

フォトグラメトリを用いた農作物の体積推定手法の提案

小野坂捺¹ 内海智仁² 峰野博史^{1,3}

概要：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。中でも2D画像のみから3D点群の構築が可能なフォトグラメトリ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、導入コストが安価であるといった長所を持つ。フォトグラメトリによって構築される3D点群は2Dにはない立体的な情報を持つことから、収量予測を目的とした農作物の体積推定などへの活用が期待されている。ただし、境界の曖昧さが原因で、点群から直接体積を推定することが難しいといった課題がある。また、農作物の果実部分のみの体積を計量したい場合、3D点群用の物体検出モデルやセグメンテーションモデルは2D画像用の学習モデルに比べ未成熟であることから、構築された3D点群から果実部分のみを自動で取り出すことは難しい。本研究では、ワインブドウ圃場の動画データから圃場画像を切り出し、フォトグラメトリを用いて構築した3D点群に対して、alpha-shapeを用いて表面形状を復元し高精度な体積推定できることを示す。更に、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用いて、圃場画像から体積を推定したい房部分のみのマスク画像を作成し、フォトグラメトリで点群を構築する際の入力に加えることで、圃場画像から房部分のみの3D点群を構築し体積を推定する手法を提案する。

キーワード： フォトグラメトリ、セマンティックセグメンテーション、スマート農業、体積推定

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。2D画像間の対応関係をソフトウェア解析することで3D点群データを構築するフォトグラメトリ(SfM/MVS)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、導入コストも安価であるといった長所を持つ。

フォトグラメトリによって構築される3D点群は2D画像にはない立体的な情報を持つことから、農業分野でも収量予測等を目的として、農作物の体積推定など様々な形での活用が期待されている[1][2]。ただし、拡大すると分かるが3D点群の境界は曖昧であるため、3D点群から直接体積を推定することは困難である。また、VoteNet[3]のように3D点群からの物体検出モデルの研究開発が注目されているが、2D画像からの物体検出モデルと比べまだまだ発展途上であり、対象とする物体も屋外の人や車、あるいは屋内の家具といった一般物体を対象としており、植物の3D点群データを対象とした研究は見当たらない。また、3D点群からの物体検出モデルは、立方体のバウンディングボックスでの検出を想定しているため、検出された物体の体積を推定するためには3D点群でセグメンテーションされたデータセットが必要となる。ただし、前述のとおり3D点群は境界が曖昧なため、3D点群から対象物体をセグメンテーションする作業は、2D画像の場合に比べて困難で労力を要する。そのため、3D点群から体積を推定したい部分(例えば、農作物の果実部分)のみを正確に抽出する技術の研究開発が望まれている。

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とし、フォトグラメトリによって構築された3D点群から農産物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。まずはワインブドウ圃場で撮影された収穫時期の動画に対し、3D点群を構築し房部分の体積を推定する手法について検討する。

以降、本稿の構成を示す。第2章では関連技術について述べ、第3章で提案手法について説明する。第4章では、机上の収穫済みワインブドウを様々な角度から撮影した画像でフォトグラメトリを行って構築した3D点群に対し、alpha-shapeを用いた体積推定の精度に関する基礎実験についてまとめる。第5章では、実圃場のワインブドウ画像に対し、房部分をアノテーションして学習させたセグメンテーションモデルを用いて生成させたマスク画像を適用し、ブドウの房部分のみの3D点群を構築した結果について説明する。最後に第6章で本稿をまとめる。

2. 関連技術

2.1 フォトグラメトリ

広義のフォトグラメトリは、2D画像を用いた測量全般を指すことが多いが、本稿では狭義のSfM(Structure from Motion)とMVS(Multi-View Stereo)[4]によるものを指すこととする。SfMでは、対象物を様々な角度から撮影した各画像から特徴点を抽出し、その特徴点を基に、カメラの位置と姿勢推定(アライメント)を行い、3次元空間上に特徴点をマッピングする。特徴点の3D点群は疎な点群であ

1 静岡大学情報学部
Faculty of Informatics, Shizuoka University

2 ヤマハ発動機株式会社
Yamaha Motor Co., Ltd.

3 静岡大学大学院情報領域/グリーン科学技術研究所
College of Informatics, Academic Institute, Shizuoka University /
Research Institute of Green Science and Technology

るため、より詳細な表面を表現する3D点群を得るために、MVSによって密な点群を獲得する。フォトグラメトリによって得られる点群は、LiDARなどで得られる点群とは異なり、2D画像のみから3D点群を構築するため基本的にはスケール情報を持たない。ただし、GPS情報を併用するなどの手法でスケール情報を組み込むことが可能である。

2.2 低解像度カメラを用いた植物の3D点群構築

Santos, T.T.ら[2]は、低解像度なカメラであっても被写体に接近して撮影するという単純な手法で高精度に3D点群を構築可能であることを示した。また、フォトグラメトリによって構築したトウモロコシの3D点群を用いて、地平面からの高さや葉長を、真値との誤差1%未満の精度で推定できることを示した。また、スケール情報を得るにあたり、点群構築対象の植物と一緒にスケールが分かるもの(テープメジャー)も撮影し点群化することで、スケール情報を獲得するという方法をとった。本研究においてもこれを参考に、ブドウと一緒に定規を撮影し、定規ごと点群化することでスケール情報付きの点群を得ることとした。

2.3 凸包を用いたブドウの体積推定

Mónica H[5]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(convex hull)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点群から手作業で構築したCADモデルに対して凸包を適用した場合と性能比較した結果、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と、CADモデルを構築する半手動的な手法に比べ相対的に高い決定係数であったと述べている。つまり、3D点群に対し凸包を用いて構築された3D形状から、比較的高精度かつ最小限の労力でブドウの房の体積、重量、粒数を推定できる可能性が示された。

2.4 関連技術のまとめ

前述のように植物の高さや葉の長さといった情報は、3D点群データから直接算出することが可能であるのに対し、3D点群の境界の曖昧さなどが原因で、体積については点群から直接推定することは困難である。そのため先行研究では、3D点群に対し凸包を適用することでブドウ房の体積を算出可能としている。しかし、凸包はアルゴリズムの特性上、窪みのある形状に対してはその窪みを埋めるような形で表面形状を構築するため、房の形状によっては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部を求めるalpha-shape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、フォトグラメトリで3D点群を構築する際に、2D画像用の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を入

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点群用の物体検出モデルやセグメンテーションモデルを使わなくとも、特定部分のみの画像群から対象物の3D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手順からなる。第一ステップでは、ブドウの房のみの3D点群データから体積を推定する体積推定モデルを構築する。この体積推定モデルを構築するために、収穫されたブドウの房の3D点群データと体積を高精度に算出する3Dモデルの作成が必要である。収穫された個々のブドウを判別できるように異なる番号のタグを付け、それぞれ順番に机の上に置いて様々な角度からフォトグラメトリ用の画像を生成する動画を撮影する。その際、ブドウの横に定規を置くことで、スケール情報を点群に付与できるようにしておく(図1)。

また、体積推定だけでなく重量推定や粒数推定も可能なように、収穫されたブドウの房の撮影が終わった後、一房毎に重量を計測するだけでなく、全ての粒を取って一房毎の粒数も記録しておく。フォトグラメトリによって生成されたスケール情報付きの3D点群に対して、alpha-shapeで表面形状を構築し高精度な3Dモデルを生成する。この3Dモデルを用いて体積の算出を行って目的変数とし、フォトグラメトリで生成された3D点群データを入力とした学習を行うことで、体積推定モデルを構築する。

第二ステップでは、圃場動画から切り出されたワインブドウ圃場画像に対し、ブドウの房部分のみの3D点群データを生成する。ここで、ブドウの房部分のみの3D点群データの生成には、別途2D画像で房部分を抽出できるように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し、ブドウの房部分のみを抽出するマスク画像として用いる。こうして生成されたブドウの房部分のみの様々な角度の2D画像を用いてフォトグラメトリすることで、ブドウの房部分のみの3D点群データを生成する。



図1 ブドウ房撮影画像の例

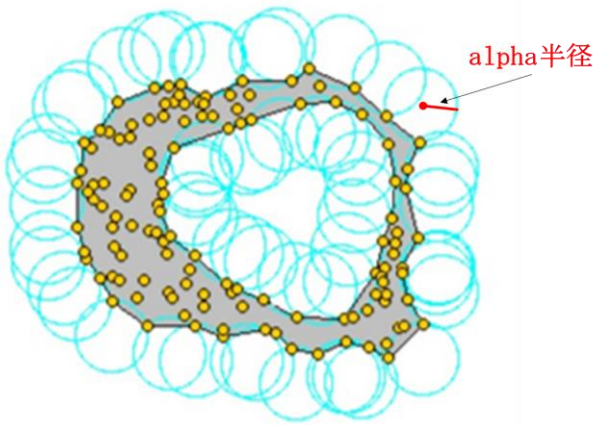


図2 alpha-shape の概要

3.2 高精度な体積データの作成

第一ステップでは、ブドウの房のみの 3D 点群データから体積を推定する体積推定モデルを構築する。収穫された個々のブドウを判別できるよう異なる番号のタグを付け、定規を横において、それぞれ順番に机の上に置いて様々な角度からフォトグラメトリ用の画像を生成する動画を撮影する。フォトグラメトリによって生成されたスケール情報付きの 3D 点群に対して、alpha-shape で表面形状を構築し高精度な 3D モデルを生成する。

図2に、alpha-shape の概要を示す[6]。alpha-shape は凸包の概念を一般化したものであり、H. Edelsbrunner[7]らによる直感的な説明を参考にすると、3D 点群からの alpha-shape の構築は、3 次元空間上に存在する複数の「点」を含む形状の不明な「物体」を「球」で切り取るイメージに例えられる。各点に対し、接触することなく到達可能な物体の全ての部分（内部を含む）を切りとった結果、最終的に残る物体の形状が alpha-shape となる。ここで、点群によって形成される物体を切り取る「球」の半径を alpha 半径と呼び、alpha 半径の値を変更することで、点群の表面形状を再現する際の精細さを調整できる。なお、alpha 半径の値を無限大にすることで構築される表面形状が、前出の凸包となる。以降、凸法と明確に区別しやすくするために、本稿では alpha 半径の値が無限大でないもののみを alpha-shape と表記することとする。Gardiner, J.D. [8]らは、alpha-shape の体積が元のデータセットの体積と一致するために必要な alpha 半径の最小化が、形状の「複雑さ」を測る指標として採用できることを示している。このことから、窪みの多い複雑な形状の房であっても alpha 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、alpha-shape で点群に対して表面形状を構築する場合、凸包と異なって表面に穴が開いてしまうことがある。穴が開いた場合、alpha-shape の中身は空洞となってしまうため、実際の体積よりも極端に小さい体積が算出されてしまう。より大きな alpha 半径を設定することで、表面に穴の空いた形状となることを塞ぐことができるが、alpha 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、本研究では alpha 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することとする。例えば、alpha 半径の値を少しずつ大きくしていくと、穴がふさがった場合、ふさがった瞬間には alpha-shape の内部に生じていた空洞が無くなるため、急激に体積が上昇する。つまり、体積の変動量に対する閾値を経験則によって設定することで、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。

3.3 ブドウの房部分のみのマスク画像の生成

第二ステップでは、実圃場を撮影した画像からブドウの房部分のみの 3D 点群データを生成する。このブドウの房部分のみの 3D 点群データを生成するために、別途 2D 画像で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し、房部分のみを抽出するマスク画像として用いる。実圃場において各ワインブドウの房は非常に密集していたり、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無く、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。

図3に、ワインブドウ圃場で撮影された動画から切り出された圃場画像(a)と、ブドウの房部分のみアノテーションしたマスク画像(b)を示す。このマスク画像を生成できるよ



(a) 元画像



(b) アノテーションしたマスク画像

図3 元画像(上)とアノテーションしたマスク画像(下)

う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。

4. 基礎実験：机上のブドウ画像での検証

4.1 概要

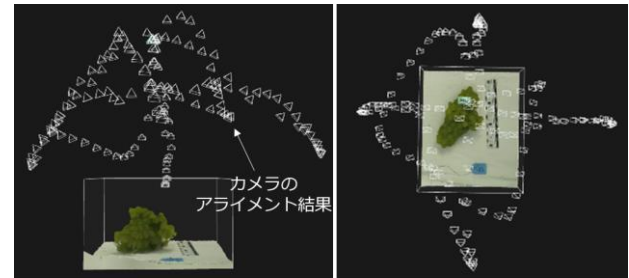
本提案手法の第一ステップでは、ブドウの房のみの 3D 点群データから高精度に体積を算出することが重要となる。凸包を用いて 3D モデルを生成する手法は、ブドウのような複雑な形状で窪みのある形状に弱いため、alpha-shape によって表面形状を構築し高精度な体積の算出ができることを確認する。また、ブドウの房を撮影した 2D 画像から房部分の面積の算出を行い、凸包や alpha-shape を用いて形成される 3D モデルを用いて重量や粒数を推定する手法が、2D 画像を用いて推定する手法に対してどの程度の優位性を持つのか定量的に比較する。

表 1 に、本基礎実験で使用した 27 房のワインブドウの房動画データの概要を示す。収穫された個々のブドウを判別できるよう異なる番号のタグを付け、それぞれ順番に机上に置いて様々な角度からフォトグラメトリ用の画像を生成する動画を撮影した。フォトグラメトリでは、一般に各画像間で一定程度のオーバーラップ率が確保されていると 3D 点群データの構築に成功しやすいとされている。そのため、様々な角度で撮影をした動画から、約 60~80% のオーバーラップ率を確保した画像を抽出した。フォトグラメトリには、高精度な 3D 点群構築が可能であった RealityCapture を採用することとした。

ここで抽出された画像には、ブドウの房以外に各房を識別するためのタグとスケール情報を付与するための定規も含まれているため、房以外の不要な 3D 点群も生成されてしまう。そのため、これらの不要な 3D 点群を除去するにあたり、python の 3D 点群処理用ライブラリである open3d[9]を用いて平面セグメンテーションを行った。これは Random Sample Consensus (RANSAC) アルゴリズムを用いたセグメンテーションであり、試行錯誤の末、引数 ransac_n の値を 3、num_iterations の値 (RANSAC のサンプリング処理と評価処理の繰り返し回数) を 500 に設定した。また、房の下部分が平面判定によって除去されないように、distance_threshold (平面のインライアとして判定するための距離の閾値) の値を 0.0017 に設定した。ここで、除去しきれなかった不要な点については手動で除去を行った。フォトグラメトリでは、画像中に写っていない部分には 3D 点群が構築されないため、ブドウの房と机が接触している部分には 3D 点群が構築されない。alpha-shape の場合、穴が開いている部分が大きくえぐられ、実際の体積の値より小さい値になることがある。そのため、平面の点群を用意することで、手動で机と接触している部分の穴をふさぐこととした。

表 1 ワインブドウの房動画データ

撮影対象	ワインブドウの房(収穫済み)
房数	27房
撮影場所	机上
機材	iphone11
画像サイズ (ピクセル)	幅: 1080 高さ: 1920
ファイル形式	MOV(学習時: JPG)



(a)横からのアングル (b)上からのアングル

図 4 点群とカメラのアライメント結果

以上の手順で生成したブドウ房の 3D 点群に対し、MATLAB を用いて凸包と alpha-shape を適用し 3D モデルを構築した。ここで、alpha-shape は、表面の穴がふさがるまで alpha 半径の値を 0.0001 ずつ上昇させ急激な体積の値の変化があった時の体積を推測値とした。

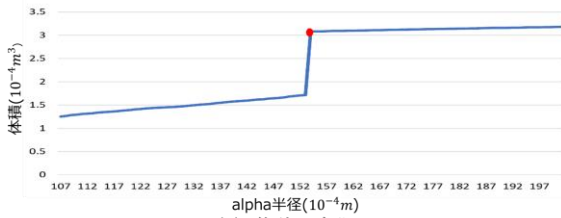
また、凸包、alpha-shape といった 3D モデルの性能比較用に、真上からブドウを撮影した画像と、机と平行に真横から撮影した画像の合計 2 パターンの角度で撮影された 2D 画像で算出される面積を用いることとした。ここで、画像へのスケール情報の付与、面積算出には、画像処理ソフトの ImageJ を用いた。

一方、凸包、alpha-shape を適用した 3D モデルの体積比較に関して、今回収穫したワインブドウはそのまま醸造に使用することによって、ワインブドウを水に沈めて正確な体積を計測することはワインの品質低下に繋がり不可ということであった。そのため、体積の真値を得ることはできなかったが、体積と強い正の相関があることが明らかとなっているブドウの重量と粒数で性能比較を行うこととした。体積および面積の推測値と、実測のブドウ重量と粒数の相関係数、決定係数をもとに精度を検証した。

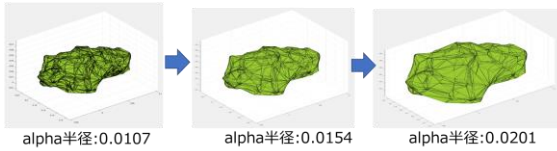
4.2 実験結果

図 7 に、フォトグラメトリによって構築された 3D 点群の例を示す。図中の白色の三角形は、カメラの撮影位置と角度のアライメント結果である。すべての房について問題なく 3D 点群を構築することができた。

図 5 に alpha 半径の値を 0.0001 ずつ上昇させたときの alpha-shape の体積と見た目の変化を示す。急激な体積変化が確認されたとき (グラフ内の赤点)、表面の穴が塞がることが確認できた。また、急激な値の変化があった後も alpha 半径の値を上昇させ続けると、alpha-shape の形状が徐々に



(a) 体積の変化



(b) 形状の変化

図 5 alpha 半径の変化の影響 (体積, 形状)

凸包に近づいていき、実際のブドウの体積から乖離していくこと、今回用いた 27 房分のブドウ全てで同様の結果となることが確認できた。

表 2 に、各房の重量、粒数、真上からの面積、真横からの面積、凸包の体積、alpha-shape の体積、間の相関行列を示す。特に alpha-shape の体積は、重量と粒数のいずれに対しても従来手法である凸包よりも高い相関 0.981 と 0.916 を示した。また、真上からの面積は、重量と粒数に対し、alpha-shape には及ばないものの 0.961 と 0.902 を示し、いずれも凸包の 0.952 と 0.888 を上回る高い値を示した。一方、真横からの面積は、それぞれ 0.865 と 0.827 となり、撮影角度の違いに相関が大きく下がる結果となった。実圃場においては、ブドウの房は様々な向きで実っており、さらに葉などの障害物に隠れていることもあるため、ある側面の 2D 画像から重量や粒数を常に高精度に推定するのは難しいことが分かった。

これらの結果から、面積といった 2D 画像ではなく、凸包や alpha-shape といった 3D モデルを活用したほうが、よ

表 2 相関行列

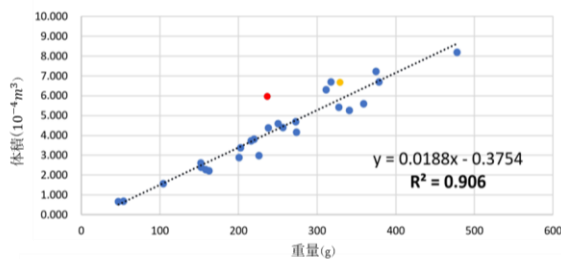
	重量	粒数	真上からの面積	横からの面積	凸包の体積	alpha-shapeの体積
重量	1					
粒数	0.938	1				
真上からの面積	0.961	0.902	1			
真横からの面積	0.865	0.827	0.887	1		
凸包の体積	0.952	0.888	0.977	0.888	1	
alpha-shapeの体積	0.981	0.916	0.978	0.868	0.979	1

り安定して高精度に体積や重量、粒数といった推定が可能である見通しを得た。

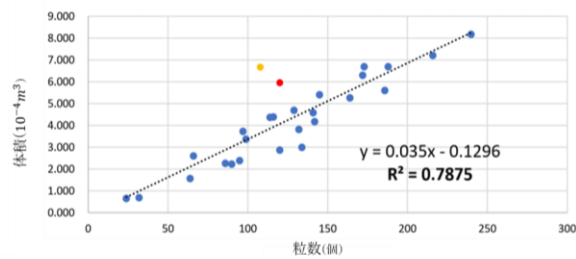
4.3 詳細分析

図 6 に、凸包ならびに alpha-shape で得られた 27 房のブドウ房体積に対するそれぞれの重量と粒数の散布図と近似直線の決定係数を示す。図中の (c), (d) に示されるように、ブドウ房の体積は重量や粒数と線形関係があり、特に alpha-shape の体積は重量に対し決定係数 0.963、粒数に対し決定係数 0.838 と、凸包の体積よりも高い決定係数を示すことが分かる。

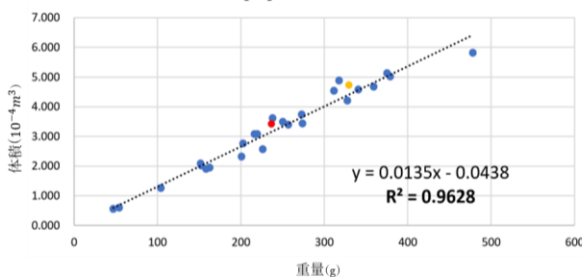
ここで図中に赤色の点として表記した ID36 のブドウ房に着目する。図 6 (a), (b) に示されるように、凸包の体積と重量ならびに粒数の散布図では、近似直線から離れた位置にプロットされているのに対し、図 6 (c), (d) に示される alpha-shape の体積と重量ならびに粒数の散布図では、近似直線の近くにプロットされている。図 7 に、ID36 のブドウ房の画像と、凸包ならびに alpha-shape で構築されたブドウ房の 3D モデルを示す。撮影画像から、このブドウ房は二股に分かれた形状をしていることが分かる。このような形状のブドウ房の 3D 点群に対して、alpha-shape は二股の形状を適切に構築できていたのに対し、凸包は二股を埋めるような形で表面形状を構築しており、実際の体積より大きな値を算出していた。つまり、複雑で窪みの大きい形状のワインブドウの房の場合、凸包では正確な値の算



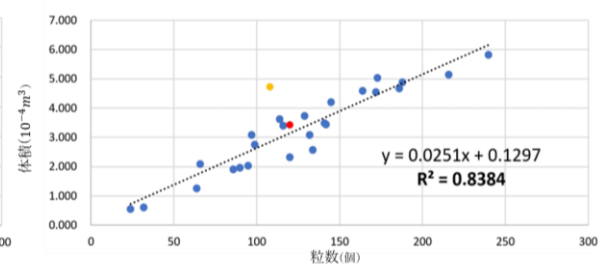
(a)凸包と重量



(b)凸包と粒数



(c)alpha-shapeと重量



(d)alpha-shapeと粒数

図 6 重量、粒数と凸包、alpha-shape の散布図と決定係数

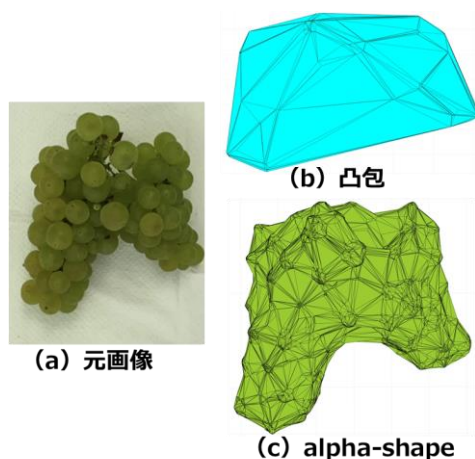


図 7 ID36 のブドウ房と構築された 3D モデル

出が困難であった。

次に、図 6 中で黄色の点で表記した ID58 のブドウ房に着目する。図 6 (a), (c) に示されるように、凸包の体積と重量ならびに alpha-shape の体積と重量の散布図では、近似直線から大きく離れた位置にプロットされていない。一方、(b), (d) に示されるように、粒数の散布図では近似直線から大きく離れた位置にプロットされている。

このようなことの生じる要因について考察する。粒数とは別に粒径という指標がある。粒径の算出方法は様々あるが、今回ブドウの房からランダムに選んだ 6 粒を直線状に並べた直径の和で算出することとした。オレンジ色の点で示した ID58 のブドウ房から得られた粒径は、12.6cm であった。全 27 房の平均粒径は 13.05cm であったため、ID58 のブドウ房における粒径は、平均より 1cm 程小さい値であった。ID58 のブドウ房の重量はほぼ真値と同じで、粒数が真値よりかなり大きかったため、感覚的には粒径が大きかったのだろうと想像できるが結果は逆であった。ID58 のブドウ房の粒数が、凸包でも alpha-shape の体積からも正確に予測できない要因として、ブドウ房における果実部分以外の穂軸や支梗などが、他の房に比べて大きく重かった可能性が考えられる。

一方、重量や粒数の推定に、粒径が与える影響はそれほど大きくないということも明らかとなった。今回、詳細な計量を行った 27 房分のブドウ房の重量ならびに粒数と粒径との相関係数を調べたところ、それぞれ 0.121, 0.052 とほぼ相関が無いことが分かった。つまり、より正確な収量予測を実現するには、穂軸及び支梗の大きさも考慮した方がよく、凸包も alpha-shape の体積のいずれにおいても穂軸や支梗の大きさまでは考慮できていないため今後の課題となる。もし今後、穂軸や支梗の大きさが分かるようになれば粒径も推測可能になる。粒径はワインブドウの品質計測の際に利用されるため、粒径の推定が可能となれば更なる実用性の向上が期待できる。

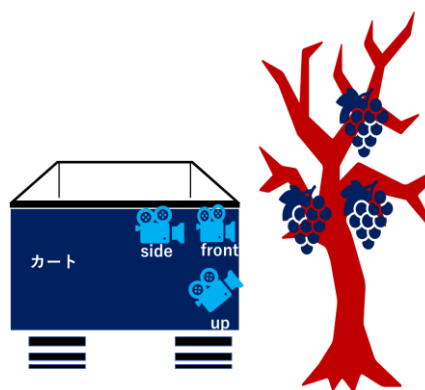


図 8 カメラアングル

5. 圃場のブドウ画像を用いた検証

5.1 アノテーション画像を用いた 3D 点群構築

収穫された状態のブドウではなく、実圃場において木に生っている状態のワインブドウの体積を算出することを目標に、セマンティックセグメンテーションでブドウ房部分のマスク画像を生成すれば、ブドウ房部分のみの 3D 点群を構築できることを確認する。

図 8 に、本検証に用いる動画データを撮影した環境の概要を示す。図 8 においてカートは図の奥方向に進み、カート後部にそれぞれ異なる角度に取り付けられた 3 台の GoPro Hero9 でワインブドウ (品種: シャルドネ) の群落の動画を撮影した。図 8 において side はカートの進行方向に対し直角、front は side に比べややカートの進行方向に傾いた向き、up は群落を見上げるような方向を向いている。表 3 に、収集された動画データの概要を示す。

撮影した動画を、おおよそ画像間で 80%以上のオーバーラップ率が確保される画像に変換する。変換された各画像に対し、撮像内の全てのブドウ房に対するアノテーションを行う。アノテーション結果の json ファイルから、図 3(b) に示すようなマスク画像を生成し、フォトグラメトリの際に元画像とともに入力として用いた。マスク画像を用いた際の点群の精度を検証するために、以下 A, B, C に示す 3 つの入力パターンで 3D 点群の構築を試した。

- A) マスク画像なし: 角度 front の画像 10 枚
- B) マスク画像あり: 角度 front の画像 10 枚
- C) マスク画像あり: 角度 front&side の画像 20 枚

図 9 に、A, B, C の各入力パターンで構築された 3D 点群の画像を示す。マスク画像なしで、角度 front の画像 10 枚

表 3 実圃場の動画データ

撮影対象	ワインブドウの木
機材	GoPro Hero9
画像サイズ (ピクセル)	幅: 5120 高さ: 2880
撮影角度	front, side, up
ファイル形式	MP4(学習時: JPG)

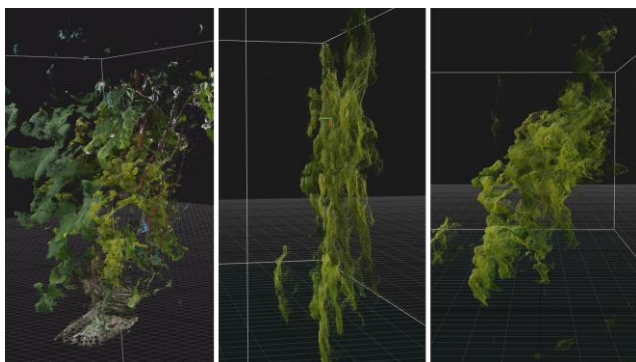


図 9 点群構築結果 (左から順に入力が A, B, C の場合)

を用いて 3D 点群を構築した入力パターン A の場合、実際のブドウの木に近い立体的で奥行のある 3D 点群が構築された。一方、入力パターン B の場合、平面的で厚みのない 3D 点群が構築された。マスク画像を用いたことで、カメラのアライメントや疎な点群の描画などに使用する特徴点が減少してしまったことが要因と考える。入力パターン C のように、異なる角度の画像も加えて 3D 点群を構築したところ、厚みのある 3D 点群が構築できることを確認した。つまり、マスク画像を用いて 3D 点群を構築する場合、より適切な座標の点群を生成するには、マスクなしの場合よりも多様かつ多数の画像が必要であるといえる。

5.2 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

アノテーション画像からではなく、アノテーション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。

画像データベースとして一般的な 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[10]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。精度の評価指標には、mIoU を用いた。

表 4 に、エンコーダとアーキテクチャの組み合わせと mIoU の結果を示す。これらの結果から、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、

表 4 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	0.712675
ResNet101	UNet	0.6988
ResNet152	UNet	0.6865
VGG19	UNet	0.747869
ResNet34	DeepLabV3+	0.870408
EfficientNetB2	DeepLabV3+	0.872241
EfficientNetB4	DeepLabV3+	0.881892

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 10 に、本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、赤い丸で囲んだ部分のように、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、房以外の不要な 3D 点群が生成されていた。

房以外の不要な 3D 点群が生成される課題の解決策として、セマンティックセグメンテーションモデル構築に用いる学習用データセットの枚数を増やすだけでなく、入力時のリサイズ時に可能な限り解像度が落ちないように工夫するといった手法が考えられる。また、現時点では発展途上である 3D 点群用のセグメンテーションモデルの性能が向上すれば、それを適用するといった手法も考えられる。



図 10 マスク画像を用いた点群構築結果
 ((a)元画像, (b)構築点群)

6. おわりに

本研究では、ワインブドウの 3D 点群から表面形状を再現する際に、従来手法の凸包ではなく alpha-shape を用いることで高精度な表面形状の再現ができ、正確な体積を算出できることを示した。また、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用いて、ブドウ房部分のマスク画像を生成し、マスク画像をフォトグラメトリで点群構築する際の入力として加えることで、圃場画像からブドウの

房部分のみの 3D 点群を構築する手法を検証した。

今後は、一部に不要部分が残ってしまった実圃場におけるブドウ房の 3D 点群に対し、3D 点群用の物体検出モデルもしくはセグメンテーションモデルを適用することで、ブドウ房部分のみの 3D 点群の構築性能の向上を図れるか検証する。また、実圃場を撮影した画像から生成した房部分の点群に対し、alpha-shape を用いて表面形状を構築し、机上条件との精度の差についても検証したい。今回用いた手法には、ブドウ房に特化した処理は含まれていないため、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。

7. 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る。

8. 参考文献

- [1] 藤原 峻, 保田 浩, 齋藤 正博ほか: 高精度測位システムを搭載した Unmanned Aerial Vehicle (UAV) によるイネ稈長推定法の検討, 育種学研究, Vol.24, No.1, pp.12-21 (2022).
- [2] Santos, T. T. and Rodrigues, G. G.: Flexible three-dimensional modeling of plants using low-resolution cameras and visual odometry, Machine Vision and Applications, Vol.27, pp.695-707 (2015).
- [3] Qi, C. R., Litany, O., He, K., et al.: Deep Hough Voting for 3D Object Detection in Point Clouds, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp.9277-9286 (2019).
- [4] 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).
- [5] 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).
- [6] Kai, T. F., Lorient, S. and Yvinec, M.: CGAL 5. 5. 1 - 3D Alpha Shapes User Manual, available from https://doc.cgal.org/latest/Alpha_shapes_3/index.html (accessed 2022-10-26)
- [7] Edelsbrunner, H., Mücke, E. P.: Three-dimensional alpha shapes, ACM Trans. Graph., 13(1), pp43-72 (1994)
- [8] Gardiner, J. D., Behnsen, J. and Brassey, C. A.: Alpha-shapes: determining 3D shape complexity across morphologically diverse structures, BMC Evolutionary Biology, Vol. 18, No.184 (2018)
- [9] Open3D, available from <http://www.open3d.org/> (accessed 2022-10-26)
- [10] Lakubovskii, P.: segmentation_models.pytorch, GitHub (online), available from <http://www.squeakland.org/community/biography/alanbio.html> (accessed 2022-12-20)