

4

3D-Pose: Uma Aplicação Exemplo

Uma aplicação de exemplo foi criada para demonstrar o método proposto e prover uma solução completa para reconstrução e registro de câmeras. A idéia é prover um sistema de catálogo e gerenciamento das fotos cujas câmeras já foram reconstruídas, bem como um novo paradigma de navegação tridimensional entre elas. Além disso, tal aplicação permite comparar detalhes entre foto e modelo através de diversas ferramentas como a capacidade de *zoom*, controle de transparência etc.

A aplicação funciona da mesma forma que o método recém-exposto, formando um fluxo de execução. Usuário pode carregar arquivos de imagem em diversos formatos, malhas geométricas no formato *.OBJ* (*Wavefront Object File*) e, ainda, utilizar decoradores *.IVE* (*OpenSceneGraph Scene File*) para visualizar um modelo *placeholder* com texturas e materiais, afim de melhorar a experiência visual e de manipulação.

Quando a malha de um modelo é carregada pela primeira vez, ela é processada pela aplicação e salva em um arquivo auxiliar para uso futuro. Desta forma cada modelo só precisa ser processado uma única vez. Tal processamento efetua as etapas expostas na Seção 3.2 e armazena o resultado. Se houver alteração na geometria do modelo, basta remover o arquivo gerado, que tem mesmo nome e se localiza no mesmo local do arquivo de modelo mas com extensão diferente (*.simp*). Esse arquivo é um arquivo texto simples que armazena a lista de arestas extraídas na etapa de processamento da geometria.

Após carregar um modelo, a aplicação permite carregar um número arbitrário de imagens, uma a uma, e efetuar o processo de reconstrução de câmera. Há uma funcionalidade para detecção de relações entre modelo e imagem a partir de uma posição de câmera inicial provida pelo usuário e é também possível inserir novas marcações na imagem para que sejam automaticamente associadas (se possível) as arestas estruturais do modelo. Algumas imagens reais da aplicação serão mostradas na Seção de resultados.

Algumas das funcionalidades principais da aplicação são apresentadas em mais detalhes a seguir.

4.1

Descarte Manual de Arestas Estruturais

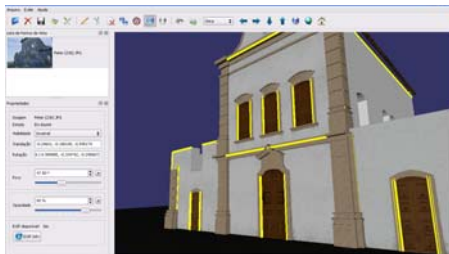
O processamento da geometria resulta em um conjunto de arestas que pode ser visualizado e tratado pela aplicação aqui implementada. Algumas arestas podem ser convenientemente eliminadas manualmente para facilitar o procedimento de localização de correspondências entre modelo e imagem. Como mencionado, as arestas estruturais são salvas em um arquivo para que o processamento da geometria ocorra apenas uma única vez. Sendo assim, as arestas removidas manualmente após o processamento da geometria também são eliminadas de tal arquivo, que é então reescrito.

4.2

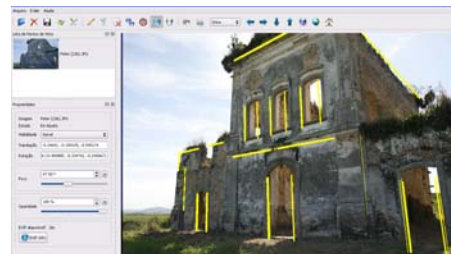
Posicionamento de Câmera

Um mecanismo foi desenvolvido para facilitar a manipulação do modelo permitindo que o usuário informe a rotação e a translação iniciais necessárias para o processo de ajuste. Tal mecanismo funciona como um manipulador *trackball*.

Nesta aplicação a imagem real tem um controle de transparência e, ainda, o modelo CAD tem linhas-guias estruturais que são desenhadas em destaque e que aparecem mesmo através da foto. A imagem 4.1 ilustra esse fato.



4.1(a): Linhas guia (em amarelo) desenhadas sobre a estrutura do modelo virtual.



4.1(b): Modelo sobreposto por uma imagem com 100% de opacidade. Mesmo com opacidade máxima as linhas-guias ainda são visíveis e o modelo pode ser manipulado.

Figura 4.1: Linhas-guias para manipulação do modelo virtual.

Assim, a aplicação permite que um usuário carregue e ajuste diversas imagens em relação a um modelo de edificação, o que resulta em uma espécie de álbum tridimensional de fotos.

4.3 Cálculo de Correspondências

É possível calcular correlações através de uma funcionalidade implementada utilizando a técnica proposta na Seção 3.6. Para isso, basta que o modelo esteja corretamente posicionado, e o detector de arestas configurado. A aplicação permite o ajuste fino do detector de arestas, que pode funcionar através do método PPHT ou HT (3.6.1).

A Figura 4.2 mostra o diálogo para configuração do detector de arestas e a Figura 4.3 mostra um exemplo real de cálculo automático de correspondências baseado em uma posição inicial do modelo.

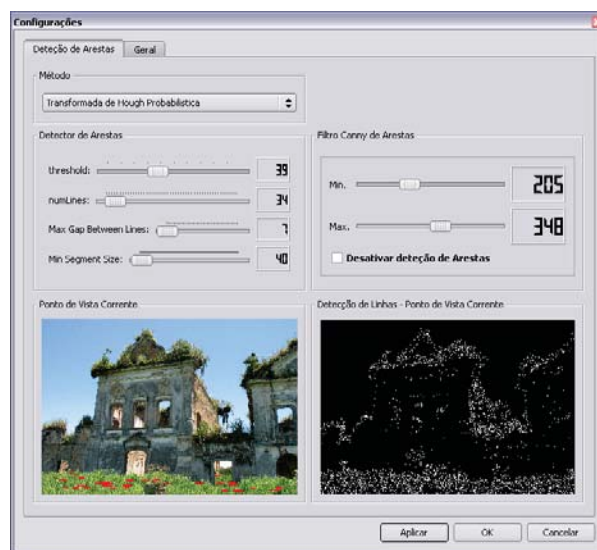


Figura 4.2: Diálogo de configuração de arestas. É possível selecionar o algoritmo de detecção (PPHT ou HT) e configurar valores de threshold para cada um deles. A imagem usada como referência para as configurações é a selecionada como corrente.

4.4 Informando Correlações Manualmente

Em casos complexos, onde as imagens possuem uma quantidade de ruído que impossibilita a detecção de arestas com qualidade — arestas que realmente correspondam ao objeto de interesse na foto —, é possível criar manualmente correlações entre modelo e imagem através de uma funcionalidade especial implementada nesta aplicação. O processo ocorre através da interação direta do usuário, que deve selecionar arestas na imagem e indicar a qual arestas no modelo ela corresponde. Isso possibilita que a aplicação seja utilizada em muitos casos em que seria inviável localizar automaticamente alguma correspondência entre imagem e modelo.



4.3(a): Correlações calculadas automaticamente entre modelo e imagem. As arestas em amarelo são as arestas do modelo, que está atrás da imagem. Os segmentos vermelhos são os segmentos do modelo que foram associados com a imagem e os verdes são suas respectivas associações com a imagem.



4.3(b): Detalhe das mesmas correlações mostradas na figura ao lado, mas com a imagem omitida e com o modelo visível.

Figura 4.3: Cálculo automático de correlações entre modelo e imagem a partir do posicionamento de câmera inicial.

4.5 Navegação Tridimensional entre Fotos

O processo de reconstrução de câmera resulta em uma câmera virtual totalmente calibrada e posicionada, que pode ser registrada em uma estrutura de dados para habilitar a navegação entre fotos. Como cada câmera i representa o ponto de vista em que a foto real foi obtida, a transformação projetiva $P_i = K_i R_i T_i$ sobrepõe perfeitamente imagem real e modelo virtual se mapearmos os pixels dessa imagem no plano de projeção da câmera i e desenharmos o modelo na cena na posição relativa correta.

Cada uma das câmeras registradas pode ser usada para prover um mecanismo simples de navegação entre as fotos no espaço 3D, o que permite catalogar as fotos e também permite saber quais partes do modelo não são cobertas por elas, se simplesmente projetarmos todas as fotos no modelo a partir de uma transformação projetiva inversa usando a matriz K^{-1} .

É comum encontrar coleções de fotos guardadas em arquivos de dados e cujos nomes sugerem pouca informação a respeito das mesmas. Isso torna difícil seu gerenciamento porque, mesmo que se saiba aproximadamente a posição de onde a foto foi tirada em relação ao modelo real — utilizando simplesmente a memória visual de como a foto se parece —, é preciso uma busca linear em todos os arquivos para encontrá-la. Com a possibilidade de registro 3D, muitas funcionalidades novas para encontrar uma foto a partir de uma região da cena virtual podem ser desenvolvidas para facilitar o gerenciamento de conjuntos

de imagens.

Esta dissertação desenvolveu uma técnica que permite navegar entre as fotos utilizando direções no espaço \mathbb{R}^3 . Dada uma posição corrente de câmera, que representa o ponto de vista de uma foto i registrada, é possível obter a foto cuja câmera está mais próxima em uma determinada direção. Esta técnica foi criada por esta dissertação e o autor desconhece trabalhos relacionados que a abordem da mesma forma.

Para saber qual a posição de câmera mais próxima em uma determinada direção, a partir do sistema de coordenadas de uma outra câmera corrente, basta percorrer as demais câmeras registradas atribuindo a seguinte classificação de distância

$$dist(\hat{\mathbf{d}}, \hat{\mathbf{v}}) = \frac{1}{\langle \hat{\mathbf{d}}, \hat{\mathbf{v}} \rangle} \|\mathbf{d}\| \quad (4-1)$$

onde o vetor \mathbf{d} é o vetor que vai da origem do sistema de coordenadas corrente até a posição de uma outra câmera e $\hat{\mathbf{v}}$ é a direção na qual se deseja obter a classificação de distância (por exemplo o eixo x do sistema atual). Como existe a possibilidade do produto interno $\langle \hat{\mathbf{d}}, \hat{\mathbf{v}} \rangle$ assumir valor zero ou muito próximo de zero, um teste é feito anteriormente para garantir que a divisão na função apresentada não gera inconsistências. Caso $p = \langle \hat{\mathbf{d}}, \hat{\mathbf{v}} \rangle < 10^{-8}$, o valor de $1/p$ é substituído por um outro valor suficientemente grande. A idéia geral dessa técnica é que quanto menor for o alinhamento entre os vetores \mathbf{d} e \mathbf{v} , maior é o valor de classificação de distância. Contudo, se duas câmeras estão na mesma direção a partir do sistema de coordenadas atual, a câmera mais próxima é escolhida, já que a classificação também é uma função da norma de \mathbf{d} . A aplicação estabelece ainda um limite para que, a partir de um determinado valor de classificação, seja possível especificar que não há câmeras próximas em uma determinada direção — por exemplo, especificar que não há câmeras próximas na direção $-x$.

Ambos os vetores \mathbf{d} e $\hat{\mathbf{v}}$ estão escritos no sistema de coordenadas da câmera corrente. Para isso é preciso antes obter o vetor \mathbf{d} no espaço do mundo e só então transformá-lo para o espaço da câmera corrente. Dado que a câmera atual é posicionada no mundo pela rotação \mathbf{R}_a e pelo vetor de translação \mathbf{t}_a (escrito no espaço do mundo) e, analogamente, uma câmera i têm vetor de translação \mathbf{t}_i , temos

$$\mathbf{d}^w = \mathbf{t}_i - \mathbf{t}_a \quad (4-2)$$

$$\mathbf{d} = \mathbf{R}_a(\mathbf{d}^w) \quad (4-3)$$

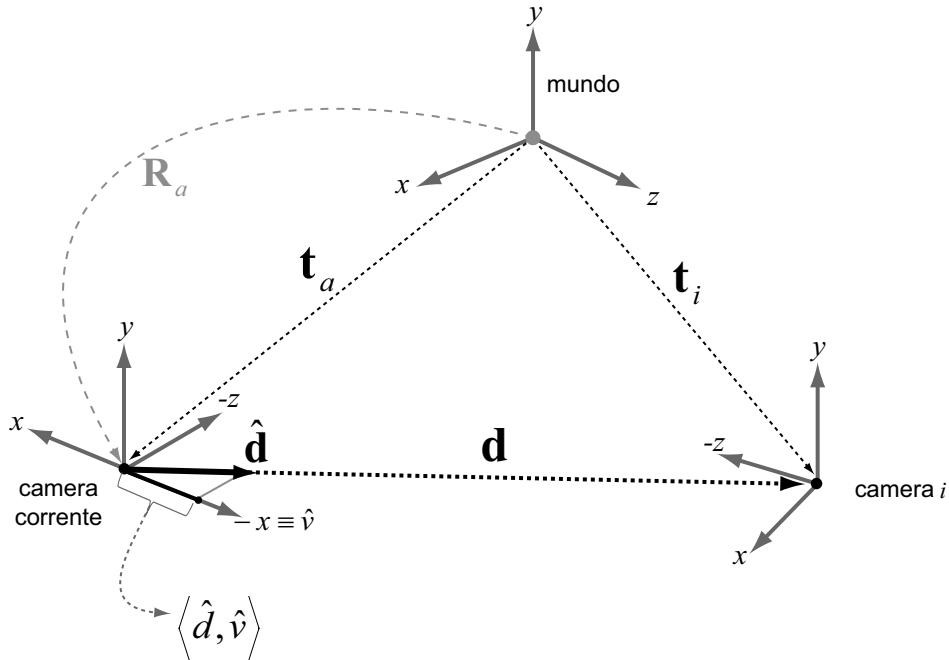


Figura 4.4: Representação dos vetores \mathbf{d} , $\hat{\mathbf{v}}$, \mathbf{t}_i e \mathbf{t}_a usados para calcular a classificação de distância a partir de uma direção.

onde o símbolo w representa um vetor no sistema de coordenadas do mundo (Figura 4.4).

4.6 Visão Geral das Funcionalidades

Esta seção apresenta resumidamente as principais funcionalidades da aplicação.

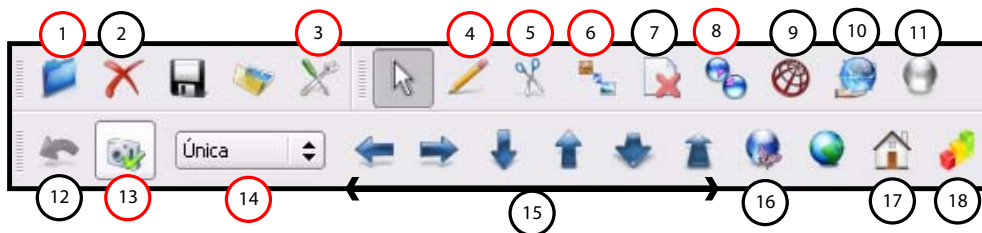


Figura 4.5: Barras de Ferramentas do sistema. As principais funcionalidades são indicadas pelos números. Os círculos vermelhos indicam as funcionalidades mais importantes.

A seguir são listadas as funcionalidades relativas a numeração na Figura 4.5.

1. Abrir e carregar imagem ou modelo;
2. Excluir imagem;

3. Configurar detector de arestas / propriedades gerais da cena;
4. Adicionar novas marcações na imagem;
5. Remover marcações / arestas detectadas da imagem;
6. Adicionar correlação modelo-imagem manualmente;
7. Remover todas as marcações;
8. Calcular automaticamente correlações a partir da posição de câmera atual;
9. Modo de exibição do modelo: *wireframe* ou normal;
10. Exibir / Esconder linhas estruturais do modelo;
11. Exibir / Esconder modelo;
12. Desfazer último ajuste;
13. Ajustar câmera utilizando correspondências calculadas / informadas;
14. Selecionar tipo de ajuste: Único, Contínuo ou por Convergência. O ajuste único ajusta o modelo apenas uma vez utilizando as correlações detectadas. O contínuo cria um ciclo de reajuste constante que utiliza uma nova busca por correlações a cada ajuste feito. Dessa forma cada posicionamento obtido pelo ajuste i serve de entrada para um novo ajuste $i+1$. Já o ajuste por Convergência funciona como o contínuo, mas realiza o processo até que a minimização deixe de convergir — quando dois ajustes consecutivos não alteram significativamente a posição da câmera. Os dois últimos tipos de ajuste permitem que o modelo seja ajustado a partir de uma posição mais distante, pois realizam uma série de ajustes antes de finalmente
15. Navegação 3D direcional entre as fotos registradas;
16. Navegação livre pela cena;
17. Posicionar modelo na posição inicial (*home position*);
18. Salvar informações sobre a imagem atual: salva dados da calibração em um arquivo texto, tais como coordenadas das arestas do modelo e da imagem que foram correlacionadas, parâmetros de câmera, dimensões do canvas, tamanho da imagem atual e outros. Esta funcionalidade é útil para comparar os resultados do ajuste e da calibração com outros métodos.