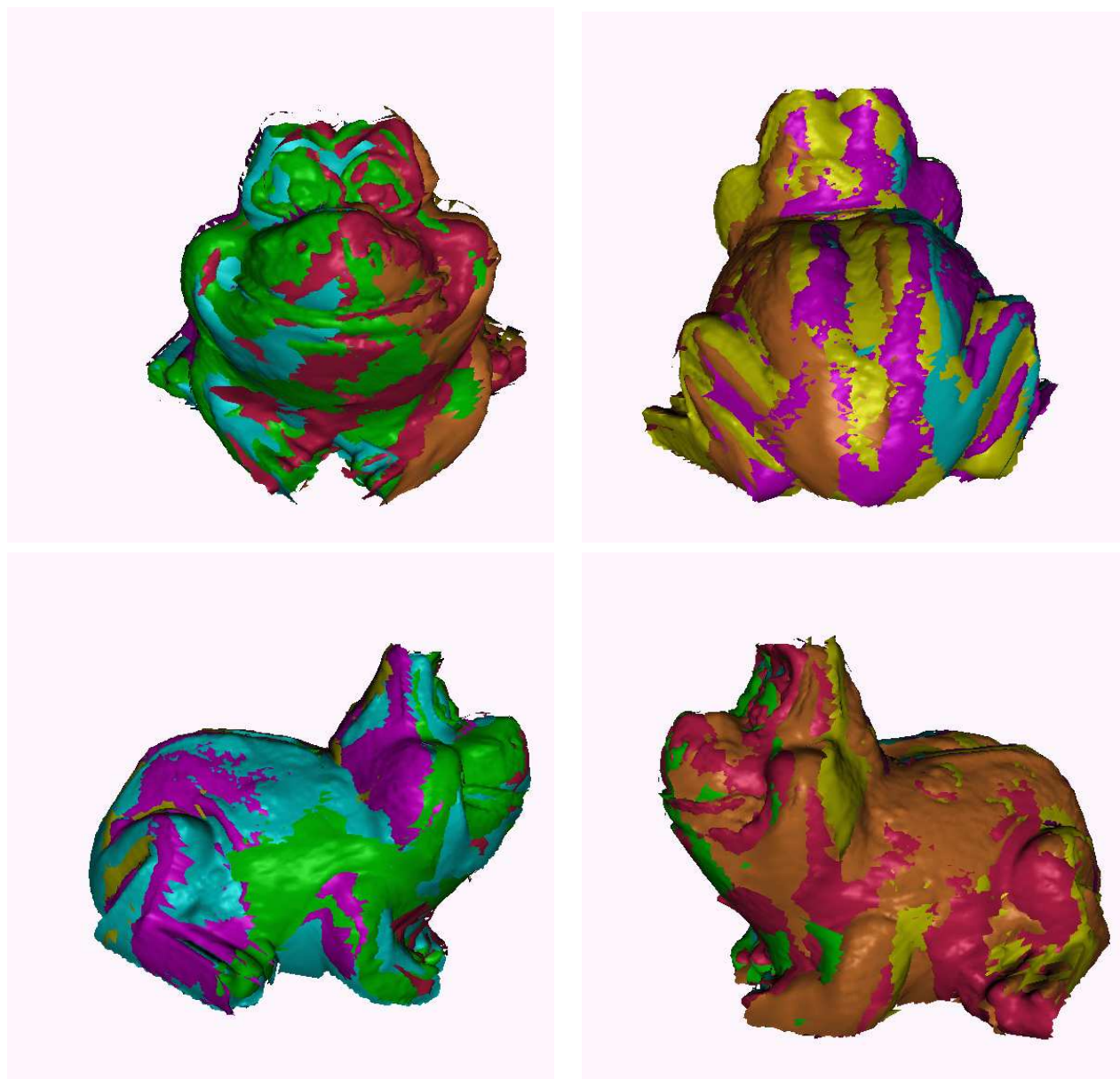


Fig. 5.14: Registro global de vistas: FROG.

5.3.3 Comparações

Planitz et al. [7, 8] elaborou um trabalho de análise e comparação de quatro técnicas que realizam pré-alinhamento e que foram descritas no Capítulo 2. Estes algoritmos são DARCES baseado em RANSAC (DBR), Casamento de Pontos orientados (CPO), Casamento de Histogramas Geométricos (CHG) e Casamento de Curvas Intrínsecas (CCI). Tendo como base esses resultados, imagens de teste similares às usadas por Planitz et al. [7, 8] são usadas para analisar o desempenho do algoritmo proposto. Em todos os casos foram usados os valores ϵ , δ , f_1 e f_2 sugeridos nos Capítulos 4 e 3. As imagens de teste (Figura 5.15(a)-(b)) variam em aquisição, tamanho, nível de detalhe, e características geométricas [7]. O par MACHINE foi obtido do repositório [51] e sua geometria apresenta quinas, arestas com descontinuidade de orientação e faces planares. Esse par foi escolhido para representar a geometria de uma cena considerada por Planitz et al, a qual possui as mesmas características geométricas de MACHINE. O par CLUB foi obtido do repositório [51], e foi escolhido para representar as geometrias com apenas uma pequena região contendo características geométricas distintas. Este par substitui o par DUCK usado por Planitz et al. para representar este tipo de geometrias. Observe-se que apenas o pescoço de CLUB se destaca como região distinta. O restante das imagens, ANGEL, BLUEDINO, HUB e BANANA, são as mesmas escolhidas por Planitz et al. para suas comparações.

O par ANGEL e o par BLUEDINO foram obtidos do site [35]. O primeiro representa geometrias com diferentes variações de curvatura e o segundo exibe uma região menor de sobreposição, limitada a apenas pequenas partes das duas pernas, cauda e apenas pequenos pedaços da cabeça e corpo. O par HUB foi criado para testar algoritmos de reconhecimento de objetos e foi obtido do repositório [51]. Este par de imagens tem um grande percentual de sobreposição, mas sua principal característica é que as imagens apresentam simetria em relação ao eixo z o que pode ocasionar falha no casamento em muitos procedimentos de alinhamento. O par BANANA também foi obtido do site [51]. As duas imagens também possuem uma grande região de sobreposição, mas a principal característica é a variação suave de curvatura em toda a superfície.

Segundo análise de Planitz et al. nos trabalhos [7, 8] os algoritmos CPO, CHG e CCI apresentam algumas restrições como mostra a Tabela 5.3. CPO não trabalha bem em casos que exibem simetria (par HUB) já que nestes casos haverá regiões localmente similares se repetindo na superfície. Como CPO (Seção 2.2.2) carece de avaliação global, se limitando a aceitar como correto qualquer casamento local, se tem como consequência um registro errado. O algoritmo proposto consegue realizar o pré-alinhamento como se pode observar na Figura 5.15(c) pois mesmo que um descritor-alvo obtido se ajuste ao descritor-referência, a avaliação global verifica o erro entre os pontos de sobreposição.

CHG (Seção 2.2.3) precisa de variações fortes de curvatura para realizar o pré-alinhamento em razão da construção dos histogramas geométricos os quais acumulam na mesma posição do histograma, os pontos que possuem mesma distância e mesma variação de vetores normais em relação a um ponto fixo. Desta forma, para duas regiões com variações suaves de curvaturas, ter-se-á o mesmo

	Simetria	Variação Suave de Curvatura	Variação Forte de Curvatura	Pouca região de sobreposição
CPO	NT			
CHG		NT		NT
CCI			NT	
DBR				
A proposta				

Tab. 5.3: Quadro Comparativo de Algoritmos (NT = Não trabalha bem).

histograma, o que ocasionará excessiva redundância (par BANANA e CLUB). Ou seja, para superfícies exibindo, na sua maioria, variações suaves de curvatura, CHG não conseguirá ser discriminante. A aplicação do algoritmo proposto foi bem sucedida pois consegue ser discriminante pelo fato de pegar um descritor globalmente mais abrangente e realizar a avaliação global (Figura 5.15(c)).

CCI precisa de variações suaves de curvatura para trabalhar bem em razão do pré-processamento de extração de curvaturas médias (Seção 2.3.2). Em casos como MACHINE e HUB o algoritmo não conseguiu realizar o pré-alinhamento devido às variações abruptas de curvatura que estas superfícies possuem. O algoritmo proposto nesta tese trabalha bem nesse tipo de geometrias pois arestas de descontinuidade e quinas são feições distintas que são sempre retidas pela malha decimada (Figura 5.15(c)).

DBR (Seção 2.2.1) destacou-se como o método que conseguiu fazer registro com sucesso em todos os casos devido à busca exaustiva. O algoritmo proposto reduz a complexidade combinatorial substituindo a imagem densa por malhas triangulares simplificadas sem perda de qualidade no registro. Adicionalmente o processo de busca se torna mais eficiente ao se considerar um descritor mais discriminante que o usado por DBR.

De novo, é usado o aplicativo Scanalyze[48], para avaliar os resultados do algoritmo proposto. O número de iterações necessárias para obter ajuste fino no registro foi 5, 7, 5, 4, 5 e 8 para as imagens ANGEL, HUB, CLUB, BANANA, BLUEDINO e MACHINE, respectivamente. Estes valores pequenos demonstram a bom desempenho da proposta apresentada neste trabalho para pré-alinhamento de uma classe de geometrias bastante diversificadas.

Fig. 5.15: (a) Imagem $\mathcal{R}_1$ (b) Imagem $\mathcal{R}_2$ (c) Pré-alinhamento.

5.4 Considerações Finais

Neste capítulo apresentam-se alguns resultados que confirmam nossa conjectura: aproximações fiéis à imagens de profundidades densas podem ser usadas para encontrar uma boa estimativa da transformação inicial de registro de duas imagens de profundidade parcialmente sobrepostas.

Os resultados deste capítulo mostram que a despeito da drástica redução do volume de dados, na ordem de 96%, e do espaço de busca, na ordem de 99%, o algoritmo proposto é ainda capaz de prover uma transformação rígida que faz o algoritmo ICP convergir em poucas iterações.

Comparações experimentais com outro métodos mostram que o algoritmo proposto possui habilidade de registro para uma grande diversidade de imagens.

Capítulo 6

Conclusões e Trabalhos Futuros

Este trabalho trata do pré-alinhamento de duas imagens de profundidade parcialmente sobrepostas de objetos rígidos. A partir da análise do problema e do estudo das principais técnicas do estado-da-arte, chegou-se a obter um conjunto de requisitos que garantem tanto a existência, no sentido de ser capaz de sempre extrair n pares de correspondências, quanto a eficiência nos procedimentos. Com base nestes requisitos, propõe-se um algoritmo eficiente de pré-alinhamento.

O principal problema do pré-alinhamento é o desconhecimento da correspondência entre as regiões de sobreposição; por tal motivo, a sua solução é baseada em tentativas e erros. Em cada tentativa assume-se um conjunto de correspondências (um par de descritores), a partir das quais uma transformação rígida é obtida. Após realizar todas as tentativas, as transformações hipotéticas passam por um processo de validação para se obter um único T . Para assegurar que um conjunto de correspondências seja determinado em ambas imagens, todos os possíveis descritores em R_1 e todos os possíveis descritores em R_2 devem ser considerados, o que gera um problema de combinatorial de alta complexidade. Mais precisamente, dois problemas combinatoriais são identificados: construir os descritores em cada imagem, individualmente, e o processo de busca. Soluções propostas para esses problemas são:

1. para aliviar a complexidade na construção de descritores, uma alternativa de solução é a substituição do conjunto denso de dados por uma representação reduzida da imagem de profundidade, e
2. para aliviar a complexidade na busca, além da redução do espaço de busca, uma alternativa é propor um descritor rico em medidas invariantes.

Várias propostas de redução de dados têm surgido nos últimos anos. No entanto, uma avaliação minuciosa das suas características levou a identificar algumas limitações, principalmente relacionadas com a perda de informação relevante para o registro o que leva à perda da robustez do método. A

conclusão foi que uma representação reduzida deve ser tal que retenha a informação essencial para o registro. Uma análise apurada permitiu elaborar mais três requisitos para satisfazer essa necessidade:

- a representação reduzida não deve estar limitada a certos domínios de superfícies (irrestrição à geometria da superfície).
- a representação reduzida deve reter vértices nas regiões de considerável variação de curvatura, mas também deve-se ter um conjunto de vértices com adequada densidade nas regiões com variações mais suaves de curvatura. Esta característica é necessária para garantir a obtenção de descritores na região de sobreposição, quando esta tiver uma geometria suave.

Assim sendo, propõe-se um novo algoritmo de pré-alinhamento. Atendendo à alternativa de solução 1, propõe-se o uso de uma malha triangular simplificada para reduzir a complexidade combinatorial na construção de descritores e contribuir com o quesito de eficiência. Para assegurar a existência de correspondências, deve-se considerar no processo de redução as seguintes características:

- obter uma densidade adequada de amostragem da malha para o tipo de geometria da imagem. A densidade é ajustada pelo parâmetro ϵ e leva em consideração a estimativa do índice de suavidade da superfície e o número de amostras.
- assegurar que hajam pares de vértices correspondentes na região de sobreposição, através do ajuste do parâmetro δ .
- usar todos os vértices da malha simplificada para assegurar que nenhuma informação essencial seja perdida.

Atendendo à alternativa de solução 2, propõe-se a construção de um descritor altamente discriminante em cada vértice da malha, para diminuir a complexidade combinatorial na busca. A discriminância do descritor está vinculada com sua configuração tridimensional, e as propriedades geométricas locais (curvaturas) e globais (distâncias) que possui.

Sobre o quesito de existência, pode-se concluir em relação ao novo algoritmo proposto

1. Extração de pelo menos um descritor-referência: Dos experimentos realizados verificou-se que atentando para os valores ϵ sugeridos sempre se consegue obter pelo menos um descritor-referência.
2. Extração de pelo menos um descritor-alvo: Dos experimentos realizados e seguindo a sugestão para os valores ϵ , δ , f_1 e f_2 , constatou-se que sempre foi possível obter um descritor-alvo.

E as seguintes afirmações podem ser colocadas em relação ao quesito de eficiência:

1. Eficiência na construção do descritor-referência: a construção da malha simplificada e a extração de curvaturas locais são de baixo custo pois é feito em um número reduzido de vértices. O fator de redução da malha, em relação à amostragem densa original da imagem de profundidade, foi aproximadamente de 4:100. Além disso, a construção do triedro é feito por cálculos simples.
2. Eficiência na busca de um descritor-alvo: o estabelecimento de correspondências segue um processo de busca exaustiva, porém eficiente, devido ao tamanho reduzido dos dados da malha simplificada e à capacidade discriminante do triedro. Os experimentos comprovam que a redução do espaço de busca é de mais de 99%.
3. Eficiência no processo de filtragem: as transformações candidatas são validadas usando um conjunto reduzido de dados que representa bem a forma global da superfície. Isso faz o processo de filtragem eficiente. Pode-se estimar esta redução, na ordem de 96%, que é a mesma ordem de redução dos dados. Devemos ainda comentar que apesar da simplificação introduzir um *overhead* de, em média, 30 s. no processo de pré-alinhamento, este é bem compensado pelos ganhos em eficiência citados.

Adicionalmente propõem-se como trabalhos futuros:

- investigar algoritmos de decimação de malha que sejam adequados para o problema de pré-alinhamento. Os requisitos para um algoritmo de simplificação não são atendidos pelos algoritmos de simplificação do nosso conhecimento. De fato, tradicionalmente, os algoritmos de simplificação têm focado na redução do número de vértices da malha e na preservação da forma geral. Já no caso do registro, existe mais o requisito de densidade amostral, como já foi anteriormente colocado.
- propor um algoritmo de integração de imagens de profundidade para, após o alinhamento, construir um modelo único 3D a partir das n imagens de profundidade registradas.

Referências Bibliográficas

- [1] A.Johnson and M.Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:433–448, 1999.
- [2] A.P. Ashbrook, R.B. Fisher, N. Werghi, and C. Robertson. Aligning arbitrary surfaces using pairwise geometric histograms. In *Proceedings of Noblesse Workshop on Non-linear Model Based Image Analysis*, pages 103–108, 1998.
- [3] B.Curless and M. Levoy. A volumetric method for building complex models from range images. In *Computer Graphics Proceedings*, pages 303–312, New Orleans, Louisiana, Agosto 1996.
- [4] O.R.P. Bellon. *Imagens de Profundidade: Segmentação e Representação por superfícies Planares*. Tese de Doutorado, Univesidade Estadual de Campinas, Campinas, São Paulo, Brasil, 1997.
- [5] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–256, 1992.
- [6] G. Blais and M.D.Levine. Registering multiview range data to create 3d computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, 1995.
- [7] B.M.Planitz, A.J. Maeder, and J.A. Williams. The correspondence framework for 3d surface matching algorithms. *Computer Vision and Image Understanding*, 97:347–383, 2005.
- [8] B.M.Planitz, A.J. Maeder, and J.A. Williams. Using the correspondence framework to select surface matching algorithms. In *Proceedings of the Australian Pattern Recognition Society (APRS) Workshop on Digital Image Computing*, pages 85–90, 2005.
- [9] B.Zitová and J.Flusser. Image registration methods : a survey. *Image and Vision Computing*, 21:977–1000, 2003.
- [10] C.Fermuller and Y.Aloimonos. Which shape from motion? In *Proceedings of the International Conference of Computer Vision*, pages 689–695, 1998.

- [11] Chu-Song Chen, Yi-Ping Hung, and Jen-Bo Cheng. RANSAC-based DARCES: A new approach to fast automatic registration of partially overlapping range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1229–1234, 1999.
- [12] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, pages 145–155, 1992.
- [13] C.Pradalier and S.Sekhvat. Concurrent matching, localization and map building using invariant features. In *Proceedings of IEEE/RSJ Int.Conf.of Intelligent Robots and System*, pages 514 – 520, 2002.
- [14] J.E. Cryer, P.S. Tsai, and M. Shah. Integration of shape from shading and stereo. *Pattern Recognition*, 28(7):1033–1043, 1995.
- [15] G. Dalley and P. Flynn. Range image registration : A software platform and empirical evaluation. In *3th International Conference on 3-D Digital Imaging and Modeling*, pages 246–253, 2001.
- [16] M.P. do Carmo. *Geometry of Curves and Surfaces*. Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1976.
- [17] C. Dorai and et.al. Registration and integration of multiple object views for 3D model construction. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(1):83–89, 1998.
- [18] R.C.Gonzalez e R.E.Woods. *Digital Image Processing*. Addison-Wesley, 1992.
- [19] R.Zhang et al. Analysis of shape from shading techniques. In *Proceedings of the International Conference of Computer Vision and Pattern Recognition*, pages 377–384, 1994.
- [20] H. Gagnon and et.al. Registration of multiple range views for automatic 3-d model building. In *IEEE Conf. on Vision and Pattern Recognition*, pages 581–586, New York, USA, 1994.
- [21] M. Garland and P.Heckbert. Simplification using quadric error metrics. In *Computer Graphics*, volume 31, pages 209–216, New York, USA, 1997.
- [22] Michael Garland. *Simplification Using Quadric Error Metrics*. PhD thesis, School of Computer Science Carnegie Mellon University, New York, USA, 1999.
- [23] J.M. Gomes and L. Velho. *Computação Gráfica*, volume 1. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, 1998.
- [24] A. Hoover. *The Space Envelope Representation for 3D Scenes*. PhD thesis, University of South Florida, <http://marathon.csee.usf.edu/hoover>, 1994.

- [25] B.K.P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society A*, 4(4):629–642, 1987.
- [26] Pavel Krsek, Tomáš Pajdla, and Václav Hlaváč. Differential invariants as the base of triangulated surface registration. *Computer Vision Image Understanding*, 87:27–38, 2002.
- [27] O.R.P. Bellon L. Silva and K.L. Boyer. Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms. *IEEE Transactions on Pattern and Machine Intelligence*, 27:762–776, 2005.
- [28] M.A.Fischler and R.C.Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. Association for Computing Machinery*, 24(6):381–395, 1981.
- [29] T. Masuda and N. Yokota. Robust method for registration and segmentation of multiple range images. *Computer Vision and Image Understanding*, 61(3):295–307, 1995.
- [30] M.Clerc and S.Mallat. Shape from texture through deformations. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 1, pages 405–410, 1999.
- [31] M.Ringer and R.D.Morris. Robust automatic feature detection and matching between multiple images. Technical report, Cambridge University Engineering Department, http://www.riacs.edu/navroot/Research/TR_pdf/TR_01.27.pdf, 2002.
- [32] Peter Neugebauer. Geometrical cloning of 3d objects via simultaneous registration of multiple range images. In *International Conference on Shape Modeling and Applications (SMA '97)*, Japan, 1997.
- [33] Next Best View Problem. <http://wwwx.cs.unc.edu/~sud/courses/comp258/NBV.html>.
- [34] F.M.A. Nogueira. *Geração Automática de Mapas de Disparidade em Visão Estéreo*. Dissertação de Mestrado, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil, 1997.
- [35] Range images. <http://sampl.eng.ohio-state.edu/sampl/data/3DDB/RID>.
- [36] David P.Luebke. A developer's survey of polygonal simplification algorithms. *IEEE Computer Graphics and Applications*, 21:24–35, 2001.
- [37] P.S.Heckbert and M.Garland. Survey of polygonal surface simplification algorithms. Technical report, Carnegie Mellon University, 1997.

- [38] K. Pulli. *Surface Reconstruction and Display from Range and Color Data*. PhD thesis, University of Washington, 1997.
- [39] Usf range image database. <http://marathon.csee.usf.edu/range/DataBase.html>.
- [40] R.H.Zhang. Mesh features and segmentation. Lecture 17 - CMPT 469: Geometric Modeling in Computer Graphics, Março 2005.
- [41] C. Rocchini, P. Cignoni, P. Pingi, and R. Scopigo. A low cost 3d scanner based on structured light. In *EUROGRAPHICS*, volume 20, 2001.
- [42] Azriel Rosenfeld. *Advances in Computer*, volume 27. Academic Press Inc., San Diego, 1988.
- [43] Gerhard Roth. Registering two overlapping range images. In *Second International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 191–200, Ottawa, 1999.
- [44] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Third International Conf. on 3D Digital Imag. and Modeling*, pages 145–152, 2001.
- [45] Szymon Rusinkiewicz. Estimating curvatures and their derivatives on triangle meshes. In *3DPVT'04: Proceedings of the 3D Data Processing, Visualization, and Transmission, 2nd International Symposium on (3DPVT'04)*, pages 486–493, Washington,DC,USA, 2004.
- [46] A. Sappa. *Automatic generator of 3D geometric model from Range Images*. Tese de Doutorado, Universidade de Catalunha, Catalunha, Barcelona, Espanha, 1999.
- [47] S.Belongie and J.Malik. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:509–522, 2002.
- [48] Scanalyze, a system for aligning and merging range data. <http://graphics.stanford.edu/software/scanalyze>.
- [49] M. Soucy and D. Laurendeau. General surface approach to the integration of a set of range views. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 17(4):334–358, 1995.
- [50] Range images. <http://graphics.stanford.edu/data/3Dscanrep>.
- [51] Range images. <http://range.informatik.uni-stuttgart.de/htdocs/html>.
- [52] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Computer Graphics Proceedings*, pp. 311–318. Florida, 1994. Siggraph 94.

- [53] J. Vanden and L. Van Gool. Automatic crude patch registration. *Computer Vision Image Understanding*, 87, 2002.
- [54] S. Weik. Registration of 3d partial surface models using luminance and depth information. In *Proceedings of the Int. Conf. on Recent Advances in 3-D Imaging and Modeling*, Canada, 1997.
- [55] Wikipedia: Enciclopédia virtual. <http://pt.wikipedia.org/wiki/>.
- [56] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *International Symposium on 3D Data Processing Visualization and Transmission*, pages 24–36, 2002.
- [57] Z.Tu, S.Zheng, and A.Yuille. Shape matching and registration by data-driven em. *Computer Vision and Image Understanding*, 109:290–304, 2008.

Apêndice A

Outros Resultados

Neste apêndice apresentam-se resultados adicionais gerados pelo algoritmo proposto. Esses resultados vêm se somar aos já apresentados no Capítulo 5. Mais três conjuntos de imagens são considerados: TELETUBBY e BIRD obtidos do site referenciado em [35] e DINO do site citado em [51]. As imagens dos primeiros dois conjuntos são apresentadas na Figura A.1 e as do terceiro conjunto são exibidas na Figura A.2. A Tabela A.1 apresenta os dados antes e depois da simplificação da malha construída a partir das amostras das imagens de profundidade.

Imagens	Amostras ($\hat{N}$)	Vértices (N)	% Redução
TELETUBBY0	6841	606	91,14
TELETUBBY60	4361	409	90,62
TELETUBBY120	3822	333	91,29
TELETUBBY180	6764	483	92,86
TELETUBBY240	5199	470	90,96
TELETUBBY300	5175	501	90,32
BIRD0	9115	326	96,42
BIRD60	6309	289	95,42
BIRD120	10502	319	96,96
BIRD180	9464	331	96,50
BIRD240	8263	289	96,50
BIRD300	10585	305	97,12
DINO25	16562	358	97,84
DINO28	11895	399	96,65
DINO31	14608	356	97,56
DINO34	16036	454	97,17
DINO37	11879	483	95,93
DINO40	14736	391	97,35

Tab. A.1: Redução do volume de dados.

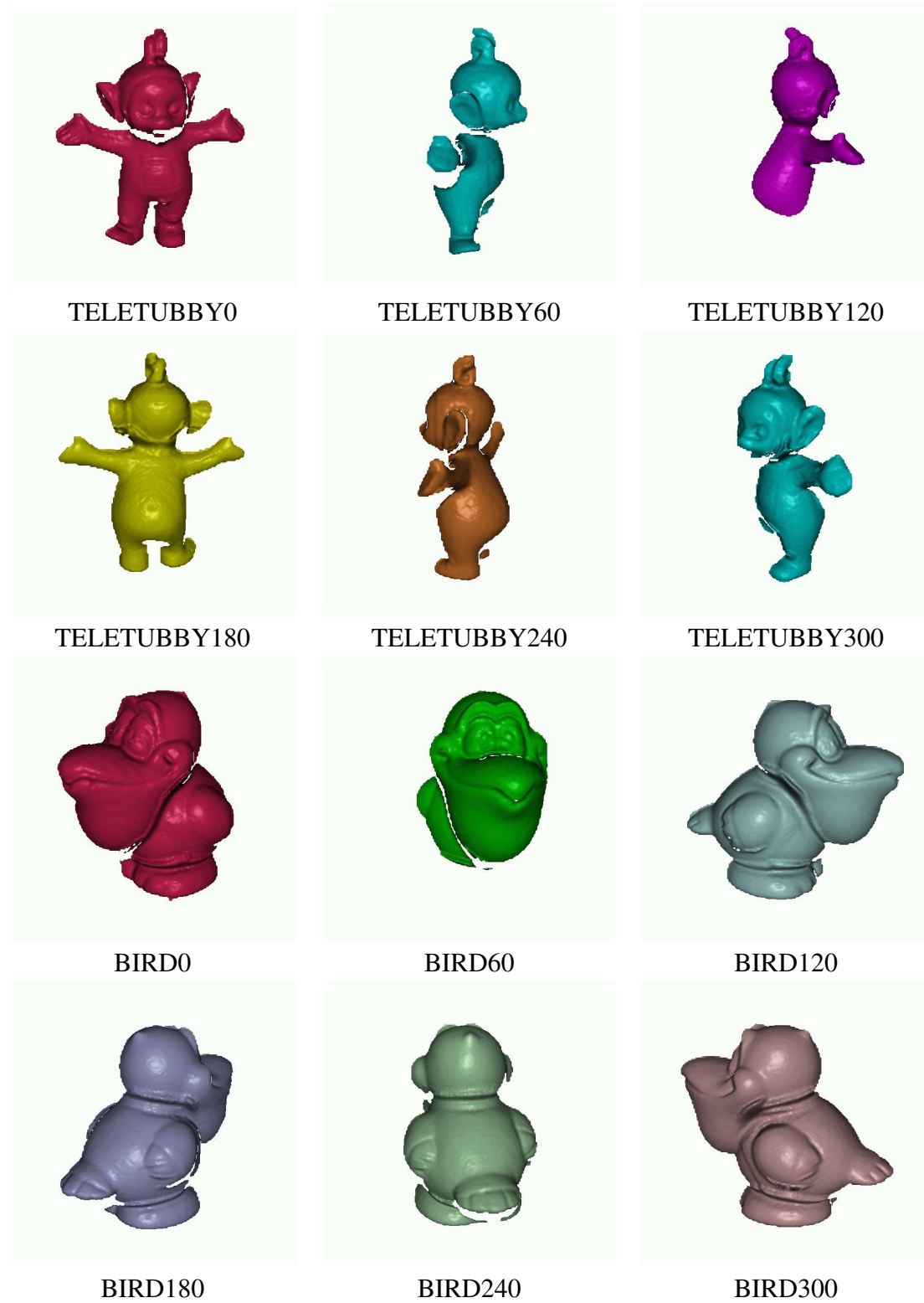


Fig. A.1: Imagens de Teste: TELETUBBY [35] e BIRD [35].

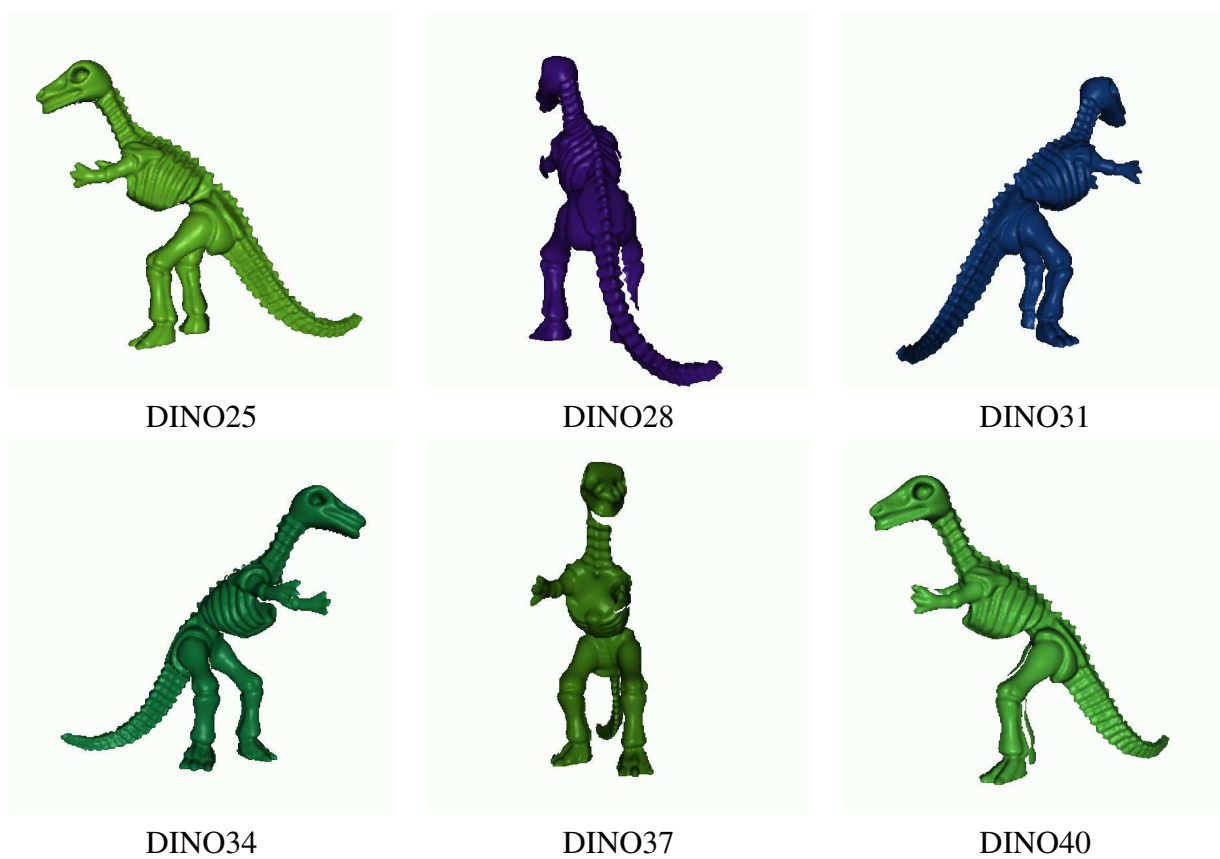


Fig. A.2: Imagens de Teste: DINO [51]

Nas Figuras A.3, A.4 e A.5 mostram-se os pares TELETUBBY320-TELETUBBY0, BIRD0-BIRD60 e DINO31-DINO34, respectivamente. Em todas as figuras, as imagens de profundidade originais são mostradas separadamente nos itens (a) e (b). Em (c) exibe-se o par de malhas triangulares simplificadas, correspondentes às imagens, registrado pela transformação rígida T_0 , obtida pelo algoritmo proposto. Em (d), a transformação T_0 é aplicada em malhas finas para efeito de comparação com o registro fino. Em (e), é apresentado o registro final das imagens. Como mencionado no Capítulo 5 pode-se avaliar visualmente a qualidade do pré-alinhamento pelo intercalamento das superfícies, como sugerido por Flynn [15].

As Figuras A.6, A.8 e A.10 mostram os seis pares de imagens pertencentes aos conjuntos TELETUBBY, BIRD e DINO, respectivamente. E finalmente, a partir das transformações de pré-alinhamento dos seis pares de cada conjunto foi realizado um registro global com uso do procedimento descrito na Seção 5.3.2 (Figuras A.7, A.9 e A.11).



(a) TELETTUBY320



(b) TELETUBBY0

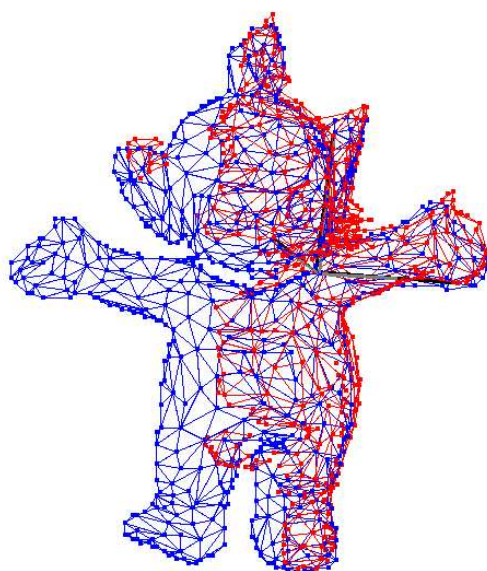
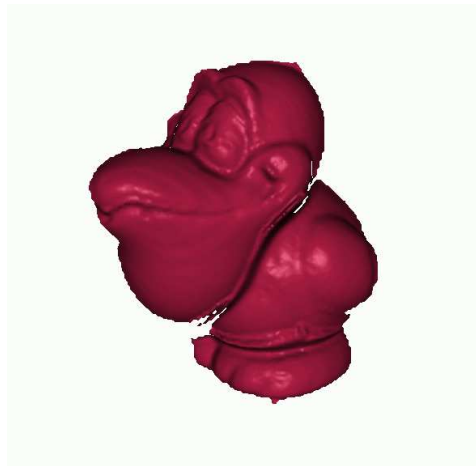
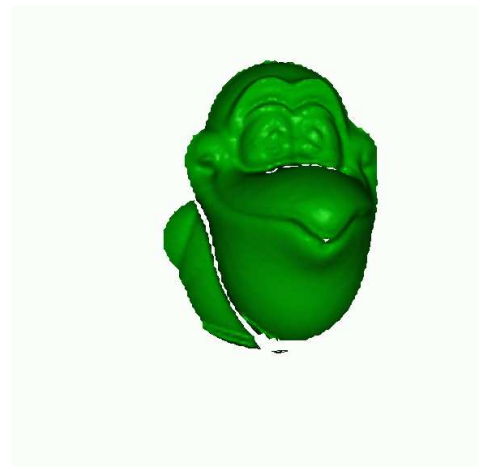
(c) Pré-alinhamento de Malhas simplificadas (T_0)(d) Pré-alinhamento de Malhas Finas (T_0)(e) Registro após refinamento de T_0 (ICP - 7 iterações)

Fig. A.3: Pré-alinhamento e Registro Fino (TELETUBBY).



(a) BIRD0



(b) BIRD60

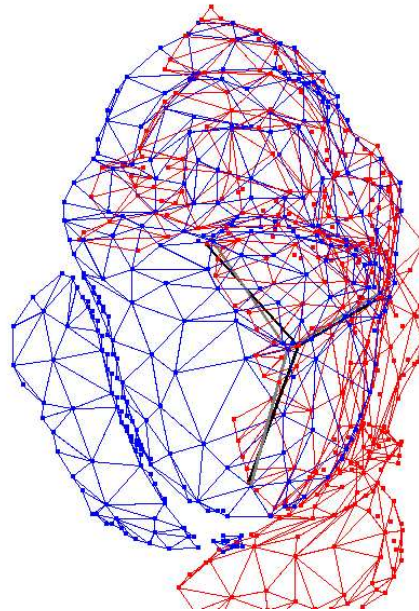
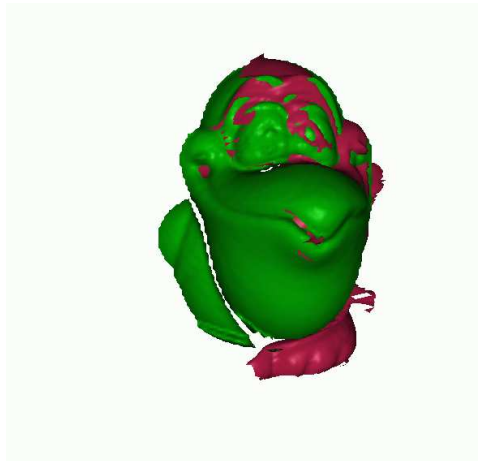
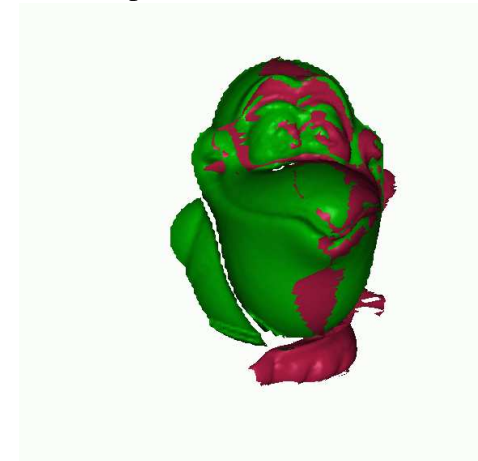
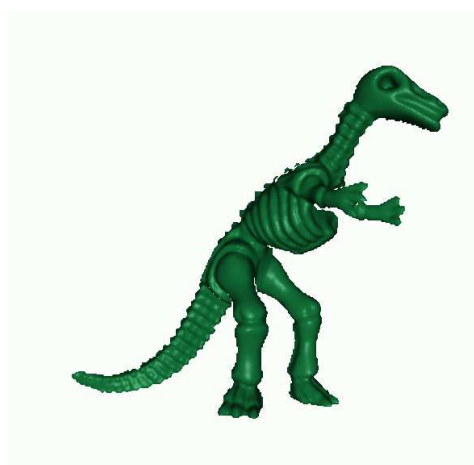
(c) Pré-alinhamento de Malhas simplificadas (T_0)Pré-alinhamento de Malhas Finas (T_0)(e) Registro após refinamento de T_0 (ICP - 7 iterações)

Fig. A.4: Pré-alinhamento e Registro Fino (BIRD).



(a) DINO31



(b) DINO34

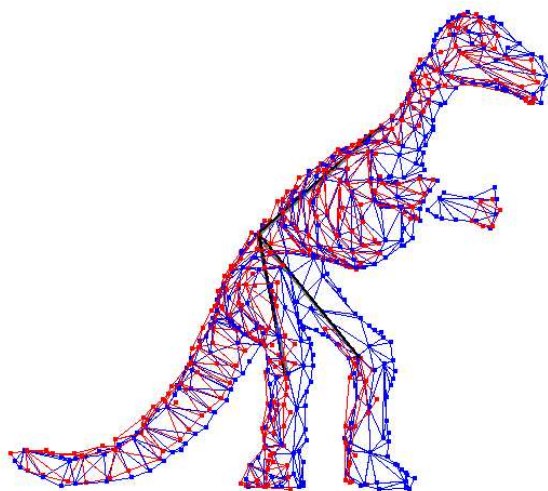
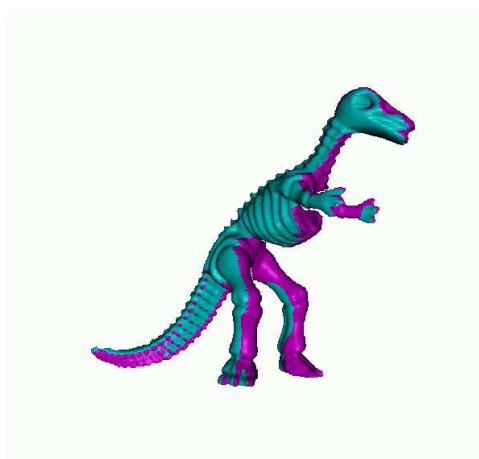
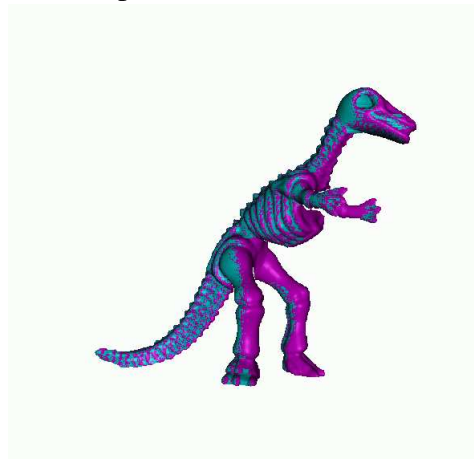
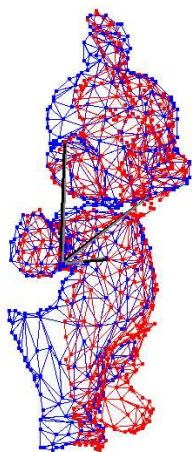
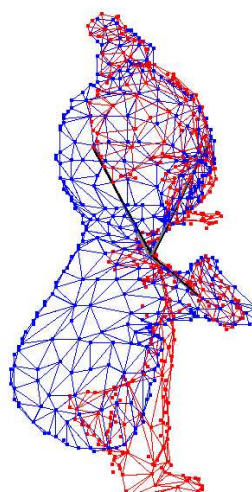
(c) Pré-alinhamento de Malhas simplificadas (T_0)Pré-alinhamento de Malhas Finas (T_0)(e) Registro após refinamento de T_0 (ICP - 4 iterações)

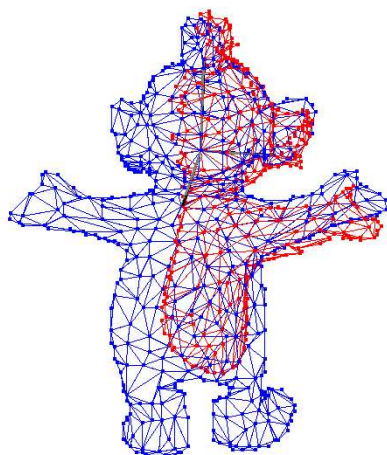
Fig. A.5: Pré-alinhamento e Registro Fino (DINO).



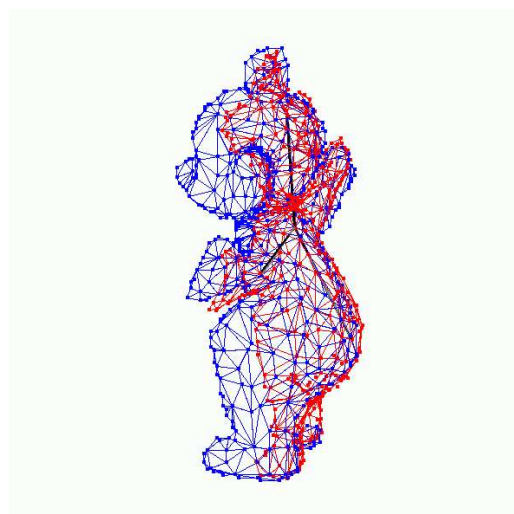
(a) 5 iterações ICP



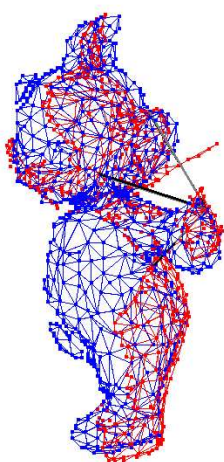
(b) 9 iterações ICP



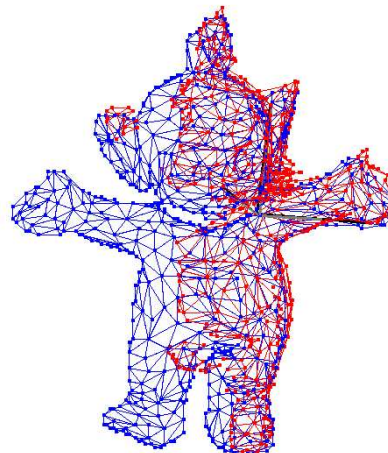
(c) 8 iterações ICP



(d) 6 iterações ICP



(e) 5 iterações ICP



(f) 7 iterações ICP

Fig. A.6: Registro de pares de vistas: TELETUBBY.

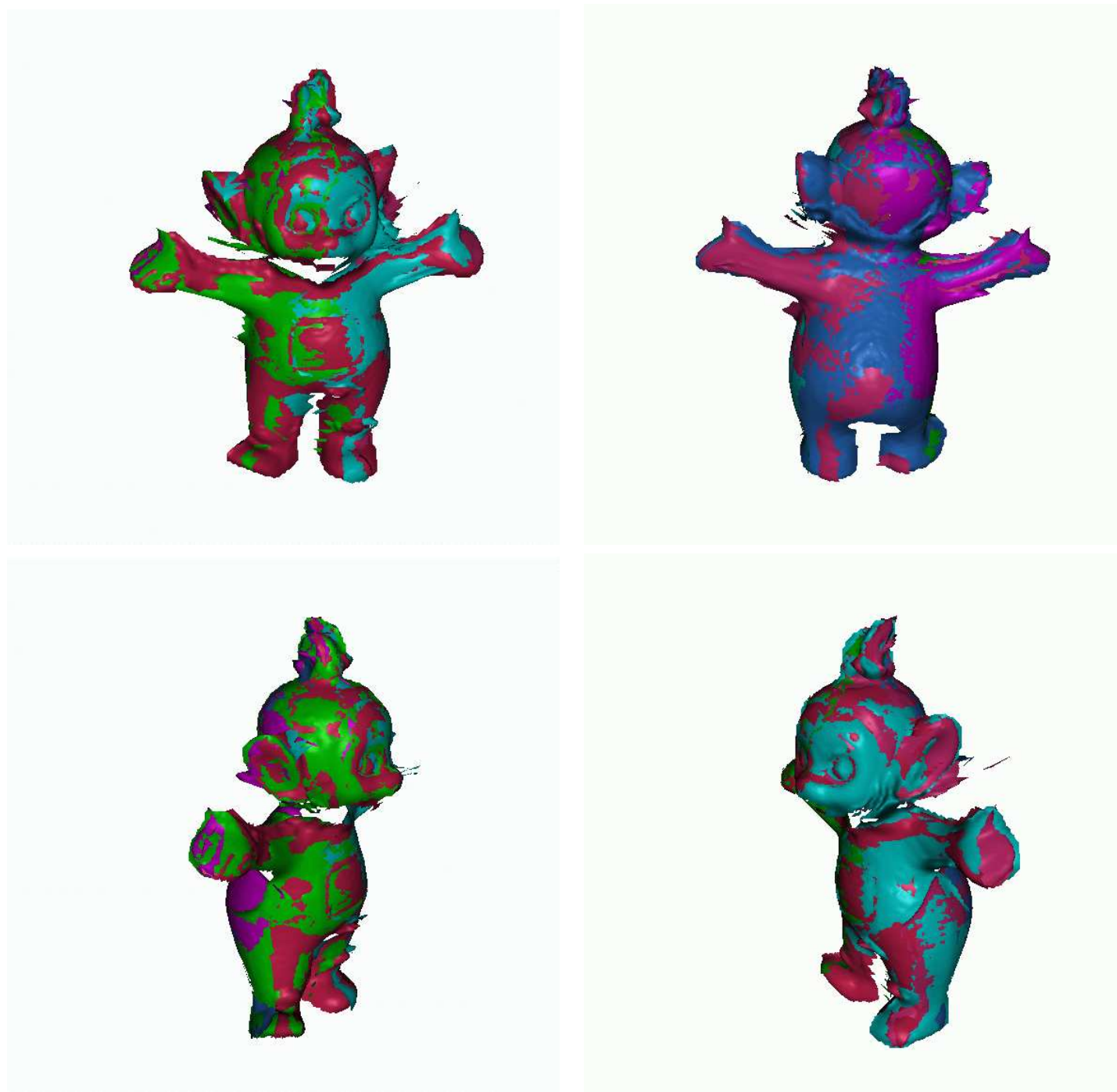
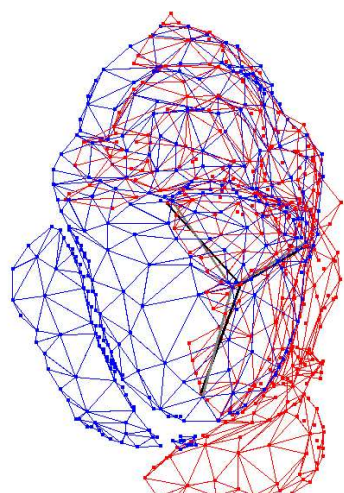
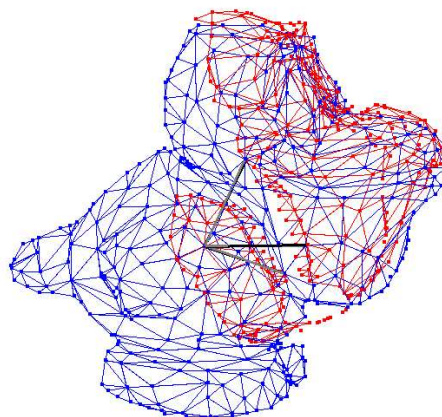


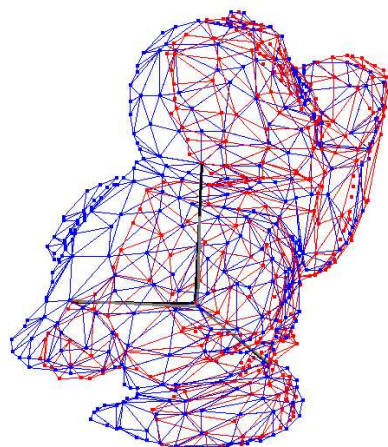
Fig. A.7: Registro global de vistas: TELETUBBY.



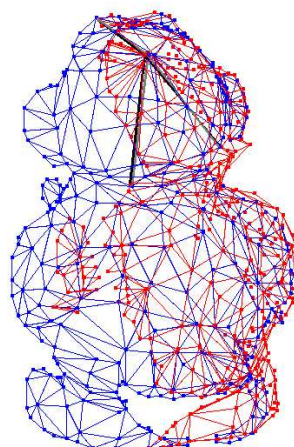
(a) 7 iterações ICP



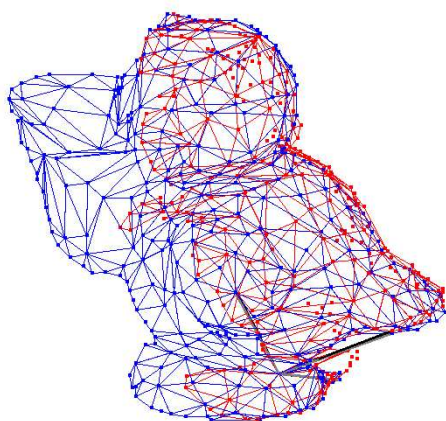
(b) 7 iterações ICP



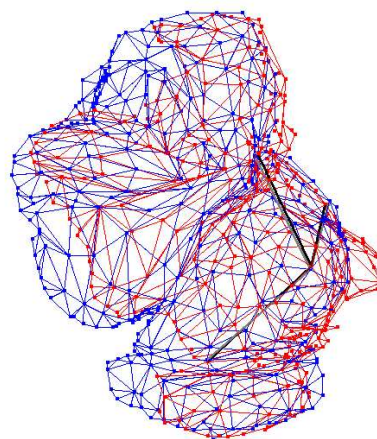
(c) 5 iterações ICP



(d) 9 iterações ICP



(e) 5 iterações ICP



(f) 6 iterações ICP

Fig. A.8: Registro de pares de vistas: BIRD.

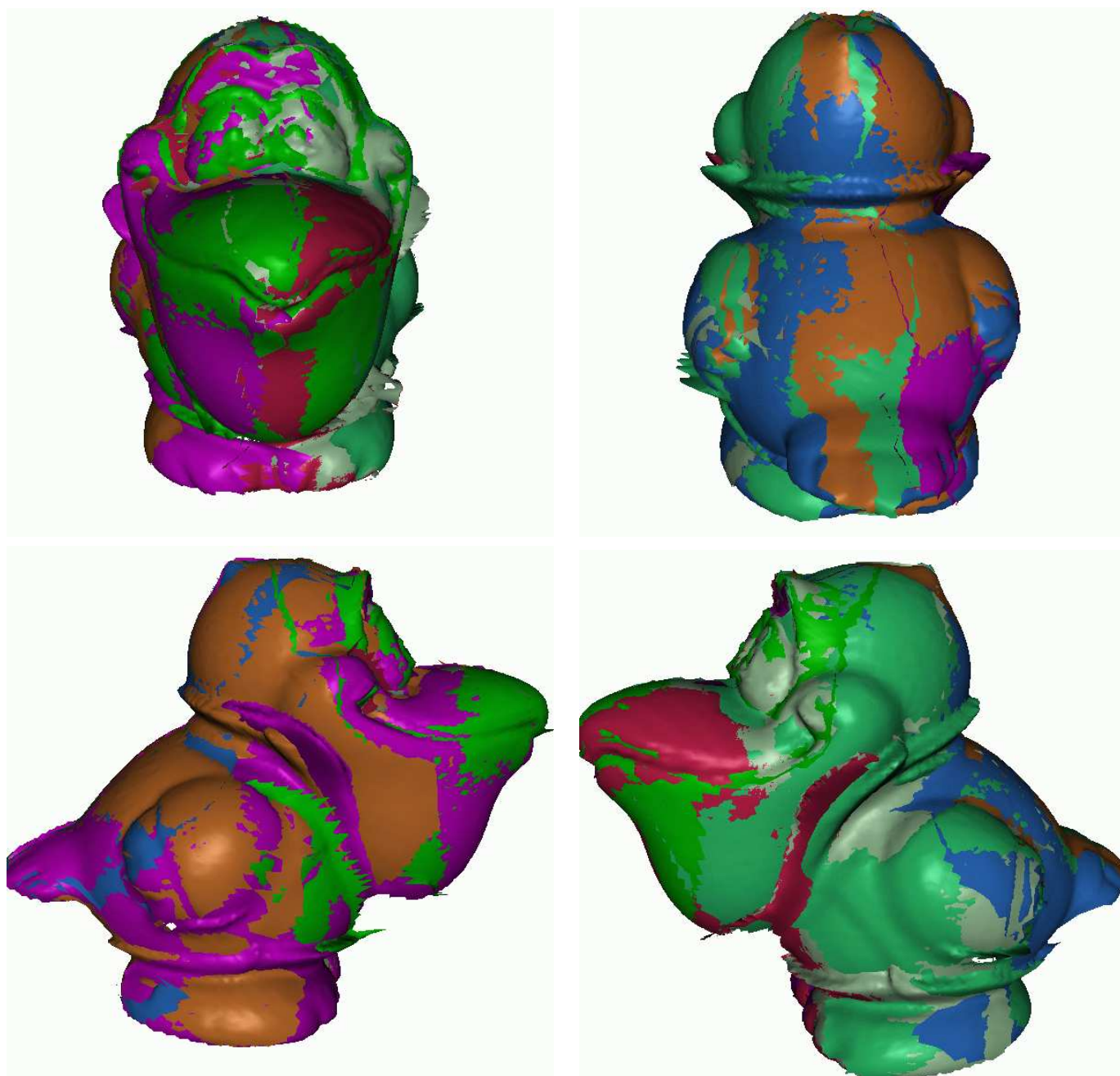
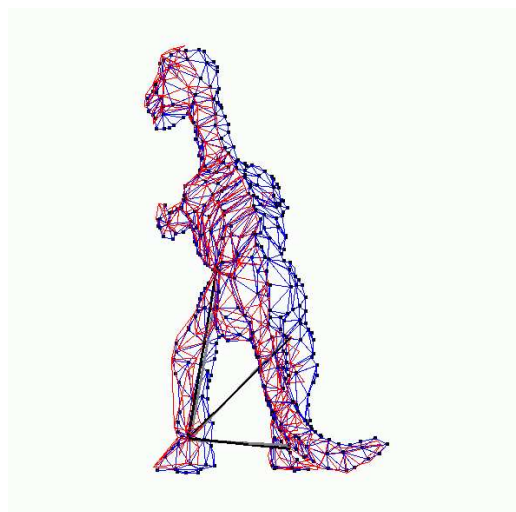
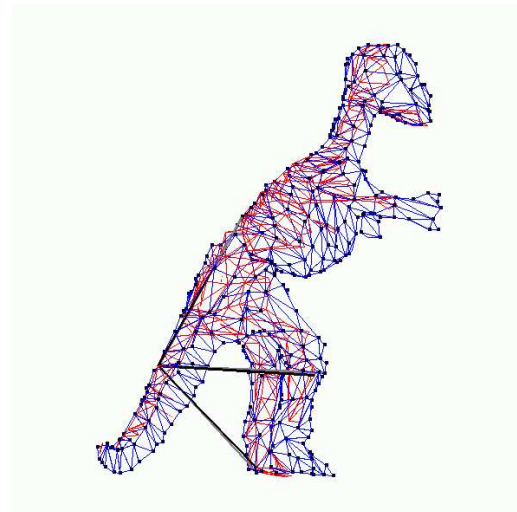


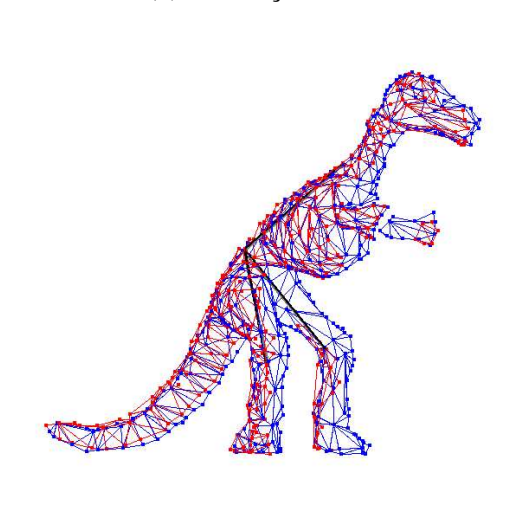
Fig. A.9: Registro global de vistas: BIRD.



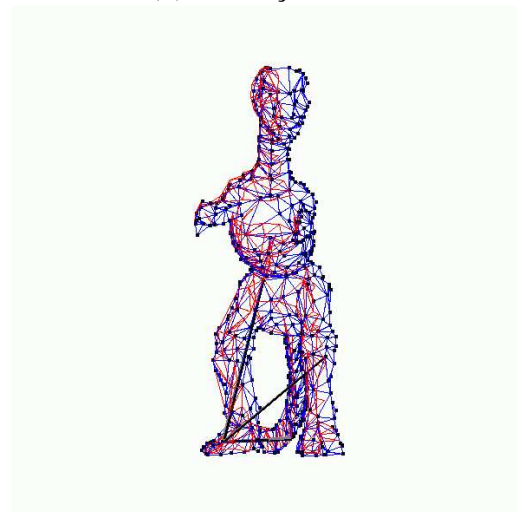
(a) 5 iterações ICP



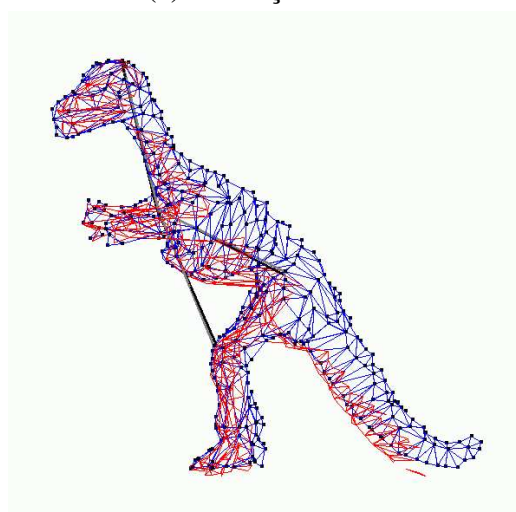
(b) 5 iterações ICP



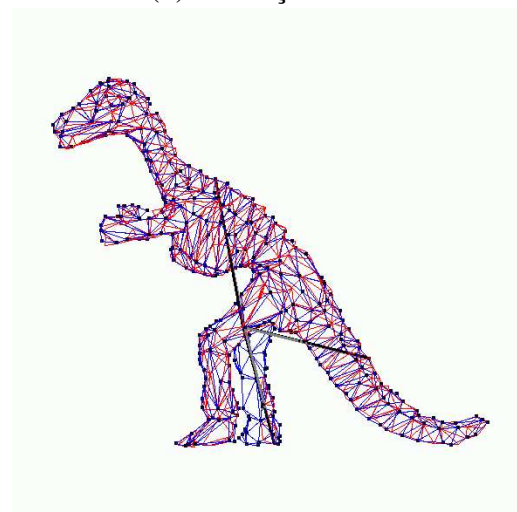
(c) 4 iterações ICP



(d) 5 iterações ICP



(e) 5 iterações ICP



(f) 4 iterações ICP

Fig. A.10: Registro de pares de vistas: DINO.

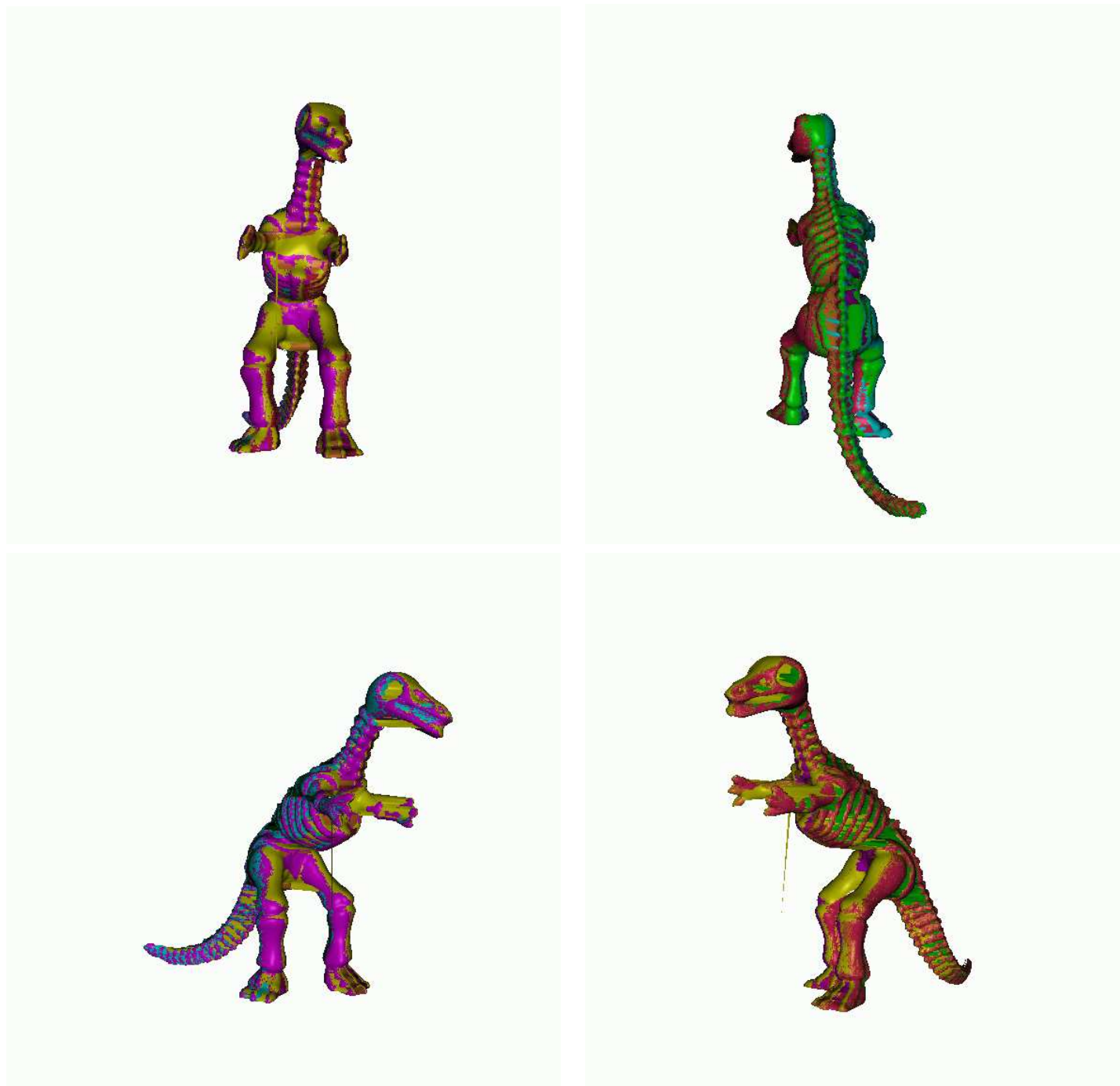


Fig. A.11: Registro global de vistas: DINO.

Apêndice B

Pseudo-código

Segue-se o pseudo-código que apresenta o algoritmo de pré-alinhamento que inclui os procedimentos descritos no Capítulo 3. Chamam-se de malha-referência e malha-alvo as malhas triangulares construídas a partir das amostras das imagens $\mathcal{R}_1$ e $\mathcal{R}_2$, respectivamente.

ALGORITMO PRÉ-ALINHAMENTO (R_1, R_2, f_1, f_2)

1. Construir a malha-referência
 2. Construir a malha-alvo
 3. Para cada vértice $\mathcal{V}$ da malha-referencia repita
 4. Calcule $H(\mathcal{V})$
 5. Para cada vértice $\mathcal{V}$ da malha-alvo repita
 6. Calcule $H(\mathcal{V})$
 7. Para cada vértice $\mathcal{V}$ da malha-referencia repita
- /*Construir um triedro t*/
8. Escolha um vértice v_1 com $f_2 \leq \text{dist}(v_1) \leq f_1$
 9. Para cada vértice v_2 com $f_2 \leq \text{dist}(v_2) \leq f_1$
 10. Calcule $c = \langle v_2 \vec{\mathcal{V}}, v_1 \vec{\mathcal{V}} \rangle$
 11. Escolha v_2 com menor c
 12. Calcule $N_1 = v_2 \vec{\mathcal{V}} \times v_1 \vec{\mathcal{V}}$
 13. Para cada vértice v_3 com $f_2 \leq \text{dist}(v_3) \leq f_1$
 14. Calcule $d = \langle v_3 - \mathcal{V}, N \rangle$
 15. Escolha v_3 com maior d
 16. Calcule as coordenadas baricêntricas a, b e c de v_3 no triângulo $\mathcal{V}v_1v_2$.
 17. Para cada vértice $\mathcal{C}$ da malha-alvo com $\text{abs}(H(\mathcal{V}) - H(\mathcal{C})) < \text{limiar}$ repita
 18. Para cada vértice c_1 da malha-alvo com $(\|c_1 - \mathcal{C}\| - \|v_1 - \mathcal{V}\|) < \delta$
 19. Para cada vértice c_2 da malha-alvo com $(\|c_2 - \mathcal{C}\| - \|v_2 - \mathcal{V}\|) < \delta$ e

$$(\|c_2 - c_1\| - \|v_2 - v_1\|) < \delta$$

20. Calcule $N_2 = c_2 \vec{\mathcal{C}} \times c_1 \vec{\mathcal{C}}$

21. Calcule $x = a * \mathcal{C} + b * c_1 + c * c_2$ e $p = x + N_2 d$.

22. Para cada vértice c_3 da malha-alvo com $abs(c_3 - p) < \delta$

23. Determine T com a equação 3.1

24. Transforme a malha-referência com uso de T

25. Para cada face $\mathcal{F}$ da malha-referência

26. Calcula-se o baricentro ba de $\mathcal{F}$

27. Calcula-se a distância D de ba até a face que é a mais próxima a $\mathcal{F}$ na

malha-alvo ¹

28. Se ($D < \text{limiar}$) $\text{cont} = \text{cont} + 1$

29. Escolhe-se T cujo cont é o maior de todos.

¹Para otimizar o procedimento e não testar contra todas as faces da malha, em um pré-processamento associa-se a cada face da malha-alvo os pixels por ela ocupados. Como conhece-se, também de antemão, o pixel ocupado por cada vértice da malha-referência, a procura da face mais próxima se restringe a identificar a face da malha-alvo, à qual pertence o pixel ocupado pelo baricentro ba .