

Reconstrução 3D para digitalização de plantas com dados fotográficos e sua validação através de visão computacional

Alberto Arruda de Oliveira¹
Thiago Teixeira Santos²

A digitalização de plantas consiste no processo de capturar e armazenar a geometria de espécimes de forma a estudar sua estrutura e caracterizar seu fenótipo. A modelagem estrutural tem por objetivo determinar o padrão de crescimento de plantas e como sua forma é afetada por esse padrão (SANTOS; OLIVEIRA, 2012). A fenotipagem estuda justamente as características físicas resultantes do crescimento dos espécimes e como elas se relacionam com o ambiente em questão (SANTOS; OLIVEIRA, 2012). Atualmente, o processo de digitalização é caro, como no caso da varredura a laser, ou demanda grande esforço e períodos de tempo, como a varredura via dispositivos de rastreamento tridimensional. Com técnicas de visão computacional como *structure from motion* (SFM) (SZELISKI, 2010) e *multiple view stereo* (MVS) (SZELISKI, 2010), é possível criar modelos tridimensionais de plantas com alta precisão utilizando apenas fotografias, cujo processo de captura é de baixo custo e automatizável. Desta forma, é preciso analisar as diversas maneiras de coletar os dados necessários para essas reconstruções, assim como as muitas formas de implementação dos algoritmos utilizados, e como validar os dados obtidos pela reconstrução. Tais estudos foram o objetivo do presente trabalho.

O processo de reconstrução 3D utilizando um conjunto de imagens fotográficas possui o seguinte *pipeline* simplificado: primeiramente, é necessário encontrar pontos de interesse com algoritmos como SIFT (SZELISKI, 2010) em cada imagem, realizando o casamento desses pontos entre cada par de fotografias. Em seguida, utilizando-se SFM,

¹ Universidade Estadual de Campinas, ra093311@students.ic.unicamp.br

² Embrapa Informática Agropecuária, thiago.santos@embrapa.br

estima-se a partir desses casamentos entre as imagens a calibração das câmeras e uma nuvem esparsa de pontos tridimensionais (SZELISKI, 2010). Por fim, é produzida uma nuvem densa de pontos utilizando MVS (SZELISKI, 2010) e gerada uma malha tridimensional que aproxima a superfície do objeto com algoritmos como *Ball pivoting*. O objetivo da etapa de validação é coletar um *ground truth* relacionado à posição das câmeras no processo de captura das fotografias utilizadas na reconstrução tridimensional, e casar essas posições com às estimadas pelo SFM.

Para a detecção, descrição e casamento de pontos de interesse, foi utilizada a biblioteca OpenCV (WILLOW GARAGE, 2012), assim como o algoritmo SIFT, em sua implementação original para detecção e descrição e o casamento implementado pelo sistema *Bundler* (SNAVELY, 2012), aplicativo utilizado para SFM. Por fim, para MVS foram utilizadas as ferramentas CMVS e PMVS2 (FURUKAWA, 2012) e a implementação da aplicação Meshlab para o *Ball pivoting*. Para a validação dos dados, deseja-se saber a posição real das câmeras no momento da captura de cada imagem. Para isso, é utilizado o sensor eletromagnético de posicionamento FASTRAK, que se fixado na câmera torna possível obter a posição desta no momento da captura da fotografia, o que será o *ground-truth* usado para validação. Este processo é feito casando-se o *ground-truth* com as coordenadas da câmera dadas como saída do algoritmo de SFM. Para essa finalidade, foi desenvolvido o *Fastraker*, aplicativo que permite manipular tanto câmeras como o sensor FASTRAK, de forma a guiar o usuário no processo de captura de imagens e gravar também a posição do sensor no momento da captura. Tendo então a nuvem de coordenadas obtida pelo *Fastraker*, o casamento com a nuvem produzida pelo *Bundler* pode ser feito com algoritmos como *iterative closest points* (ICP) (SZELISKI, 2010).

A Figura 1 mostra alguns resultados da reconstrução 3D de um manjeriçõ e algumas das fotografias utilizadas. Foram usadas 143 imagens em um processo de reconstrução que durou aproximadamente duas horas, excluindo-se o tempo para a geração de malha por *Ball Pivoting*. A imagem mostra como o modelo tridimensional obtido possui alta precisão, mesmo tendo sido gerado apenas utilizando fotografias 3648 x 2736 pixels, obtidas por uma câmera comum. Algumas limitações dessa técnica estão na grande quantidade de fotografias necessárias e na forma de captura destas, e no tempo para a geração do modelo final. Entre as

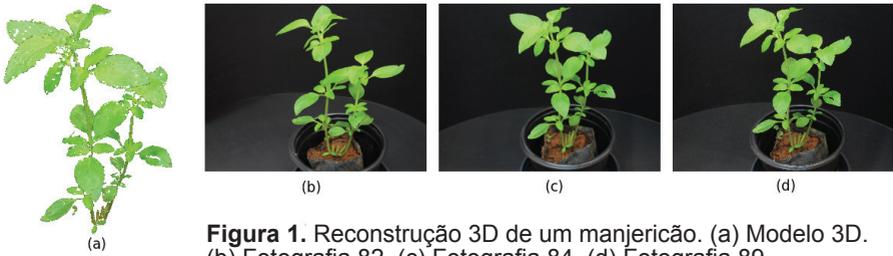


Figura 1. Reconstrução 3D de um manjeriço. (a) Modelo 3D. (b) Fotografia 82. (c) Fotografia 84. (d) Fotografia 89.

melhorias previstas para o trabalho, estão formas mais rápidas de realizar o processo, algoritmos melhores para geração de malhas e na automatização do processo de captura.

A interface do *Fastraker* (Figura 2) foi desenvolvida de modo a guiar o usuário no processo de captura das fotos, exibindo *streams* de vídeo oriundas da câmera e do sensor Kinect. Além disso, o programa exibe também em uma área de visualização a posição relativa dos pontos lidos pelo FASTRAK. Futuramente, espera-se utilizar a ferramenta para adquirir dados de um número maior de plantas, assim como dados de *laser-scanning* para serem usados como *ground-truth*.



Figura 2. Interface da aplicação *Fastraker*, com região de visualização de pontos e *streaming* de câmeras.

O processo de reconstrução 3D utilizando um conjunto de imagens fotográficas possui o seguinte pipeline simplificado: primeiramente, é necessário encontrar pontos de interesse com algoritmos como SIFT (SZELISKI, 2010) em cada imagem, realizando o casamento desses pontos entre cada par de fotografias. Em seguida, utilizando-se SFM, estima-se a partir desses casamentos entre as imagens a calibração das

câmeras e uma nuvem esparsa de pontos tridimensionais (SZELISKI, 2010). Por fim, é produzida uma nuvem densa de pontos utilizando MVS (SZELISKI, 2010) e gerada uma malha tridimensional que aproxima a superfície do objeto com algoritmos como Ball pivoting. O objetivo da etapa de validação é coletar um ground truth relacionado à posição das câmeras no processo de captura das fotografias utilizadas na reconstrução tridimensional, e casar essas posições com às estimadas pelo SFM.

Referências

FURUKAWA, Y.; PONCE, J. Accurate, Dense, and Robust Multiview Stereopsis. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 32, n. 8, p. 1362-1376, Aug. 2010.

SANTOS, T. T.; OLIVEIRA, A. A. D. Image-based 3d digitizing for plant architecture analysis and phenotyping. In: CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, 25.; WORKSHOP ON INDUSTRY APPLICATIONS, 2012, Ouro Preto, MG. **Proceedings...** 2012. Disponível em: <www.decom.ufop.br/.../wgari/104836_2.pdf>. Acesso em: 27 set. 2012.

SNAVELY, N. **Bundler**: Structure from Motion (SfM) for Unordered Image Collections. Disponível em: <<http://phototour.cs.washington.edu/bundler/>>. Acesso em: 27 de set. 2012

SZELISKI, R. **Computer Vision**: algorithms and Applications. London: Springer, 2010. 812 p.

WILLOW GARAGE. **OpenCV (Open Source Computer Vision)**. Disponível em: <<http://opencv.willowgarage.com>>. Acesso em: 27 de set. 2012.