

ブドウ果粒の三次元位置情報に基づく 摘粒対象提示アプリケーションの開発

菅谷 友稀^{1,a)} 内海 ゆづ子^{2,b)} 三輪 由佳^{3,c)} 岩村 雅一^{2,d)}

概要

本研究では、ブドウ摘粒作業の習得支援を目的として、動画撮影から摘粒対象の提示までを自動的に完結させる摘粒支援アプリケーションを開発した。三次元点群処理による三次元再構成と球体検出によって各果粒の三次元位置を推定し、摘粒すべき果粒の選択を組合せ最適化問題として定式化して遺伝的アルゴリズムを適用することで摘粒対象を決定する。決定した摘粒対象をインタラクティブな 3D モデル上に可視化することで、ユーザが直感的に摘粒箇所を確認できる。評価実験では、アプリを使用することで初心者でも熟練者と同程度の果粒間距離を実現できることを確認した。

1. はじめに

日本におけるブドウ栽培では、収穫物の味の良さだけでなく、房全体の外観や形状にも極めて高い品質が求められる。高品質なブドウを栽培するために不可欠なプロセスが摘粒である。摘粒とは房から過剰な果粒を取り除くことで、残った果粒の肥大促進や糖度向上を図り、房の形状を理想的な逆円錐形などに整える作業である。摘粒作業は目標とする果粒数や果粒同士の適切な距離、房の内部構造、さらには将来的な成長スペースの確保といった複数の要素を考

慮しながら、瞬時に間引くべき果粒を判断する必要がある。加えて、この作業は約 10 日間という極めて短い期間内に全ての房で行わなければならない、迅速な作業が求められる。

前述の通り、摘粒の基準は複雑であり、かつ素早い作業が求められるため、新規就農者やブドウ栽培初心者が技術を習得するには相当な時間を要する。また、農業従事者の高齢化に伴い、摘粒技術の次世代への継承も課題となっている。そのため、情報技術を用いた摘粒技術習得支援が期待されている。

摘粒作業の支援を目的とした研究として、Buayai ら [1] は、ディープニューラルネットワークを用いて二次元画像から摘粒すべき果粒を識別し、その結果を AR 技術によってスマートグラスに提示する摘粒支援システムを提案した。このシステムは二次元画像から摘粒すべき果粒を識別するものであり、果粒間の三次元的な空間配置を考慮した最適化は行われていない。また、専用のスマートグラスが必要であり、その購入・維持管理コストが農業現場への普及における大きな障壁となる。

そこで本研究では、新たにブドウ栽培を始める人の摘粒技術習得を支援することを目的として、図 1 のようなブドウ房の果粒の三次元配置に基づいて摘粒対象を選択する摘粒支援システムを開発する。このシステムでは、動画から各果粒の三次元位置を推定することで果粒間の空間的な配置を正確に把握し、間引く果粒の選択を組み合わせ最適化問題として定義し、遺伝的アルゴリズムによって果粒が均等に分散するような最適な摘粒対象を決定する。さらに、決定した摘粒対象をスマートフォン上の 3D モデルで可視化することで、専用デバイスを必要とせず誰でも直感的に摘粒箇所を確認できる。

2. 提案システム

2.1 システム概要

本研究で提案する摘粒支援システムは、スマートフォンでブドウを撮影するだけで、摘粒すべき果粒を 3D モデルで確認できる摘粒支援システムである。ユーザはまずアプリの設定画面で解析する品種と摘粒後に残す目標粒数を設定する。次に、ArUco マーカ [3] をブドウ房近傍に設置した



図 1: 摘粒支援システム

¹ 大阪公立大学工学部情報工学科

² 大阪公立大学大学院情報学研究科

³ 大阪府立環境農林水産総合研究所

a) sp23904i@st.omu.ac.jp

b) yuzuko@omu.ac.jp

c) MiwaY@knsk-osaka.jp

d) masa.i@omu.ac.jp

状態でスマートフォンをブドウ房の周囲に一周させながら動画を撮影する。撮影後、アプリ上の解析ボタンを押すだけで、動画の送信から三次元再構成・果粒検出・摘粒対象の決定までの一連の処理が自動的に実行される。処理が完了するとアプリは自動的に 3D ビューア画面に遷移し、ブドウ房の 3D モデルが表示される。3D モデル上では摘粒対象の果粒に赤いマークが配置されており、ユーザはモデルをタッチ操作で回転・拡大縮小して摘粒箇所を直感的に確認しながら、摘粒作業を行うことができる。

2.2 システム構成

本システムは、スマートフォン、クラウド上のデータ管理基盤、GPU サーバの 3 つの構成要素から成る。三次元再構成や最適化処理は計算コストが高くスマートフォン単体では処理できないため、撮影と結果表示をスマートフォンが、重い計算処理を GPU サーバが担う構成とし、両者の間にクラウド上のデータ管理基盤を介して大容量データを非同期にやり取りする。スマートフォンアプリは Flutter で実装し iPhone 15 Pro (iOS 26.5) 上で、クラウド基盤には Firebase (Firebase Storage) を、GPU サーバには NVIDIA H100 PCIe を使用した。

2.2.1 スマートフォン

スマートフォンはユーザとのインターフェースを担う構成要素であり、房の動画撮影、撮影データのクラウドサービスへの送信、および解析結果の表示を担う。撮影では、ArUco マーカを房上部の軸に設置した状態でスマートフォンを房の周囲に一周させながら動画を撮影することで、三次元再構成に必要な多視点画像を取得する。三次元再構成・果粒検出・摘粒対象の決定が完了すると、GPU サーバ上で決定された摘粒対象の果粒が 3D モデル上に提示され、ユーザは摘粒箇所を確認することができる。

2.2.2 クラウド上のデータ管理基盤

クラウド上のデータ管理基盤はスマートフォンと GPU サーバの間でデータの保存と受け渡しを担う。スマートフォンで撮影した動画ファイルをクラウド上のデータ管理基盤に保存し、GPU サーバがそれを取得して処理を実行する。GPU サーバでの処理が完了すると、解析結果の三次元点群データと摘粒対象の情報がクラウド上のデータ管理基盤に保存され、スマートフォンがそれを取得して表示する。両者が直接通信するのではなくクラウド上のデータ管理基盤を介してデータの受け渡しを行うことで、スマートフォンと GPU サーバはそれぞれ独立して動作できる。これにより、ユーザは GPU サーバ上での三次元再構成・果粒検出・摘粒対象の決定の完了を待つことなく、スマートフォンで他の操作を行うことができる。

2.2.3 GPU サーバ

GPU サーバは山崎らの手法 [7] に基づき、ブドウ房の三次元再構成、スケール推定、果粒検出、摘粒対象の決定の 4

段階の処理を行う。以下にそれぞれの処理を示す。

まず、撮影した動画から抽出した多視点画像を用いてブドウ房の三次元再構成を行う。Structure from Motion (SfM) [6] と Eliminating Densification for Efficient Convergence of 3DGS (EDGS) [4] を用いて、高密度な三次元ガウシアン点群を生成する。また、Density-Based Spatial Clustering of Applications with Noise (DBSCAN)[2] の密度に基づくクラスタリングによって、最も点の多いクラスタをブドウの房とみなして抽出する。

次に、RANSAC による球体検出には果粒の実際の大きさに対応した半径範囲を実寸で指定する必要があるため、スケール推定を行う。三次元再構成によって得られた点群は任意のスケールで表現されているため、実寸が既知の ArUco マーカを基準にスケール係数 s [m/unit] を求める。具体的には、ArUco マーカが検出された各フレームについて SfM によるカメラ位置を ArUco 座標系と三次元再構成座標系の両方で取得し、検出フレームから選んだ 2 フレームの全組み合わせについて実世界座標系でのカメラ間距離 d_{aruco} [m] と三次元再構成座標系でのカメラ間距離 d_{recon} [unit] の比から

$$s = \frac{d_{\text{aruco}}}{d_{\text{recon}}} \quad (1)$$

としてスケール係数を求め、推定誤差を軽減するため全組み合わせの中央値を最終的なスケール係数として採用し、点群の各座標値に乘じることで実寸法に変換する。

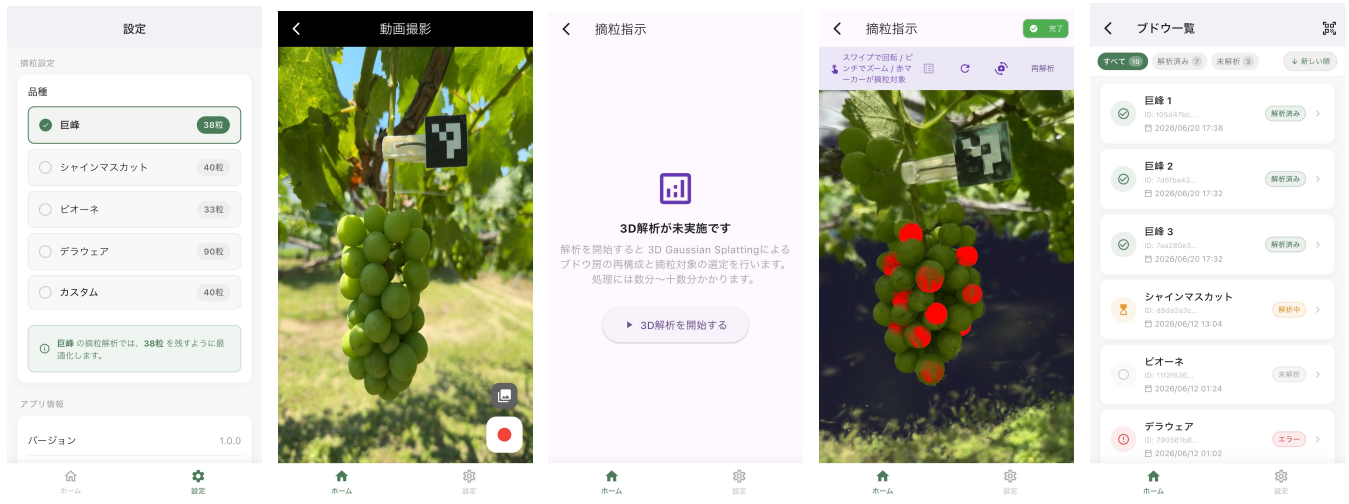
続いて、Random Sample Consensus (RANSAC) [5] による球体検出を適用し、各果粒の三次元位置と半径を推定する。点群中から球面に適合する点の集まりを繰り返し探索し、果粒の半径として想定される範囲内かつフィッティング誤差が一定の閾値以下となる球体のみを有効な果粒として採用する。

最後に、検出された果粒群から間引く果粒を選択する問題を組合せ最適化問題として定式化し、遺伝的アルゴリズムを適用することで摘粒すべき果粒を決定する。果粒間の距離の総和と目的粒数との差から構成される評価関数を定義し、この評価関数の最大化を図ることで、最適な摘粒対象を決定する。

3. アプリケーションの実装

本システムのアプリは、目標果粒数設定画面、撮影画面、摘粒指示・3D ビューア画面、ブドウ房一覧画面の 4 つの主要画面から構成される。以下では、各画面の機能と操作の流れを説明する。

ブドウ房を撮影する前に、まず図 2a で示す目標果粒数設定画面で品種と目標果粒数を設定する。摘粒後に房に残す果粒数は、品種や農家の好み、また房の状態によって異なるため、本アプリではユーザが任意に設定できる設計とした。シャインマスカット・巨峰・ピオーネなどの代表的な



(a) 目標果粒数設定画面 (b) 撮影画面 (c) 摘粒指示画面 (d) 3D ビュー画面 (e) ブドウ房一覧画面

図 2: アプリケーション画面

品種がプリセットとして用意されており、各品種に対して摘粒後に残す標準的な目標果粒数があらかじめ設定されている。プリセットにない品種はカスタム品種として登録でき、目標果粒数を数値入力で任意に設定できる。設定した目標果粒数は解析リクエスト時に GPU サーバへ送信され、摘粒対象の決定に使用される。

図 2b のように撮影画面でスマートフォンのカメラを用いてブドウ房の動画を撮影する。撮影の際は ArUco マーカをブドウ房の近傍に設置した状態でスマートフォンをブドウ房の周囲に一周させながら撮影する。ライブラリから動画を選択する機能も用意されており、事前に撮影済みの動画を使用することも可能である。撮影完了後、動画は自動的にクラウド上のデータ管理基盤に保存される。

図 2c で示す摘粒指示画面より、3D 解析を開始するボタンを押すと、GPU サーバで三次元再構成・果粒検出・摘粒対象の決定が自動的に実行される。処理中は「三次元再構成中」「果粒検出中」「摘粒対象の決定中」といった進捗状況が画面上にリアルタイムで表示されるため、ユーザは処理の進捗を随時確認できる。解析が完了すると 3D ビュー画面に自動遷移し、図 2d のようにブドウ房の 3D モデルがスマートフォン上に表示され、摘粒対象の果粒には赤い球体マークで提示される。カメラは初期表示時に全果粒の重心が画面中心に来るように自動的に調整されるため、果房全体を見渡ししやすい視点でモデルが表示される。ユーザは 1 本指のスワイプ操作でモデルを回転、2 本指のピンチ操作で拡大縮小しながら、さまざまな角度から摘粒箇所を確認できる。摘粒対象の果粒を実際に摘粒した後は、対応するマークをタッチすることで非表示にでき、残りの摘粒対象を把握しながら作業を進めることができる。マークへのタッチ操作は二段階方式を採用しており、一度タッチすると点滅して確認状態となり、もう一度タッチするとマ

カが非表示になる。一度のタッチで即座に摘粒済みと記録されない設計とすることで、作業ミスの低減を図っている。

各房の撮影データおよび解析結果はすべて図 2e で示すブドウ房一覧画面で管理される。登録した果房はリスト形式で表示されており、各果房の名前・撮影日時・解析ステータス（未解析・解析中・解析済み・エラー）を一覧で確認できる。一覧はフィルタリングおよびソート機能を備えており、解析済み・未解析での絞り込みや撮影日時の昇順・降順での並び替えが可能である。また、各果房には名前とメモを自由に記録できる編集画面があり、圃場内の場所や作業時の気づきをアプリケーション上で果房ごとに管理できる。

4. 実験

提案するブドウ摘粒支援アプリケーションの有効性を検証するため、評価実験を行った。開発したアプリケーションを使用することで、熟練者に近い品質や効率で摘粒作業を行えるかを評価する。

4.1 実験条件

本実験は、2026 年 6 月 11 日に大阪公立大学農学部附属教育研究フィールドにて実施した。巨峰を対象とし、摘粒後の目標粒数は 38 粒に統一した。実験では、アプリを使用して摘粒を行う実験協力者（以降、アプリと呼ぶ）、摘粒経験のない初心者（以降、初心者と呼ぶ）、および摘粒経験を持つ熟練者（以降、熟練者と呼ぶ）の 3 名で行った。各実験協力者が 3 房ずつ、合計 9 房に対して摘粒作業を行った。RANSAC による球体検出では、巨峰の果粒サイズに基づき半径の探索範囲を最小 4mm、最大 10mm として設定した。

4.2 実験方法

実験にあたり、各実験協力者には作業内容に関する指示

表 1: アプリ、初心者、熟練者の実験結果比較

実験協力者	摘粒前粒数	摘粒後粒数	絶対誤差	時間	秒/粒	房の長さ	最近傍平均	最近傍標準偏差
アプリ	67.3	38.0	2.0	11:33	23.6	142.0mm	20.8mm	2.8mm
初心者	61.0	37.0	2.3	9:02	22.6	124.6mm	17.6mm	2.0mm
熟練者	54.3	38.0	1.3	6:57	25.5	139.8mm	20.7mm	2.1mm

を以下の通り与えた。アプリには事前にアプリケーションで房の解析を行い、3D ビューア上に提示された摘粒対象の指示に従って作業を行うよう指示した。初心者には摘粒のコツや方法を一切与えず、38 粒を残して摘粒するよう指示した。熟練者にも同様に 38 粒を残すように摘粒するよう指示した。具体的な実験の手順は以下の通りである。

- (1) 各実験協力者は、与えられた指示に従って、用意された巨峰の房の摘粒作業を行う。その際、作業開始から終了までの時間を計測する。
- (2) 摘粒完了後、スマートフォンを用いて各房の周囲を一周する動画を撮影する。
- (3) 撮影した動画を本システムの解析パイプラインで処理し、EDGS による三次元再構成および RANSAC による球体検出を適用することで、摘粒後に残った各果粒の三次元座標を算出し、摘粒後の房の状態を評価する。

4.3 評価方法

得られた果粒の三次元座標をもとに、以下の指標で摘粒後の房の状態を評価する。摘粒精度の指標として、目標粒数 $N_t = 38$ に対する摘粒後の粒数 N_r との誤差 $|N_r - N_t|$ を用いた。この指標はアプリの提示する摘粒対象の数が目標粒数に対してどれだけ正確かを評価する。作業効率の指標として、房ごとに摘粒する粒数が異なる条件を考慮し、果粒 1 粒を処理するのに要した平均時間 [秒/粒] を用いる。この指標はアプリを使用することで作業者の摘粒速度がどのように変化するかを評価する。さらに、摘粒後の果粒が空間的に均等に分散しているかを評価するため、各果粒の最近傍の距離平均 $\bar{d}_{nn}$ および最近傍距離の標準偏差を算出した。なお、房の長さが異なると果粒間距離の比較に影響が生じるため、各房の長さを全 9 房の平均値に正規化した。 $\bar{d}_{nn}$ は、 $\mathbf{g}_i$ を i 番目の果粒の三次元座標とし、検出された N 個の果粒のそれぞれについて最も近い果粒との距離を求め、その平均値として以下のように定義される。

$$\bar{d}_{nn} = \frac{1}{N} \sum_{i=1}^N \min_{j \neq i} \|\mathbf{g}_i - \mathbf{g}_j\|_2 \quad (2)$$

$\bar{d}_{nn}$ が大きいほど果粒が密集せず均等に配置されていることを示し、残存果粒の肥大促進という摘粒の目的に沿った配置が実現されているかを評価できる。また、最近傍距離の標準偏差が小さいほど各果粒の間隔のばらつきが小さく、

房全体で均一な果粒配置が実現されていることを示す。

4.4 実験結果

実験の結果を表 1 に示す。表は、アプリ、初心者、熟練者の 3 名それぞれの平均値を示している。最近傍の距離平均は、アプリが 20.8mm、熟練者が 20.7mm、初心者が 17.6mm の順となり、アプリと熟練者がほぼ同等の値を示すことができた。これはアプリが果粒間距離の総和を最大化するように最適化した摘粒対象を提示した結果、ユーザが指示に従って作業するだけで果粒が房全体に均等に分散する配置が実現できたためである。

目標粒数 38 粒に対する摘粒後の絶対誤差を見ると、熟練者が 1.3 粒、アプリが 2.0 粒、初心者が 2.3 粒であった。アプリは初心者に比べて絶対誤差が小さく、目標粒数により近い粒数にできた。アプリで誤差が出るのは、ブドウの房の解析の誤差や作業者のミスのためである。

1 粒あたりの摘粒時間は、初心者、アプリ、熟練者の順となった。アプリを使用した場合、初心者と同程度の作業時間となり、熟練者と比較しても大きな差は見らず、アプリを使用することによる作業時間の短縮は認められなかった。アプリを使用した摘粒では、3D ビューア上の摘粒対象と実際のブドウ房の果粒を対応付ける位置合わせ作業が必要となり、この作業がユーザの負担となり、時間がかかったと考えられる。アプリは摘粒の判断をシステムに委ねることで判断の迷いをなくすることができる一方、3D モデルと現実の房との対応付けというアプリ固有の作業が新たに生じるため、作業時間の短縮には至らなかった。

5. まとめ

本研究では、ブドウ摘粒作業の習得支援を目的として、動画撮影から摘粒対象の 3D モデル提示までを自動的に完結させる摘粒支援アプリケーションを開発した。評価実験の結果、果粒が均等に分散しているかの指標においてアプリと熟練者がほぼ同等の値を示し、アプリを使用することで熟練者に近い品質の摘粒が実現できることが示された。また、目標粒数との誤差においてもアプリは初心者を上回り、アプリの指示に従うことで摘粒精度の向上が確認された。今後の課題として、隠蔽された果粒の検出精度の向上、3D モデルと実際の房との対応付けの自動化が挙げられる。

参考文献

- [1] Buayai, P., Yok-In, K., Inoue, D., Nishizaki, H., Makino, K. and Mao, X.: Supporting table grape berry thinning with deep neural network and augmented reality technologies, *Computers and Electronics in Agriculture*, Vol. 213, p. 108194 (2023).
- [2] Ester, M., Kriegel, H.-P., Sander, J. and Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, AAAI Press, p. 226–231 (1996).
- [3] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. and Marín-Jiménez, M.: Automatic generation and detection of highly reliable fiducial markers under occlusion, *Pattern Recognition*, Vol. 47, No. 6, pp. 2280–2292 (2014).
- [4] Kotovenko, D., Grebenkova, O. and Ommer, B.: EDGS: Eliminating Densification for Efficient Convergence of 3DGS, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 41065–41076 (2026).
- [5] Schnabel, R., Wahl, R. and Klein, R.: Efficient RANSAC for Point-Cloud Shape Detection, *Computer Graphics Forum*, Vol. 26, No. 2, pp. 214–226 (2007).
- [6] Schönberger, J. L., Zheng, E., Frahm, J.-M. and Pollefeys, M.: Pixelwise View Selection for Unstructured Multi-View Stereo, *Computer Vision – ECCV 2016* (Leibe, B., Matas, J., Sebe, N. and Welling, M., eds.), Cham, Springer International Publishing, pp. 501–518 (2016).
- [7] Yamasaki, T., Ryo, S., Utsumi, Y., Tadokoro, K., Miwa, Y. and Iwamura, M.: Automatic Thinning Selection Based on 3D Grape Bunch Reconstruction, *SSRN Electronic Journal* (2026).