Relatório de Trabalho No 2

Mercedes Gonzales Márquez

13 de novembro de 2002

Sumário

1 Introdução

Este relatório visa a apresentar as atividades durante o período nov/2001-out/2002 no contexto do projeto de doutorado intitulado "Reconstrução de um modelo 3D a partir de múltiplas imagens de profundidade". Além das atividades a serem detalhadas na Seção 5 foram cursadas 2 disciplinas para completar os créditos exigidos pelo programa de doutorado da FEEC.

Período	Nome da disciplina	Créditos	Carga horária	Conceito
ag/01-dez/01	Estudos Especiais IV	12	180H	S
mar-jun/02	Inteligência Artificial	12	180H	A

2 Plano Inicial do Projeto de Pesquisa

O Método Fluxo Radial, a ser sintetizado na Seção 4, é um método de reconstrução que gera uma malha triangular de arbitrária topologia a partir de uma imagem de profundidade através de deformações. A malha reconstruída, apesar de ser suficientemente próxima da superfície amostrada, apresentava triângulos degenerados (muito longos e finos). Além do mais, a malha representava apenas parcialmente a superfície do objeto, uma vez que uma única imagem de profundidade só pode capturar parcialmente a sua geometria. Em vista disso, o objetivo da nosso projeto de pesquisa é o aprimoramento e a extensão do Método Fluxo Radial, de maneira a termos um algoritmo de reconstrução de um modelo 3D a partir da inclusão incremental de outras imagens de profundidade. Desta forma, assumindo que as imagens não estejam registradas a priori, acreditamos que o nosso objetivo pode ser atingido através das seguintes tarefas:

- Proposta de um mecanismo de controle de faces degeneradas, possibilitando a construção de uma malha de melhor qualidade.
- Proposta de um algoritmo de registro e integração incremental de imagens.
- Reavaliação da estratégia de deformação do RFM.

Foi estabelecido o seguinte cronograma para a realização destas tarefas

Assunto	1S	2S	3S	4S	5S	6S	7S	8S
Disciplinas	x	x						
Aprimoramento da malha	x	x						
Estudo das técnicas registro e integração		x	x					
Algoritmo de registro e integração			x	x	x			
Algoritmo de deformação								
incorporando as novas vistas				x	x	x		
Estágio Sandwich							x	
Conclusões, escrita da tese e defesa								x

3 Etapas já realizadas

No primeiro relatório apresentado à FAPESP, apresentou-se um algoritmo de subdivisão baseado no esquema de triangulação Delaunay incremental, junto com uma estratégia conveniente de inserção de vértices na malha. Este algoritmo permitiu a geração de malhas de melhor qualidade que as obtidas pelo algoritmo de subdivisão anteriormente usado pelo Modelo de Fluxo Radial. Os resultados obtidos foram aceitos na categoria de artigos completos no "XIV Brazilian Symposium on Computer Graphics and Image Processing". Foi também feito um levantamento bibliográfico inicial de algumas técnicas de registro de imagens de profundidade. Paralelamente, foram cursadas 5 disciplinas necessárias para o cumprimento de créditos exigidos pelo programa de doutorado da FEEC.

4 Método de Fluxo Radial

Para ser auto-contido incluímos nesta seção uma descrição sucinta do algoritmo RFM, a base do nosso projeto.

O Modelo de Fluxo Radial (RFM) é um método iterativo de reconstrução parcial da superfície de um objeto a partir de uma única imagem de profundidade R. Seu domínio de aplicação é superfícies simplesmente conexas, orientáveis e de arbitrária topologia. O processo de reconstrução inicia com uma malha inicial simples (decaedro) posicionada num primeiro sistema de referência que é determinado por um centro radial e um eixo z na direção de profundidade da imagem. Os vértices desta malha crescem na direção radial para pontos correspondentes em R, fazendo com que a malha se aproxime às amostras de R. Enquanto a malha cresce, suas faces subdividem gerando novos vértices e novos correspondentes radiais em R, propiciando uma adaptação mais fina à geometria espacial das amostras. O processo crescimento-subdivisão se repete até a malha aproximar à região estrelada correspondente na imagem R sob uma tolerância de erro λ. Se a imagem apresentar uma geometria não estrelada, outras frentes de crescimento devem ser determinadas e novos centros radiais devem ser criados para atingir essas regiões. Em paralelo, *testes de proximidade* entre pares de frentes devem ser realizados para, através de cirurgias topológicas, possibilitar a geração de uma malha com gênero compatível ao gênero de R. A seguir uma breve descrição dos sete estágios do método.

1. Pré-processamento
2. Inicialização
3. Crescimento
4. Erro de reconstrução e ancoramento
5. Refinamento adaptativo da malha
6. Determinação de novas frentes de crescimento
7. Adaptações topológicas

4.1 Pré-processamento

Os passos a serem realizados nesta etapa são:

- Segmentação das componentes maximamente conexas da imagem, guardando a componente C de maior tamanho em pixels.
- Geração da imagem binária I, através da atribuição de 0 para cada pixel inválido (com profundidade zero em R) de C e 1 para cada pixel válido (com profundidade diferente de zero em R) de C.
- Determinação da borda externa de I, e as internas se houverem.

4.2 Inicialização

- Posicionamiento de um vértice v_0 da malha inicial (decaedro) num primeiro sistema de referência determinado pelo eixo z_0 na direção de profundidade, e um centro radial C_0 cuja coordenada z é o menor valor de profundidade das amostras, e cujas coordenadas (x, y) correspondem à posição 2D de um pixel válido em R.
  O vértice v_0 é chamado de *vértice inativo* pois permanecerá fixo durante o crescimento, e a suas faces incidentes serão chamadas de *faces inativas* pois não serão subdivididas.

- Determinação da primeira frente de crescimento A_0 constituída pelas faces que não são inativas.
- Determinação da região estrelada[1] R_0 correspondente à primeira frente de crescimento A_0.
- Associação de um ponto correspondente em R_0 para cada vértice na malha. Estes pontos são tomados na direção radial dos seus vértices associados.

Na apresentação dos próximos estágios do método usaremos a notação $M_{k,i}$ para referir à malha contendo as faces da frente A_k e na i-ésima iteração de crescimento. A região r pertencente à região estrelada R_k e determinada pelos pontos correspondentes dos vértices da face f, será chamada de região correspondente da face f.

4.3 Crescimento

Para a malha $M_{k,i}$ poder crescer em direção da sua região estrelada R_k, os seus vértices são deslocados radialmente na direção das amostras. O passo de deslocamento será considerado como o menor erro de reconstrução das faces em relação a cada amostra p_i da sua região correspondente. Isto é para não ultrapassar a imagem. Além disso, para impedir a interpolação de pontos ruidosos (com profundidade muito distante das profundidades dos seus pontos vizinhos) é feita uma diminuição da velocidade dos vértices que tiverem algum dos seus vizinhos suficientemente próximos dos seus pontos correspondentes.

4.4 Erro de reconstrução e ancoramento

Devido ao deslocamento radial dos vértices do modelo até seus correspondentes na imagem, o erro de reconstrução de um vértice v_i do modelo é simplesmente calculado em coordenadas esféricas por: $e_{v_i} = \rho_{p_i} - \rho_{v_i}$, onde ρ_{v_i} é a coordenada ρ do vértice v_i, e ρ_{p_i} é a coordenada do ponto correspondente p_i. Também o erro radial e_{mi} de uma face f_m em relação a um ponto p_i pertencente a sua região r_m é expresso por $e_{mi} = \|\vec{Op_i} - (\vec{Op_i} \bigcap f_m)\|$, onde $\vec{Op_i} \bigcap f_m$ é o vetor definido pelo ponto interseção da face f_m e o raio $\vec{Op_i}$. Quando $e_{v_i} < \lambda$ dizemos que v_i é ancorado, e quando o erro de reconstrução de f_m em relação a pelo menos $\eta\%$ dos pontos amostrados na sua região

[1] Chama-se de região estrelada quando um raio saindo da origem O até cada ponto $p \in R$ só intersecta com R em p

correspondente r_m for menor que λ, dizemos que a face f_m está ancorada. O parâmetro λ mede a tolerância do erro de proximidade da malha às amostras, e o parâmetro η controla o grau de fidelidade da reconstrução.

Consideremos agora a amostragem de uma superfície com geometria não estrelada. Nesta situação quando a malha $M_{k,i}$ ancorar na sua região estrelada R_k, faces paralelas à direção radial dos seus vértices surgirão, não havendo portanto nenhuma região correspondente para ser associada. Estas faces serão chamadas de *faces paralelas* e, embora não tenham regiões correspondentes, estão com erro e portanto deve-se abrir novas frentes de crescimento.

4.5 Refinamento da malha

Nesta etapa a malha $M_{k,i}$ é ajustada à geometria do objeto amostrado através de:

- Determinação de faces válidas para subdivisão.- Uma face f_m é subdividida se ela não for paralela e se tiver erro de reconstrução em relação a sua região correspondente r_m.
- Inserção de vértices na malha $M_{k,i}$.- Quando uma face f_m é marcada para subdivisão, insere-se um novo vértice na malha, de maneira que seja o mais próximo ao baricentro da face e que possua um correspondente radial em R_k.
- Subdivisão de faces.- As faces são subdivididas de acordo com o esquema de triangulação incremental Delaunay (algoritmo de Guibas [1]) segundo descrito no relatório científico anterior apresentado à FAPESP.

4.6 Determinação de novas frentes de crescimento

Quando o refinamento não é mais possível, as faces com erro de reconstrução acima da tolerância são agrupadas em conjunto maximamente conexos de faces que definirão as novas frentes de crescimento. Para cada conjunto A_k um sistema de referência é definido. Há uma diferença na determinação, dependendo de A_k possuir ou não vértices de borda (vértices da malha associados a amostras da borda da imagem).

4.6.1 Frentes que não incluem vértices de borda

- A direção do eixo z (Z_k) é a direção do vetor resultante da diferença do centróide do conjunto de faces de A_k, e o centróide das amostras da região correspondente às faces de A_k.

- A origem C_k é um ponto sobre Z_k tal que todos os pontos de borda do conjunto de faces da frente formem em relação a Z_k, um ângulo menor ou igual a 45^o.

4.6.2 Frentes que incluem vértices de borda

- A frente A_k contém uma ou várias faces paralelas com vértices de borda (faces adjacentes às faces inativas). Tais faces chamaremos de *faces-borda paralelas.*
- A direção de Z_k será a direção de vista.
- A origem C_k será um ponto sobre o ponto médio dos vértices de borda associados à face-borda paralela, tal que todos os vértices da borda de A_k formem com relação a Z_k um ângulo menor ou igual a 45^o. Se houver mais de uma face-borda paralela, C_k será a média dos valores computados para cada face-borda paralela.

A partir do centro radial C_k determina-se a região estrelada R_k e realiza-se a seqüência cíclica crescimento-subdivisão sobre A_k até que não seja mais possível o crescimento.

4.7 Adaptações topológicas

Quando existirem duas frentes A_k e A_h com faces-borda f_1 e f_2 paralelas e conectadas por uma aresta-borda interna, avalia-se o seguinte teste de proximidade:(1) Se os vetores normais de f_1 e f_2 são quase paralelos e opostos, e (2) se os dois segmentos definidos pelos vértices de borda estão contidos em R. Se estas duas condições forem satisfeitas realiza-se a cirurgia topológica, que compreende dois estágios

1. Corte: Remova-se as faces limite f_1 e f_2
2. Costura
   - As duas bordas resultantes da eliminação das faces são conectadas por oito faces triangulares.
   - Faz-se a soma conexa entre A_k e A_h resultando uma única frente $A_k = A_k \bigcup A_h$

5 Atividades de Pesquisa

Em relação ao plano inicial listado na Seção 2 realizamos ao longo deste período as seguintes tarefas de pesquisa

- Aprimoramento do procedimento de subdivisão do RFM
- Revisão bibliográfica das técnicas de registro e integração. Um resumo de cada artigo é apresentado no apêndice.
- Esboço de um algoritmo de integração de duas imagens que se adequa a RFM
- Projeto de algoritmo de registro de duas imagens

5.1 Subdivisão da malha

Uma mudança na subvisão foi feita em relação ao apresentado no relatório anterior. Faces ao invés de arestas são subdivididas com o intuito de simplificar o nosso algoritmo de subdivisão. Para isso deve-se rotular faces no lugar de arestas como válidas para subdivisão. Constatamos que o procedimento de determinação de arestas válidas é mais complexo que o de identificação de faces válidas. Uma aresta válida para subdivisão deve satisfazer

1. O seu comprimento e for maior que L_k (limiar estimado),
2. pelo menos uma das suas faces adjacentes tiver erro de reconstrução e não for paralela
3. o comprimento da maior aresta das suas faces adjacentes exceder no máximo 25% do seu comprimento.

Enquanto para validar uma face somente uma condição precisa ser verificada: se ela tiver erro de reconstrução e não for paralela.

5.2 Integração

Embora a seqüência usual para reconstrução de uma malha triangular a partir de duas ou mais imagens é registro seguido de integração, decidimos fazer primeiro alguns experimentos simples para certificar a factibilidade da reconstrução incremental de uma malha usando o princípio de deformação do modelo RFM.

Diferentemente das técnicas encontradas na literatura (Apêndice B) nós propomos fazer uma integração incremental de uma imagem de profundidade a uma malha triangular fechada existente usando o modelo RFM. Para isso identificamos os seguintes estágios a serem incluídos no nosso modelo RFM para viabilizar a nossa proposta.

- Registro da malha triangular com a nova imagem de profundidade (seção 5.3)
- Identificação da região de sobreposição
- Ajuste dos valores de profundidade na região de sobreposição
- Reposicionamento dos vértices dos triângulos na região de sobreposição de forma que cada um deles tenha uma amostra correspondente na nova imagem de profundidade
- Determinação de novas frentes de crescimento

Como testes preliminares implementamos um algoritmo simples de integração considerando que a malha M e a nova imagem R_2 sejam registrados. A seguir descrevemos sucintamente as principais idéias a serem incorporadas em cada estágio.

- Primeiro Estágio : Região de sobreposição
  Para descarte dos triângulos de M que não pertencem à região de sobreposição podemos usar o algoritmo z-buffer. Com este algoritmo um triângulo não é considerado visível a partir do ponto de vista v de R_2 se um raio saindo de v intercepta M em outro ponto com profundidade menor, além dos pontos do triângulo.

- Segundo Estágio: Ajuste dos valores z
  Para cada amostra a_i redundante em R_2 o triângulo t associado a ela é projetado sobre a grade G_2 da imagem R_2 e as coordenadas baricêntricas de a_i são determinadas. O valor de profundidade $z1$ é interpolado a partir dos valores de profundidade dos vértices de t e as coordenadas baricêntricas determinadas de acordo com a sugestão de Soucy e Laurendeau [19, 20]. O valor z_2 é o próprio valor da profundidade de a_i. Finalmente, o valor z da amostra é calculado através da média ponderada de z_1 e z_2, onde as ponderações são os valores de incerteza correspondentes.

- Terceiro estágio: Reposicionamento dos vértices da região de sobreposição
  Associar a cada vértice v_i de M na região de sobreposição o ponto correspondente na imagem R_2 à amostra mais próxima à projeção de v_i sobre R_2.

- Quarto estágio: Determinação de novas frentes
  Determina-se novas frentes a partir das faces da região de sobreposição. Para cada frente de crescimento deve-se ainda reavaliar os erros de reconstrução das faces na região de sobreposição.

O algoritmo apresentado têm sido proposto de maneira intuitiva. Só o testamos na integração das amostras de duas imagens de uma esfera (Figura 1),

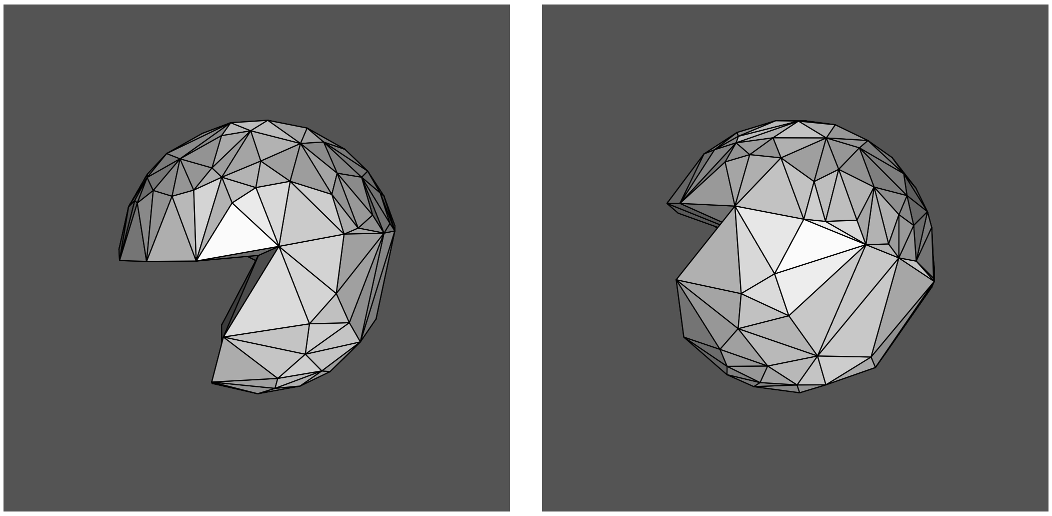

Figura 1: Testes preliminares

cujo registro é dado pela transformação

$$T = \begin{pmatrix} -0,500 & 0,0 & -0,866 & 0,0 \\ 0,0 & 1,0 & 0,0 & 0,0 \\ 0,866 & 0,0 & -0,500 & 0,0 \\ 0,0 & 0,0 & 0,0 & 1,0 \end{pmatrix}$$

Testes exaustivos devem ainda ser feitos.

5.3 Registro

Propomos fazer alguma modificação no algoritmo DARCES [4] e combiná-lo com um algoritmo ICP [3] para obtermos uma transformação de registro. O algoritmo DARCES nos permite obter uma transformação inicial requerida pelo algoritmo ICP e este por sua vez faz ajustes finos para convergir tal transformação a uma transformação mais precisa. No entanto, um dos pontos críticos do algoritmo DARCES é a determinação de pontos representativos ou pontos de controle para estabelecer triplos de controle planares que são usados no processo de registro. Isso é feito de forma totalmente aleatória. A nossa hipótese é que tendo as informações topológicas de uma das imagens através de uma malha de baixa resolução, podemos não só escolher de forma mais criteriosa e automática estes pontos, como também utilizar quádruplos espaciais no lugar de triplos planares. Os quádruplos podem refletir melhor a geometria (local) espacial da superfície, pois eles indicam a orientação dos pontos de controle no espaço. Implementamos o seguinte algoritmo para registrar duas imagens R_1 e R_2

1. Reconstruir a estrutura topológica M da imagem R_1 com uso do RFM

2. Avaliar as variações dos vetores normais das faces adjacentes de cada vértice p_i da malha M reconstruída

3. Para cada vértice p_i cuja variação dos vetores normais é maior que um limiar pré-definido α

    4 Escolher três pontos na direção de três arestas adjacentes ao ponto p_i

    5 Determinar com p_i e os três pontos um quádruplo espacial q_{1i}

    6 Determinar todos os conjuntos de quatro pontos em R_2 que são similares a q_{1i} a menos de um movimento rígido (translação e rotação). O princípio de busca destes quádruplos é similar ao algoritmo originalmente proposto por Chu-Song e outros [4].

    7 Determinar dentre os quádruplos identificados em R_2 aquele que tiver associada uma transformação T_i que leve a malha M mais próxima possível das amostras de R_1, ou seja com o menor erro de registro.

    Utilizamos o critério de distância ponto-ponto para avaliação destes erros. Por simplicidade, determinamos somente para os vértices adjacentes a p_i o seu correspondente usando a técnica de projeção [3] e avaliamos o erro em função da distância destes pares de pontos.

8 Escolher dentre as transformações T_i aquela que leve a um registro com o menor erro.

Alguns experimentos preliminares foram feitos para avaliar a nossa proposta. Usamos duas imagens R_1 e R_2 de duas vistas mostradas na Figura 2. Reconstruímos a partir da imagem da Figura 2a a malha com uma tolerância de erro ($\tau = 2.5$) através do modelo RFM. A partir desta malha foi possível determinar 4 quádruplos cujos vértices geradores apresentam faces adjacentes com um limiar $\alpha = 56,5^o$ nos seus vetores normais. Na Figura 3 estes vértices

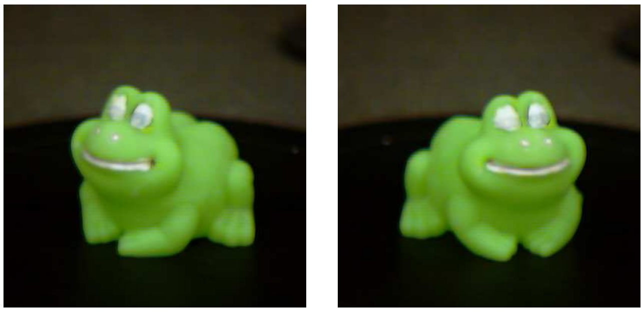

Figura 2: (a) R_1, 0^o em torno ao eixo y (b) R_2, 60^o em torno ao eixo y

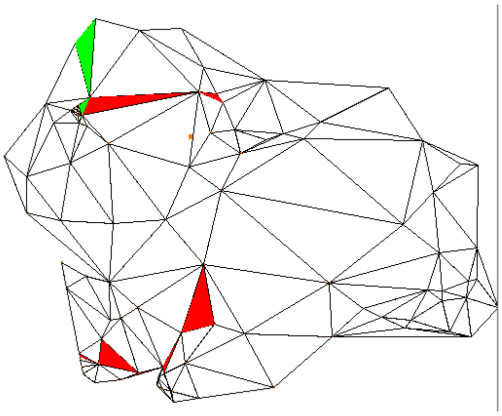

Figura 3: Quádruplos

geradores tem duas das suas faces adjacentes preenchidas.

Aplicamos para o vértice de maior variação o algoritmo DARCES modificado e obtivemos um registro razoável, como ilustra a Figura 4.

Em seguida, aplicamos o algoritmo ICP [3] usando a matriz de transformação inicial

$$T_0 = \begin{pmatrix} 0,554 & 0,352 & -0,753 & 0,0 \\ -0,305 & 0,928 & 0,209 & 0,0 \\ 0,773 & 0,114 & 0,623 & 0,0 \\ -761,318 & -111,901 & 367,021 & 1,0 \end{pmatrix}$$

dada pelo algoritmo DARCES modificado, e obtivemos a seguinte matriz de registro

$$T = \begin{pmatrix} 0,521 & 0,362 & -0,773 & 0,0 \\ -0,316 & 0,923 & 0,220 & 0,0 \\ 0,793 & 0,130 & 0,595 & 0,0 \\ -779,554 & -126,711 & 393,947 & 1,0 \end{pmatrix}$$

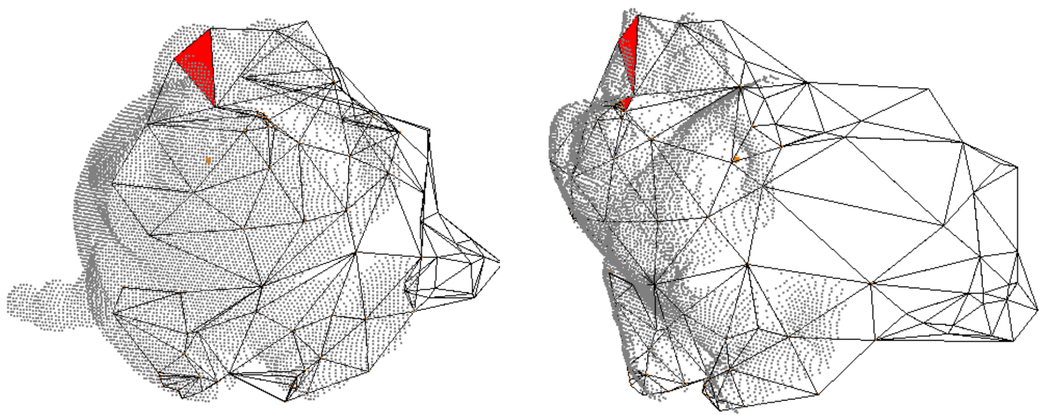

Figura 4: Testes preliminares

Estes resultados animadores suportam a nossa hipótese de que o conhecimento de uma malha triangular de baixa resolução nos permite restringir o domínio de busca das amostras na imagem R_1. Enquanto Chu-Song e outros no pior dos casos, usariam a amostragem completa, nós usamos uma subamostragem representativa da amostragem original, que são os vértices da malha reconstruída. No exemplo da Figura 4 a amostragem é de 9649 pontos e a malha gerada com $\lambda = 2.5$ e $\eta = 0.8$ é de 92 vértices, o que representa uma subamostragem do 0.95% dos dados originais. Vários experimentos precisam ser conduzidos ainda para validar a nossa hipótese.

6 Conclusões

Neste período do projeto foram dedicados mais de seis meses para estudar e analisar as principais técnicas de registro e integração encontradas na literatura. Uma síntese destas técnicas é incluída no apêndice deste relatório.

Com base nestas técnicas foi esboçado um algoritmo de registro de duas imagens. A nossa idéia é combinar o algoritmo DARCES com um algoritmo ICP para obter uma boa matriz de transformação de registro. Pensamos ainda explorar as informações topológicas contidas numa malha reconstruída pelo RFM para melhorar o procedimento de procura do algoritmo DARCES. Conduzimos alguns testes preliminares com quádruplos no lugar de triplos de controle. Fizemos ainda um experimento simples para verificar a factibilidade de um procedimento de integração incremental entre uma malha triangular fechada e uma imagem de profundidade.

Em relação à previsão inicial apresentada na Seção 2 estamos um pouco defasadas e pretendemos cobrir parcialmente esta diferença no próximo período do projeto conforme o cronograma que se segue

Assunto	11/12	01/02	03/04	05/06	07/08	09/10
Depurar módulo de segmentação inicial do RFM	X	X	X	X		
Conduzir testes mais conclusivos em relação ao algoritmo de registro	X	X				
Instalação/adaptação do algoritmo de Turk e Levoy		X				
Proposta com fundamentação de um algorimo de registro Implementação		X	X			
Preparo de um artigo para o Sibgrapi			X			
Análise dos problemas referentes à integração com uso do RFM e estudo de possíveis soluções			X	X	X	
Validação das diferentes soluções			X	X	X	
Proposta com fundamentação de um algorimo de integração Implementação				X	X	X
Participação no Sibgrapi					X	

A Técnicas de registro

Dadas n imagens de profundidade $R_1, R_2, \dots R_n$ de um objeto, cada uma contendo pontos 3D no sistema local de coordenadas do escaneador, que descrevem a geometria do objeto visto a partir de um particular ponto de vista. O problema de registro dessas vistas consiste em encontrar n transformações rígidas $T_1, T_2, \dots, T_n$ que especifiquem as corretas posições do escaneador com relação a um único sistema de referência global (arbitrariamente escolhido e usualmente o referencial de uma das vistas).

O registro de n vistas é o objetivo final, porém para chegar a atingi-lo precisa-se do processo intermediário de registro entre duas vistas R_i e R_j (ou R_1 e R_2). Assumiremos que a transformação procurada T levará R_1 em alinhamento com R_2.

Pode-se expressar a transformação T entre R_1 e R_2 como uma operação linear usando uma representação homogênea de matriz 4x4. Ela é composta por $R_{3\times3}$ e $t_{3\times1}$. $R_{3\times3}$ é composta por uma seqüência de 3 rotações de ϕ, θ e ψ graus com relação aos eixos x, y e z respectivamente, e $t_{3\times1}$ por 3 translações com relação aos eixos x, y e z, respectivamente. A notação matricial seguinte é usada T pode ser expressa como

$$T = \begin{pmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \\ T_{41} & T_{42} & T_{43} & T_{44} \end{pmatrix} = T(R,t) = \begin{pmatrix} R_{11} & R_{12} & R_{13} & t_1 \\ R_{21} & R_{22} & R_{23} & t_2 \\ R_{31} & R_{32} & R_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = T(\phi, \theta, \psi, t_x, t_y, t_z)$$

As técnicas de registro existentes na literatura podem ser classificadas em duas categorias:

1. dependentes da transformação inicial
2. independentes da transformação inicial

No registro dependente parte-se da suposição que a transformação entre as duas vistas é aproximadamente conhecida. Ela é utilizada como estimativa inicial é no algoritmo ICP (Iterative closest point) para obter uma transformação mais justa. Quando não se consegue obter um chute inicial recomenda-se o uso do registro independente. Este método se baseia no fato de que a distância entre os pontos são preservadas nas transformações rígidas. Apresentamos treze algoritmos de registro, entre eles os 11 primeiros são dependentes da transformação inicial, e os dois últimos são independentes dela. Eles são sintetizados a seguir:

A.1 Algoritmo ICP [2]

Como foi mencionado, o Algoritmo ICP (Ponto mais próximo iterativo) é o método dominante na tarefa do alinhamento de duas ou mais vistas quando uma estimativa inicial é fornecida. Ele foi originalmente proposto por Besl e McKay em 1992. Várias novas sugestões foram apresentadas desde então para aprimorá-lo.

Sejam $R_1 = \{p_i\}_1^{N_1}$ e $R_2 = \{q_i\}_1^{N_2}$ dois conjuntos de pontos 3D que representam duas posições no espaço diferentes de uma superfície. Assuma-se que o conjunto de pontos $\mathcal{S}_{R1}$ da superfície representados por R_1 está contido no conjunto de pontos $\mathcal{S}_{R2}$ que R_2 representa na superfície ($\mathcal{S}_{R1} \subset \mathcal{S}_{R2}$), e portanto cada ponto p_i em R_1 tem um correspondente q_i em R_2 que ocupa a mesma localização da superfície que p_i, satisfazendo a seguinte equação

$$q_i = Tp_i \qquad i = 1, \ldots, N_1$$

Conhecemos p_i (amostras de R_1), e se conhecêssemos q_i o problema de registro estaria resolvido pois T pode ser facilmente encontrada através de um método de mínimos quadrados ou média de mínimos quadrados

$$\min f \qquad \text{onde } f = \sum_{i=1}^{N_1} \frac{1}{N_1} \|q_i - Tp_i\|,$$

No entanto os $q_i \in R_2$ correspondentes aos p_i não são conhecidos pois dependem da transformação T que estamos pretendendo encontrar.

Besl e McKay apresentam uma proposta com uma tentativa de estimar de maneira simples o ponto correspondente q_i em R_2 para dado p_i em R_1: O ponto $q_i \in R_2$ mais próximo a p_i. Para calcular a transformação eles iniciam um processo iterativo fornecendo uma transformação inicial T_0 que pode ser obtida manualmente pelo usuário ou por casamento de características na imagem. Em cada iteração k uma transformação intermediária T_k é computada e novos pontos q_i são determinados para novas posições dos pontos p_i que resultam de aplicar a transformação T_{k-1}. Após sucessivas iterações os autores esperam que finalmente os q_i determinados sejam os pontos que ocupam (com uma margem pequena de erro) a mesma posição na superfície amostrada. O algoritmo ICP é o seguinte:

1. Inicialize com T_0, $k = 1$,
2. Atualize p_i aplicando T_{k-1} $\qquad p_i = T_{k-1}p_i$
3. Encontre o ponto mais próximo q_i para o ponto p_i usando a fórmula $\min_{q \in R_2} \|q - p_i\|$
4. Compute a transformação T_k que minimiza a função erro

$$f_k = \sum_{i=1}^{N_1} \frac{1}{N_1} \|q_i - T_k p_i\|$$

5. Se $(f_k - f_{k-1}) < \tau$ (tolerância pré-definida), páre, caso contrário vá a (2)

No método ICP usa-se o método de Horn [9] para minimizar a função erro f e achar T. Horn propõe um método direto ao invés de métodos numéricos iterativos para solucionar o problema de mínimos quadrados para 3 ou mais pontos. Seu trabalho usa quatérnios unitários para representar rotações e é descrito brevemente a seguir:

Um quatérnio unitário é um vetor $\vec{q} = [q_0 q_1 q_2 q_3]$ ou $\vec{q} = q_0 + q_1 i + q_2 j + q_3 k$ com $q_0 \geq 0$ e $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$, e a matriz 3x3 de rotação gerada por $\vec{q}$ é

$$R_q = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{pmatrix}$$

A transformação procurada é representada pelo vetor $\vec{q} = [\vec{q}_R|\vec{q}_t]$ onde $\vec{q}_R$ é um quatérnio unitário que gera uma matriz de rotação e $\vec{q}_t = [q_4 q_5 q_6]^t$ o vetor traslação. No contexto de quatérnios, a função erro a ser minimizada é:

$$f(\vec{q}) = \sum_{i=1}^{N_1} \frac{1}{N_1} \|q_i - R(\vec{q}_R)p_i - \vec{q}_t\|^2$$

Horn demonstrou que o quatérnio $\vec{q}_R$ que minimiza f é o mesmo quatérnio $\vec{q}_R$ que maximiza a seguinte forma quadrática

$$\sum_{i=1}^{N_1} (p_i - \vec{\mu}_{R_1}) R_{\vec{q_R}} (q_i - \vec{\mu}_{R_2})$$

onde $\vec{\mu}_{R_1}$ e $\vec{\mu}_{R_2}$ são os centróides dos N_1 pontos de R_1 e dos N_1 pontos de R_2 respectivamente. Provou ainda que o quatérnio unitário solução $\vec{q}_R$ é o autovetor que corresponde ao máximo autovalor da matriz simétrica 4x4 Q cujas componentes são geradas pela matriz covariante C entre os pares de pontos dados

$$C = \frac{1}{N_1} \sum [(p_i - \vec{\mu}_{R_1})(q_i - \vec{\mu}_{R_2})^t]$$

As componentes da matriz $A = C - C^t$ são usadas para formar o vetor coluna $\Delta = [A_{23} A_{31} A_{12}]^t$ que é usado para formar Q da seguinte forma

$$Q = \begin{pmatrix} tr(C) & \Delta^t \\ \Delta & C + C^t - tr(C) I_3 \end{pmatrix}$$

onde I_3 é a matriz identidade 3x3.

Finalmente o vetor translação $\vec{q}_t$ pode ser facilmente obtido com a seguinte fórmula

$$\vec{q}_t = R_{\vec{q_R}} \vec{\mu}_{R_1} - \vec{\mu}_{R_2}$$

Concluindo, o problema de achar a $T(R, t)$ que minimiza a função erro f se resume a encontrar o autovetor pertencente ao maior autovalor da matriz Q.

Antes de passar para a análise do método ICP apresentaremos alguma terminologia que será usada no decorrer desta relatório. O conjunto $S_1 \subset R_1$ e o conjunto $S_2 \subset R_2$ que correspondem à mesma área da superfície amostrada, serão chamados de *região de sobreposição* de R_1 e *região de sobreposição* de R_2 respectivamente, e a união desses dois conjuntos será chamada apenas *região de sobreposição*. O par de pontos (p_i, q_i) onde $p_1 \in S_1$ e $q_1 \in S_2$ que ocupam exatamente a mesma posição na superfície amostrada será chamado de *par ideal*, e o par será chamado *par conveniente* quando ambos pontos pertencem à região de sobreposição e q_i ocupa uma posição suficientemente próxima da localização de p_i na superfície amostrada. Por outro lado o par será um *par inconveniente* quando pelo menos um dois pontos não pertence à região de sobreposição ou q_i não ocupam uma posição suficientemente próxima da localização de p_i na superfície. Chamaremos de *par de controle* ao par de pontos (p_i, q_i) a ser usados na construção da função erro f que será minimizada para computar a transformação de registro T. O ponto p_i será chamado *ponto de controle* e q_i ponto de controle correspondente ou simplesmente *ponto correspondente*.

Analisando ICP

- Convergência do método
  Os autores demonstram que o algoritmo ICP sempre converge monotónicamente ao mais próximo mínimo local da função erro f, e verificam através de experimentos que a razão dessa convergência é rápida. No entanto um mínimo local não satisfaz pois ele não da a solução.

  Portanto somente dispondo de uma transformação inicial T_0 suficientemente próxima a T pode-se garantir que f atinja o seu mínimo global e conseqüentemente a ICP converja na solução. A transformação T_0 é usualmente obtida manualmente, por casamento de algumas características relevantes na imagem ou por interação com o usuário.

- Desconhecimento da região de sobreposição ou incerteza na determinação dos pares de controle
  A condição que $\mathcal{S}_{R1} \subset \mathcal{S}_{R_2}$ imposta pelo ICP facilita a determinação de pontos p_i que podem ser levados em correspondência com pontos de R_2, de fato nesse caso $R_1 = S1$ e portanto qualquer ponto de R_1 pode ser um ponto de controle em R_1. O mesmo não pode ser dito da determinação dos pontos em R_2 correspondentes aos pontos de controle em R_1, já que a região de sobreposição S_2 é desconhecida. A proposta de Besl e McKay de considerar o ponto mais próximo como ponto correspondente é porém uma tentativa de encontrar pontos correspondentes que determinem pares de controle convenientes. Veremos mais adiante que isso nem sempre é possível na prática.

  O problema fica mais sério ao considerarmos que no nosso contexto R_1 e R_2 representam duas imagens de profundidade <u>parcialmente sobrepostas</u>, ou seja $\mathcal{S}_{R1} \not\subset \mathcal{S}_{R2}$ e portanto a região de sobreposição S_1 também é desconhecida. Isso nos leva a precisar de mais uma estratégia para determinar pontos de controle em R_1. Em princípio todos os pontos de R_1 podem ser considerados (como proposto por Besl e McKay) o que resultaria numa grande quantidade de pares de controle e, na determinação de pontos fora da região de sobreposição S_1. Esses pontos não serão em absoluto boas referências para o alinhamento. Além do mais ao acharmos um correspondente mais próximo para um p_i fora de S_1 estaremos determinando um par de controle potencialmente inconveniente para o alinhamento.

A incerteza na escolha de pontos de controle e determinação da correspondência tem sido alvo de muito pesquisa por tornar-se um problema crítico quando uma transformação inicial suficientemente boa não está disponível.

- Sensibilidade da minimização ponto-ponto da função erro
  Este problema tem a ver com a determinação de pares inconveniente e é decorrente do problema (2). A minimização ponto-ponto proposta pelos autores pode ser visualizada imaginando que os pares de pontos determinados são conectados por molas que puxam as regiões comuns das imagens até o alinhamento. A abordagem trabalha bem se os pares de pontos são ideais, e neste caso o registro correto é obtido numa única iteração. Porém, como analisado, pares inconvenientes poderão facilmente ser determinados fazendo que a convergência do algoritmo seja muito lenta ou fazendo que ele simplesmente diverja.

  Na Figura 5 vemos que duas superfícies estão muito próximas do al-

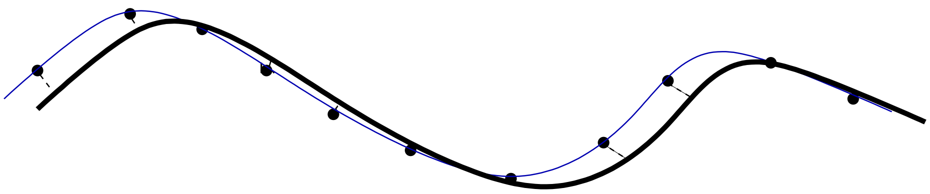

Figura 5: Distância Ponto - Ponto

inhamento correto, faltando só que S_1 "escorregue" ligeiramente para casar com S_2. Quando pontos mais próximos em S_2 são associados a pontos de controle p_i de S_1, ainda que os p_i pertençam à região de sobreposição pares inconvenientes podem ser gerados. Observe que as molas entre pares convenientes puxam a solução na direção certa, porém, os pares inconvenientes já estão muito próximos e portanto resistem a se movimentar mais que um pequeno paso, fazendo muito lenta a convergência à solução exata. O outro problema que surge, mais crítico ainda é quando os pontos de controle são escolhidos fora da região de sobreposição como os da figura 11, ou pontos correspondentes são determinados fora desta região. Observe as associações indevidas

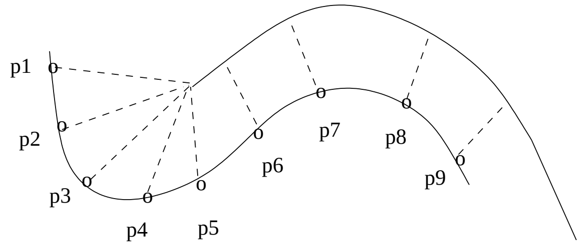

Figura 6: Associações indevidas

indevidas de p_1, p_2, p_3, p_4 feitas pelo ICP, que dão origem a pares inconvenientes. Quando a média dos residuais quadrados seja minimizada a superfície poderá facilmente divergir.

Com estes dois exemplos apresentamos uma séria conseqüencia da determinação de pares inconvenientes quando a função erro é representada pela média dos residuais quadrados. Apenas um par de controle inconveniente poderá redundar na lentidão da convergência e no pior dos casos até na divergência do ICP.

Com estes dois problemas básicos em aberto a comunidade envolvida na pesquisa de registro de imagens de profundidade partiu na busca de aprimoramentos da técnica ICP. A seguir apresentamos algumas variantes dele.

A.2 Chen e Medioni [5]

Eles ao invés de minimizar a distância entre pares de pontos, propõem minimizar a distância entre um ponto $p_i \in R_1$ e o plano tangente no ponto $q_i \in R_2$ mais próximo na direção do seu vetor normal. Desta maneira eles pretendem associar cada ponto de controle em R_1 a uma aproximação de superfície continua e não apenas discreta como no caso da minimização ponto-ponto. A função erro é a seguinte para N_s pares de controle

$$f = \sum_{i=1}^{N_1} d^2(S_i, Tp_i)$$

onde d representa a distância entre um ponto e um plano, S_i é o plano tangente no ponto mais próximo q_i a p_i na direção do vetor normal de p_i. O ponto mais próximo q_i a p_i é computado achando a interseção entre a reta normal a p_i e a superfície respresentada pelo conjunto de pontos de R_2.

Os pontos de controle são escolhidos interativamente clicando sobre pontos numa grade regular. Estes pontos devem pertencer à áreas suave curvatura e a suavidade da região é verificada ajustando um plano a uma vizinhança (por exemplo uma janela 9 x 9) e usando um método mínimos quadrados para obter o desvio residual do ajuste.

Pontos de controle sobre áreas de suave curvatura são escolhidos devido a duas razões

1. As vizinhanças correspondentes em R_2 dos pontos de controle em R_1 provavelmente serão também suaves e conseqüentemente o algoritmo que computa a interseção da reta com a superfície fornecerá uma solução mais confiável, e

2. o vetor normal à superfície pode ser computado mais confiávelmente.

Análise

- Quanto a correspondência
  Com a proposta dos autores, a superfície representada pelo conjunto R_1 e movimentada talque o seus pontos se aproximem aos planos tangentes dos seus correspondentes na outra imagem. Usando novamente a analogia com a mola, podemos ver que um extremo dela estaria no ponto de controle p_i e o outro estaria livre para "escorregar" ao longo da superfície. A única restrição para o movimento seria a direção na qual a distancia seria reduzida, ficando os dois graus de liberdade dado pelas direções sobre o plano, para que a superfície possa se movimentar com mais chance sem criar conflito com outros pares. Após aplicar esta técnica em algumas imagens teste, os autores verificaram que o mínimo global foi atingido mais rapidamente que o algoritmo ICP. No entanto, em presença de ruídos ou regiões de alta curvatura poderá haver erro no cómputo do vetor normal e conseqüentemente erros no computo da interseção e do plano tangente. Esses erros poderão influenciar na estimação dos parâmetros da transformação T, ao alterar a distância d que está sendo minimizada para estimar T. Este método poderá facilmente falhar nesse contexto.

- Quanto à escolha dos pontos de controle acreditamos também que ao restringir a escolha de pontos de controle em regiões suaves está-se desprezando a informação muito rica de variação de curvatura da superfície, pois pontos de controle com estas características relevantes podem ser vitais para determinar um correto alinhamento. Na Figura 12 há dois planos com uma ranhura no centro em forma de X, com uma diferença ligeira de ângulos entre eles. Como pode ser observado, para alinhar a geometria complexa das ranhuras, pontos sobre elas devem ser escolhidos, pois pontos fora de delas não fornecerão informação relevante para o sucesso do alinhamento.

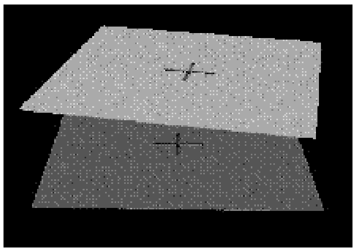

Figura 7: Plano com ranhura con centro em forma de X

A escolha restrita proposta faz que só poucos pontos significativos possam ser escolhidos, impedindo o correto alinhamento como foi experimentado por Rusinkiewicz e Levoy [15].

Nenhuma restrição é feita para a escolha de pontos de controle fora da região de sobreposição.

A.3 Gagnon e outros [8]

Eles tentam aprimorar o algoritmo de Chen e Medioni [5] e realizam modificações basicamente em dois aspectos

1. Na escolha dos pontos de controle .- Enquanto Chen e Medioni [5] consideram apenas pontos localizados em áreas suaves, os autores reduzem ainda mais esse domínio descartando pontos com descontinuidade de profundidade e incorporando um teste de visibilidade. Um ponto q_i em R_2 não é visível na vista R_1 quando o produto interno entre o vetor normal de q_i e a vetor que indica a profundidade de R_1 é negativo, o ponto q_i não é usado como ponto de controle. O teste é aplicado em ambas as imagens.

2. Na determinação do plano tangente .- Para encontrar o plano tangente em R_2 correspondente ao ponto $p_i \in R_1$, Chen e Medioni [5] escolhem o ponto p_2 na grade discreta na vista R_2 mais próximo à interseção entre o segmento de linha 3D normal a p_i e a superfície representada por R_2. Com o intuito de melhorar a precisão da transformação T, eles propõem fazer uma interpolação subpixel para computar a interseção ao invés de usar uma aproximação discreta.

Análise

- Quanto à região de sobreposição
  Os autores aprimoram a técnica de Chen e Medioni [5] no sentido de

introduzir dois critérios para evitar prováveis pares inconvenientes, um deles é a desconsideração dos pontos com descontinuidade de profundidade e o outro é a determinação das amostras em R_2 que são visíveis a partir da linha de vista de R_1. O teste de visibilidade aplicado na imagem R_1 fornece uma estimativa da provável região de sobreposição de R_1. Embora muita informação em R_1 que é visível a partir da linha de vista de R_2 possa não pertencer à região de sobreposição das imagens, este teste é uma boa primeira tentativa de redução de informação desnecessária. Os domínios para pontos de controle e pontos correspondentes são reduzidos com este teste.

- Nenhuma melhora substancial é feita em relação à determinação de correspondência.

A.4 Turk e Levoy [22]

Eles propõem minimizar a distância entre duas malhas triangulares M_1 e M_2 construídas a partir das imagens de profundidade R_1 e R_2 respectivamente. Os pontos de controle são vértices sobre a malha e os pontos correspondentes são pontos mais próximos sobre M_2. Usa-se a técnica de triangulação restrita à descontinuidade no passo que gera triângulos aproveitando a estrutura 2D da imagem da seguinte forma: Pixels adjacentes na imagem são conectados sempre que a distância euclideana entre os dois pontos correspondentes no espaço seja menor que um pré-definido limiar.

Constrói-se uma hierarquia de malhas de menor a maior detalhamento sendo a malha na hierarquia i 4 vezes maior em número de vértices que a da hierarquia $i-1$. A escolha de pontos é feita seguindo uma distribuição uniforme no plano da imagem. O objetivo da hierarquia é de acelerar o registro realizando o casamento sobre uma malha com poucos vértices e usando incrementalmente malhas mais detalhadas para maior refinamento.

Um limiar para desprezar pares de pontos muito distantes é introduzido e nenhum ponto de borda é usado como elemento do par de pontos de controle. Considerando que nem todos os pontos amostra tem o mesmo erro de aquisição já que isso depende da sua posição em relação ao escaneador, os autores introduzem um valor que mede o grau de confiabilidade da amostra. Altera-se a função erro, incorporando o valor de confiança w_i, resultando numa função erro expressa pelo sumátorio de residuais quadrados ponderados. Assim

$$f = \sum_{i=1}^{N_s} \frac{1}{N_s} w_i \|Tp_i - q_i\|^2.$$

Análise

- Quanto à região de sobreposição
  Dado o problema crucial de desconhecimento da região de sobreposição e portanto um risco latente de escolha de pares inconvenientes que prejudicariam o computo da transformação, consideram-se importantes as duas contribuições em relação à rejeição de pares de pontos potencialmente inconvenientes. (a) Rejeição de pontos muito afastados e (b) Rejeição de vértices sobre a borda da malha.

  Para com maior precisão representar um par de controle os autores introduzem um valor de grau de confiabilidade das amostras dependente da posição delas em relação ao escaneador.

- Quanto à correspondência e pontos de controle
  A estratégia de escolher o ponto mais próximo encima de uma malha aumenta a precisão da correspondência pois a malha representa uma superfície continua ajustada às amostras e não apenas pontos esparcidos no espaço. É inteligente também o esquema de hierarquia de malhas para evitar o uso de malhas muito densas que fariam muito custoso o processamento. Porém, até que resolução as malhas terão que ter para fornecer amostras significativas para o computo da função erro?. Lembremos como discutido na análise da técnica de Chen e Medioni [5] que em caso de geometrias complexas, apenas amostras com características relevantes favorecem um bom registro.

A.5 Blais e Levine [3]

Eles propõem basicamente três mudanças na técnica ICP

1. Determinar uma subamostragem com distribuição uniforme no plano da imagem (análogo a Turk e Levoy [22], mas sem hierarquias incrementais de malhas).

2. Usa-se um método de projeção para determinar os pontos correspondentes aos pontos de controle. O ponto de controle $p_i \in R_1$ é projetado sobre o plano da imagem R_2, se o ponto cair fora dos limites da plano da imagem ou no "background" da imagem esse ponto é desconsiderado como potencial ponto de controle, caso contrário determina-se o ponto com coordenada (x, y) mais próximo à coordenada (x, y) do ponto projetado. Pode-se acelerar o computo da projeção tornando-o bem simples através da determinação de equações que relacionem as coordenada (x, y) de um ponto 3d com os índices da imagem R_2. A

função custo fica então determinada como

$$f = \sum_{i=1}^{N_s} \frac{1}{N_s} \|Tp_i - P(p_i)\|$$

onde $P()$ é a função projeção descrita.

3. Introduze-se também o limiar de Turk e Levoy para descartar prováveis pares de pontos inconvenientes.

Análise

- Quanto a nova proposta de correspondência, o método de projeção é vantajoso em termos de simplicidade implementacional e custo de tempo em relação às duas proposta de correspondência já mencionadas (a) Ponto mais próximo (Blais e Levine [3]) e (b) Ponto mais próximo na direção do vetor normal (Chen e Medioni [5]). O custo de achar a projeção do ponto é constante uma vez que as equações que relacionam as coordenadas (x, y) com os índices da imagem são determinadas. No entanto a pesar que os autores afirmem que a distância entre uma amostra da imagem R_1 e a sua projeção no plano da imagem de R_2 é a melhor estimativa da distância entre vistas, nós temos encontrado na prática que isso não é bem a verdade. Quando as superfícies representadas pelas duas vistas estão desalinhadas em alguma direção k do plano ortogonal à direção de projeção n_z, a minimização das distancias entre os pontos e suas projeções não resultará num bom alinhamento na direção k.

A.6 Masuda e Yokota [10]

Eles tentam fazer o etapa de minimização de Horn do ICP mais robusta usando o método Mediana mínima de quadrados (Least median of squares - LMS) na função erro ao invés da média mínima de quadrados usada por Horn. Horn minimiza a soma dos residuais quadrados (Least squares - LS) significando que ainda um único par inconveniente dos N_s pares pode gerar uma transformação errada. Já LMS seria mais robusto porque minimizaria a mediana dos residuais quadrados podendo dar uma resposta correta ainda que quase o 50% dos pares sejam inconvenientes. No entanto a minimização da função $f = med\|Tp_i - q_i\|$ não pode ser resolvida em forma analítica como no caso da minimização de $f = \sum_{i=1}^{N_s} \frac{1}{N_s} \|Tp_i - q_i\|$, portanto uma técnica estocástica como amostragem randómica deve ser incorporada. O algoritmo proposto será então iterativo, considerando em cada passo um processo de subamostragem aleatória, estimativa da transformação e avaliação da transformação. O número de iterações N_t do algoritmo é pre-determinado pelo usuário. A estimação da tranformação T_{ICP} é feita usando o algoritmo ICP, a avaliação do erro da transformação estimada é feita usando a função mediana dos quadrados MS

$$MS = \sqrt{med_{1 \leq i \leq N_s} \|Tp_i - P(Tp_i)\|}$$

que retorna o residual mdiano entre pontos de controle transformados por T_{ICP} e as suas projeções em R_2. Finalmente determina-se a transformação T_{ICP} que tenha o menor erro MS entre todas as transformações encontradas nas N_t iterações.

O algoritmo é o seguinte:

1. Inicializa-se como $T_{LMS} = T_0$ transformação inicial
2. Calcule $MS_{LMS} = \sqrt{med_{1 \leq i \leq N_s} \|T_{LMS} p_i - P(T_{LMS} p_i)\|}$
3. Para i=1 até N_t faça
   (a) Extraia-se aleatóriamente um conjunto S_{NS} de N_s pontos de controle de R_1.
   (b) Use-se o conjunto S_{NS} com R_2 para calcular a transformação T_{ICP} através do algoritmo ICP.
   (c) Calcule-se a mediana dos residuais quadrados
   $$MS_{ICP} = \sqrt{med_{1 \leq i \leq N_s} \|T_{ICP} p_i - P(T_{ICP} p_i)\|}$$
   (d) Se $MS_{ICP} < MS_{LMS}$ então $T_{LMS} = T_{ICP}$

A transformação resultante T_{LMS} satisfaz

$$MS_{LMS} = min_{1 \leq k \leq N_t} MS_{ICP}$$

Os autores esperam que através das N_t tentativas pelo menos uma subamostragem exenta de pontos inconvenientes possa ser determinada para possibilitar o computo de uma transformação correta (Masuda e Yokota [10]) usam $N_t = 200$ com $S_{NS} = 5$).
Análise

- Quanto à sensibilidade da função erro
  Os autores contribuem diminuindo a sensibilidade da função erro ao número de pares inconvenientes. Ao ser avaliada apenas a mediana dos residuais, um resultado confiável pode ser obtido sem usar um limiar para descartar pares de pontos inconvenientes, mas isso sujeito a que pelo menos 50% dos pares de pontos sejam convenientes.

- Quanto a pontos de controle
  Uma subamostragem aleatória pode ser um problema quando amostras significativas são excluídas na seleção. Neste caso a transformação resultante que registra as duas imagens poderá não ser tão precisa como deveria.

- Por outro lado esta técnica baseada em N_t escolhas aleatórias introduz preocupação com a lentidão do algoritmo. Em cada uma das N_t tentativas, diferentes subconjuntos de pontos de controle são escolhidos randomicamente, o que gera N_t transformações entre as quais apenas uma será a saída do algoritmo. Isto pode ser traduzido como uma busca exaustiva da matriz de transformação T num espaço de transformações encontradas usando combinações randômicas de S_{NS} pares de pontos.

A.7 Neugebauer [11]

Similarmente a Turk e Levoy [22], ele introduz uma hierarquia de resolução na seleção de pontos de controle, mas a diferença deles nenhuma triangulação é feita. Nas primeiras iterações apenas poucos pontos são usados, e posteriormente pontos adicionais são incluídos. Para montar a hierarquia de resolução escolhe-se pontos cuja coordenada na grade da imagem seja uma potência de 2^d, e dobra-se a resolução a cada novo nível da hierarquia. Como Blais e Levine [3] ele determina como ponto correspondente para cada ponto de controle p_i a projeção desse ponto $P(p_i)$ na imagem R_2, e como Chen e Medioni [5] considera a minimização ponto-plano tendo a mesma função erro

$$f = \sum_{i=1}^{N_s} \frac{1}{N_s} d^2(Sq_i, Tp_i)$$

onde Sq_i é o plano tangente estimado em q_i.

Uma estratégia diferente de rejeição de prováveis pares inconvenientes é proposta baseada na suposição que o residual dos elementos de um par inconveniente se encontra acima da média dos residuais dos outros pares. Calcula-se então em cada iteração do método a média σ dos residuais e usa-se esse valor para detectar prováveis pares inconvenientes na seguinte maneira. Se $\|p_i - q_i\| > 3\sigma$ então despreza-se o par.

Análise

- Quanto à região de sobreposição
  A proposta de um critério diferente de rejeição de pontos é mais uma tentativa de rejeitar pares inconvenientes de pontos, porém não garante uma eliminação completa deles.

A.8 Weik [23]

Ele requer como entrada além das duas imagens de profundidade R_1 e R_2 as duas imagens de intensidade I_1 e $I2$ correspondentes. Começa selecionando pontos de controle bem distribuídos na imagem R_1 e encontra os seus pontos correspondentes em R_2 seguindo o seguinte processo que considera a linearização da informação de luminância. Seja o ponto 2D p_i pertencente ao plano da imagem R_1 e o seu associado ponto 3D, P_i. Seja ainda p_i' a projeção sobre R_2 do ponto 2D p_i. Partindo do presuposto que a intensidade associada a P_i (ponto de controle em R_1) é a mesma que a associada a seu ponto correspondente Q_i em R_2 então a igualdade de diferença de intensidades $\Delta = I_1(p_i) - I_2(p_i') = I_1(q_i) - I_2(p_i')$ é satisfeita, onde q_i é a posição 2D de Q na imagem R_2. Por tanto com essa igualdade e a equação que lineariza I, $\Delta I = \nabla I_2(p_i').d$, podemos calcular $q_i = p_i' + d$ (Figura 13) e desta

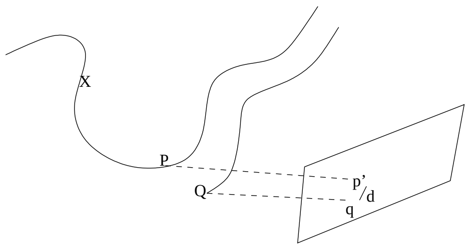

Figura 8: Técnica de Weik

maneira o ponto correspondente para P_i será Q_i, o ponto 3D associado ao ponto imagem q_i na imagem R_2.

Para rejeitar prováveis pares inconvenientes aproxima-se a imagem R_2 por uma malha triangular. Um ponto p_i é desconsiderado se um raio saindo do ponto de vista de R_1 até ele intercepta a malha em outro ponto com profundidade menor, como no caso do ponto X na Figura 13.

Análise

- Quanto à correspondência
  Tendo a informação de luminância e seguindo o método proposto, a determinação dos pontos correspondentes é direta, mais simples e menos custosa que a procura do ponto mais próximo de Besl e McKay [2] e a determinação do ponto correspondente proposta por Chen e Medioni [5], porém isso esta sujeito à boa qualidade das imagens de intensidade que devem ter idénticas intensidades nos pontos comuns. Um erro nesse sentido pode levar à determinação errada das correspondências.

- Quanto à região de sobreposição
  O teste de visibilidade de Gagnon e outros [8] é aprimorado nesta proposta. De fato, enquanto os primeiros realizam apenas um teste "front-face"que considera ainda pontos que podem estar "ocultos"por regiões da superfície com profundidade menor, o teste de Weik os elimina.

A.9 Pulli [13]

Pulli considera que nenhuma estratégia de rejeição de pontos satisfaz e por isso orienta seus esforços para determinar pares potencialmente convenientes de pontos utilizando as duas imagens coloridas correspondentes às imagens de profundidade. Primeiro projeta a amostragem da imagem R_1 sobre o plano da imagem R_2 e vice-versa, registra a informação 2D alinhando as características de cor das imagens e depois toma N_s pares de pontos 3D de controle a partir das suas posições 2D alinhadas na imagem. Com isso os pares determinados terão muitas boas chances de serem pares convenientes, mas não certeza. Poderá haver alguns pares com distancia muito grande entres seus elementos, como no caso de uma borda de descontinuidade de profundidade, onde um pixel sobre a região de maior profundidade pode (por um ligeiro erro de registro 2D) ser alinhado com o pixel na região de menor profundidade. Para eliminar esses pares inconvenientes opta-se por uma nova estratégia de rejeição que não é precisamente o critério de limiar. Similarmente a Masuda e Yokota [10] que propõem LMS, Pulli propõe uma mudança na função erro usando LTS (Mínimos quadrados recortados). Enquanto que LMS minimiza a mediana dos residuais quadrados e pode dar uma resposta correta ainda que quase o 50% dos pares sejam inconvenientes, LTS minimiza a soma dos h residuais quadrados menores e pode dar o resultado correto quando mais do 50% dos pares são inconvenientes.

$$TS = \sum_{i=1}^{h} (r^2)$$

onde $(r^2)_{(1:N_s)} \leq \ldots \leq (r^2)_{(N_s:N_s)}$ são os residuais quadrados ordenados.

$$r = \|Tp_i - P(Tp_i)\|$$

Uma vez obtida a transformação T_{LTS} usando LTS com N_t tentativas, subseleção randômica e LTS com $h = 0.5 * Ns$ para avaliar o erro das transformações intermediárias (num algoritmo similar a LMS), Pulli refina ainda esta estimativa T_{LTS} através da aplicação de LTS com atualizações dinâmicas do valor h. O valor h é visto como uma razão entre uma estimativa do número de pares convenientes e o numero N_s das amostras do conjunto aleatório escolhido inicialmente. A estimativa do numero de pontos convenientes é feita usando o seguinte critério: Contabiliza-se os pares de pontos cujo residual é menor ou igual a 2.5 vezes o desvio padrão dos N_s residuais. O algoritmo completo é dado a seguir:

1. Inicializa-se com $T_{LTS} = T_0$
2. Faça $p = 0.5$
3. Calcula-se a função erro com $h = p * N_s$ (f_{LTS})
4. Enquanto (o erro LTS reduz ou p incrementa) fazer
    (a) Casamento
        i. Para cada ponto de vista (1 e 2) fazer
            A. Projetar a malha da outra vista
            B. Registro 2D das imagens coloridas
            C. Pegar pontos no mesmo pixel
    (b) Minimização
        i. LTS
            A. Repetir N_t vezes (número de tentativas)
                - Escolher S_{N_s} pares
                - Usa-se os pares para computar a transformação T_{ICP} usando ICP.
                - Calcula-se a função erro com $h = p * N_s$ (f_{ICP})
                - Se $f_{ICP} < f_{LTS}$ então $T_{LTS} = T_{ICP}$.
        ii. Refinamento LTS
            A. Aplicar T_{LTS}
            B. Calcula residuais r para os N_s pontos
            C. Estimar desvio padrão τ dos N_s residuais

D. Rejeitar os pares cujo residual excede $2.5 * \tau$, ficando S_τ pares.
E. Calcular novo $p = S_\tau / N_s$
F. Usa-se os N_s pares para calcular a nova transformação T_{LTS}
G. Aplicar o novo T_{LTS}

Análise

Pela sua semelhança com LMS a analise será feita tendo esse método como referência.
Quanto ao grau de confiança da transformação, vemos que o fato de h ser livre nos permite calcular a transformação com mais segurança em função dos prováveis pares convenientes pertencentes ao conjunto dos N_s pares. Se por exemplo 80% dos pares são convenientes, para determinar a qualidade da T_{LTS} é melhor somar o 80% dos residuais quadrados menores que considerar apenas a mediana desses residuais. A pesar que o valor estimado do número de pares inconvenientes depende de um limiar heurístico, a idéia é diferente e representa uma contribuição.

A determinação do conjunto de N_s pares de pontos é feito mais consistentemente pois são elegidos a partir de um registro $2D$ das imagens coloridas e não randomicamente como Masuda e Yokota [10] propõem. De fato os N_s pares de pontos tem mais chance de ser convenientes, que os apenas escolhidos aleatoriamente.

Quanto ao tempo, considerando que os pares de pontos são na sua maioria pares convenientes devido à escolha e o refinamento, tem-se um algoritmo que converge mais rapidamente que LMS. Segundo os experimentos de Pulli, LTS converge em 3 ou 5 iteracoes, enquanto LMS converge em aproximadamente 50.

Porém cabe salientar que a determinação de pares convenientes depende fortemente das imagens coloridas ter boa qualidade e do registro 2D ser confiável, e sem essas condições pouco será garantido.

A.10 Dorai et.al. [7]

Eles tentam aprimorar o trabalho de Chen e Medioni [5] introduzindo um pré-procesamento para eliminar ruído e um novo esquema de rejeição de prováveis pares de pontos inconvenientes baseado no princípio de rigidez: Uma transformação rígida preserva distâncias. Seguindo este princípio define-se o conceito de compatibilidade entre pontos. Dois pares (p_i, q_i) e (p_j, q_j) são compatíveis se a distância $\|p_i - p_j\|$ e aproximadamente igual à distância $\|q_i - q_j\|$.

Na sua implementação é usado o seguinte teste de compatibilidade

$$-0.1 \leq \frac{\|p_i - p_j\| - \|q_i - q_j\|}{max(\|p_i - p_j\|, \|q_i - q_j\|)} \leq 0.1$$

Usando o conceito de compatibilidae, propõe-se o seguinte algoritmo que valida a conveniência de N_s pares de pontos

1. Para cada $i, j \in \{1, 2, ..., N_s\}$ e $i \neq j$ verifique se os pares de pontos (p_i, q_i) e (p_j, q_j) são compatíveis
2. Remova-se o par de pontos correspondentes que não são compatíveis com a maioria de pares
3. Itere passos (1) e (2) até
   (a) Todos os pares de pontos sejam compatíveis uns com outros
   (b) Acima de um pré-definido limiar máximo (50% de N_s) seja removido
   (c) O número de pares de pontos restantes seja menor que um pré-definido limiar

Análise

- Quanto ao teste de conveniência proposto pelos autores, devido a estar baseado no princípio de rigidez parece ser o mais consistente do que os até o momento mencionados, fornecendo pares mais confiáveis. No entanto é exaustivo (verificação de um par de pontos com todos os outros pares).

A.11 Rusinkiewicz e Levoy [15]

Eles classificam e comparam diversas variantes do algoritmo ICP, se preocupando na velocidade de convergência. Baseados na sua análise, constroem um algoritmo ICP altamente veloz combinando algumas das variantes analisadas. Quanto à seleção de pontos, eles por economia de tempo propõem seleção randómica, porém não deixam de enfatizar que este tipo de escolha pode levar ao ICP a não convergir no alinhamento correto quando as amostras significativas são desconsideradas (Figura 12). Para resolver este problema eles propõem escolher os pontos de controle tal que a distribuição dos vetores normais entre as amostras selecionadas seja tão espaçada como possível. A Figura 14 apresenta uma ilustração comparativa da escolha randômica e a

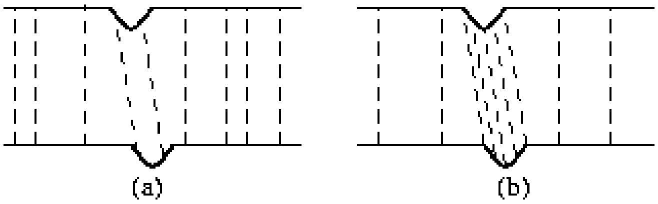

Figura 9: (a) Escolha randômica (b) Escolha no espaço dos normais

escolha de pontos com boa distribuição de vetores normais numa seção 2D da imagem da Figura 12.

O fato de escolher os pontos de controle no espaço do vetores normais faz que o ICP atinja a convergência na solução exata no caso apresentado, onde a seleção randômica fracassa.

Em relação à determinação do ponto correspondente, sempre preocupados com o custo de tempo optam pelo método de projeção usado por Blais e Levine [3] e Neugebauer [11]. Como já mencionado o custo de tempo para procura do correspondente neste caso é constante.

Para rejeitar prováveis pares inconvenientes eles optam pelo critério do limiar proposto por Turk e Levoy [22] por ser menos custoso de implementar que as opções LMS, LTS e o esquema de rejeição de Dorai e outros [7] baseado no princípio de rigidez.

Quanto à métrica usada para avaliar o erro da transformação eles escolhem a métrica ponto-plano por ter demonstrado converger mais rapidamente que a métrica ponto-ponto.

Análise

- O método proposto pelos autores é uma combinação de algumas das variantes já apresentadas da técnica ICP com a diferencia na estratégia de seleção.

- Quanto à seleção de pontos de controle
  Embora a idéia de seleção com boa distribuição de normais represente uma boa contribuição à pesquisa sobre aprimoramento da escolha, podemos distinguir a seguinte restrição: Já que a informação espacial não é levada em conta, esta estratégia de escolha pode deixar de considerar *características significativas* em outras regiões da superfície, como podemos observar na Figura 15

- Quanto à distância ponto-plano
  Até aqui todas as variantes de ICP que usam minimização ponto-ponto

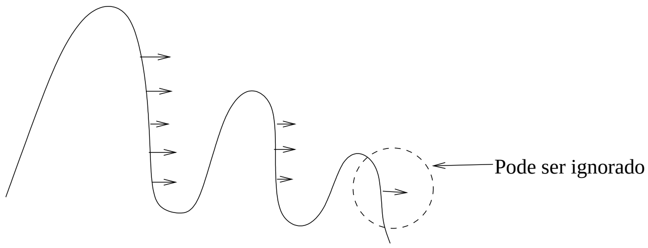

Figura 10: Escolha de pontos de controle no espaço de vetores normais

utilizam a distância euclideana para avaliar o erro da transformação, e só apenas aquelas que usam a minimização ponto-plano como Chen e Medioni [5] e Neugebauer [11] usam a distância ponto-plano para esta avaliação. A proposta dos autores é uma tentativa de considerar não apenas um ponto no espaço mas uma representação continua da superfície para calcular a distância. Desta forma, devido a envolver informação da superfície, temos confirmado com experimentos que converge mais rapidamente.

A.12 Chu-Song e outros [4]

Propõem um novo método chamado DARCES (data-aligned rigidity-constraint exhaustive search - Busca exaustiva restrita a rigidez de dados alinhados). Este método baseia-se no princípio de rigidez que afirma que sob transformações rígidas as distâncias são preservadas.

A partir de três pontos de controle em R_1 (triplo de controle tc) com distâncias entre os seus vértices d_1, d_2, d_3, realiza-se uma busca exaustiva de todos os possíveis triplos em R_2 que possuam a mesma distância d_1, d_2, d_3 entre seus vértices. Em outras palavras determina-se um triplo de controle com o qual é feita uma varredura completa na amostragem da imagem R_2 em todas as possíveis direções procurando um encaixe (triplo correspondente) o mais exato possível.

Neste varredura muito triplos serão encontrados. Para cada cada triplo t_i encontrado, computa-se a tranformação que leva o triplo de controle tc em alinhamento com t_i. Entre todas essas possibilidades determina-se a transformação que menor erro de ajuste reporte. O algoritmo é o seguinte:

1. Selecione um conjunto de pontos de referência em R_1

2. Selecione entre esse conjunto, três pontos de controle (triplo de controle)

3. Determine todos os possíveis triplos em R_2 que satisfaçam o princípio de rigidez

4. Para cada triplo t_i em R_2

    5 Encontre a transformação T_i de registro entre o triplo de controle tc e o triplo t_i

    6 Transforme os pontos de referência aplicando T_i

    7 Conte o número (n_{T_i}) de pontos suficientemente próximos à imagem R_2

8 Escolha $T = T_i$ com maior número de pontos n_{T_i}

A procura do triplo correspondente é feita da seguinte forma: Sejam S_p, S_s, S_a os pontos do triplo de controle, e M_p, M_s, M_a os pontos correspondentes a ser determinados em R_2. Devido a que nenhum conhecimento se tem da transformação, o ponto M_p pode ser qualquer ponto em R_2, no entanto usando o princípio de rigidez uma procura numa área mais restrita pode ser feita para encontrar M_s e M_a. Assumindo que a distância entre S_p e S_s é d_{ps}, entre S_p e S_a é d_{pa} e entre S_s e S_a é d_{sa}, o ponto M_s se encontrará na superfície da esfera C_{ps} centrada em M_p e com raio d_{ps}. Da mesma forma, o ponto M_a pode ser encontrado na interseção das superfícies das esferas C_{pa} com centro em M_p e raio d_{pa} e C_{sa} com centro em M_s e raio d_{sa}. Na implementação os pontos sobre a superfície de uma esfera podem ser procurados no menor quadrilátero que contém a região circular resultante de projetar a esfera sobre o plano da imagem R_2.

O algoritmo para a determinação dos triplos em R_2 (passo (3) do algoritmo anterior) é

1. Escolha os três pontos do triplo de controle (S_p, S_s, S_a)

2. Para (M_p=cada ponto contido em R_2)

    (a) Para ($M_s =$ cada ponto no quadrilátero Q correspondente à projeção da esfera C_{ps})

        i. Se a distância entre M_p eM_s é aproximadamente d_{ps}

            A. Para (M_a = cada ponto na interseção dos quadriláteros correspondentes à projeção das esferas C_{pa} e C_{sa})

                - Se a distância entre M_p e M_a é aproximadamente d_{pa} e a distância entre M_s e M_a é aproximadamente d_{sa} Define-se o triplo candidato de pontos M_p, M_s, M_a

Já que a busca é exaustiva apenas um triplo de controle seria suficiente para achar a solução, mas para isso acontecer o triplo precisaria estar na região de sobreposição. Como esta região é desconhecida os autores propõem uma escolha aleatória dos pontos do triplo esperando que após n_t tentativas o algoritmo determine os pontos do triplo na região de sobreposição. O mesmo pode ser dito ao respeito da escolha dos pontos de referência. Como eles determinarão a qualidade da transformação devem também fazer parte da região de sobreposição, caso contrário não serão uma boa referência para a medição do erro. Tendo em conta isso, propõe-se que o algoritmo inicial apresente as seguintes modificações nos passos 1, 2 e 8.

1. Os pontos de referência são randomicamente escolhidos
2. Entre eles escolhe-se os três pontos do triplo de controle

8 Se o número de pontos n_{T_i} é maior que um pré-definido limiar Ω, então a saída é $T = T_i$, caso contrário tente outro conjunto de pontos de referência (Vá a 1.)

O limiar Ω será uma estimativa da razão de sobreposição mínima permitida dos dois conjuntos parcialmente sobrepostos e deve ser pré-definida pelo usuário.

Análise

Não havendo nenhum conhecimento prévio da transformação, os autores optam por um método intuitivo, a construção de um espaço das possíveis transformações do ponto de controle, entre as quais encontra-se a transformação T. Vários pontos de esta proposta merecem ser analisados. Antes disso definimos alguma terminologia coerente com a já existente. Chamamos de triplo correspondente ideal do triplo de controle tc ao triplo cujos pontos são correspondentes ideais dos pontos de controle do triplo tc. Chama-se de triplo conveniente ao triplo cujos pontos ocupam aproximadamente a mesma localização da superfície amostrada, que os pontos de tc.

- Quanto à correspondência
  Em principio se temos que tc pertence à região de sobreposição de R_1 podemos afirmar que o triplo ideal de tc pode ser obtido através da transformação T aplicada em tc. No caso de ruídos ou erros de aquisição, a mesma afirmação é valida para um triplo conveniente. Analisando o método vemos que a busca do triplo correspondente conveniente é feita exaustivamente no espaço de todas as possíveis transformações do triplo de controle, portanto pelo dito pelo parágrafo ante-

rior a determinação de um triplo correspondente conveniente é perfeitamente viável. Porém a principal desvantagem desta procura é o custo muito grande de construção de todos os possíveis correspondentes, o computo da transformação que alinha o triplo de controle com cada candidato e a avaliação de erro de cada transformação candidata.

- Quanto à escolha do triplo de controle
  Foi dito na análise anterior que o resultado mais exato pode ser garantido quando o triplo de controle pertencer à região de sobreposição e ruídos ou erros de aquisição da amostragem não distorcem significativamente o triplo correspondente ideal. O desconhecimento da região de sobreposição pode em principio orientar a propor todas as possíveis combinações de três pontos como triplos de controle, porém a preocupação dos autores de não onerar mais o tempo os levou a propor uma escolha randômica. No entanto nada garante que após um número pequeno de tentativas o triplo seja escolhido na região de sobreposição. No pior dos casos (região de sobreposição muito pequena) todas as possíveis combinações $n_c$de três pontos serão consideradas, e neste caso o grande custo de tempo de processamento faria o algoritmo muito impratico. O mesmo pode se dizer em relação à escolha dos pontos de referência, nada garante que após $n_t < n_c$ tentativas eles serão escolhidos na região de sobreposição. Se eles não fizerem parte dela a avaliação do erro da transformação candidata não é confiável.

A.13 Gerhard Roth [14]

Ele usa como entrada além das imagens de profundidade R_1 e R_2, as suas correspondentes imagens de intensidade I_1 e I_2. A idéia básica é construir um espaço de possíveis transformações de onde a transformação procurada T possa ser extraída. As possíveis transformações são determinadas considerando pares de triângulos da triangulação Delaunay de pontos de intensidade relevante nas imagens. Para calcular os pontos de interesse determinam-se primeiro os máximos locais do gradiente da intensidade, sendo considerados aqueles acima de um determinado limiar são classificados como pontos de interesse, com estes pontos realiza-se uma triangulação Delaunay dos pontos 3D correspondentes aos pontos de interesse em ambas as imagens. Antes de partir para a construção do espaço de transformações tenta-se aliminar os pares de triângulos (t_{1i}, t_{2i}) onde $t_{1i} \in R_1$ e $t_{2i} \in R_2$ que não tem chances de gerar a transformação T, ou seja um proceso de filtragem baseado no princípio de rigidez é feito para determinar apenas os possíveis pares de triângulos convenientes. Para saber a correspondência entre vértices dos pares de triângulos, um teste é feito sobre os vértices baseado nos seguintes princípios derivados do princípio de rigidez.

- A diferença angular entre os vetores normais de cada par de vértices deve ser invariante sob transformação rígida
- As distâncias de um ponto de interesse até os outros pontos de interesse (se eles estiverem na região de sobreposição) deve ser aproximadamente a mesma.

Depois de que N_s pares de triângulos forem filtrados, calcula-se a transformação que leva um triângulo do par em alinhamento com o outro e determina-se a transformação que alinha o maior número de pontos de interesse.

Análise
Este é outro método de busca no espaço de transformações, porém a diferencia da técnica anterior, os triângulos de controle e o seu correpondente foram determinados usando a informação de intensidade. Identificamos dois problemas que o método enfrenta

1. Dependência forte da determinação precisa das características relevantes de luminância, as quais devem ser idênticas em ambas as imagens.
2. Falha do registro se nenhuma informação de luminância relevante existe na região de sobreposição.

B Técnicas de Integração

Uma vez que as imagens são alinhadas num único sistema de referência o passo seguinte é a fusão delas fazendo um devido tratamento nas regiões redundantes e gerando finalmente uma única representação não ambígua. (Figura 11).

Métodos para integrar imagens de profundidade podem ser classificados em duas abordagens. A primeira é a abordagem baseada na representação malha triangular, e a segunda é a abordagem volumétrica. Estes pela sua vez podem ser subdivididos em métodos dinâmicos ou estáticos, os primeiros permitindo a incorporação incremental de uma imagem num modelo constituído por uma ou várias imagens já integradas. E os estáticos que requerem apriori a informação de todas as imagens a serem integradas. Cabe

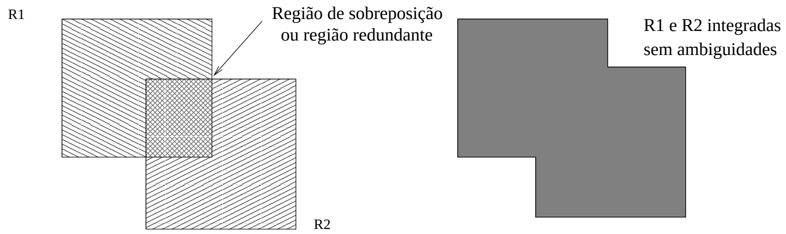

Figura 11: Integração de duas imagens

mencionar também que usualmente trabalha-se com a malha triangular representando a imagem de profundidade, ao invés de simplesmente com as amostras no espaço. No processo de triangulação, por simplicidade a maioria das vezes usa-se a técnica de triangulação restringida à descontinuidade no passo [22, 16, 20, 18, 21, 6] que gera os triângulos aproveitando a estrutura 2D da imagem. Malhas adaptativas à geometria dos dados podem também ser usadas [17].

A abordagem malha triangular é chamada assim por que o processo completo de integração é feito considerando uma representação de bordo triangulado das vistas do objeto amostrado a serem integradas.

Na abordagem volumétrica o modelo inicial (malha triangular) passa num estágio intermediário a se tornar um modelo volumétrico, aonde a cada voxel é atribuído um valor que representa a distância de um raio a saindo do ponto de vista da imagem R_i, atravesando o voxel e chegando até a malha representando R_i. A função distância define uma função implícita e um algoritmo de poligonização é usado para obter a malha final.

A seguir apresentamos sete técnicas de integração, as seis primeiras correspondentes à abordagem malha triangular e a última à abordagem volumétrica.

B.1 Turk e Levoy [22]

Usam uma técnica incremental de erosão parcial na qual os triângulos na região de sobreposição são eliminados a partir das bordas das malhas M_1 e M_2 até que elas fiquem apenas "ligeiramente" sobrepostas (Figura 12a). Então os pontos de interseção entre arestas da malha M_1 e a borda de M_2 são calculados e inseridos na borda de M_2 (Figura 12b). Após isso, as áreas de triângulos redundantes de M_1 são eliminadas (Figura 12c) e uma retriangulação considerando os novos vértices é realizada obtendo uma única malha (Figura 12d). Posterior processamento é necessário para remover triângulos pequenos e finos.

Depois de todas as malhas forem seqüencialmente integradas, a posição

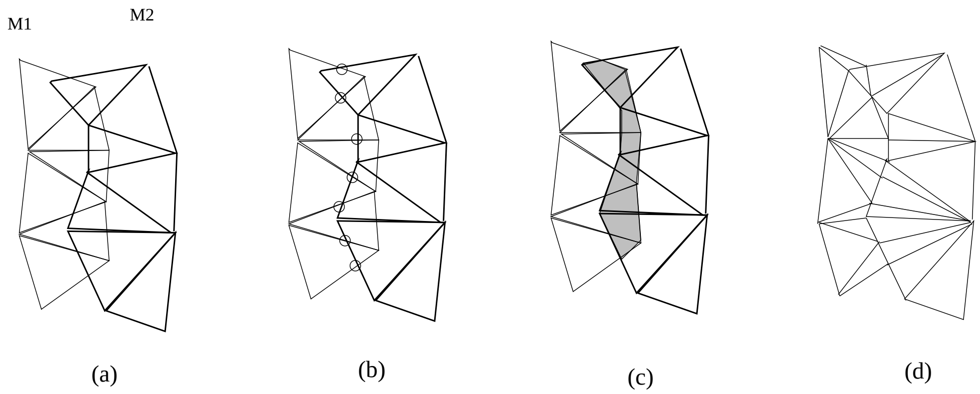

Figura 12: Integração de duas imagens segundo Turk e Levoy

definitiva de um vértice é considerada como a média das posições em cada imagem de profundidade em que ele foi amostrado.

A determinação dos triângulos redundantes é feita da seguinte maneira: Um triângulo em M_1 é redundante se os seus vértices se encontram a uma distância menor que um limiar d em relação a M_2. O triângulo redundante é removido se pelo menos dois dos seus vértices tem menor grau de certeza que os seus correspondentes em M_2.

B.2 Soucy e Laurendeau [19, 20]

Apresentam um método estático de integração que permite a entrada de n imagens de profundidade registradas, e consiste em modelar (identificar, triangular e conectar) conjuntos de informação redundante entre vistas (subconjunto canônico do diagrama de Venn).

Um subconjunto canônico consiste de um conjunto de elementos que pertencem exclusivamente a k imagens de profundidade, e a união dos subconjuntos canônicos representa o diagrama de Venn (Figura 13). Para obter o conteúdo de todos os subconjuntos canônicos, deve-se determinar a informação comum entre pares de imagens, através dos dois seguintes testes:

- Teste de Vizinhança espacial.- Um ponto p em M_i pertence à região de sobreposição se ele está suficientemente próximo à uma região espacial em M_j definida pelos três vizinhos do ponto p projetado na grade paramétrica de R_j.

- Teste de Visibilidade.- p é visível a partir de M_j se o ângulo entre o normal em p e o vetor orientação de M_j é menor que 90^o.

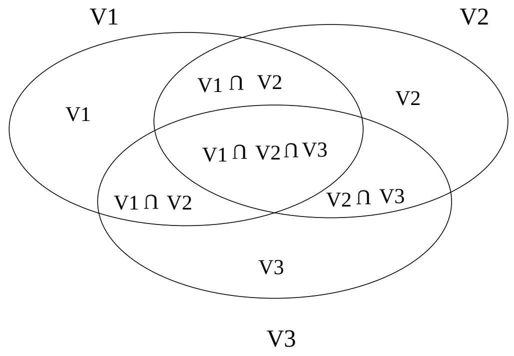

Figura 13: Subconjuntos canônicos do Diagrama de Venn de três imagens

A modelagem ou triangulação de um subconjunto canônico redundante C_k é feita em duas etapas:

- Reparametrização das k malhas redundantes.- O objetivo é construir uma representação que fusione a contribuição das k malhas.

  Uma grade paramétrica planar G é definida para C_k com um vetor de orientação igual a média dos vetores de orientação das k imagens. Os triângulos que tem um vértice k vezes redundante são projetados sobre G e os valores de profundidade da grade G são então interpolados (Figura 14). Esses k valores são integrados usando uma média pon-

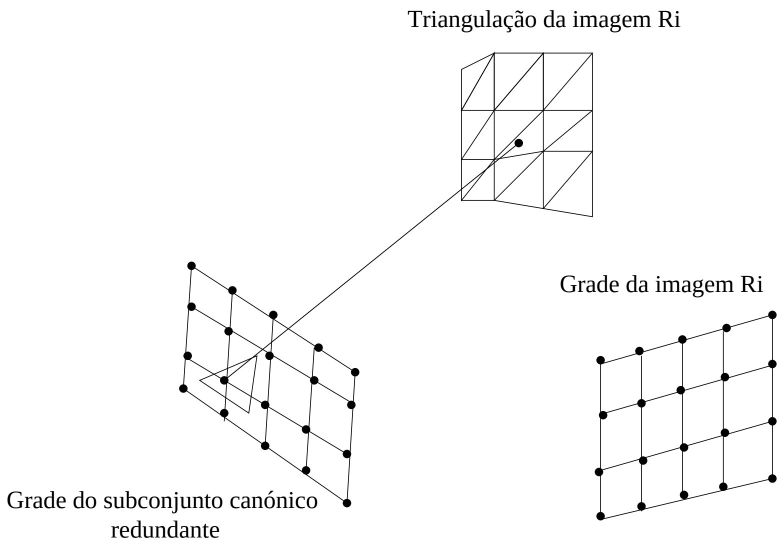

Figura 14: Reparametrização da um conjunto canônico redundante

  derada onde os pesos são estimativas do grau de incerteza do dispositivo de adquisição.

- Triangulação restringida do subconjunto canônico reparametrizado.

A etapa de costura de modelos parciais consiste em três passos:

- Impõe-se uma distância paramétrica mínima entre as triangulações dos subconjuntos canônicos. Esta distância tem como objetivo separar as triangulações para que a costura seja mais eficiente.
- Triangula-se o espaço vazio resultante, usando triangulação Delaunay.
- Elimina-se os triângulos finos e longo que tem seu ângulo interior menor que um certo limiar.

B.3 Pito [12]

Na sua proposta de integração de duas malhas M_1 e M_2, Pito determina os triângulos redundantes a partir de informações de posição e orientação do escaneador. Uma vez que os triângulos redundantes são detectados, aquele de maior grau de incerteza é retirado. O resultado é um conjunto de malhas desconexas que cobrem as áreas amostradas do objeto. Relações de vizinhança dos pontos da borda de cada malha desconexa são usadas para costurar as brechas entre elas, tomando cuidado de que cada triângulo novo não cause autointerseção na malha (Figura 15).

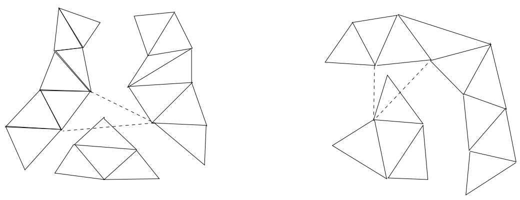

Figura 15: Triângulos adicionados não devem causar autointerseção

O processo de costura pode ser realizado em duas formas

- A brecha é prenchida começando com duas arestas sementes em opostas direções, para então criar dois novos triângulos entre elas (Figura 16a e b). As arestas que unem ambos lados da costura são chamadas de arestas-fronteira. A costura prossegue escolhendo o melhor entre os dois triângulos que podem ser criados a partir da aresta-fronteira e o ponto da sua aresta adjacente, de qualquer dos lados da costura (Figura 16c). O melhor triângulo é aquele com menor razão de proporção entre o comprimento de suas arestas. A aresta fronteira é atualizada e o processo repetido.

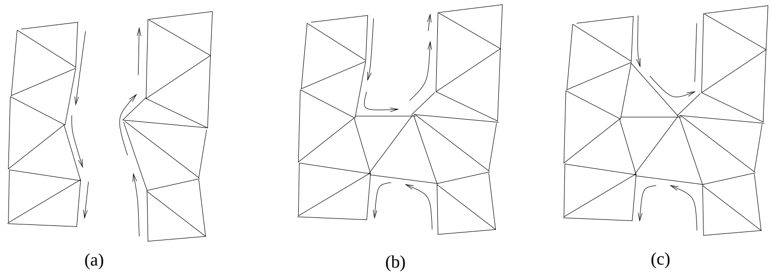

Figura 16: Costura segundo Pito

- Quando o algoritmo de costura (a) não pode ser usado, adicionam-se pontos vizinhos às arestas de borda, criando novos triângulos da seguinte maneira:

  1. Escolhe-se a malha M_1 como modelo inicial M_0
  2. Inicia com uma aresta pertencente à borda de M_0 (chamada aresta-semente)
  3. Obtém os k pontos mais próximos à aresta-semente.
  4. Determine entre os k pontos, àquele p' que enxerga a aresta-semente sob o ângulo maior
  5. Se o triângulo formado com a inserção de p' forma com o triângulo adjacente que compartilha a aresta-semente, um ângulo menor que 90^o, então rejeita p' e repita 4.
  6. Teste se o triângulo não sobrepõe o modelo, se assim repita 4.
  7. Adicione o triângulo válido ao modelo e atualize a sua borda.
  8. Escolhe a próxima aresta-semente
  9. Repita 3-9 até que novos triângulos não possam ser mais acrescentados no modelo.

B.4 Shutz e outros. [18]

Propõem uma técnica de integração incremental de novas vistas. Na deteção das redundâncias, eles usam os pares de pontos de controle (p, q) onde $p \in M_1$ e $q \in M_2$ obtidos no processo de registro (um algoritmo ICP) com limiar d para rejeitar pontos muito afastados. Eliminam-se os pontos correspondentes em M_2 aos pontos de controle pertencentes a M_1 obtendo uma brecha de tamanho aproximado d entre as malhas M_1 e M_2. O costura sobre a brecha é feita com um algoritmo similar ao de Pito [12].

B.5 Sappa [17]

Propõe uma técnica de integração que similarmente à proposta de Soucy e Laurendeau, realiza a projeção planar da região redundante para reduzir o tratamento dessa região numa dimensão.

O seu trabalho consiste dos seguintes estágios: Primeiro ele remove as regiões de sobreposição (Figura 17a) e tomando em consideração essas regiões

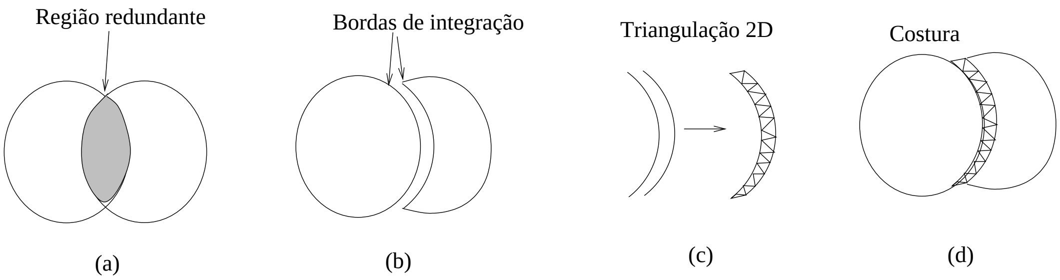

Figura 17: Algoritmo de integração incremental segundo Sappa

define as bordas de integração (Figura 17b) as quais são projetadas num plano de integração. Os pontos pertencentes a estas bordas são então triangulados no espaço 2D (Figura 17c) e finalmente a triangulação obtida é levada ao espaço 3D. As bordas trianguladas ao serem levadas ao espaço "costuram" as malhas (Figura 17d).

No algoritmo de deteção de triângulos redundantes Sappa considera que um triângulo t de M_2 é redundante (a) se os seus vértices se sobrepõem com algum dos seus triângulos vizinhos em M_1, e (b) o vetor normal tem uma orientação próxima a do vetor normal do triângulo sobreposto em M_1.

Uma vez que os triângulos vizinhos a t em M_1 são determinados, procede-se a verificar se os vértices de t se sobrepõem a algum deles.

Um vértice de t é considerado a ser sobreposto com o triângulo T se pertence ao politopo de sobreposição superior ou inferior de T. O politopo superior de sobreposição de um triângulo é a interseção entre o *half-space* positivo do plano de T e três planos superiores de sobreposição associados com as arestas de T. Dada uma aresta E de T, o plano superior de sobreposição associado a E é o plano que contém E e é ortogonal ao plano do triângulo T' adjacente a T ao longo de E, se T' existir, ou ortogonal ao plano que contém T se T' não existir ou se os triângulos T e T' formam uma superfície côncava.

Por outro lado o politopo inferior de sobreposição é o simétrico ao politopo superior em relação ao plano que contém T.

Na Figura 18 observa-se uma seção 2D de três triângulos com o politopo

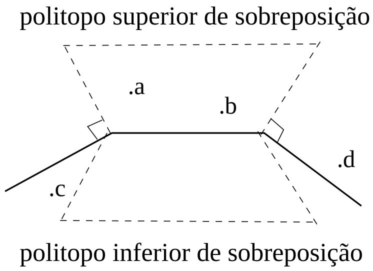

Figura 18: Politopos superior e inferior de sobreposição

superior e inferior de sobreposição do triângulo central. Neste exemplo pode-se ver que os pontos a e b são considerados sobrepostos por estar contidos no politopo superior, enquanto c e d não o são.

Antes que os triângulos redundantes em M_2 sejam retirados determina-se a borda de integração de M_1 pelo conjunto orientado de arestas que conectam vértices da borda de M_1 que são sobrepostos por triângulos em M_2. Depois de remover a informação redundante, a borda de integração de M_2 será determinada pelas arestas que ligam os vértices que passam a ser da nova borda de M_2. Estas bordas de integração são então projetadas sobre o plano de referência cujo vetor normal é a média dos vetores normais associados como os triângulos de M_1 que tem alguma das suas arestas pertencentes às bordas de integração.

O conjunto de pontos que define as bordas de integração são triangulados através de uma triangulação Delaunay 2D. Logo, as arestas de borda da triangulação 2D são removidas, levando finalmente os triângulos ao espaço 3D.

B.6 Yiyong e outros [21]

Apresentam um método de integração, resultando num trabalho similar ao de Pito [12] com diferenças no algoritmo de deteção de ambigüidades e com mais simplicidade no algoritmo de costura.

Na deteção de sobreposição cada triângulo t de M_1 é projetado sobre o plano de referência P_{M2} de M_2, e se ele ficar contido dentro de P_{M2} e estiver "de frente" a P_{M2} (teste de visibilidade) verifica-se a possibilidade de sobreposição da seguinte maneira:

- Calcula-se a caixa limitante da sua projeção em P_{M2},

- Determina-se os triângulos pertencentes a M_2 que estão dentro da caixa limitante

- Verifica-se a interseção com cada um deles.

Há dois casos de interseção entre triângulos como ilustra a Figura 19.

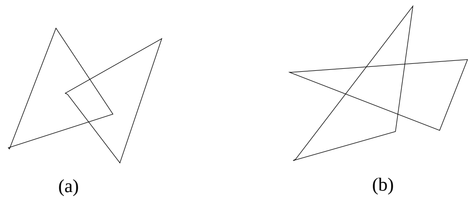

Figura 19: Casos de interseção de dois triângulos

Uma vez que todos os triângulos da caixa limitante são verificados, procede-se a eliminação do(s) triângulos de maior incerteza. Para isso calcula-se a média do grau de incerteza dos triângulos sobrepostos em M_2, se ela for menor que o grau de incerteza de t, elimina-se t, caso contrário eliminam-se os triângulos sobrepostos pertencentes a M_2.

Para garantir maior precisão no processo de deteção, a sobreposição é também verificada projetando os triângulos da caixa limitante no plano de referência de t.

Uma vez que os triângulos redundantes de menor confidência são retirados, um algoritmo de costura preenche a brecha entre as malhas. Para isso rotula-se triângulos candidatos chamados triângulos ativos cujas arestas combinadas com outros pontos possam originar novos triângulos. Esses triângulos encontram-se sobre as bordas da malha e podem ter uma, duas ou ainda três arestas que devem encontrar um ponto para construir um novo triângulo. Para cada uma destas "arestas ativas", encontram-se os k pontos mais próximos como candidatos. Escolhe-se o ponto que enxerga a aresta sob o maior ângulo e verifica-se se não interceptará os triângulos já existentes. O ponto validado será usado para criar o novo triângulo. O processo é repetido até esgotar as arestas ativas.

B.7 Curless e Levoy [6]

Apresentam um método de integração incremental. Cada imagem de profundidade é triangulada e um peso é atribuído a cada um dos vértices em relação ao grau de incerteza. Então uma função distância com sinal D_{i+1} definida para cada vértice v é atualizada na integração da (i+1)-ésima imagem de profundidade da seguinte forma:

$$D_{i+1}(v) = \frac{W_i(v)D_i(v) + w_{i+1}(v)d_{i+1}(v)}{W_i(v) + w_{i+1}(v)}$$

onde $w_{i+1}(v)$ e $d_{i+1}(v)$ são respectivamente o peso e a distância do ponto v até a (i+1)-ésima imagem de profundidade (representada pela malha triangular) e $W_i(v)$ e $D_i(v)$ são a função peso e a função distância acumuladas após a integração da i-ésima imagem. E

$$W_{i+1}(v) = W_i(v) + w_{i+1}(v)$$

Wu Shin-Ting	Mercedes Gonzales Márquez
Orientadora	Bolsista

Referências

[1] M. Berg and et.al. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, 2 edition, 2000.

[2] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[3] 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.

[4] Chu-Song Chen and et.al. Ransac-based darces:a new approach to fast automatic registration of partially overlapping range images. Technical report, Institute of Information Science, Taiwan, 1998.

[5] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, pages 145–155, 1992.

[6] 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, August 1996. SIGGRAPH 96.

[7] C. Dorai, G. Wang, A.K. Jain, and C. Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(1):83–89, 1998.

[8] 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.

[9] B.K.P. Horn. Closed-form solution of absolute orientation using unit quaternios. *Jornal of the Optical Society A*, 4(4):629–642, 1987.

[10] 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.

[11] Peter Neugebauer. Geometrical cloning of 3d objects via simultaneous registration of multiple range imagens. In *International Conference on Shape Modeling and Applications (SMA '97)*, Aizu-Wakamatsu, JAPAN, 1997.

[12] R. Pito. Mesh integration based on co-measurements. In *Proceedings of ICIP*, volume 2, pages 397–400, Laussane, Switzerland, 1996.

[13] K. Pulli. *Surface Reconstruction and Display from Range and Color Data*. PhD thesis, University of Washington, 1997.

[14] 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.

[15] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Third International Conference on 3D Digital Imaging and Modeling*, pages 145–152, 2001.

[16] M. Rutishauser, M. Stricker, and M. Trobina. Merging range images of arbitrarily shaped objects. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Patter Recognition*, pages 573–580, Seattle, WA, 1994.

[17] A. Sappa. *Automatic generator of 3D geometric model from Range Images*. Tese de Doutorado, Universidade de Catalunha, Catalunha, Barcelona, Espanha, 1999.

[18] C. Schültz, T. Jost, and H. Hügli. Semi_automatic 3D object digitizing system using range images. In *Proceedings Lecture Notes in Computer Science, Computer Vision - ECCV'98*, volume pp. 490-497. Springer-Verlag, 1998.

[19] M. Soucy and D.Laurendeau. Multi-resolution surface modelling from multiple range images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 348–353, 1992.

[20] 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.

[21] Yiyong Sun, C. Dumont, and M. Abidi. Mesh-based integration of range and color images. In *Proceedings SPIE Aerosense, the International Society for Optical Enginnering*, volume 4051, pages 110–117, 2000.

[22] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Computer Graphics Proceedings*, pp. 311-318. Orlando, Florida, Julho 1994. Siggraph 94.

[23] S. Weik. Registration of 3d partial surface models using luminance and depth information. In *Proceedings of the International Conference on Recent Advances in 3-D Imaging and Modeling*, Ottawa - Canada, 1997.