

DIGITALIZAÇÃO TRIDIMENSIONAL DE CAFEIROS EM CAMPO POR VISÃO ESTÉREO MÚLTIPLA¹

Thiago Teixeira Santos²; Miroslava Rakocevic³

¹ Trabalho financiado pelo Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café – Consórcio Pesquisa Café

² Pesquisador, PhD, Embrapa Informática Agropecuária, Campinas-SP, thiago.santos@embrapa.br

³ Pesquisadora, PhD, Bolsista Consórcio Pesquisa Café, Embrapa Informática Agropecuária, Campinas-SP, miroslava.rakocevic@colaborador.embrapa.br

RESUMO: O registro da arquitetura de uma planta, isto é, a descrição da topologia e da geometria de sua estrutura, é empregado em estudos funcionais-estruturais variados como interceptação solar, interações bióticas e abióticas e dispersão de defensivos, entre outros. Porém, a realização desse registro é extremamente laboriosa, especialmente em plantas de maior porte e complexidade como o cafeeiro, o que restringe sua adoção e, conseqüentemente, o volume de dados disponíveis a tais estudos. Neste trabalho, mostramos que técnicas de visão computacional podem ser empregadas na construção de modelos tridimensionais de cafeeiros com o uso de uma única câmera comum e software adequado. Esses resultados indicam a possibilidade de automação do registro *in silico* da estrutura tridimensional de cafeeiros, permitindo um maior volume na obtenção de dados da arquitetura das plantas. Tal metodologia de aquisição de dados tridimensionais está ao alcance de qualquer grupo de pesquisa, devido ao baixo custo do equipamento necessário e implementações em software livre. Exemplos de reconstruções tridimensionais para oito cafeeiros adultos estão disponíveis em uma base de dados pública.

PALAVRAS-CHAVE: reconstrução 3-D, arquitetura de planta, FSPM, visão computacional

THREE-DIMENSIONAL DIGITISING OF COFFEE PLANTS ON FIELD BY MULTIPLE STEREO VISION

ABSTRACT: The representation of a plant architecture, that is, the description of its structure's topology and geometry, is employed in various functional-structural studies such as solar interception, biotic and abiotic interactions, and pesticide dispersion, among others. However, this registration is extremely laborious, especially for larger and more complex plants such as coffee, which restricts its adoption and, consequently, the volume of data available to such studies. In this work, we show that computer vision techniques can be employed in the construction of three-dimensional coffee tree models using a single common camera and suitable software. These results indicate the possibility of automating the *in silico* representation of three-dimensional coffee trees structure, allowing a larger volume on obtaining plant architecture data. Such a three-dimensional data acquisition methodology is within the reach of any research group, due to the low cost of required equipment and open source implementations. Examples of three-dimensional reconstructions for eight adult coffee trees are available in a public database.

KEY WORDS: 3-D reconstruction, plant architecture, FSPM, computer vision

INTRODUÇÃO

Modelos funcionais-estruturais de plantas (*functional-structural plant models* – FSPM) unem representações da estrutura tridimensional da planta a modelos de suas funções fisiológicas. Esses modelos se mostram úteis a diversos estudos em ciência de plantas: interceptação solar, troca de gases, difusão de pesticidas, poda e interações entre plantas e agentes biológicos (Vos et al., 2010). Modelos tridimensionais de plantas podem ser pensados como uma representação *in silico* da arquitetura da planta, permitindo aos pesquisadores avaliar detalhadamente seu estado passado e, ao longo do tempo, estudar a dinâmica de seu desenvolvimento e resposta ao ambiente. Um exemplo é o trabalho recente desenvolvido por Rakocevic et al. (2018), que empregou modelos funcionais-estruturais para avaliar alterações em plantas de café submetidas a taxas elevadas de CO₂.

A representação da estrutura tridimensional deve registrar tanto a topologia da planta (como as várias estruturas da planta estão conectadas entre si) quanto sua geometria (forma, dimensões e orientação no espaço) (Vos et al., 2010). A construção dessas representações necessita de um registro acurado de dados topológicos e geométricos, comumente codificados na forma de um grafo em árvore multiescalar (*multiscale tree graph* – MTG), como proposto por Godin & Caraglio (1998). Uma vez que a arquitetura é registrada na forma de MTGs, ela pode ser analisada com o auxílio de software específico como o AMAPmod (Godin et al., 1999) e visualizada na forma de *mock-ups* tridimensionais (3-D), como apresentado por Matsunaga et al. (2016) para o caso do cafeeiros. Para a obtenção desses dados, rastreadores tridimensionais como o FastTrack (Polhemus Inc., Cochester, VT, USA) são comumente empregados no registro da posição tridimensional de pontos de interesse na arquitetura vegetal (Costes et al., 2008; Sinoquet & Rivet, 1997), posteriormente transcritos manualmente pelo pesquisador na forma de MTGs. Infelizmente, esse processo de aquisição de dados é lento, trabalhoso, custoso (devido ao valor elevado dos rastreadores) e deve ser realizado em campo. Essas

dificuldades restringem a escala de obtenção de dados, principalmente para plantas de maior porte e complexidade como os cafeeiros.

Uma alternativa ao uso de rastreadores e outras formas trabalhosas de registro é o emprego de imageamento tridimensional (Vázquez-Arellano et al., 2016). Dentre as várias opções de imageamento 3-D, uma delas se destaca pelo seu baixo custo, facilidade operacional e pelo nível fino de detalhe dos modelos gerados: a combinação de *structure-from-motion* (SfM) com visão estéreo múltipla (*multiple-view stereo*). Oriundas do campo da visão computacional (Hartley & Zisserman, 2003; Triggs et al., 2000) e da robótica (Scaramuzza & Fraundorfer, 2011), essas técnicas permitem que modelos tridimensionais sejam construídos a partir de apenas um conjunto de imagens sobrepostas dos objetos de interesse, virtualmente transformando um câmera comum em um *scanner 3-D*. Inicialmente utilizadas na reconstrução 3-D de plantas de pequeno porte em vasos (Kumar et al., 2014; Lou et al., 2014; Santos & de Oliveira, 2012; Santos & Rodrigues, 2016) e no campo (Jay et al., 2015), essas técnicas se mostraram capazes também de capturar a estrutura tridimensional de plantas de maior porte no campo, como por exemplo videiras (Santos et al., 2017).

Neste trabalho, apresentamos uma metodologia para a produção de modelos tridimensionais de cafeeiros com o uso de técnicas de visão computacional. Mostramos que modelos tridimensionais detalhados podem ser obtidos com o uso de uma webcam de baixo custo e software livre que implementa rotinas de SfM e MVS. Esses modelos tridimensionais, armazenados na forma de nuvens de pontos (*point clouds*), podem ser utilizados como um registro *in silico* da estrutura tridimensional da parte aérea de cafeeiros e empregados em atividades de inspeção, medição e análise das plantas, bem como na geração de MTGs para pesquisas em modelagem funcional-estrutural de cafeeiros.

MATERIAL E MÉTODOS

Os dados experimentais foram obtidos em cafeeiros Arábica (*Coffea arabica* L.), cultivar 'Catuaí Vermelho IAC 144', localizados na Embrapa Meio Ambiente (22°43'S, 47°01'W, 570 m de altitude). Os resultados apresentados neste trabalho são oriundos de capturas de imagens realizadas em março e maio de 2018, de cafeeiros de sete anos de idade. As imagens foram obtidas em campo com uma câmera C920 (Logitech Inc., Lausanne, Suíça), comumente encontrada no mercado brasileiro, e conectada a um computador notebook. Na captura das imagens, foi utilizado o programa *3dmcap*¹, software livre (GPLv3) desenvolvido pela Embrapa Informática Agropecuária e introduzido por Santos et al. (2017). O software realiza uma seleção automática de imagens a partir da sequência de vídeo produzida pela câmera, de modo que seja possível encontrar um número suficiente de correspondências entre as imagens necessárias ao *pipeline* de *structure-from-motion*. Internamente, *3dmcap* emprega o sistema de SfM em tempo real ORB-SLAM, desenvolvido por Mur-Artal & Tardos (2017), que computa a posição da câmera no espaço no momento de cada captura de imagem. A interface do *3dmcap* permite ao usuário acompanhar o processo de aquisição e relocalizar automaticamente a câmera em caso de falhas. O uso do programa é descrito com detalhes em Santos (2018).

Durante a aquisição, uma placa rígida contendo um padrão de dimensões conhecidas é posicionada próximo à base da planta, como exibido na Figura 1 (a). Além de marcar o plano do solo, útil para realização de medições de altura das plantas, o padrão fornece uma referência de escala, permitindo que o modelo tridimensional seja transformado para uma escala de milímetros, permitindo a realização de medições a partir do modelo. Sem uma referência de escala, o *pipeline* de SfM e MVS ainda seria capaz de obter estruturas tridimensionais de modo que forma, ângulos e proporções estejam corretos, mas as dimensões reais dos objetos permaneceriam desconhecidas.

A etapa de visão estéreo múltipla é computacionalmente custosa e é realizada posteriormente, a partir das imagens e dos dados de posição da câmera obtidos pelo *3dmcap* (odometria visual, descrito em Scaramuzza & Fraundorfer, 2011). Dois sistemas diferentes para MVS foram testados: o PMVS, proposto por Furukawa & Ponce (2010) e integrado ao *3dmcap*, e o COLMAP², introduzido por Schönberger et al. (2016). Ambos os sistemas são capazes de utilizar as imagens e os dados de posição da câmera para gerar *point clouds*, conjuntos de pontos tridimensionais com informação de cor, representados como vetores (X, Y, Z, R, G, B) em que X, Y e Z informam a posição tridimensional do ponto e R, G e B a informação de cor. Tais pontos correspondem a uma amostragem da superfície visível dos objetos: solo, folhas, frutos e demais estruturas visualizadas pela câmera. As nuvens de pontos são armazenadas em arquivos no formato PLY³.

RESULTADOS E DISCUSSÃO

A Tabela 1 exhibe os resultados obtidos para oito plantas observadas. O software *3dmcap* selecionou automaticamente as imagens enquanto o operador movimentava a câmera ao redor da planta, como ilustrado na Figura 1 (b). As imagens obtidas apresentaram resolução de 1920 x 1080 pixels. Cabe ao operador movimentar a câmera de modo a prover observações das estruturas de interesse em diferentes ângulos, o que permite aos sistemas de visão estéreo obter reconstruções acuradas dessas estruturas. No geral, o operador levou de 1 a 2 minutos para obter cerca de 100 imagens de cada planta (Tabela 1). A Tabela 1 também evidencia a capacidade do sistema de visão estéreo implementado no

¹ <https://github.com/thasant/3dmcap>

² <https://colmap.github.io>

³ Ver <http://paulbourke.net/dataformats/ply> para uma descrição detalhada do formato.

COLMAP de obter nuvens de pontos mais densas. Comparadas aos resultados obtidos pelo PMVS, as nuvens produzidas pelo COLMAP apresentaram cerca de 10 vezes mais pontos.



Figura 1. (a) Padrão de escala utilizado em cafeeiros. (b) Exemplo de aquisição - a posição e a orientação da câmera em cada uma das 167 imagens utilizadas neste modelo são indicadas pelas pirâmides vermelhas. Este modelo 3-D foi composto por 5.081.529 pontos.

A Figura 2 exibe um exemplo de nuvem de pontos tridimensional para uma das plantas do estudo – visualizações interativas deste mesmo espécime estão disponíveis *on line*⁴. O modelo está na escala de milímetros, exibida nos eixos da figura. Sistemas de visualização 3-D⁵ permitem que medições, análises e manipulações sejam realizadas em nuvens de pontos como essas. Os dados das oito plantas, incluindo imagens e modelos tridimensionais, se encontram disponíveis *on line* (Santos, 2019).

As nuvens de pontos foram capazes de capturar acuradamente as superfícies visíveis nas imagens, incluindo folhas, ramos e frutos. Porém, estruturas mais internas no dossel foram oclusas e não foram reconstruídas. Este é um problema frequente no registro da arquitetura de qualquer planta de maior porte e complexidade. Uma alternativa a ser considerada é o emprego de um segundo operador, que mecanicamente suspenderia os ramos superiores enquanto o operador da câmera realiza o imageamento. Isto levaria a criação de várias reconstruções de uma mesma planta, que poderiam ser facilmente integradas com o auxílio do padrão no solo (Figura 1 (a)).

Trabalhos recentes têm demonstrado que *redes neurais convolutivas* (LeCun et al., 2015) podem ser treinadas para o reconhecimento de estruturas de plantas em imagens (Bargoti & Underwood, 2017; Pound et al., 2017). Uma possível linha de pesquisa para trabalhos futuros seria a utilização de redes neurais para a identificação de pontos de interesse da arquitetura da planta como os pontos de incisão de folhas, pontos de ramificação e frutos. Uma vez identificadas nas imagens (2-D), essas localizações poderiam ser projetadas no modelo tridimensional, tendo sua localização 3-D determinada. Tal inovação possibilitaria a criação de um sistema para produção semiautomática de MTGs a partir de imagens, o que levaria a grandes ganhos na captura de informações para FSPM e fenotipagem de arquitetura de plantas complexa como os cafeeiros Arábica.

⁴ Visualizações interativas de uma mesma planta em duas datas diferentes estão disponíveis na Web. Em <https://skfb.ly/6MBLB> encontra-se um modelo obtido em março de 2018 e em <https://skfb.ly/6MBKV> a mesma planta em maio do mesmo ano.

⁵ Um exemplo é o software livre *Meshlab*, disponível em <http://www.meshlab.net>.

Tabela 1: Comparação entre os métodos de visão estéreo múltipla (PMVS e COLMAP), informando o número de imagens disponível para cada planta de cafeeiros. As duas colunas da direita informam o total de pontos 3-D obtidos por cada um dos métodos.

Planta	Núm. imagens	PMVS	COLMAP
<i>20/03/2018</i>			
Planta 1	167	429.867	5.081.529
Planta 2	142	472.469	4.024.688
<i>14/05/2018</i>			
Planta 1	73	313.433	2.327.210
Planta 3	115	338.961	3.871.258
Planta 4	85	275.487	2.688.350
Planta 5	72	216.967	2.105.871
Planta 6	114	350.196	3.361.511
Planta 7	101	272.363	3.383.105
Planta 8	111	344.790	3.048.892

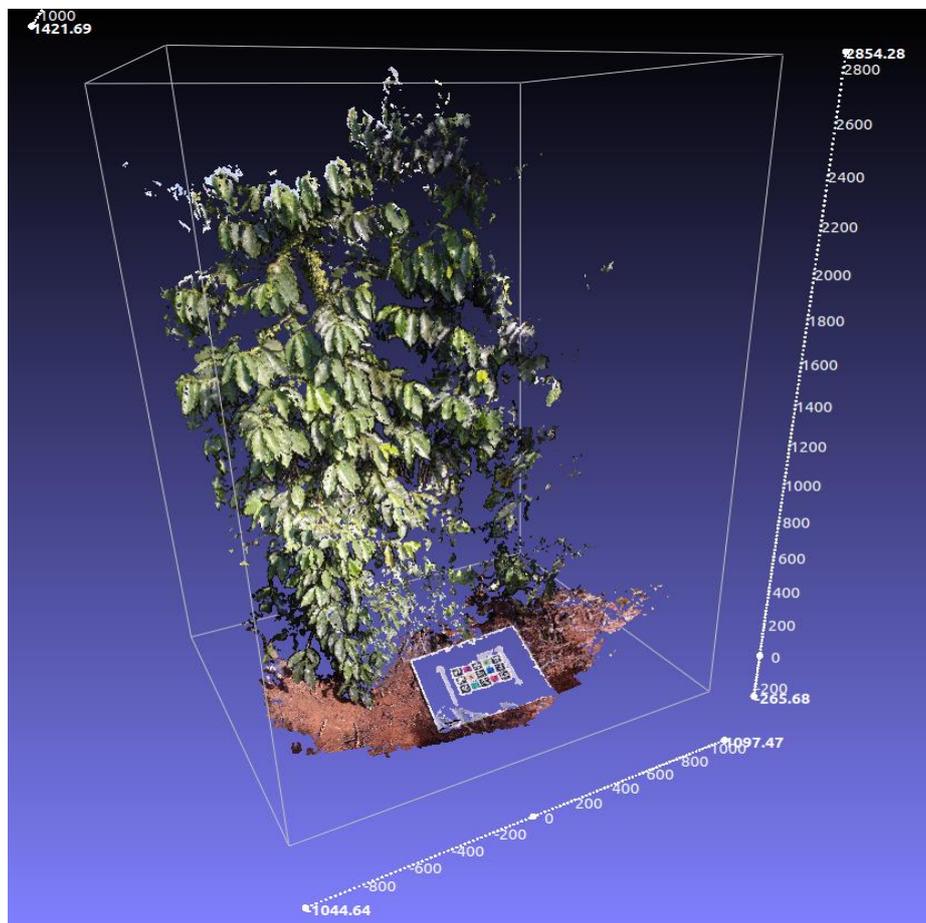


Figura 2. Exemplo de modelo 3-D de cafeeiros com escala em milímetros. Uma versão interativa está disponível *on line*.

CONCLUSÕES

1. Reconstrução tridimensional utilizando câmeras RGB e software livre para visão estéreo múltipla é uma alternativa de baixo custo para registro da estrutura tridimensional de cafeeiros em campo;
2. Modelos estruturais da planta podem ser registrados a partir desses modelos 3-D, substituindo alternativas mais trabalhosas como rastreadores de movimento comumente utilizados pela comunidade de FSPM;
3. Projetos de pesquisa em FSPM e fenotipagem de plantas podem utilizar a reconstrução tridimensional como um método para o registro das características morfológicas das plantas em estudo *in silico*, armazenando características do espécime para consulta e análise.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARGOTI, S. & UNDERWOOD, J. (2017). Deep fruit detection in orchards, in: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE. pp. 3626-3633. doi: 10.1109/ICRA.2017.7989417.
- COSTES, E.; SMITH, C.; RENTON, M.; GUÉDON, Y.; PRUSINKIEWICZ, P.; GODIN, C. (2008). MAppleT: simulation of apple tree development using mixed stochastic and biomechanical models. *Functional Plant Biology* 35:936-950. doi: 10.1071/FP08081.
- FURUKAWA, Y. & PONCE, J. (2010). Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 1362-1376. doi: 10.1109/TPAMI.2009.161.
- GODIN, C. & CARAGLIO, Y. (1998). A multiscale model of plant topological structures. *Journal of Theoretical Biology* 191:1-46. doi: 10.1006/jtbi.1997.0561.
- GODIN, C.; COSTES, E.; SINOQUET, H. (1999). A method for describing plant architecture which integrates topology and geometry. *Annals of Botany* 84:343-357. doi: 10.1006/anbo.1999.0923.
- HARTLEY, R. & ZISSERMAN, A. (2003). Multiple view geometry in computer vision. Cambridge University Press.
- JAY, S.; RABATEL, G.; HADOUX, X.; MOURA, D.; GORRETTA, N. (2015). In-field crop row phenotyping from 3D modeling performed using Structure from Motion. *Computers and Electronics in Agriculture* 110:70-77. doi: 10.1016/j.compag.2014.09.021.
- KUMAR, P.; CONNOR, J.; MIKIAVCIC, S. (2014). High-throughput 3D reconstruction of plant shoots for phenotyping. In: 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), IEEE. pp. 211-216. doi: 10.1109/ICARCV.2014.7064306.
- LECUN, Y.; BENGIO, Y.; HINTON, G. (2015). Deep learning. *Nature* 521:436-444. doi: 10.1038/nature14539.
- LOU, L.; LIU, Y.; HAN, J.; DOONAN, J.H. (2014). Accurate Multi-View Stereo 3D Reconstruction for Cost-Effective Plant Phenotyping. In: 2014 - International Conference Image Analysis and Recognition Springer, Cham, pp. 349-356. doi: 10.1007/978-3-319-11755-3_39.
- MATSUNAGA, F.T.; TOSTI, J.; ANDROCIOLI-FILHO, A.; BRANCHER, J.; COSTES, E.; RAKOCEVIC, M. (2016). Strategies to reconstruct 3D *Coffea arabica* L. plant structure. *SpringerPlus* 5, 2075. doi: 10.1186/s40064-016-3762-4.
- MUR-ARTAL, R. & TARDOS, J.D. (2017). ORB-SLAM2: An Open-Source SLAM System for Monocular Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics* 33:1255-1262. doi: 10.1109/TRO.2017.2705103.
- POUND, M.P.; ATKINSON, J.A.; TOWSEND, A.J.; WILSON, M.H.; GRIFFITHS, M.; JACKSON, A.S.; BULAT, A.; TZIMIROPOULOS, G.; WELLS, D.M.; MURCHIE E.H.; PRIDMORE, T.P.; FRENCH, A.P. (2017). Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *GigaScience* 6. doi: 10.1093/gigascience/gix083.
- RAKOCEVIC, M.; RIBEIRO, R.V.; RIBEIRO-MARCHIORI, P.E.; FILIZOLA, H.F.; BATISTA, E.R. (2018). Structural and functional changes in coffee trees after 4 years under free air CO₂ enrichment. *Annals of Botany* 121:1065-1078. doi: 10.1093/aob/mcy011.
- SANTOS, T.T. (2019). Coffee3D - 3-D models for coffee trees using multiple-view stereo [Data set]. doi: 10.5281/zenodo.3364776.
- SANTOS, T.T. (2018). 3dmcap documentation: release 1.0. Documentos, 158 ed. Embrapa Informática Agropecuária. Campinas. URL: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/188876/1/3DMCAP-Doc158.pdf>.
- SANTOS, T.T.; BASSOI, L.H.; OLDONI, H.; MARTINS, R.L. (2017). Automatic grape bunch detection in vineyards based on affordable 3D phenotyping using a consumer webcam, in: Anais do XI Congresso Brasileiro de Agroinformática (SBIAgro 2017), Unicamp, Campinas. pp. 89-98.
- SANTOS, T.T. & de OLIVEIRA, A.A. (2012). Image-based 3D digitizing for plant architecture analysis and phenotyping, in: SAÚDE, A.V., GUIMARÃES, S.J.F. (Eds.), *Workshop on Industry Applications (WGARI) in SIBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images)*, Ouro Preto. doi: 10.13140/2.1.4576.1608.
- SANTOS, T.T. & RODRIGUES, G.C. (2016). Flexible three-dimensional modeling of plants using low-resolution cameras and visual odometry. *Machine Vision and Applications* 27: 695-707. doi: 10.1007/s00138-015-0729-3.

- SCARAMUZZA, D. & FRAUNDORFER, F. (2011). Visual Odometry [Tutorial]. *IEEE Robotics & Automation Magazine* 18:80-92. doi: 10.1109/MRA.2011.943233.
- SCHÖNBERGER, J.L.; ZHENG, E.; POLLEFEYS, M.; FRAHM, J.M. (2016). Pixelwise view selection for unstructured multi-view stereo. In: 2016 - European Conference on Computer Vision (ECCV), Springer, Cham., p. 501-518. doi: 10.1007/978-3-319-46487-9_31.
- SINOQUET, H. & RIVET, P. (1997). Measurement and visualization of the architecture of an adult tree based on a three-dimensional digitising device. *Trees - Structure and Function* 11:265-270. doi: 10.1007/s004680050084.
- TRIGSS, B.; MCLAUCHLAN, P.F.; HARTLEY, R.I.; FITZGIBBON, A.W. (2000). Bundle Adjustment – A Modern Synthesis, Springer, Berlin, Heidelberg, pp. 298-372. doi: 10.1007/3-540-44480-7_21.
- VÁZQUEZ-ARELLANO, M.; GRIEPENTROG, H.; REISER, D.; PARAFOROS, D. (2016). 3-D Imaging Systems for Agricultural Applications - A Review. *Sensors* 16, 618. doi: 10.3390/s16050618.
- VOS, J.; EVERS, J.B.; BUCK-SORLIN, G.H.; ANDRIEU, B.; CHELLE, M.; DE VISSER, P.H.B. (2010). Functional-structural plant modelling: a new versatile tool in crop science. *Journal of Experimental Botany* 61:2101-2115. doi: 10.1093/jxb/erp345.