



Title	PlantPose: Plant Skeleton Estimation in 2D and 3D
Author(s)	劉, 新鵬
Citation	大阪大学, 2025, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105269
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

論文内容の要旨

氏 名 （ 劉 新 鵬 ）	
論文題名	PlantPose: Plant Skeleton Estimation in 2D and 3D (PlantPose : 2次元および3次元における植物骨格推定)
論文内容の要旨 Accurately extracting plant skeletal architecture is a fundamental component for high-throughput phenotyping, growth modelling, and agricultural robotics. A plant skeleton is represented as a graph whose nodes lie on branches and whose edges encode connectivity between node pairs. Plant skeleton estimation is typically approached in two main domains: (i) single-view 2D and (ii) multi-view 3D skeleton estimation. Unlike human poses with a fixed and well-defined joint structure, plant skeletons can contain an unknown number of joints and must follow a tree topology without loops or disconnections. Additionally, in 2D estimation, models trained on one type of plant often fail to generalize to other species or imaging conditions. In 3D, existing methods rely on point clouds from light detection and ranging (LiDAR) or multi-view stereo (MVS), which frequently suffer from uneven density and occlusion, resulting in broken or fragmented connectivity, violating the desired tree structure. To address these challenges, this dissertation introduces PlantPose, a unified framework that (1) constructs a diverse and large-scale dataset to enable generalizable 2D skeleton estimation across different plant types, (2) produces tree-constrained skeletons in an end-to-end manner, and (3) proposes a mask-guided 3D optimization method that avoids point cloud fragmentation by directly aligning 3D graphs with multi-view masks. First, we assemble the PlantPose dataset, the largest and most diverse benchmark for plant skeleton estimation, comprising ten subsets including lab seedlings, greenhouse crops, orchard trees, web-crawled photographs, hand-drawn sketches, and photorealistic synthetics. In the branch-level raw annotations, branch segments are labeled independently as sequences of nodes and edges, without guaranteed cross-branch connectivity, leaving many disconnected components. We therefore propose a graph-recovery pipeline that merges disconnected branch segments into a single tree using a minimum spanning tree (MST). The resulting supervision for each image consists of pixel-level joint locations and an adjacency matrix that encodes a tree topology. Second, we present PlantPose2D, an end-to-end model that predicts graph nodes and edges from a single image while enforcing tree constraints during learning. The central component is a Selective Feature Suppression (SFS) layer. We project the model's unconstrained edge scores at each training step onto an MST defined over the current node set using a non-differentiable solver. SFS layer then suppresses the internal features that create cycles or isolated parts, aligning the network's gradients with the tree structure projection. This reparameterization enforces the constraint directly in the learning loop without requiring a differentiable MST. Leveraging the generalized dataset, PlantPose2D achieves the highest proportion of valid tree-structured skeletons and lower topology-aware error and generalizes well to unseen species, lighting conditions, and drawing styles. Third, we address the limitations of 3D skeleton estimation from multi-view images. Point-cloud pipelines built on LiDAR or MVS often exhibit uneven density on thin, low-texture branches, which breaks connectivity. We introduce PlantPose3D, a mask-guided graph optimization method that reduces reliance on point-cloud quality by aligning a 3D graph's observed masks with refined multi-view masks. We first represent the plant as a graph whose nodes store 3D positions and branch radii, with connectivity encoded by an adjacency matrix. Then, using 3D Gaussian Splatting (3DGS) for differentiable silhouette rendering, we optimize node positions, radii, and adjacency matrix to match the observed masks across views. In addition, we reuse the SFS layer from PlantPose2D to regularize the graph to a tree structure graph during optimization. This method overcomes the limitations of uneven point-cloud density while guaranteeing a tree-structured graph.	

論文審査の結果の要旨及び担当者

氏 名 (劉 新 鵬)			
論文審査担当者	(職) 氏 名		
	主 査	准教授	大倉 史生
	副 査	教授	前川 卓也
	副 査	准教授	佐々木 勇和
	副 査	教授	原 隆浩
	副 査	教授	五十部 孝典
	副 査	教授	鬼塚 真
	副 査	教授	中野 珠実
	副 査	教授	村田 忠彦
	副 査	教授	伊達 進
論文審査の結果の要旨 本論文は、植物の骨格構造（スケルトン）を高精度に推定するための新しい枠組み「PlantPose」を提案し、その理論的基盤、データセット構築、ならびに2次元および3次元におけるアルゴリズム開発を体系的に示したものである。植物の骨格構造は、高速フェノタイピング、成長モデリング、農業ロボティクスなどの基盤技術として極めて重要であるが、人間の骨格推定と異なり、節点数が未知であり、閉ループのない樹状構造（tree topology）を保ちながら推定する必要がある。この特性により、既存の深層学習モデルや3D復元手法の多くは構造的破綻や接続欠損を生じるという課題があった。 本論文ではまず、多様な植物種・撮影環境を含む世界最大規模のデータセット PlantPose dataset を構築し、分岐単位で付与された生データを最小全域木（MST）により統合することで、教師データとして整合性のある樹状グラフを再構成した。 次に、単一画像から節点と枝の接続関係を同時に推定するモデル PlantPose2D を提案し、Selective Feature Suppression（SFS）層を導入することで、学習過程において非微分的な木構造制約を直接反映させる新しい学習手法を確立した。その結果、未知の植物種や照明条件、描画スタイルに対しても高い汎化性能を実現し、有効な樹状構造を最も高い割合で生成することを実証した。 さらに、3次元骨格推定における既存点群ベース手法（LiDARや多視点ステレオ）による密度不均一・遮蔽問題に対して、著者は PlantPose3D を提案した。この手法では、3D Gaussian Splatting を用いた微分可能なシルエットレンダリングを通じて、多視点マスク情報を直接整合させながら節点位置・枝半径・接続行列を最適化する。これにより、点群品質への依存を大幅に低減しつつ、木構造制約を保持した一貫性のある3次元骨格モデルの生成を実現した。 以上の成果は、深層学習・画像解析・グラフ最適化・3D再構成など複数分野の技術を高いレベルで統合し、植物構造のモデリングという難題に対して新しい原理的アプローチを提示するものである。また、本研究で開発されたアルゴリズム群およびデータセットは、植物フェノタイピングや農業ロボット、デジタルツイン技術などへの応用可能性が高く、情報科学および植物科学の双方において学術的・実用的に大きな波及効果を持つ。 本論文は学術的独創性、技術的完成度、応用的有用性のいずれの観点からも極めて優れた内容を有しており、よって、博士（情報科学）の学位論文として価値のあるものと認める。			