

ブドウの 3 次元構造把握のための球体検出を用いた果粒検出

山崎 匠真^{1,a)} 内海 ゆづ子^{1,b)} 三輪 由佳^{2,c)} 岩村 雅一^{1,d)}

概要

ブドウ生産における摘粒作業は、果粒の肥大促進や房の整形を目的として成長前のブドウの粒数を減らす重要な作業である。しかし、複雑な判断基準と迅速な作業が求められるため、未経験者の習得には相当な時間を要する。本研究では、新規就農者の摘粒作業習得を支援するため、ブドウの 3 次元構造把握に基づく果粒検出システムを提案する。従来の深度センサーや画像ベースの 3 次元復元手法は、照明環境の変動や果粒のテクスチャレスな表面特性により、圃場環境での実用が困難であった。そこで本研究では、3D Gaussian Splatting を用いて疎な初期点群から高密度な 3 次元表現を生成し、RANSAC により果粒を検出する新しいアプローチを提案する。提案手法では、3D Gaussian Splatting の高品質な 3 次元復元能力と RANSAC の外れ値に対するロバスト性を組み合わせることで、密集・遮蔽・類似形状というブドウ果粒特有の困難な条件下でも高精度な果粒検出を実現する。シャインマスカットを対象とした実験において、実際の果粒数 39 個に対して 38 個を検出し、果粒数検出における相対誤差 2.56% の高精度な結果を得た。本手法により、テクスチャレスなブドウ果粒の検出が可能となった。

1. はじめに

ブドウ生産の作業過程の中で、特に生産者への負担が大きいのが房から果粒を間引く摘粒作業である。摘粒とは、果粒の肥大促進、房の整形、実割れ防止を目的として、1 房あたりの果粒数がそれぞれの品種ごとに適正になるよう、房の形を整えながら粒数を減らす作業である。摘粒する果粒の選定には、果粒の間隔や房の各部位の果粒数など、複数の基準がある。この摘粒作業は 6 月上旬から下旬の間に全ての房で行わなければならない、迅速な作業が必要である。摘粒の基準が複雑であり、かつ素早い作業が求められるため、未経験者が作業を習得するには相当な時間を要する。

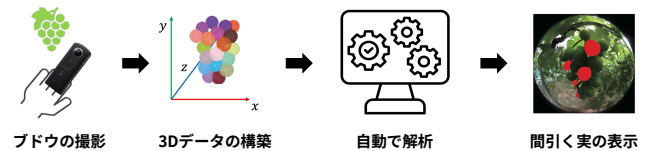


図 1: 摘粒作業習得支援システム

そのため、情報技術を用いた習得支援が期待される。

本研究では、新たにブドウ栽培を始める人の摘粒作業習得を支援することを目的として、摘粒すべき果粒を自動で判別し、利用者に指示するシステムを構築する。図 1 に本研究で開発するシステムを示す。このシステムでは、利用者がブドウの房を撮影し、その房における摘粒すべき果粒を画像で示す。

ブドウの適切な摘粒管理を実現するためには、個々の果粒の正確な 3 次元空間位置の推定が不可欠である。3 次元情報を取得する手法として、深度センサーの利用が一般的であるが、既存研究 [7], [11] で提案された手法は、照明環境が制御された室内での使用を前提としており、日照条件が常に変動する圃場環境下での実用は困難である。これらの制約を考慮すると、環境変動に対してより堅牢で、かつ導入コストの低い単眼カメラベースの手法が適切である。

単眼カメラベースの一般的な 3 次元復元手法として Structure from Motion (SfM) [1] や Visual Simultaneous Localization and Mapping (Visual SLAM) [3], [8] が提案されているが、ブドウ果粒のような表面テクスチャに乏しい物体においては、特徴点の抽出およびマッチング処理の信頼性が著しく低下し、場合によっては処理が困難となる。この問題により、3 次元点群データの密度が不十分となるか、あるいは位置精度が劣化する。また、深層学習ベースの手法として、田村ら [15] は教師なし単眼深度推定を用いてブドウの 3 次元復元を試みているが、深度推定の精度やフレーム間での特徴点の対応付けに課題が残されている。

近年、Kerbl ら [5] によって提案された 3D Gaussian Splatting は、SfM で得られる初期点群を入力とし、3D Gaussian のパラメータを最適化することで、高品質な自由視点画像を得る手法として注目されている。この手法は、初期点群の精度が十分でない場合でも、最適化の過程で点群の数を調整し、その結果、幾何学的に正確な点群を得

¹ 大阪公立大学大学院情報学研究所

² 大阪府立環境農林水産総合研究所

a) sc24951c@st.omu.ac.jp

b) yuzuko@omu.ac.jp

c) MiwaY@knsk-osaka.jp

d) masa.i@omu.ac.jp

ることが可能である．本研究では特に，3次元データ構築の部分に焦点を当て，3D Gaussian Splatting と RANdom SAmple Consensus (RANSAC) [10] を組み合わせた新しい手法を提案する．これにより，ブドウの房の正確な3次元構造の把握を目指す．

2. 関連研究

ここでは，3次元復元手法について，深度センサーを用いた手法，画像ベースの3次元復元手法，そしてテクスチャレスな物体の3次元復元手法について，それぞれ説明する．

2.1 深度センサーを用いた3次元復元

深度センサーを用いた3次元復元は，直接的に対象物体の3次元情報を取得できる利点がある．しかし，既存手法 [7], [11] では室内環境での比較的大きな深度センサーの使用を想定しており，照明条件が安定しない圃場での使用は困難である．また，ブドウ果粒のような光沢のある物体に対しては，赤外線反射の影響により深度情報の取得が不安定になる．

2.2 画像ベースの3次元復元

複数視点の画像や動画を使用して被写体の3次元復元を行う手法として，SfM [1], Multi-view stereo (MVS) [14], Visual SLAM [3], [8] などが広く研究されている．SfMは，複数の異なる視点から撮影された画像間の対応点を基に，カメラの姿勢と被写体の3次元構造を同時に推定する手法である．この手法では，まず画像間の特徴点マッチングを行い，次に Bundle Adjustment によりカメラパラメータと3次元点座標を最適化する．MVSは，SfMにより得られたカメラ姿勢を基に，ステレオマッチングを用いて密な3次元点群を生成する手法である．一方，Visual SLAMは，動画のシーケンスからリアルタイムでカメラの自己位置推定と環境の3次元地図構築を同時に実行する技術である．これらの手法では，複数の画像や動画から抽出した局所特徴量を用いて，被写体の対応点を検出することが基本となる．代表的な局所特徴量としては，スケール不変性を有する SIFT [6]，高速処理が可能な SURF [2]，バイナリ記述子を用いる ORB [13]，コーナー検出に特化した FAST [12] などが挙げられる．しかしながら，ブドウ果粒のようなテクスチャレスな被写体においては，表面の一様性により形状や輝度勾配といった識別可能な特徴量が極めて少ない．このため，特徴点の抽出密度が著しく低下し，さらに抽出された特徴点においても類似性が高いため，フレーム間での正確な対応付けが困難となる．特に，球状で表面が滑らかなブドウ果粒では，局所的な特徴の区別が困難であり，誤対応や対応点の欠失が頻発する問題がある．これらの要因により，従来の画像ベースの3次元復元手法では，ブドウのような農作物の高精度な3次元モデル生成が極めて困

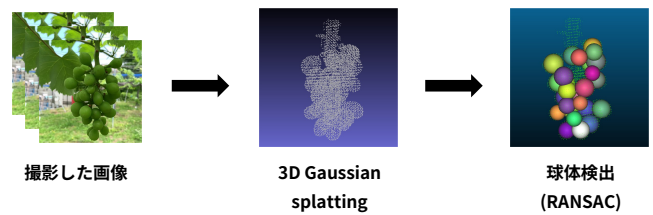


図 2: 提案手法の流れ

難である．

また，深層学習ベースの手法として，田村ら [15] は教師なし単眼深度推定を用いてブドウの3次元復元を試みているが，深度推定の精度や，フレーム間での特徴点の対応付けに課題が残されている．

2.3 テクスチャレスな物体の3次元復元

テクスチャレスな物体に対する3次元復元手法として，物体の輪郭やその表面から検出される曲線を使用した SfM [4], [9], [16] が提案されている．これらの手法では，2つの画像でのマッチングを可能にするため，画像上から検出される曲線が一意に定まることを想定している．しかし，ブドウ果粒のような遮蔽が多く，似た形状を持つ物体には曲線の対応が取れないため，適用が困難である．

近年，Kerblら [5] によって提案された 3D Gaussian Splatting は，これらの従来手法の課題を解決する新たなアプローチとして注目されている．この手法は，SfMにより得られる疎な初期点群を入力とし，各点を3次元ガウシアン関数でモデル化する．具体的には，各ガウシアンの位置，共分散行列，不透明度，色情報をパラメータとして定義し，これらを画像の再投影誤差を最小化するように適応的に最適化する．この過程において，ガウシアンの分割や統合，密度調整が自動的に実行されることで，テクスチャレスな物体表面においても滑らかで高品質な3次元表現が可能となる．特に，3D Gaussian Splatting は従来の特徴点に依存しない表現手法であるため，ブドウ果粒のような表面特徴に乏しい物体に対しても，連続的で詳細な3次元モデルの生成が可能である．

3. 提案手法

本研究では，テクスチャレスなブドウ果粒の3次元構造復元という従来手法では困難な課題に対して，3D Gaussian Splatting と RANSAC ベース球体検出を組み合わせた新しいアプローチを提案する．図 2 に提案手法の全体構成を示す．

従来の特徴点ベース手法では，ブドウ果粒の滑らかで一様な表面から十分な特徴点を抽出できず，正確な3次元復元が困難であった．本手法では，この問題を以下の2段階アプローチで解決する．

第1段階：3D Gaussian Splatting により，疎な初期点

群から高密度で滑らかな 3 次元表現を生成し、テクスチャレスな表面の詳細な幾何学情報を復元する。

第 2 段階: 生成された高密度点群に対して RANSAC ベース球体検出を適用し、果粒の位置・サイズを自動抽出する。

本手法では、3D Gaussian Splatting の高品質な 3 次元復元能力と、RANSAC の外れ値に対するロバスト性を組み合わせることで、密集・遮蔽・類似形状というブドウ果粒特有の困難な条件下でも高精度な果粒検出を実現する。

3.1 3D Gaussian Splatting による密な 3 次元点群生成

3D Gaussian Splatting では、シーン内の各点を 3 次元 Gaussian 分布 $G(x) = e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$ として表現する。ここで μ は位置、 Σ は共分散行列である。3D Gaussian を 2 次元画像平面に投影する際、 W を視点変換、 J を射影変換のヤコビ行列とすると、共分散行列 Σ' は以下のように計算される。

$$\Sigma' = JW\Sigma W^T J^T \quad (1)$$

c_i を各 Gaussian の色、 α_i を不透明度、 \mathcal{N} を画素に影響を与える Gaussian の集合とすると、画像の色 C は投影された 2 次元 Gaussian の重み付き和として以下のように計算される。

$$C = \sum_{i \in \mathcal{N}} c_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j) \quad (2)$$

3D Gaussian Splatting が持つ適応的密度制御では、SfM により得られた疎な点群から開始し、Under-reconstruction 領域での Gaussian 複製と Over-reconstruction 領域での分割により、段階的に表面詳細を学習する。この機構は、滑らかで一様な果粒表面の詳細な形状変化を効果的に捉えることができる。この過程により、初期の疎な点群は最終的に果粒表面の微細な形状変化まで捉えた高密度点群へと変換される。最適化により得られる各 3 次元 Gaussian の位置パラメータを μ_i 、総 Gaussian 数を N とすると、高密度点群 \mathcal{P} は以下のように表現される。

$$\mathcal{P} = \{\mu_1, \mu_2, \dots, \mu_N\} \quad (3)$$

3.2 RANSAC による果粒の球体検出

3D Gaussian Splatting によって推定した点群から個別の果粒を検出するため、球体モデルベースの RANSAC アルゴリズムを適用する。RANSAC は、入力点群に外れ値が多数含まれる場合でも頑健なパラメータ推定を行う手法である。ブドウ果粒を球と仮定して、RANSAC を使って点群から球を検出することで、果粒を検出する。

本手法では果粒を検出するため、Schnabel ら [10] による Efficient RANSAC for Point-Cloud Shape Detection (Efficient RANSAC) を適用する。Efficient RANSAC

Algorithm 1 RANSAC による球検出

Require: 点群 $\mathcal{P} = \{\mu_1, \mu_2, \dots, \mu_N\}$
Ensure: 検出された球集合 $\Psi = \{\psi_1, \dots, \psi_k\}$

- 1: $\Psi \leftarrow \emptyset$ /* 検出された球 */
- 2: $\mathcal{C} \leftarrow \emptyset$ /* 球候補 */
- 3: **repeat**
- 4: $\mathcal{C} \leftarrow \mathcal{C} \cup \text{newCandidates}()$ /* 新候補生成 */
- 5: $m \leftarrow \text{bestCandidate}(\mathcal{C})$ /* 最良候補選択 */
- 6: **if** $P(|m|, |\mathcal{C}|) > p_t$ **then**
- 7: $\mathcal{P} \leftarrow \mathcal{P} \setminus P_m$ /* インライア点除去 */
- 8: $\Psi \leftarrow \Psi \cup m$ /* 球を結果に追加 */
- 9: $\mathcal{C} \leftarrow \mathcal{C} \setminus C_m$ /* 無効候補除去 */
- 10: **end if**
- 11: **until** $P(\tau, |\mathcal{C}|) > p_t$
- 12: **return** Ψ

表 1: 果粒検出精度の評価結果

評価項目	値
実際の果粒数 (N_t)	39 個
検出された果粒数 (N_p)	38 個
相対誤差 (δ)	2.56%
検出率	97.44%

アルゴリズムを Algorithm 1 に示す。まず検出球集合 Ψ と球候補集合 \mathcal{C} を空集合として初期化する。処理の各反復では、点群から 2 点とその法線をランダムサンプリングし、球のパラメータである中心座標と半径を計算して候補集合 \mathcal{C} に追加する。次に、現在の候補集合 \mathcal{C} から最良の候補 m を選択する処理を行う。最良候補の選択は、各候補に対してインライア数を計算し、最大値を持つ候補を選ぶことで行われる。インライア数の計算では、点群中の各点 p_i と球体候補との距離を評価し、距離が許容誤差 ϵ 以内にある点をインライアとして判定する。選択された候補 m に対して検出成功確率 $P(|m|, |\mathcal{C}|)$ を計算し、この確率が要求信頼度 p_t を上回る場合、候補 m を有効な球として検出する。検出時には、球 m に適合するインライア点群 P_m を入力点群から除去し、検出された球 m を結果集合 Ψ に追加し、除去された点を含む無効な候補 C_m を候補集合から削除する。処理の終了判定では、最小インライア数閾値 τ に対する検出成功確率 $P(\tau, |\mathcal{C}|)$ を評価し、この確率が要求信頼度 p_t を上回った場合、残りの点群から球が検出される可能性が低いと判断して処理を終了する。

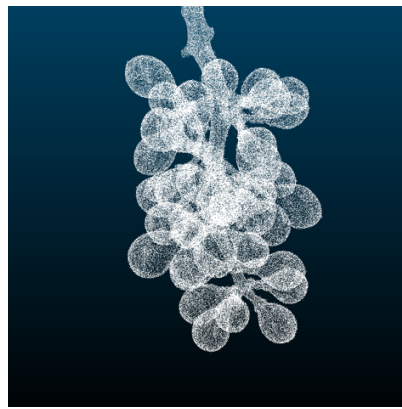
4. 実験

4.1 実験条件

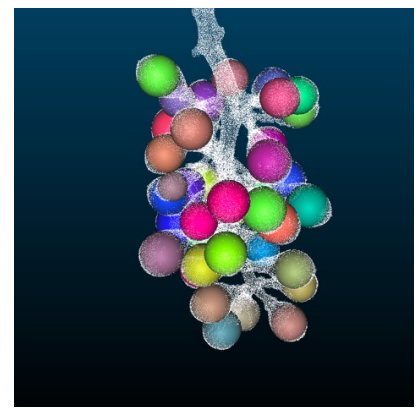
本実験では、提案手法を評価するために用いる画像を、iPhone XS を使用して撮影した。撮影は大阪公立大学農学部附属教育研究フィールドにおいて 2025 年 6 月 6 日に実施し、シャインマスカット 1 房を対象として動画撮影を行い、その動画から 1297 枚の画像を取得した。撮影時の解像度は 1080×1920 ピクセル、フレームレートは 60fps に



(a) 3D Gaussian Splatting による復元結果



(b) 復元結果の房部分の点群



(c) RANSAC による果粒検出結果

図 3: RANSAC による果粒検出結果

設定し、房の周囲を一周回りながら撮影することで、3次元復元に必要な十分な情報を収集した。

3次元復元処理には複数のソフトウェアを組み合わせで使用した。3D Gaussian Splatting の処理には Postshot^{*1}を採用し、点群データの処理および RANSAC アルゴリズムによる果粒検出には Cloud Compare^{*2}を使用した。

果粒検出精度の評価には、検出された果粒数と実際の果粒数との相対誤差を用いる。 N_t を実際の果粒数（手動カウントによる真値）、 N_p を提案手法により検出された果粒数とすると、相対誤差 δ は以下の式で表現される。

$$\delta = \frac{|N_p - N_t|}{N_t} \times 100 \quad (4)$$

相対誤差が小さいほど、提案手法の検出精度が高いことを示す。

4.2 実験結果

提案手法をシャインマスカット房に適用した結果を表 1 に示す。実際の果粒数 39 個に対して 38 個の果粒が検出され、相対誤差 δ は 2.56% という高精度な結果を得た。図 3 に示すように、提案手法による検出結果では、密集したブドウ果粒の大部分が正確に球体として検出されている。これは、3D Gaussian Splatting による高密度点群生成と RANSAC ベース球体検出の組み合わせが、テクスチャレスなブドウ果粒に対して有効であることを示している。検出に失敗した果粒は、図 4 の赤色で塗り潰した部分で示すように、房の内部に位置し他の果粒によって大部分が隠蔽されている。この果粒は隠蔽のため、撮影した複数の視点の画像にほとんど写り込んでおらず、その結果、3D Gaussian Splatting においても当該領域の 3次元復元の精度が悪く、後の RANSAC 球体検出に十分な点群が推定で



図 4: 検出出来なかった果粒

きなかったと考えられる。

5. まとめ

本研究では、ブドウの摘粒作業習得支援を目的として、3D Gaussian Splatting と RANSAC ベース球体検出を組み合わせた新しい果粒検出手法を提案した。提案手法は、テクスチャレスなブドウ果粒において特徴点抽出が困難という従来手法の限界を克服している。シャインマスカットを対象とした実験では、果粒の検出率 97.44% という高い性能を達成し、相対誤差 2.56% という結果を得た。今後の課題として、実験規模の拡大と、大部分が隠蔽された果粒の検出精度向上が挙げられる。また、多様な品種への適用が課題である。

謝辞 この研究は日本生命財団の支援を受けて実施された。

*1 <https://www.jawset.com/>

*2 <https://github.com/cloudcompare/cloudcompare>

参考文献

- [1] Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M. and Szeliski, R.: Building Rome in a Day, *Commun. ACM*, Vol. 54, No. 10, p. 105–112 (2011).
- [2] Bay, H., Tuytelaars, T. and Van Gool, L.: SURF: Speeded Up Robust Features, *Proceedings of European Conference on Computer Vision 2006*, pp. 404–417 (2006).
- [3] Caruso, D., Engel, J. and Cremers, D.: Large-scale direct SLAM for omnidirectional cameras, *Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 141–148 (2015).
- [4] Furukawa, Y., Sethi, A., Ponce, J. and Kriegman, D.: Structure and Motion from Images of Smooth Textureless Objects, *Proceedings of European Conference on Computer Vision 2004*, pp. 287–298 (2004).
- [5] Kerbl, B., Kopanas, G., Leimkühler, T. and Drettakis, G.: 3D Gaussian Splatting for Real-Time Radiance Field Rendering, *ACM Transactions on Graphics*, Vol. 42, No. 4 (2023).
- [6] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol. 60, pp. 91–110 (2004).
- [7] Mack, J., Lenz, C., Teutrine, J. and Steinhage, V.: High-precision 3D detection and reconstruction of grapes from laser range data for efficient phenotyping based on supervised learning, *Computers and Electronics in Agriculture*, Vol. 135, pp. 300–311 (2017).
- [8] Mur-Artal, R., Montiel, J. M. M. and Tardos, J. D.: ORB-SLAM: A Versatile and Accurate Monocular SLAM System, *IEEE Transactions on Robotics*, Vol. 31, No. 5, pp. 1147–1163 (2015).
- [9] Nurutdinova, I. and Fitzgibbon, A.: Towards Pointless Structure from Motion: 3D Reconstruction and Camera Parameters from General 3D Curves, *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 2363–2371 (2015).
- [10] R., S., R., W. and R., K.: Efficient RANSAC for Point - Cloud Shape Detection, *Computer Graphics Forum*, Vol. 26, No. 2, pp. 214–226 (2007).
- [11] Rist, F., Herzog, K., Mack, J., Richter, R., Steinhage, V. and Töpfer, R.: High-Precision Phenotyping of Grape Bunch Architecture Using Fast 3D Sensor and Automation, *Sensors*, Vol. 18, No. 3 (2018).
- [12] Rosten, E. and Drummond, T.: Machine Learning for High-Speed Corner Detection, *Proceedings of European Conference on Computer Vision 2006*, pp. 430–443 (2006).
- [13] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G.: ORB: An efficient alternative to SIFT or SURF, *Proceedings of 2011 International Conference on Computer Vision*, pp. 2564–2571 (2011).
- [14] Schönberger, J. L., Zheng, E., Frahm, J.-M. and Pollefeys, M.: Pixelwise View Selection for Unstructured Multi-View Stereo, *Proceedings of European Conference on Computer Vision*, Vol. 9907 (2016).
- [15] Tamura, Y., Utsumi, Y., Miwa, Y., Iwamura, M. and Kise, K.: Unsupervised monocular depth estimation with omnidirectional camera for 3D reconstruction of grape berries in the wild, *PLOS ONE*, Vol. 20, No. 2, pp. 1–18 (2025).
- [16] Wong, K.-Y. K. and Cipolla, R.: Structure and motion from silhouettes, *Proceedings of Eighth IEEE International Conference on Computer Vision. ICCV 2001*,