

フォトグラメトリを用いた 収穫前ワインブドウ房の体積推定手法の評価

小野坂捺[†] 内海智仁^{††} 峰野博史^{†‡}

静岡大学情報学部[†] ヤマハ発動機株式会社^{††} 静岡大学大学院情報領域/グリーン科学技術研究所[‡]

1. はじめに

近年、農業分野において 3D データの活用が注目を集めている。画像間の対応関係をソフトウェア解析することで 3D 点群を構築するフォトグラメトリ (SfM/MVS) [1]は、非接触かつ非破壊で対象物の測定が可能な点や、レーザー測量と比較してモデルの生成が早く、導入コストも安価といった長所を持つ。一方で、構築される 3D 点群の境界は曖昧であるため、点群から直接体積を推定することは困難である。本研究では、農業用無人走行車両 (Unmanned ground vehicle, UGV) を用いて圃場で撮影される栽培データを対象に、フォトグラメトリによって構築された 3D 点群に対し、 α -shape で表面形状を構築することでワインブドウの果実部分の体積を高精度に推定することを目的とする。

2. 関連研究

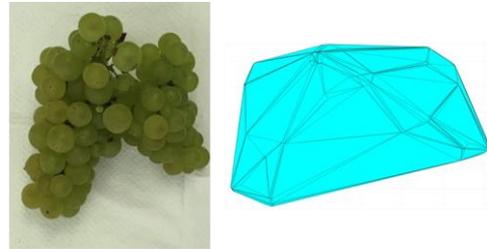
Mónica H[2]らは、フォトグラメトリを用いて構築されたブドウの房の 3D 点群へ凸包を適用することで、非可視部分を含めた表面形状の再現を試みている。3D 点群から直接凸包を適用した手法は、体積、重量の真値との決定係数が約 0.77、粒数については約 0.79 の精度を示した。

凸包は窪みのある形状に対しては、図 1 (b)のように窪みを埋めるような形で表面形状を構築するため、実際の体積の値よりも大きい値が算出されるという課題がある。そこで、有限の点群の形状に付随する単純で区分線形な曲線の一群を求める α -shape[3]を用いれば、より正確な体積の算出が可能になると考える。さらに、房部分の 3D 点群を構築する際に、2D 画像用のセグメンテーションモデルによって生成されたマスク画像をフォトグラメトリの入力として加えることで、現時点では発展途上で性能の低い 3D 点群用のセグメンテーションモデルを使わなくとも、房部分の点群を構築できると考える。

3. 提案手法

3.1 概要

提案手法は大きく二つのステップからなる。第一ステップでは、圃場動画から切り出されたワインブドウのマスク画像を用いて、フォトグラメトリでブドウの房部分のみの 3D 点群データを生成する。ここで、マスク画像の生成には、



(a) 元の形状 (b) 凸包

図 1 窪みの大きい形状に対する凸包

別途 2D 画像から房部分を抽出できるよう学習させたセマンティックセグメンテーションモデルを使用する。

第二ステップでは、ブドウの房のみの 3D 点群データに対し、 α -shape によって表面形状を再現し体積の算出をする。

3.2 ブドウの房部分のみの点群構築

第一ステップでは、実圃場を撮影した画像からブドウの房部分のみの 3D 点群を構築する。房部分のみの点群を構築するために、別途 2D 画像用のセマンティックセグメンテーションモデルを使用して、房部分を抽出するためのマスク画像を生成する。本研究では既存のセマンティックセグメンテーション用の学習済みモデルをファインチューニングしてマスク画像生成モデルを構築することで、少ない学習データでの高精度なマスク画像生成を可能にする。

実圃場においてブドウの房は密集していたり、形状も多様であったりするため、各房を個別に検出することは難しい。また、実際にブドウの収量を計測する際はブドウ一房一房といった単位ではなく、畑全体という単位で計量を行うことが多い。そのため、実用性も考慮し、各房を個別に識別するインスタンスセグメンテーションではなく、個々の境界までは判断せずにブドウの房をまとめて検出するセマンティックセグメンテーションを採用することとした。

3.3 高精度な体積算出

第二ステップでは、フォトグラメトリによって生成された 3D 点群に対して、 α -shape で表面形状を構築して体積を算出する。 α -shape は凸包の概念を一般化したものであり、 α 半径と呼ばれるパラメータを変更することで、構築する表面形状の精細さを調整できる。このことから、

Evaluation of volume estimation method of pre-harvest wine grape bunches using photogrammetry
Natsu Onozaka[†], Tomoyoshi Utsumi^{††}, Hiroshi Mineno^{†‡}
[†]Faculty of Informatics, Shizuoka University

^{††}YAMAHA Motor Co., Ltd.
[‡]College of Informatics, Academic Institute, Shizuoka University/ Research Institute of Green Science and Technology, Shizuoka University

窪みの多い複雑な形状の房であっても α 半径を小さくすることで高精度に体積を推定することが可能と考える。

ここで、 α -shape で点群に対して表面形状を構築する場合、表面に穴が開くことがあり、その際 α -shape の中身は空洞となるため、実際の体積よりも極端に小さい体積が算出される。より大きな α 半径を設定することで、表面の穴を塞ぐことができるが、 α 半径の値を大きくすればするほど構築される表面形状は徐々に凸包に近づいていき精細さが失われる。そのため、この課題を解決するために、 α 半径の値を変化させたときの体積の変化量に閾値を設定し、閾値を超える体積の変化が現れたときの体積を推測値として採用することとした。

4. 実験

4.1 セマンティックセグメンテーションを用いた 3D 点群構築

ブドウの房部分をアノテーションした画像で学習したセマンティックセグメンテーションモデルの構築によって未学習の実圃場画像からブドウ房のマスク画像を生成し、ブドウ房のみの 3D 点群を構築可能か検証する。まず、幅 5120、高さ 2880 ピクセルの元画像およびマスク画像を 8 等分し、幅 1280、高さ 1440 ピクセルの画像に変換して 256×256 にリサイズした。元画像とマスク画像を 8 分割した画像 488 枚を、学習用 368 枚と検証用 40 枚、テスト用 80 枚に分け、これらをセマンティックセグメンテーションモデル構築のためのデータセットとした。

画像データベースの ImageNet で学習済みのバックボーンを持つセグメンテーションモデル[4]でファインチューニングを行い、エンコーダとアーキテクチャの組み合わせを変更しながら、どの組み合わせが最も高精度にブドウ房のセマンティックセグメンテーションを行えるか検証した。検証した組み合わせの一部を表 1 に示す。様々な組み合わせで検証した結果、エンコーダに EfficientNetB4、アーキテクチャに DeepLabV3+ を採用したセマンティックセグメンテーションモデルが最も高い性能を示した。

このセマンティックセグメンテーションモデルを用いて、実圃場でブドウの木を片側から撮影した画像からマスク画像の生成を行った。図 2 に、セマンティックセグメンテーションモデルによって生成されたマスク画像から構築した 3D 点群の例を示す。概ねブドウ房の 3D 点群が適切に構築されていることが確認できたが、元画像で枝や葉などによるオクルージョンが多い部分では、マスク画像の精度が落ちることもあり、図で赤い丸で示したように、房以外の不要な点群が生成される場合があった。

4.2 α -shape を用いた体積算出

構築した点群に対し、 α -shape を用いて体積を算出する。精度の評価指標として体積の真値を得ることができなかったため、体積と強い正の相関を持つ重量を用いる。点

表 1 セグメンテーションモデルの比較

エンコーダ	アーキテクチャ	mIoU
ResNet34	UNet	0.713
VGG19	UNet	0.748
EfficientNetB4	DeepLabV3+	0.882

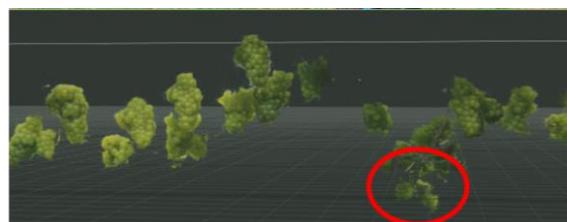


図 2 マスク画像を用いた点群構築結果

群から任意の房部分を切り出して、 α -shape で体積を算出した。ここで、収穫済みワインプドウの画像から構築した 27 房分の点群に対して、それぞれ α -shape を用いて体積を算出した際の重量と体積の散布図の近似直線の式を用いて重量の推測値を算出したところ、真値 247.3g に対し、推測値は 236.6g で約 10g 程度の誤差であった。凸包でも同様の方法で重量の値を算出したところ、値は 285.9g となり、この結果から α -shape を用いることで既存手法の凸包より高精度な重量、体積計測が可能であることがわかった。

5. おわりに

本研究では、ワインプドウ圃場の画像からセマンティックセグメンテーションを用いてマスク画像を生成し、フォトグラメトリで房部分のみの点群が構築可能であることを示した。また、3D 点群に対し、 α -shape を用いて表面形状を構築することで体積を算出できることを示した。

今後は、他の農作物に対する 3D 点群構築の性能評価など汎用性についての検討も進めていく。

6. 謝辞

本研究の一部は、JST 創発的研究支援事業 (JPMJFR201B) の支援を受けたものである。また、データセットの収集にご協力いただいた中伊豆ワイナリー、ヤマハ発動機株式会社の皆様に感謝の意を表す。

7. 参考文献

- [1] 布施考志：解説：Structure from Motion (SfM) 第二回 SfM と多視点ステレオ、写真測量とリモートセンシング、Vol. 5, No. 4, pp. 259-262 (2016).
- [2] Mónica, H. H., Diego, H. A., Pablo R. G., et al.: Vineyard yield estimation by automatic 3D bunch modelling in field conditions, Computers and Electronics in Agriculture, Vol. 110, pp. 17-26 (2015).
- [3] Kai, T. F., Loriot, S. and Yvinec, M.: CGAL 5. 5. 1 - 3D Alpha Shapes User Manual, available from <https://doc.cgal.org/latest/Alpha_shapes_3/index.htm> (accessed 2023-01-13).
- [4] Lakubovskii, P.: segmentation_models.pytorch, GitHub(online), available from <https://github.com/qubvel/segmentation_models.pytorch> (accessed 2023-01-13).